Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **422,** 1460–1488 (2012) doi:10.1111/j.1365-2966.2012.20719.x

# Axial symmetry breaking of Saturn's thermosphere

C. G. A. Smith[1]★† and N. Achilleos[2]

[1]*The Brooksbank School, Victoria Road, Elland, West Yorkshire HX5 0QG*
[2]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT*

Accepted 2012 February 8. Received 2012 February 4; in original form 2011 December 18

**ABSTRACT**

The source of the various planetary-period signals in Saturn's magnetosphere is at present unknown. We investigate the possibility that the source of these signals is an axially asymmetric wind system in the thermosphere. We describe a feedback mechanism that has the potential to drive such axially asymmetric wind systems. The proposed mechanism relates thermospheric winds to heating from particle precipitation, via the generation of horizontal and field-aligned currents. The relevant physical processes are investigated using a highly simplified general circulation model of Saturn's thermosphere and ionosphere. Our principal result is that the feedback mechanism is effective in permanently breaking the axial symmetry of the thermosphere, generating a drifting vortex-like structure at high latitudes. However, the precipitating electron energies required to power this structure are of the order of 5 MeV, 2–3 orders of magnitude greater than the observed auroral electron energies, and the highly axially asymmetric distribution of precipitation required across the polar regions of the planet is also inconsistent with observations. Despite these flaws, the model qualitatively explains several features of the observed variation in the pulsing of SKR emissions; in particular, the seasonal variation and the faster rotation rate in the winter hemisphere. We cannot reproduce the apparent 7 month lag in the response of the Saturn Kilometric Radiation (SKR) rotation rate to seasonal variation, but instead suggest the possibility that this effect may have its origin in long chemical time-scales in the upper atmosphere. We also predict the possible existence of secondary periodic features in the SKR emissions with periods of ∼15 planetary rotations, driven by complex wave behaviour in the thermosphere.

**Key words:** magnetic fields – methods: numerical – planets and satellites: aurorae – planets and satellites: atmospheres – planets and satellites: individual: Saturn – planets and satellites: magnetic fields.

## 1 INTRODUCTION

The *Voyager* missions determined the internal rotation period of Saturn by measuring the pulsing of kilometric radio emissions (SKR) (Desch & Kaiser 1981). The assumption that the period of this pulsing was directly related to the internal rotation period was called into question when it was found to vary by ∼1 per cent on a time-scale of years (Galopeau & Lecacheux 2000). The variation of the radio period has subsequently been very well characterized (Gurnett et al. 2010). The periods of radio emissions from the Northern and Southern hemispheres are now known to be different and to vary seasonally, with the period in the summer hemisphere found to be slightly longer than that in the winter hemisphere. Surprisingly, the periods appear to equalize not at the exact time of the equinox but ∼7 months later. Both the north–south asymmetries, the seasonal variation and the ∼7 month time lag remain unexplained.

A similar periodicity was first detected in magnetic field data by Espinosa, Southwood & Dougherty (2003b) and has subsequently been very well characterized and shown to be correlated with the SKR period (Andrews et al. 2008). North–south asymmetries in the magnetic field periodicities have also been shown to match the north–south asymmetry in the SKR period (Andrews et al. 2010; Provan et al. 2011; Southwood 2011). Other phenomena in the magnetosphere exhibit related periodicities, for example particle distributions (e.g. Burch et al. 2009) and the location of the main UV auroral oval (Nichols et al. 2008).

There are two main classes of models for explaining the magnetospheric periodicities and their variability: internal models and external models. Internal models propose that Saturn's magnetic field contains differentially rotating magnetic anomalies (Dessler 1985) that are ultimately responsible for periodically perturbing the magnetosphere, perhaps via the 'camshaft' model described by Espinosa, Southwood & Dougherty (2003a) and Southwood &

★E-mail: cgasmith@gmail.com
†Visiting Research Fellow, Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT

Kivelson (2007). The main problems with this model lie in the nature of the seasonal variation in the periodicities: (i) the angular momentum of the deep layers of the planet where a magnetic anomaly might be located is apparently too great to permit changes in rotation period on the required time-scales; (ii) there is no clear mechanism for producing seasonal variation, since absorption of solar radiation is expected to have a negligible effect on the dynamics of the deep layers of the planet.

External models propose that the periodicities are a consequence of rotating structures in the magnetosphere itself. For example, Goldreich & Farmer (2007) described a mechanism in which a large-scale axially asymmetric plasma convection system is set up in the magnetosphere, providing the source of the various periodic phenomena, while Gurnett et al. (2007) showed a correlation between the rate of rotation of magnetospheric plasma and the observed SKR rotation rate. External models do not suffer from either of the problems described above for internal models: (i) the angular momentum of the magnetospheric plasma is small enough to vary plausibly on the required time-scales and (ii) seasonal variation can be introduced via coupling to the ionosphere, whose conductance is expected to vary with season. However, there is a significant further problem: (iii) it is difficult to explain the existence of differing rotation rates in the Northern and Southern hemispheres, since structures in the magnetosphere are expected to be 'frozen in' to flux tubes that extend through both hemispheres [although, as pointed out by Gurnett et al. (2009), parallel electric fields could decouple the rotation rates of the ionospheric plasma in each hemisphere from their magnetospheric counterpart and from each other].

A compromise that is potentially able to overcome all three of these problems is an atmospheric model. An atmospheric model proposes that the periodicities originate in wind systems, in the outermost layers of the planetary atmosphere. These models are distinct from internal models (that postulate deep magnetic anomalies) in that the periodicity is communicated to the magnetosphere, not by a rotating anomaly in the magnetic field itself, but by an anomaly in the magnetosphere–atmosphere coupling currents that flow through the thermosphere and mesosphere.

An atmospheric source potentially addresses all of the problems outlined above: (i) the angular momentum of the atmosphere, in particular the upper atmosphere, is sufficiently small to vary on seasonal time-scales [e.g. Smith & Aylward (2008) found that the dynamical equilibration time-scale for the thermosphere was much less than 400 planetary rotations]; (ii) seasonal variation of the ionospheric conductance offers a simple explanation for seasonal variability [ionization by solar extreme-UV makes a significant contribution to the electron density, and therefore the conductance, at all latitudes (Moore et al. 2010), so seasonal variation of insolation implies seasonal variation of conductance]; (iii) independent rotation rates in the Northern and Southern hemispheres, dependent upon season, are probably inevitable, since the northern and southern polar regions that connect to the magnetosphere are largely hydrodynamically isolated due to the strong Coriolis forces on Saturn, which prevent strong interhemispheric winds. In the literature, there has been some discussion of the possible influence of seasonally variable atmospheric winds and ionospheric conductance (e.g. Southwood & Kivelson 2007, 2009; Gurnett et al. 2009, 2010). However, the key component of an atmospheric explanation is the existence of axial asymmetries (Smith 2006, 2011; Southwood & Kivelson 2007; Southwood 2011).

The possible importance of such axial asymmetries was first discussed by Smith (2006), who proposed not just variability of zonal winds, but the existence of large-scale axially asymmetric wind systems, either in the form of global-scale 'planetary waves' or more localized vortices. Smith (2011) subsequently showed that, if such an asymmetry did exist in the upper atmosphere, it could produce approximately the required current systems in the magnetosphere to explain some of the magnetic field observations [in particular, the quasi-uniform equatorial perturbation field implied by the camshaft model of Southwood & Kivelson (2007)]. This latter study, while demonstrating the plausibility of the concept, had two important caveats: (i) the calculated magnetic field perturbations in the equatorial magnetosphere were $\sim$100 times smaller than those that are observed; (ii) the asymmetry was artificially generated by an imposed distribution of thermosphere heating, and there was therefore no explanation of the true origin of the asymmetry. This study aims to investigate (ii) further by modelling a possible axial symmetry breaking mechanism that could naturally generate a persistent asymmetry in the thermosphere. There are two additional questions that we are able to investigate. (iii) Can the seasonal variation of insolation produce differential rotation? If so, (iv) how can a $\sim$7 month delay occur between the seasonal variation of insolation and the seasonal variation of the rotation rate?

In Section 2 we first outline a possible feedback process that could generate a persistent asymmetry, and in Section 3 we describe a numerical model designed to investigate this feedback process. In Section 4 we then describe the development of axial asymmetries in the model, and in Sections 5 and 6 we investigate the rotation rate of axial asymmetries and the sensitivity of our feedback model. Finally, in Sections 7 and 8, we discuss the implications of our results and conclude.

## 2 A POSSIBLE FEEDBACK MECHANISM

It has already been shown by Smith (2011) that an axial asymmetry in Saturn's thermosphere is capable of driving current systems that reproduce some of the observed periodic phenomena in Saturn's magnetosphere. The main question left unanswered by this previous study was the source of the axial asymmetry itself. What seems to be required is a symmetry-breaking mechanism that allows an axial asymmetry to persist, at least semi-permanently. Such a mechanism has already been discussed for Saturn's magnetosphere (Goldreich & Farmer 2007), although this model does not explain the existence of the separate rotation rates in the Northern and Southern hemispheres.

Saturn's thermosphere is a heavily damped system, in that gradients in temperature or wind speed are removed relatively rapidly by thermal conduction and viscosity. This means that for an axial asymmetry to persist it must be continuously driven by an external source of energy. One way to provide such a source of energy is to invoke a feedback process in which a small axial asymmetry in an external energy source drives an asymmetric structure in the thermosphere, which in turn strengthens the initial asymmetry. A good analogy for this concept is the feedback process that drives a terrestrial hurricane (e.g. Holton 1992). The circulation of the winds in a hurricane extracts thermal energy from the ocean surface; this thermal energy in turn sustains the circulation. To produce such a system in Saturn's thermosphere, we need a mechanism by which the winds can impose structure upon the thermospheric heating. The sources of heating in Saturn's thermosphere are not well understood, but most studies agree that the main energy sources are probably a mixture of breaking gravity waves (e.g. Matcheva & Strobel 1999), Joule heating (e.g. Cowley, Bunce & O'Rourke 2004a) and particle precipitation (e.g. Galand et al. 2011). Here we will focus on particle precipitation.

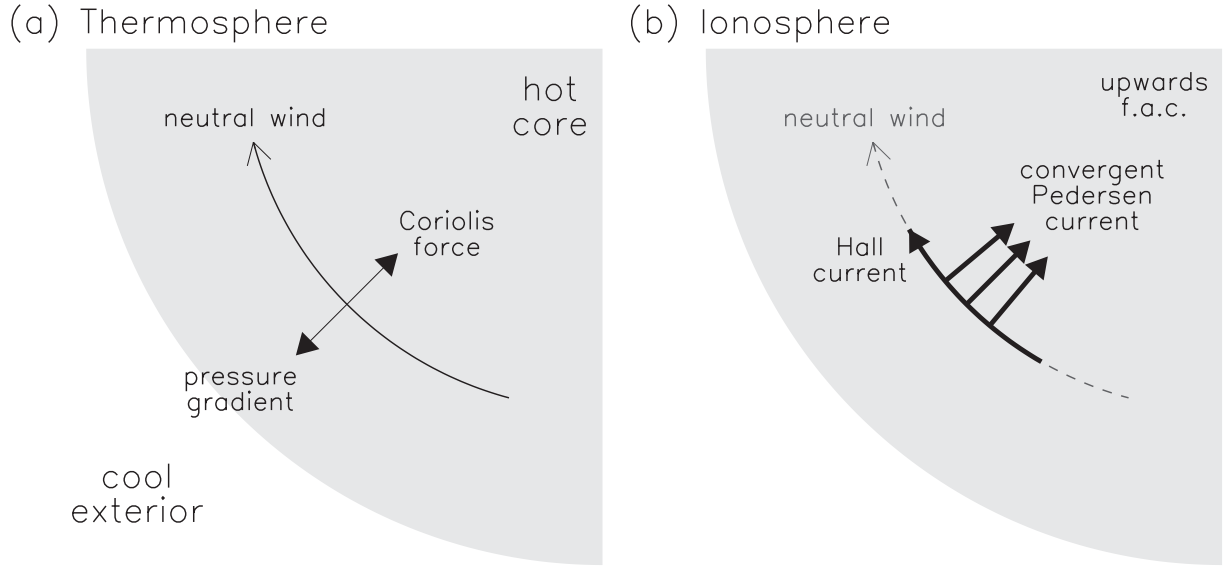


**Figure 1.** Sketch of winds and currents related to a thermospheric vortex with a hot core in the Northern hemisphere of Saturn. (a) A relatively hot region of the thermosphere produces an outward radial pressure gradient. In a simplified situation in which we neglect drag, this leads to clockwise winds as the outward pressure gradient is in geostrophic balance with the Coriolis force. (b) The clockwise winds drive clockwise Hall currents and inward radial Pedersen currents. The convergent Pedersen currents may close via upward field-aligned currents (f.a.c.) at the core, which in turn may lead to particle precipitation that sustains the high temperature of the core. In the Southern hemisphere, the signs of the Coriolis force and the Pedersen current are both reversed, and the same physical principles apply, but in mirror image.

Energy is being continually deposited in Saturn's polar regions by particle precipitation. We know this because of the bright UV auroras (Clarke et al. 2005) and the existence of IR auroras at mid-latitudes (Stallard et al. 2008), both of which are associated directly or indirectly with particle impacts. The particle precipitation itself is associated with field-aligned current systems (Cowley et al. 2004a). These current systems are in turn associated with horizontal currents in the thermosphere–ionosphere that are partially driven by plasma motions in the magnetosphere and partially driven by thermospheric winds (Smith & Aylward 2008). If the horizontal currents are convergent, then an upward field-aligned current may flow in order to conserve charge. This upward field-aligned current, flowing out of the atmosphere, may be associated with downward precipitation of accelerated electrons, travelling into the atmosphere (Cowley et al. 2004a).

We will now describe a possible feedback process based on interaction of particle precipitation and thermospheric winds. To help us imagine how such a process may occur, we consider a simplified situation, consisting of an isolated, circularly symmetric vortex, somewhat analogous to a hurricane. The behaviour of a such a vortex with a high pressure or 'hot' core is sketched in Fig. 1. (We describe here the behaviour of a Northern hemisphere vortex. The sign of the Coriolis force and the Pedersen current are both reversed in the Southern hemisphere, so exactly the same physical principles apply, but in mirror image.) In the absence of Coriolis force, the outward radial pressure gradient would simply drive outward radial winds. However, including the Coriolis effect – and ignoring drag forces – geostrophic balance produces a pure anticlockwise vortex, as sketched in Fig. 1(a). The resulting ionospheric currents, driven by these winds, are sketched in Fig. 1(b). A clockwise Hall current is driven directly by the clockwise neutral wind. This is not divergent so does not drive any field-aligned current. However, an inward radial Pedersen current is also driven by the clockwise neutral wind. This is convergent and may thus be associated with upward field-aligned currents and particle precipitation. The particle precipitation may in turn heat the core and thus sustain the vortex.

It is of course possible that the convergent currents will close not through field-aligned currents flowing into the magnetosphere, but by inducing horizontal electric fields in the thermosphere (radial with respect to the core of the vortex) that oppose the converging currents. However, unless the plasma flows in the ionosphere are completely decoupled from the magnetosphere by parallel electric fields, such an electric field would imply a rotational motion of the connected flux tube, including plasma in the equatorial magnetosphere. Thus, in order for the currents to close locally, plasma in the magnetosphere must first be set in motion, and field-aligned currents must flow while this process takes place. Furthermore, we would not expect a specific location in the (neutral) thermosphere to remain connected to the same part of the magnetosphere, as the two systems rotate about the axis of the planet at slightly different angular velocities. Even if the time-scale to 'spin up' a flux tube in this way was very short, there would still be field-aligned currents present as the thermospheric vortex drifted, relative to the magnetosphere, and became connected to fresh flux tubes that had not yet been 'spun up'.

We have so far neglected drag forces, which may arise due to viscosity and ion drag. These will both tend to produce an outward radial component to the neutral flow, in addition to the circulating clockwise component. This outward radial component is associated with an outward Hall current and a clockwise Pedersen current. Provided the ratio of Pedersen-to-Hall conductance is greater than the ratio of radial-to-clockwise wind speeds, the resultant distribution of currents will be convergent and have the potential to produce upward field-aligned currents.

The precise mechanism that we have described – that of a localized vortex driven directly by particle precipitation – seems unlikely to be responsible for the global-scale asymmetry that seems to be required to explain the periodicities observed in the magnetosphere. We would require a persistent region of relatively intense particle

precipitation across a broad range of longitudes for such a vortex to exist on a global scale. We would expect this phenomenon to be observable in either UV or IR measurements of the polar auroras as a clear longitudinal brightness asymmetry, rotating close to the planetary period. Although planetary-period oscillations have been observed in the location (Nichols, Cowley & Lamy 2010b) and integrated brightness (Nichols et al. 2010a) of the UV emissions, no rotating, planetary scale asymmetry in the brightness of the UV or IR emissions has been reported.

Instead of direct generation of particle precipitation by a thermospheric vortex, we thus propose a more subtle mechanism. The pattern of particle precipitation across the polar regions of Saturn consists of relatively small-scale structures, for example the main auroral oval is typically ∼500 km wide compared to its polar diameter of ∼30 000 km (Cowley, Bunce & Prangé 2004b). These small-scale structures are also not static. The latitude of the auroral oval shifts, and the structure of the oval is not a uniform ring but a chain of moving blobs, arcs, spiral structures and transient regions of bright emission within the oval (Clarke et al. 2005; Grodent et al. 2011; Radioti et al. 2011).

The consequence of this behaviour is that any particular region of the thermosphere will encounter variable particle precipitation, which is presumably related to variable current systems driven by small-scale magnetospheric plasma flows. Auroral emission will occur when the upward currents exceed a critical level, requiring electron acceleration to supply the necessary current. Fig. 2(a) is a schematic of how the upward and downward field-aligned currents might vary across a region of the polar thermosphere. The critical level for electron acceleration – on the arbitrary scale shown – is taken to be $j = 1$, and the curve is filled in black when it exceeds this level. The dashed line shows a larger scale current system that might be driven by the planetary-scale thermospheric asymmetry in which we are interested. In the schematic, the large-scale component does not exceed the threshold for electron acceleration.

Fig. 2(b) shows the situation when these two current systems are superposed. It is clear that across the left-hand side of the schematic – where the large-scale current system produces weak upward currents – the small-scale upward currents are enhanced and exceed the threshold for electron acceleration more often. On the right-hand side, the opposite happens. Thus, the large-scale current system leads to relative increases and decreases in the pattern of particle precipitation by modulating the existing small-scale currents.

