Modelling high-latitude ionosphere for time-varying plasma convection

M. Lockwood, PhD

Indexing terms: Radiowave propagation (ionospheric), Convection

Abstract: The paper discusses how variations in the pattern of convective plasma flows should be included in self-consistent time-dependent models of the coupled ionosphere-thermosphere system. The author shows how these variations depend upon the mechanism by which the solar wind flow excites the convection. The modelling of these effects is not just of relevance to the polar ionosphere. This is because the influence of convection is not confined to high latitudes: the resultant heating and composition changes in the thermosphere are communicated to lower latitudes by the winds which are also greatly modified by the plasma convection. These thermospheric changes alter the global distribution of plasma by modulating the rates of the chemical reactions which are responsible for the loss of plasma. Hence the modelling of these high-latitude processes is of relevance to the design and operation of HF communication, radar and navigation systems worldwide.

1 Introduction

The solar wind plasma flows around the Earth's magnetic field, giving a lower density cavity called the magnetosphere. As it does so, it generates large scale electric fields which cause circulation of plasma in the outer magnetosphere. Averaged over periods of at least several hours, the magnetospheric circulation maps down the magnetic field lines into the high-latitude ionosphere. (On shorter times scales, the ionospheric flow is not a simple image of the magnetospheric flow because the magnetosphere stores energy and then releases it as explosive events called substorms) [1]. The average circulation of high-latitude ionospheric plasma is termed 'convection' and is now known to control the three-dimensional distribution of plasma and its temporal variations. Its influence is greatest at high latitudes but extends to cover the globe. Because HF systems make use of ionospheric reflections to give over-the-horizon propagation, the operations of HF communication, radar and navigation systems worldwide will therefore be influenced by convection.

In the lower ionosphere (the E region), the lifetime of free ion–electron pairs is short (of order seconds). On application of a typical convection electric field, free charges move distances which are usually less than about one kilometre during this lifetime. As a result, plasma decays in roughly the same location as it was produced and the spatial distribution of E-region plasma density reflects that of the production sources of solar photon (EUV and X-ray) radiations and auroral particle precipitation. However, at the greater altitudes of the F-region, the plasma lifetimes are typically several hours, during which the plasma will convect over several thousand kilometres, taking it between and across the various regions of enhanced plasma production (the dayside and the auroral zones) and loss (the nightside). However, its density will evolve only slowly because the plasma lifetime is long. Hence the spatial distribution of the plasma will be considerably more complex than that in the E-region and will be directly controlled by the convection pattern. The importance of convection to the high-latitude plasma distribution was first pointed out by Knudsen in 1974 [2] and modelled numerically by Knudsen *et al.* in 1977 [3]. Subsequently, various models of the convecting high-latitude ionosphere have been developed and these allow for the chemical and dynamic coupling with the neutral thermosphere with varying degrees of self consistency [4–6]. The general principles on which these models are based are demonstrated by Fig. 1 (after Quegan *et al.* [7]). The solid lines are F-region plasma streamlines and the distances between the dots on each streamline are those travelled in 1 h periods, commencing from the larger dots in the dawn sector (near 06 h, local time). This steady-state flow pattern is shown in a frame fixed with respect to the sun, and the sunward direction is to the top of the Figure. If transposed into a frame which is corotating with the Earth, this flow pattern would be broadly symmetric and twin-celled (as shown in the centre of the top panel of Fig. 4). The nature and origin of this flow will be discussed in Section 3. The equator-most streamline is a closed loop containing 25 smaller dots, i.e. the plasma corotates with roughly a 25.5 h period. At lower (middle) latitudes this

The author thanks many scientists for discussions of this work, in particular: Dr. S. Quegan of Sheffield University; Dr. T.J. Fuller-Rowell of the Solar Environment Laboratory, Boulder; and Prof. S.W.H. Cowley of Imperial College, London. He is grateful to the first two of the above for simulations carried out on the Cray XMP supercomputer at RAL, using the UCL/Sheffield University time-dependent, three-dimensional coupled ionosphere–thermosphere model. Funding for this work was provided in part by the National Radio Propagation Programme and in part by the Science and Engineering Research Council.

Paper 9252H (E11), first received 9th January 1991 and in revised form 4th September 1992

The author is with the Space Science Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon., OX11 0QK, United Kingdom

period reduces to 24 h. As this plasma moves between the dayside and the nightside it undergoes a regular diurnal cycle of enhancement and decay. This behaviour is expected throughout the lower latitudes from which the

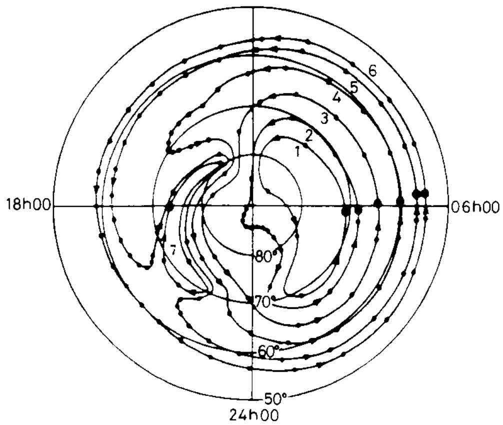

Fig. 1 *Steady-state pattern of ionospheric plasma convection in a frame of reference fixed with respect to the sun*

The dots on the plasma streamlines define distances travelled in 1-hour periods, commencing from the larger dot near dawn. The geomagnetic and rotational poles are both placed at the centre of the plot for simplicity and local noon is at the top of the Figure. After Quegan *et al.* [7]

electric fields generating convection are shielded and hence where plasma corotates with the Earth. However, the other streamlines shown [1–5 and 7] are all strongly influenced by convection and all traverse the polar cap in an antisunward direction. Although the plasma still follows closed circuits in this steady-state model, the time to complete one circuit varies and the plasma moves between night and day with a wide variety of periods.

As an example of the possible effects of these more complex flow paths, consider streamline 5 which passes close to a stagnation point in the postdusk sector (the bottom left quadrant of the Figure) where convection and corotation are equal in magnitude and opposite in direction. The proximity of the dots on the streamline indicates that the plasma remains in this region for a prolonged period. If this stagnation region is in darkness (which it will be at all times other than in summer), the plasma density will decay and very low values will be seen in, and on streamlines emanating from, this region. Such a depletion is indeed observed and is termed the 'main', or 'midlatitude' ionospheric trough.

Conversely, if we consider streamline 1, we find the plasma completes the closed circuit in only 5 h. After moving through the dayside and being enhanced by solar photoionisation, the plasma passes into the polar cap through a constriction termed the 'throat'. In the vicinity of this convection throat dayside auroral particles precipitate into the ionosphere [10]. These 'cusp' and 'cleft' auroral particles originate from the solar wind and precipitate without significant acceleration. Their energy is such that they would produce enhanced ionisation most readily in the F-region. On exiting the polar cap, the streamlines pass through the nightside auroral precipitation region. This second population of auroral particles has been energised within the magnetosphere to rather greater energies than the cusp particles and they penetrate to lower altitudes. Hence they produce plasma predominantly in the E-region, but do also cause some F-region enhancement. The lack of time for significant plasma decay, and the passage through regions of significant plasma production, mean that plasma on streamline 1 would always have relatively high densities.

