The High-Latitude Ionosphere and its Effects on Radio Propagation

R. D. Hunsucker, J. K. Hargreaves

Book DOI: http://dx.doi.org/10.1017/CBO9780511535758

Online ISBN: 9780511535758

Hardback ISBN: 9780521330831

Paperback ISBN: 9780521041362

Chapter

Chapter 7 - The high-latitude D region pp. 337-416

Chapter DOI: http://dx.doi.org/10.1017/CBO9780511535758.009

Cambridge University Press

Chapter 7

The high-latitude D region

7.1 Introduction

The differences between the E and D regions in middle latitudes hold also at high latitude. The E region is characterized by relatively simple photochemistry and high electrical conductivity, whereas the D region below it has a complex and less well-known chemistry, the electric currents and plasma motions being inhibited by the higher atmospheric pressure. What they have in common at high latitude is the importance of ionization by energetic particles. Typical spectra include particles with energies such that they are stopped and ionize in both regions, the lower energies (for example, electrons of a few kilo-electron volts) affecting the E region and the higher ones (e.g. electrons with energies of tens of kilo-electron volts) penetrating into region D. Figure 7.1 shows electron-density profiles between 65 and 110 km due to representative spectra of ionizing electrons incident on the atmosphere from above. Increasing the characteristic energy of the spectrum lowers the peak of the layer, increasing the electron density in the D region but reducing it in the E region.

At middle latitude the D region's role in radio propagation is a secondary one. The main parameters of HF propagation are determined by the E and the F regions, and the D region acts mainly as an absorbing layer, reducing the strength of the signal but seldom preventing communications for any long period. At high latitude the D region may be much enhanced and then absorption becomes a considerable problem. There are two principal phenomena, each peculiar to high latitude. The first is *auroral radio absorption (AA)*, which occurs only in the auroral regions and is due to fluxes of energetic electrons precipitated from the magneto-

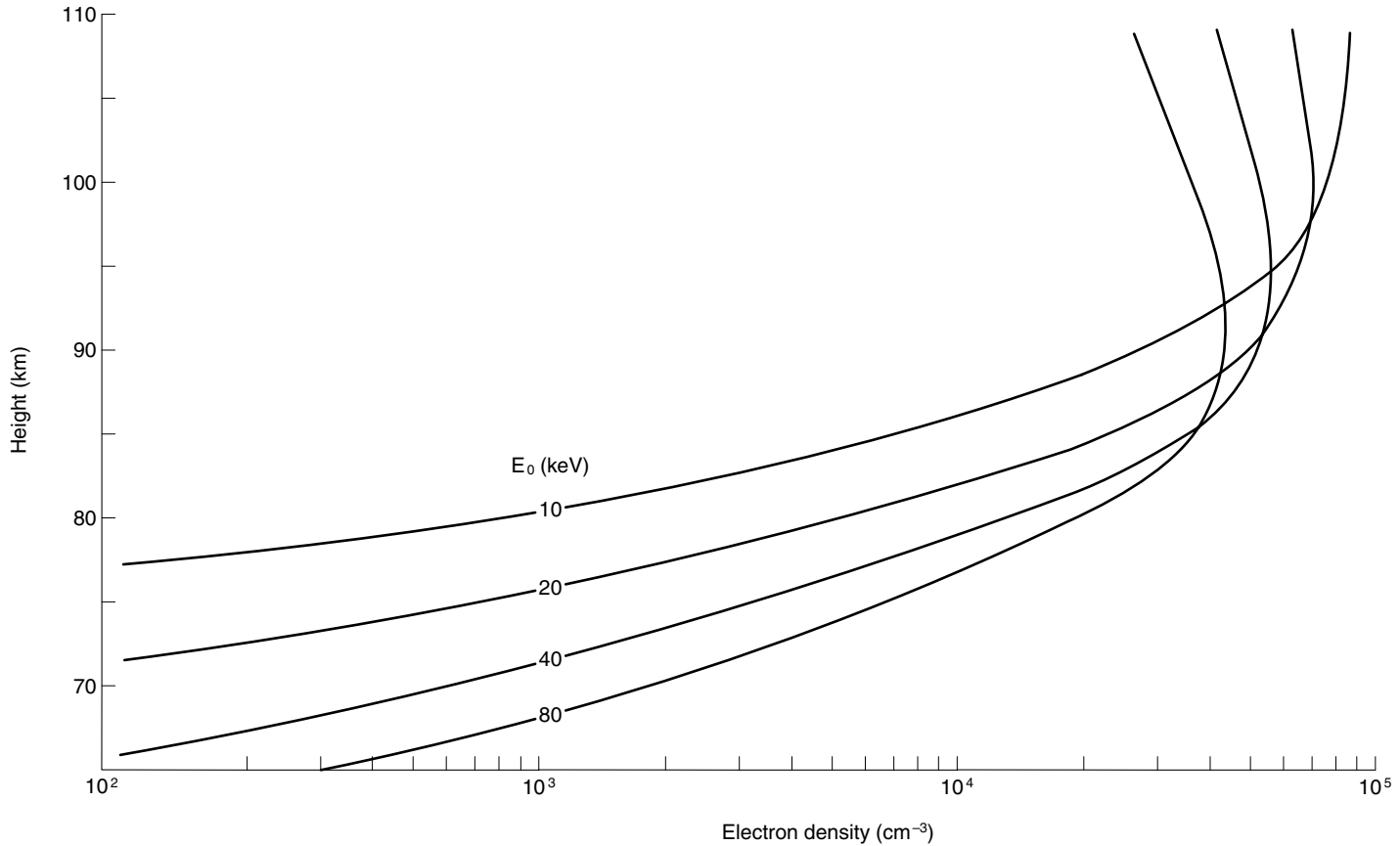


Figure 7.1. Specimen profiles of the D and lower E regions due to electron precipitation. The incident electron flux is assumed to be of the form $\exp(-E/E_0)$, E being the energy in kilo-electron volts and E_0 the characteristic energy, the total energy flux being 4×10^7 keV s^{-1} cm^{-2} sr^{-1} in each case. Production rates were worked out using the method outlined in Section 2.6.1, and a profile of the effective recombination coefficient was assumed.

sphere sporadically during periods of auroral activity. The second is *polar-cap absorption (PCA)*, which is caused by energetic protons emitted from the Sun, usually at the time of a major solar flare.

These two kinds of phenomenon have rather different properties. The PCA is relatively infrequent, there being only about one event per month on the average in a year of high solar activity; many less when the Sun is quiet. However, when an event does occur, the absorption may be very strong. The absorbing region is relatively uniform over the whole polar cap, leading to HF black-out over a wide area. Auroral absorption is more common, but it is confined to the auroral zones and is generally more structured. Though the amount of absorption does not rise to the intensity sometimes seen in PCA, the spatial structure, which is generally not known in detail, adds to the difficulty of predicting the effects on HF propagation.

AA and PCA are discussed in Sections 7.2 and 7.3, respectively. The chapter concludes with an introduction to a phenomenon that is still not well understood, the *polar mesosphere summer echo*.

7.2 Auroral radio absorption

7.2.1 Introduction – history and technique

Auroral radio absorption was discovered by Appleton and colleagues (Appleton *et al.*, 1933) during an expedition to Tromsø during the International Polar Year (1932–1933), when it was observed that ionosondes were blacked out during periods of auroral and magnetic activity. The earliest studies of the phenomenon were performed using ionosondes, but this method is not entirely satisfactory because all that can be measured is the incidence of black-out, which in any case might not always be due to absorption; furthermore, ionosondes are not all equally sensitive. The absorption due to proton events, which has different properties, could also have confused the early results since the proton event was not at that time recognized as a separate entity.

Since the International Geophysical Year of 1957–1958, auroral absorption (AA) has generally been studied with a riometer – and preferably with a group of riometers covering a range of latitude and/or longitude. Since it uses trans-ionospheric propagation, the riometer does not say at what height the absorption occurs, but various studies have left little doubt that most of the absorption detected in the auroral regions arises in the D region of the ionosphere and is caused by energetic electrons arriving from the magnetosphere.

Riometer technique is outlined in Section 4.2.4. We should recall here that the absorption is not measured directly but requires first the determination of a *quiet-day curve (QDC)*, an estimate of the signal level in the absence of absorption.

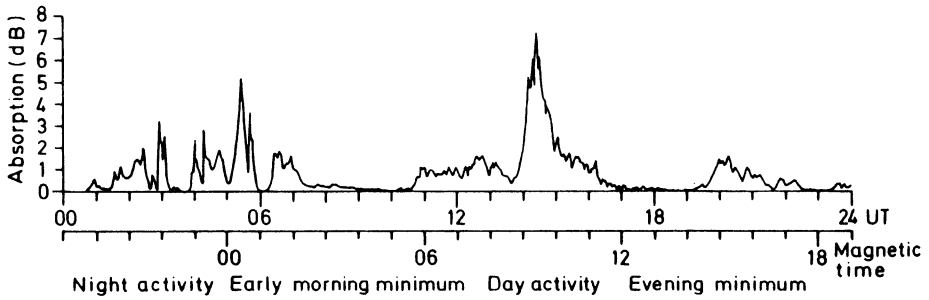


Figure 7.2. Auroral radio absorption observed with a 30-MHz riometer on 15 October 1963 at Byrd Station, Antarctica. The descriptions below the axis refer to the typical behavior; the evening minimum was not respected on the day shown! Note the difference of structure between the night and day activity. (After J. K. Hargreaves, *Proc. Inst. Electr. Electron. Engineers* **57**, 1348, 1969, © 1969 IEEE.)

Although the idea is simple enough, the accurate derivation of the QDC can be the most difficult part of absorption measurement by the riometer technique.

Most riometer-based absorption data come from instruments using a simple antenna, which, therefore, has a wide beam – e.g. 60° between half-power points – projecting onto a region about 100 km across in the D region. Therefore a standard riometer installation has limited spatial resolution. In recent years, however, there has been an increase in narrow-beam work and the use of imaging riometers. We shall quote results from both wide-beam and narrow-beam instruments.

Since the absorption is strongly frequency-dependent (an inverse square law in most circumstances), the observing frequency must also be stated. The reduction of absorption with increasing frequency is one factor determining the optimum observing frequncy. At the lower frequencies the antenna is larger and also there is more interference from ionospherically propagated signals. The compromise has generally led to use of the band 30–50 MHz. When data are obtained at several frequencies, it is usual to reduce them to 30 MHz for comparison purposes:

$$A(30\ \text{MHz}) = A(f)(30)^2/f^2. \tag{7.1}$$

7.2.2 Typical auroral-absorption events and their temporal and spatial properties

One notable fact about auroral absorption is its temporal structure, distinguishing it from other major varieties of radio absorption, which generally vary more gradually. In the example of Figure 7.2 it is seen that the absorption tends to occur in bursts (or *events*); these show preferences for certain times of day, and they

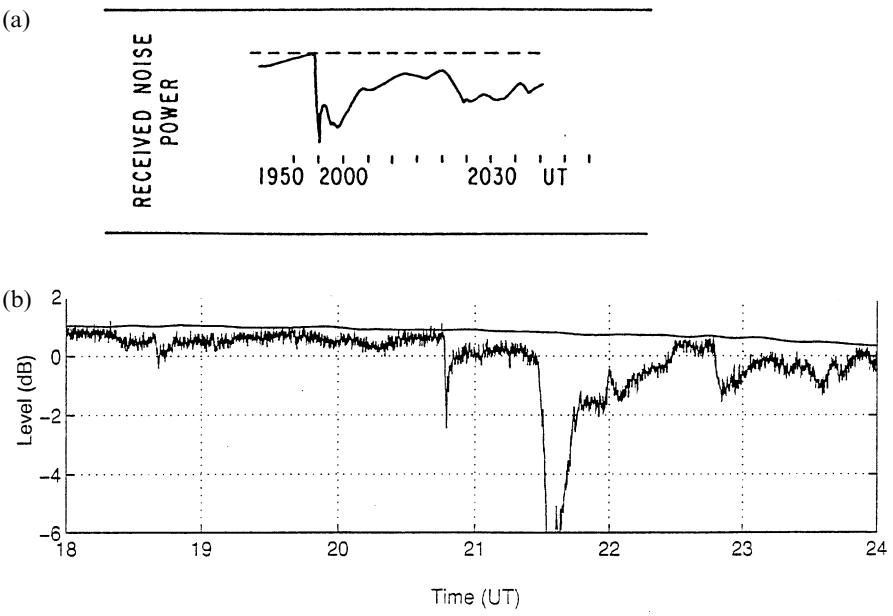


Figure 7.3. Examples of sharp-onset night events. (a) Skibotn ($L = 6.0$), 4 November 1975 (J. K. Hargreaves *et al.*, *J. Geophys. Res.* **84**, 4225, 1979, copyright by the American Geophysical Union.) (b) Kilpisjärvi ($L = 5.9$), 6 October 1994. (Reprinted from J. K. Hargreaves *et al.*, *J. Atmos. Solar–Terr. Phys.* **59**, 872, copyright 1997, with permission from Elsevier Science.). The first is in the form of a traditional riometer chart, with the level proportional to the received power. The lower one was reconstructed from digital data and the signal power is plotted on a scale of decibels. The absorption is reckoned from the marked quiet-day curve.

change character between day and night. While there is no general classification of auroral absorption that covers all events, there are some recognized types that occur frequently.

Sharp-onset and spike events at night

Occurring near magnetic midnight (and more before than after) are *sharp-onset events*. Here the event rises in a couple of minutes or less (Figure 7.3). The duration is tens of minutes to an hour. Some of these events appear isolated, but others are followed by continuing activity lasting for several hours. The continuing activity tends to be less prominent at the higher latitudes. At some other latitudes what appears to be the same event may begin with a more gradual onset. Many sharp onsets, though not all, coincide with the beginning of a substorm.

At the beginning of the event there may be a "spike," as in the examples of Figure 7.3, in which case that feature is a *spike event*. The occurrence of spike events at $L = 5.6$ is shown in Figure 7.4. At that location, half occurred in a 3-h period up to local magnetic midnight. Usually, an individual spike is seen over a

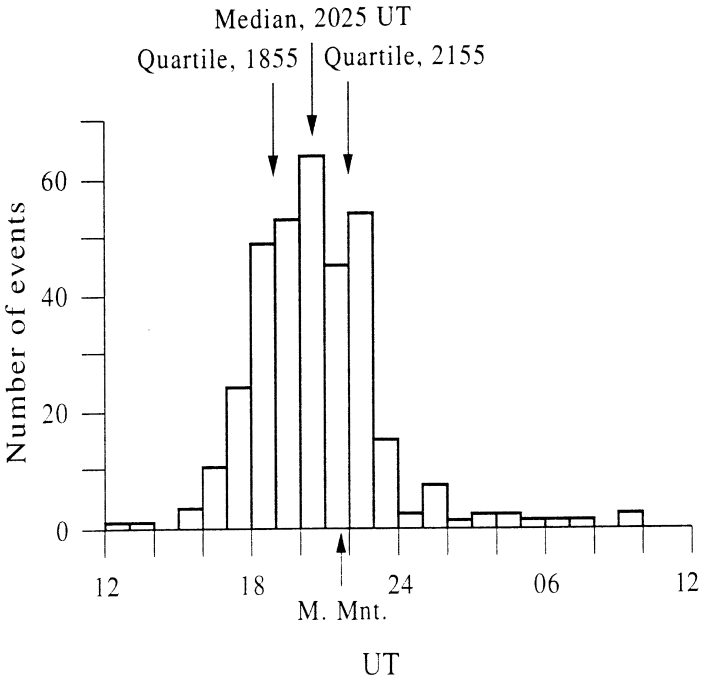


Figure 7.4. The occurrence of spike events (1 dB) at Abisko, 1980–1985. Magnetic midnight is about 2130 UT. The riometer frequency was 30 MHz. (Reprinted from J. K. Hargreaves *et al.*, *J. Atmos. Solar–Terr. Phys.* **59**, 872, copyright 1997, with permission from Elsevier Science.)

more limited range of latitude (probably less than 200 km) than the onset itself, which in some cases may be tracked over a wide range of L values if some time differences are allowed.

Figure 7.5 illustrates the spatial confinement of the spike event determined with an imaging riometer. At $L = 5.9$ the typical spike event is elliptical in shape, the major axis being generally east–west. Typical dimensions are 190 km by 80 km, and the axial ratio is about 2.5 (Hargreaves *et al.*, 1997). The properties of spike events at the South Pole ($L \sim 13$) have been found to be remarkably similar in form though they are generally smaller in magnitude. The spike event lasts for 1–2 min only, and is dynamic (Section 7.2.4).

The main part of the night event is considerably more widespread than the spike. As an example, Figure 7.6 shows the distribution before, at, and after the intense peak (2132 UT) in the event of Figure 7.3(b). During the main part of the event (which peaked at about 9 dB at 38.2 MHz) the absorption covers the whole area at a substantial level, although spatial structure is also present. The properties of these events have not yet been worked out in full, but they are clearly distinct from those of spikes and arcs (see below).

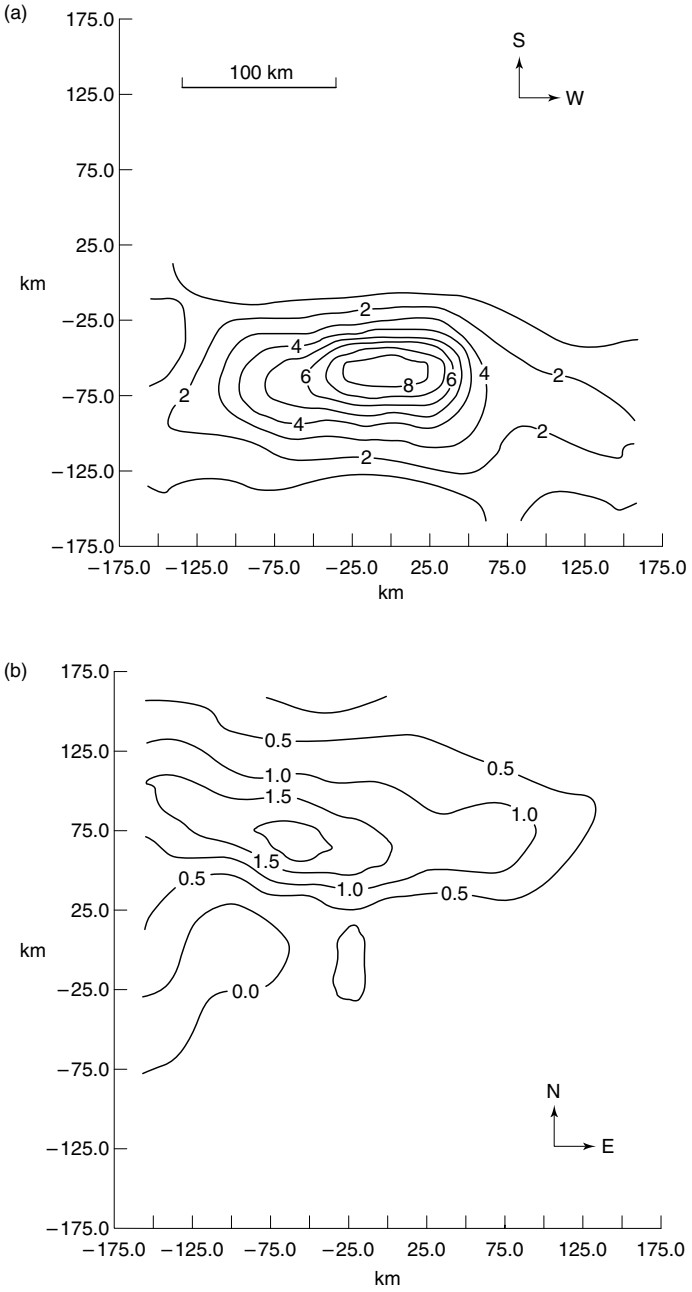
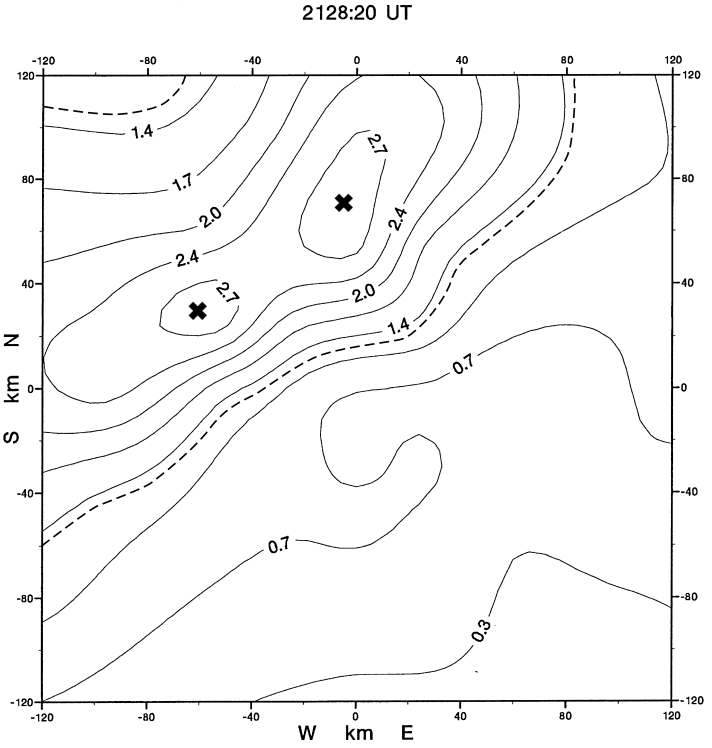
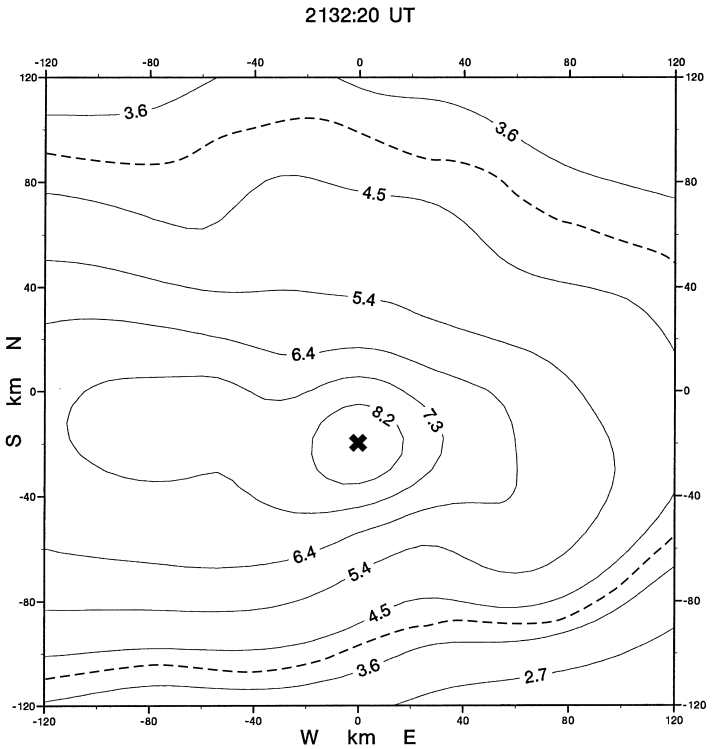


Figure 7.5. Examples of spike events at (a) the South Pole, 22 July 1988 at 2042:50 UT, and (b) Kilpisjärvi, 14 November 1994 at 2015:10 UT showing typical dimensions. A height of 90 km is assumed. Contours are of absorption in decibels at 38.2 MHz, and the time resolution was 10 s. This South-Pole ($L \sim 13$) event was an exceptionally intense one, but that at Kilpisjärvi ($L = 5.9$) is more typical for that latitude. ((a) J. K. Hargreaves *et al.*, *Radio Sci.* **26**, 925, 1991, copyright by the American Geophysical Union; (b) reprinted from J. K. Hargreaves *et al.*, *J. Atmos. Solar–Terr. Phys.* **59,** 872, copyright 1997, with permission from Elsevier Science.)





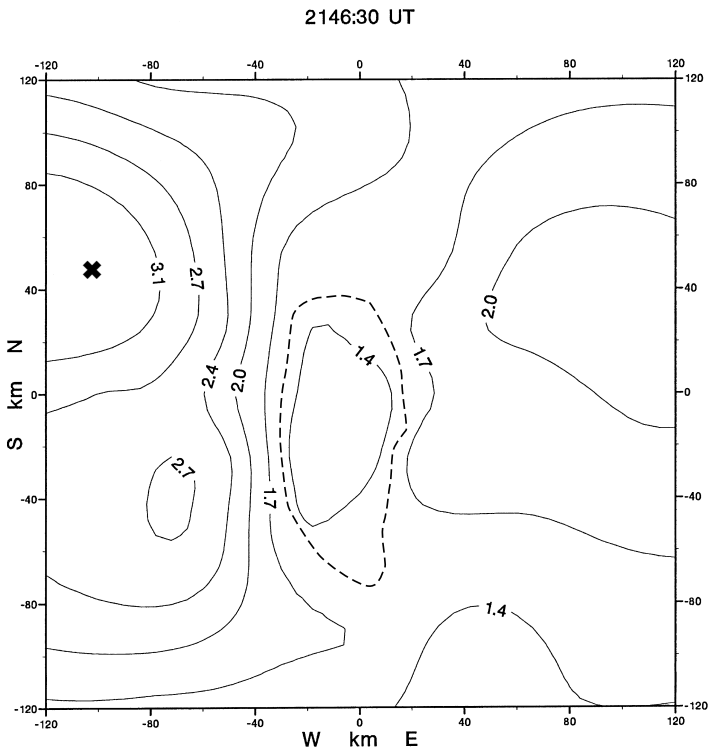
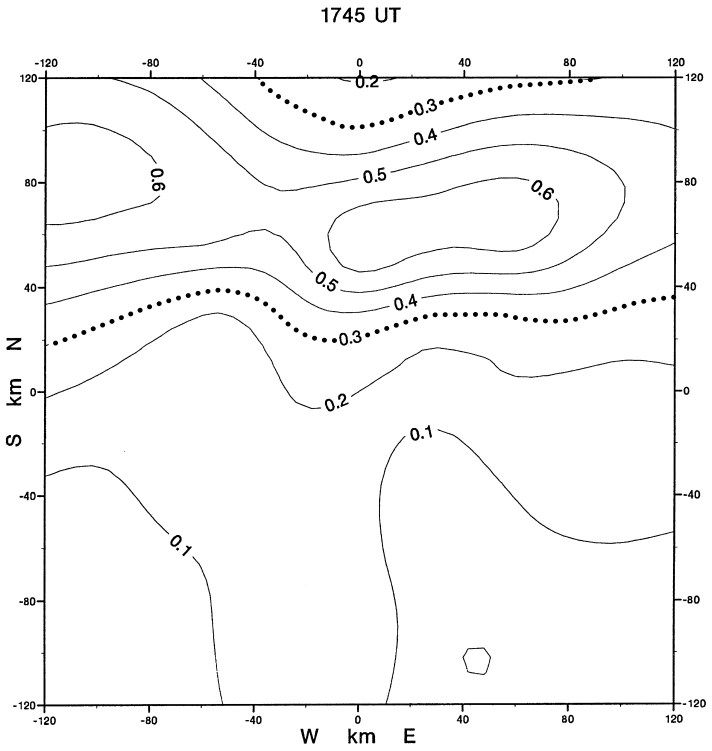


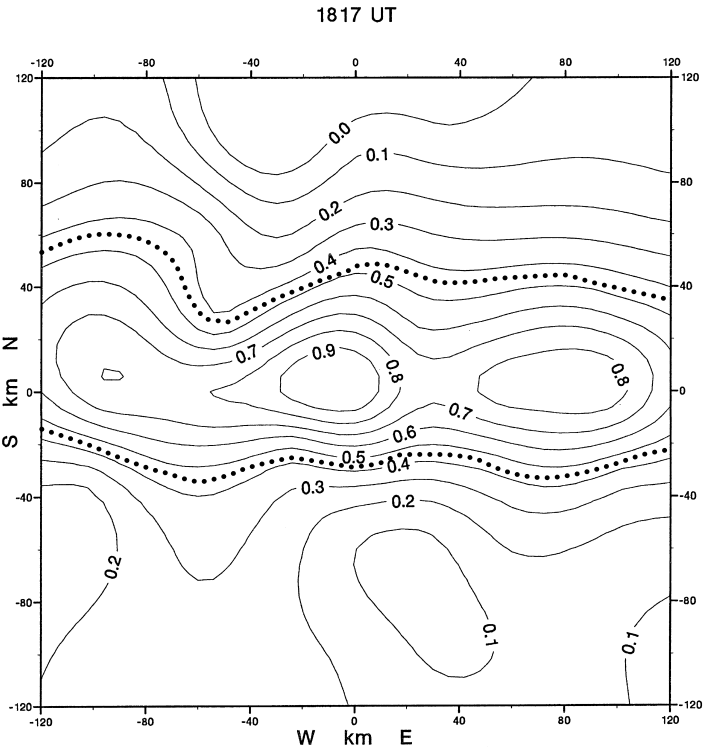
Figure 7.6. Absorption distributions during the main part of a night event at Kilpisjärvi on 6 October 1994. The highest contours are 2.7, 8.2, and 3.1 dB at 2128:20, 2132:20, and 2146:30 UT, respectively. The frequency is 38.2 MHz. The maxima are marked x and the contours at half the maximum are dotted. These are 10-s averages. The event came into view from the north-west, peaked overhead, and then drifted westward.

Apparently distinct from the spike is another short-duration feature, which appears for only a short time because it rapidly crosses the field of view in a westward direction, namely the *westward surge*. It extends 75–85 km north–south but its east–west dimension is not known. It may be related to the westward-traveling surge in the luminous aurora (Section 6.4.2).