Combining the vortex model sketched in Fig. 1 with the mechanism for enhanced particle precipitation sketched in Fig. 2, we can now outline a possible model for generating a persistent global-scale asymmetry. A global-scale vortex system across the polar cap may drive a weakly convergent current at its core that drives weak, upward, field-aligned currents. These currents lead to enhancement of the existing – and highly variable – particle precipitation in this region that provides the energy input to sustain the vortex. In this situation, there would presumably be a corresponding region of divergent currents in the opposite longitude sector in which the particle precipitation would be reduced, increasing the temperature contrast available to drive the vortex wind system. If this mechanism

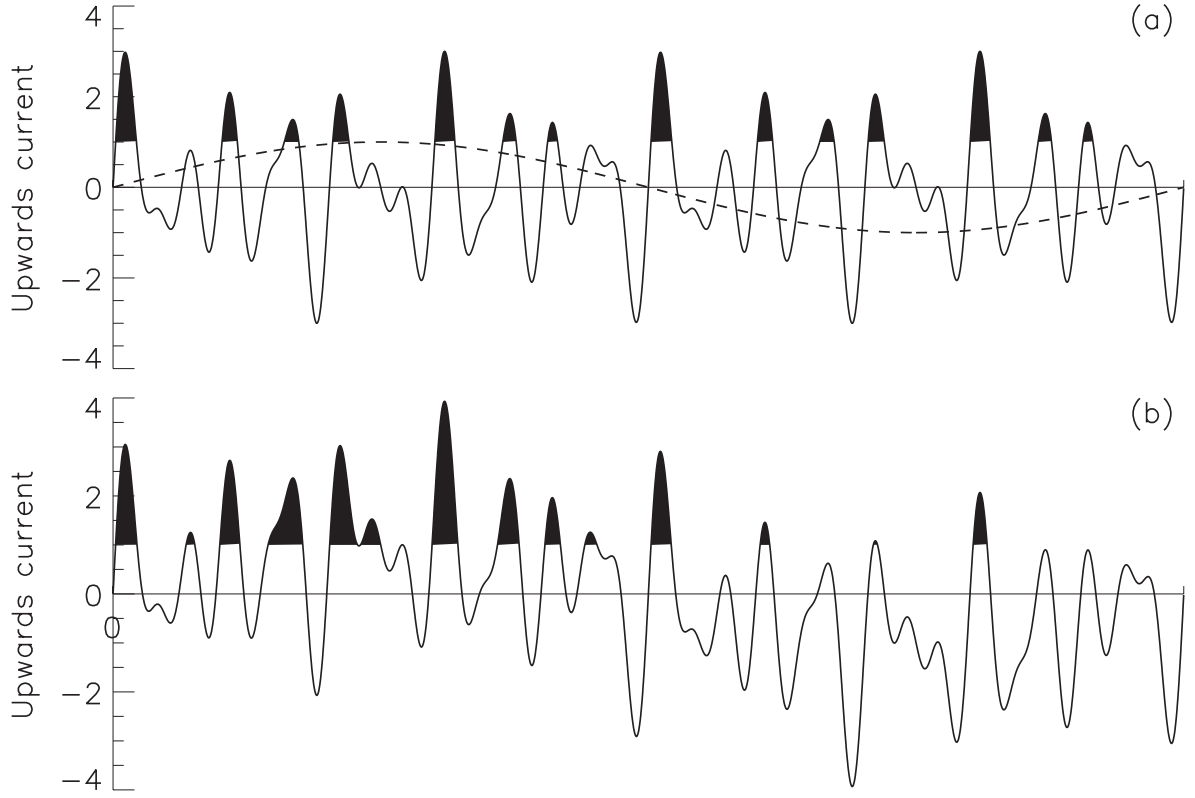


**Figure 2.** Illustration of the effect of a large-scale field-aligned current system on electron precipitation. (a) The solid line is a sketch of small-scale field-aligned currents, in arbitrary units. When the magnitude of these currents exceeds a threshold, $j = 1$, field-aligned acceleration of electrons occurs. These regions are shaded black. The dashed line shows a large scale system of currents, perhaps driven by thermospheric winds, that is of insufficient magnitude to cause field-aligned acceleration. (b) Total current when these two current systems are superposed. On the left-hand side, the upward currents everywhere are enhanced, and it is more likely for field-aligned acceleration to occur. On the right-hand side, the upward currents are everywhere reduced, and field-aligned acceleration is less likely.

is at work, the enhancement would not be immediately obvious in UV images of the polar cap. However, it could be detected by an analysis of the integrated brightness in each longitude sector, which in this model should show a maximum at a particular longitude, rotating with the same period as the SKR emissions.

## 3 NUMERICAL MODEL

To assess the plausibility of the feedback process discussed in Section 2, we employ a numerical model of Saturn's thermosphere. Our approach to this modelling effort is to construct as simple a model as possible in order to gain insight into the behaviour of the thermosphere. Our experience modelling the thermospheres of Jupiter and Saturn (e.g. Achilleos et al. 1998; Smith et al. 2007) indicates that including too much detail in a model may, when interpreting simulation results, obscure the basic physics at work. There is an essential mismatch between the level of complexity that we could, in principle, include – for example full calculations of ionospheric chemistry and transport – and our ability to do justice to this complexity within the framework of a global model with limited spatial resolution and limited data available for validation. Our intention is therefore to analyse the physical processes at work using what is essentially a 'toy model' of Saturn's thermosphere, and thus draw some general conclusions about the potential for feedback processes to drive thermospheric axial asymmetries on the long time-scales that are evidently required.

The core of our effort to model the feedback process described in Section 2 is the same thermosphere–ionosphere model used by Smith (2011). This is a version of the model originally described by Müller-Wodarg et al. (2006). Briefly, the model solves the Navier–Stokes equation on a fixed (Eulerian) latitude–longitude-pressure grid, including terms due to advection, curvature, Coriolis forces, viscosity, pressure gradients and ion drag. Neutral temperatures are calculated taking into account advection, viscous heating, Joule heating, thermal conduction and adiabatic heating and cooling, in addition to a series of additional heating distributions, labelled $\mathcal{Q}_X$, which will be described below.

We do not explicitly calculate plasma flows either in the ionosphere or magnetosphere. We assume a simple, fixed, axially symmetric plasma flow distribution for the magnetosphere, sufficient to drive a realistic baseline distribution of Joule heating and ion drag in the thermosphere. The electric fields associated with the plasma flows are assumed to penetrate unmodified across the full depth of the region that we model. The thermosphere and magnetosphere models are coupled using a simple, fixed model of ionospheric conductance derived from an ionosphere model (Moore et al. 2004). The plasma flow and conductance models will be described in more detail below.

For this study we use the same basic model parameters as Smith (2011): a latitude resolution of 2°, a longitude resolution of 10° and 20 pressure levels at a resolution of 0.5 scale heights. The base of the model is at a pressure of 100 nb, with the base temperature held at 143 K consistent with the temperature profile presented by Moses et al. (2000). The highest altitude pressure level lies at 0.0075 nb. Approximate thermal and dynamical equilibrium of the model is established on a time-scale of ∼400 planetary rotations (Smith et al. 2005; Smith & Aylward 2008). For this reason, we always start the model 'cold' at a global temperature of 143 K, and run it forward for these 400 planetary rotations with axially symmetric heating and conductance parameters, to ensure that an approximate, axially symmetric global equilibrium is in place before introducing any axial asymmetries.

Since the focus of this study is rotational periods, it is important to be clear about how the planetary rotation is communicated to the model. The model is cast in a rotating frame with the wind speeds at the base of the model set to zero. In practice, this means that the base of the model is assumed to rotate rigidly at $\Omega_S$, where this rotation rate is the same as that used to calculate Coriolis forces. An alternative means to communicate the planetary rotation would be to assume a free-slip condition at the lower boundary and allow the thermosphere to evolve in isolation from lower regions of the atmosphere, with a conserved total angular momentum. This situation would be physically unrealistic, because at high latitudes ion drag causes angular momentum to be transferred to the magnetosphere. In steady state, this angular momentum is replaced by upward transfer from the lower regions of the atmosphere. Thus, if the thermosphere was effectively isolated from the lower atmosphere, via a free-slip boundary condition, its angular momentum would not be conserved, but would ultimately all be transferred to the magnetosphere. The zero-slip boundary condition thus expresses the requirement for upward transfer of angular momentum by viscous or advective processes.

We use $\Omega_S = 1.6378 \times 10^{-4}\,\mathrm{rad\,s^{-1}} = 810.8\,\mathrm{d^{-1}}$, where $1\,\mathrm{d} = 24\,\mathrm{h}$ and $1^\circ\mathrm{d}^{-1} = 2.02 \times 10^{-4}\,\mathrm{rad\,s^{-1}}$. This is the rotation rate consistent with the rotation period of Desch & Kaiser (1981). Although it is now known that this rotation rate is not definitive, we retain this number for consistency with previous versions of the model. Alternative determinations of the internal rotation period of Saturn (Anderson & Schubert 2007; Read, Dowling & Schubert 2009) differ by ∼1 per cent. For practical purposes, the exact number that we choose should make very little difference: changes of a few per cent to the rotation rate are highly unlikely to have a significant effect on thermospheric dynamics, and in any case such a small change is almost certainly minor compared to the much larger uncertainties and simplifications elsewhere in our modelling. Indeed, rather than being interested in absolute rotation speeds, what we are interested in is how the rotation rate of structures in the thermosphere varies *relative* to the fixed rotation speed of the lower boundary of the model. Since the actual rotation speed of the lower boundary of the thermosphere is unknown, it is these relative rotation speeds that we can reasonably compare to the observations.

In carrying out such comparisons, we will use the conditions at 74° latitude as representative of the thermospheric behaviour. We choose this latitude for a number of reasons.

(i) It is approximately in the centre of the latitude range to which we apply the feedback process (60°–90°).

(ii) It is close to the shear in the magnetospheric plasma flow (Cowley et al. 2004a) which has an important influence on thermospheric dynamics (Smith & Aylward 2008).

(iii) This latitude was previously used as a diagnostic location by Smith (2011) because, in a dipole magnetic field model, it maps to $L \sim 13$ in the equatorial magnetosphere, which lies in the region of the cam currents proposed by Southwood & Kivelson (2007).

(iv) In most of the runs presented here, it corresponds closely with the central region of the axially asymmetric structures that develop as a result of the feedback process.

### 3.1 Thermospheric heating

As already mentioned in Section 2, the energy balance of Saturn's thermosphere is not well understood. In particular, the high exosphere temperature of ∼400 K is unexplained. Smith (2011) reproduced the approximate thermal structure by imposing a globally

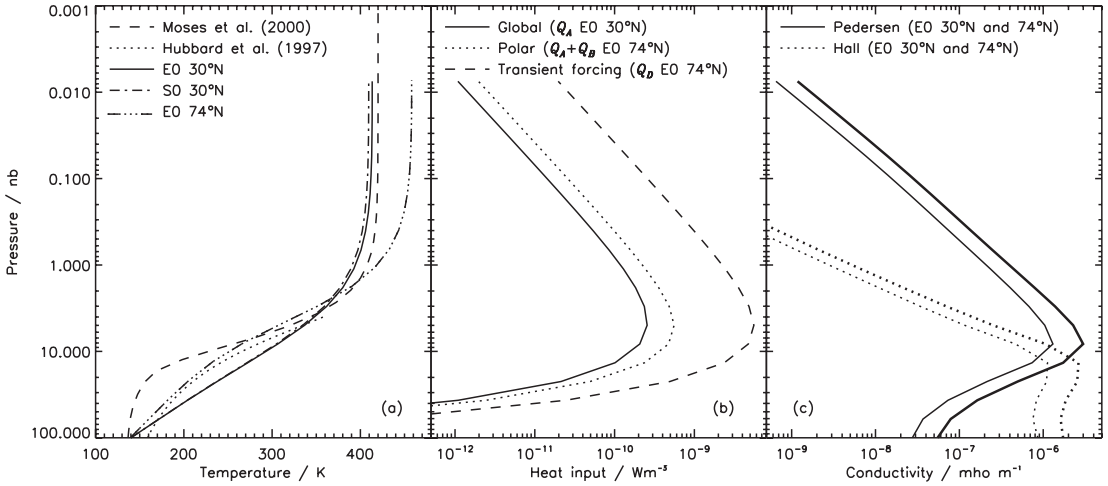


**Figure 3.** Profiles of temperature, heating and conductance. (a) Comparison between observed and calculated temperature profiles. Observations presented by Hubbard et al. (1997) are shown with the dotted line and those presented by Moses et al. (2000) by the dashed line. The thermospheric part of the Moses profile is based on the data of Smith et al. (1983). The calculated profile at 30°N in our unperturbed thermosphere model is shown with solid and dot–dashed lines for the equinox (E0) and solstice (S0) models, respectively. The profile at 74°N in the equinox model (E0) is shown with the triple-dot–dashed line. The profiles at 74°N for the solstice model (S0) are not shown, for clarity. They vary significantly from the equinox profile only above the 4 nb level, by ∼10 K. (b) Heating profiles per unit volume. Solid line: global baseline heating rate $\mathcal{Q}_A$, calculated at 30°N for equinox conditions (E0); dotted line: enhanced heating rate $\mathcal{Q}_B$ polewards of 60° latitude in unperturbed conditions, calculated for equinox conditions (E0) at 74°N; dashed line: transient heating $\mathcal{Q}_D$, also calculated for equinox conditions (E0) at 74°N. (c) Conductivity profiles in mho m$^{-1}$. Solid lines show Pedersen conductivity and dotted lines show Hall conductivity. The thicker lines show the conductivities calculated at 74°N using the equinox temperature profile (E0); the thinner lines show conductivities calculated at 30°N using the equinox temperature profile (E0).

uniform heat source with the following vertical distribution as a function of pressure $p$:

$$q_m = q_{m\infty} \exp\left(-\frac{p}{p_0}\right), \tag{1}$$

where $q_m$ is the heating rate per unit mass in W kg$^{-1}$, $q_{m\infty}$ is the limiting value at the top of the thermosphere, and $p_0$ is the characteristic pressure that determines the peak altitude of the heating rate per unit volume. Smith (2011) used parameters of $q_{m\infty} = 1.6$ W kg$^{-1}$ and $p_0 = 4$ nb to produce a reasonable match to the occultation data of Smith et al. (1983) and Hubbard et al. (1997), the former being represented by the temperature profile constructed by Moses et al. (2000). These temperature data are shown on Fig. 3(a) by the dashed and dotted lines, respectively.

We use a more complex distribution of heating in order to provide a spatially and temporally variable component at high latitudes, related to particle precipitation. This heating distribution generically represents the effects of all external sources of thermospheric heating, and thus in the analysis that follows we will refer to it for brevity as 'external heating' or simply $\mathcal{Q}$. It is defined as follows, in terms of the heating rate per unit mass $q_m$:

$$\begin{aligned} q_m = {} & q_A(\theta, t) \exp\left(-\frac{p}{p_A}\right) \\ & + q_B(\theta, t) \exp\left(-\frac{p}{p_B}\right) \\ & + q_C(\theta) \exp\left(-\frac{p}{p_C}\right) \times \frac{\delta j_{\parallel}(\theta, \phi, t)}{j_{\parallel 0}} \\ & + q_D(\theta, \phi, t) \exp\left(-\frac{p}{p_D}\right). \end{aligned} \tag{2}$$

The parameters $q_X$ and $p_X$ represent four independent components of the form described by equation 1. In all of the experiments described here, we use $p_A = p_B = p_C = p_D = 4$ nb, consistent with the distribution used by Smith (2011). The factors that are dependent on time $t$, colatitude $\theta$ and longitude $\phi$ are indicated by parentheses. We now discuss each of the four components in turn. For brevity we will refer to each component as $\mathcal{Q}_X$, to the heating rate per unit mass associated with each component as $q_X$ and to the height-integrated heating per unit area associated with each component as $Q_X$.

### *3.1.1 Global heating, $\mathcal{Q}_A$*

Component $\mathcal{Q}_A$ is a global heating distribution whose primary purpose is to establish a global thermal structure that approximately matches the observations. We find that setting $q_A(\theta, t) = q_{A0} = 2.5$ W kg$^{-1}$ provides a temperature of ∼400 K at 30°N, as observed by *Voyager 2* (Smith et al. 1983).[1] To represent non-equinox conditions, we introduce a linear north–south bias in the global heating:

$$q_A(\theta, t) = q_{A0} \times \left[1 + S(t) \times \left(\frac{\theta - 90^\circ}{90^\circ}\right)\right]. \tag{3}$$

Here, $S$ is a north–south bias factor that we set to 10 per cent in our initial non-equinox experiments. This factor can be varied with time to represent seasonality – this is the only time-dependence of this heating component.

The solid line in Fig. 3(b) shows the heating profile for $\mathcal{Q}_A$ at 30°N for the baseline equinox run (E0) to be described below. This profile shows heating rates per unit volume. It is necessary to calculate the heating rate per unit volume for a particular temperature profile, since the quantity that is fixed is the heating rate per unit mass, as a function of pressure. The heating rate per unit volume then depends

[1] Note that the global heating value used by Smith (2011) was 1.6 W kg$^{-1}$. The value used here is slightly greater due to the correction of a minor bug in the thermosphere model that caused the thermal conductivity to be slightly underestimated. We thus require slightly greater heating to achieve similar temperatures.

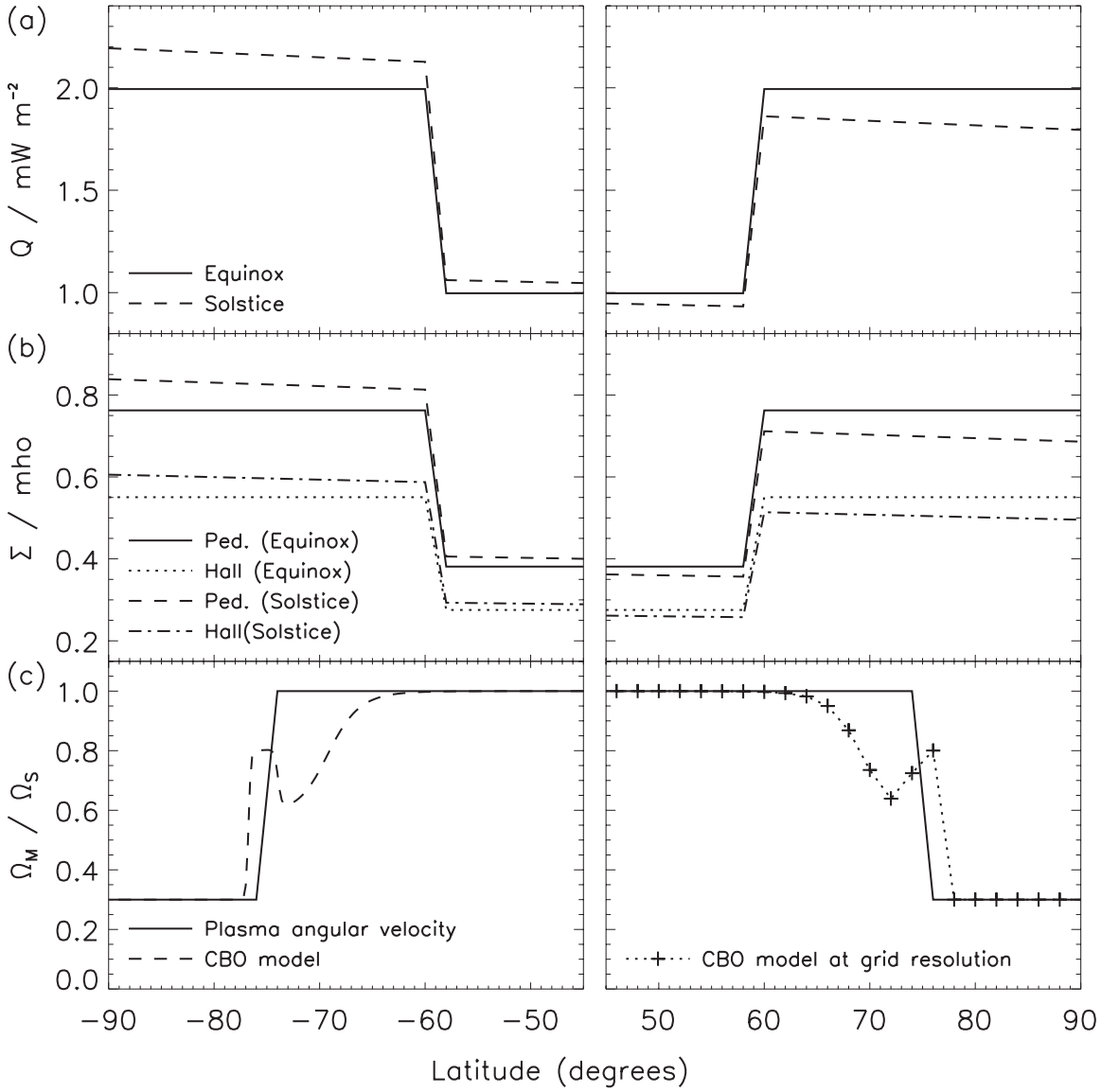


**Figure 4.** Latitude profiles of heating, conductance and plasma angular velocity. (a) Unperturbed column-integrated thermospheric heating. The solid line shows equinox conditions and the dashed line the north–south asymmetric heating intended to represent solstice-type conditions. (b) Column-integrated Pedersen and Hall conductances. Solid and dashed lines show Pedersen conductance (dotted and dot–dashed lines show Hall conductance) for equinox and solstice conditions, respectively. (c) The solid line shows our simplified plasma angular velocity model. For comparison, the left-hand panel shows with a dashed line the detailed empirical model of Cowley et al. (2004a), as mapped to the thermosphere by Smith & Aylward (2008). The right-hand panel shows with a dotted line and crosses the same model at the resolution of our thermospheric grid, illustrating the difficulty in resolving the fine structures in the plasma flow within our thermosphere model.

on the local conditions, decreasing with rising temperature as the atmosphere expands.