For simplicity, Fig. 1 colocates the magnetic and rotational poles, whereas in reality there is an offset of about 10° (slightly larger in the southern hemisphere). The introduction of the diurnal orbit of the magnetic pole (about which the convection pattern is centred) around the rotational pole causes some interesting and complex effects. For example, at around 17 UT, the magnetic pole (and hence the entire convection pattern) in the northern hemisphere is tipped sunward. This means that the throat region is in the sunlit hemisphere and plasma previously enhanced by solar photoionisation will be convected antisunward, producing a tongue of high-density plasma over the pole or more localised 'patches' (see review by Tsunoda [8]). However, 12 hours later, the throat will be tipped antisunward and in winter would not be sunlit: at such times neither the tongue nor patches are present. Initially it was thought that this high density plasma in the polar cap was produced by the cusp precipitation, but more detailed analysis has shown that it is produced by photoionisation on the dayside and subsequently convected poleward [9]. The motion of the day–night terminator means that behaviour will depend strongly on the season and the altitude profile of photoionisation introduces a major solar cycle variation through the effects of atmospheric heating.

Further, important effects are produced by the fact that the convecting ions collide with the neutral gases of the thermosphere, causing heating and upwelling of molecular gases, as well as driving winds in the thermosphere. Such winds are in addition to those driven by the day–night pressure difference due to solar heating effects. Winds influence the loss rate of the plasma by altering the rate of the chemical reactions responsible for plasma decay by blowing plasma up/down field lines (depending on the direction of the winds) to where plasma loss is slower/faster, respectively. However, the winds driven depend not only on the ion flow speed but also on the plasma densities themselves. Hence it can be seen that the ionosphere–thermosphere system is highly coupled. The neutral winds are subject to Coriolis and centrifugal forces, which are ordered about the rotational pole, whereas the ion flows and particle precipitations are ordered by the magnetic field about the magnetic poles. Consequently, complex UT effects are introduced by the offset of these two poles.

The convection electric fields also modulate the ionospheric densities directly, by heating the ionospheric and neutral gases and thereby increasing the rate at which the plasma loss reactions proceed. The increase of the loss rate with increasing electric field can result in localised troughs forming where the plasma drift is large (where streamlines are bunched together in Fig. 1).

However, in all the above discussion, the convection pattern was assumed to be nonvarying. The time dependence arose purely from the offset of the magnetic and rotation poles. In such cases, the ionosphere undergoes regular 24 h period oscillations, i.e. steady state, in a diurnal sense, is established. However, recent observations have shown that the convection pattern should not be thought of as being steady; indeed it appears steady convection can be present for only 15% of the time at the very most [1, 11, 12]. This paper discusses how departures from steady-state convection should be allowed for in numerical models of the high-latitude ionosphere.

2 Variability of the high-latitude ionosphere and some effects on radio systems

The numerical models of the high-latitude ionosphere predict that the distribution of the plasma will be complex and variable as a result of all the competing processes described in the previous Section. Fig. 2 shows some data from the European Incoherent Scatter (EISCAT) radar in northern Scandinavia which demonstrate that this is indeed the case in practice. The plot shows contours of plasma density at 325 km altitude, observed on four successive days, each panel showing the data for a full day as a function of latitude and Universal Time (UT). The local time is roughly (UT + 2) hours. On 12–13 February (top panel) the main trough is seen moving equatorward in the premidnight sector (16–22 UT), this motion reflecting the orientation of the auroral oval and the convection pattern. High densities are seen around local noon (10 UT). This behaviour is not uncommon in EISCAT data; similar examples observed near the autumn equinox have been presented by Collis and Häggström [13]. The trough is seen on the next day (second panel), but is less deep and moves equatorward earlier and more rapidly, indicating an expanded auroral oval and convection pattern. A patch of plasma with densities as high as on the dayside is then seen in the postmidnight sector, and the dayside peak is shorterlived and less dense than on the previous day. The third day is similar to the first, other than the midlatitude trough being seen to migrate poleward again in the postmidnight sector, whereas the fourth day is considerably different, with highly structured patches and troughs before midnight, a depletion at all latitudes after midnight, a small dayside peak in the midafternoon sector (13 UT) followed by an early appearance of the equatorward-moving main trough.

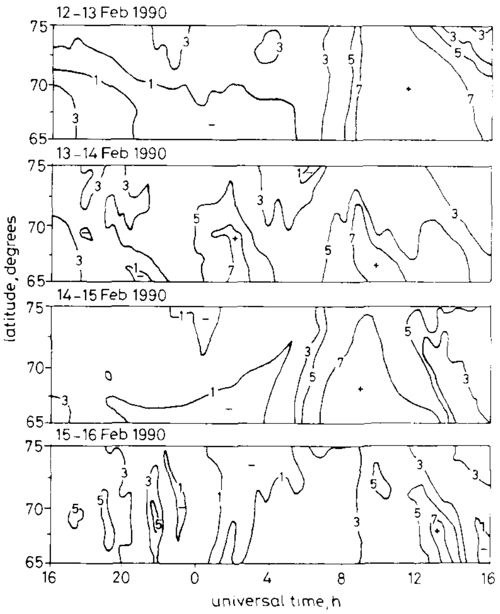

Fig. 2 *EISCAT observations of the plasma density at 325 km altitude on four successive days*

Plasma density contours (in units of 10^9 m^{-3}) are shown as a function of latitude and Universal Time (UT)

This sequence illustrates that the high-latitude ionosphere varies on a day-to-day basis and over shorter periods, but that some features, although variable, are present on a regular basis (like the midlatitude trough). The major reason for this is the variability of the plasma convection. The convection flows are observed by EISCAT, by receiving the radar echoes at three sites. The flow data corresponding to the densities shown in Fig. 2 are plotted in Fig. 3. It can be seen that the flows in the first and third days were mainly weak and irregular,