Daytime spike events

The larger spike events are never observed in the day sector, but smaller ones have been observed by day at high latitude in the northern hemisphere (Stauning and Rosenberg, 1996). At Sondrestrom (invariant latitude 73.7°, $L \sim 13$) they have durations of less than 5 min, with the most probable value 1–2 min. The distribution of magnitude (at 38 MHz) peaks at 0.2–0.3 dB, and most occur between 1200 and 1800 local magnetic time with the mode at 1500–1600. Their spatial extent is 50–100 km. On present evidence these events are distinct from the larger spikes typical of the night sector both at this site and at lower latitudes.





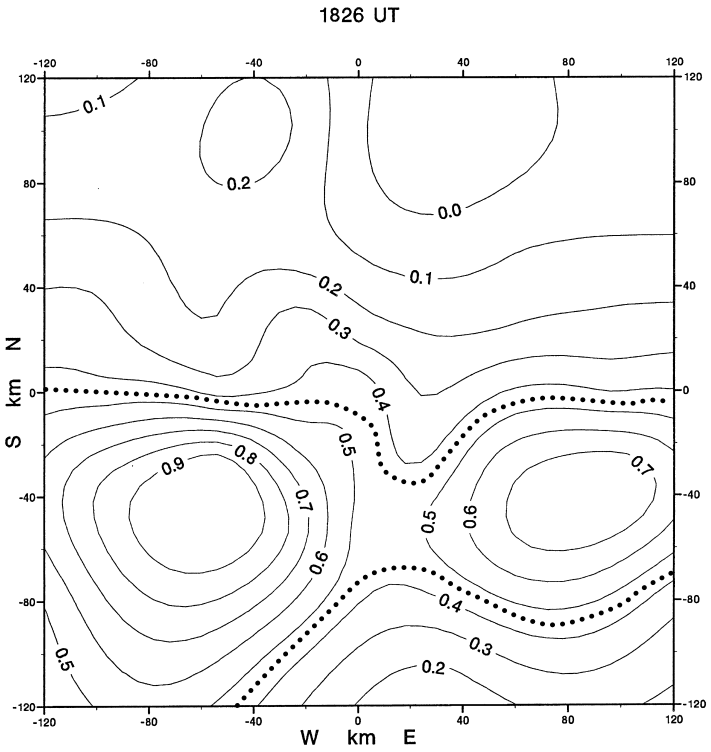


Figure 7.7. An absorption arc observed at Kilpisjärvi ($L = 5.9$) on 11 April 1995, at 1745, 1817, and 1826 UT. Each picture is a 1-min average and the dotted lines show the absorption arc defined by absorption half that at the peak. There is marked spatial structure along the arc. Each picture is 240 km on the side, and a height of 90 km is assumed. The feature was visible for an hour, and it drifted equatorward at less than 10 m s^{-1} initially but then more rapidly at 130 m s^{-1}. At 1845 UT it was followed by a sharp-onset event.

Preceeding bays

Starting 1–1.5 h before the onset there may be observed a weak absorption event lasting 40–60 min (*pre-onset absorption* or the *preceding bay*), and there can be little doubt that this feature is in some way related to the main event which follows. Imaging riometers have identified the form of the preceeding bay as an arc, extended east–west but only 60–100 km wide north–south (Figure 7.7). The whole feature tends to be weak, and it contains embedded structure. The arc normally undergoes a slow equatorward drift, and, as we shall see (Section 7.2.4), there are cases in which a sharp-onset event appears to grow from it. It is possible that this absorption feature is connected with the luminous auroral arc.

Slowly varying events and pulsations

Then, in the morning sector, between about 0600 and 1200 local magnetic time, there occur *slowly varying events (SVAs)*. These last for an hour or two and are

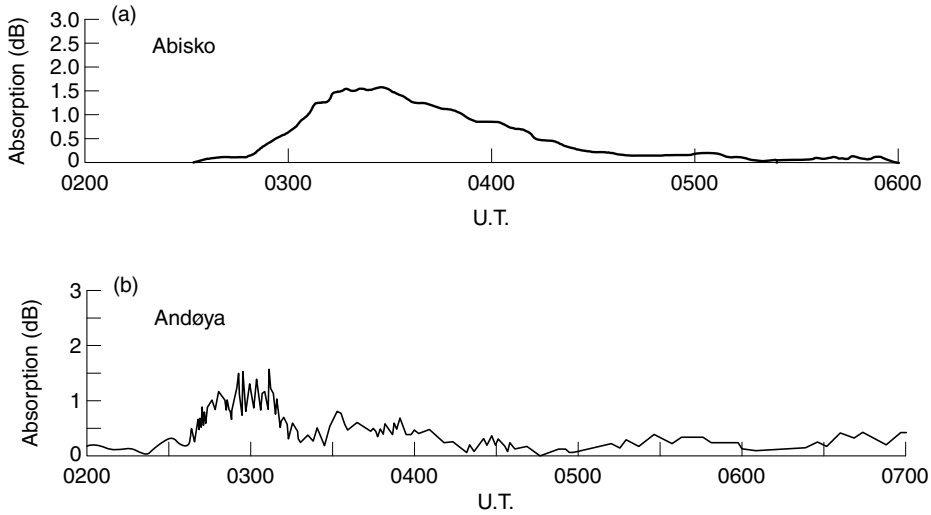


Figure 7.8. Events in the morning sector: (a) a slowly varying event at Abisko ($L = 5.6$), 23 March 1985 and (b) an event with pulsations at Andøya ($L = 6.2$), 23 August 1985 (Reprinted from J. K. Hargreaves and T. Devlin. *J. Atmos. Terr. Phys.* **52**, 193, copyright 1990, with permission from Elsevier Science.)

smooth, with little structure (Figure 7.8(a)). Some of these are modulated with quasi-periodic pulsations having a period of several minutes (Figure 7.8b), in the range of Pc4 and Pc5 (Section 2.5.6). The SVA is spatially more widespread than the spike and the preceeding bay.

Relativistic electron-precipitation events

As long ago as 1965 it was realized that some of the absorption events affecting radio circuits (particularly in VHF forward-scatter communications) were due to electrons of unusually high energy and relativistic speed (Bailey, 1968). *Relativistic electron precipitation (REP)* is a daytime phenomenon, and more events are observed at the equinoxes than at the solstices. The events may be intense, and they are geographically widespread according to the reports of the 1960s.

Since a riometer does not determine the height of the absorption, it is not immediately apparent which of the events detected are in the REP category, but simultaneous incoherent-scatter measurements have shown that, in some cases, the absorption was indeed at unusually low altitude (Collis *et al.*, 1996), and these are almost certainly due to relativistic electrons. (An electron of energy 100 keV travels at just over half the speed of light and one of energy 500 keV travels at $0.86c$. Electrons more energetic than 250 keV penetrate below 67 km and produce maximum ionization at heights below 75 km – Figure 2.26.) In at least some of these cases the event is confined to a small area in the D region, less than 100 km north–south though more extended east–west. Figure 7.9 is an example. In this

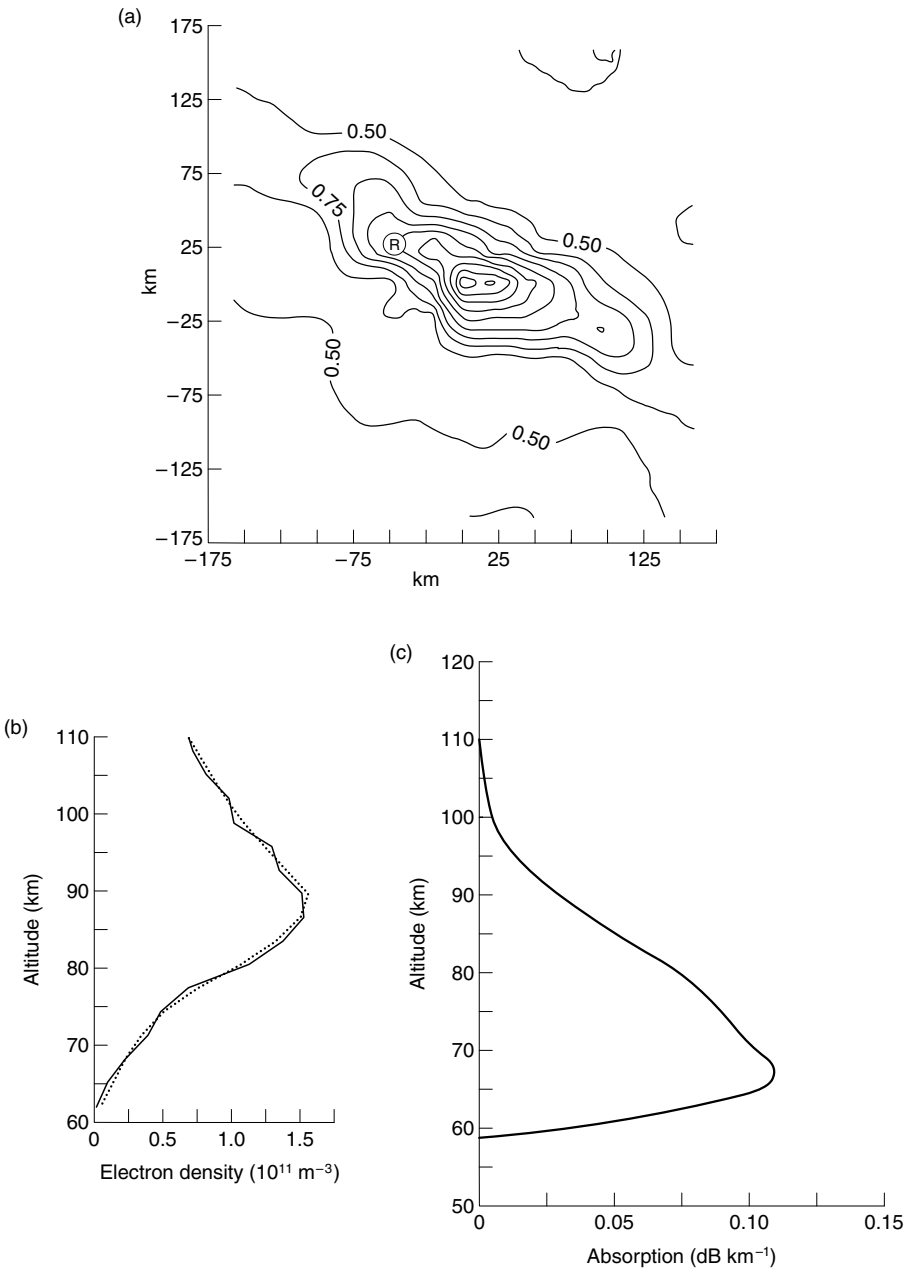


Figure 7.9. Properties of a co-rotating daytime event observed at 1204 UT on 1 March 1995 by an imaging riometer and the EISCAT radar simultaneously. (a) The spatial structure, assuming an altitude of 90 km. The contours are in decibels and the time resolution is 2 min. The radar beam intersected the event at R, somewhat away from the maximum of 3.5 dB. (b) The vertical profile of electron density, at 1-min resolution. (c) The vertical profile of the incremental absorption computed from (b). The electron-density peak below 90 km and the absorption peak below 70 km identify this as an event due to electron precipitation of unusually high energy. (P. N. Collis *et al.*, *Ann. Geophysicae* **14**, 1305, 1996, copyright notice of Springer-Verlag.)

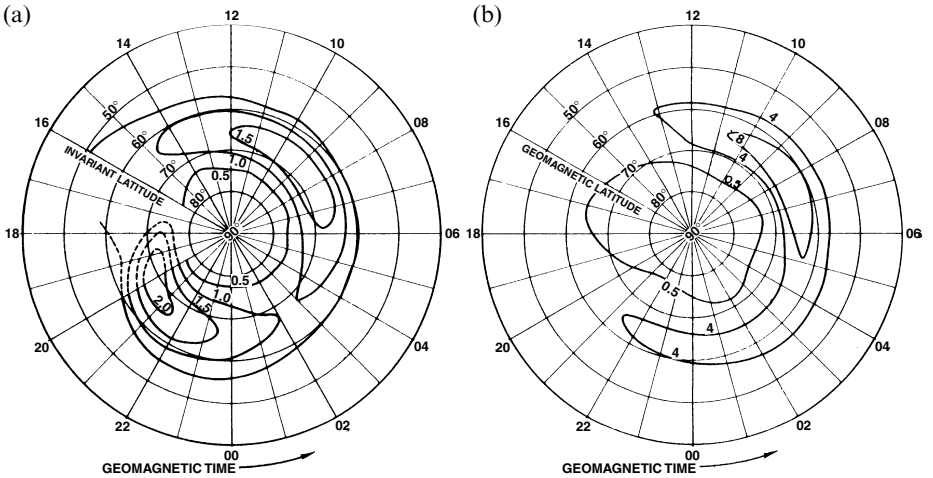


Figure 7.10. (a) The median intensity of AA events in decibels at 30 MHz. (Reprinted from J. K. Hargreaves and F. C. Cowley, *Planet. Space Sci.* **15**, 1571, copyright 1967, with permission from Elsevier Science.) (b) the percentage occurrence of 30-MHz absorption exceeding 1 dB. (After T. R. Hartz *et al.*, *Can. J. Phys.* **41**, 581, 1963.) The diagrams differ because the night events are shorter than the day events.

event the electron density due to electron precipitation peaked below 90 km, and the absorption peaked at 67 km. These are unusually low altitudes for an auroral absorption event (see Section 7.2.6).

It is interesting to note the similarity in the size and shape of some of these types of event, even when they are seen in different circumstances. This suggests that they have a common underlying physical cause.

7.2.3 General statistics in space and time

Latitude and longitude distributions

Two versions of the overall global occurrence of AA are shown in Figure 7.10 with respect to magnetic latitude and time. Most obvious in the second diagram is the morning peak around 0600–1000 magnetic time, where 1 dB (at 30 MHz) is exceeded for 8% of the time. This does not mean that AA is only a daytime phenomenon, however. We have already described some night-time events, and there is just as much absorption activity in the night as there is in the day sector, a fact that shows up more clearly in the first diagram which plots the median intensity of those events which happen to peak at a given time and latitude. The daytime events dominate in the other kind of statistics because they tend to be of longer duration. However, the night events can be just as intense. There is a deep minimum in the pattern of occurrence at around 1600–1700 magnetic time.

These distributions reveal a zonal phenomenon, having a maximum at about 67° geomagnetic latitude (corresponding to $L = 7$), though the details vary somewhat. Hartz *et al.* (1963) found the maximum at 67° in Canada, but Holt *et al.* (1961) found it at 62° in Norway. Using data from Canada and conjugate stations in the Antarctic, Hargreaves and Cowley (1967a) found small daily variations in the latitude of the maximum, in particular a decrease of 2° or 3° over the few hours up to about midnight, a recovery to 67°–68° during the morning, and a further decrease after noon. The latitudinal distribution of absorption may be approximated by a Gaussian curve,

$$A_m = A_0 \exp\left(-\frac{(\lambda - \lambda_0)^2}{2\sigma^2}\right) \tag{7.2}$$

where A_m is the median absorption, λ the invariant latitude, and σ the half width (or "standard deviation") of the absorption zone. The half width is several degrees: for example 4.5° (Hartz *et al.*, 1963) or 3.7° (Holt *et al.*, 1961).

Despite its daily variation, it is clear that the absorption zone is not the same as the auroral oval defined by the occurrence of luminous aurorae, but corresponds to the more circular zone – the "outer zone" – discussed in Section 6.3.5. The luminous oval and the absorption zone coincide (or at least are very close together) near midnight, but the absorption zone is more circular than the oval and lies at lower latitude on the day side. It is instructive to compare the incidence of AA in Figure 7.10 with the distribution of energetic electron precipitation observed from satellites in Figure 6.6.

The spatial extent

The horizontal extent of individual absorption events in kilometers is obviously an important matter practically as well as scientifically. Should they be very small, their effect on HF propagation could be reduced by space-diversity reception. If, on the other hand, they should blanket very large areas, it is difficult to see what could be done. (The same is the case with the polar-cap absorption due to protons – Section 7.3). Various measurements using groups of wide-beam riometers (Table 7.1) do not agree with each other completely, but the general indication is that the events cover a few hundred kilometers. Some reports suggest that there is a degree of elongation in the east–west direction, but other studies have come out in favor of more or less circular patches of absorption.

The results of Table 7.1 come from the earlier period of observations, and we must remember that the broad beam of the antenna prevents structure smaller than about 100 km being detected. As pointed out in Section 7.2.2, work using narrower beams is finding events narrower than 100 km. Spike events, which are only some tens of kilometers across, tend to appear in isolation and their magnitude is considerably underestimated in broad-beam measurements. However, since they occur for such a short time, they will not have much effect on the general

Table 7.1. *Spatial properties of Auroral Absorption (Hargreaves, 1969)*

Author	L	Data	Values correlated	Results	Shape of correlation pattern
Little and Leinbach (1958)	5.5 5.5–9	3 days, summer 1 month, March	Hourly values Hourly values	Regions at least 200 km N–S and 90 km E–W Day: $\rho = 0.57$ at 800 km N–S Night: $\rho = 0.43$ at 800 km N–S	
Holt *et al.* (1961)	6	12 selected periods	?	$\rho = 0.5$ at 380 km	Circular
Kavadas (1961)	4	?	?	High correlation over 10 km N–S	
Jelly *et al.* (1961)	4–8	?	—	Can be similar over 380 km or different over 35 km, N–S	
Leinbach and Basler (1963)	5.5	49 days, January–March	Hourly values	$\rho = 0.70$ at 250 km N–S $\rho = 0.74$ at 800 km E–W	Elliptical
Little *et al.* (1965)	4	Daytime events, 54 days, December–February	2-min values if absorption $\geqslant$0.3 dB	$\rho = 0.5$ at 650 km N–S and $\rho =$ 700 km E–W	Circular
Parthasarathy and Berkey (1965)	5.5	Sudden-onset events	Peak absorption	$\rho = 0.26$ at 250 km N–S (90 events) $\rho = 0.41$ at 800 km E–W (35 events)	Elliptical
Ansari (1965)	5.5	Slowly varying events, about 60 examples	Peak absorption	Fair agreement over 350 km N–S	
Berkey (1968)	5.5	Disturbed nights	? $\geqslant$0.5 dB	$\rho = 0.9$ at 20 km N–S	
Ecklund and Hargreaves (1968)	4	17 months, August–January	Hourly values if $\geqslant$0.3 dB	Night: $\rho = 0.5$ at 750 km E–W and 155 or 465 km[a] N–S	Elliptical
				Day: $\rho = 0.5$ at 365 km E–W and 170 or 300 km[a] N–S	Elliptical or circular
Bewersdorff *et al.* (1968)	5–7	Slowly varying events, four examples	–	Little variation over 300–400 km E–W	
Hargreaves and Ecklund (1968)	7	12 months	Hourly values if $\geqslant$0.3 dB	Night: $\rho = 0.5$ at 160 km Day: $\rho = 0.5$ at 250 km	Circular Circular

Note:

[a] Measurements poleward and equatorward of the central station, respectively.

statistics. At other times, particularly during the longer-lived activity, whether by night or by day, the smaller structures are only a component of the total distribution, and the results of Table 7.1 still have significance in showing that, even when finer structure is present, at least some part of the absorption extends for 200–300 km.

Durations

It is not as easy as one might imagine to determine the durations of individual absorption events. Some events are isolated and thus easily recognized, whereas others run into each other and it might not be obvious whether any such case should be described as one long event or several short ones. Furthermore, many events fade away rather gradually, so the end is not always clear. (Onsets tend to be sharper.)

Some values and the variation of duration from event to event are indicated by Figure 7.11, which shows the relative distribution of duration in the day and night groups at one station near the occurrence maximum and one just equatorward of it. Note that the durations are shorter at the higher L value, and are shorter by night than they are by day at both sites (Table 7.2). The groups of day and night events are specifed in UT. Add 2.5 and 0.5 h to get local magnetic time at Kiruna and Siglufjordür, respectively.

7.2.4 Dynamics

The dynamic nature of AA events is a property that is often not appreciated. The movements have been investigated using chains or groups of geographically separated riometer stations, supplemented more recently by imaging riometers. The results reveal a good degree of consistency in the movement of events of a given type, implying that the motion has some physical significance – even if we are not yet sure what that significance is! Some examples are given below.

The onset and main event in the night sector

The sharp-onset event occurs in the pre-midnight sector. As pointed out above, it may but need not include a spike, and it may appear as a more gradual onset at some latitudes. The onset of this type of event usually appears first at an L value between 5 and 6, which is somewhat equatorward of the statistical absorption maximum (at $L \sim 7$). From that latitude it spreads both poleward and equatorward. The poleward section is the more often observed; the velocity is between 0.5 and 3 km s^{-1} in most cases. There is a clear demarkation between the poleward and the equatorward sections, as may be seen in Figure 7.12(b).

However, this is not the case for absorption peaks subsequent to the onset (Figure 7.12(a)): they can move in either direction, rather more than half moving equatorward. In the example of Figure 7.6 the event arrives over Kilpisjärvi

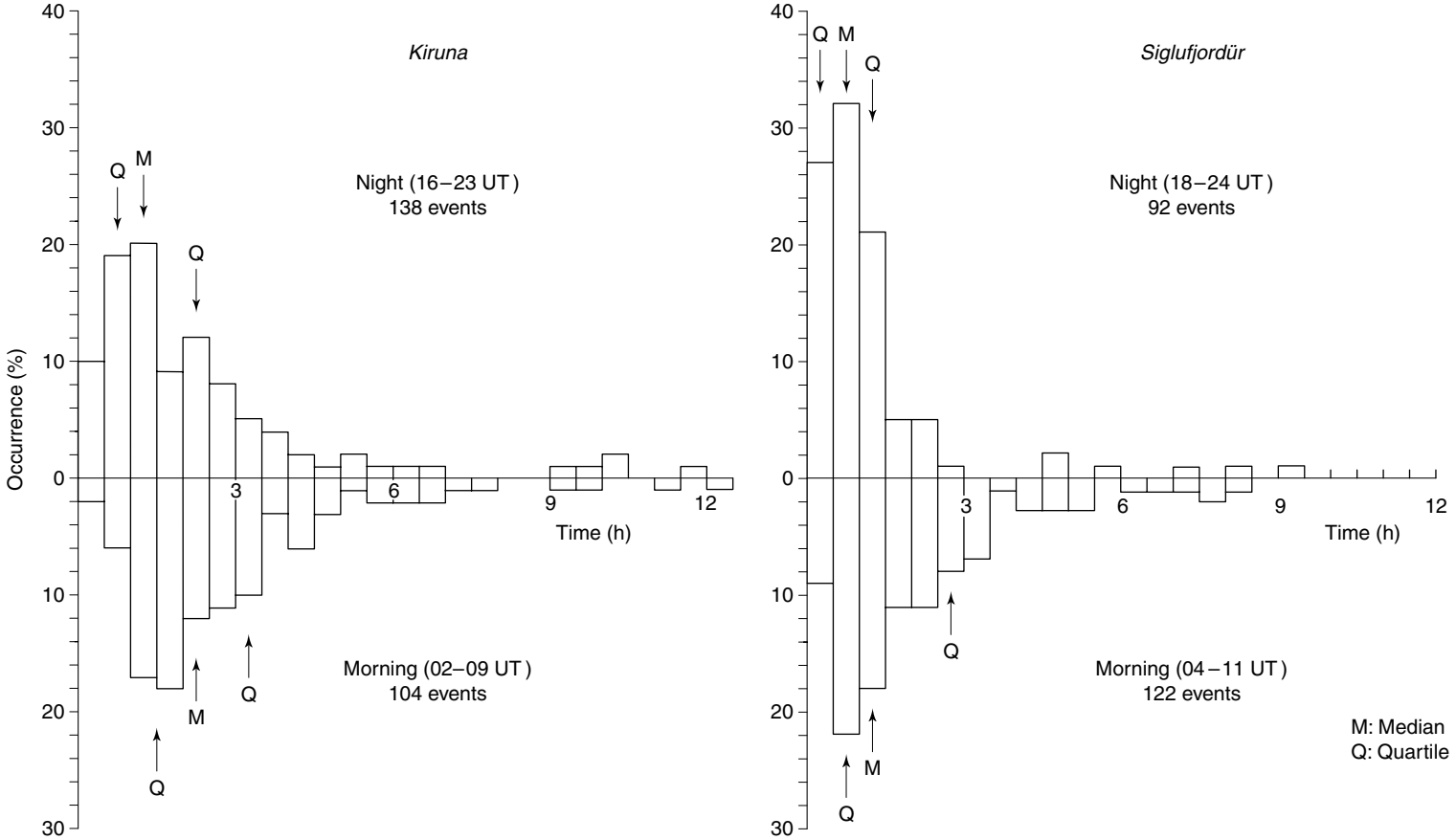


Figure 7.11. Durations of events starting in the night and morning sectors at Kiruna ($L = 5.4$) and Siglufjordür ($L = 6.9$).

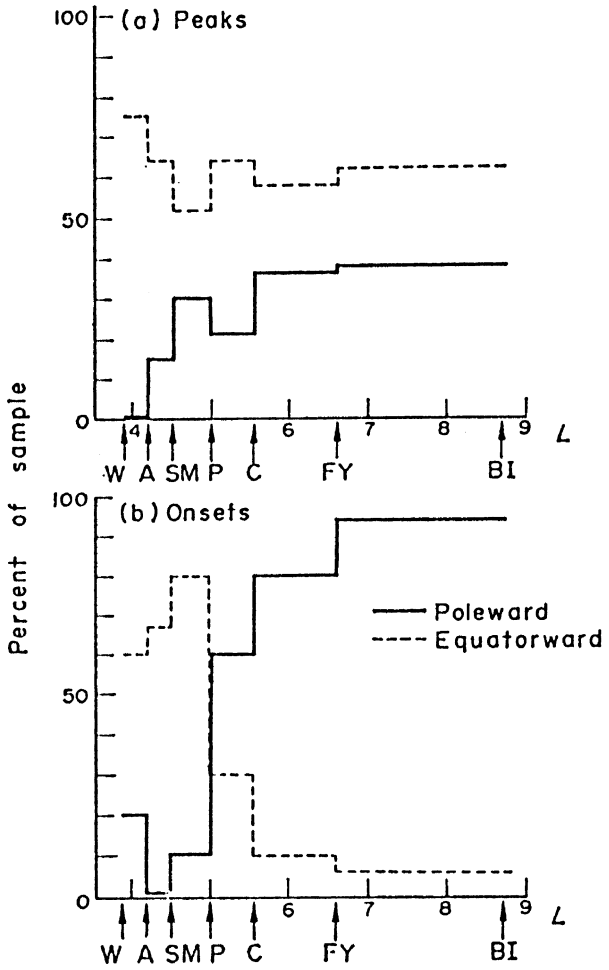


Figure 7.12. Relative frequencies of poleward and equatorward movements along a meridian through Alaska: (a) peaks, and (b) onsets. At all latitudes most of the peaks move equatorward, whereas the onset tends to move equatorward at $L < 5$ but poleward at $L > 6$. (Reprinted from J. K. Hargreaves. *Planet. Space Sci.* **22**, 1427, copyright 1974, with permission from Elsevier Science.)

Table 7.2. *Medians and quartiles of the distributions of durations of events at Kiruna ($L = 5.4$) and Siglufjordür ($L = 6.9$) (durations are given in hours)*

	Kiruna		Siglufjordür	
	Day	Night	Day	Night
Lower quartile	1.5	0.7	0.8	0.3
Median	2.2	1.5	1.5	0.9
Upper quartile	3.5	2.9	2.9	1.3

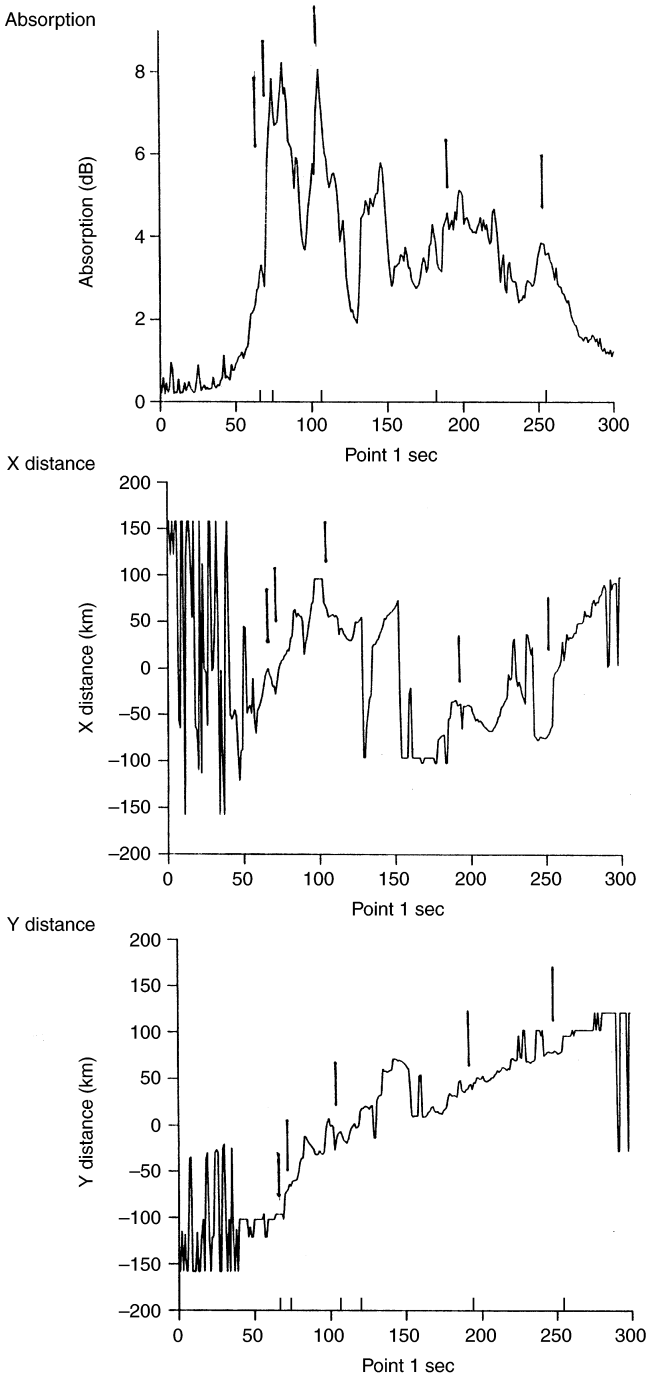


Figure 7.13. Maximum absorption, and distances in the X (west–east) and Y (south–north) directions of the location of the maximum from overhead at Kilpisjärvi, for the spike event at 2046–2051 UT on 6 October 1994 (see Figure 7.3). The time resolution is 1 s. Despite the time structure revealed at this resolution and erratic movement west–east, the poleward progression is remarkably persistent. These features are typical of night-time spike events.