### *3.1.2 Axially symmetric particle heating, $\mathcal{Q}_B$*

Component $\mathcal{Q}_B$ represents heating due to particle precipitation present at high latitudes. We have no intention to represent any of the complexity observed in the various UV and IR auroras that stretch from the pole to approximately 60° latitude. We therefore simply add a component with the same functional form as $\mathcal{Q}_A$ (equation 3) at latitudes polewards of 60° latitude. We use $q_{B0} = 2.5\,\mathrm{W\,kg^{-1}}$, thus doubling the existing background heating $\mathcal{Q}_A$ in the high-latitude regions. We expect this heating to be seasonally variable, since higher conductance is expected in the summer hemisphere implying larger currents. We therefore also scale this term using the same bias factor as $\mathcal{Q}_A$ (as expressed in equation 3) – this is the only time-dependence of this heating component. The combined latitude distributions of components $\mathcal{Q}_A$ and $\mathcal{Q}_B$ are indicated in Fig. 4(a).

The dotted line in Fig. 3(b) shows the combined heating profile for $\mathcal{Q}_A$ and $\mathcal{Q}B$ at 74°N, calculated for run E0, as described above for $\mathcal{Q}_A$.

### *3.1.3 Axially asymmetric particle heating, $\mathcal{Q}_C$*

Component $\mathcal{Q}_C$ represents changes to the high-latitude heating due to our feedback mechanism. Since this represents a modification to $\mathcal{Q}_B$, this heating is only applied polewards of 60° latitude. In

the presence of axial asymmetries, this component must increase or decrease in response to large scale systems of field-aligned currents, as discussed in Section 2. Representing this relationship accurately is challenging. The distribution of heating associated with particle precipitation depends on the spectrum of energies and pitch angles of the incident particles at any particular location, which in turn will depend on the size of the field-aligned current and on the distribution of plasma along the relevant flux tube in the magnetosphere. In principle, we could estimate these parameters and then carry out a full calculation of electron energy degradation, secondary electron production and multiple other processes. However, in practice, this simply adds a huge amount of complexity to the situation that is unlikely to help shed light on the basic physical processes involved.

In line with our general approach, we therefore seek a very simple formulation of the dependence of heating on field-aligned current. First, we note that some of the calculated field-aligned current is associated with heating that occurs even when the thermosphere is axially symmetric, in particular with precipitation in the region of the main auroral oval. We therefore assume that axially symmetric components of the field-aligned current are already represented in the heating distribution, and only modify the heating in response to axially asymmetric field-aligned currents, which we label $\delta j_\parallel$. These field-aligned currents are calculated simply by finding the divergence of the horizontal current $\boldsymbol{J}$ and removing the longitude-averaged part:

$$\delta j_\parallel = -(\nabla \cdot \boldsymbol{J} - \overline{\nabla \cdot \boldsymbol{J}}), \tag{4}$$

where the overline indicates an average over longitude and a positive value of $\delta j_\parallel$ represents an upward field-aligned current. The heating component $\mathcal{Q}_C$ is then scaled linearly with $\delta j_\parallel$ as shown in equation (2), with parameters $q_C(\theta) = q_{C0} = 2.5\,\mathrm{W\,kg^{-1}}$ polewards of 60° latitude and $j_{\parallel 0} = 0.02\,\mathrm{nA\,m^{-2}}$. We set $q_C(\theta) = 0$ equatorwards of 60° latitude. We do not introduce a seasonal bias into this parameter, since it directly represents a change in the heating due to a change in the current, which should maintain the same form regardless of the background conductance and heating. Furthermore, $\mathcal{Q}_C$ may represent an increase or decrease in the heating due to particle precipitation. However, it can only decrease this heating to zero, not cause it to become negative. We therefore do not allow the sum of components $\mathcal{Q}_B$ and $\mathcal{Q}_C$ to become negative.

Our choices for $q_C$ and $j_{\parallel 0}$ are essentially arbitrary. A sensitivity study of different values of $j_{\parallel 0}$ is presented in Section 6, and the physical significance of the chosen values of $j_{\parallel 0}$ and $q_C$ will be discussed further in Section 7. However, it is worth emphasizing immediately that the values chosen for $j_{\parallel 0}$ and $q_C$ imply precipitating electron energies much greater than those that are observed. This aspect is at present the major discrepancy in matching our model to the observations.

#### *3.1.4 Transient heating, $\mathcal{Q}_D$*

Finally, component $\mathcal{Q}_D$ is a transient heating distribution that is used to introduce an initial asymmetry to the model. We use a distribution that is identical to the axially asymmetric heating imposed by Smith (2011): $q_D = 50\,\mathrm{W\,kg^{-1}}$ in the longitude range 180°–360° and the latitude range 72°–78°N. This is imposed for a single rotation at the start of some model runs in order to 'kick-start' an asymmetry. As we will show, under certain conditions the thermosphere needs only a very small initial asymmetry for it to grow and become self-sustaining, so the exact distribution of this transient heating term appears to be relatively unimportant.

The dashed line in Fig. 3(b) shows the heating profile for $\mathcal{Q}_D$ at 74°N, calculated for run E0, as described above for $\mathcal{Q}_A$.

### 3.2 Ionospheric conductivities

Thermosphere–ionosphere coupling is implemented using the same simplified scheme of Smith (2011), which was described in detail by Smith & Aylward (2008). This entails using data from an ionosphere model to produce a fixed, global map of Pedersen and Hall conductivities and then prescribing these in such a way that the height-integrated conductivities are independent of variations in the thermal structure. We use the same conductivity profiles as Smith (2011), which were derived from a version of the Moore et al. (2004) ionosphere model (Luke Moore, private communication). Smith (2011) averaged the model ionospheres from 70°N and 70°S (in longitude and across the two hemispheres) to produce a single representative ionosphere profile that was applied globally. We use precisely this profile, however we choose to represent the higher conductance at high latitude by reducing the conductances by a factor of 2 at latitudes below 60°. This latitude is the same as the boundary of heating due to precipitation, representing our expectation that conductances at high latitudes are elevated due to particle precipitation (Moore et al. 2010). The latitudinal profiles of Pedersen and Hall conductance are shown by the solid and dotted lines in Fig. 4(b). To represent non-equinox conditions, we scale the conductances in exactly the same way as we scale heating components $A$ and $B$, as represented by equation (3). The resulting latitudinal profiles of Pedersen and Hall conductance are shown in Fig. 4(b) by the dashed and dot–dashed lines. It should be emphasized that although we have modified the conductance distribution of Smith (2011) so that it varies in latitude, it remains axially symmetric.

Fig. 3(c) shows the profiles of conductivity for the baseline equinox run (E0) to be described below. These show conductivity in units of mho $\mathrm{m^{-1}}$. As for the heating profiles shown in Fig. 3(b), these conductivity profiles must be calculated for a particular temperature profile due to the procedure by which they are fixed, as described in detail by Smith & Aylward (2008). The solid and dotted lines show Pedersen and Hall conductivities; the thick and thin lines show values at 74°N and 30°N, respectively. The structure of these conductivity profiles is unsurprising. The Pedersen conductivity peaks at a higher altitude than the Hall conductivity and declines more gradually with increasing altitude, such that Pedersen conductivity is most important for most of the altitude range studied.

### 3.3 Magnetospheric plasma flows

Smith & Aylward (2008) presented a detailed analysis of the interaction between the thermosphere–ionosphere and magnetosphere at flow shears in the rotational profile of the magnetosphere, as represented by the model of Cowley et al. (2004a). Such an analysis is not the purpose of this study, nor is it possible with the reduced latitudinal resolution necessitated by the use of a three-dimensional model. We therefore use the same very simplified model of the rotational plasma flows in the magnetosphere as Smith (2011). This is enough to represent the overall forcing of the thermosphere–ionosphere by the magnetosphere without introducing unnecessary complications. To this end, we assume that the plasma in the magnetosphere rigidly corotates with the planetary angular velocity (which, for the purposes of the thermosphere model, means rigid corotation with the lower boundary pressure surface) at colatitudes greater than 15°.

Polewards of this colatitude, we assume 30 per cent of rigid corotation, an appropriate value according to observations of the polar cap region (Stallard et al. 2004). Our identification of this colatitude as a nominal polar cap boundary is justified by studies which use realistic models of the planetary magnetic field, and which identify colatitude $\sim 15^\circ$ as being magnetically conjugate to the region of the day-side magnetopause (Cowley et al. 2004a; Achilleos et al. 2010). These plasma flows are shown in Fig. 4(c). In the left-hand panel, the Cowley et al. (2004a) model is shown as a dashed line for comparison. In the right-hand panel, the same model is shown at the resolution of our model grid to illustrate the difficulty in resolving these structures within our low-resolution, three-dimensional model.

The plasma flows then interact with the thermosphere using the same formulations of Joule heating and ion drag described by Smith & Aylward (2008). Following Smith (2011), we assume a constant vertical magnetic field at all latitudes, taking a round value of 60 000 nT, consistent with the fields observed in Saturn's polar regions (Davis & Smith 1990). The advantage of assuming a constant field is that our model of conductivity can be held fixed at all latitudes while retaining a reasonable degree of self-consistency. The advantage of assuming a vertical field is that we can use very simple expressions for Joule heating and ion drag, and that all altitudes at a particular latitude–longitude grid point correspond uniquely to a single magnetic field line. The assumption of a constant vertical field is good at high latitudes (the dip angle at $60^\circ$ latitude in a dipole field is $\simeq 74^\circ$) but becomes progressively weaker moving towards the equator. However, at latitudes below $60^\circ$, the horizontal currents do not contribute to the feedback effect, since this only applies to the heating component $\mathcal{Q}_C$ which is only included polewards of $60^\circ$ latitude. Currents in this region therefore only have an effect on local Joule heating and ion drag, which are expected to be relatively unimportant at low latitudes given the lack of an external driver in the form of large-scale magnetospheric electric fields.

### 3.4 Baseline results with no axial asymmetries

Fig. 3(a) shows the temperature profiles from our baseline equinox and solstice models with no axially asymmetric forcing, which we label E0 and S0. respectively (see Table 1 for a summary of all of the model runs described in this paper). As discussed in Section 3, the models were run from 'cold', starting at 143K globally, for 400 rotations to establish approximate equilibrium. The predicted temperature profiles are compared to the Moses et al. (2000) profile (solid line) and the Hubbard et al. (1997) data (dotted line). The solid line and dot–dashed lines respectively show the temperature profiles in the equinox (E0) and solstice (S0) models at $30^\circ$N latitude. Overall, it is clear that our baseline heating profile does provide a global temperature that is broadly supported by the data for both equinox and solstice profiles. The triple-dot–dashed line shows the temperature in the equinox model at $74^\circ$N, which, unsurprisingly, is somewhat higher than the low latitude temperatures, in line with the greater heating in the polar regions. For clarity, high-latitude profiles in the solstice model are not shown. They only differ significantly from the equinox profile above the 4 nb level, by no more than $\sim 10$ K.

The dynamical structure of the northern polar region in the baseline equinox model (E0) is shown in Fig. 5(a). This is in line with the basic dynamical structure described by Smith et al. (2007): a broad region of sub-corotation at latitudes where ion drag dominates; a region of super-corotation equatorwards of these latitudes; a general poleward flow across the polar region; and a resultant convergence and downwelling at the pole. Due to the three-dimensional nature of our model necessitating a relatively low-latitude resolution, the details of our baseline model differ slightly from those of Smith et al. (2007), but it is clear that the overall behaviour is qualitatively identical.

## 4 SYMMETRY BREAKING EXPERIMENTS

We have now established baseline axially symmetric model atmospheres within which we can study the feedback effects proposed above. In the following we describe a further 400 rotation model run in the equinox model, which we label E1 (see Table 1). We apply the axially asymmetric forcing $\mathcal{Q}_D$ for the first planetary rotation only, and then allow the model to develop freely with our feedback mechanism active.

### 4.1 Structure of asymmetry after 400 rotations

After 400 rotations, clear and persistent axially asymmetric structures have developed in the Northern and Southern hemispheres. We will first analyse the Northern hemisphere asymmetry (as it stands after the full 400 rotations of run E1) before looking in more detail at the development and evolution of the axial asymmetries in both hemispheres. The axially asymmetric components of the thermal

**Table 1.** Summary of model runs.

| Input | Run | Rotations | $S(t)$ | $\mathcal{Q}_D$ | Description |
|---|---|---|---|---|---|
| 143 K global | E0 | 400 | 0.0 | No | Baseline equinox run with no axial asymmetries. |
| E0 output | E1 | 400 | 0.0 | Rot 1 | Equinox run with transient asymmetric forcing $\mathcal{Q}_D$ during Rotation 1. |
| E1 output | E2 | 800 | 0.0 | No | Continuation of E1 for comparison with S2. |
| 143 K global | S0 | 400 | 0.1 | No | Baseline solstice run with no axial asymmetries. |
| S0 output | S1 | 400 | 0.1 | Rot 1 | Solstice run with transient asymmetric forcing $\mathcal{Q}_D$ during Rotation 1. |
| S1 output | S2 | 800 | $0.1 \times \cos \omega t$ | No | Continuation of S1 to test seasonal variation. $\omega = 2\pi/T$, where $T = 800$ rotations. |
| E0 output | E1a | 400 | 0.0 | Rot 1 | As E1 with $j_{\|0} = 0.015\,\mathrm{nA\,m^{-2}}$. |
| E0 output | E1b | 400 | 0.0 | Rot 1 | Identical to E1, ($j_{\|0} = 0.020\,\mathrm{nA\,m^{-2}}$). |
| E0 output | E1c | 400 | 0.0 | Rot 1 | As E1 with $j_{\|0} = 0.025\,\mathrm{nA\,m^{-2}}$. |
| E0 output | E1d | 800[a] | 0.0 | Rot 1 | As E1 with $j_{\|0} = 0.030\,\mathrm{nA\,m^{-2}}$. |
| E0 output | E1e | 400 | 0.0 | Rot 1 | As E1 with $j_{\|0} = 0.035\,\mathrm{nA\,m^{-2}}$. |
| E0 output | E1f | 400 | 0.0 | Rot 1 | As E1 with no feedback ($j_{\|0} \to \infty$). |

[a]The run time of 800 rotations for this run (E1d) appears anomalous. This longer run time was necessary to ensure that the rotation rates of the axial asymmetries had sufficient time to stabilize (see Section 6 and Fig. 15).