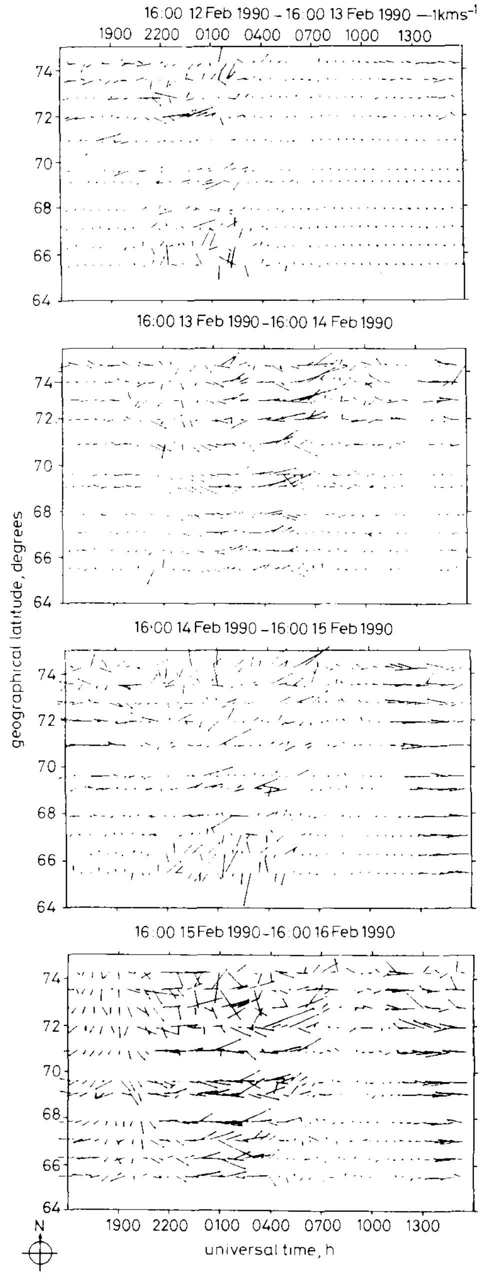

Fig. 3 *Plasma convection observed by EISCAT during the intervals shown in Fig. 2*

explaining the similarity of the densities observed on these two days. However, the convection is considerably stronger and more regular on the night of 13–14 February and after 13 UT on 15 February. These results do not prove that the different density behaviours result from the different convection conditions, but they do illustate the variability of both.

There are a number of ways in which such density variations influence the operation of a number HF, VHF and UHF systems. For example, the maximum usable frequency (MUF) of a point-to-point HF communications link depends upon the plasma density at the F-layer peak and varies at high latitudes during phenomena such as substorms [14], when nightside convection is greatly enhanced. In addition, the density contours shown in Fig. 2 reveal many steep gradients in the plasma density at a fixed height. Such F-layer tilts cause deviations of HF raypaths [15]. In addition to causing errors in locations determined by over-the-horizon radars [16], F-layer tilts can induce great changes in HF communications links, for example by making raypaths pass through regions of higher plasma density in the lower D-region [17]: such regions are caused by auroral precipitation or by energetic precipitation in the polar cap ('PCA events') and greatly increase the absorption of the radiowaves. Large and persistent features like the midlatitude trough have also been shown to focus HF waves [18], increasing the received signal strengths and lowering the lowest usable frequency.

In addition to these raypath effects, large convection drifts and spatial plasma gradients are known to cause a wide spectrum of intense, small-scale irregularities in the high-latitude ionosphere. These scatter HF, VHF and UHF waves and hence have a number of effects on both terrestrial and satellite systems. They cause amplitude and phase scintillation of transionospheric propagation, reducing the data communication rate with satellites. They also cause clutter on over-the-horizon radar data and can extend the maximum usable frequency of communication systems to higher values. A full review of the causes and effects of high-latitude irregularities has recently been published by Tsunoda [8].

3 Nature and causes of convection

Comparison with data from satellites monitoring the interplanetary medium has shown that the form of the convection pattern and the strength of the flows are strongly dependent on the orientation of the interplanetary magnetic field (IMF) embedded within the solar wind flow. This was found using convection measurements by polar-orbiting satellites [19–21], by ground-based radars [22], by ground-based magnetometers [23] and by combinations of these [24]. These results are summarised in Fig. 4 which shows the average patterns of convection for different orientations of the IMF vector in the Geocentric Solar Magnetospheric (GSM) coordinate system: The B_z component of the field is that aligned with the northward direction along the Earth's magnetic dipole axis, the B_x component is sunward, and B_y makes up the right-handed set and is therefore duskward. The Figure shows flow streamlines in invariant magnetic latitude (Λ): Magnetic Local Time (MLT) frames with local magnetic noon at the top and the magnetic pole ($\Lambda = 90°$) at the centre. It should be noted that the flow patterns are as would be seen by an observer who is corotating with the Earth: Fig. 1 shows the pattern for $B_z < 0$, $B_y = 0$ (centre, top panel of Fig. 4) transposed into a reference frame which is fixed with respect to the sun. The dashed lines show the poleward and equatorward edges of the zone of auroral particle precipitation. It can be seen that the IMF B_y component introduces dawn–dusk asymmetries, while B_z controls the size, strength and number of the convection cells. In par-

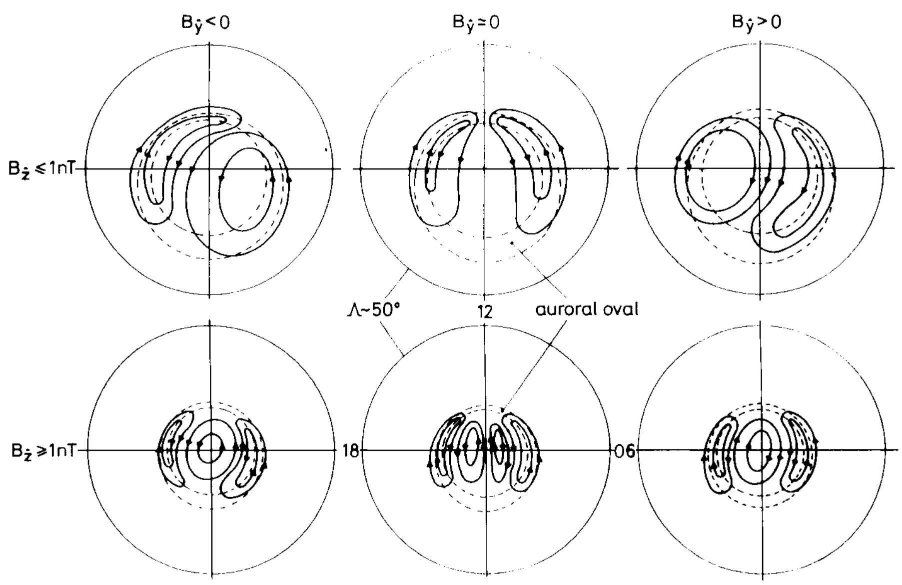

Fig. 4 *Inferred steady-state convective flow patterns in the northern hemisphere for various orientations of the IMF, as seen by an observer corotating with the Earth*

The patterns are given in invariant magnetic latitude (Λ) — magnetic local time (MLT) frames with the magnetic pole at the centre and noon at the top
⊙ B_z = northward
← B_y = dawn-to-dusk component of IMF

ticular, flows are faster and more extensive when the IMF points southward ($B_z < 0$).