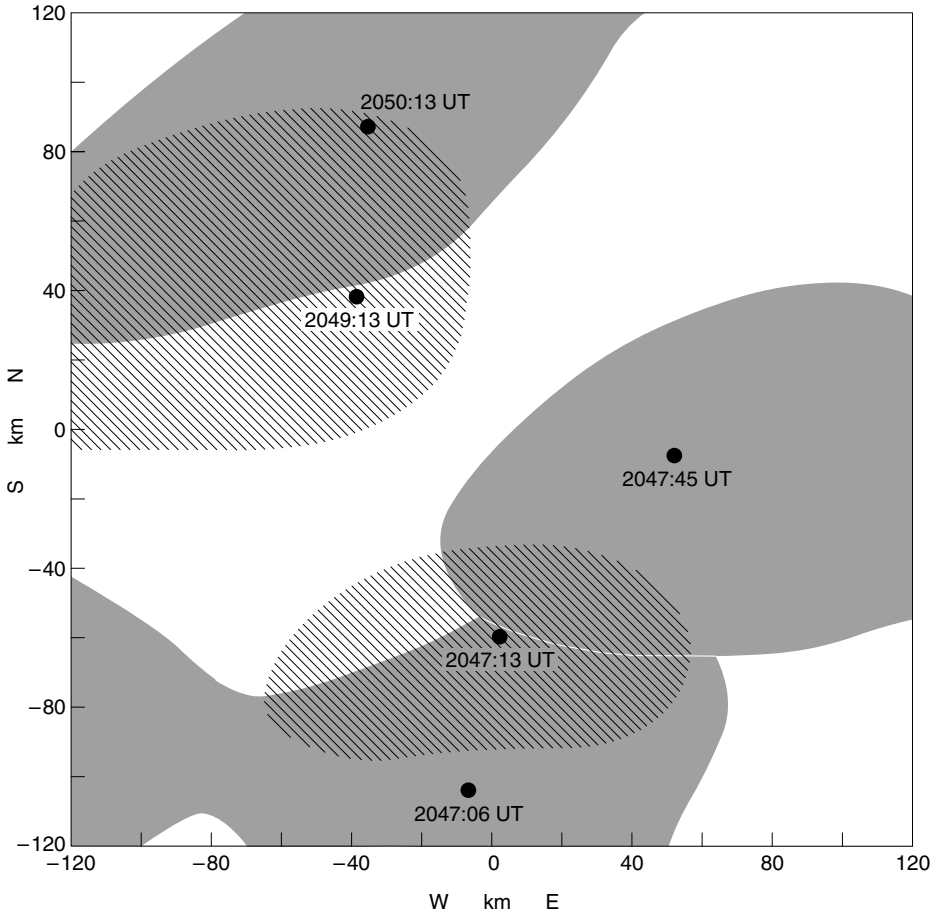


Figure 7.14. The spike event of Figure 7.13 at five selected times. The absorption was maximum at the points marked by black circles, and the shaded areas show where the absorption was greater than half the maximum. Note the tendency towards an elliptical shape.

($L = 5.9$) from the north and peaks almost overhead, but then moves off to the west. It is not clear whether this is typical, but previous observations (Hargreaves, 1970) using wide-beam riometers over a 250-km baseline at $L = 7$ have revealed a westward component in the night sector up to about 2 h after midnight.

When an event begins with a spike, this invariably moves poleward, as in the example from the Kilpisjärvi imaging riometer in Figure 7.13. This details the spike event at 2046–2051 UT in Figure 7.3(b), and shows the value of the maximum absorption and its location within the field of view, all at 1-s resolution. Note that the magnitude of the absorption varies with quasi-period 30–60 s. East–west motions are rather irregular, but there is a poleward progression overall. This is typical of spikes occurring at the beginning of a night event. Figure

7.14 shows the position of the absorption patch at five selected times (which are also marked on Figure 7.13). The maximum moved by 200 km in just over 3 min, an average speed of 1 km s^{-1}, though the speed was greater to begin with. The dimensions of the patch are changing, but the tendency towards east–west extension is maintained.

Motions on the global scale

The onset also propagates eastward and westward from its first appearence in the night sector. There is some variation among published results, partly due, no doubt, to actual variability in the movements from one instance to another, and perhaps also to observational selection. For stations near the centre of the absorption zone, observations between stations separated by some thousands of kilometers (for example over 90° longitude) indicate median speeds of about 4° longitude min^{-1} (or 2.8 km s^{-1}), eastward between midnight and 1400 LT, westward otherwise (Hargreaves, 1967; Pudovkin *et al.*, 1968; Jelly, 1970). The figure of 4° min^{-1} is for specific features recognized at both the stations. If absorption events are compared without consideration of form, the median speed comes out smaller by a factor of three. Hajkowicz (1990) found westward speeds of 2.7–4.5 km s^{-1} for pre-midnight sudden onsets at L values of 5.2–6.1.

A simple model of the longitudinal movements at $L \sim 7$ is given in Figure 7.15. From its first appearance near midnight it takes (on average) about 20 min for an onset to travel 5 h of local time and 30–40 min to reach the morning sector. The eastward and westward sections are not alternatives; observations have verified that they occur simultaneously. It is generally thought that the energetic electrons which are precipitated by day actually originated in the night sector and then drifted eastward as particles trapped in the geomagnetic field (Section 2.3.4 and Figure 2.14). This cannot explain the westward motion before midnight, which is presumably governed by other factors in the magnetospheric tail.

Combining results from studies of latitudinal and longitudinal movements gives the overall global picture of Figure 7.16. This makes no attempt to distinguish among different types of event.

The most comprehensive investigation of absorption movements on the global scale was performed by Berkey *et al.* (1974), who analyzed 60 substorm events at 40 riometer stations. Some of the main points, which confirmed the earlier work and added some new results, are as follows.

(a) The activity most frequently begins near midnight.
(b) The onset is earlier and at lower latitude when the level of magnetic disturbance is greater.
(c) The longitudinal velocity is in the range 0.7–7 km s^{-1}.
(d) The westward part of the expansion (usually seen before midnight) sometimes follows the auroral zone (i.e. the outer zone) and sometimes follows the auroral oval.

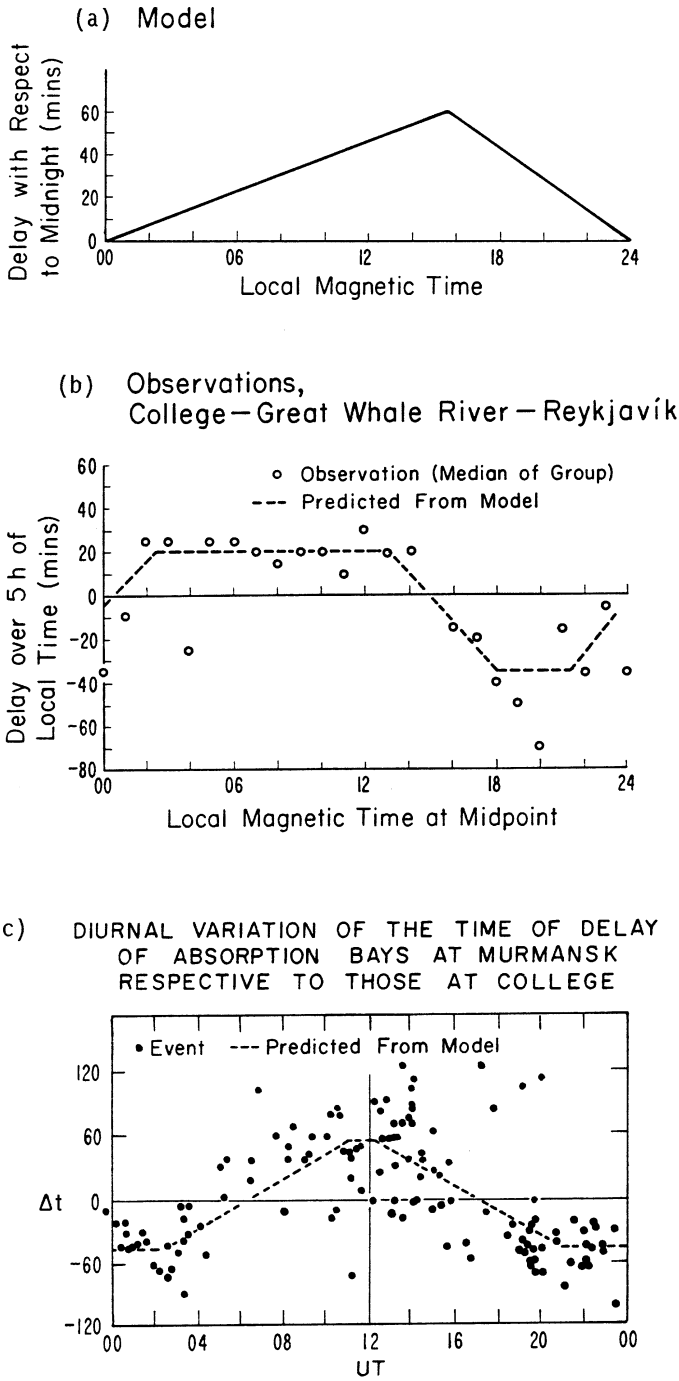


Figure 7.15. Longitudinal time delays. (J. K. Hargreaves. *Proc. Inst. Electr. Electron. Engineers* **57**, 1348, 1969a, © 1969 IEEE.) (a) A simple model of delay with respect to an onset at midnight. (b) The delay over 5 h of local time, compared with observations from Hargreaves (1967), the basis of the model. (c) Time delays between events at College and Murmansk (Pudovkin *et al.*, 1968) compared with predictions from the model.

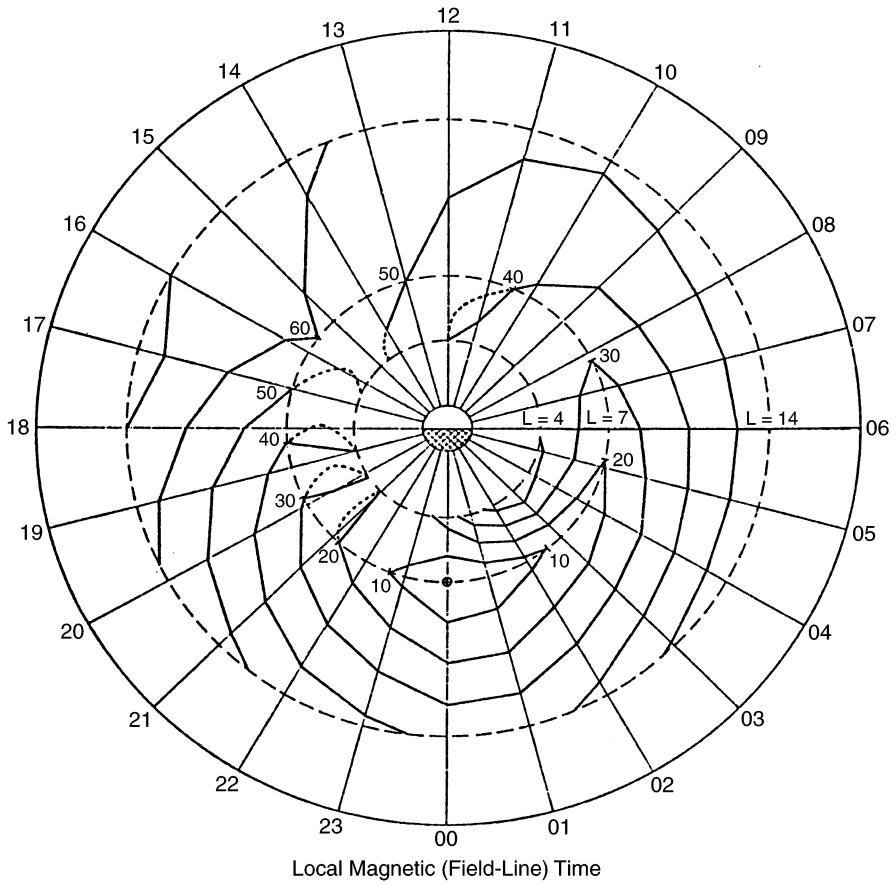


Figure 7.16. The progression of the onset of absorption projected onto the equatorial plane (assuming a dipole field). The wavefronts are drawn at 10-min intervals. (Reprinted from J. K. Hargreaves. *J. Atmos. Terr. Phys.* **30**, 1461, copyright 1968, with permission from Elsevier Science.)

(e) The speed of the westward expansion is about 1 km s^{-1} when it expands along the oval and about twice that when it expands along the zone.

(f) There is much variability among individual substorms.

(g) Some maps show morning activity (pre-noon) following a midnight onset, but with little or no activity between the midnight and day regions.

The drift of the pre-onset bay

The weak absorption bay that may precede an onset moves equatorward at a typical speed of a few hundred m s^{-1}. (See also Figure 7.27 later.) Many of these are so weak that they may be detected only by the practiced eye, but Figure 7.17 shows one that was unusually strong.

That bay was clearly seen in the sectors of Finland, Sweden/Norway, and

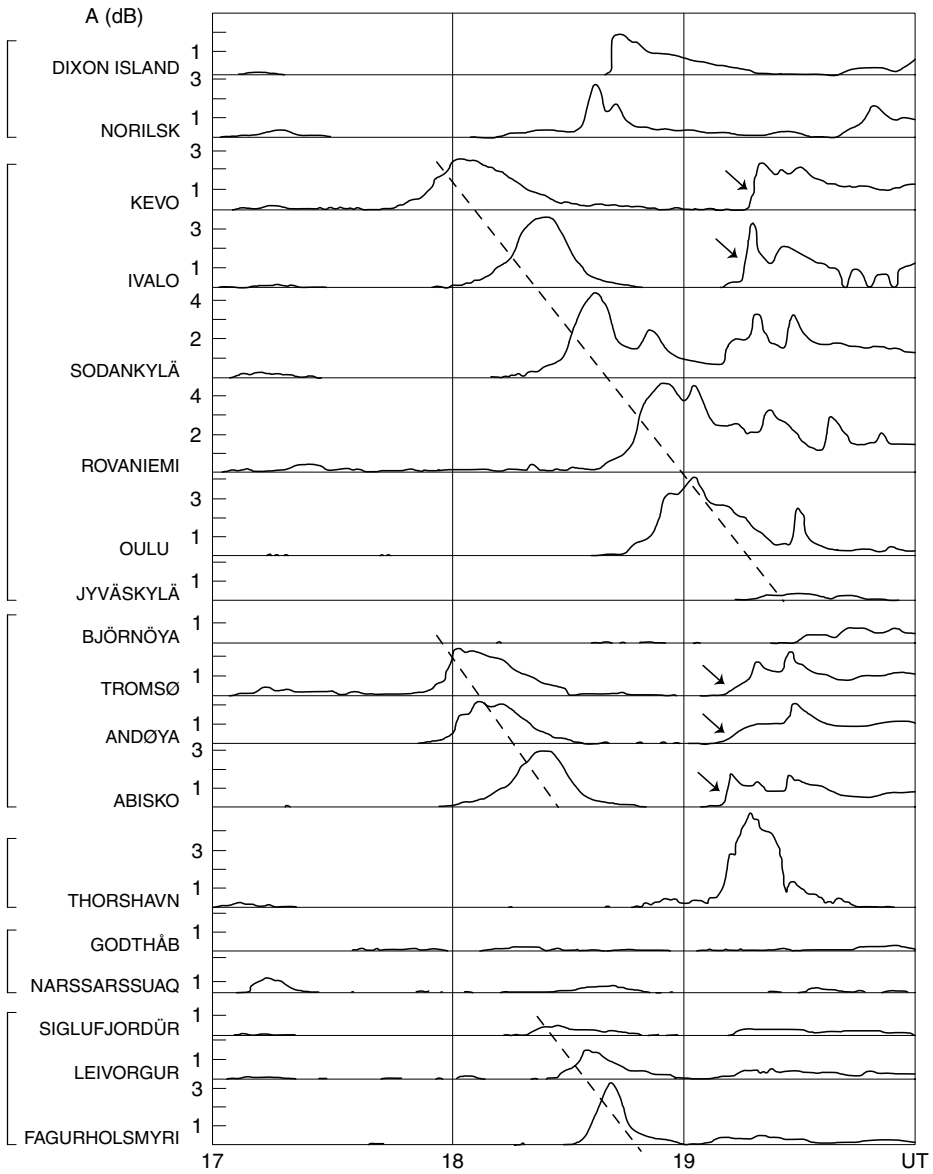


Figure 7.17. Equatorward motion of a bay preceeding an onset in the European sector on 4 May 1977. The diagram includes chains of stations at various longitudes, and the motion is clearly seen in the data from Finland, Norway/Sweden and Iceland. The related onset which followed is marked by arrows. (Reprinted from H. Ranta *et al*., *Planet. Space Sci.* **29**, 1287, copyright 1981, with permission from Elsevier Science.)

Iceland. It turns out that the imaging riometer is able to observe these moving arcs in greater detail (as in Figure 7.7). The speed is not always uniform, and the arc may fade and strengthen during its passage across the field of view. It is tempting to relate the movement to that of auroral arcs (Section 6.4.2), which is also equatorward in most cases.

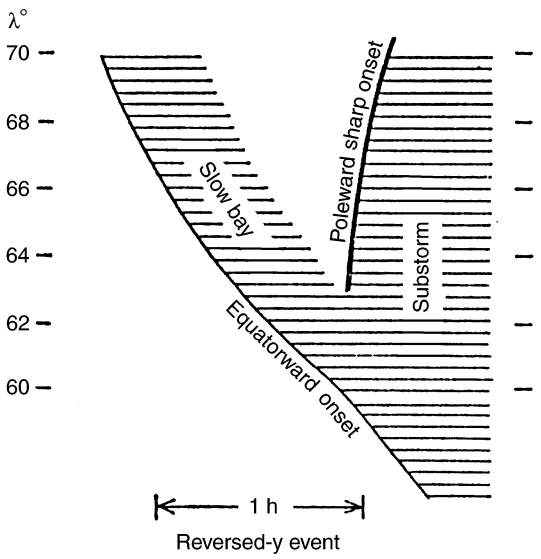


Figure 7.18. The connection between preceeding bay and sharp onset idealized as a "reversed-y" event. (Reprinted from J. K. Hargreaves *et al. Planet. Space Sci.* **23**, 905, copyright 1975, with permission from Elsevier Science.)

The relation between the bay and the onset

Ranta *et al.* (1981) have studied the incidence of these bays in relation to the sharp-onset events which follow. Most of the bays occur between L values of 4 and 9, and individual examples cover between one and five or six L units. They have not been reported from the South Pole (at $L = 13$). In longitude they can extend more than 90°. The onset may be seen over a larger range of L, from 4 to 16 or more, and individuals have been observed to cover ten units of L. It can exceed 150° in longitude.

The onset often appears first at or near the eastern end of the preceeding bay, which means that, statistically, the bays occur earlier in the day (in the afternoon and evening sectors) than do the sharp onsets whose preference is for the hours just before and up to midnight. The event following the bay of Figure 7.17 was observed in the sectors of Finland and Sweden/Norway, and exhibited poleward motion.

The relationship between bay and sharp-onset event may be summarized as the "reversed-y" event of Figure 7.18.

Figure 7.19 illustrates what appear to be the typical dynamics of a night-time event at $L = 5.9$, somewhat equatorward of the zone maximum. The record is from a wide-beam 38.2-MHz riometer but the movements have been identified using an imaging riometer at the same site. The spike had a rapid poleward movement, but an arc preceeding it and patches following drifted equatorward. The main event, which was widespread, came into view from the poleward side, but then drifted out of view to the west.

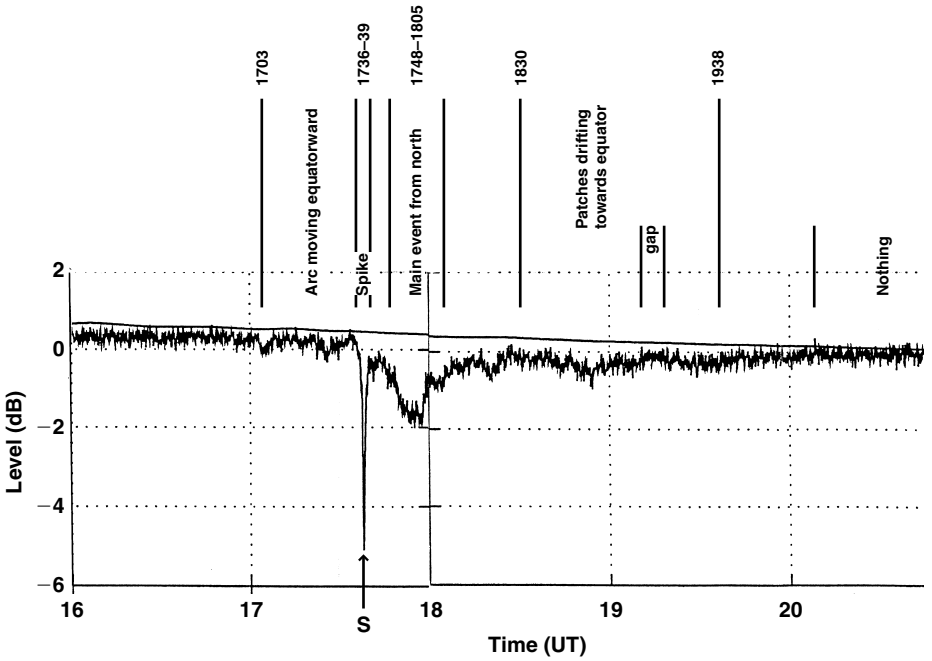


Figure 7.19. Night activity on 30 January 1995 observed with a wide-beam 38.2-MHz riometer at Kilpisjärvi, noting the main features and their movement. Excepting the spike, the dominant movement was equatorward.

The slowly varying event

The slowly varying event in the morning sector typically exhibits an eastward motion when it is observed by riometers 250 km apart (Hargreaves, 1970; Hargreaves and Berry, 1976). Each of these studies gave a median eastward speed just below 40 km min^{-1} (about 620 m s^{-1}), but with a large variation in individual cases. (Half the eastward speeds were between 20 and 80 km min^{-1}.) It will be noticed that these speeds are considerably below those determined from widely spaced observations.

Co-rotation

A tendency towards co-rotation has been noted in some events of the morning and dayside (Hargreaves *et al.*, 1994; Collis *et al.*, 1996). In particular, it appears that the spatially restricted, very energetic type of event described in Section 7.2.2. can remain virtually fixed with respect to the rotating Earth for a long period. The event shown in Figure 7.9 remained in the field of view for more than 1.5 h. According to radar measurements, the meridional F-region ion drift during the event varied from 100 m s^{-1} westward to approximately zero – in agreement with the motion of the absorption event. This clear example, taken with the evidence

from the previous paragraph, indicates that more than trapped-particle drift is involved in the longitudinal motion of auroral absorption.

7.2.5 The relation to geophysical activity, and predictions of auroral absorption

A relation to A_p

Auroral absorption (AA) is more frequent and more intense when geomagnetic activity is high. One might also expect that AA would be stronger and occur more often at times of high sunspot number, but that is not necessarily the case. In some solar cycles the magnetic activity does not rise and fall in the same way as the number of sunspots, and in such a case the AA is seen to go with the magnetic activity rather than with the sunspot number. We shall return to this point.

In the shorter term, it is possible to show relations between the absorption on single days and the A_p index (Section 2.5.4). For example, the probability of there being at least one event of at least 1 dB during a period of 24 h rises almost linearly with A_p, becoming virtually unity at $A_p > 15$ for a station at $L = 5.6$ (Figure 7.20(b)). The rate of occurrence is smaller at higher latitude (Figure 7.20(a)) but still increases approximately linearly with A_p. The average number of events per day also increases with A_p (Figure 7.20(c)), rising from one to three over the A_p range 8–25.

One observation that this association explains is the tendency for AA to be intense for several days at a time, often then followed by a week or more when it is very low. The pattern tends to repeat from one month to the next. This behavior just mirrors that of magnetic disturbance, and is due to the rotation of the Sun which carries active regions out of view after a few days and tends to bring them round again a month later.

The latitude of the absorption zone (Section 7.2.3) also shifts with the intensity of magnetic activity (Hargreaves, 1966). During that period of the day when AA is most significant, the latitude of the maximum (λ_0) decreases from approximately 70° to 66° as K_p increases from 0 to 5, and at values of 6 or 7 it may be as low as 60°. At the same time the half width of the absorption zone (σ) increases somewhat (from about 4° to 5.5°), with even greater broadening at the largest values of K_p.

A relation to HF radio propagation

This brings us to the topic of predictions. First, though, to get some idea of the importance of AA in HF propagation, we must consider the magnitude of the absorption involved. The intensity of the absorption, which is usually measured in decibels, depends inversely on the square of the radio frequency. In a

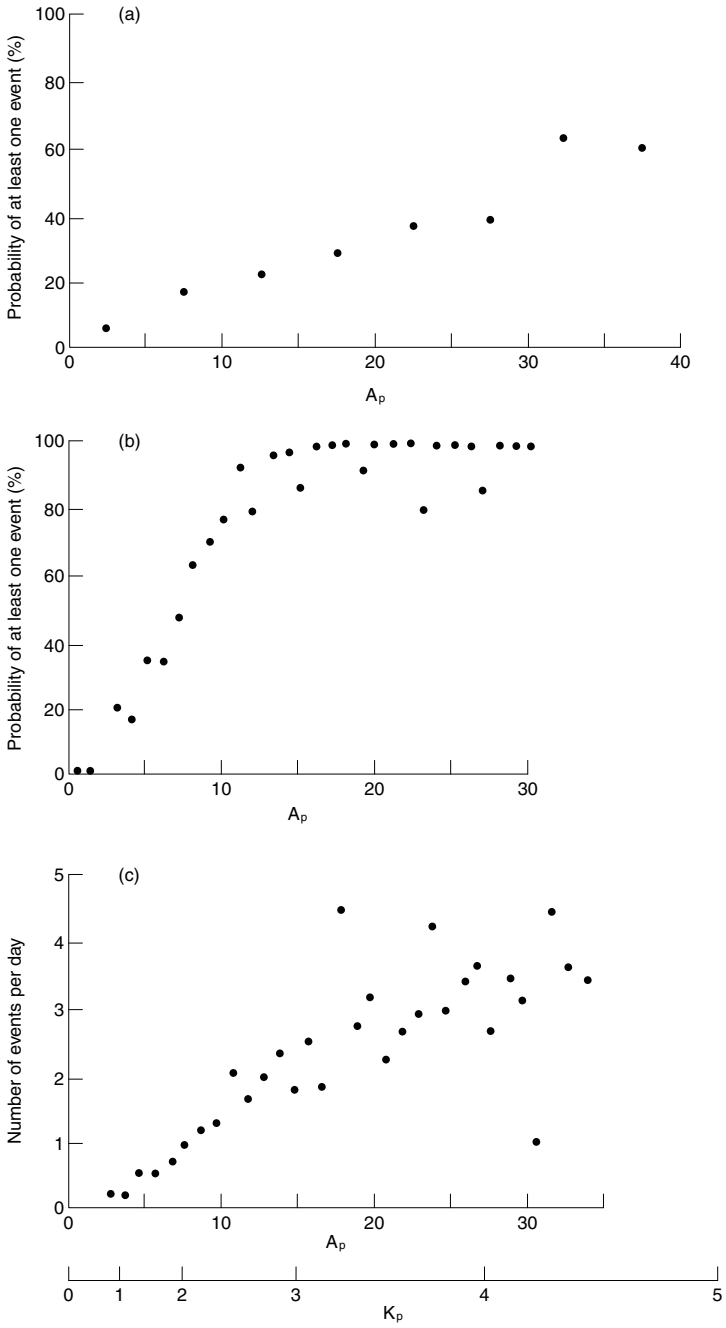


Figure 7.20. Effects of magnetic activity on the incidence of AA events. (a) The probability that at least one event of at least 0.3 dB at 51.4 MHz occurs at the South Pole ($L = 13$) within 24 h. This level is equivalent to 0.88 dB at 30 MHz. The data covered the three years 1990–1992 inclusive. (b) Similarly for 1-dB events at 30 MHz at Abisko ($L = 5.6$). The data covered the two years 1976 and 1977. (c) The average number of 30-MHz, 1-dB events per day at Abisko. (South Pole data from T. J. Rosenberg, private communication, and Abisko data compiled from Hargreaves *et al.*, *Report UAG-84*, 1982.)

radio-communication circuit it also depends on the geometry (specifically, the angle at which the ray passes through the D region). Nonetheless, in round numbers, if a 30-MHz zenithal riometer detects 1 dB absorption, an oblique HF path will suffer about 20 dB (Agy, 1970). Thus, absorption greater than 1 dB on a 30-MHz riometer is likely to be of practical significance to HF propagation (especially between 3 and 15 MHz), and the statistics regarding the occurrence of absorption are often presented in terms of the 1-dB level. At some latitudes there are many days per month when this level is reached or exceeded at least once.

AA predictions

Clearly, predictions of AA are going to be statistical in nature because the phenomenon is essentially sporadic; but at least there is a good base for the statistics because large quantities of riometer data are available. If the predictions are required for radio propagation, the task has two parts: first, to specify from existing riometer data the statistics regarding the occurrence of absorption and the effects of independent variables such as latitude, season, time of day, and solar and magnetic activity; and second, from propagation experiments to observe how the events detected by the riometer are related to circuit effects. Here we consider the matter of absorption statistics, in which there have been some useful developments. Relations to communications circuits are considered in Sections 8.2 and 8.4.