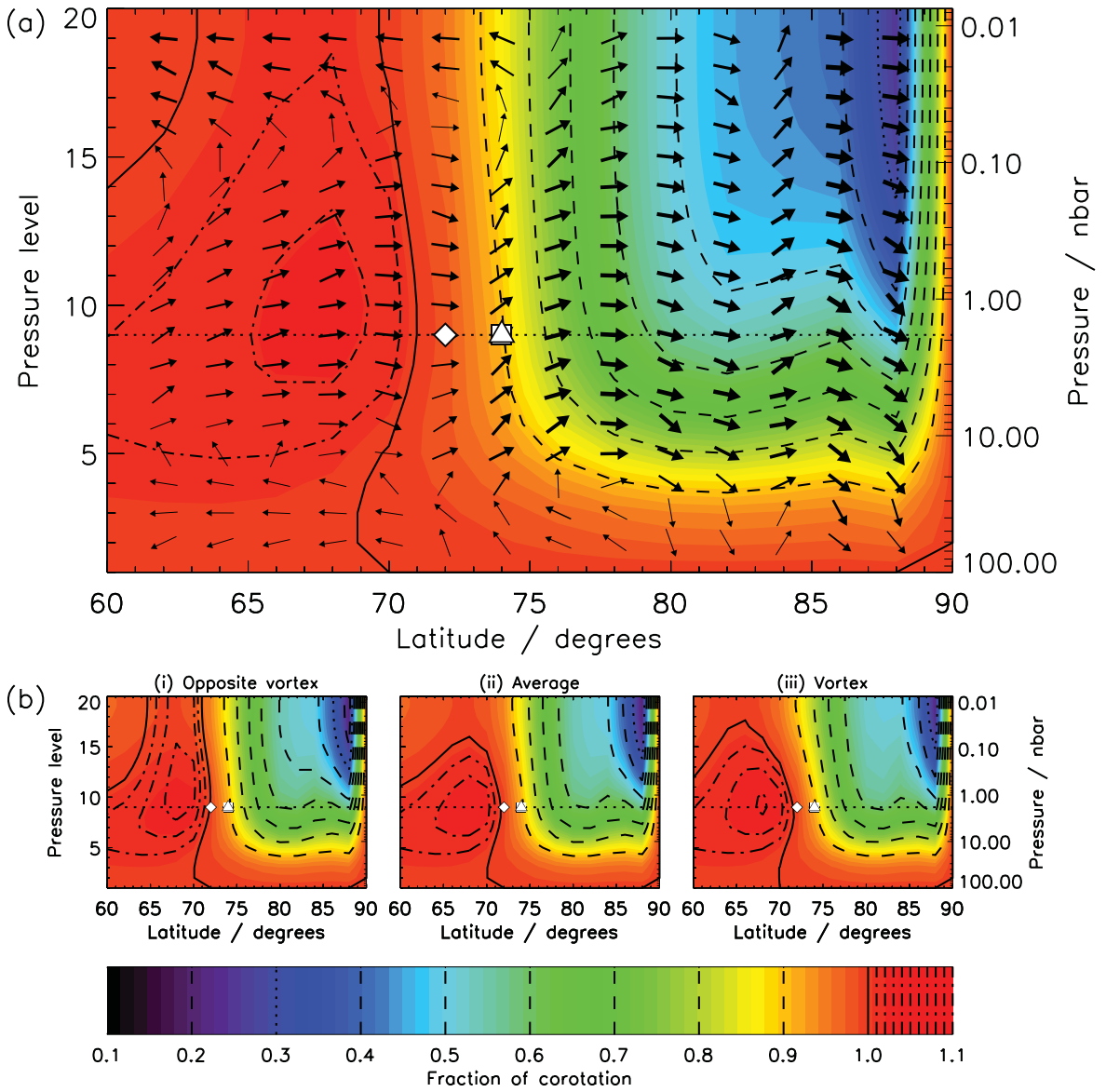


**Figure 5.** Panel (a): circulation in baseline equinox model (E0). The vertical range is equivalent to ∼2000 km. This varies slightly across the range shown due to the variation of the thermal structure with latitude. The latitude range shown is equivalent to a distance of ∼30 000 km, so the plot has been horizontally 'squashed' by a factor of ∼15 to accommodate both dimensions in a legible format. The direction of the arrows has been similarly scaled, thus exaggerating the size of the vertical winds but preserving the overall flow pattern. The colour scale and contour lines illustrate the zonal winds compared to the planetary angular velocity. The solid contour shows gas rotating at $\Omega_S$ (with the planet) and the dotted contour (only visible in the top right of the figure) shows gas rotating at the polar plasma sub-corotation speed of $0.3\Omega_S$. The dashed lines show winds rotating between these extremes at 10 per cent increments of $\Omega_S$ and the dot–dashed lines show super-corotating winds at 1 per cent increments of $\Omega_S$. The arrows show the meridional circulation. The thickness of the arrows indicates the overall speed at each location, most of which is in the meridional direction due to the necessary distortion of the plot. The thickest arrows show speeds of $\sim 100\,\mathrm{ms}^{-1}$. The horizontal dotted line marks pressure level $n = 9$ (1.8 nb) for which we plot information in Figs 6–8 and the plot symbols mark the latitude of various features of the vortex shown in these figures, as described in the caption to Fig. 6. Panel (b): each plot shows the same information about the zonal winds as panel (a), but for the equinox model after the development of axial asymmetries (E1). The central panel (ii) shows the longitude averaged circulation. The panel to the right (iii) shows the circulation at the central longitude of the vortex (where the temperature peaks at $n = 9$, as marked by the square and dot–dashed line in Fig. 6) and the panel to the left (i) the same information at the opposite longitude (as marked by the triple-dot–dashed line in Fig. 6).

structure and winds after 400 rotations at pressure level $n = 9$ (1.8 nb) are shown in Fig. 6. These components have been calculated by longitude averaging the temperature and winds at each latitude and pressure level and then subtracting this longitude average from the original values. The corresponding axially asymmetric horizontal currents associated with this pressure level are shown in Fig. 7 and the axially asymmetric components of the principal terms in the energy equation at the same level are shown in Fig. 8. This level is chosen because the horizontal currents and divergences at this level are representative of their height-integrated values: we can thus discuss the relationship between a specific temperature and wind field and the currents that they drive, while also giving a description of the overall height-integrated behaviour. In Figs 6–8 the square and triangle indicate, respectively, the location of the maximum

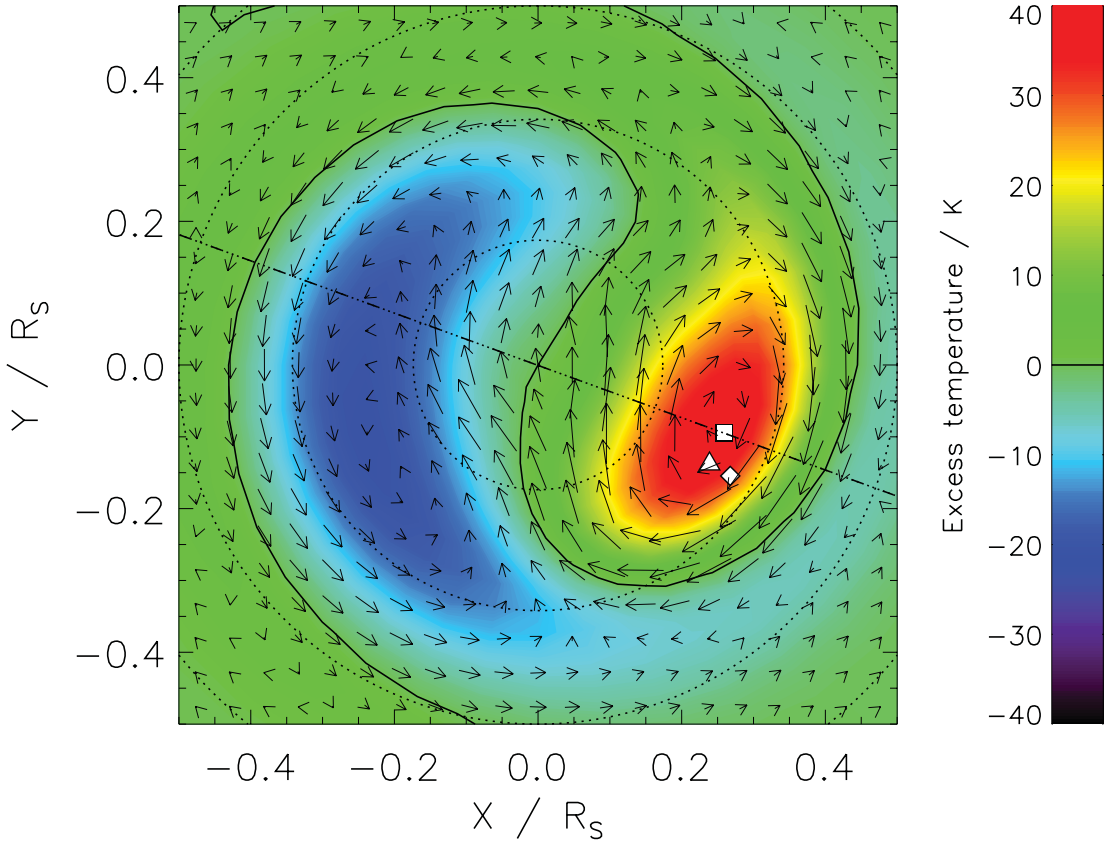


**Figure 6.** Axially asymmetric components of temperature and wind in the northern polar cap at pressure level $n = 9$ (1.8 nb). The dotted lines show circles of constant latitude at 10° separations. The colours show the axially asymmetric temperature, with zero excess temperature indicated by the solid contour. Arrows show horizontal winds. The longest arrows indicate an excess horizontal wind of $\sim$12 m s$^{-1}$. The plot symbols show the locations of the maximum temperature (square), current convergence (triangle) and external heating (diamond) at this pressure level. Note that the maximum of the external heating lies at the same latitude and longitude at all altitudes. The triple-dot–dashed and dot–dashed lines indicate the locations of the meridional slices shown in Figs 5(a) and 5(c), respectively.

values of the temperature and current convergence at $n = 9$. The diamond indicates the location of the maximum value of the total external heating $\mathcal{Q}$. The only axially asymmetric component of the external heating that is present at the end of the model run is $\mathcal{Q}_C$, which depends directly on the current convergence via the axially asymmetric field-aligned current $\delta j_\parallel$. The maxima of the horizontal current convergence and external heating at $n = 9$ are not at precisely the same location because the external heating is dependent on the height-integrated, rather than the local, current convergence.

Looking first at Fig. 6, it is dominated by a striking vortex-like structure. However, this is *not* a true vortex, in that, as discussed above, the figure shows only the axially asymmetric components of the winds. Looking also at Fig. 5(b), which shows the full zonal winds in a meridional plane at the longitude of the vortex-like structure and at the opposite longitude, the overall zonal wind structure is very similar at all longitudes. The vortex-like structure is thus a relatively minor modification to a wind field dominated by strong sub-corotational zonal winds. Indeed, it is quite astonishing how clear the vortex-like structure appears when the axially symmetric components of the winds are removed. The obvious conclusion is that this is a coherent structure that is being uniformly advected by the sub-corotational winds. However, as is clear from Fig. 5, the sub-corotational winds vary greatly across the 70–80°N latitude range of the vortex-like structure, from full corotation at 70°N to $\sim$60 per cent of corotation at 80°N. The structure is thus being continually 'smeared out' in longitude by zonal wind shear. Individual parcels of gas do not orbit around the centre of the structure as we might expect in a true vortex, but instead experience a perturbation to their largely zonal motion as they pass through the region of the vortex-like structure.

Despite the fact that the structure is not a true vortex, it does show behaviour very similar to that described in Section 2 and sketched in Fig. 1. Looking at Fig. 6, it is clear that the structure does indeed have a 'hot' core. The winds circulate clockwise around this core, and, close to the centre, they flow almost along isothermals such that there is very little outflow from the core itself. Fig. 7 shows the currents convergent at the centre of the structure, again as described in Section 2. We have thus developed a structure in the thermosphere whose flow pattern resembles that of a vortex, but is not a vortex in the true sense. We will thus refer to it as a *pseudo-vortex*.

Fig. 8 illustrates the energy balance that sustains this pseudo-vortex, in the form of the axially asymmetric components of various energy terms. Panel (a) shows the external heating $\mathcal{Q}$, as defined by equation (2). The only axially asymmetric component of $\mathcal{Q}$ is $\mathcal{Q}_C$, which is determined directly by the divergence of the horizontal currents. It is thus unsurprising that the shape of the external heating region is similar to the shape of the region of current convergence in Fig. 7. However, it is perhaps surprising that the main process balancing this heating is vertical advection (which, as

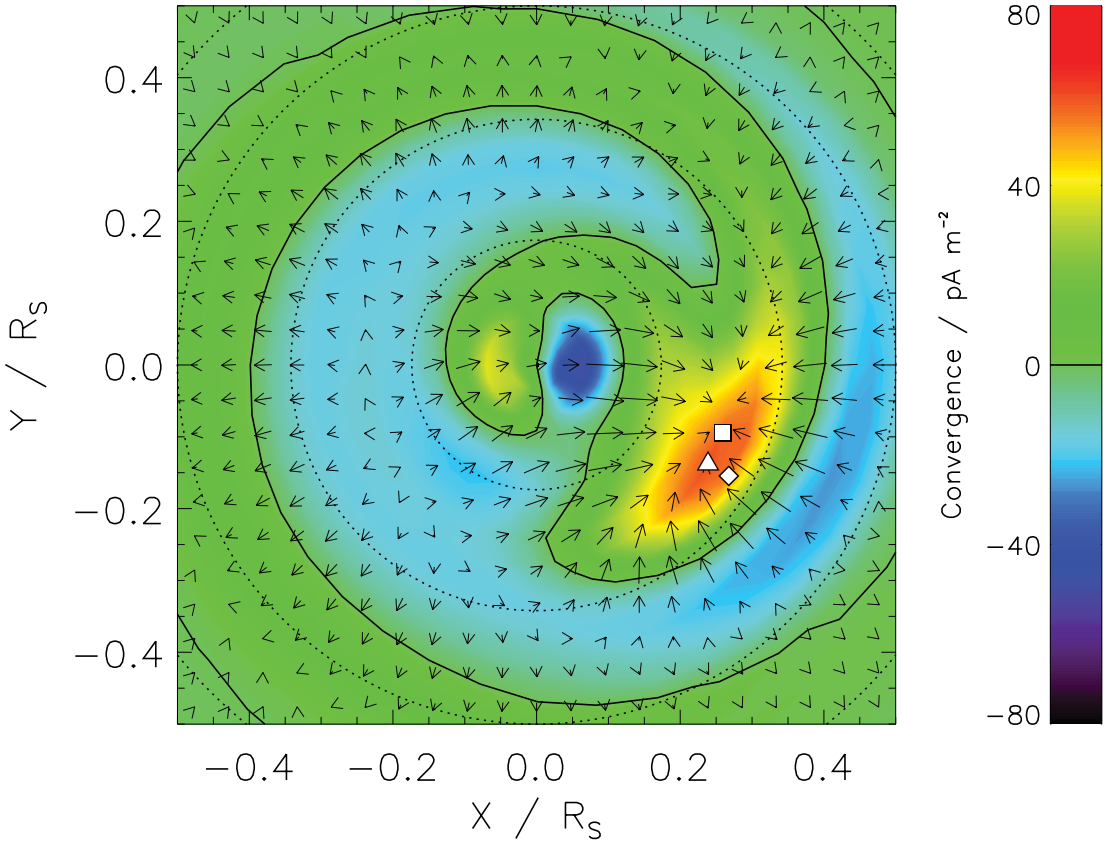


**Figure 7.** Axially asymmetric components of currents in the northern polar cap at pressure level $n = 9$ (1.8 nb). The dotted lines show circles of constant latitude at 10° separations. The arrows show the axially asymmetric horizontal currents at this level. The longest arrows represent horizontal currents of $\sim$0.2 mAm$^{-1}$. The colour scale shows the convergence/divergence of the axially asymmetric horizontal currents at the $n = 9$ pressure level. Positive values represent convergent horizontal currents, implying upward field-aligned currents in these regions.

plotted, includes a small contribution from adiabatic heating/cooling of falling/rising gas), as shown in panel (b). The importance of this term is due to the slight divergence of horizontal winds in the core of the pseudo-vortex, leading to upwelling of gas. The upwelling gas not only comes from cooler altitudes, but also cools adiabatically as it rises.

Panels (c) and (d) show advection of energy due to meridional and zonal winds, respectively. There is substantial advection of the axially asymmetric structures by the axially symmetric components of the winds. Since the axially symmetric winds are dominated by a sub-corotational zonal component, this explains the dominance of the zonal advection component, which is consistent with shifting the temperature peak westwards.

Panel (e) shows the component due to Joule heating and ion drag. This is most significant at latitudes greater than 75° where there is substantial plasma sub-corotation, but even in this region it is a relatively insignificant contributor to the energy balance. Finally, panel (f) shows the effect of vertical conduction. This is clearly correlated with the temperatures shown in Fig. 6, in that hotter regions are cooled the most by thermal conduction and vice versa; however, comparing it to panel (b) it is relatively unimportant as a mechanism for cooling the core of the pseudo-vortex. This is perhaps surprising given that thermal conduction is the most important cooling process when taking account of the energy budget of the thermosphere as a whole. The dominant role of advective heating and cooling in this context illustrates that a localized, vortex-like flow perturbation can give rise to a very different hierarchy of heating and cooling mechanisms compared to an atmosphere which is dominated by symmetric rotational flows.

The plot symbols on Figs 6–8 indicate the locations of the maximum values of the temperature (square), current convergence (triangle) and external heating $\mathcal{Q}$ (diamond) at pressure level $n = 9$. It is significant that the maximum values of these three quantities do not occur at the same location. The difference between the location of maximum current convergence and the location of maximum external heating $\mathcal{Q}$ arises because the heating depends on the height-integrated current convergence; slightly different distributions of horizontal current at other altitudes shift the location of the maximum slightly compared to the maximum at $n = 9$. However, the fact that the location of the maximum external heating is west of the location of the maximum temperature indicates that the pseudo-vortex may naturally shift westwards over time. It seems likely that zonal advection will have a similar effect as the entire pseudo-vortex will be moved westwards by the prevailing winds. The result of these two effects is that the pseudo-vortex does *not* rotate with the same angular velocity as the base of the model, but rather drifts at a rate determined by the relative position of the heating maximum and the speed of the background zonal winds.

To further analyse these effects, Fig. 9 shows energy terms at latitude 74°N and pressure level $n = 9$ for the longitudes coincident with the pseudo-vortex. Fig. 9(a) shows the same terms as plotted in Fig. 8, alongside the axially asymmetric component of temperature (thick grey line). Fig. 9(b) shows various combinations of these terms, to be discussed below. Looking first at Fig. 9(a), it is clear

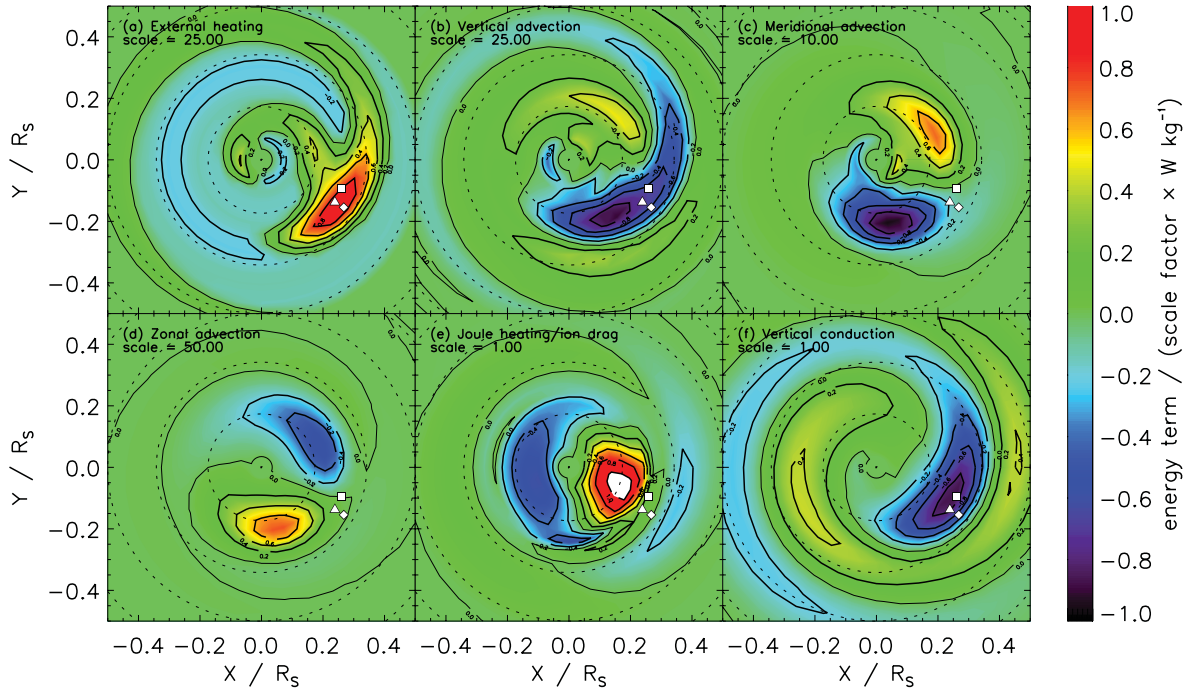


**Figure 8.** Axially asymmetric components of energy terms in the northern polar cap at pressure level $n = 9$ (1.8 nb). The dotted lines show circles of constant latitude at 10° separations. The labelled solid lines and colours show the size of each energy term in units of W kg$^{-1}$ when multiplied by the scale factor shown for each plot. (a) External heating $\mathcal{Q}$. This shows only $\mathcal{Q}_C$ because $\mathcal{Q}_A$ and $\mathcal{Q}_B$ are axially symmetric by definition, and $\mathcal{Q}_D$ is only applied during the first rotation of the model run. (b) Vertical advection and adiabatic heating/cooling due to vertical motion of the gas. (c) Southward advection. (d) Eastward advection. (e) Total thermal and kinetic energy input due to Joule heating and ion drag. (f) Vertical thermal conduction. Other energy terms are relatively unimportant and are not shown.

again that the contribution of vertical thermal conduction and Joule heating/ion drag to the axially asymmetric energy terms is negligible. However, the more significant energy terms can be placed into two groups. First, the external heating and vertical advection peak at approximately the same location, ∼20° west of the temperature peak. Secondly, the two horizontal advection terms peak at approximately the same location, ∼50° west of the temperature peak.