Recently, there has been much interest in how the pattern of flow associated with a particular IMF orientation evolves into that for a different orientation following a sudden change between the two. Lockwood *et al.* [11] have discussed these observations and their implications for how ionospheric convection is excited. The results show that convection is very variable on short (subhour) time scales and that a steady-state convection pattern, of the types shown in Fig. 4, can only be established after more than about 2 h of stable IMF. However, a survey of 24 years' data on the IMF has shown that the IMF is stable for more than 2 h for only 15% of the time [12]: hence steady convection will be relatively rare. More often, the convection pattern at any instant will be considerably different from the steady-state patterns shown in Fig. 4, which represent long-term averages [25].

Fig. 5, adapted from Cowley (1984) [26], illustrates the two main types of mechanism which drive ionospheric convection. In Fig. 5*a* we see the Earth and the GSM reference frame where x is sunward and z is northward. The dashed lines mark the boundary of the magnetosphere, the magnetopause, in the xy and xz planes. The tube marked 'o' is an open magnetic flux tube which contains geomagnetic field lines which have been interconnected with the IMF by the process of magnetic reconnection [27] at some point near the subsolar magnetopause. Open field lines therefore thread the magnetopause. The reconnection takes place at a line in the magnetopause, called the dayside 'x-line' and marked D in Fig. 5. The flux tube marked 'c' is an example of a closed tube which means that it contains geomagnetic field lines which connect the inospheres of opposite hemispheres and does not thread the magnetopause. Magnetic reconnection near the subsolar magnetopause (at D) is most efficient if the IMF is directed southward. Once connected, an open tube is dragged antisunward by the solar wind flow and stretched out in the geomagnetic tail. Eventually, it reconnects with the corresponding open flux tube in the other hemisphere (at a second X-line in the nightside magnetosphere marked N) to form a closed field line again and returns sunward under the influence of the stored magnetic tension. As first suggested by Dungey [27], this cycle produces the basic two-celled convection cells in the ionosphere, shown in Fig. 5*c* on unshaded areas. The ionospheric projections of the reconnection x-lines (called the 'merging gaps') are marked D and N in Fig. 5*c*. Fig. 5*b* shows the flows in the equatorial (xy) plane (only sunward flows are seen in the unshaded regions here as the antisunward motions of open field lines are out of this plane, as shown in Fig. 5*a*. We call these flows 'reconnection-driven'.

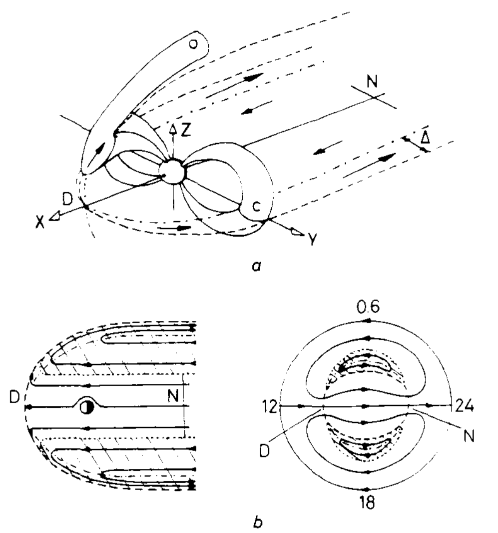

Fig. 5 *Illustration of the two processes which excite convection. Dayside and nightside reconnection x-lines (and in the case of Fig. 5c, the merging gaps) are labelled D and N, respectively. After Cowley [26]*

a 'Open' flux tube o interconnected with a southward IMF by magnetic reconnection and hence dragged antisunward by the solar wind flow. Closed flux tube c on the flanks of the magnetosphere and moved antisunward by viscous-like interaction across the magnetopause

b Flows in the equatorial plane of the magnetosphere

c Flows in the ionosphere: viscously driven flows are in the cross-hatched areas whereas reconnection driven flows are in the unhatched areas

- - - - magnetopause (in the case of Fig. 5*c*, its projection along magnetic field lines into the ionosphere)

–·–·– convection reversal boundaries

The second type of mechanism acts only on closed flux tubes on the flanks of the magnetosphere (like that marked c in Fig. 5*a*). Here momentum is transferred across the magnetopause from the solar wind and closed field lines are dragged into the tail. There they may drift away from the boundary and become less influenced by the solar wind flow and eventually return sunward under the influence of magnetic tension. Alternatively, the closed field lines in the tail could be 'pinched off' by reconnection, forming less extended closed field lines and magnetic loops which drift antisunward into interplanetary space [28]. The resulting flows in the ionosphere and in the equatorial magnetosphere are shown in the shaded regions in Figs. 5*c* and 5*b*, respectively. There have been a number of mechanisms proposed whereby momentum is transferred from the magnetosheath to closed field lines, including: Kelvin–Helmholtz waves on the magnetopause; penetration of solar wind plasma onto closed field lines; wave-driven anomalous diffusion and buffeting by solar wind dynamic pressure changes. We refer to the resulting flows as 'viscous-like' (even though for some of the proposed mechanisms the interaction is not strictly analogous to viscosity). Note in Fig. 5*a* the effect of this interaction penetrates a distance Δ into the magnetosphere, this being the width of the region of antisunward-moving closed field lines on the flanks of the geomagnetic tail.

An important difference between these two types of excitation mechanism is that the motion of a flux tube in the reconnection-driven cells is dependent on whether it is open or closed. In the ionosphere, open field lines are constrained to move generally antisunward until they are closed again in the geomagnetic tail at N. (This is true for southward IMF but some sunward motion in circulation of very high latitude open field lines can be generated by reconnection away from the subsolar magnetopause if the IMF is northward, as shown in the lower panel of Fig. 4.) Similarly, closed field lines generally move sunward. On the other hand, a flux tube in a viscously driven cell is always closed and its motion depends only upon its position, i.e. its proximity to the magnetopause.

The relative importance of these two types of mechanism has been investigated using the transpolar voltage inferred from satellite passes through the convection pattern. This is derived by integrating the electric field observed along the satellite track between the flow reversals shown by dot–dash lines in Fig. 5*c*. We shall refer to this boundary as the 'convection reversal boundary' (which should be distinguished from the dashed line in Fig. 5*c* which is the boundary of open and closed field lines and which we will call the 'polar cap boundary'). It is found that these transpolar voltage estimates (which, incidentally, assume that the pattern of flow did not alter during the satellite transit time of order 10 min [25]) depend upon the IMF. In particular, when the IMF is northward and efficient reconnection is not expected at *D*, The transpolar voltage is something less than 30 kV, whereas when it is strongly southward, values of 100–120 kV are observed [21, 26]. It is thought that the viscous-like interaction is independent of the IMF orientation, in which case reconnection contributes in excess of 70–90 kV to the transpolar voltage during strongly southward IMF, whereas viscous-like interactions contribute less than 30 kV at all times. Note that the IMF is southward for exactly half the time [12].