The representation of absorption statistics can be taken in two stages. First, having decided our significant absorption level – for instance, 1 dB at 30 MHz – we can then inspect the data from various riometer stations and count up the probability of 1 dB being exceeded as a function of the various external parameters. The probability that A dB will be exceeded is generally called $Q(A)$.

Calculation of Q(1)

Foppiano and Bradley (1985) published a formula (Table 7.3) for calculating $Q(1)$, based on an extensive study involving many sources and taking in data from several longitude sectors and different years. The formula is written as the sum of day and night contributions, each comprising the product of terms for the variation with magnetic latitude, time of day, solar activity, longitude, and season. The latitude variations are of Gaussian form (similar to Equation (7.2)) with the night peak at 67° and the day peak at 68° at low sunspot number. The time-of-day terms are also Gaussian, the night activity peaking at midnight and the day activity at 1000 local time – compare with Figure 7.10. The dependence on solar activity is expressed in a table, and there are empirical formulae for the dependences on longitude and season.

Some of these terms are better established than others. Some seasonal variation probably occurs (see Section 7.2.6), but the question of a longitudinal effect has not been investigated thoroughly. However, the greatest problem with the formula of Table 7.3 is in the assumed dependence on solar activity expressed by

Table 7.3. *The Foppiano–Bradley formula for Q(1)*

Total	$Q(1) = Q_{1d} + Q_{1s}$
Day component	$Q_{1d} = K_d d_\lambda d_T d_R d_\theta d_M$
Night component	$Q_{1s} = K_s s_\lambda s_T s_R s_\theta s_M$

λ = magnetic latitude; T = local time in hours; R = sunspot number; θ = geomagnetic longitude; M = month; K_d and K_s are constants

Latitude terms

$$d_\lambda = \exp[-(\lambda - \lambda_m)^2/(2\sigma_\lambda^2)] \qquad s_\lambda = \exp[-(\lambda - \lambda'_m)^2/(2\sigma'^2_\lambda)]$$

λ_m and λ'_m = geomagnetic latitudes of maxima for day and night components:

$$\lambda_m = 68(1 - 0.0004R) \qquad \text{for } R \leqslant 100;$$
$$\lambda_m = 65.28 \qquad \text{for } R \geqslant 100;$$
$$\lambda'_m = 67(1 - 0.0006R) + 0.3(1 + 0.012R)|t|$$

where $t = (T - 3)$ for $0 \leqslant T \leqslant 15$; $t = (T - 27)$ for $15 < T < 24$

σ_λ and σ'_λ = widths of latitude distributions for day and night components:

$$\sigma_\lambda = \sigma'_\lambda = 3(1 + 0.004R) \qquad \text{for } R \leqslant 100;$$
$$\sigma_\lambda = \sigma'_\lambda = 4.2 \qquad \text{for } R \geqslant 100.$$

Time-of-day terms

$$d_T = \exp[-(T - T_m)^2/(2\sigma_T^2)] \qquad s_T = \exp[-(T - T'_m)^2/2\sigma'^2_T]$$

T_m and T'_m = local times of maxima for day and night components:

$$T_m = 10(1 - 0.002R); \qquad T'_m = 0$$

σ_T and σ'_T = widths of time distributions for day and night components:

$$\sigma_T = \sigma'_T = 2.8$$

Solar-activity terms

$d_R = s_R = (1 + aR)$, the values of a being from the following table:

T (h)	00	02	04	06	08	10	12	14	16	18	20	22
a	0.0032	0.0025	0.0141	0.0048	0.0149	0.0146	0.0142	0.0090	0.0037	0.0156	0.0206	0.0092

Longitude terms

$$d_\theta = s_\theta = 0.58 - 0.42\sin[0.947(\theta + 85)] \qquad \text{for } 0° \leqslant \theta < 10°$$
$$= 0.16 \qquad \text{for } 10° \leqslant \theta < 80°$$
$$= 0.58 + 0.42\sin[1.80(\theta - 130)] \qquad \text{for } 80° \leqslant \theta < 180°$$
$$= 0.58 - 0.42\sin[0.947(\theta - 275)] \qquad \text{for } 180° \leqslant \theta < 360°$$

where θ = corrected geomagnetic longitude in °E

Seasonal terms

$$d_M = 1 - 0.3\sin(3.86\delta); \qquad s_m = 1$$

where δ = solar declination angle in degrees (positive in summer, negative in winter).

Constants

$$K_d = 21; \qquad K_s = 12$$

These give $Q(1)$ values in percent

Notes:

The use of *d* and *s* for the day and night components derives from Hartz and Brice's (1967) "drizzle" and "splash" terminology.

the sunspot number. In a subsequent study of absorption over Finland during a whole solar cycle (1972–1983), Hargreaves *et al.* (1987) concluded that the sunspot term of the formula was not very accurate, and they proposed an alternative based on the monthly mean value of A_p ($\bar{A}_p$):

$$Q(1) = (\bar{A}_p - 30\cos^2\lambda)\exp[-(\lambda - 65)^2/25]. \tag{7.3}$$

This formula gives the average $Q(1)$ over all times of day.

The significance of Equation (7.3) is that, over the long term, the absorption probability is proportional to the mean A_p ($\bar{A}_p$) above a threshold. The results may also be represented by a Gaussian variation with latitude in which the peak value (Q_0), the latitude of the peak (λ_0), and its width (σ) depend on $\bar{A}_p$ as in Table 7.4. Note in particular that, with increasing $\bar{A}_p$, the maximum probability increases linearly and the position of the maximum moves equatorward. (The foregoing analysis is based on observations covering only the equatorward side of the absorption zone.)

The log-normal distribution

The second stage is to consider the form of $Q(A)$. That is, if we can predict $Q(1)$, can we say what $Q(2)$ or $Q(0.5)$ will be? This means knowing the probability distribution for the occurrence of absorption. Foppiano and Bradley (1984) assumed a log-normal distribution for the occurrence of absorption; that is, that the logarithm of the absorption follows a normal distribution:

$$f(\log A)\,\mathrm{d}(\log A) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(\log A - \log A_m)^2}{2\sigma^2}\right)\mathrm{d}(\log A), \tag{7.4}$$

where A is the absorption in decibels, A_m is the median absorption ($\log A_m$ being the mean of $\log A$), and σ is the standard deviation of $\log A$. The probabilility of A being exceeded is then

$$Q(A) = \int_{\log A}^{\infty} f(\log A)\,\mathrm{d}(\log A). \tag{7.5}$$

Table 7.4. *Parameters of Gaussian curves fitted to latitudinal variations of Q(1) in the Finland sector*

$\bar{A}_p$	λ_0 (degrees)	σ (degrees)	Q_0 (%)
0–10	68.1	3.8	5.7
10–15	67.8	3.9	9.3
15–20	66.9	3.6	13.3
>20	65.6	3.6	17.4

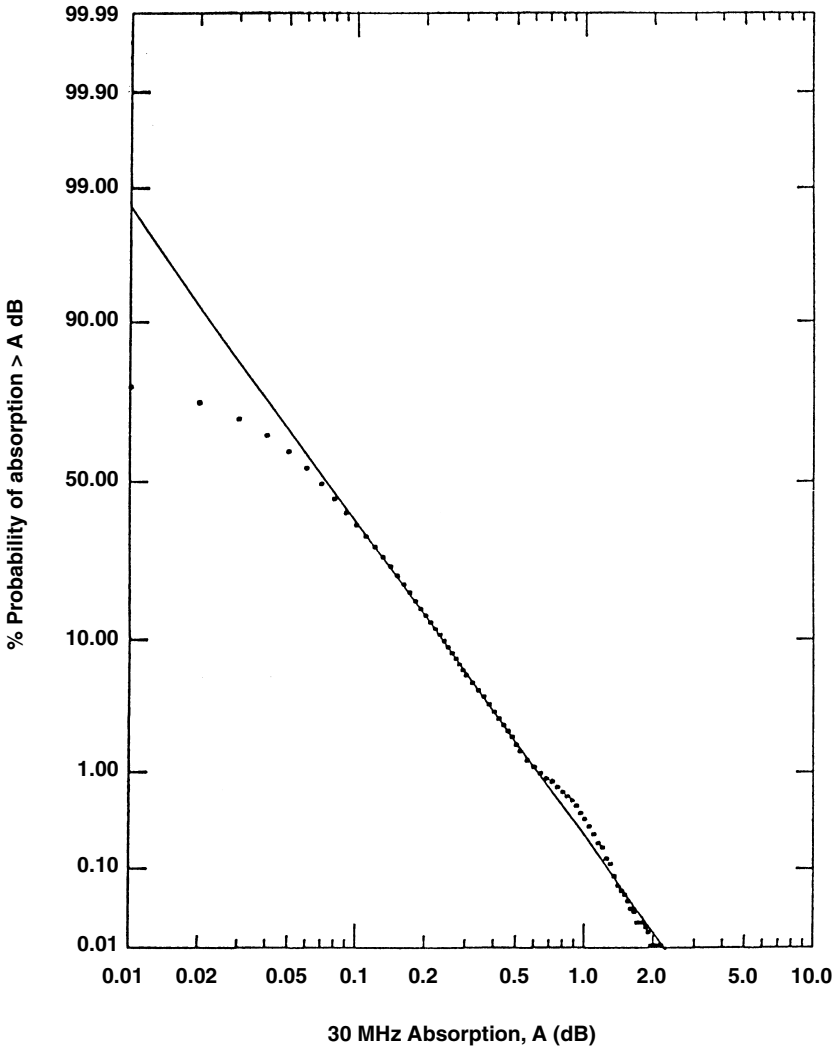


Figure 7.21. The log-normal distribution of $Q(A)$, South Pole, March 1982. (Data from T. J. Rosenberg, private communication.)

The cumulative distribution $Q(A)$ should come out as a straight line on log-probability paper. In most cases this appears to be so (Figure 7.21), at least if the range of absorption is restricted to values between a few tenths of a decibel and a few decibels – that is, to the range for which riometer data are accurate and most plentiful. (Whether very large or very small values obey the same distribution is not really known.)

The log-normal distribution is described by just two parameters, one of which can be $Q(1)$. There is something to be said, also, for using A_m and σ, the median and the standard deviation (the second giving the slope on the log-probability plot), since both of these appear explicitly in the formula. It is not wise to extrap-

olate the log-normal law below 0.1 or 0.2 dB. In most sets of riometer data these small values are much affected by any error in the quiet-day curve, and also there are theoretical reasons why the log-normal form cannot continue indefinately towards ever smaller values.

Predicting events

The foregoing approach aims to predict the likehood of a certain level of absorption at a given site if the level of geomagnetic disturbance (or solar activity) is known – the latter, of course, also being a quantity requiring prediction. No account is taken of the event aspect: the fact that the absorption occurs in bursts; that, once started, it is likely to continue for some time but that there are also long periods with no significant absorption at all.

Relatively little appears to have been done on prediction of AA using an event approach, though elements are implicit in some of the foregoing account. A comprehensive event description would specify the magnitude, duration, structure, etc., of which (apart from magnitude) the statistical approach takes no account. An event description would also include an element of short-term forecasting. One example, based on the data of Berkey *et al.* (1974), is shown in Table 7.5. Medians and deciles are given for the absorption at various local times for every 15 min after the onset of a substorm (Elkins, 1972). These are interesting for showing how the distribution of absorption develops in a substorm as a function of the local time, and also for what seems to be the first use of the log-normal distribution to describe absorption statistics. The actual magnitudes depend, of course, on the original selection of substorms. In this set the maximum absorption was found to be related to the AE index by the empirical formula

$$(\text{Absorption})_{\max} \sim 0.008(\text{AE})_{\max}, \tag{7.6}$$

but this is not of much help in a prediction because AE is not a predicted quantity. It is better to relate absorption to the daily index A_p, predictions of which are published a month in advance.

7.2.6 The wider geophysical significance of auroral absorption events

The immediate implication of auroral radio absorption for high-latitude propagation is simply the resulting loss of signal. However, since we know that the absorption is due to additional ionization in the lower ionosphere, which in turn is produced by energetic electrons entering the atmosphere from above, these events clearly have deeper implications. In this section we review some contributions of riometer studies to geophysical topics.

Table 7.5. *Medians and deciles of absorption at various local times during a substorm (Elkins, 1972)*

Substorm time	LT	Median (dB)	Upper decile (dB)	Lower decile (dB)
$T+15$	00	0.75	2.6	0.22
	03	0.70	2.2	0.22
	06	0.54	1.8	1.17
	09	0.28	1.25	0.066
	12	0.10	0.56	0.019
	15	(0.14)	(0.38)	(0.052)
	18	0.10	0.56	0.019
	21	0.37	1.5	0.090
$T+30$	00	0.94	3.7	0.24
	03	1.1	4.0	0.32
	06	1.1	4.5	0.30
	09	0.64	2.8	0.14
	12	0.42	1.7	0.10
	15	(0.20)	(0.78)	(0.052)
	18	(0.17)	(0.58)	(0.050)
	21	0.50	1.9	0.13
$T+45$	00	1.1	3.5	0.34
	03	1.3	4.0	0.43
	06	1.6	4.5	0.54
	09	1.4	6.0	0.32
	12	0.67	2.8	0.16
	15	0.28	1.3	0.064
	18	(0.20)	(0.84)	(0.049)
	21	0.44	1.8	0.11
$T+60$	00	1.0	3.2	0.30
	03	1.3	3.8	0.40
	06	1.6	4.5	0.56
	09	1.6	4.5	0.56
	12	1.1	3.6	0.35
	15	0.38	1.5	0.096
	18	(0.25)	(1.0)	(0.063)
	21	0.50	2.0	0.070

Notes:

1. Parentheses ( ) indicate values with large uncertainties due to the small sizes of statistical samples.
2. Time is to be interpreted as follows: local time "00" means the hours 0000–0259 and so forth.

Electron-density profiles

That there are, indeed, relationships between total absorption and electron density at various heights has been shown by direct measurements. Friedrich and Torkar (1983) collected electron-density data from rockets flown into absorption events and thus "calibrated" the riometer in terms of the electron-density profile from 70 to 110 km. This has recently been extended (Friedrich and Kirkwood, 2000) using electron densities from the EISCAT incoherent-scatter radar in Scandinavia (Figure 7.22). This comparison provides an estimate of the electron density at given height for a given intensity of AA, though with considerable scatter about the average. Some 50% of values lie within a factor of two of the average plotted.

Some of the scatter is no doubt due to real changes in the profile during events, and from one event to another. Figure 7.23 illustrates the changing electron-density profile during a morning event of the slowly varying type in which the height of the peak lifts as the event decays. (The growth was more complicated.)

Absorption profiles

Since the electron–neutral species collision frequency is known (for Equation (3.95)), the absorption profile may be computed from the electron-density profile. In most cases the computed and observed absorptions agree well enough to serve as confirmation that the reductions in signal recorded by the riometer are indeed due to non-deviative absorption.

The height of the absorbing layer and its thickness are of direct interest in HF propagation. The heights of absorption maxima computed from rocket profiles of electron density range over 90 to 95 km at night, but may be lower (75 km) by day (Hargreaves, 1969a). The calculated absorption peaked between 88 and 95 km in the event of Figure 7.23 which occurred during the early morning, and at 67 km in the hard, daytime event of Figure 7.9. The absorbing layer is quite thick: generally 15–20 km between points where the incremental absorption is half the maximum. About 80% of the total absorption is produced in this slab. The specific absorption coefficient increases downward, and the absorption peaks some 5–15 km (depending on the spectrum) lower than the electron density; as a rough guide it would be true to say that the absorption layer occurs in the underside of the electron-density layer, starting just below the peak.

Incoming electron fluxes

Since the AA event is due to precipitating electrons, the calibration can in principle extend back a stage further so that we may infer something about the intensity and spectra of the electrons precipitated during AA events. Proceedures for inverting the electron-density profile (Kirkwood, 1988; Hargreaves and Devlin, 1990; Osepian *et al.*, 1993) involve routines that give the rate of production from an assumed incoming electron spectrum (Section 2.6.1 and Figure 2.26), and by some means adjust the spectrum until the computed electron-density profile

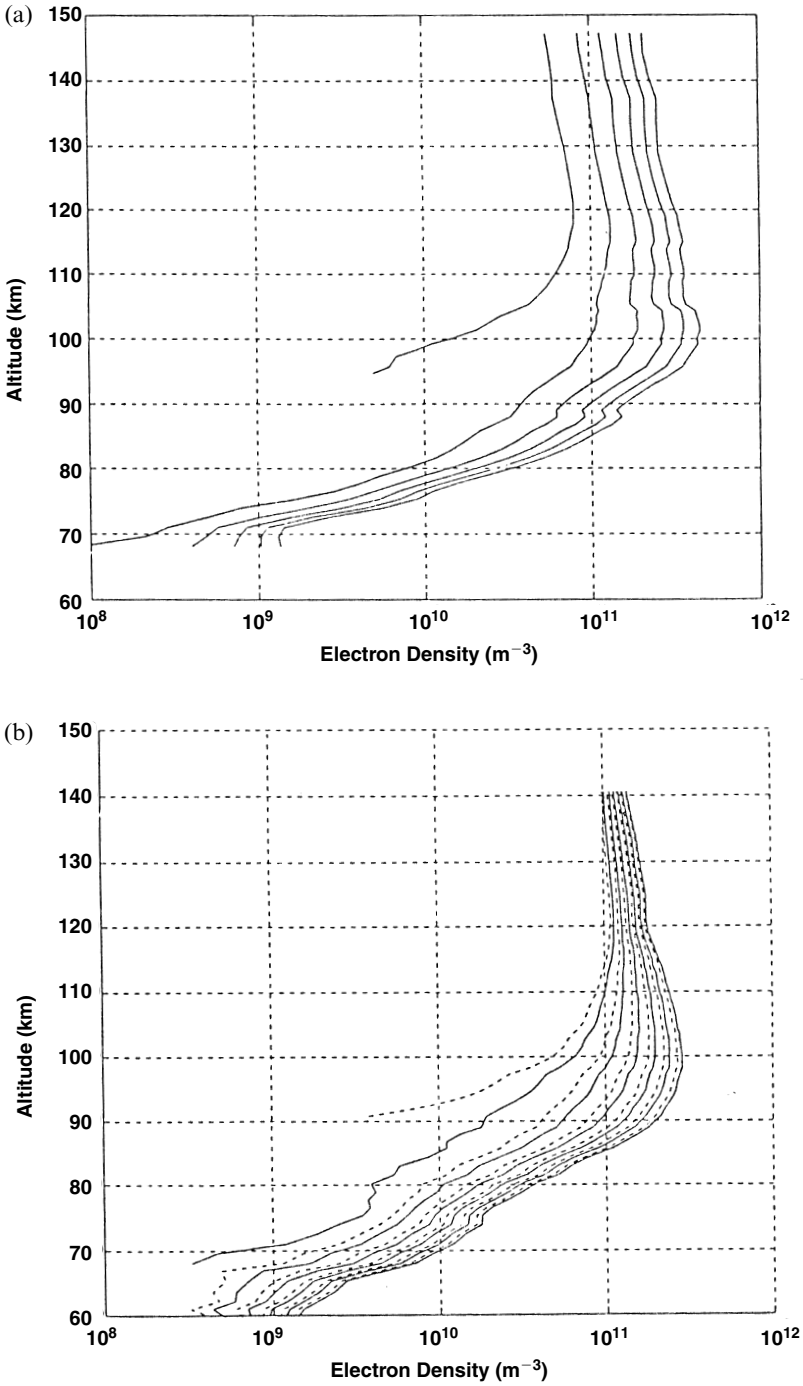


Figure 7.22. A riometer "calibration" against electron-density profiles measured by rockets and incoherent-scatter radar. (a) with the Sun below the horizon and (b) solar zenith angles 90° (dashed line) and 60° (solid line). In each case the curves are given for every 0.5 dB from 0 to 2.5 dB at 27.6 MHz. All seasons and times of day are included. (M. Friedrich and S. Kirkwood, *Advances in Space Research*, **25**, 15 (2000).)

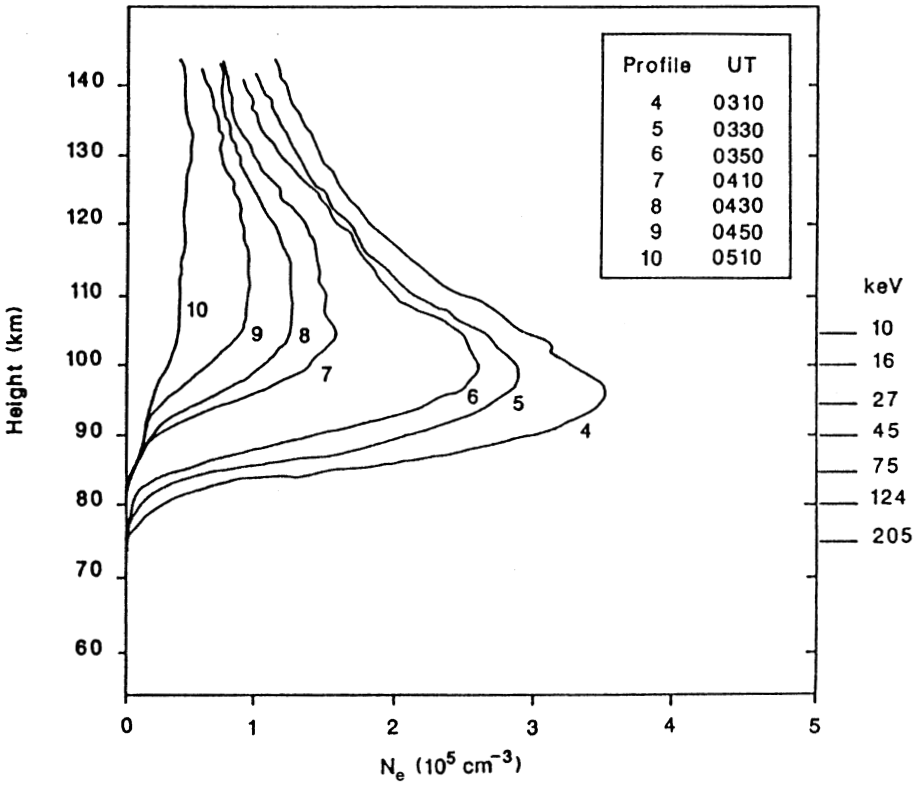


Figure 7.23. Electron-density profiles measured by the EISCAT incoherent-scatter radar during a slowly varying event in the morning of 23 March 1985. The heights of maximum ion production for electrons of the stated initial energy are marked on the right-hand axis. (Reprinted from J. K. Hargreaves and T. Devlin, *J. Atmos. Terr. Phys.* **52**, 193, copyright 1990, with permission from Elsevier Science.)

matches the observed one. An essential element is the effective recombination coefficient as a function of altitude (Section 1.3.3), which relates the rate of electron–ion production to the resulting electron density. This may be taken from other experimental results or computed from the known chemistry of the D region (Section 1.4.3) – neither approach being entirely satisfactory. Figure 7.24 shows electron-density profiles measured by incoherent-scatter radar and the corresponding spectra of incoming electrons computed from them. Note that the daytime spectrum is "harder" (contains a greater proportion of more energetic particles) and that the resulting electron-density profile peaks at a lower altitude than does the night-time spectrum.

Some direct comparisons have also been made. Using particle fluxes measured on low-orbit satellites, Jelly *et al.* (1964), Hargreaves and Sharp (1965), and Parthasarathy *et al.* (1966) obtained, respectively, the following empirical relations:

$$A = 4 \times 10^{-3} J^{1/2}, \tag{7.7a}$$

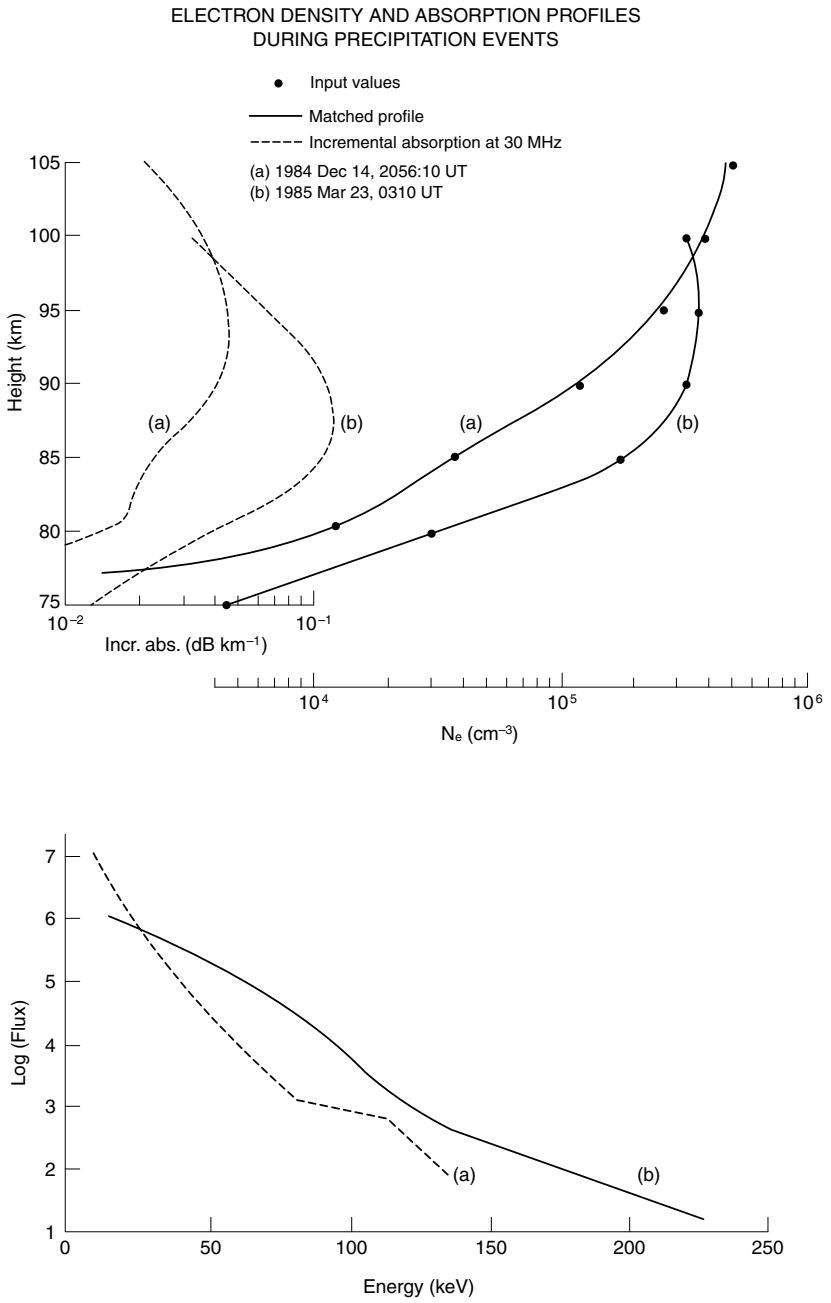


Figure 7.24. Electron-density and absorption profiles for typical night and morning events, and the estimated spectra of incoming energetic electrons: (a) 14 December 1994 at 2056:10 UT; (b) 23 March 1985 at 0310 UT. In the upper panel the solid lines show the electron-density profiles computed from the spectra in the lower panel, the black circles being observed values. (Flux is in units of cm^{-2} sr^{-1} s^{-1} keV^{-1}.) The morning event has some ten times the flux of the night event between about 40 and 80 keV, whereas the night event has a greater flux of softer (<25 keV) particles. The daytime absorption peaks at 87–88 km, the night absorption some 5 km higher.

$$A = 2 \times 10^{-3} J^{1/2}, \tag{7.7b}$$

$$A = 0.40 Q^{1/2}, \tag{7.8}$$

$$A = 3.3 \times 10^{-3} J^{1/2} \tag{7.9}$$

where A is the 30-MHz absorption in decibels, J is the flux of electrons of energies above 40 keV in $\mathrm{cm^{-2}\ s^{-1}\ sr^{-1}}$, and Q is the total energy (above 80 eV) in $\mathrm{erg\ cm^{-2}\ s^{-1}}$. Equation (7.7a) is for day and (7.7b) is for night.

The absorption is also significantly correlated to the energy flux over some energy ranges of electrons detected at geosynchronous orbit (Figure 7.25). Here the flux was taken only when the detector pointed into the loss cone; at other angles the electrons would mirror before reaching the D region. These, and other (Penman *et al.*, 1979), comparisons have indicated that the absorption correlated best to the energy influx in the bands 40–80 keV and 80–160 keV. As might be expected, the rate of production calculated from the particle flux also correlated well at some heights. Schematic production-rate profiles derived from that comparison are shown in Figure 7.26.

It should be stressed that these results and those of Figures 7.25 and 7.22 and Equations (7.7)–(7.9) make no distinction between types of event and should be taken as no more than indicative in any single instance.

The onset and dynamics of the substorm

The night event which often begins with a sharp onset, and probably a spike too, is a consequence of the substorm in the magnetosphere. The riometer is therefore a useful monitor of the occurrence of substorms at the site of the riometer. This aspect was referred to in Section 6.4.6. Furthermore, the dynamics of absorption events which may be observed using a network (Section 7.2.4) relate to the development of particle precipitation in the substorm.