Fig. 9(b) shows these terms grouped together. The dashed line shows the sum of the external heating and vertical advection. The effect of vertical advection is to balance ∼80 per cent of the external heating. However, while the total of these heating rates is reduced relative to the external term, the peak remains in approximately the same location. The dot–dashed line shows the sum of the two horizontal advection terms, with the other (relatively insignificant) terms also included, for completeness. The effect of the meridional advection term is to balance ∼30 per cent of the zonal advection. Again, although the total heating is reduced the peak remains in approximately the same location. Thus, there are two overall peaks of heating, one driven by zonal advection and the other by the external heating. Both peaks are located to the west of the temperature peak and thus both will have the effect of causing the temperature peak to drift westwards. In the following, we will refer to these two distinct drift effects as *advection drift* and *heating drift*. The sum of all terms (solid line) shows that combining these two effects produces a single peak westwards of the temperature peak, to which advection drift is the more important contributor. The overall drift is therefore mostly an advection effect, but this is modified by the heating drift. The importance of this insight is that the rotation rate of the asymmetry is more than just a signature of the zonal wind speed, but is due to a combination of the zonal wind speed with other factors, the most important of which in our model scenario is heating due to particle precipitation.

### 4.2 Development of Northern hemisphere asymmetry

While we have analysed the energy balance of the pseudo-vortex as it exists after 400 rotations, it is interesting also to describe its development. Fig. 10 shows this in terms of the axially asymmetric temperature and winds at pressure level $n = 9$ in the Northern hemisphere during the first 10 rotations of the model. The first panel shows the situation after one quarter rotation. There is a clear semicircular temperature structure generated directly by the transient heating. By the end of the first rotation (second panel), this semicircular structure has intensified. However, it has also rotated clockwise, opposite to the rotation direction of the planet. By the end of the second and third rotations (third and fourth panels), this westward motion clearly continues, and the semicircular structure has begun to break down. However, the warm and cool regions of the asymmetric pattern remain approximately the same size and shape. The final two panels show the situation at the end of the sixth and 10th rotations. The warm region clearly intensifies under the influence of the feedback effect and simultaneously shrinks, leading to an asymmetry in the shapes of the warm and cool regions. By the end of the 10th rotation, a distinctive spiral structure has begun to develop in the warm region, and the structure is essentially identical to that present after 400 rotations, as shown in Fig. 6.

### 4.3 Development of Southern hemisphere asymmetry

The transient heating $\mathcal{Q}_D$ is only applied in the Northern hemisphere and causes an asymmetry to develop as described above. However,

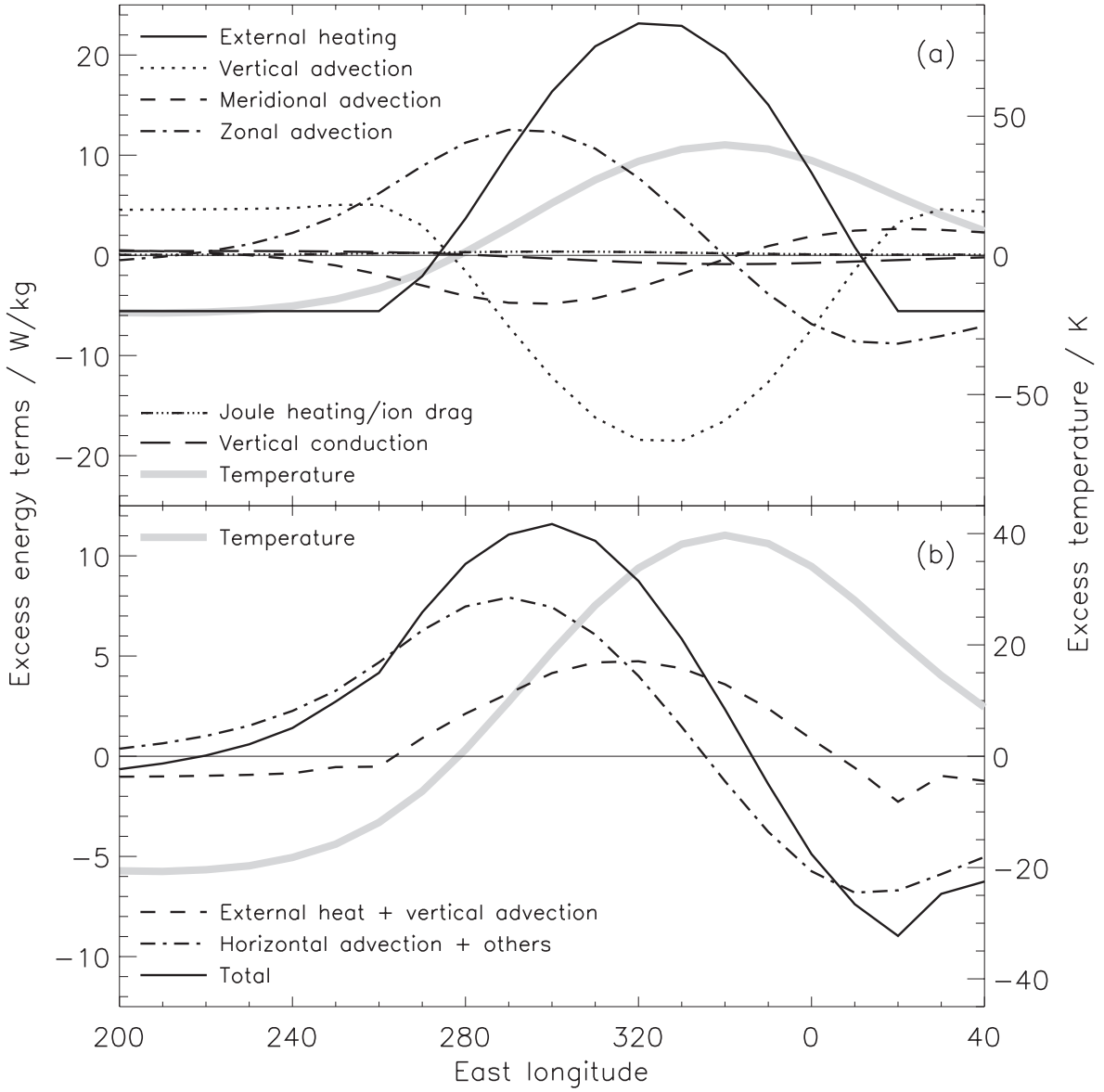


**Figure 9.** Axially asymmetric energy terms and temperatures at 74°N and at pressure level $n = 9$ (1.8 nb) at longitudes coinciding with the core of the pseudo-vortex. The line formats are as shown in each plot. Panel (a) shows the same energy terms as Fig. 8. Panel (b) shows various groups of these terms, as described in the text. The discontinuities in the first derivative of the external heating term at 20° and 260° longitudes are due to the external heating from particle precipitation ($\mathcal{Q}_B + \mathcal{Q}_C$) falling to zero at these locations, as the downward field-aligned currents driven by thermospheric winds become great enough to entirely 'switch off' heating due to particle precipitation.

after 400 rotations, a clear, and almost exactly identical, asymmetry has developed in the Southern hemisphere. This is interesting because the asymmetry in the north does not significantly extend towards the equator – this suggests that the southern asymmetry has developed from a very small initial asymmetry in the south that must, ultimately, have been driven in some way by the asymmetry in the north.

To analyse how this occurs, Fig. 11 summarizes the development of axial asymmetries at all latitudes for the first 100 rotations of run E1. To represent the degree of asymmetry at each latitude, we have calculated the difference between the maximum and minimum temperatures present at that latitude at the $n = 9$ pressure level. Thus, a completely axially symmetric atmosphere at any latitude would be represented by zero. These values for the entire latitude range are represented by the colours in Fig. 11(a), using a log-scale. The values for the equator are shown in Fig. 11(b), again using a log-scale.

A prominent feature of Fig. 11(a) is the large axial asymmetry close to the north pole that develops from the first rotation onwards and persists. This is the asymmetry described above. Also prominent is the development of an almost identical axial asymmetry close to the south pole between 10 and 20 rotations, developing to its full magnitude by 40 rotations. In both of these major axial asymmetries, there is a ∼15 rotation oscillation in the size of the asymmetry, in particular at ∼60° latitude, but also clearly visible at the equator. The cause of this ∼15 rotation oscillation is unclear. It will be discussed further in Section 5.

More interesting than these large-scale features is the detailed behaviour of very small axial asymmetries. Throughout Fig. 11(a), in particular in the first 50 rotations, there are clear diagonal features.

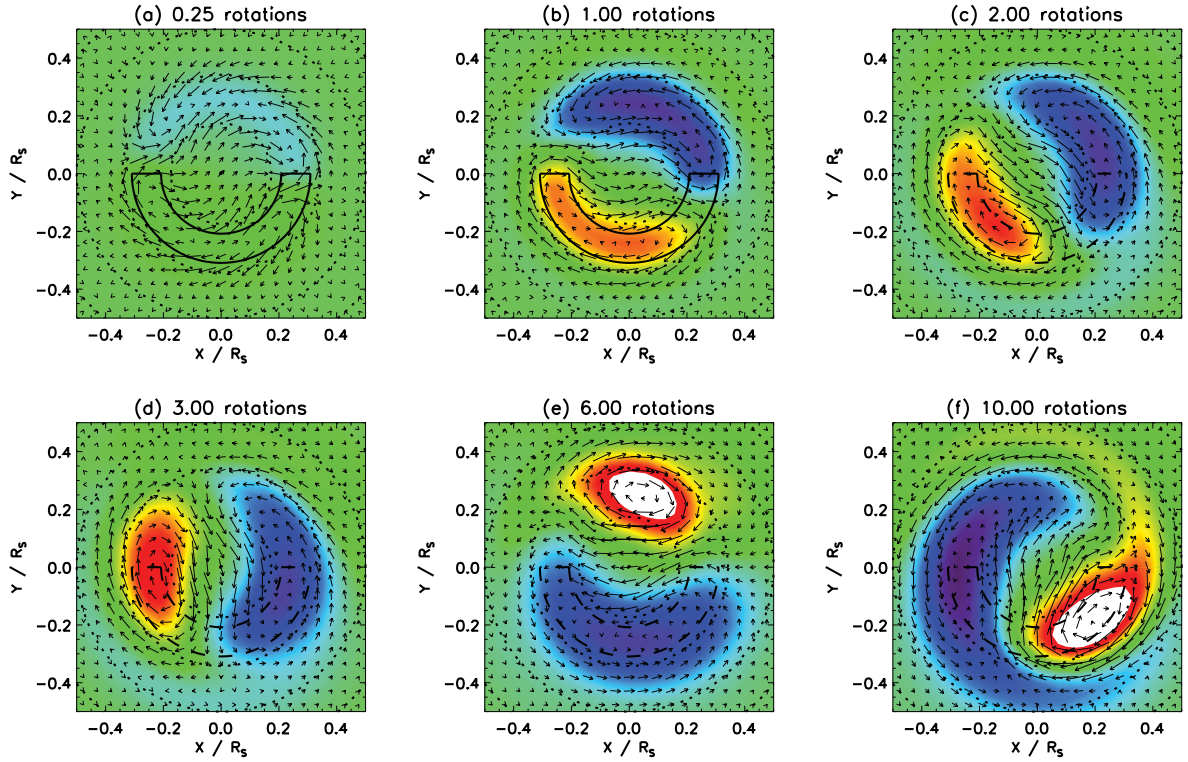


**Figure 10.** Development of the Northern hemisphere pseudo-vortex. Each figure shows the same information as Fig. 6 at the marked time during the first 10 rotations of run E1. The colour scale for temperature is the same as in Fig. 6. White regions indicate values greater than 40 K. The arrows in panels (b)–(f) are scaled such that the longest arrows show wind speeds of $\sim$80 m s$^{-1}$. The arrows in panel (a) are shown at three times this scale. The semicircular box in each panel shows the location of the transient axially asymmetric heating $\mathcal{Q}_D$. The line is broken in panels (c)–(f) to indicate that after the first rotation this heating component is no longer applied to the model.

These are an indication of axially asymmetric structures, which we may reasonably characterize as waves, propagating across the planet. Initially these clearly propagate in a north–south direction. The gradient of these features on Fig. 11(a) indicates the speed of wave propagation. This is easily measured in the equatorial region within the first 20 rotations, when the wave structures are relatively clear. The typical speeds are $\sim$1200 m s$^{-1}$, such that the pole-to-pole travel time for a wave is $\sim$4.4 rotations.

Looking at the first five rotations, the first axial asymmetries to develop in the Southern hemisphere consist of a clear diagonal feature $\sim$3 rotations wide. This appears to be directly related to the initial forcing of the Northern hemisphere with the transient and localized heating $\mathcal{Q}_D$, which lasts for one rotation. Looking at this structure in detail (Fig. 11c), there are two groups of peaks, each approximately one rotation wide and about one rotation apart. The cross polar cap travel time for the waves is of the order of one planetary rotation. It thus seems likely that the first group is due to axially asymmetric waves propagating directly from the location of the heating component $\mathcal{Q}_D$, and the second, delayed group, is due to axially asymmetric waves propagating across the north pole and down the far side of the planet. The period and angular frequency $\omega_1$ of these waves can be estimated by simply measuring the locations of the peaks on Fig. 11(c). The crosses indicate the peaks that we have taken into account for this calculation, which yield $\omega_1 \simeq 4.7\Omega_S$, where $\Omega_S$ is the angular velocity of the lower boundary of the model.

After this initial train of waves, there follows continual wave activity for the whole of the range shown. These waves all have broadly similar characteristics. Fig. 11(d) shows one of these structures in detail in the range 5–18 rotations. We can again estimate the angular frequency of these waves, using the peaks indicated by crosses on Fig. 11(d). This yields $\omega_2 \simeq 2.1\Omega_S$. Not only do waves with this, or a very similar, angular frequency appear throughout the range shown, but beyond the first 20 rotations they no longer show clear north–south diagonal structures, but show either a mixture of north–south and south–north diagonals or patterns indicative of interference. This suggests that these waves are being generated naturally by the large-scale axial asymmetries in both the Northern and Southern hemispheres and are then propagating in both directions across the equatorial region.

On the basis of these results, it seems highly probable that the first set of waves is specifically related to the transient heating, whereas the second set is a natural feature of the thermosphere. A very simple analysis of the force balance in the thermosphere casts some light on this. The winds at high latitudes are almost in geostrophic balance. Usually this means that the force balance is between the Coriolis force and a pressure gradient, producing circulation of winds around high- or low-pressure regions. Polewards of 75° latitude, where there is significant plasma sub-corotation, ion drag is also an important force, and the balance can be between the Coriolis force and a mixture of pressure gradients and ion drag. If we consider a wind in this situation, with components $u$ and $v$, the component $u$ being parallel to the combined driving force per unit mass $F$, then neglecting all viscous drag and advection:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -fu \qquad \frac{\mathrm{d}u}{\mathrm{d}t} = F + fv, \tag{5}$$

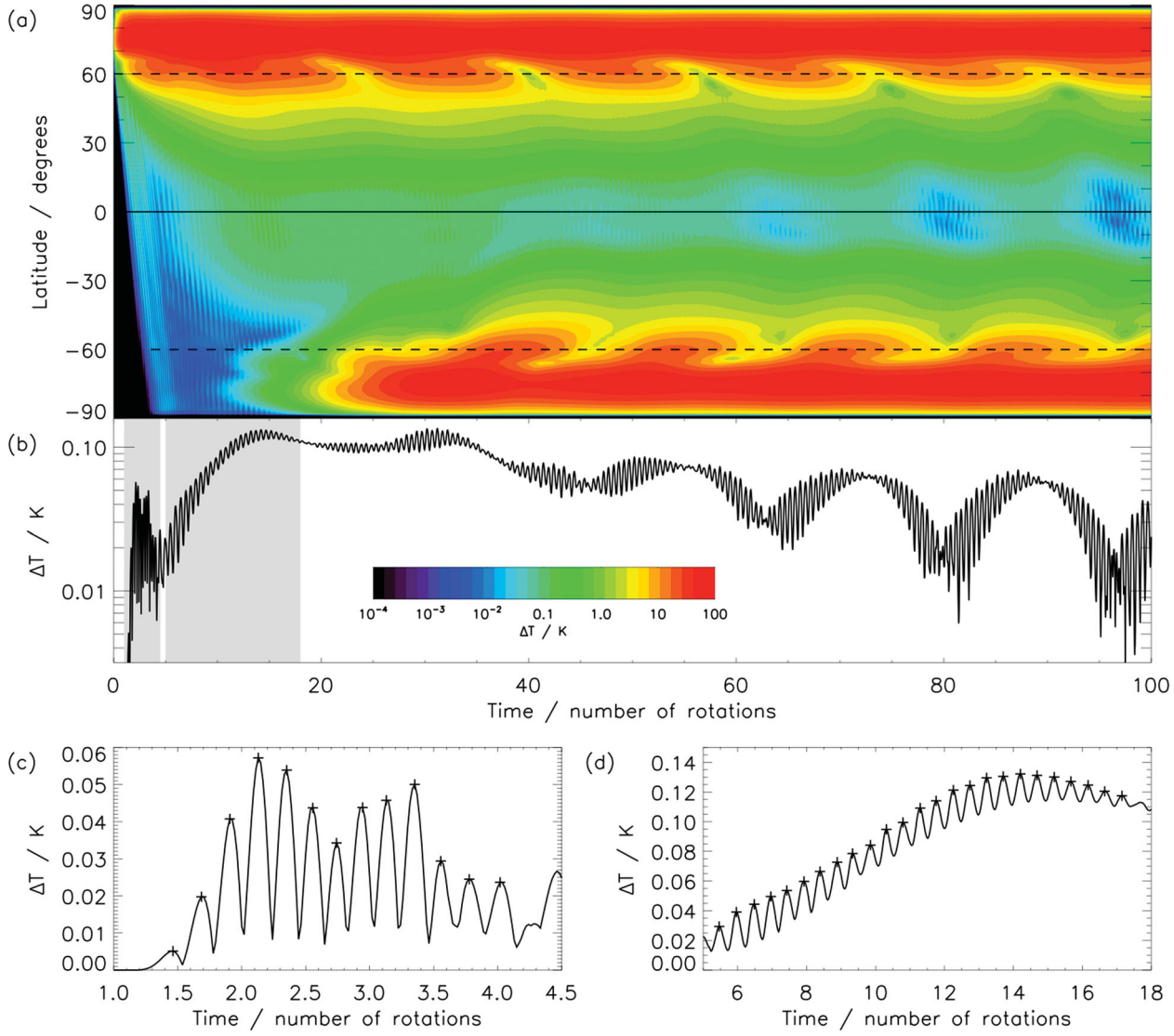


**Figure 11.** Analysis of the development of global axial asymmetries. Panel (a) shows the evolution of the maximum temperature contrast at the $n = 9$ pressure level at each latitude over the first 100 rotations of run E1. The temperature contrast is shown on a log-scale. The horizontal solid line indicates the equator and the horizontal dashed lines show 60° latitude, the cut-off latitude for the feedback effect. Panel (b) shows a cut through panel (a) at the equator, again on a log-scale. The grey shaded regions show the parts of this plot that are expanded, using a linear scale, in panels (c) and (d). In panels (c) and (d), the crosses shows the wave peaks that have been used to calculate the angular frequency.

where $f = 2\Omega_S \cos\theta$ is the Coriolis parameter. It is easily shown that the solution to this equation consists of a constant speed $v = -F/f$ plus a sinusoidal component with angular velocity $\omega = f$. This means that a natural oscillation frequency of winds in geostrophic balance is given by the Coriolis parameter, which at high latitudes $\sim 2\Omega_S$. Thus it seems reasonable to suggest that the second set of waves, which appear throughout the model, may be generated by oscillations of the high-latitude thermospheric winds at their natural frequency. We cannot yet offer a similar simple explanation of the frequency of the first set of waves.