The ionospheric circulation described in Fig. 5*c* is a steady-state one. For the reconnection-driven flows, this means that the rate at which field lines are opened at the dayside magnetopause (at *D*) is the same as the rate at which they are closed again in the tail (at *N*) and sunward and antisunward transfer rates of magnetic flux are equal. For viscous-like flows it means there is no accumulation of closed field line in the tail. As we discussed earlier, viscous-like flows can be envisaged as a steady circulation (field lines moving away from the tail magnetopause and drifting sunward at a rate which balances the rate at which closed field lines are moved into the tail by the viscous interaction) or a non-steady one (closed field lines accumulate in the tail and are pinched off in bursts in substorms).

Because the magnetic field in the ionosphere is virtually incompressible, any inbalance between the rate at which field lines are opened and the rate at which they are closed will be reflected in an expansion or contraction of the polar cap, in accordance with Faraday's induction law [1, 11]. Such expansions and contractions have been observed using precipitation boundaries and shown to be consistent with the inferred reconnection rates at the dayside magnetopause and in the geomagnetic tail [29]. The changes in polar cap area take place over several hours, because it is such a large reservoir of open flux compared with the rates of flux transfer into and out of it (i.e. the reconnection rates).

In the steady-state case, it is possible to envisage the dawn-to-dusk convection electric field, associated with the solar wind flow, as mapping down all open field lines and causing antisunward flow in the ionospheric polar cap. However, this is not consistent with the short (minute) time scales on which dayside convection has been observed to respond to changes in the IMF [1, 11]: were all open field lines equally effective in exciting ionospheric flows, then convection could only vary on the longer (hour) time constants of expansion and contraction of the polar cap. The results are, however, well explained in terms of the non-steady-state expanding/contracting polar cap model of ionospheric convection proposed by Siscoe and Huang [30]. This model is illustrated by Fig. 6*a* which shows schematically the flows which would result from opening field lines at twice the rate at which they are closed on the nightside (i.e. the polar cap is expanding, as denoted by the heavy black arrows). In Fig. 6*b*, reconnection proceeds twice as quickly at *N* as at *D*, and a different (but still two-celled) pattern is pedicted as the polar cap contracts. The merging gaps are shown as dashed lines and, as in Fig. 5*c*, are labelled *D* and *N*. The solid portions of the polar cap boundary are called 'adiaroic' (meaning 'not flowing across') because no plasma nor magnetic flux is transferred across them. Hence a closed field line near such a boundary must remain closed and an adjacent open field

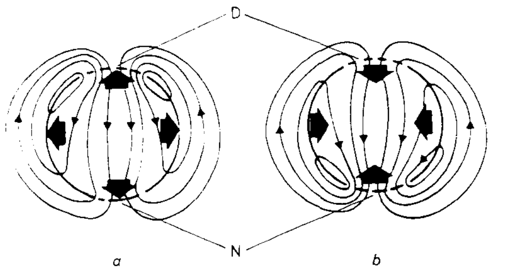

Fig. 6 *Non steady state ionospheric flow snapshots for reconnection driven flows*
a Expanding polar cap
b Contracting polar cap
motion of the polar cap boundary (shown here as circular for simplicity). Solid segments of that boundary are 'adiaroic'
---- merging gaps, labelled *D* and *N* as in Fig. 5*c*. After Lockwood [25]

line must remain open and both will move equatorward during the polar cap expansion, with the boundary. Likewise they will move poleward with the boundary during a polar cap contraction when reconnection in the tail proceeds faster than at the dayside magnetopause. Note that the viscously driven cell will be moved equatorward/poleward during these expansions/contractions of the polar cap.

Lastly, we should note that the schematic flow patterns presented here address the principles and not the details of ionospheric plasma convection. For example, strong flow shears are known to be associated with auroral arcs, and the precipitation caused by auroral particles modifies the patterns of flow by increasing the ionospheric conductivities. There are other significant features: for example, channels of extremely rapid flow at relatively low latitudes, called subauroral ion drifts. The low-latitude boundary of the flow pattern is also variable: middle latitudes are largely shielded from the effects of high-latitude electric fields by energetic particles in the magnetosphere called the ring current; however, fluctuations in this effect allow the electric field to penetrate to lower latitudes to a varying extent.

4 Implications for motion of flux tubes in time-dependent modelling of high-latitude ionosphere

Steady-state models of the convection pattern have been used as input to the numerical models of the high-latitude ionosphere and have had considerable success in predicting some regular and average features of the high-latitude ionosphere. For example, Farmer *et al.* [31] have shown that the seasonal changes of the average diurnal variation in plasma density observed by EISCAT are well explained by the UCL/Sheffield University model. Individual days on which the convection pattern is relatively steady have also been modelled successfully [32].

However, on most days the convection is highly variable. Several attempts to model the time-dependent behaviour of the ionosphere–thermosphere system have treated the temporal variation as a series of quasi-steady states. In such studies, steady-state models of convection have been employed, but the transpolar voltage has been increased (while the polar cap is expanded and the flux of auroral particles enhanced), in accordance with observed statistical regressions [4, 5]. However, the use of these regressions can lead to serious errors, because of the departures from steady state which were discussed in Section 3. For example, consider an increase in transpolar voltage. This can be produced in three ways: by an increase in the dayside reconnection rate; by an increase in the nightside reconnection rate; or by an enhancement of the viscous momentum transfer [25]. If it is due to an increase in the reconnection rate at the subsolar magnetopause (D), the polar cap will expand. However, if it is due to a similar increase in the tail (at N), the polar cap, will contract. Only when averaged over the cycles of polar cap expansion and contraction does the transpolar voltage increase with polar cap area. Hence, the statistical relationship between the polar cap size and transpolar voltage only applied to averages over long periods (greater than several hours) and should not be used to model time-dependent behaviour [1]. Similarly, although nightside auroral precipitation will be enhanced in association with an increase in the reconnection rate in the tail, this will not initially be true for an increase in reconnection rate at the subsolar magnetopause. In addition, the relationship between the auroral boundaries and the convection reversal boundary at any instant of time is not known.

As mentioned above, the convection boundary has been expanded equatorward in some time-dependent simulations, using an average model of the convection pattern [19, 20]. In effect, this is using a sequence of steady-state model convection patterns (like those in Fig. 4), which may be considerably different from the evolution of the true nonsteady flow pattern (like those in Fig. 6). In this Section, we consider two implications of the use of steady-state models in the light of the previous discussion of how convection is excited.