The equatorward movement of the absorption arc preceeding an onset probably reflects the inward drift of an active region in the magnetosphere. Hargreaves *et al.* (1975) suggested that the motion may be $\mathbf{E} \times \mathbf{B}$ drift due to the magnetospheric electric field, in which case the value of the field can be estimated from the relation (Ranta *et al.*, 1981)

$$E(\mathrm{mV\ m^{-1}}) = \left(5.88 \times 10^{-4} \frac{\mathrm{d}t}{\mathrm{d}(1/L^2)} \right)^{-1}. \tag{7.10}$$

Equatorward drifts measured using a chain of riometers in Alaska (Figure 7.27) had speeds of several hundred $\mathrm{m\ s^{-1}}$, greatest at the highest latitudes. Interpretation in terms of a magnetospheric electric field gives a median value of $1.3\ \mathrm{mV\ m^{-1}}$. That the deduced field is independent of L supports the hypothesis, but the procedure has yet to be verified by direct comparisons with the electric field measured by other means. A later study using data from the Scandinavian sector revealed speeds mostly in the range 0–300 $\mathrm{m\ s^{-1}}$ with a peak at 100–200 m

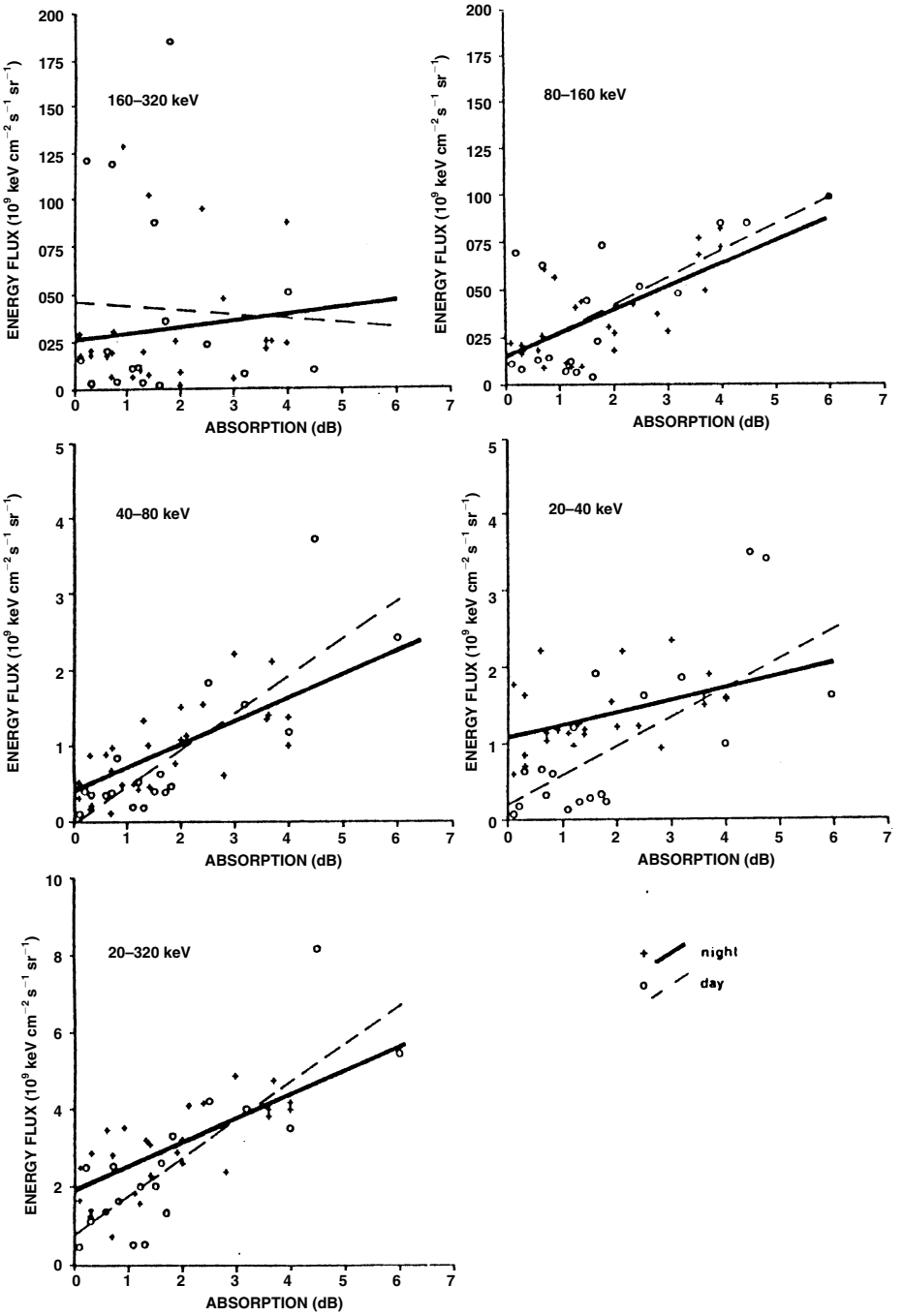


Figure 7.25. Relations between energy flux in selected bands measured on the GEOS-2 geosynchronous satellite, and radio absorption at 30 MHz observed in the auroral zone. The correlation is best in the two middle bands, indicating that the greatest contribution to the absorption comes from electrons of energy 40–160 keV. (Reprinted from P. N. Collis *et al.*, *J. Atmos Terr. Phys.* **46**, 21, copyright 1984, with permission from Elsevier Science.)

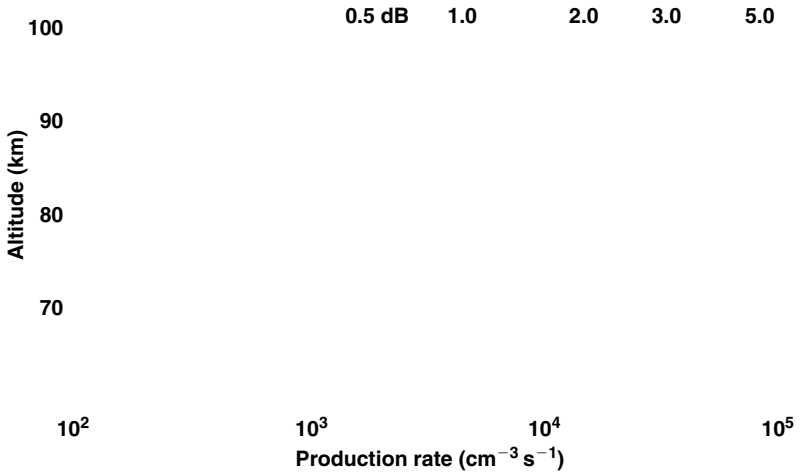


Figure 7.26. Schematic production-rate profiles for a range of values of 30-MHz radio absorption. The error bars are for one standard deviation. (Reprinted from P. N. Collis et *al., J. Atmos Terr. Phys.* **46**, 21, copyright 1984, with permission from Elsevier Science.)

s^{-1}. Table 7.6 lists some values of the cross-tail electric field derived from this set of data. In this case the median value is 0.63 mV m^{-1}.

On the other hand, the poleward motion of the spike event is not an $\mathbf{E} \times \mathbf{B}$ drift (Nielsen, 1980).

The morning events referred to in Section 7.2.2 typically move eastward, from the night towards the day side of the Earth. Velocities measured over wide baselines are not inconsistent with the concept that the electrons precipitated in the morning sector were originally injected into the closed magnetosphere near midnight and then moved eastward by gradient-curvature drift: an 80-keV electron would drift eastward at 2.6° min^{-1} and cover 90° of longitude in 35 min. (Compare with Figure 7.16.) However, motions over smaller baselines tend to be significantly slower (even to the point of mere co-rotation) and some other mechanism is plainly at work as well.

The westward movement before local midnight also requires some other explanation.

Conjugate behavior

Auroral radio absorption is particularly well suited to studies of the relative behaviors of auroral phenomena at magnetically conjugate points: that is, at the northern and southern ends of the same field-line. In the first instance, one would expect to see the same intensity of absorption and the same patterns of variation. In fact these expectations are rarely met. For instance, the absorption tends to be stronger in the winter hemisphere (Figure 7.28), and there is considerable variation in individual cases, even to the extent that an event is seen at one station but not at all at

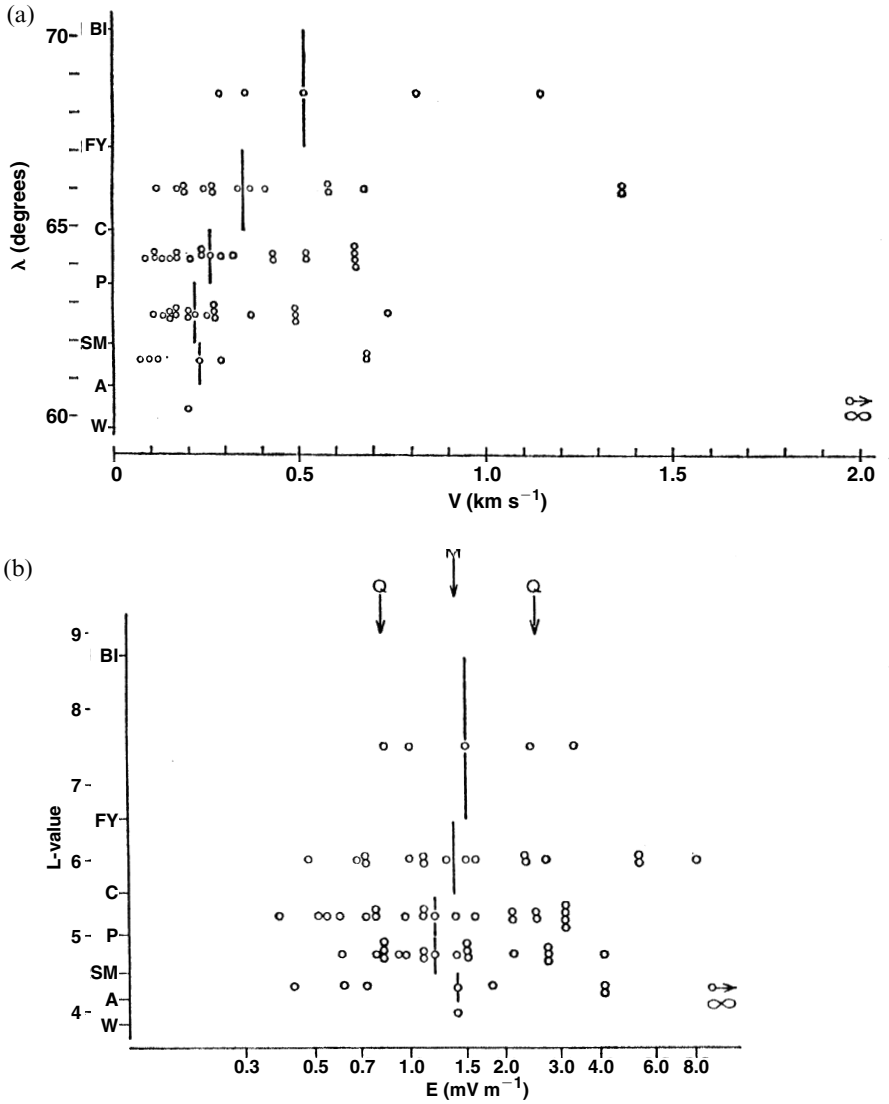


Figure 7.27. Equatorward drift of absorption bays preceeding onsets in Alaska. The locations of the riometer stations are shown by letters. (a) The speeds determined between pairs of stations (increasing with latitude). (b) The deduced magnetospheric electric field on the assumption that the motion is E × B drift. The estimated field is independent of *L*. (Reprinted from J. K. Hargreaves *et al.*, *Planet. Space Sci.* **23**, 905, copyright 1975, with permission from Elsevier Science.)

the other. The night-time events, particularly at the higher latitudes, exhibit time differences between the peaks of events in conjugate regions, the event which appears first being of greater intensity than its counterpart in the conjugate region.

One particularly interesting result, which so far lacks an explanation, is that the interhemispheric ratio depends on the direction of the interplanetary magnetic field carried by the solar wind (Figure 7.29).

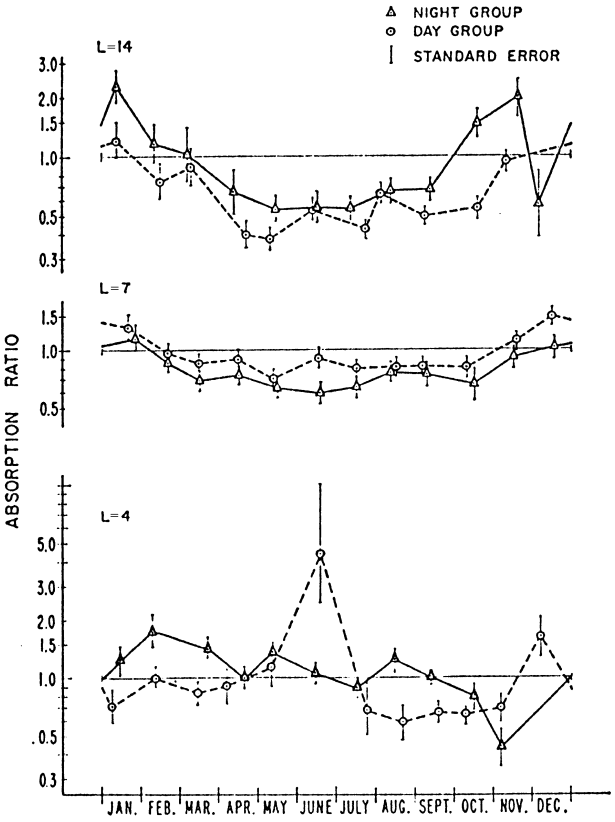


Figure 7.28. The seasonal variation of the ratio of absorption in northern and southern conjugate regions at L values of 14, 7, and 4. At both of the higher latitudes the absorption tends to be larger in the winter hemisphere. There is some difference between the day and night events. (Reprinted from J. K. Hargreaves and F. C. Cowley (1967b), *Planet. Space Sci.* **15**, 1585, copyright 1967, with permission from Elsevier Science.)

Table 7.6. *Values of the cross-tail magnetotail electric field deduced from the equatorward drift of absorption arcs (Ranta* et al.*, 1918)*

Date	UT	E(mV m^{-1})
27 March 1975	1430–1530	2.4
2 May 1975	1300–1500	0.74
	1720–1900	0.45
3 November 1975	1700–1800	1.1
2 March 1976	1600–1730	0.43
	1930–2030	0.94
2 May 1976	1900–2000	0.94
22 May 1976	1800–1900	0.44
	2000–2100	0.62
29 May 1977	1900–2000	0.58
4 May 1977	1800–1930	0.63

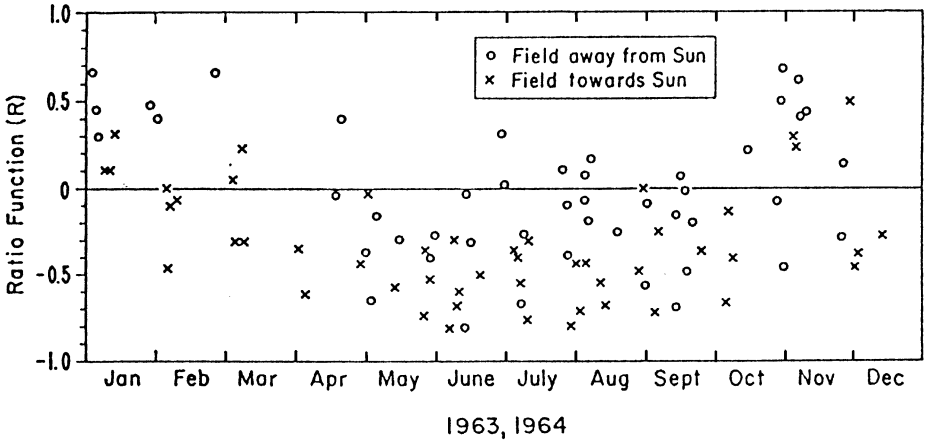


Figure 7.29. The variation of the north-to-south absorption ratio between the conjugate stations Frobisher bay and South Pole (expressed as the "ratio function" $(r-1)/(r+1)$). In addition to the seasonal variation, note that the ratio is greater when the interplanetary magnetic field is pointed away from the Sun. (Reprinted from J. K. Hargreaves and F. C. Cowley (1976b), *Planet. Space Sci.* **15**, 1585, copyright 1967, with permission from Elsevier Science.)

During the 1960s, spaced riometers were deployed around Byrd station in the Antarctic and its computed conjugate point (Great Whale River) in the Canadian Arctic, and these produced evidence that the conjugate point may be displaced by up to ± 85 km north–south with respect to the computed conjugate point, depending on the season and time of day (Hargreaves, 1969b).

The absorption pulsations in the Pc4 and Pc5 bands, which appear as a modulation of slowly varying events in the morning sector, are observed to be in phase in magnetically conjugate regions (Chivers and Hargreaves, 1964). See Figure 7.30. This indicates that the modulation is symmetrical between hemispheres and is imposed in the magnetosphere. From the electron-density profiles observed during pulsations, it appears that the modulation involves the energy of the particles, not just their flux (Hargreaves and Devlin, 1990).

7.3 The polar-cap event

7.3.1 Introduction

In the history of ionospheric studies the polar-cap event is a relatively recent discovery. On 23 February 1956 there occurred a major solar flare that was followed by polar radio blackouts lasting for several days. At the same time, cosmic-ray monitors detected a large increase in the intensity of cosmic rays at ground level. The effects on VHF forward-scatter circuits operating at that time were studied by D. K. Bailey, who showed that the cause of the blackout was an enhancement of ionization in the D region of the polar ionosphere (Bailey, 1959). He deduced

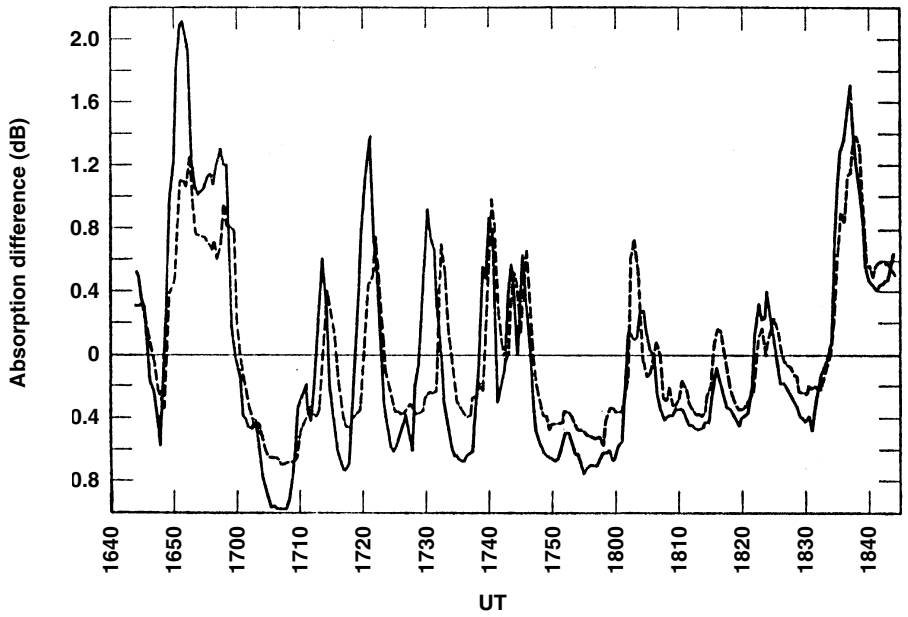


Figure 7.30. Conjugate pulsations at Great Whale River (– – –) and Byrd (——) (L = 7). The mean trend has been removed. The pulsations are essentially in phase at the conjugate stations. (Reprinted with permission, from J. K. Hargreaves and H. J. A. Chivers, *Nature* **203**, 963, copyright 1964, Macmillan Magazines Limited.)

further that the most likely cause of this added ionization was a flux of energetic protons released from the Sun at the time of the flare.

As with auroral absorption, studies based on the occurrence of the blackout condition have their limitations. Most scientific studies of the phenomenon have therefore made use of riometers, which give a quantitative measure of the absorption. These showed that the effects were confined to high magnetic latitudes, but, unlike AA, covered the whole polar cap. Thus they became known as *polar-cap absorption (PCA) events*.

PCA events are much less frequent than auroral events, there being several each year on average. However, their effects, when they do occur, are more severe both because they blanket a large region of the Earth and because of the magnitude of the absorption. The most energetic events are also detected at the ground by cosmic-ray counters, and there is about one of these events each year on average. (In total 34 were noted in the 30 years between 1955 and 1985 – Smart and Shea, 1989.) A PCA that is also recorded by a cosmic-ray detector at the ground is called a *ground-level event (GLE)*. The first recognized GLE occurred on 28 February 1942; it was identified in retrospect, of course, since the PCA was not yet a known phenomenon. The flare responsible for that event has another claim to fame as the source of the first solar radio noise to be recorded at the Earth.

Since the early 1960s it has been possible to observe solar protons in space, and

the monitoring of energetic protons from satellites has now become a matter of routine. As might be expected, the satellite monitors find some events that are not seen by ground-based methods. In fact, most solar flares emit protons at the lower energies – that is, up to 10 MeV. At energies of several tens of mega-electron volts the flux of protons reaching the Earth's vicinity far exceeds that from galactic cosmic rays, though at the highest energies, greater than 1 GeV, the galactic particles dominate.

7.3.2 Observed properties of PCA events

Occurrence and duration

There is no doubt that the PCA event is due to energetic (1–1000 MeV) protons emitted from the Sun, usually during a solar flare. The occurrence of PCA therefore depends strongly on the sunspot cycle. There can be more than ten events in an active year, and none at all – though more often one to three – near solar minimum. The long-term average is about six events per year. The numbers detected depend, of course, on what detection threshold has been selected.

As an example, the occurrence of PCA events that reached at least 1 dB on a 30-MHz riometer situated within the polar cap is shown in Figure 7.31(a). This covers the period 1962–1972, which included the end of solar cycle 19 and the first eight years of cycle 20. Some of these events were much larger than 1 dB; 12% of them reached 10 dB or more. Those events recorded as $\geqslant$5 dB are indicated. It will be noted that none of these larger events occurred during the quiet years 1962–65.

The durations of the events of magnitude $\geqslant$1 dB are shown in Figure 7.31(b). The median was about 2.5 days. The main group in the histogram spans 12–108 h. Those occurring within the narrow range 120–132 h appear as a separate group, but an alternative explanation may be that these long events were actually composed of several shorter ones. Be that as it may, we can summarize the distribution by saying that, once a PCA event has started, it is most likely to last for 1–4 days but in some cases may continue for a week or more.

Figure 7.32(a) shows the occurrence of proton events in relation to the maximum flux of protons with energy at least 10 MeV. The plot covers the years 1976–1989, which included solar cycle 21 and the beginning of cycle 22. The general influence of the sunspot cycle is seen again, except that there is a remarkable dearth of events near the peak of the cycle in 1979–1980. This looks like an extreme case of a well-reported effect. The correlation between the solar cycle and the occurrence of PCA events is imperfect, but it has often been noticed that there are fewer events than might be expected at the maximum of the solar cycle. (Alternatively, there might be too many as the cycle begins and during its decline.) The pattern of occurrence varies from cycle to cycle, but this may be due in part to the statistics of small numbers. Although the number of sunspots is a guide, it is not safe, therefore, to try to predict from previous cycles too exactly.

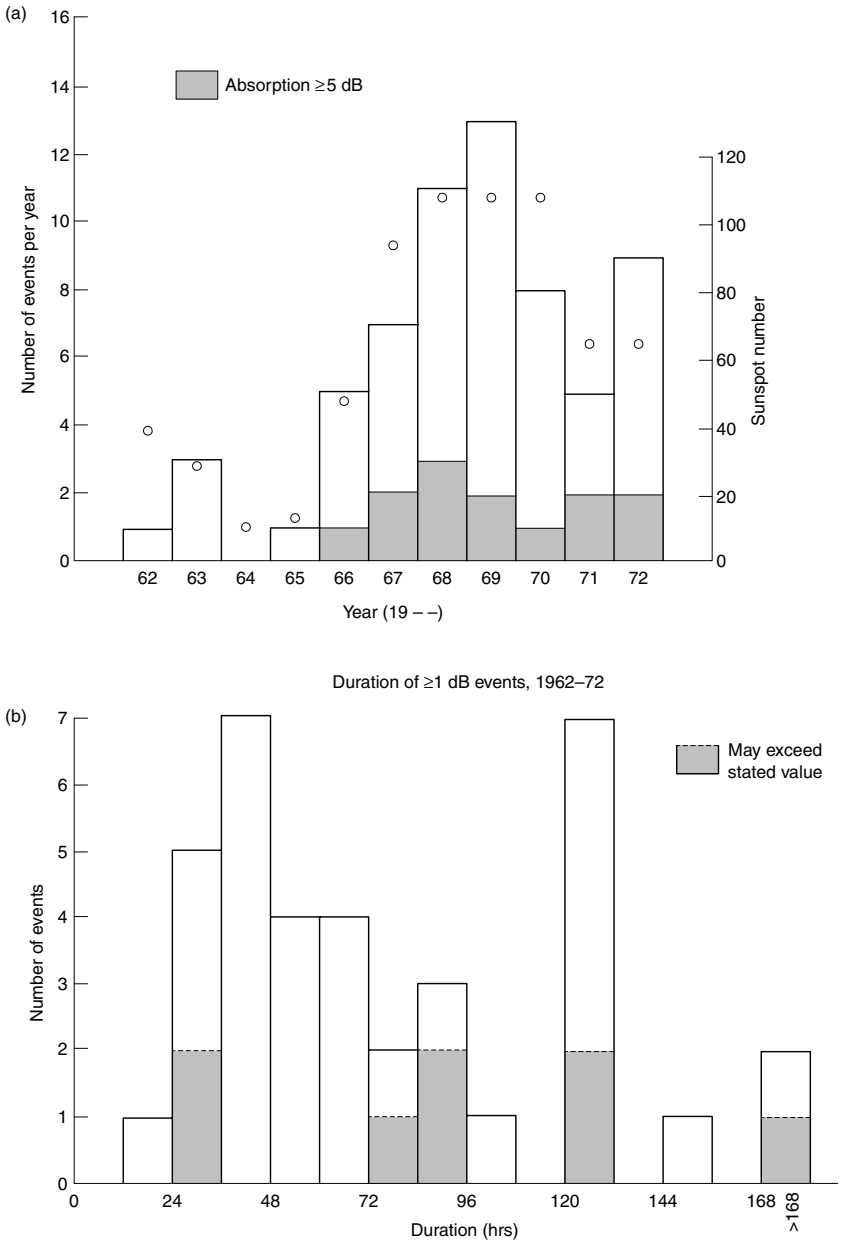


Figure 7.31. Occurrence and duration of PCA events producing at least 1 dB of absorption on a 30-MHz riometer in the polar cap. The period covered is 1962–1972. (a) The annual rate of occurrence, related to the 12-month running mean sunspot number. The incidence of events exceeding 5 dB is indicated. (b) Durations of ⩾1-dB events. In some cases it was only recorded that the duration exceeded some value, and these are indicated by shading. The median duration was 62 h (about 2.5 days). It is possible that some of the longer events were composed of several shorter but overlapping events. (After M. A. Shea and D. F. Smart, *Solar–Terrestrial Physics and Meteorology: SCOSTEP Working Document II, 1997; SCOSTEP Working Document III*, 1979.)

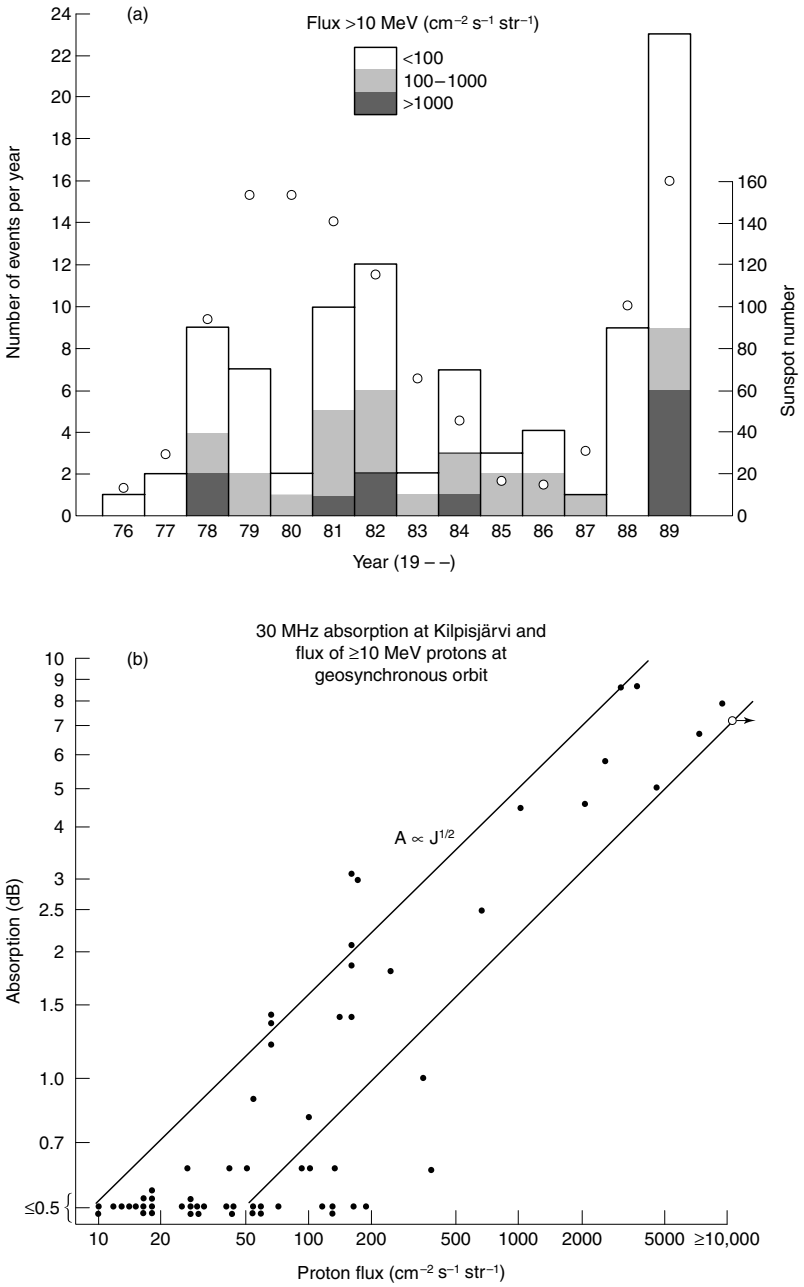


Figure 7.32. Some properties of solar-proton events recorded by a geosynchronous satellite and a riometer in the auroral zone. (a) The incidence of proton events according to the maximum proton flux at synchronous orbit. The sunspot numbers are shown as the 12-month running mean. The data cover the years 1976–1989. (b) The relation between 30-MHz absorption at Kilpisjärvi and the flux of protons of energy ≥10 MeV at geosynchromous orbit. 60% of the points lie between the straight lines marked, representing $J = 37A^2$ and $J = 200A^2$ – compare with equation (7.12). (Data from H. Ranta *et al.*, *J. Atmos. Terr. Phys.* **55**, 751 (1993).)