These considerations indicate a rich variety of wave behaviour in the thermosphere. Some of the observed wave behaviour may represent the thermospheric equivalent of Rossby waves in the deeper atmosphere. It is beyond the scope of this paper to investigate this further, although a future study could usefully assess the full range of possible global-scale wave modes in this unique environment.

In the present context, the discussion of waves is significant because the Southern hemisphere oscillation clearly develops from the very small axial asymmetries that have propagated as waves within the first 10 planetary rotations. The initial structures at 80°S that begin to grow at around 10 rotations emerge just a few rotations after the first waves arrived. Since the Southern hemisphere axial asymmetries are already growing at ∼10 rotations, the asymmetry must have been transmitted by wave action, because to transmit it hydrodynamically from pole to pole on this time-scale would require meridional wind speeds of the order of $\sim 500\,\mathrm{m\,s^{-1}}$, which is far greater than the meridional wind speeds that are observed. For example, the highest meridional winds shown in Fig. 5 in the northern polar region are only $\sim 100\,\mathrm{m\,s^{-1}}$, and the speeds at mid-latitudes are much smaller in magnitude. It is therefore clear that only very low amplitude wave structures are required to trigger the feedback effect in the Southern hemisphere. This indicates that, provided the feedback mechanism itself is of sufficient

magnitude, it takes only a very small initial perturbation to develop a structure with substantial axial asymmetry. In other words, the axial symmetry of the thermosphere is not a naturally stable state, and is easily broken in the particular scenario that we have modelled.

## 5 ROTATION RATES

In order to compare the behaviour of our atmospheric axial asymmetries with the observations, we need to calculate their overall rotation rates about the axis of the planet. As already discussed, we are not interested in the absolute rotation rate of an asymmetry but in its rotation rate relative to the lower boundary of the model. Gurnett et al. (2010) found that the SKR rotation rates were all slower than independent measurements of the internal rotation rate of $\sim 820\,^{\circ}\mathrm{d}^{-1}$, and ranged from approximately this value in the winter hemisphere to $\sim 800\,^{\circ}\mathrm{d}^{-1}$ in the summer hemisphere. The important features here are the near equality of the winter rotation rate and the internal rotation rate and the difference of $\sim 20\,^{\circ}\mathrm{d}^{-1}$ between the rotation rates of the two hemispheres at solstice.

To analyse the rotation speed of the pseudo-vortex, we need a simple way of estimating its longitude at each time-step. To achieve this, we select as a proxy for the location of the pseudo-vortex the longitude of the maximum in the current convergence at our usual diagnostic latitude of 74°. The maximum in the current convergence is selected because it consistently exhibits a clear peak which, once the pseudo-vortex is well established, must be correlated with its location. We calculate the average angular velocity of the pseudo-vortex for each planetary rotation by comparing the longitude of the current convergence peak at the beginning and end of each rotation. It is difficult to do this accurately from the raw model output, because the 10° longitude resolution of the model places a 10° limit on the accuracy with which we can directly locate the peak. To overcome this, we find the location of the maximum from the raw model output and then perform a spline fit to this point and the two grid points either side using the SPLINE function in IDL. The maximum of this spline fit is then used to establish the location of the maximum current convergence. This produces a much more accurate location for the pseudo-vortex that allows a more precise estimate of its rotation rate.

### 5.1 Equinox model

Fig. 12(a) shows the angular velocities so calculated for the 400 rotation equinox run (E1) described above. The top panel shows the first 40 rotations on a large scale, showing large fluctuations in the first 20 rotations as the axial asymmetries in the Northern and Southern hemispheres establish themselves. After 40 rotations, both asymmetries have settled down to closely similar angular velocity, both of which are below the angular velocity of the base of the model. This is consistent with the presence of the westward drifts described above. The results of Gurnett et al. (2010) also show the SKR rates as consistently lower than the independently measured internal rotation rates of the planet (Anderson & Schubert 2007; Read et al. 2009).

Fig. 12(b) shows the angular velocities across the whole 400 rotation run. This shows transient oscillations with a ∼15 rotation period continuing past rotation 200, with a decay time-scale of ∼50 rotations. Oscillations with a ∼15 rotation period are also present in Fig. 11: the fact that we observe oscillations with this period in two independent analyses of the model output suggests that they are a genuine feature of the thermospheric response, rather than an aliasing effect derived from the processing of the data. It seems likely that the long thermal time-scales of the thermosphere have a role to play in generating these long-period oscillations, but it is beyond the scope of this study to theorize further about their precise cause. Although we would not expect oscillations with this exact period to be present in Saturn's thermosphere, this does raise the possibility of secondary variations in the SKR period superimposed on the seasonal variation that has already been described.

Towards the end of the 400 rotation run, the northern and southern axial asymmetries have settled down to very closely similar angular velocities. Even after 400 rotations, there remains a residual jitter ($\sim 0.2\,^{\circ}\mathrm{d}^{-1}$) in the calculated angular velocities. This appears to be an aliasing effect due to the 10° resolution of the model, which has not been completely eliminated by our spline fitting procedure.

We have considered a situation in which the thermosphere is forced with axially asymmetric heating for one planetary rotation period, and then left alone to develop in isolation. In reality, we would expect there to be continual axially asymmetric forcing of the thermosphere due, for example, to varying activity in the magnetosphere. This may have the effect of continually stimulating behaviour such as the 15 rotation oscillation, such that it never dies away completely. If the thermosphere is responsible for the periodicities in the magnetosphere, we may thus expect to find secondary, longer period signals superimposed on the strong ∼10 h signals in the magnetosphere.

### 5.2 Solstice model

Fig. 13 shows the same analysis as Fig. 12 but for the solstice model S1 (see Table 1). Exactly the same procedure has been applied to initiate an asymmetry in the Northern hemisphere. We recall that the solstice model increases the relative heating and conductance in the Southern hemisphere, thus broadly simulating conditions immediately before the recent equinox in 2009. Once again there is considerable variation in the rotation rate to begin with, followed by a clear sequence of ∼15 rotation period oscillations which die away on a ∼100 rotation time-scale.

After the rotation rates have settled down there is a clear north–south asymmetry. The form of this asymmetry qualitatively matches that observed by Gurnett et al. (2010) in the SKR rates in that the northern rate is higher than the southern rate, and both lag the rotation of the lower layers of the atmosphere. The quantitative match is less good: the difference between the two rates is $\sim 7^{\circ}\mathrm{d}^{-1}$, compared to a maximum value of $\sim 20^{\circ}\mathrm{d}^{-1}$ observed by Gurnett et al. (2010), and the lag of the northern rate compared to the lower boundary rotation rate of $\sim 810^{\circ}\mathrm{d}^{-1}$ is $\sim 60^{\circ}\mathrm{d}^{-1}$, much greater than the near corotation observed by Gurnett et al. (2010). Thus the quantitative details of the calculated rotation rates do not match the observations well, but the qualitative behaviour and north–south 'ordering' of our predictions are perfectly in line with what is observed. The latter point is more significant, given that our model of solstice conditions crudely applies an arbitrary 10 per cent north–south bias: the actual difference between Northern and Southern hemispheres could be much more or less than this, altering our results quantitatively but probably not qualitatively.

To investigate the cause of the seasonal variation, we have repeated run S1 twice: once with the north–south bias applied only in the heating, and once with it applied only in the conductance. We find that the overwhelming majority of the seasonal variation is due to the variation in the conductance. As shown in Section 4, the rotation rate of the asymmetry is a combination of advection

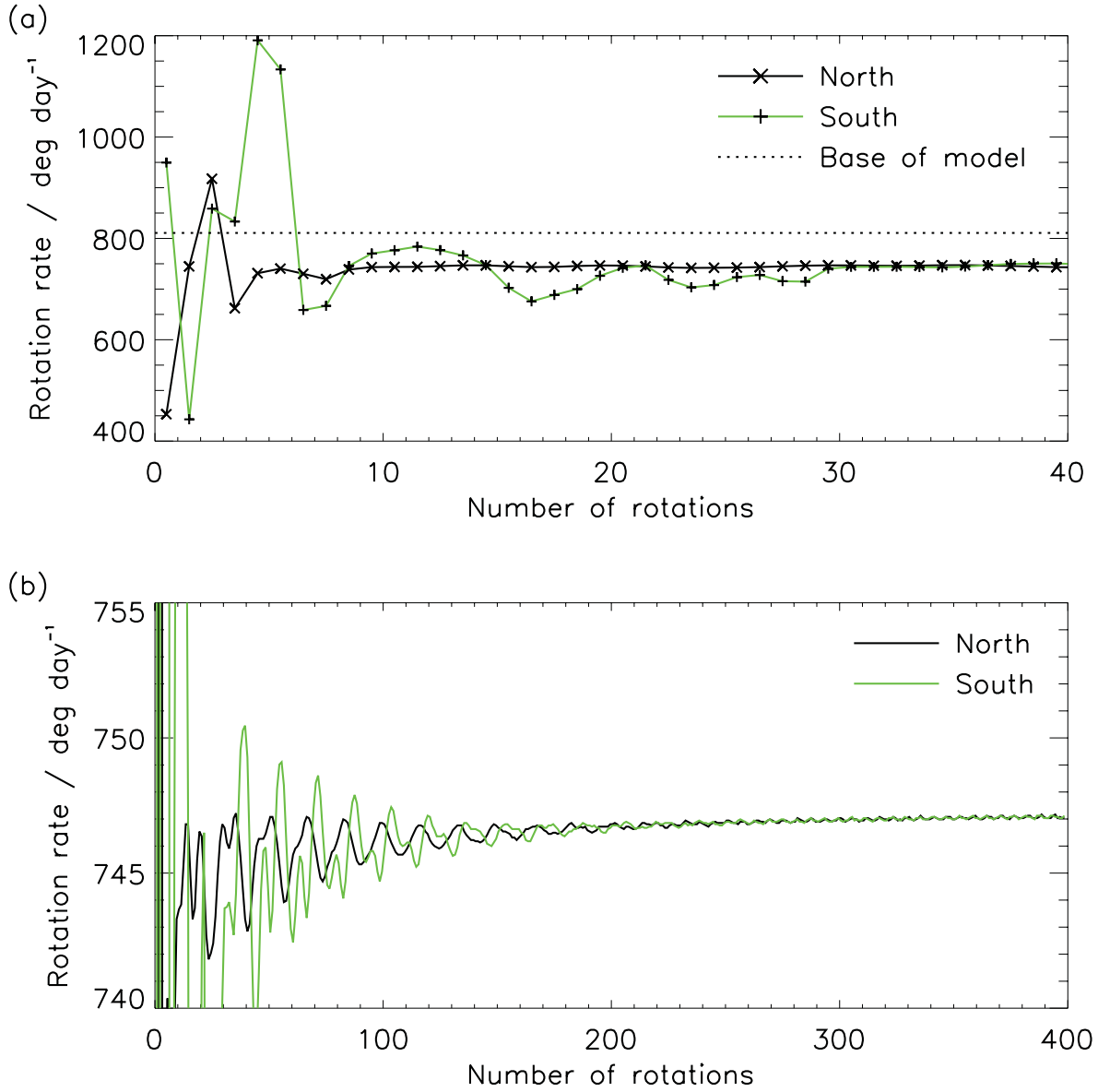


**Figure 12.** Evolution of northern and southern rotation rates in model run E1. Panel (a) shows the first 40 rotations of the model run. The black and green solid lines and crosses show the rotation rates of the northern and southern axial asymmetries, respectively. The crosses are placed at the mid-point of the time period that they represent. The horizontal dotted line shows the rotation rate of the lower boundary of the model (Desch & Kaiser 1981). Panel (b) shows the full 400 rotations of the run, using the same line formats but omitting the crosses.

drift and heating drift. The small effect of changing the heating is probably because the seasonal change is in the axially symmetric component of the heating which mostly affects the axially symmetric component of the zonal winds – however, the axially symmetric component of the zonal winds at high latitudes is strongly controlled by ion drag from sub-corotating plasma, such that small changes in heating have a relatively small effect on the overall zonal wind field. Thus, the seasonal change in heating will have only a small effect on the size of the advection drift. In contrast, the large effect of changing only the conductance is probably a signature of the heating drift becoming correspondingly more or less important, since the size of the axially asymmetric heating $\mathcal{Q}_C$ is correlated with the currents and therefore with the conductance. The sense of the change is as we would expect: a low conductance in winter leads to less heating drift such that the overall rotation rate becomes faster in winter.

### 5.3 Seasonal variation

Our model is thus able to qualitatively explain one puzzling aspect of the observed rotation rates, specifically the north–south difference and its seasonal dependence. However, a further aspect of this seasonal dependence is the apparent delay in the response of the SKR rotation rates to the season. In the case of the SKR rotation rates, the northern and southern rates have been observed to equalize ∼7 months after equinox (Gurnett et al. 2010); in the case of the magnetic field perturbations, at the time of writing, the northern and southern rates have not yet been observed to equalize, over ∼20 months after equinox. A period of 1 month ∼75 planetary rotations, so to explain the ∼7 month delay alone would require a delay of ∼500 planetary rotations. This seems highly unlikely given that 400 rotations is a sufficient time-scale for the rotation rates in each hemisphere to stabilize. Nevertheless, we have tested the variation of the

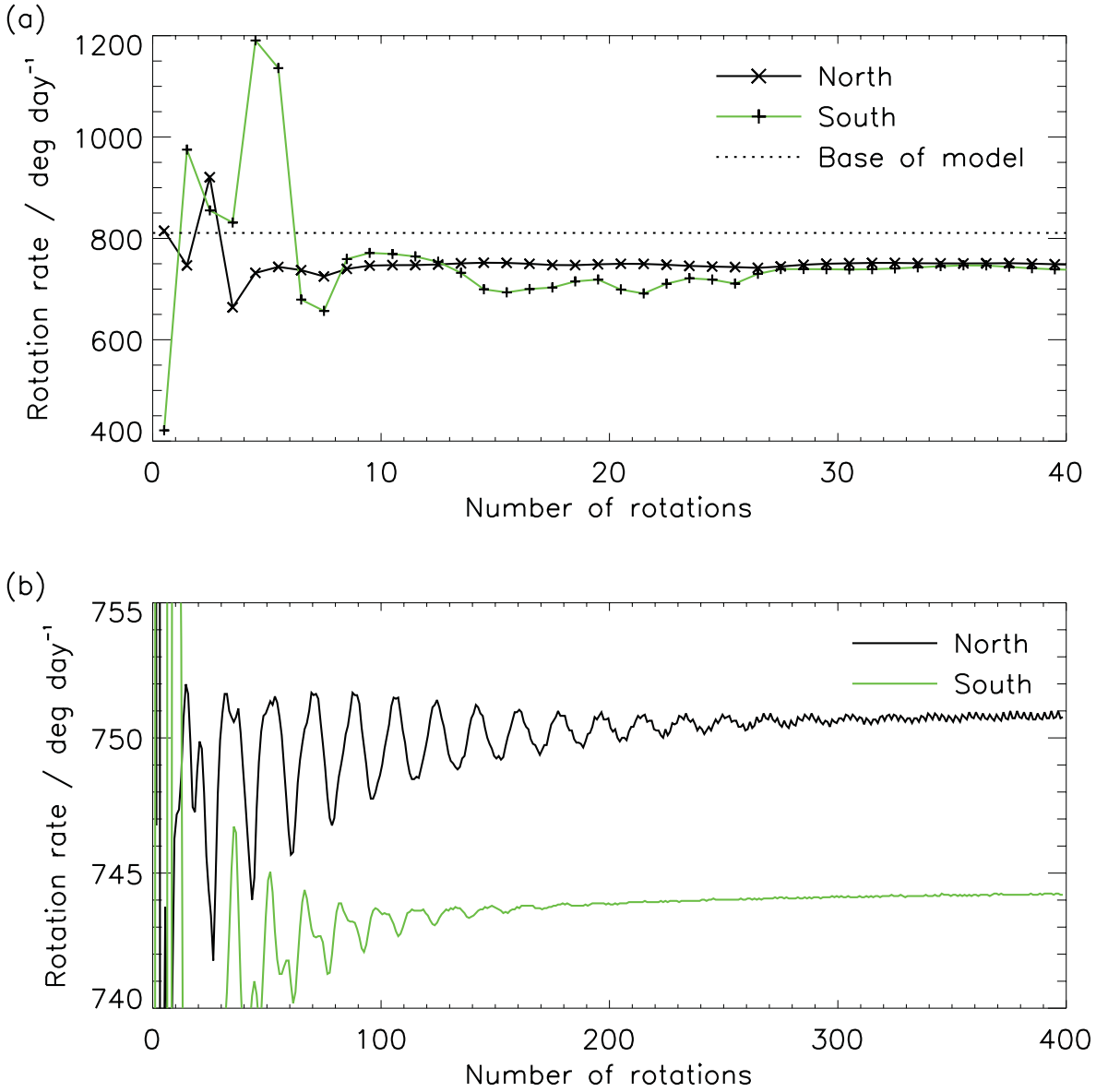


**Figure 13.** Evolution of northern and southern rotation rates in model run S1, in the same format as Fig. 12.

rotation rates with varying season. To achieve this, we sinusoidally vary the bias factor $S(t)$ in equation 3, setting $S(t) = S_0 \cos \omega t$ where $T = 2\pi/\omega$ is the time for one whole seasonal cycle. To accurately model the real situation, we should set $T = 29.5$ yr to match Saturn's orbital period. However, in practice this represents ∼24 000 planetary rotations, which implies unmanageably long model run times. To allow practical simulations, we set $T = 800$ planetary rotations, equivalent to ∼1 terrestrial year. This allows the practical simulation of one whole seasonal cycle.