4.1 Tracking flux tubes in the ionosphere

On the left-hand side of Fig. 7a, we consider the loci of flux tubes near an adiaroic segment of the polar cap boundary in the dawn sector, during a polar cap expansion. In this case the expansion has been caused by an increase in the reconnection rate at the dayside magnetopause (due to a southward swing of the IMF). The flux tube marked o is open and that marked c is closed. The polar cap boundary locations at the start and end of the expansion are given as dashed lines. If we consider only reconnection driven flows, the convection reversal lies on this open/closed field line boundary. Between the pairs of dots on the paths of o and c, the polar cap expands equatorward (in the y direction) at a constant speed V_b from the left-hand to the right-hand dashed line. At the onset of this expansion, the speed of motion of both o and c along the polar cap boundary (parallel to the x direction) is shown as increasing, which would be the expected situation if the polar cap expansion were due to a southward turning of the IMF and the consequent increase in the dayside reconnection rate. This behaviour is shown by the plot of the plasma flow speed along the boundary, V_x, on the right-hand side of the Figure. Note that for illustrative purposes the antisunward motion of o is shown as being faster than the sunward motion of c. However, both o and c must move equatorward with the boundary (at speed $V_y = V_b$) during the expansion. Conversely, we would expect o and c to move poleward, but continue to convect in the same direction, during a polar cap contraction. If this were caused by a substorm the flow speeds along the boundary, V_x, would increase, but if it were due to a decrease in reconnection rate at the subsolar magnetopause, the flow speeds would decrease. The important general point is that, in both cases, o continues to move antisunward (but at a different latitude)

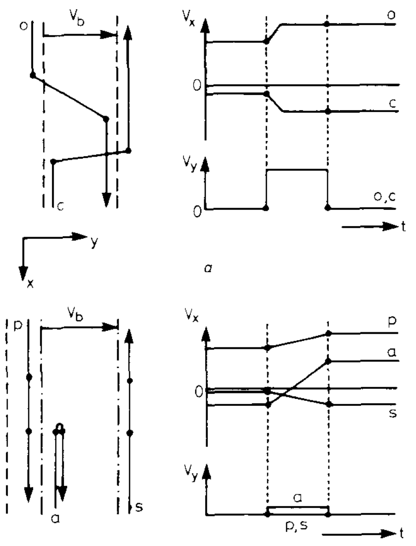

Fig. 7 *Motion of flux tubes near the polar cap boundary in the dawn sector for an expansion of the polar cap (at speed V_b)*

a Enhanced dayside reconnection
b Enhanced viscous-like interaction
Left-hand panels show the boundary motion in the xy frame which is stationary with respect to the Earth and where x is aligned with the boundary and y is normal to it
The open closed field line boundry (dashed line) which in Fig. 7b differs from the convection reversal boundary (dot dash line) because of viscous-like interaction. Motions of open, o, and closed, c, field lines are shown in a and of (initially) polar cap, p, auroral, a, and subauroral, s, field lines in Fig. 7b
Right-hand panels show the evolution of the velocity components along and normal to the boundary (V_x and V_y, respectively) for each of these field lines

and c continues to move sunward (again at a different latitude). Because the polar cap expansion/contraction moves closed field lines equatorward/poleward, any viscously driven cells will also move in latitude and hence so will the convection reversal boundary. Any plasma density features on the flow streamlines close to the boundary will also migrate in latitude.

This behaviour should be compared with a second type of expansion of the convection boundary, this time caused by an increase in the efficiency of the viscous-like interaction, as depicted in Fig. 7b. This could cause an increase of the thickness Δ of the region of antisunward-moving flux tubes in the equatorial plane (see Fig. 5a) and of the separation of polar cap and convection boundaries in the ionosphere, shown here as dashed and dot–dash lines, respectively (as in Fig. 5). Fig. 7b shows that the behaviour of flux tubes near the convection boundary is very different to that in Fig. 7a. Initially, p is a polar cap flux tube (closed but convecting antisunward) whereas a and s are auroral and subauroral, sunward-convecting (and closed) tubes, respectively. The increase in momentum transfer at the magnetopause will cause p to move faster, as shown in the variation of V_x on the left hand side, but cause a to cease moving sunward and turn antisunward. No flux tubes show any appreciable V_y.

Lastly, we must consider what will happen if we take a model convection pattern and attempt to simulate a polar cap expansion by increasing the radius of the convection boundary reversal. Because no attention is paid to which flux tubes are open and which are closed by such a procedure, the effect will be similar to that in Fig. 7*b*. In other words, what were sunward convecting auroral field lines will become antisunward convecting field lines as the convection reversal boundary is expanded over them. The loci of flux tubes near the boundary is as shown in Fig. 7*b*. Hence we must conclude that such a procedure is correct for expansions of the convection pattern caused by an enhancement of viscous-like interaction but will incorrectly predict how flux tubes move in cases when the polar cap expansion is due to reconnection rate changes. Because the later cases include southward and northward turnings of the IMF and substorms, the expansion of empirical convection models will frequently incorrectly predict flux tube paths. Furthermore, the dominant driving mechanism for convection is reconnection and most of the variability in transpolar voltage appears to arise from variations in reconnection rates. Hence we would expect the situation shown in Fig. 7*a* to apply more generally than that in Fig. 7*b*.

Given that the numerical models predict the spatial distribution of ionospheric densities by following the convecting plasma, the two types of polar cap motion described above could lead to very different predictions of the ionospheric density distribution and its evolution. The precise implications of this are not yet clear. However, consider a longitudinally extended plasma depletion lying equatorward of the convection boundary reversal in Fig. 7 (i.e. the midlatitude trough). In Fig. 7*a*, this trough would move as a whole to lower latitudes during the polar cap expansion and such behaviour is indeed observed. However, in Fig. 7*b* the trough would not migrate equatorward and would be changed in character by the changes in the flow directions (one possible effect is that the trough depletion would be progressively smoothed out). This example alone is enough to suggest that the distinction between the two types of change described in Fig. 7 is of vital importance.

In addition, Fig. 7*a* illustrates that flow paths of flux tubes can cross in cases where the convection pattern varies with time, something which cannot happen for steady-state convection. This generates a problem for time-dependent simulation models. The time interval used between steps in the time-dependent simulation, and the spatial grid size, must be small enough to ensure that the flux tubes are correctly tracked at all times.

4.2 Ion–neutral frictional heating effects

The time dependence of convection also influences the prediction of thermospheric perturbations due to ion–neutral frictional heating. As the ion gas drifts through the neutrals, collisions between the two convert the drift energy into thermal energy and heat both the ion and neutral gases. The rate of frictional heating of the ion gas is

$$Q_i = n_i m_R \nu_{in} |\boldsymbol{V} - \boldsymbol{U}|^2$$

where n_i is the ion density, m_R is the reduced mass and ν_{in} is the ion–neutral collision frequency: $\boldsymbol{V}$ and $\boldsymbol{U}$ are the ion and neutral velocity vectors, respectively. The neutral heating rate is similarly proportional to $|\boldsymbol{V} - \boldsymbol{U}|^2$; here we shall refer to the relevant constant of proportionality (which depends on the density) as α.