The direct connection between proton flux and PCA is confirmed by Figure 7.32(b), which plots the absorption at Kilpisjärvi, Finland, against the proton flux detected on a satellite in synchronous orbit. The straight lines indicate the law

$$\text{absorption} \propto (\text{flux})^{1/2}, \tag{7.11}$$

which is to be expected if the electron-production rate is proportional to the particle flux. Note that fluxes greater than 100 $\text{cm}^{-2}\,\text{s}^{-1}\,\text{sr}^{-1}$ are likely to produce a significant PCA. Since Kilpisjärvi is in the auroral zone (at $L = 5.9$) rather than the polar cap, the absorption recorded there may at times be reduced by the proximity of the edge of the polar cap.

An approximate rule that is sometimes used to deduce the proton flux from the radio absorption (Smart and Shea, 1989) is

$$J = 10A^2, \tag{7.12}$$

where J is the flux (in $\text{cm}^{-2}\,\text{s}^{-1}\,\text{sr}^{-1}$) of protons with energy exceeding 10 MeV, and A is the absorption (in decibels) measured with a 30-MHz riometer in the sunlit polar cap.

The statistics of the occurrence of PCA is complicated by episodic behavior. An individual proton event is generally recognized by noting an increase in proton flux or by virtue of radio absorption having the established PCA characteristics (i.e. a smooth event of long duration). However, an active solar region may well persist long enough to produce two or more proton flares and it is not unusual, therefore, for two or more PCAs to occur within a few days of each other. Since an event may last for several days, some events run into each other. The data set shown in Figure 7.31 contained 63 events. Of these, 25 occurred within one of ten groups of events, the criterion for a group being that events occurred within 5 days of each other. The count of groups is of course less than the count of individual events. To take an example, 1968 had 11 PCA events, but eight of them occurred in three groups and only three events were isolated. Perhaps 1968 should be credited with six PCA-producing regions, therefore, instead of with 11 PCAs. 1969 was also significantly affected in this way: in February of that year four events occurred on four successive days! Beyond a general impression that more events fall within groups in the more active years, it is difficult to draw general conclusions because of the small numbers involved.

Variation from month to month

One of the puzzles regarding the occurrence of PCA which came to light early was what appeared to be a seasonal effect: it was observed that fewer events occurred during the northern hemisphere winter than at other times of year. There is no reason to suppose that the Sun becomes less active in December and January, but, taking type-IV radio bursts (see Section 7.3.3) as a reference, there is evidence that the protons were taking longer to reach the Earth at those times – see Figure 7.33.

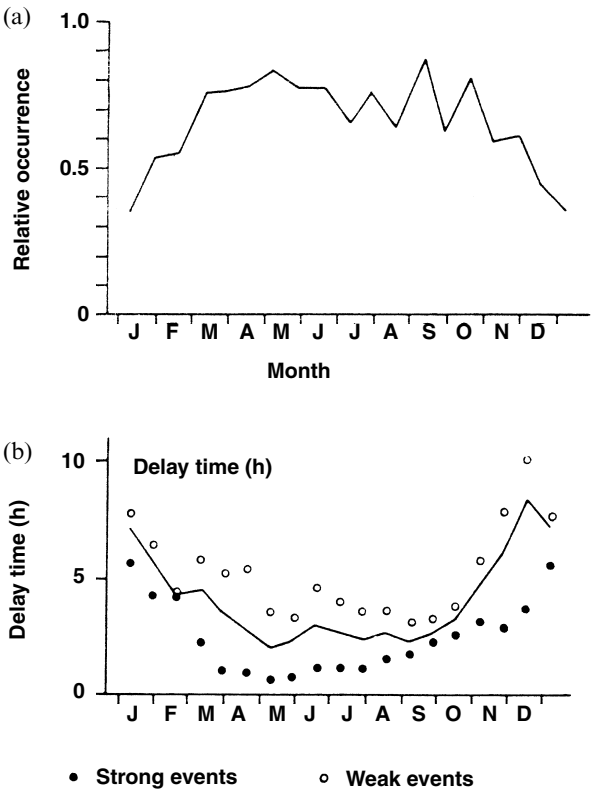


Figure 7.33. Seasonal effect in PCA: (a) the fraction of flares having Type-IV radio bursts which also produce PCA, and (b) the seasonal variation of the time delay between a radio burst and commencement of PCA. (After B. Hultqvist, *Solar Flares and Space Research* (eds. de Jager and Svestka), p. 215, North-Holland, 1969.)

It has also been argued that the effect may be artificial and due to some observational bias. The most likely cause of bias is that, since the absorption is weaker in a dark ionosphere (Section 7.3.6), the ionosphere is dark for more of the time in winter, and more of the early riometer stations were in the northern hemisphere, then the detection of PCAs by radio would be less sensitive overall in the northern winter.

Supporting this view (which probably holds sway at present) is the fact that the anomaly in the seasonal occurrence seems to be one of those effects which becomes less convincing the more intensively they are studied. It has tended to vanish as the data base has grown with the passing of the years! Thus the sets of data used for Figures 7.31 and 7.32 both show the incidence varying considerably from month to month, but they contain no evidence for any significant seasonal effect. Indeed, the monthly distribution of proton events measured on a satellite appears to show some preference for the equinoxes (Smart and Shea, 1989). Since the question of seasonal effects remains in doubt, it is probably best to assume for prediction purposes that the incidence of PCA has no seasonal dependence beyond ordinary statistical variations.

That assumption being made, the probability that a stated number of events will occur in one month may be calculated from the Poisson distribution which

describes the frequency of occurrence of independent events within a given period of time. We have to know (or assume) the average rate of occurrence. Table 7.7 gives the monthly statistics for annual rates of occurrence of two, six, and ten, corresponding approximately to low, average, and high PCA activity. Since events occurring in a group (as defined above) are probably not independent, a group should be counted as one event for this purpose.

Magnitude

Not surprisingly, there are more small PCA events than large ones. Table 7.8, taken from the data of Shea and Smart (1977; 1979), shows how many events exceeded various absorption thresholds during the 11-year period 1962–1972. Note that, of the events of magnitude $\geqslant$0.5 dB, about half reach 1 dB, about one fifth of those reach 5 dB, and about one third of those reach 15 dB. An approximate rule that appears to satisfy the limited information available is that the number of events exceeding a stated threshold varies in inverse proportion to that threshold value.

The review by Smart and Shea (1989) discusses the incidence of proton events in some detail.

7.3.3 The relation to solar flares and radio emissions

In fact, not all large flares give rise to proton events and there are some proton events that have not been associated with any known flare. However, although the correlation might not be 100%, there is no doubt that, as a general rule, proton

Table 7.7. *The probability that the stated number of PCA events will occur in one month, given the annual rate*

Expected annual rate	Probability of the stated number occurring in one month			
	0	1	2	3
2	0.846	0.141	0.012	0.001
6	0.607	0.303	0.076	0.013
10	0.435	0.362	0.151	0.042

Table 7.8. *The distribution of magnitude of PCA events*

Threshold (dB) (30-MHz riometer)	0.5	1.0	2.0	5.0	10.0	15.0
Total number of events exceeding threshold	113	63	36	13	8	3
Percentage of total	100	56	32	11.5	7.1	2.7

events are associated with the larger solar flares. Those flares that produce protons are often called *proton flares* and they are recognized as a distinct class in the flare predictions which are issued regularly by various national and international warning agencies.

The solar radio emissions known as *type IV* are useful for predicting which flares emit protons. The type-IV emission is a radio burst of long duration that follows some flares and covers a wide band of radio frequency. (It is attributed to synchrotron radiation from high-energy particles gyrating in the solar magnetic field.) The bursts associated with proton emission are characterized by a U-shaped spectrum in which the intensity is smaller in the middle than it is at the ends. For example, if the spectrum covers the range from a few hundred megahertz to 10 GHz, it will be relatively strong at the high- and low-frequency ends but weaker at the middle frequencies around 1 GHz. From the spectral characteristics of the radio burst it is possible to predict the flux of protons with energy exceeding 10 MeV (Castelli *et al.*, 1967) and also the proton spectrum (Bakshi and Barron, 1979). Since the radio burst is received at the Earth some time before the protons are due to arrive, the association obviously has some practical importance.

The association between proton ejection and the radio burst is also useful for identifying the flare responsible and for timing the flight of the proton cloud to Earth. This time appears to be shorter (about 1 h) for strong events and longer (about 6 h) for weak ones.

7.3.4 Effects arising during the proton's journey to Earth

The production and release of energetic protons appears to be a normal part of the solar-flare phenomenon, and flares causing PCA and GLE at the Earth probably differ from others more in degree than in nature. The journey from Sun to Earth may be considered in three parts:

(a) propagation from Sun to Earth – i.e. from the Sun to the boundary of the magnetosphere;
(b) motions within the magnetosphere; and
(c) the interaction between protons and the atmosphere.

Effects in interplanetary space

The propagation of charged particles in the space between the Sun and the Earth is affected by the form of the interplanetary magnetic field. As illustrated in Figure 2.3, the field has a spiral form due to the rotation of the Sun, and this affects the propagation of solar protons despite the weakness of the field and the high energy of the protons. The gyroradius of a 1-GeV proton in a field of 5 nT is less than a hundredth of the distance between the Sun and the Earth; hence there is time for the IMF to act on even an energetic proton. Those with less energy gyrate in

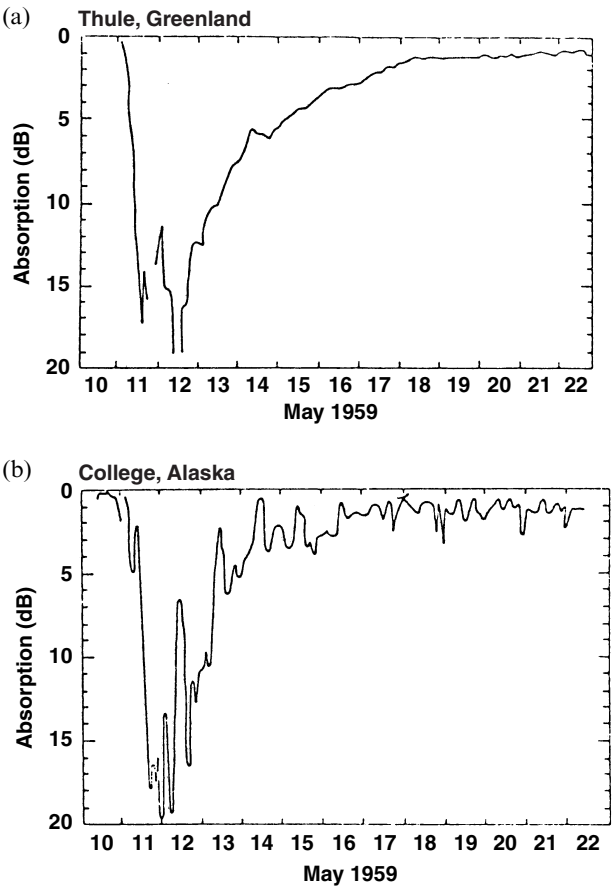


Figure 7.34. A PCA recorded by riometers at (a) Thule, Greenland and (b) College, Alaska. Thule is in the polar cap and College in the auroral zone. The event lasted for a week at both sites, but was modulated at the lower latitude. (G. C. Reid, in *Physics of Geomagnetic Phenomena* (eds. Matsushita and Campbell), Academic Press, 1967.)

smaller loops (gyroradius $\propto$ energy$^{1/2}$) and are more tightly controlled, responding to irregularities in the IMF as well as to its general form. Hence there is scattering, and the protons appear to be coming from all directions by the time they reach the Earth.

Scattering in the interplanetary medium, since it also provides a mechanism for storing particles in space, can account for the observed time delay between a flare and the beginning of the PCA, and for the duration of PCA events. A proton of energy 10 MeV would reach the Earth in only 1 h if it traveled in a straight line, and the duration of a flare is typically some tens of minutes only. In fact the delay before an event begins is typically several hours, and the event due to one flare may last for several days (Figure 7.34).

Further evidence for the role of the IMF is as follows.

(a) Flares near the eastern limb of the Sun rarely give rise to PCA events, whereas some events seem to be associated with flares that are out of sight around the western limb. This is illustrated by Figure 7.35, which gives the positions of solar flares associated with those proton events energetic enough to be detected at ground level (i.e. GLEs). (Note that the western

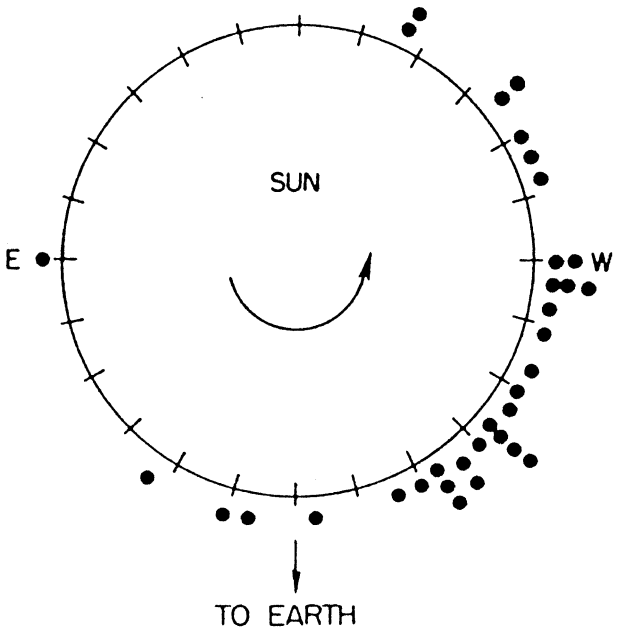


Figure 7.35. Solar longitudes of flares associated with ground-level events. The Sun is happy about this. (D. F. Smart and M. A. Shea. *J. Spacecraft Rockets*, **26**, 403, 1989.)

limb of the Sun is on the right-hand side as seen from the terrestrial northern hemisphere.) It is obvious that the distribution of these flares with solar longitude is significantly biased towards one side of the central meridian. The heliolongitudinal distribution of the source flares broadens for protons of lower energy (Smart and Shea, 1995).

(b) The time delay between a flare and the related PCA increases with the eastern longitude of the flare.

(c) The delay between flare and PCA is greatest at times of high solar activity, and this is also when the IMF is most irregular.

Recent improvements in detecting structures in the interplanetary medium have focused attention on the role of coronal mass ejections (Section 2.2.2). It is found that some PCAs may be attributed not directly to flares but to the shock wave related to a coronal ejection of mass from the Sun (Shea and Smart, 1995).

Effects in the magnetosphere

On reaching the magnetopause the protons must then pass through the geomagnetic field to reach the atmosphere. To a first approximation this problem may be treated by *Störmer theory*. The theory describing the trajectories of charged particles in a dipolar magnetic field, which C. Störmer worked out in connection with his studies of the aurora, does not actually apply to auroral particles (because their energies are too low)! However, the theory is valid for cosmic rays and for solar protons.

In a magnetic field a charged particle tends to follow a spiral path whose radius

of curvature ($r_B = mv/(Be)$) is directly proportional to its velocity and inversely proportional to the magnetic flux density. Because solar protons are of relatively high energy, the magnetic field changes significantly over one gyration, and therefore trapping theory (as outlined in Section 2.3.4) does not apply. Nonetheless, it is still true that particles traveling almost along the magnetic field undergo the least deviation. To reach the equator, which is the least accessible region, a proton has to cross field-lines all the way down to the atmosphere. Charged particles may do this only if they are sufficiently energetic, and the equatorial region is effectively forbidden to typical protons of solar origin. However, most of the particles in a proton event can penetrate into the atmosphere over a polar cap extending down to about 60° magnetic latitude.

Since the radius of gyration in a given magnetic field depends on the momentum per unit charge (mv/e), it is convenient to discuss particle orbits in general in terms of a parameter called *rigidity*:

$$R = Pc/(ze), \tag{7.13}$$

where P is the momentum, c the speed of light, z the atomic number, and e the electronic charge taken positive. The advantage of this parameter is that all particles with the same value of R will follow the same path in a given magnetic field.

Although the trajectory of a proton in the geomagnetic field can be very complicated, Störmer's analysis simplified matters by defining "allowed" and "forbidden" regions that could and could not, respectively, be reached by a charged particle approaching the Earth from infinity. To reach magnetic latitude λ_c in a dipole field, the rigidity of the particle must exceed a *cutoff rigidity*, R_c:

$$R_c = 14.9 \cos^4 \lambda_c, \tag{7.14}$$

where R_c is measured in gigavolts (10^9 V). That is, particles of rigidity R_c reach latitudes λ_c and above. Conversely, a place at latitude λ_c would receive only those particles with rigidities equal to and greater than R_c. Figure 7.36(a) plots the Störmer cutoff latitude against energy both for protons and for electrons.

To perform an exact calculation of the trajectory of a proton through the geomagnetic field, the procedure is to imagine that a proton with negative charge is projected upward from the point of impact, since the trajectory of such a particle is exactly the reverse of that of an incoming positively charged particle having the same rigidity. From a set of computations of this kind it is possible to work out the directions in space from which the particles reaching a given place at a given time must have come. Results confirm other evidence that, whereas most protons are isotropic near the Earth, the more energetic ones (those exceeding 1 GeV which are responsible for ground-level events) originate from the western side of the Sun. (See Figure 7.35.)

During the main part of a typical PCA event the absorption region is essentially

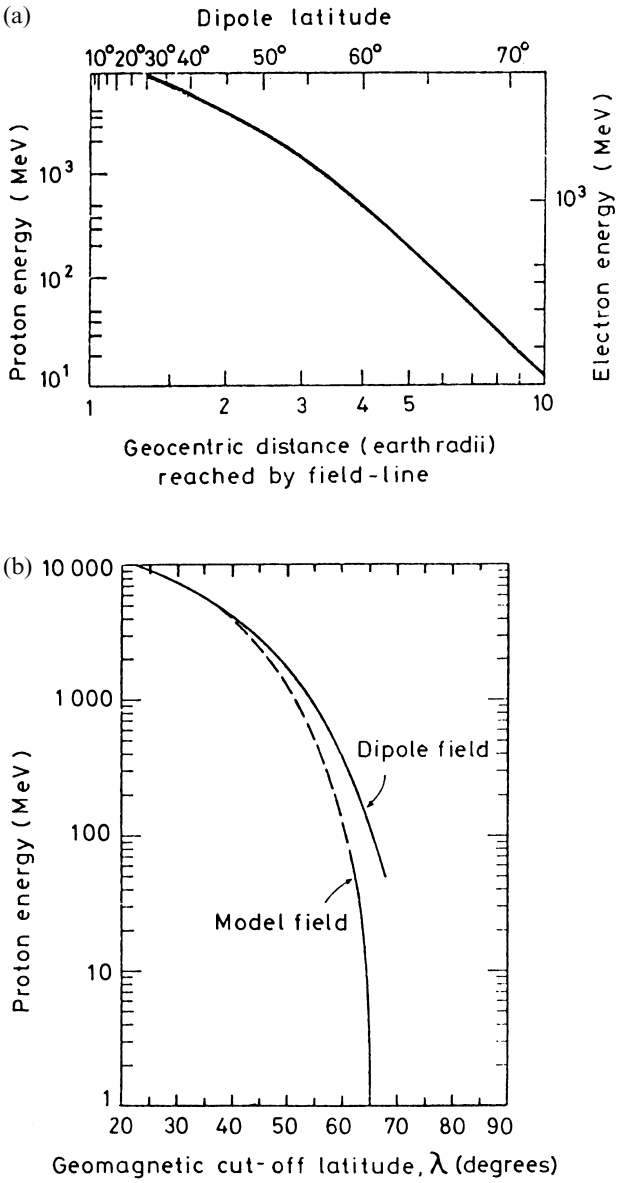


Figure 7.36. (a) The Störmer cutoff latitude for protons and electrons. (S.-I. Akasofu and S. Chapman, *Solar–Terrestrial Physics*. Oxford University Press, 1972, by permission of Oxford University Press. After T. Obayashi, *Rep. Ionosphere Space Res. Japan*, **13**, 201, 1959.) (b) Cutoff latitudes for dipole and realistic geomagnetic fields (G. C. Reid and H. H. Sauer, *J. Geophys. Res.* **72**, 197, 1967, copyright by the American Geophysical Union.)

uniform and symmetrical over the polar caps down to about 60° geomagnetic latitude. According to Störmer theory these protons should have energies exceeding 400 MeV, but direct observations of the particles have shown that the cutoff rigidity at the edge of the polar cap is significantly less than the Störmer value. The situation appears to be that there is a main polar cap surrounding the geomagnetic pole that is open to solar protons of all energies, and then at slightly lower latitude the cutoff reverts fairly abruptly to the Störmer value. Much of this effect (though perhaps not all) may be explained by taking account of the tail of the magnetosphere which connects directly to the polar caps and presumably provides an easy

route even for protons of low energy. Figure 7.36(b) shows the difference in cutoff energy between dipolar and more realistic geomagnetic fields. The cutoff is reduced still further if a magnetic storm (Section 2.2.3), which enhances the ring current (Section 2.3.5) and moves the magnetopause inward, occurs while a PCA event is in progress. The geographic regions most affected by PCA are illustrated in Figure 7.37 in general terms. The boundaries may be several degrees nearer the equator during a magnetic storm.

7.3.5 Non-uniformity and the midday recovery

Non-uniformity

The spatial distribution of radio absorption is not always uniform, particularly during the early and late phases of a PCA. The absorption usually appears first near the geomagnetic poles and then spreads to cover the polar caps some hours later. Towards the end of the event there is likely to be contamination by auroral electrons related to a magnetic storm, and a concentration of absorption into the auroral zone is then to be expected. In addition, the polar cap expands during the storm, moving the PCA equatorward.

Midday recovery

Some events exhibit a reduction in the absorption for several hours near local noon. This effect is known as the *midday recovery (MDR)*, and its main properties are as follows (Leinbach, 1967).

(a) They occur during about 20% of PCA events.
(b) They are usually pronounced on the first day of the event only.
(c) They peak between 0800 and 1500 LT, most of them between 1000 and 1200.
(d) They may last as long as 6–10 h, most being remarkably symmetrical about the peak.
(e) They are strongest near the equatorward boundary of the polar cap, and are not evident at locations well within the polar cap.
(f) When the polar cap expands during a magnetic storm, the recovery region remains at its equatorward edge.

Figure 7.38 illustrates some of these features during a PCA observed at the Alaskan stations College ($L = 5.5$), Farewell ($L = 4.3$) and King Salmon ($L = 3.3$). The time scale is given in UT, from which 10 h should be subtracted to obtain Alaskan time. MDRs occurred between 0800 and 1000 LT at the first two stations on the first day of the event. On the second day a magnetic storm extended the polar cap to lower latitude and a MDR was observed at King Salmon, but was not (College) or was barely (Farewell) seen at the higher latitudes. (The horizontal bars on Figure 7.38 indicate night-time recoveries; these are different and will be considered in Section 7.3.6.)

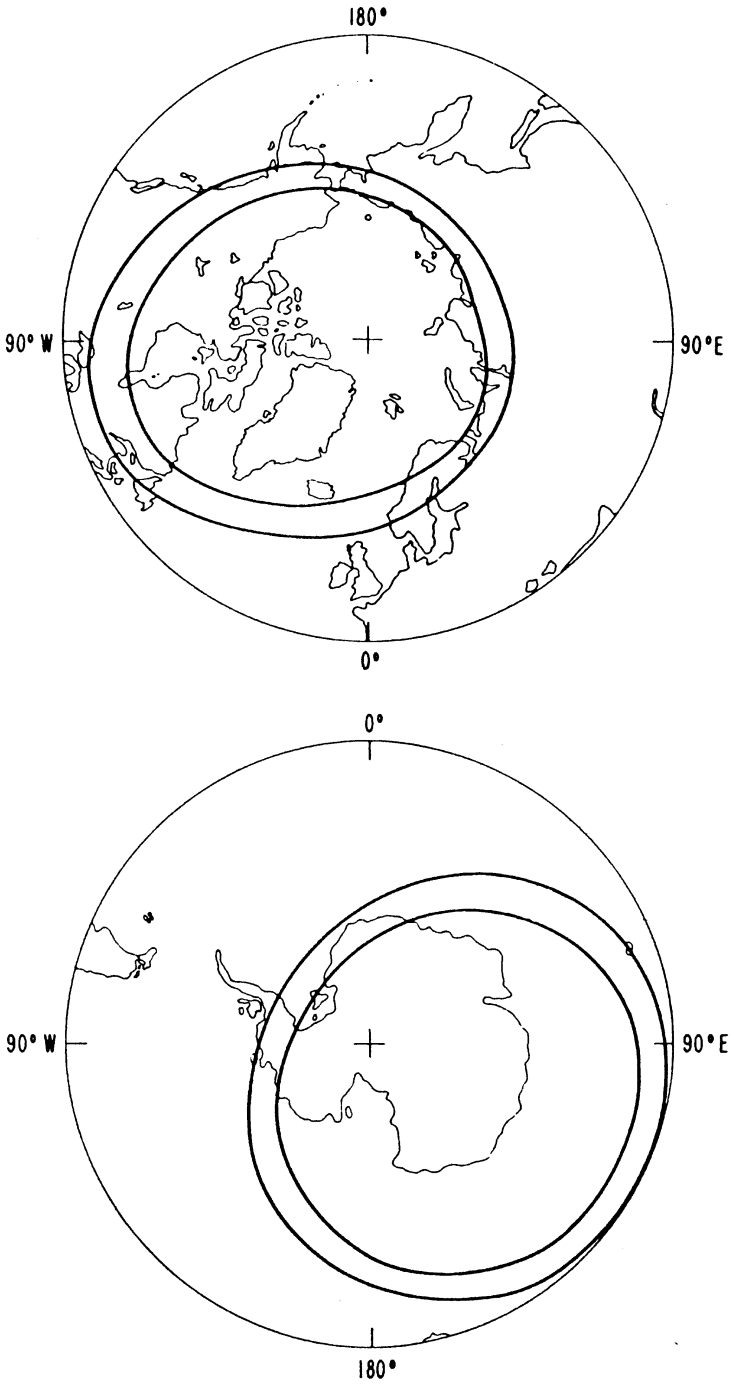


Figure 7.37. The polar areas normally affected by polar-cap absorption. The regions inside the inner curves may be considered as "polar plateaux", whereas regions outside the outer curves are usually not affected except during severe geomagnetic disturbance. The outer edges of the diagrams are at latitude 45°. (G. C. Reid, *Physics of the Sun* (ed. P. A. Sturrock), **3**, 251, Reidel, 1986, with kind permission from Kluwer Academic Publishers.)

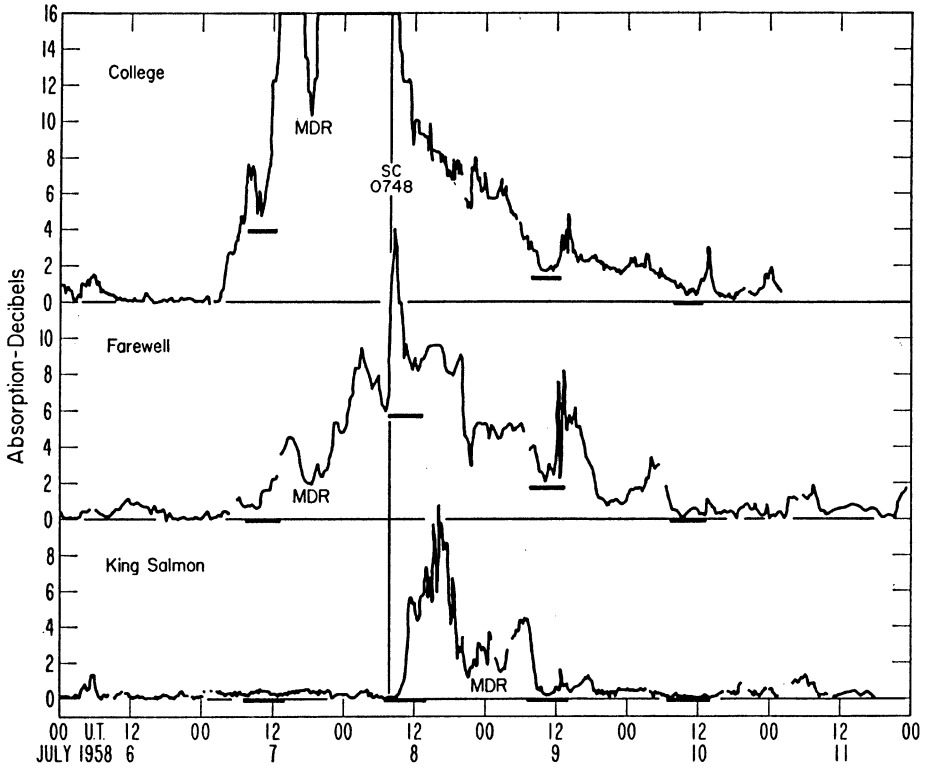


Figure 7.38. The PCA event of 7 July 1958, seen at College ($L = 5.5$), Farewell ($L = 4.3$), and King Salmon ($L = 3.3$). The horizontal bars indicate night recoveries and MDR marks midday recoveries. All observations were at 27.6 MHz. (H. Leinbach. *J. Geophys. Res.* **72**, 5473, 1967, copyright by the American Geophysical Union.)