We implement this experiment with an 800 rotation run following directly on from the solstice run S1 already shown in Fig. 13. For comparison, we also implement an 800 rotation run with fixed equinox conditions (see Table 1). The results of these runs, which we label S2 and E2, are shown in Fig. 14. Fig. 14(a) is equivalent to Fig. 13(b). Fig. 14(b) is an enlargement of the section marked with the broken rectangle. The blue solid curve shows a perfect sinusoid exactly in phase with $S(t)$, with the maximum and minimum values scaled by eye to match the output of S2 as closely as possible. The vertical blue solid lines mark the equinoxes. It is clear that the rotation rate of the pseudo-vortex does vary roughly sinusoidally, almost exactly in step with $S(t)$. However, there are two clear deviations from the perfect sinusoid that are most clearly visible in Fig. 14(b). First, the rotation rate at which the northern and southern rates cross is slower than the rotation rate at which the perfect sinusoids cross. This is unsurprising since we would not expect the variation of rotation rate to vary exactly linearly with $S(t)$, and therefore we would not expect the variation of rotation rate to be symmetric about the equinox rotation rate. However, the two rates do cross at a slightly greater rotation rate than the equinox rotation rate, which is harder to explain. Secondly, and more interestingly, the crossing occurs ∼5 rotations *before* equinox, which is the opposite of what is observed.

To investigate whether these two deviations are specific to the particular value of $T = 800$ rotations, we have conducted a series of runs with $T$ varying in the range 100–6400 rotations, with each run continued long enough to reach the first equinox crossing. The

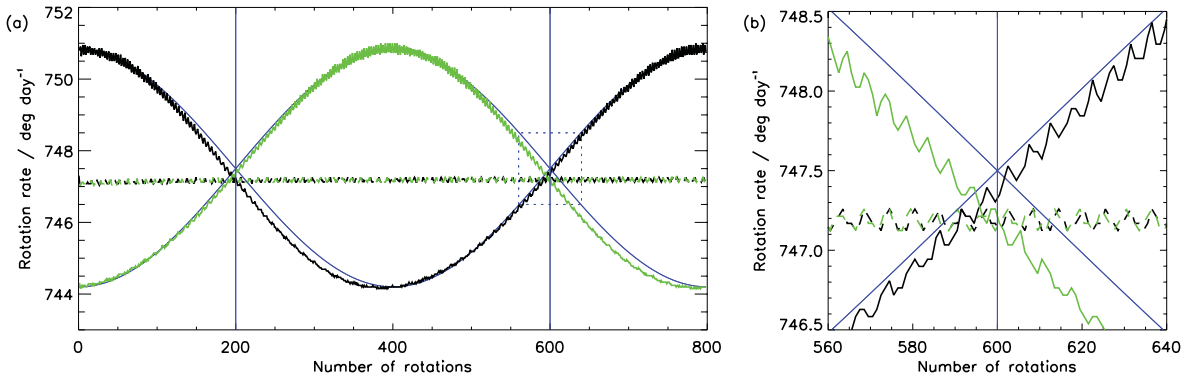


**Figure 14.** Evolution of northern and southern rotation rates due to seasonal variation. The solid black and green lines in panel (a) show the variation of the rotation rates of the northern and southern axial asymmetries in run S2. The dashed lines show the equivalent information for run E2. The blue solid curves are perfect sinusoids matched to the maximum and minimum values of the rotation rates in run S2. The vertical blue lines mark the location of the equinoxes. The broken rectangle marks the section shown at a larger scale in panel (b).

general trend in the value at which the north and south rates cross is that for longer values of $T$ it moves closer to the equinox rate. This indicates that the difference is probably a small non-equilibrium effect, and that for the very long value of $T = 29.5$ yr that applies to the real planet, this anomaly would not be present. The second anomaly also shows clear trends. First, the rates always cross before equinox. Secondly, the number of rotations by which the crossing leads equinox increases as $T$ increases from 100 to 400 rotations, but for values of $T$ greater than 400 rotations the lead time levels off, with some variation due to the residual jitter in the calculated values of the rotation rate, at $\sim 5 \pm 1$ rotations. Thus the crossing occurs with an approximately constant lead time of five rotations for large values of $T$.

A possible explanation for this behaviour lies in the way that we calculate the rotation rate. For this purpose we have used an essentially arbitrary indicator of the location of the pseudo-vortex – specifically, the location of the peak in the current convergence at a particular latitude and altitude. There is no guarantee that this – or any other chosen indicator – is representative of the rotation rate of the structure as a whole. However, suppose that it were possible to find an indicator, whose location we label $\alpha$, that did represent the overall rotation rate. If the structure is in a slowly evolving near-equilibrium state, then, assuming for simplicity a linear response to the seasonal variation, we would expect the rotation rate $\Omega_\alpha$ of this location to be proportional to $S(t)$ and thus to vary approximately as $\cos \omega t$:

$$\Omega_\alpha = \Omega_\alpha^{\rm eq} + \Delta\Omega_\alpha \cos \omega t, \tag{6}$$

where $\Omega_\alpha^{\rm eq}$ is the rotation rate of $\alpha$ at equinox and $\Delta\Omega_\alpha$ is the amplitude of its seasonal variation. Our chosen indicator – the location of the maximum in the current convergence at a specific latitude and altitude – whose location we label $\beta$ will lie at some longitudinal separation $\phi_{\alpha\beta}$ from $\alpha$. We might expect this distance $\phi_{\alpha\beta}$ to vary with $S(t)$ as the size and structure of the pseudo-vortex vary seasonally, and it seems reasonable to assume that to a first approximation this might also vary in proportion to $S(t)$, so that

$$\phi_{\alpha\beta} = \phi_{\alpha\beta}^{\rm eq} + \Delta\phi_{\alpha\beta} \cos \omega t. \tag{7}$$

We would then expect the rotation rate of $\beta$ to be a combination of $\Omega_\alpha$ and the time variation of $\phi_{\alpha\beta}$:

$$\Omega_\beta = \Omega_\alpha + \frac{{\rm d}\phi_{\alpha\beta}}{{\rm d}t}. \tag{8}$$

If the contribution from this second term is small, then we can combine these three equations to give

$$\Omega_\beta \simeq \Omega_\alpha^{\rm eq} + \Delta\Omega_\alpha \cos \omega(t + \tau), \tag{9}$$

where $\tau$ is a small lead time given by

$$\tau = \frac{\Delta\phi_{\alpha\beta}}{\Delta\Omega_\alpha}. \tag{10}$$

This expression is completely independent of the time period of the seasonal variation $T$. Thus this origin for a phase lead is consistent with the observation that for values of $T$ greater than 400 rotations it tends towards an approximately constant value of five rotations $\sim$2 d. The smaller phase lead for values of $T$ less than 400 is probably because the variation over these time-scales is sufficiently rapid that non-equilibrium effects are important. The value of $\Delta\phi_{\alpha\beta}$ consistent with our lead time of 2 days and our seasonal variation of $\sim 3.5\,^\circ{\rm d}^{-1}$ is $\Delta\phi_{\alpha\beta} \sim 7^\circ$. Given that our longitudinal grid spacing of $10^\circ$ is of a similar order of magnitude, this model is a plausible explanation of the modelled lead time (although, of course, it is difficult to be certain that it is the correct explanation without increasing the longitudinal resolution of the model and investigating these effects in more detail).

There is nothing in this analysis that states whether $\beta$ should lag or lead $\alpha$. Indeed, it is possible that this model could also explain the observed $\sim$7 month time lag in the seasonal variation of the SKR rotation rate, if we suppose that the feature that triggers SKR emission is not the central feature $\alpha$ but a secondary feature $\beta$ that drifts seasonally relative to $\alpha$. In this case we have a lead time $\tau \sim 7$ months $\sim 200$ days and an amplitude of variation in the rotation rate of $\Delta\Omega_\alpha \sim 10\,^\circ{\rm d}^{-1}$, giving a required $\Delta\phi_{\alpha\beta} \sim 2000^\circ$ of longitude. This is clearly an absurdly large angular size for a structure in the thermosphere, since there are only $360^\circ$ of longitude available for a structure to occupy. It thus seems unlikely that the effect described can be responsible for the observed 7 month phase lag in the SKR rotation rate.

## 6 SENSITIVITY TO MAGNITUDE OF FEEDBACK EFFECT

There are many aspects of the model that we could alter or improve to investigate their influence on the feedback effect between the magnetosphere and thermosphere. The most obvious alteration would be to substantially improve our physical modelling. In Section 7 we will discuss the four possible improvements that we

believe to be most significant: full ionosphere calculations, improved spatial resolution, an improved precipitation heating model and the addition of radiative cooling.

Within the existing model, we can also vary a number of background parameters such as the latitudinal distribution of conductance, or the plasma sub-corotation speed, which are broadly fixed by the observations but most likely vary somewhat from our assumed values. Preliminary investigations indicate that this sort of modification does not have a significant qualitative effect on the predictions of the model: for most combinations of parameters that we have tried, it is possible to generate asymmetric structures provided that the parameter $j_{\parallel 0}$ – which determines the magnitude of the feedback effect – is chosen appropriately. It is thus clear that this parameter is crucially important, and in the following section we will examine the sensitivity of our results to its value.

Because we have set $q_{C0} = q_{B0}$, $j_{\parallel 0}$ physically measures the field-aligned current density required to either double (when the field-aligned current is upwards) or reduce to zero (when the field-aligned current is downwards) the high-latitude heating due to precipitation. A high value of $j_{\parallel 0}$ thus represents a weaker feedback effect. Our choice of $j_{\parallel 0} = 0.02\,\mathrm{nA\,m^{-2}}$ is essentially arbitrary and has been chosen because it is small enough to make the feedback effect drive self-sustaining wind systems, but large enough not to produce unphysically large thermospheric temperatures. It is thus interesting to analyse how the feedback effect changes when we modify $j_{\parallel 0}$. The physical plausibility of the value $j_{\parallel 0} = 0.02\,\mathrm{nA\,m^{-2}}$ will be discussed further in Section 7.

To investigate different values of $j_{\parallel 0}$ we have repeated run E1 six times, labelling these runs E1$x$. The values of $j_{\parallel 0}$ used for each run are listed in Table 1. Fig. 15 then shows the results of these runs in terms of the development and evolution of axial asymmetries. These plots show the same asymmetry parameter as shown for the equator in Fig. 11(b), but for 74°N (black line) and 74°S (green line). Note that the range of times shown is different for each panel of Fig. 15, showing only as many rotations as are required for the axial asymmetries to either stabilize [panels (a)–(d)] or show a clear declining trend [panels (e) and (f)].

Looking first at panel (b), this shows the behaviour of run E1b, which is identical to our standard run E1. This shows the behaviour that has already been extensively discussed: the Northern hemisphere asymmetry settles quickly; the Southern hemisphere asymmetry begins with low level wave behaviour and then grows to the same magnitude as the northern asymmetry within ∼40 rotations; both hemispheres show ∼15 rotation oscillations. Panel (a) shows the effect of making $j_{\parallel 0}$ marginally smaller (E1a), increasing the feedback effect slightly. The axial asymmetries in both hemispheres stabilize much more quickly and with a temperature asymmetry greater than 100 K. Global temperatures are also much higher, exhibiting exosphere temperatures of ∼600 K at the equator and >1000 K at high latitudes. These high temperatures are inconsistent with observed temperatures of ∼400 K (Smith et al. 1983). Both the relative increase in gross temperatures and the temperature asymmetry are much greater than the corresponding relative change in $j_{\parallel 0}$. At these high temperatures we might expect factors not included in our model to moderate the feedback. For example, infrared radiative cooling due to the $H_3^+$ molecular ion would become more important at these higher temperatures.

Panel (c) shows the results of E1c, for which the feedback effect is slightly reduced compared to our standard value. This takes longer to stabilize, but by ∼150 rotations both hemispheres have stable axial asymmetries with a magnitude of ∼1 K. However, this run is interesting in that the Southern hemisphere shows variation in the size of the asymmetry about this value. Inspection of the Southern hemisphere shows that there are two axial asymmetries present. One is more or less identical to that in the Northern hemisphere, peaking at ∼78°S, and accounts for the majority of the axial asymmetries in the temperature at 74°S. The second consists of two similar anomalies at diametrically opposite longitudes – i.e. approximate $m = 2$ symmetry – which travel eastwards rather than westwards and are most significant at ∼66°S. The temperature asymmetries for these two separate structures are shown in Fig. 16 for three times towards the end of the model run when the asymmetries are reasonably stable. Both the development of a clear $m = 2$ asymmetry and the opposite direction of drift of this anomaly are clear.

To analyse the reasons for this opposite drift, Fig. 17 shows the same information as Fig. 9, but for 66° S at the end of run E1c. The same overall behaviour is observed in Fig. 17(a): large peaks of external heating that almost coincide with each temperature peak and are largely balanced by vertical advection, and similarly large peaks of heating due to zonal advection. The other terms are relatively unimportant. The principal difference is that the zonal advection peaks lie east of the temperature peaks. This is because at these latitudes there is a weak westward flow, as shown in Fig. 5 for runs E0 and E1. Fig. 17(b), which shows the combined terms, illustrates the effect of this. There is a heating peak to the east of the temperature peak (driven by the external heating $\mathcal{Q}$) and a

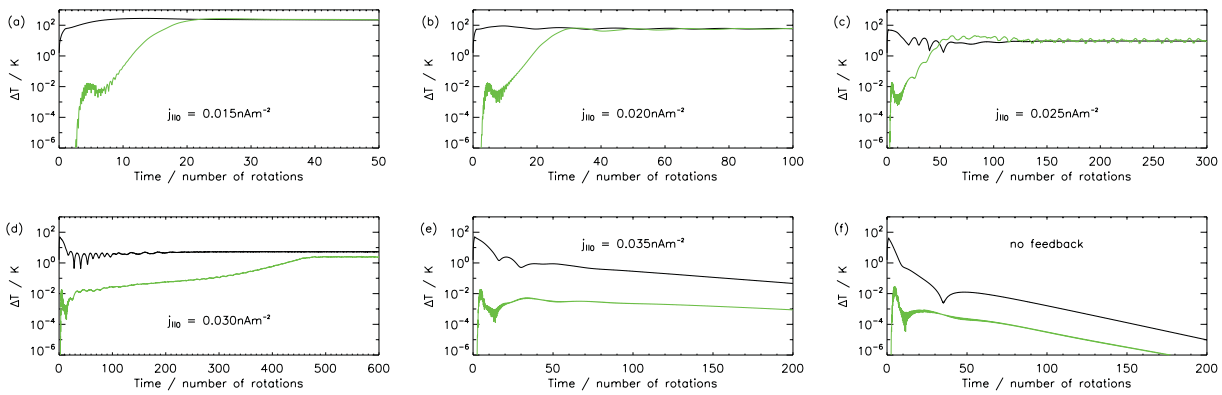


**Figure 15.** Sensitivity of the development of axial asymmetries to the magnitude of the feedback parameter $j_{\parallel 0}$. Panels (a)–(f) show output from model runs E1a–E1f (see Table 1). Each plot is equivalent to Fig. 11(b), which shows axial asymmetries at the equator in run E1, but instead shows axial asymmetries at 74°N (black lines) and 74°S (green lines). Note that the number of rotations shown are not the same for all of the panels (a)–(f), reflecting the different times taken for the rotation rates of the axial asymmetries in each run to stabilize.

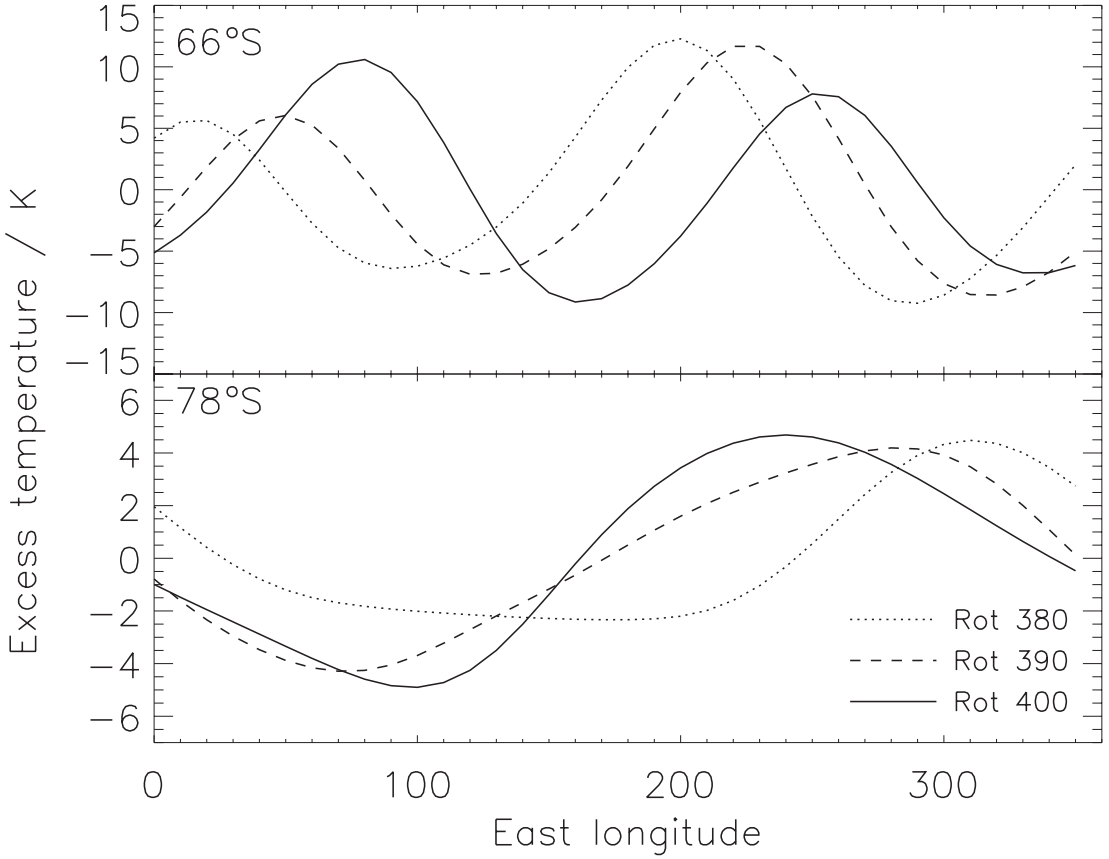


**Figure 16.** Axially asymmetric temperatures at 66°S and 78°S in model run E1c. Dotted, dashed and solid lines show the conditions after 380, 390 and 400 rotations, respectively.

heating peak to the west of the temperature peak (driven by zonal advection). When combined (solid line) the peak lies just to the east of the temperature peak. Thus, the overall effect is eastward drift of the $m = 2$ asymmetry. We therefore have a situation which contrasts with that discussed earlier in Section 4 with reference to Fig. 9: in this case, the two drift mechanisms oppose each other, and the more significant of the two (advection drift) determines the direction of drift of the anomaly.