The thermospheric wind is the resultant of many forces on a packet of neutral air, but at high latitudes it is strongly influenced by the ion drift. However, because of the large number of neutral atoms, compared to the number of ions, the response time of the wind, U, to changes in V is typically 10–100 min; hence U will not be altered by fluctuations in convection velocities on time scales shorter than this. As a result, the heating will be greater than would be calculated from the mean ion flow. To illustate this point, consider a wind, U, which is at an angle β to a steady ion flow, V_m. The heating rate is then equal to $\alpha(V_m^2 + U^2 - 2V_m U \cos\beta)$. Now consider ion flows in the same direction but of speed V with a square-wave oscillation of amplitude $2\,\Delta V$, such that the mean is still V_m. The wind is unaltered because the fluctuations in V are too rapid. The time-averaged heating rate is then found to be larger by $\alpha\,\Delta V^2$. If the changes in V are slow enough that the winds can partially respond, or if the convection direction varies, the result is more complicated.

This 'AC' heating component will be absent in predictions based on convection models which are averages over the considerable short-term variability. The models of convection that are presently used as inputs to the numerical ionosphere–thermosphere simulations are derived from averaged satellite data (which pass through the convection pattern of a given hemisphere about once every hour) or averaged radar data (which have typically employed half-hour elevation of azimuth scans). Thus, in general they average over the fluctuations in convection on time scales below about 1 hour, which are apparent in higher resolution ground-based data (radar and magnetometer). Consequently, we expect these models to underestimate the heating of the ionospheric and thermospheric gases.

5 Discussion and conclusions

We have considered the implications of variations in the pattern of high-latitude plasma convection on short time scales (less than several hours). Comparisons of model predictions with data from the high-latitude EISCAT radar have shown that the models are very effective under steady or average convection conditions [31], but that problems arise when the convection pattern alters [6, 32]. We have shown that, for expansions and contractions of the polar cap caused by reconnection, some form of the expanding/contacting convection model must be used; flux tubes will follow incorrect paths if average convection models are used with a time-varying convection reversal boundary. The paths followed in such simulations may be quite close to those expected for an increase in size of the region of antisunward flow (the convection polar cap) due to enhanced viscous interaction at the magnetopause; however, there is good evidence that the dominant oscillations in the convection polar cap are due to differences in reconnection rate at the dayside magnetopause and in the geomagnetic tail [1, 11, 29]. Hence, in most cases, the use of average flow models will produce erroneous flux tube paths and inaccurately predict plasma density distributions and HF radiowave propagation conditions.

This paper has not considered the magnitude of these effects. Indeed, such an evaluation would require full simulations using numerical models which can successfully track flux tubes whose paths cross. However, we can note that the whole concept of numerical high-latitude ionospheric models is based on the idea that flux tube density at a given point depends upon the path taken by the tube to reach that point and we consider certain situations and the likely implications. In the example given, we considered the midlatitude trough at an MLT where the polar cap boundary is adiaroic (e.g. near dawn or dusk). If the polar cap expands due to an increase in dayside reconnection, the flux tubes on which the poleward wall of the trough has been formed by auroral precipitation will simply be moved equatorward and one might expect the midlatitude trough to become thinner in latitudinal width and the minimum density to move to lower latitudes. On the other hand, an expansion of the convection polar cap due to enhanced viscous-like interaction would not cause any equatorward motion of the poleward trough wall (although it would change the zonal motion of that plasma). Observational evidence suggests that the poleward wall does indeed move equatorward on the time scales of polar cap expansion (of order an hour). However, numerical simulations are required to see if the complex coupled system really behaves in the way qualitatively suggested above. Effects away from the convection boundary will not be as obvious as those illustrated in Fig. 7. Note that errors in the predicted paths of plasma will be compounded for as long as the expansion/contraction continues.

The other important effect is the additional ion and neutral heating which rapid convection changes will cause. Again, the implications of this effect in terms of the magnitude of possible changes in plasma densities will not be known until full numerical simulations can be done. However, the heating does induce thermospheric composition changes which do modulate the rate of loss of the plasma by chemical reactions. As discussed earlier, these effects are not restricted to high latitudes because thermospheric changes and winds propagate to lower latitudes.

In conclusion, the variability of ionospheric convection poses some difficult problems for modelling of the coupled ionosphere–thermosphere system, and has some interesting implications globally. In general, an averaged steady-state convection pattern will not yield the correct corresponding average results for ionospheric and thermospheric densities and composition, nor for themospheric winds. However, the benefits of modelling the effects of time-dependent convection are very great, once these difficulties can be overcome. As an example, recent work by Lockwood and Carlson [33] has shown how application of the concepts discussed here can explain the production of polar cap density patches. This approach provides the potential of not only understanding and predicting the behaviour of the high-latitude ionosphere, but also of forecasting the day-to-day variability and storm phenomena in the midlatitude ionosphere. At present, the coupled ionosphere–thermosphere models required considerable allocations of time on large supercomputers. However, looking to the not-too-distanct future, the rapid increase in power of smaller, parallel-processing computers will soon allow the implementation of such models on relatively inexpensive, dedicated computers. This opens up potential applications for routine operations of HF systems (e.g. frequency management), in addition to the simulations of extreme conditions which would be of use in system design.

6 References

1 LOCKWOOD, M., and COWLEY, S.W.H.: 'Ionospheric convection and the substorm cycle', *in* MATTOCK, C. (Ed.): 'Substorms 1, proceedings of the first international conference on *substorms*, ICS-1', ESA SP (European Space Agency Publications, Nordvijk, The Netherlands, 1992), pp. 99–109

2 KNUDSEN, W.C.: 'Magnetospheric convection and the high-latitude F2 region ionosphere', *J. Geophys. Res.*, 1974, **79**, pp. 1046–1055

3 KNUDSEN, W.C., BANKS, P.M., WINNINGHAM, J.D., and KLUMPAR, D.M.: 'Numerical model of the convecting F2 ionosphere at high latitudes', *J. Geophys. Res.*, 1977, **82**, pp. 4784–4792

4 SOJKA, J.J.: 'Global scale, physical models of the F-region ionosphere', *Rev. Geophys.*, 1989, **27**, pp. 371–403

5 REES, D., and FULLER-ROWELL, T.J.: 'Geomagnetic response of the polar thermosphere and ionosphere', *in* SANDHOLT, P.E., and EGELAND, A. (Eds.): 'Electromagnetic coupling in the polar clefts and caps' (Kluwer Academic Publishers, 1989), pp. 355–391

6 QUEGAN, S.: 'The influence of convection on the structure of the high-latitude ionosphere', *Phil. Trans. R. Soc. London, A*, 1989, **328**, pp. 119–137