In a recent case study using data from 25 stations including some in the southern hemisphere (Uljev et al., 1995) the maximum effect was found slightly before local noon, covering a range of magnetic latitude approximately from 60° to 70° (Figure 7.39). The effect seems to occur simultaneously and with the same magnitude in magnetically conjugate regions, and it is confirmed that the effect is not seen at stations well inside the polar cap (at latitudes greater than 70°).

Two possible explanations were put forward by Leinbach (1967): a local change of cutoff, and the development of anisotropy in the pitch-angle distribution of the incoming protons. More recent studies have suggested that both effects may occur. There is evidence that a change of cutoff near noon is indeed one factor (Hargreaves et al., 1993), and modeling studies (Uljev et al., 1995) suggest that anisotropy of the pitch-angle distribution also occurs but only over the latitude range 65°–70°.

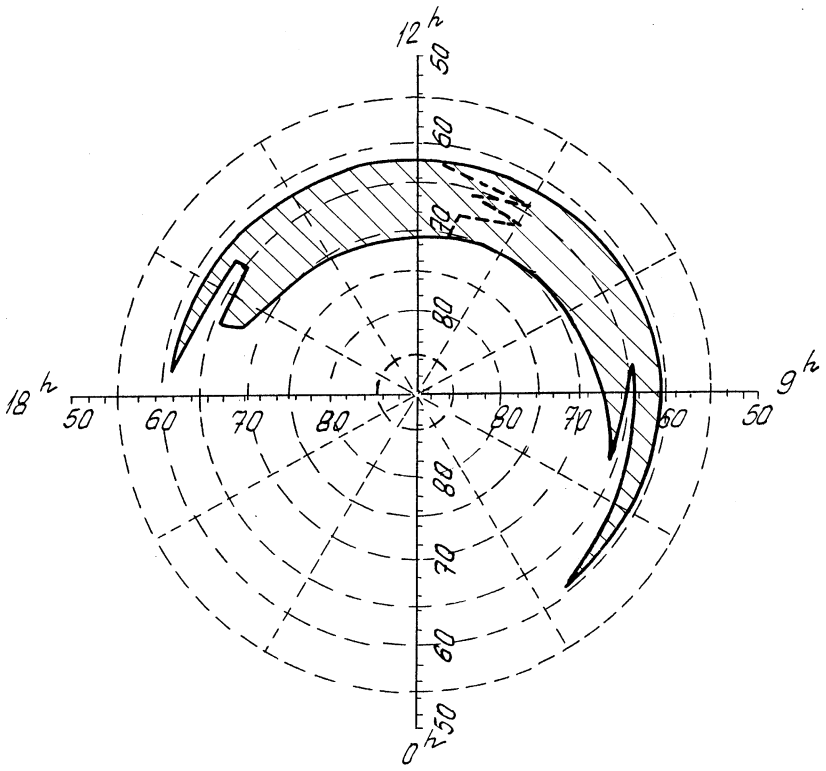


Figure 7.39. The region affected by midday recovery during the event of 20 March 1990. The coordinates are invariant latitude and magnetic LT. The broken line between 10 and 12 h marks the times of minimum absorption at each station. (Reprinted from V. A. Uljev *et al.*, *J. Atmos. Terr. Phys.* **57**, 905, copyright 1995, with permission from Elsevier Science.).

7.3.6 Effects in the terrestrial atmosphere

Upper-atmosphere ionization during a proton event

Energetic protons entering the terrestrial atmosphere lose energy in collisions with the neutral molecules and leave behind an ionized trail. In order to reach an altitude of 50 km, a proton must have an initial energy of 30 MeV, and to reach the ground (to cause a GLE) the energy must be over 1 GeV. (Refer to Figure 2.28.) An example of proton spectra observed at geosynchronous orbit during a proton event in 1984 is shown in Figure 7.40(a). Despite the name *solar proton event*, it should be appreciated that other particles, α-particles and heavier nuclei, also arrive (in proportions typical of the solar atmosphere). However, their contribution to the ionization is small relative to that of the protons. The computation of ionization by protons and α-particles was discussed in Section 2.6.3.

Having computed the rate of production of electrons at a given height, knowledge of the effective recombination coefficient allows one to calculate the resulting

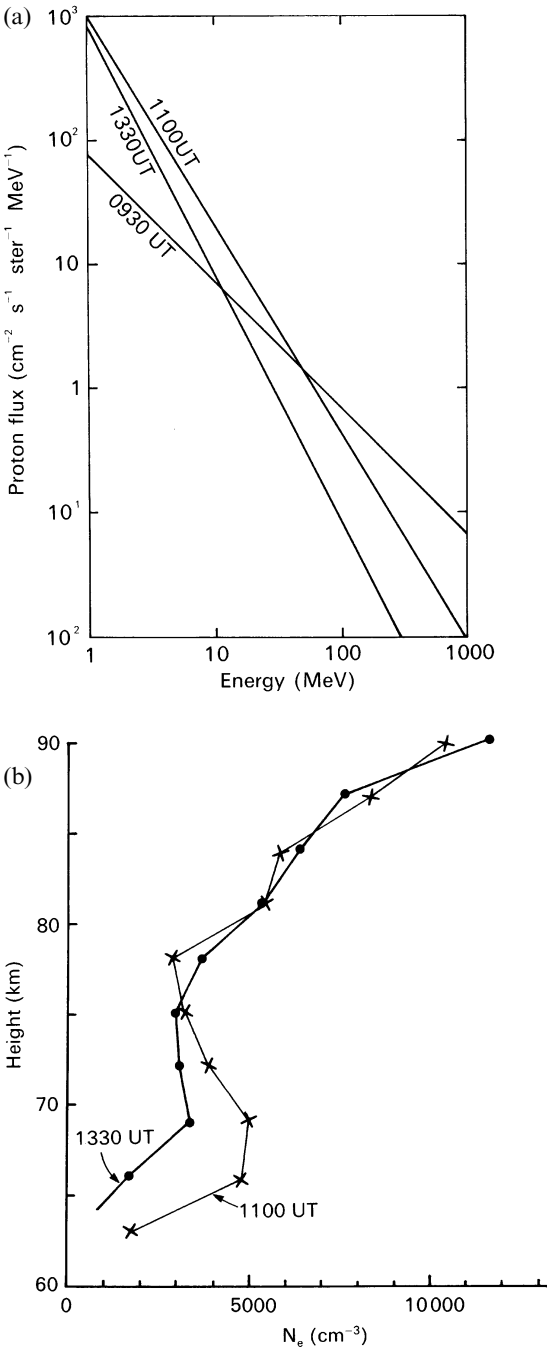


Figure 7.40. (a) Proton fluxes measured by the geosynchronous satellite GOES-5 on 16 February 1984, fitted by spectra of form $E^{-\alpha}$. (Data from F. C. Cowley, NOAA, Boulder, Colorado, private communication.) (b) Electron-density profiles measured by incoherent scatter radar during the same event. (Reprinted with permission from J. K. Hargreaves *et al.*, *Planet. Space Sci.* **35**, 947, copyright 1987, with permission from Elsevier Science.)

electron density. If an event contains particles of energy 1–100 MeV, the effects should appear within the height range 35–90 km (Figure 2.28). Effects due to the higher energies tend to be smaller because the flux is smaller and the rate of recombination is greater at lower height. Nevertheless, in some events substantial ionization is created down to 50 km.

Determination of the recombination coefficient

In fact, the recombination coefficient in the lower ionosphere is not a well-established quantity, and one use of PCA events is to measure the recombination coefficient and its variations over a range of heights in the mesosphere. The proton spectrum may be measured from a geosynchronous satellite, and from it the production rate can be computed over a range of heights using a model of the neutral atmosphere. Electron-density profiles can be determined from rocket measurements or by incoherent-scatter radar. An example of the latter is shown in Figure 7.40(b). Some studies have used riometer data, which are more readily available, though in that case only the integrated absorption can be compared.

Values of the effective recombination coefficient obtained using electron densities from incoherent-scatter radar are shown in Figure 7.41. Most striking about these values is their large spread. There is a major difference between day and night, and also between different determinations of daytime values. It is possible that there are seasonal variations caused by seasonal changes in the concentrations of minor species (Reagan and Watt, 1976). These differences will be considered in the next section.

Day–night variation and twilight effects

Because the proton influx during a PCA decays relatively slowly, the effects of daily variations in the complex chemistry of the region may also be detected. The most obvious effect is a large diurnal variation in the absorption, which is typically about five times as large by day as it is by night, though the ratio can be as small as three or as large as ten. The critical factor in this is whether the lower ionosphere is sunlit. *Night-time recoveries* were marked on Figure 7.38, and they also account for the daily absorption recoveries in Figure 7.34(b). (Thule, Figure 7.34(a), was illuminated continuously and the recoveries did not occur there.)

The effect is perhaps seen most clearly by comparing the absorption at magnetically conjugate stations, one in the summer and the other in winter, as in Figure 7.42. Over the Spitzbergen station the ionosphere was illuminated continuously, whereas at Mirnyy the Sun was above the horizon for only a few hours of the day. We expect the proton fluxes at each place to be similar, and, indeed, the absorption was almost the same when both stations were sunlit. However, the absorption fell to a considerably smaller value at Mirnyy during each night period.

The cause of the day–night modulation is without doubt a variation in the ratio of the concentrations of electrons and negative ions, λ (defined in Section 1.3.3). In a dark ionosphere, electrons become attached to oxygen molecules to form

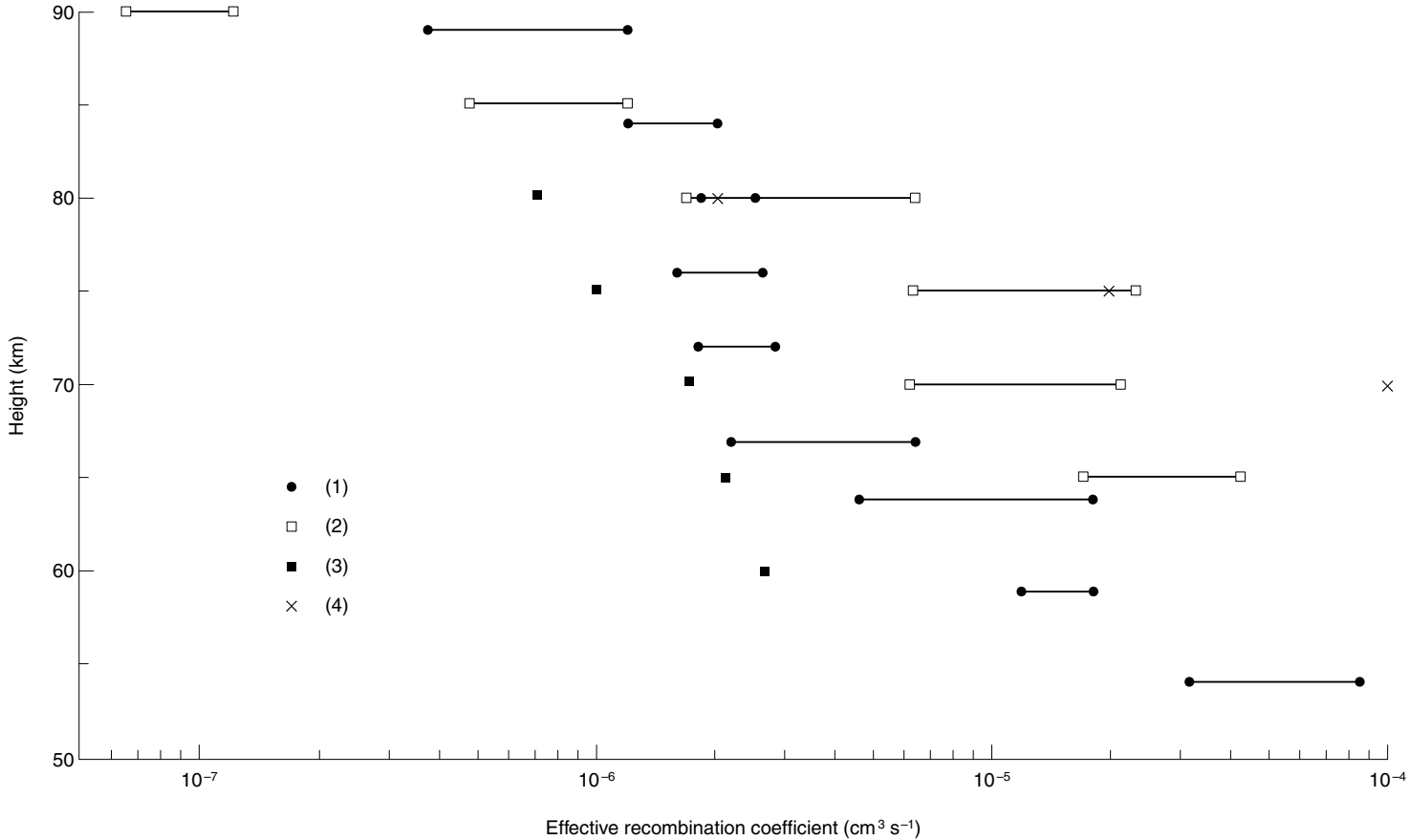


Figure 7.41. Effective recombination coefficients determined from PCA observations, using electron densities measured by incoherent-scatter radar. Key: (1) Daytime (range of values over several days). Summer (August). (Data from J. B. Reagan and T. M. Watt. *J. Geophys. Res.* **81**, 4579, 1976.) (2) Daytime (range of values over 3 hours near noon). Winter (February). (Data from J. K. Hargreaves, H. Ranta, A. Ranta, E. Turunen, and T. Turunen. *Planet. Space Sci.* **35**, 947, 1987.)
(3) Daytime (Afternoon). Spring (March). (Data from J. K. Hargreaves, A. V. Shirochkov, and A. D. Farmer. *J. Atmos. Terr. Phys.* **55**, 857, 1993).
(4) Night. (Same source as (3).)

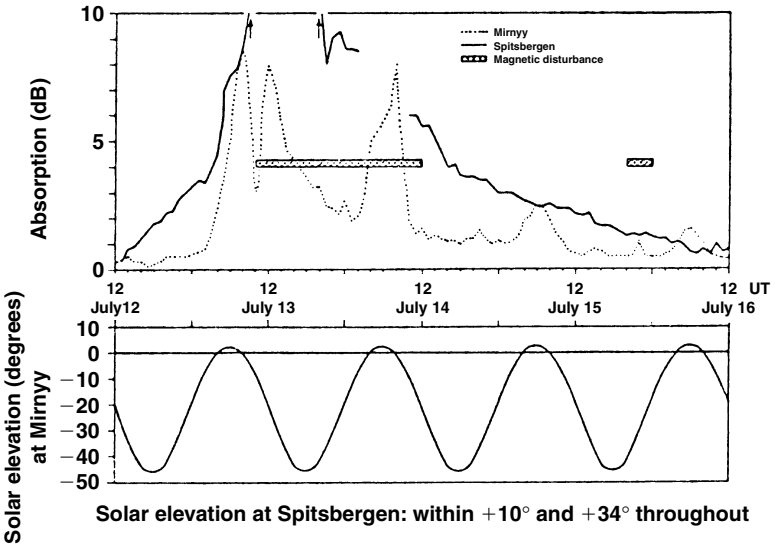


Figure 7.42. Polar-cap absorption at magnetically conjugate stations 12–16 July 1966, Spitsbergen in the northern hemisphere and Mirnyy in the Antarctic. (Reprinted from C. S. Gillmor, *J. Atmos. Terr. Phys*. **25**, 263, copyright (1963), with permission from Elsevier Science.)

negative oxygen ions (O_2^-), as Equation (1.61), but in sunlight the electrons are detached again by visible light (Equation (1.62)) or through other chemical reactions. (See Section 1.4.3.) Since only the ionospheric electrons contribute to the absorption, a variation of λ leads to a variation of absorption even though the production rate, q, remains constant.

The changes between night and day take place over the twilight periods at sunrise and sunset, and the details are of particular interest. The timing of the change in relation to the elevation angle of the Sun indicates the presence of a screening layer, probably ozone. Since ozone does not absorb in the visible, the solar radiation that detatches electrons from negative ions must be in the ultraviolet rather than the visible region of the spectrum (Reid, 1961). The effect is confined to altitudes below 80 km (Figure 7.43), which explains why it does not appear in AA (most of which occurs at a higher level).

When the details are examined it becomes apparent that some other factors are also at work.

(a) There is an asymmetry between the sunrise and sunset changes. The increase of absorption over sunrise is slower than the decrease over sunset (Chivers and Hargreaves, 1965). This means that the absorption is larger at sunset than it is at sunrise for the same solar zenith angle. The effect may be seen by plotting the absorption at a station passing through twilight periods as a ratio to that at one that is constantly illuminated. The result is a hysteresis curve like Figure 7.44, in which the curve is described counterclockwise. The same effect is present in the profiles of Figure 7.43, where

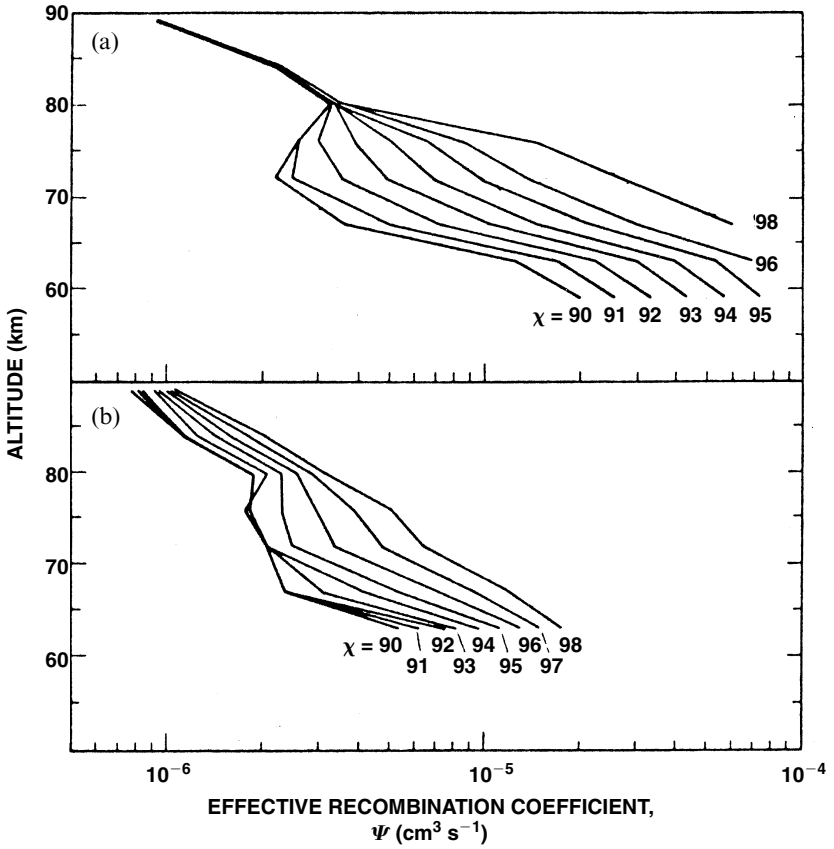


Figure 7.43. Effective recombination coefficients at various solar zenith angles over sunrise and sunset during the major proton event of August 1972. (J. B. Reagan and T. M. Watt. *J. Geophys. Res.* **81**, 4579, 1976, copyright by the American Geophysical Union.)

the smaller recombination coefficients at given zenith angle at sunset imply greater electron densities. Collis and Rietveld (1990) showed that the timing of the day–night transition depends also on the altitude, and suggested by way of explanation that different processes control the electron density above 70 km (photodetachment from O_2^-) and below 66 km (collisional detachment due to $O_2(^1\Delta_g)$) during the twilight period.

(b) There is evidence that the effective recombination coefficient varies with time even within the day and night periods. Reagan and Watt (1976) found that its value declined gradually during the sunlit period (i.e. between sunrise and sunset) by as much as a factor of two (at some heights). On the other hand, Hargreaves *et al.* (1993) reported a gradual increase of the effective recombination coefficient throughout the night. The reasons for these slow changes are not known, though they presumably lie in the chemistry of the mesosphere.

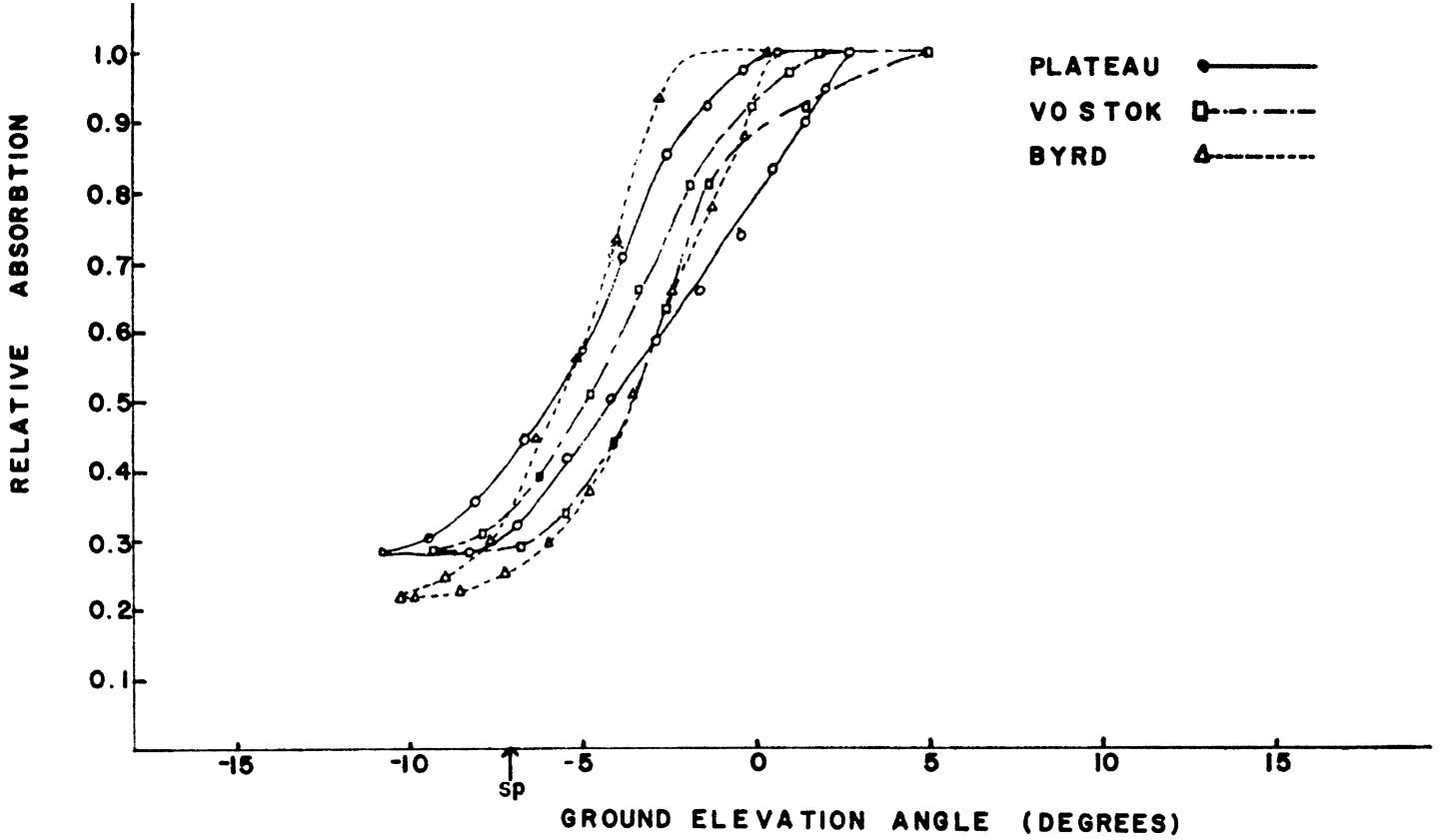


Figure 7.44. Twilight variations expressed as hysteresis curves. The reference station is South Pole (solar elevation −7°), and the curves are described counter-clockwise, implying larger absorption at sunset than at sunrise for the same solar elevation. (H. H. Sauer. *J. Geophys. Res.* **73**, 3058, 1968, copyright by the American Geophysical Union.)

Assuming that the day–night change which occurs rapidly over twilight is indeed due entirely to a variation in the ratio of negative ions to electrons (λ), and that recombination of positive and negative ions is negligible, then a simple application of Equation (1.39) – remembering also that the effective recombination coefficient $\alpha_{\text{eff}} = q/N_e^2$ by definition – gives a relation between night and day values of λ at a given height:

$$\frac{1+\lambda(\text{night})}{1+\lambda(\text{day})} = \frac{\alpha_{\text{eff}}(\text{night})}{\alpha_{\text{eff}}(\text{day})}. \tag{7.15}$$

If we take typical estimates of λ(day) of $\ll 1$, 0.25, and 0.68 at 80, 75, and 70 km, respectively, the results of Hargreaves *et al.* (1993), to take an example, give λ(night) values of 1.7, 20, and 100 at the same heights. There is, however, no generally agreed set of values for this quantity.

Effects on the neutral-species composition

Influxes of energetic particles have another important effect in that they may produce changes in the chemical composition of the atmosphere. As long ago as 1969 it was observed in rocket measurements that the ozone in the mesosphere (at heights of 54–67 km) was depleted during a PCA event by a factor between two and four depending on the height (Weeks *et al.*, 1972). The mechanism is as follows. One consequence of the ionization processes is the formation of hydrated ions ($O_2^+.H_2O$), which then undergo further reactions leading to "odd hydrogen" species such as H and OH. These radicals then react with ozone to produce molecular oxygen:

$$\begin{aligned} H + O_3 &\rightarrow OH + O_2 \\ OH + O &\rightarrow H + O_2 \\ \hline O_3 + O &\rightarrow 2O_2 \end{aligned} \tag{7.16a}$$

$$\begin{aligned} OH + O_3 &\rightarrow HO_2 + O_2 \\ HO_2 + O_3 &\rightarrow OH + 2O_2 \\ \hline 2O_3 &\rightarrow 3O_2 \end{aligned} \tag{7.16b}$$

$$\begin{aligned} OH + O_3 &\rightarrow HO_2 + O_2 \\ HO_2 + O &\rightarrow OH + O_2 \\ \hline O_3 + O &\rightarrow 2O_2 \end{aligned} \tag{7.16c}$$

In each case the odd-hydrogen radical is a catalyst; it is destroyed in the first reaction of the pair but regenerated in the second. These processes require a sufficient concentration of water vapor and therefore they are confined to the region below the mesopause. They are thought to be important over the height range 50–90 km. Several hours to a day after the precipitation event, the odd-hydrogen species reform into stable molecules; then the above reactions cease and the concentration

of ozone recovers. However, since the H atoms tend to recombine to form H_2 rather than H_2O, the water vapor may remain depleted for some time. There may be an increase in the concentration of ozone during this period.

More serious from the point of view of ozone is the effect of "odd-nitrogen" species. These have a much longer lifetime (amounting to several years) in the stratosphere, which is also the site of most of the ozone. The ionization processes produce secondary electrons with energies of tens and hundreds of electron volts, and these can dissociate and ionize molecular nitrogen to produce atoms and ions of atomic nitrogen. The N and N^+ then react with O_2 to give nitric oxide, NO, which in turn acts to destroy ozone as follows:

$$\begin{array}{l} NO + O_3 \rightarrow NO_2 + O_2 \\ NO_2 + O \rightarrow NO + O_2 \\ \hline O_3 + O \rightarrow 2O_2 \end{array} \qquad (7.17)$$

Here the NO is the catalyst. This reaction is important up to 45 km, and the long lifetime of NO at those levels means that a given molecule may pass through the reaction many times, converting one O_3 at each pass.

The above reactions do not depend on the nature of the primary ionizing radiation, but they are of particular importance in PCA because the more energetic protons ionize at particularly low altitudes and down into the stratosphere. The processes actually go on continuously with the arrival of galactic cosmic rays, but it has been estimated that the total production of NO during one major PCA event can be very great, even exceeding the annual production by cosmic rays. The great proton event of August 1972 had a measurable effect on the ozone concentration in the stratosphere, which fell by 15%–20% at latitudes 75°–80°. In the event of July 1982 ozone was depleted between 55 and 85 km. This was a relatively "soft" event, which explains why the effects were higher up. A series of PCA events that occurred in 1989 is also thought to have affected the ozone content. A computation of the effect of the events during that year is shown in Figure 7.45. The O_3 was depleted by more than 10% over a limited height range for several months in 1989, and small effects continued for a year or more. Significant though these effects are, they have no known effect on high-latitude radio propagation. Further information is given in papers by Reid (1986) and Jackman (2000).

7.4 Coherent scatter and the summer mesopheric echo

Incoherent and coherent scattering of radio waves in the ionosphere exploit different phenomena, as a result of which the second process is much the stronger (see Section 4.2.2). Given the utility of incoherent-scatter radar in ionospheric studies at high latitude, it would be a great pity if the weak signals which it uses were to be swamped by coherent echoes from the same region. Yet this is just what may occur.

Coherent echoes from the high-latitude D region were first detected in the VHF

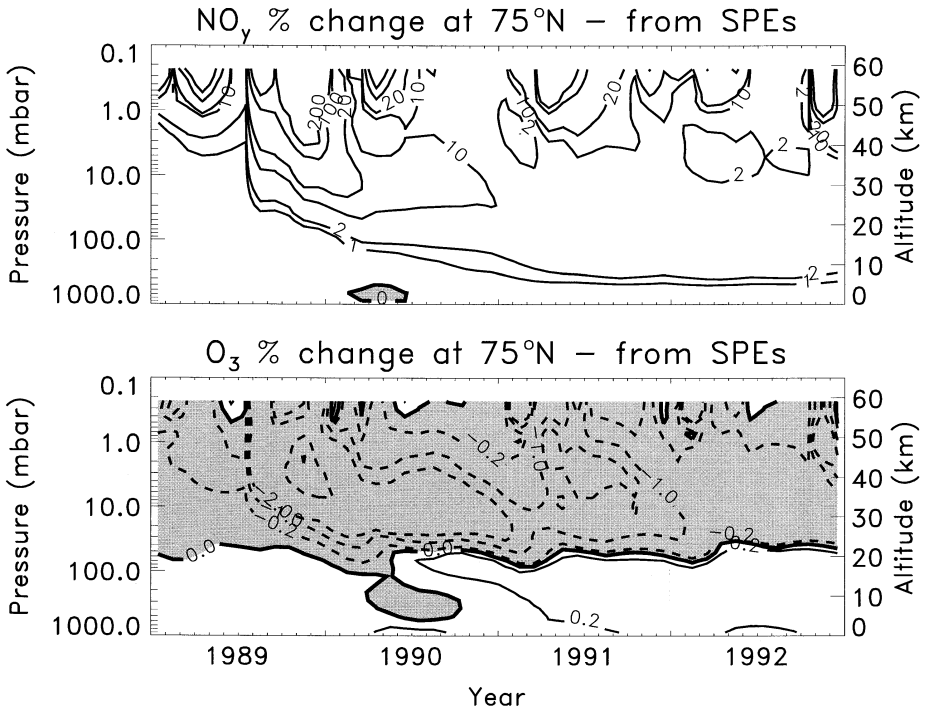


Figure 7.45. Computed variation of NO_y and O_3 concentrations at 75° north due to the solar proton events of 1989. The contours for NO_y are 0, 1, 2, 10, 20, 100 and 200%. For O_3 they are +2, +1, +0.2, 0, −0.2, −1, −2, −10 and −20%. The concentration of NO_y is increased but that of O_3 is decreased. Note the long duration of the effects. (C. H. Jackman *et al.*, *J. Geophys. Res.* **105**, 11659, 2000, copyright by the American Geophysical Union.)

band (at 50 MHz) in Alaska (Ecklund and Balsley, 1981), and subsequently in Norway at 53.5 MHz (Czechowsky *et al.*, 1989) and with the EISCAT 224-MHz radar (Hoppe *et al.*, 1988). They have also been observed, though less frequently, with the EISCAT UHF system at 933 MHz (Röttger *et al.*, 1990). Other observations cover the range 2.27 MHz to 1.29 GHz (Röttger, 1994). These strong echoes occur only in summer and are now usually called *polar mesosphere summer echoes (PMSEs)*. They are a nuisance to IS radar but constitute an interesting topic in their own right, particularly since they have proved to be something of a mystery.

Their characteristics are very different from those of the incoherent echoes received from the D region during particle precipitation. Not only are they much more intense, but also they are much narrower, usually less than 1.5 km deep, though there can also be multiple layers (Figure 7.46). The height range is more restricted, too, peaking at 84–86 km (Figure 7.47), an altitude close to the mesopause. When the echoes are present their height fluctuates (Figure 7.48), which is thought to indicate the passage of acoustic-gravity waves (Section 1.6). The spectrum of PMSE is considerably narrower than that of IS returns (Figure 7.49); even without the other evidence this point alone would be sufficient proof that quite different mechanisms are responsible.

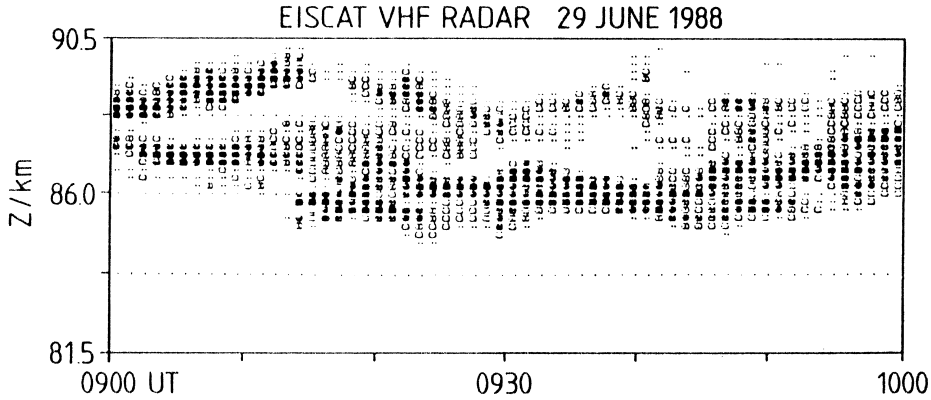


Figure 7.46. An example of PMSE observed at 224 MHz on 29 June 1988 using the EISCAT VHF radar. The density of blob suggests the strength of the echo. Note the height variations and the multiple layers. (Reprinted from P. N. Collis and J. Röttger, *J. Atmos. Terr. Phys.* **52**, 569, copyright 1990, with permission from Elsevier Science.)

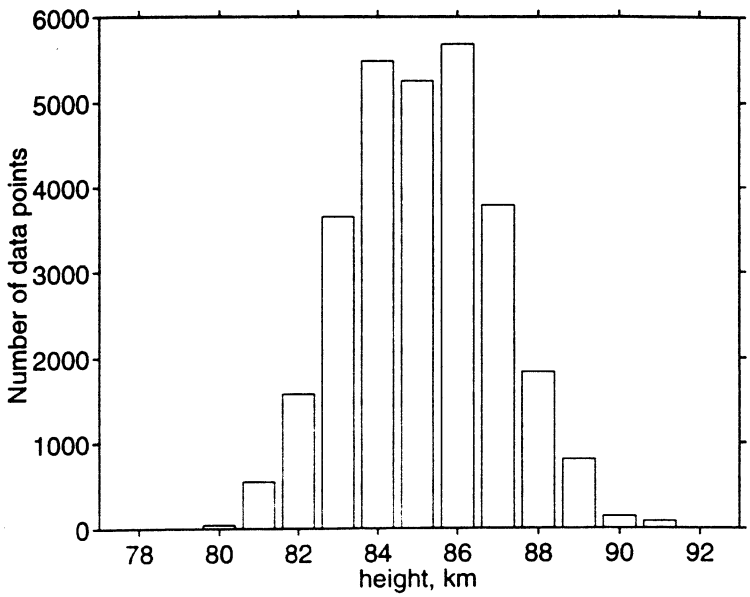


Figure 7.47. A histogram of the height distribution of PMSE observed with the EISCAT VHF radar. (Reprinted from J. R. Palmer *et al.*, *J. Atmos. Terr. Phys.* **58**, 307, copyright 1996, with permission from Elsevier Science.)

Most strikingly, the echoes are clearly a summer phenomenon, occurring from June to August only, with a maximum in July in the northern hemisphere (Palmer *et al.*, 1996). The occurrence also varies during the day. There are maxima near noon and midnight, and minima in the morning and the evening hours. The percentage occurrence, though not very well established, is some 50%–75% of days at

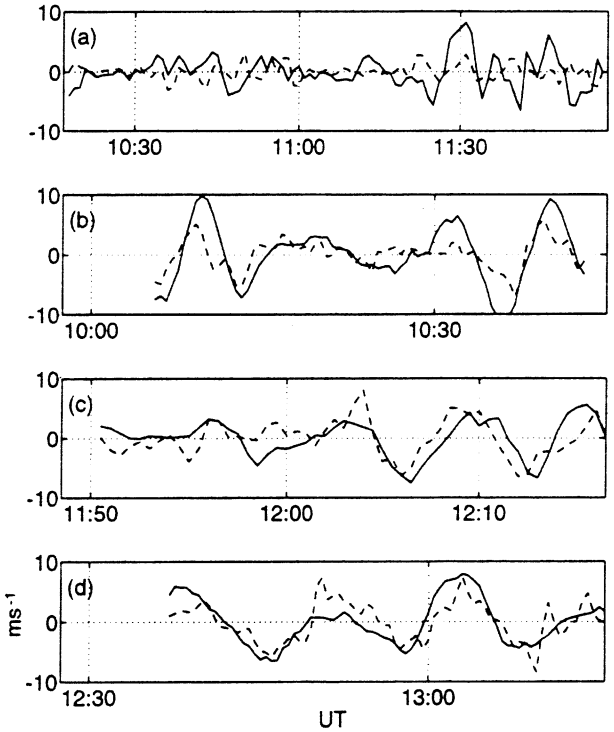


Figure 7.48. Rapid height fluctuations in PMSE, consistent with acoustic-gravity waves, on various dates in 1988 and 1991. The dashed curves show the rate of change of altitude, and the solid curves the vertical velocity derived from the Doppler shift of the echoes. (Reprinted from J. R. Palmer *et al.*, *J. Atmos. Terr. Phys.* **58**, 307, copyright 1996, with permission from Elsevier Science.)

the maxima and 10%–50% at the minima. The daily variations are most marked in June and August.

The polar mesosphere is particularly cold in the summer, and this may be the key to the mechanism. It has been proposed (Kelley *et al.*, 1987) that water-cluster ions, whose formation is favored by low temperature, reduce the diffusion coefficient of electrons and so extend the scale of turbulence, allowing coherent scatter to occur at shorter wavelengths. However, other mechanisms have also been proposed. The development of PMSE studies and the relevent theories have been reviewed by Cho and Kelley (1993) and by Röttger (1994).

7.5 Summary and implications

At middle and equatorial latitudes D-region absorption has only a minor effect on HF propagation, but at high latitude it can affect the signal strength profoundly. There are two basic types at high latitude, each having a separate cause and morphology. In its effect on radiowave propagation, auroral absorption (AA)

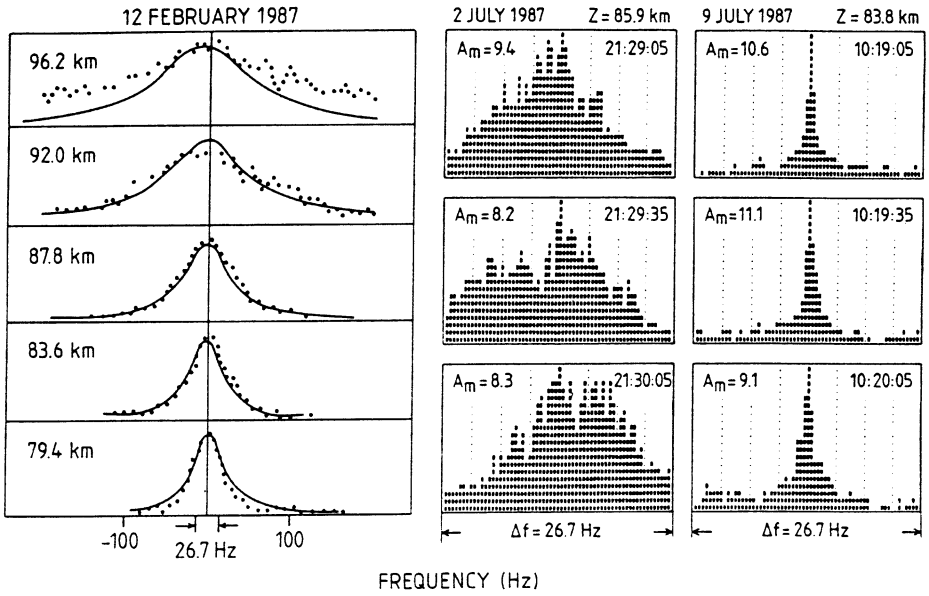


Figure 7.49. Spectra of incoherent scatter (IS) and PMSE obtained with the 224-MHz EISCAT VHF radar. The left-hand panels show typical IS spectra fitted by Lorentzian curves, and the centre and right-hand panels show broad and narrow PMSE spectra. Even the broadest PMSE spectra are considerably narrower than the IS spectra from the same height. (Reprinted from P. N. Collis and J. Röttger. *J. Atmos. Terr. Phys.* **52**, 569, copyright 1990, with permission from Elsevier Science.)

is of first-order importance. It may occur over a range of geomagnetic latitude from below 60° to above 75°, with a statistical maximum near 67°, and is patchy in its horizontal extent. The patches are tens to hundreds of kilometers in extent, and any elongation tends to be east–west. The diurnal occurrence peaks just before magnetic midnight and again in the morning sector between 0700 and 1000 magnetic LT.

Most of our knowledge about high-latitude absorption has come from several decades of observation by standard riometers (having beams about $\pm 60°$ between half-power points), though the earliest studies were based on ionosonde data. This information probably describes AA sufficiently well for the purposes of those HF communication systems which also use relatively broad beams. However, some modern HF systems (such as over-the-horizon radars and direction finders) require information on the finer structure of D-region absorption. The imaging riometers developed during the 1980s (and further deployed in the 1990s) have improved the spatial resolution considerably, and have the potential to provide information relevent to the high-resolution HF systems.

AA is a dynamic phenomenon, related, at least in part, to the auroral substorm; though almost certainly involving particle precipitation from the outer Van Allen belt in the day sector. The particles are electrons with energies from tens to hun-

dreds of kilo-electron volts – generally greater than those which produce the visual aurora. As with the aurora, there is probably some measurable AA somewhere in the auroral zone in any given period of 24 h. AA is essentially conjugate, occurring almost simultaneously (though not necessarily with the same intensity) in magnetically conjugate regions.

The other significant D-region absorption event at high latitude is polar-cap absorption (PCA), which may produce higher overall values of absorption than does AA but occurs much less frequently, only several times a year on the long-term average. PCA events are caused by the precipitation of 1–1000-MeV protons of solar origin into the polar D region. The occurrence and severity of PCA increases from solar minimum to maximum, and there may be ten or a dozen events in an active year. They produce a fairly uniform blanketing of the polar cap down to about 60° geomagnetic, and have been known to black out trans-polar HF propagation for 10 days at a time.

Both AA and PCA affect the lower frequencies more than they do the higher ones because the absorption varies (to a first approximation) as f^{-2}. At ELF and VLF, propagating in the waveguide mode, an increase in precipitation causes significant variation in the dimensions of the waveguide and thereby produces both amplitude and phase changes in the received signals.

7.6 References and bibliography

7.2 Auroral radio absorption

Agy, V. (1970) HF radar and auroral absorption. *Radio Sci.* **5**, 1317.

Ansari, Z. A. (1965) A peculiar type of daytime absorption in the auroral zone. *J. Geophys. Res.* **70**, 3117.

Appleton, E. V., Naismith, R., and Builder, G. (1933) Ionospheric investigations in high latitudes. *Nature* **132**, 340.

Bailey, D. K. (1968) Some quantitative aspects of electron precipitation in and near the auroral zone. *Rev. Geophys.* **6**, 289.

Berkey, F. T. (1968) Coordinated measurements of auroral absorption and luminosity using the narrow beam technique. *J. Geophys. Res.* **73**, 319.

Berkey, F. T., Driatskiy, V. M., Henriksen, K., Hultqvist, B., Jelly, D. H., Schuka, T. I., Theander, A., and Yliniemi, J. (1974) A synoptic investigation of particle precipitation dynamics for 60 substorms in IQSY (1964–65) and IASY (1969). *Planet. Space Sci.* **22**, 255.

Bewersdorff, A., Kremser, G., Stadnes, J., Trefall, H., and Ullaland, S. (1968) Simultaneous balloon measurements of auroral X-rays during slowly varying ionospheric absorption events. *J. Atmos. Terr. Phys.* **30**, 591.

Collis, P. N., Hargreaves, J. K., and Korth, A. (1984) Auroral radio absorption as an indicator of magnetospheric electrons and of conditions in the disturbed auroral D-region. *J. Atmos. Terr. Phys.* **46**, 21.

Collis, P. N., Hargreaves, J. K., and White, G. P. (1996) A localised co-rotating auroral absorption event observed near noon using imaging riometer and EISCAT. *Ann. Geophysicae* **14**, 1305.

Ecklund, W. L. and Hargreaves, J. K. (1968) Some measurements of auroral absorption structure over distances of about 300 km and of absorption correlation between conjugate regions. *J. Atmos. Terr. Phys.* **30**, 265.

Elkins, T. J. (1972) *A Model of Auroral Substorm Absorption*. Report AFCRL-72-0413. Air Force Cambridge Research Laboratories, Bedford, Massachusetts.

Foppiano, A. J. and Bradley, P. A. (1984) Day-to-day variability of riometer absorption. *J. Atmos. Terr. Phys.* **46**, 689.

Foppiano, A. J. and Bradley, P. A. (1985) Morphology of background auroral absorption. *J. Atmos. Terr. Phys.* **47**, 663.

Friedrich, M. and Torkar, K. M. (1983) High-latitude plasma densities and their relation to riometer absorption. *J. Atmos. Terr. Phys.* **45**, 127.

Friedrich, M. and Kirkwood, S. (2000) The D-region background at high latitudes. *Adv. Space Res.* **25**, 15.

Hajkovicz, L. A. (1990) The dynamics of a steep onset in the conjugate auroral riometer absorption. *Planet. Space Sci.* **38**, 127.

Hargreaves, J. K. (1966) On the variation of auroral radio absorption with geomagnetic activity. *Planet. Space Sci.* **14**, 991.

Hargreaves, J. K. (1967)Auroral motions observed with riometers: movements between stations widely separated in longitude. *J. Atmos. Terr. Phys.* **29**, 1159.

Hargreaves, J. K. (1968) Auroral motions observed with riometers: latitudinal movements and a median global pattern. *J. Atmos. Terr. Phys.* **30**, 1461.

Hargreaves, J. K. (1969a) Auroral absorption of HF radio waves in the ionosphere: a review of results from the first decade of riometry. *Proc. Inst. Elect. Electronics Engineers* **57**, 1348

Hargreaves, J. K. (1969b) Conjugate and closely-spaced observations of auroral radio absorption – I. Seasonal and diurnal behaviour. *Planet. Space Sci.* **17**, 1459.

Hargreaves, J. K. (1970) Conjugate and closely-spaced observations of auroral radio absorption – IV. The movement of simple features. *Planet. Space Sci.* **18**, 1691.

Hargreaves, J. K. (1974) Dynamics of auroral absorption in the midnight sector – the movement of absorption peaks in relation to the substorm onset. *Planet. Space Sci.* **22**, 1427.

Hargreaves, J. K. and Chivers, H. J. A. (1964) Fluctuations in ionospheric absorption events at conjugate stations. *Nature* **203**, 963.

Hargreaves, J. K. and Sharp, R. D. (1965) Electron precipitation and ionospheric radio absorption in the auroral zones. *Planet. Space Sci.* **13**, 1171.

Hargreaves, J. K. and Cowley, F. C. (1967a) Studies of auroral radio absorption events at three magnetic latitudes. 1. Occurrence and statistical properties of the events. *Planet. Space Sci.* **15**, 1571.

Hargreaves, J. K. and Cowley, F. C. (1967b) Studies of auroral radio absorption events at three magnetic latitudes. 2. Differences between conjugate regions. *Planet. Space Sci.* **15**, 1585.

Hargreaves, J. K. and Ecklund, W. L. (1968) Correlation of auroral radio absorption between conjugate points. *Radio Sci.* **3**, 698.

Hargreaves, J. K., Chivers, H. J. A., and Axford, W. I. (1975) The development of the substorm in auroral radio absorption. *Planet. Space Sci.* **23**, 905.

Hargreaves, J. K. and Berry, M. G. (1976) The eastward movement of the structure of auroral radio absorption events in the morning sector. *Ann. Geophysicae* **32**, 401.

Hargreaves, J. K., Taylor, C. M., and Penman, J. M. (1982) *Catalogue of Auroral Radio Absorption During 1976–1979 at Abisko, Sweden*. World Data Center A, US Department of Commerce, Boulder, Colorado.

Hargreaves, J. K., Feeney, M. T., Ranta, H. and Ranta, A. (1987) On the prediction of auroral radio absorption on the equatorial side of the absorption zone. *J. Atmos. Terr. Phys.* **49**, 259.

Hargreaves, J. K. and Devlin, T. (1990) Morning sector precipitation events observed by incoherent scatter radar. *J. Atmos. Terr. Phys.* **52**, 193.

Hargreaves, J. K., Detrick, D. L., and Rosenberg, T. J. (1991) Space-time structure of auroral radio absorption events observed with the imaging riometer at South Pole. *Radio Sci.* **26**, 925.

Hargreaves, J. K., Browne, S., Ranta, H., Ranta, A. Rosenberg, T. J., and Detrick, D. L. (1997) A study of substorm-associated nightside spike events in auroral absorption using imaging riometers at South Pole and Kilpisjärvi. *J. Atmos. Solar–Terrestrial Phys.* **59**, 853.

Hartz, T. R., Montbriand, L. E. and Vogan, E. L. (1963) A study of auroral absorption at 30 Mc/s. *Can. J. Phys.* **41**, 581.

Hartz, T. R. and Brice, N. M. (1967) The general pattern of auroral particle precipitation. *Planet. Space Sci.* **15**, 301.

Holt, O., Landmark, B., and Lied, F. (1961) Analysis of riometer observations obtained during polar radio blackouts. *J. Atmos. Terr. Phys.* **23**, 229.

Jelly, D. H., Matthews, A. G., and Collins, C. (1961) Study of polar cap and auroral absorption. *J. Atmos. Terr. Phys.* **23**, 206.

Jelly, D. H., McDiarmid, I. B., and Burrows, J. R. (1964) Correlation between intensities of auroral absorption and precipitated electrons. *Can. J. Phys.* **42**, 2411.

Jelly, D. H. (1970) On the morphology of auroral absorption during substorms. *Can. J. Phys.* **48**, 335.

Kavadas, A. W. (1961) Absorption measurements near the auroral zone. *J. Atmos. Terr. Phys.* **23**, 170.

Leinbach, H. and Basler, R. P. (1963) Ionospheric absorption of cosmic radio noise at magnetically conjugate auroral zone stations. *J. Geophys. Res.* **68**, 3375.

Little, C. G. and Leinbach, H. (1958) Some measurements of high-latitude ionospheric absorption using extraterrestrial radio waves. *Proc. IRE* **46**, 334.

Little, C. G., Schiffmacher, E. R., Chivers, H. J. A., and Sullivan, K. W. (1965) Cosmic noise absorption events at geomagnetically conjugate stations. *J. Geophys. Res.* **70**, 639.

Nielsen, E. (1980) Dynamics and spatial scale of auroral absorption spikes associated with the substorm expansion phase. *J. Geophys. Res.* **85**, 2092.

Parthasarathy, R. and Berkey, F. T. (1965) Auroral zone studies of sudden onset radio wave absorption events using multiple station and multiple frequency data. *J. Geophys. Res.* **70**, 89.

Parthasarathy, R., Berkey, F. T., and Venkatesan, D. (1966)Auroral zone electron flux and its relation to broadbeam radiowave absorption. *Planet. Space Sci.* **14**, 65.

Penman, J. M., Hargreaves, J. K., and McIlwain, C. E. (1979) The relation between 10 to 80 keV electron precipitation observed at geosynchronous orbit and auroral radio absorption observed with riometers. *Planet. Space Sci.* **27**, 445.

Pudovkin, M. I., Shumilov, O. I., and Zaitseva, S. A. (1968) Dynamics of the zone of corpuscular precipitations. *Planet. Space Sci.* **16**, 881.

Ranta, H., Ranta, A., Collis, P. N., and Hargreaves, J. K. (1981) Development of the auroral absorption substorm: studies of the pre-onset phase and sharp onset using an extensive riometer network. *Planet. Space Sci.* **29**, 1287.

Stauning, P. and Rosenberg, T. J. (1996) High-latitude daytime absorption spike events. *J. Geophys. Res.* **101**, 2377.

7.3 The polar cap event

Akasofu, S.-I. and Chapman, S. (1972) *Solar–Terrestrial Physics*. Oxford University Press, Oxford.

Bailey, D. K. (1959) Abnormal ionization in the lower ionosphere associated with cosmic-ray flux enhancements. *Proc. IRE* **47**, 255.

Bakshi, P. and Barron, W. (1979) Prediction of solar proton spectral slope from radio burst data. *J. Geophys. Res.* **84**, 131.

Castelli, J. P., Aarons, J., and Michael, G. A. (1967) Flux density measurements of radio bursts of proton-producing flares and nonproton flares. *J. Geophys. Res.* **72**, 5491.

Chivers, H. J. A. and Hargreaves, J. K. (1965) Conjugate observations of solar proton events: delayed ionospheric changes during twilight. *Planet. Space Sci.* **13**, 583.

Collis, P. N. and Rietveld, M. T. (1990) Mesospheric observations with the EISCAT UHF radar during polar cap absorption events: 1. Electron densities and negative ions. *Ann. Geophys.* **8**, 809.

Gillmor, C. S. (1963) The day-to-night ratio of cosmic noise absorption during polar cap absorption events. *J. Atmos. Terr. Phys.* **25**, 263.

Hargreaves, J. K., Ranta, H., Ranta, A., Turunen, E., and Turunen, T. (1987) Observation of the polar cap absorption event of February 1984 by the EISCAT incoherent scatter radar. *Planet. Space Sci.* **35**, 947.

Hargreaves, J. K., Shirochkov, A. V., and Farmer, A. D. (1993) The polar cap absorption event of 19–21 March 1990: recombination coefficients, the twilight transition and the midday recovery. *J. Atmos. Terr. Phys.* **55**, 857.

Hultqvist, B. (1969) Polar cap absorption and ground level effects. *Solar Flares and Space Research* (eds. C. de Jager and Z. Svestka), p. 215. North-Holland, Amsterdam.

Jackman, C. H., Fleming, E. L., and Vitt, F. M. (2000) Influence of extremely large proton events in a changing stratosphere. *J. Geophys. Res.* **105**, 11659.

Leinbach, H. (1967) Midday recoveries of polar cap absorption. *J. Geophys. Res.* **72**, 5473.

Obayashi, T. (1959) Entry of high energy particles into the polar ionosphere. *Rep. Ionosphere Space Res. Japan* **13**, 201.

Ranta, H., Ranta, A., Yousef, S. M., Burns, J., and Stauning, P. (1993) D-region observations of polar cap absorption events during the EISCAT operation in 1981–1989. *J. Atmos. Terr. Phys.* **55**, 751.

Reagan, J. B. and Watt, T. M. (1976) Simultaneous satellite and radar studies of the D-region ionosphere during the intense solar particle events of August 1972. *J. Geophys. Res.* **81**, 4579.

Reid, G. C. (1961) A study of the enhanced ionisation produced by solar protons during a polar cap absorption event. *J. Geophys. Res.* **66**, 4071.

Reid, G. C. (1967) Ionospheric disturbances. In *Physics of Geomagnetic Phenomena* (eds. Matsushita and Campbell), p. 627. Academic Press, New York.

Reid, G. C. (1986) Solar energetic particles and their effects on the terrestrial environment. In *Physics of the Sun* (ed. P. A. Sturrock), vol. 3, p. 251. Reidel, Dordrecht.

Reid, G. C. and Sauer, H. H. (1967) The influence of the geomagnetic tail on low-energy cosmic-ray cutoffs. *J. Geophys. Res.* **72**, 197.

Sauer, H. H. (1968) Nonconjugate aspects of recent polar cap absorption events. *J. Geophys. Res.* **73**, 3058.

Shea, M. A. and Smart, D. F. (1977) Significant solar proton events, 1955–1969. In *Solar–Terrestrial Physics and Meterology: Working Document II*, p. 119. SCOSTEP.

Shea, M. A. and Smart, D. F. (1979) Significant solar proton events, 1970–1972. In *Solar–Terrestrial Physics and Meterology: Working Document III*, p. 109. SCOSTEP.

Shea, M. A. and Smart, D. F. (1995) Solar proton fluxes as a function of the observation location with respect to the parent solar-activity. *Adv. Space Res.* **17**, 225.

Smart, D. F. and Shea, M. A. (1989) Solar proton events during the past three solar cycles. *Spacecraft and Rockets* **26**, 403.

Smart, D. F. and Shea, M. A. (1995) The heliolongitudinal distribution of solar-flares associated with solar proton events. *Adv. Space Res.* **17**, 113.

Uljev, V. A., Shirochkov, A. V., Moskvin, I. V., and Hargreaves, J. K. (1995) Midday recovery of the polar cap absorption of March 19–21, 1990: a case study. *J. Atmos. Terr. Phys.* **57**, 905.

Weeks, L. H., CuiKay, R. S., and Corbin, J. R. (1972) Ozone measurements in the mesosphere during the solar proton event of 2 November 1969. *J. Atmos. Sci.* **29**, 1138.

7.4 Coherent scatter and the polar mesosphere summer echo

Cho, J. Y. N. and Kelley, M. C. (1993) Polar mesosphere summer radar echoes: observations and current theories. *Rev. Geophys.* **31**, 243.

Collis, P. N. and Röttger, J. (1990) Mesospheric studies using EISCAT UHF and VHF radars: a review of principles and experimental results. *J. Atmos. Terr. Phys.* **52**, 569.

Czechowsky, P., Reid, I. M., Ruster, R., and Schmidt, S. (1989) VHF radar echoes

observed in the summer and winter polar mesosphere over Andøya, Norway. *J. Geophys. Res.* **94**, 5199.

Ecklund, W. L. and Balsley, B. B. (1981) Long-term observations of the Arctic mesosphere with the MST radar at Poker Flat, Alaska. *J. Geophys. Res.* **86**, 7775.

Hoppe, U.-P., Hall, C., and Röttger, J. (1988) First observations of summer polar mesospheric back-scatter with a 224 MHz radar. *Geophys. Res. Lett.* **15**, 28.

Kelley, M. C., Farley D. T., and Röttger, J. (1988) The effect of cluster ions on anomalous VHF back-scatter from the summer polar mesosphere. *Geophys. Res. Lett.* **14**, 1031.

Palmer, J. R., Rishbeth, H., Jones, G. O. L., and Williams, P. J. S. (1996) A statistical study of polar mesosphere summer echoes observed by EISCAT. *J. Atmos. Terr. Phys.* **58**, 307.

Röttger, J. (1994) Polar mesosphere summer echoes: dynamics and aeronomy of the mesosphere. *Adv. Space Res.* **14**, 123.

Röttger, J., Rietveld, M. T., La Hoz, C., Hall, T., Kelley, M. C., and Swartz, W. E. (1990) Polar mesosphere summer echoes observed with the EISCAT 993-MHz radar and the CUPRI 46.4-MHz radar, their similarity to 224-MHz radar echoes, and their relation to turbulence and electron density profiles. *Radio Sci.* **25**, 671.