7 QUEGAN, S., BAILEY, G.J., MOFFETT, R.J., HEELIS, R.A., FULLER-ROWELL, T.J., REES, D., and SPIRO, R.W.: 'A theoretical study of the distribution of ionization in the high-latitude ionosphere and the plasmasphere: first results on the mid-latitute trough and the light ion trough', *J. Atmos. Terr. Phys.*, 1982, **44**, pp. 619–640

8 TSUNODA, R.T.: 'High-latitude F-region irregularities: a review and synthesis', *Rev. Geophys.*, 1988, **26**, pp. 719–760

9 LOCKWOOD, M.: 'Incoherent scatter radar measurements of the cusp', *in* BARRON, C.I. (Ed.): 'CLUSTER — dayside polar cusp', ESA SP-330 (European Space Agency Publications, Nordvijk, The Netherlands, 1991), pp. 57–66

10 COWLEY, S.W.H., MORELLI, J.P., and LOCKWOOD, M.: 'Displacements of the polar cusp and polar cap related to the X and Y components of the IMF', *J. Geophys. Res.*, 1991, **96**, pp. 5557–5564

11 LOCKWOOD, M., COWLEY, S.W.H., and FREEMAN, M.P.: 'The excitation of plasma convection in the high latitude ionosphere', *J. Geophys. Res.*, 1990, **95**, pp. 7961–7971

12 HAPGOOD, M.A., BOWE, G.A., LOCKWOOD, M., TULUNAY, Y., and WILLIS, D.M.: 'Variability of the interplanetary magnetic field over 24 years: 1963–1986', *Planet. Space Sci.*, 1991, **39**, pp. 411–423

13 COLLIS, P.N., and HÄGGSTRÖM, I.: 'Plasma convection and auroral precipitation processes associated with the main ionospheric trough at high latitudes', *J. Atmos. Terr. Phys.*, 1988, **50**, pp. 389–404

14 BLAGOVESHCHENSKIJ, D.V., EGROVA, L.V., and LUKASHKIN, V.M.: 'High latitude ionospheric phenomena diagnostics by high frequency radiowave propagation observations', *Radio Sci.*, 1992, **27**, pp. 267–274

15 DAVIES, K., and RUSH, C.M.: 'High-frequency ray paths in ionospheric layers with horizontal gradients', *Radio Sci.*, 1985, **20**, pp. 95–103

16 TEDD, B.L., STRANGEWAYS, H.J., and JONES, T.B.: 'Systematic ionospheric electron density tilts at mid-latitudes and their associated HF bearing errors', *J. Atmos. Terr. Phys.*, 1985, **47**, pp. 1085–1092

17 CANNON, P.C.: 'Morphology of the high-latitude ionosphere and its implications for HF communications systems', *IEE Proc. I*, 1989, **136**, pp. 1–10

18 LOCKWOOD, M.: 'A simple model of the effects of the mid-latitude total ion trough in the bottomside F layer on HF radiowave propagation', *Radio Sci.*, 1981, **16**, pp. 385–391

19 HEELIS, R.A.: 'The effects of interplanetary magnetic field orientation on dayside high-latitude convection', *J. Geophys. Res.*, 1984, **89**, pp. 2873–2880

20 HEPPNER, J.P., and MAYNARD, N.C.: 'Empirical high-latitude electric field models', *J. Geophys. Res.*, 1987, **92**, pp. 4467–4477

21 REIFF, P.H., and LUHMANN, J.G.: 'Solar wind control of the polar cap voltage', *in* KAMIDE, Y., and SLAVIN, J.A. (Eds.): 'Solar wind-magnetosphere coupling' (Terra Scientifica, Tokyo 1986), pp. 453–476

22 HOLT, J.M., WAND, R.H., EVANS, J.V., and OLIVER, W.L.: 'Empirical models for the plasma convection at high latitudes from Millstone Hill observations', *J. Geophys. Res.*, 1987, **92**, pp. 203–212

23 FRIIS-CHRISTENSEN, E., KAMIDE, Y., RICHMOND, A.D., and MATSUSHITA, S.: 'Interplanetary magnetic field control of high-latitude ionospheric fields and currents determined from Greenland magnetometer data', *J. Geophys. Res.*, 1985, **90**, pp. 1325–1338

24 RICHMOND, A.D., KAMIDE, Y., AHN, B.-H., AKASOFU, S.-I., ALCAYDE, D., BLANC, M., DE LA BEAUJARDIERE, O., EVANS, D.S., FOSTER, J.C., FRIIS-CHRISTENSEN, E., FULLER-ROWELL, T.J., HOLT, J.M., KNIPP, D., KROEHL, H.W., LEPPING, R.P., PELLINEN, R.J., SENIOR, C., and ZAITZEV, A.N.: 'Mapping electrodynamic features of the high-latitude ionosphere from localised observations: combined incoherent-scatter radar and magnetometer measurements from 18 19 January, 1984', *J. Geophys. Res.*, 1988, **93**, pp. 5760–5776

25 LOCKWOOD, M.: 'On flow reversal boundaries and cross-cap potential in average models of high latitude convection', *Planet. Space Sci.*, 1991, **39**, pp. 397–409

26 COWLEY, S.W.H.: 'Solar wind control of magnetospheric convection', *in* 'Achievements of the international magnetospheric study, IMS'. ESA SP-217, ESTEC, Noordwijk, The Netherlands, 1984, pp. 483–494

27 DUNGEY, J.W.: 'Interplanetary magnetic field and the auroral zones', *Phys. Rev. Lett.*, 1961, **6**, pp. 47–48

28 LOCKWOOD, M., COWLEY, S.W.H., TODD, H., WILLIS, D.M., and CLAUER, C.R.: 'Ion flows and heating at a contracting polar cap boundary', *Planet. Space Sci.*, 1988, **36**, pp. 1229–1253

29 HOLZER, R.E., McPHERRON, R.L., and HARDY, D.A.: 'A quantitative model of the magnetospheric flux transfer process', *J. Geophys. Res.*, 1986, **91**, pp. 3287–3293

30 SISCOE, G.L., and HUANG, T.S.: 'Polar cap inflation and deflation', *J. Geophys. Res.*, 1985, **90**, pp. 543–547

31 FARMER, A.D., FULLER-ROWELL, T.J., and QUEGAN, S.: 'Comparing numerical simulations of the high-latitude ionosphere to an empirical mean model based on EISCAT', *Adv. Space Res.*, 1990, **10**, (6), pp. 43–48

32 QUEGAN, S., GILL, R.S., and LOCKWOOD, M.: 'Comparisons between EISCAT observations and model calculations of the high latitude ionosphere', *J. Atmos. Terr. Phys.*, 1988, **50**, pp. 1057–1076

33 LOCKWOOD, M., and CARLSON, H.C. Jr.: 'Production of polar cap electron density patches by transient magnetopause reconnection', *Geophys. Res. Lett.*, 1992, **19**, pp. 1731–1734