Returning now to Fig. 15, panel (d) shows the results of E1d for which the feedback is further reduced compared to E1c. This takes even longer (∼500 rotations) to stabilize and does so with different axial temperature asymmetries in the north and south. In this case, the northern temperature asymmetry consists of the usual westward drifting $m = 1$ asymmetry and the southern temperature asymmetry consists of the eastward drifting $m = 2$ asymmetry present in E1c. It seems likely that these different combinations of axial asymmetries at different latitudes are simply contingent on the initial conditions of each run and their subsequent evolution: there seem to be a range of possible self-sustaining axial asymmetries which may or may not be excited within the model. It thus seems reasonable to suppose that higher order axial asymmetries may be possible, but that the relatively low-longitude resolution of our model does not allow those of high azimuthal wavenumber to develop.

The final panels of Fig. 15, (e) and (f), show the situation with a further reduced feedback effect (Ele) and with no feedback whatsoever (E1f). Both these runs show a clear declining trend in the axial asymmetries present in the thermosphere, although the presence of weak feedback in E1e does slow this trend. Significantly, the difference between the value of $j_{\|0}$ in E1a (which shows unphysical overheating of the thermosphere by the feedback effect) and E1e (which shows a failure by the feedback effect to produce a permanent asymmetry) is only just greater than a factor of 2. There is thus a relatively small window of values within which this feedback model can produce axially asymmetric structures consistent with the observed temperatures.

## 7 DISCUSSION

### 7.1 Plausibility of feedback effect

The basic feedback model discussed in Section 2 seems reasonably well founded. However, there are two major flaws with the applicability of our model results to the actual situation observed at Saturn: first, the energies of the precipitating electrons required to produce the calculated heating and secondly the observational consequences of this precipitation.

The first problem becomes clear if we examine the expression for calculating $\mathcal{Q}_C$, the external heating due to the feedback effect (equation 2). Height-integrating the expression for the heating per unit mass $q_C$ with respect to the column mass per unit area $\sigma$ gives an expression for the column-integrated heating per unit area $Q_C$:

$$Q_C = \int q_C \mathrm{d}\sigma = \frac{q_{C0}}{g} \int \exp\left(-\frac{p}{p_C}\right) \mathrm{d}p = \frac{p_C q_{C0}}{g}, \qquad (11)$$

where we have used the identity $\mathrm{d}p = g\mathrm{d}\sigma$ and $g \sim 10\,\mathrm{ms}^{-2}$ is the effective acceleration due to gravity. As described in Section 3, we used values of $p_C = 4\,\mathrm{nb}$ and $q_{C0} = 2.5\,\mathrm{W\,kg}^{-1}$. These parameters give $Q_C \sim 0.1\,\mathrm{mW\,m}^{-2}$. This is scaled using the parameter

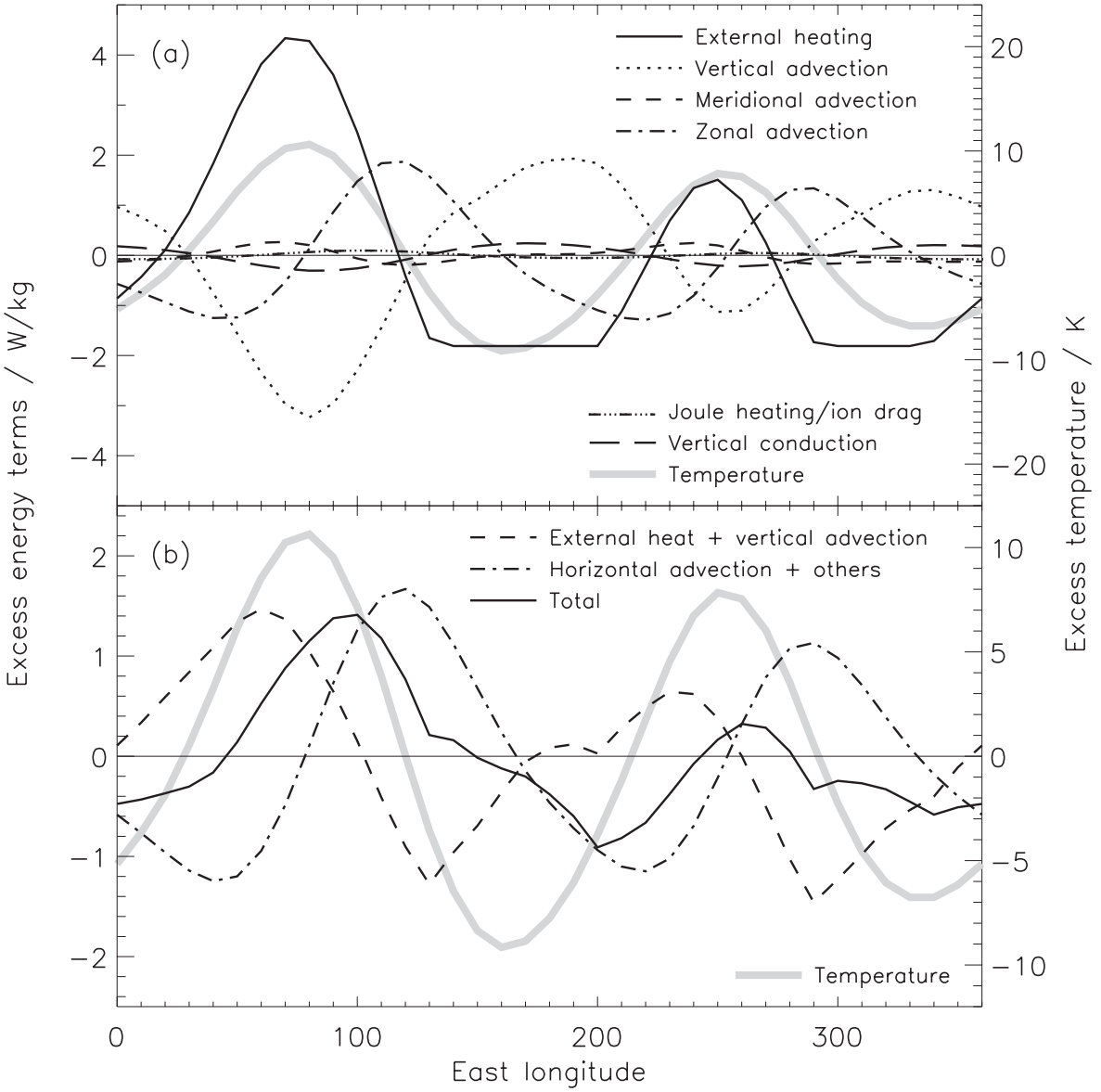


**Figure 17.** The same information as Fig. 9, but for 66°S at the end of model run E1c. The line formats in both panels are the same as those in Fig. 9.

$j_{\parallel 0} = 0.02\,\mathrm{nA\,m^{-2}}$. The ratio of these two values is ∼5 MV. This implies that ∼5 MeV electrons would be required to deliver the energy required to sustain the asymmetry – and if the heating efficiency of the precipitating electrons is less than 100 per cent, the required energies would be even higher. Typical values of precipitating electron energies estimated for the auroral regions of Saturn are ∼10 keV (Gérard et al. 2004). Thus, the electron energies required are ∼500 times greater than those that are observed. Furthermore, we require these precipitating electron energies distributed across the polar cap, whereas in reality the largest electron energies are confined to the main auroral oval during normal conditions.

The second problem is related, in that the distribution of particle precipitation [as implied by the heating distribution, e.g. Fig. 8(a)] is highly axially asymmetric. The heating distribution calculated by the model implies a greatly enhanced particle precipitation in the region of the pseudo-vortex, and almost zero at longitudes opposite to the pseudo-vortex. Even taking into consideration the model sketched in Fig. 2, in which the thermospheric winds increase or decrease precipitation by modifying pre-existing small scale patterns of precipitation, a large-scale asymmetry of this nature is not observed. However, it is worth noting that the 'filled in' polar cap (Clarke et al. 2005, fig. 1j) observed in some images is somewhat similar to the distribution that we calculate. It is possible that if an asymmetry similar to the one described in this paper does exist in the polar regions – although powered by lower energy electrons or by some other mechanism – it could induce large-scale axial asymmetries in electron precipitation during a period in which the conditions in the magnetosphere temporarily changed. Nevertheless, our principal observation is that the large-scale asymmetry in electron precipitation predicted by our model is not consistent with the observations.

Given these two major caveats to the work presented above, it is natural to question the value of our conclusions. However, our results are not physically impossible or self-contradictory – they just do not match the conditions that are observed at Saturn. Thus, while we cannot draw specific conclusions from our results, they

do establish several ideas in principle. In particular, they represent a 'proof of concept' of the idea of a persistent thermospheric asymmetry first proposed by Smith (2006). Overall, therefore, we believe that our results are worthwhile as a stepping stone to further work on this topic, but due to the inconsistency with the observations they should be considered to be highly provisional.

### 7.2 Omissions from model

One reason that our feedback model requires higher than observed precipitating electron energies may be that the numerical model that we have used contains many omissions and simplifications. It is possible that adding physics to the model, or resolving existing physics in greater detail, may reduce the energies required to power an axially asymmetric structure in the thermosphere. It is challenging to speculate on the extent to which such a more complete model would affect our results. There is effectively no limit to the extent to which the resolution and complexity of our model can be improved, only the practical limit of computation time. The most detailed available thermospheric General Circulation Models (GCMs) for gas giant planets are both for Jupiter (Achilleos et al. 1998; Bougher et al. 2005), and while both include more physics than the model discussed here, neither has a substantially higher spatial resolution. Replacing our model with a modified version of one of these GCMs would likely have a negligible effect on our results if the same parameters for the feedback process were used, because the feedback is mostly driven by a simple interaction between thermospheric dynamics and heating, but the added complexity (in aspects such as local time variation of the ionospheric conductance) would make it much more difficult to isolate and analyse any asymmetric structures.

Ideally, rather than our approach of *imagining* a feedback effect and then deliberately modelling the relevant processes, it would be preferable to employ a model with substantially higher resolution and more complete physics, and then hope to observe asymmetric structures evolve naturally, via unexpected feedback processes. Such an approach would require a vast range of physical processes and length scales to be modelled, to ensure that any unexpected feedback process could function, a requirement that seems unlikely to be met, in the immediate future at least, by the current generation of thermospheric GCMs.

Although speculating on the predictions of such a full model is beyond our scope, we will in the following sections discuss four of the most significant omissions and simplifications involved in the model, all of which could practically be improved upon, in the short term, without imposing on the model an impractically onerous computational burden.

#### *7.2.1 Ionization*

A key omission from the model is a self-consistent calculation of the electron and ion densities in the ionosphere. In one sense, it is an advantage not to carry out such calculations in that it avoids overcomplicating the model. However, we can envisage several interesting effects that more detailed ionization calculations may allow us to study.

For example, the effect of ion drifts associated with Pedersen and Hall currents may structure the ionosphere, providing the potential for further feedback effects. In the thermosphere, the overwhelming contribution to horizontal currents is motion of ions due to ion-neutral collisions. Thus, in our idealized vortex (Fig. 1) the ions should drift towards its centre. This may tend to enhance the ion density moving towards the centre of the vortex, leading to a corresponding increase in the current convergence and field-aligned current. The precipitation expected in the centre of the vortex should also increase the ion density there, further enhancing the feedback. Structuring of the ionosphere in this way is one possible mechanism for providing a functioning feedback effect with a lower energy input from precipitating electrons. However, the enhanced ion densities may also have the opposite effect by increasing ion drag and thus reducing the circulation speed of the vortex.

The influence of full ionosphere calculations is thus complex and requires detailed modelling. In preliminary experiments with simple one-dimensional models of ionization feedbacks, we have found that a gradient in the ion density sufficiently steep to induce precipitation initiates a feedback effect that tends to steepen that gradient further. This ultimately produces a shock-like structure that is difficult to resolve in a general circulation model. Modelling possible feedback effects due to ionization may therefore require a different approach.

#### *7.2.2 Model resolution*

We are concerned with the interaction between the thermosphere and the magnetosphere, and ideally our model should be able to resolve structures in both. In terms of the former, our model resolution is similar to that of other thermospheric GCMs for gas giants (e.g. Achilleos et al. 1998), and should perform equally as well in resolving the overall structure of the thermosphere. However, the most important aspect of the work presented in this paper is the generation of field-aligned currents by gradients in thermospheric winds. Clearly a limited resolution limits the size of any possible gradients, which may limit the extent of any feedback effects.

Probably more significant than this is the difficulty in resolving structures in the magnetosphere. Due to the geometry of the magnetic field, very large structures in the magnetosphere map to small structures in the thermosphere. The main auroral oval is an excellent illustration of this point: given that this region experiences the greatest intensity of particle precipitation it is almost certainly crucial to the energy balance and ionospheric structure of the high latitudes, and the presence of a planetary-period oscillation in the location of the oval (Nichols et al. 2010b) suggests it may play an important role in understanding the periodicities. However, the $2^\circ$ latitude resolution of our model does not resolve this $\sim 1^\circ$ wide structure and we are thus unable to draw any valid conclusions about how the main oval may influence or be influenced by feedback effects.

#### *7.2.3 Precipitation heating model*

Our heating model is deliberately very simple, although the essential structure – a peak with a flat bottom-side and exponential topside – is a good representation of the overall heating profiles calculated by full electron transport models (e.g. Galand et al. 2011). An interesting possibility to consider is that there could be secondary feedback effects that 'tune' the energy of the precipitating particles to produce heating at a particular altitude. If we suppose an existing pseudo-vortex of the type that we have considered intensifies, it will drive more intense field-aligned currents and thus may require higher energies of precipitating electrons to carry the required current, shifting the peak altitude of energy deposition and thus the pseudo-vortex to lower altitudes. This process will be opposed by three effects: (i) deeper in the atmosphere the greater heat capacity reduces the relative heating effect of electron precipitation, (ii) below the homopause the ionospheric densities and conductances

drop considerably such that our feedback mechanism would not function, and (iii) if precipitating electrons have sufficiently high energy, they will penetrate below the pseudo-vortex and deposit most of their energy at lower altitudes. There is thus a possible feedback process for pushing a pseudo-vortex deeper, and several mechanisms for regulating this feedback. The result could be that axially asymmetric structures, once formed, create for themselves the optimum physical conditions that will allow them to intensify. This kind of secondary feedback cannot be represented in our relatively inflexible heating model.

Another possible subtlety follows from the fact that there are local time asymmetries in the existing pattern of precipitation, due to the Sun's influence on ionospheric conductances and magnetospheric plasma flows. The UV auroral power is already known to be correlated with the pulsing of the SKR emission (Nichols et al. 2010a). Thus the precipitation driving any axially asymmetric structure may effectively be 'pulsed' at the planetary period, for example increasing on the dayside due to the greater ionospheric conductance. It may be that this simply increases the energy that needs to be delivered by precipitating electrons to power an asymmetry. Alternatively, there may be resonance effects associated with such pulsed precipitation that may act to encourage the development of an asymmetry.

#### *7.2.4 Radiative cooling*

An important phenomenon that has been completely ignored in this study is radiative cooling due to both the $H_3^+$ molecular ion in the ionosphere and hydrocarbons below the homopause. The gross effect on the thermal structure is probably unimportant, since we have artificially generated a global thermal structure. Indeed, if we had including global radiative cooling in the model we would simply have been forced to increase the global heating to produce the correct global thermal structure. However, it is also possible that radiative cooling may be a component of a feedback effect.

If we consider a cool region of the thermosphere and apply the opposite arguments to those sketched in Fig. 1, it seems that we would observe inflowing, anticlockwise winds and outward flowing Pedersen currents. As already mentioned above, the Pedersen currents are mostly representative of the drift motion of ions. Thus in this situation, ions will be driven away from the core of the vortex. The major ions present in the Pedersen conducting layer are expected to be $H^+$ and $H_3^+$. The crucial difference between these ions is that $H_3^+$ has a very short recombination time-scale, of the order of minutes, whereas $H^+$ has a recombination time-scale of the order of tens of hours (Achilleos et al. 1998). Thus advection will only significantly affect the $H^+$ densities, and we would therefore expect lower $H^+$ densities at the core of the vortex. By contrast, the $H_3^+$ density will be determined directly by a balance between ionization processes and recombination. The depleted $H^+$ density at the core of the vortex means depleted electron densities and therefore slower recombination of $H_3^+$. The resulting higher $H_3^+$ densities in the core will be responsible for greater radiative cooling that could maintain the cool temperature of the core.

In practice, it is perhaps unlikely that this precise feedback effect will be important, due to the relatively low temperatures in Saturn's thermosphere making $H_3^+$ radiative cooling a less important process for the thermal energy balance than it is at Jupiter. It may be that the distribution of hydrocarbons below the homopause – or the altitude of their penetration above the homopause associated with upwelling due to thermospheric heating – could also drive feedback effects related to radiative cooling, although a possible mechanism presently defies our imagination.

### 7.3 Comparison to observations

#### *7.3.1 SKR rotation rate*

Our calculated axial asymmetries clearly do not directly represent the source of SKR emissions. However, SKR emissions, the strongest of which occur in the morning sector of the planet, must be triggered by a rotating structure moving into this sector once per rotation (Andrews et al. 2011). We will simply assume in the following that the axially asymmetric structures calculated by our model can fulfil this role, and that we can therefore equate the rotation period of these structures with the period of SKR emissions.

As already discussed in Section 5, the magnitude of the difference between the northern and southern rotation rates in our 'solstice' model is only $\sim 7\,^\circ d^{-1}$, somewhat smaller than the $20\,^\circ d^{-1}$ difference observed. However, this difference is more or less meaningless, in that our representation of 'solstice' crudely used an arbitrary, linear 10 per cent north–south bias in the heating and conductance. We do not know enough about thermospheric heating or the seasonal variability of the ionosphere to judge the realism of this model of seasonal variation. However, considering that the average lag of the northern and southern rates is $\sim 60\,^\circ d^{-1}$, and the 10 per cent north–south bias alters this by $\pm 3\,^\circ d^{-1} \sim 5$ per cent, we can at least conclude that the percentage change in the rotation rates is of the same order as the percentage change in the conductance and heating.

The crossing of the SKR rotation rates $\sim$7 months prior to equinox is not explained by our modelling. The time-scale for the thermosphere to reach dynamical equilibrium is apparently much too short to develop a time lag of this magnitude. One factor that we have neglected is the north–south asymmetry in the planetary magnetic field, which, to a sufficient approximation, is effectively a dipole shifted slightly towards the north pole of the planet (Davis & Smith 1990). Gurnett et al. (2009) discussed this as a possible source of the asymmetry, stating that the weaker magnetic field at the south pole would increase the conductance there, which is reasonable given that if all other things remain equal, Pedersen and Hall conductances are inversely proportional to magnetic field strength. However, this is not the whole story: a weaker magnetic field raises the altitude at which the ion-neutral collision frequency equals the ion gyrofrequency, which determines the peak altitude of the Pedersen conductance. How this altitude varies relative to the altitude of the peak ionosphere will have a secondary effect on the conductance. Thus it is difficult to state with certainty the effect of the magnetic field asymmetry on the conductance.

Supposing that such an asymmetry did weaken the Northern hemisphere conductances, this could explain the delay in the most recent equinox crossing, since the Northern hemisphere would require greater insolation than the south for the conductances to equalize, and therefore this would occur some time (possibly 7 months) after equinox. However, at the other equinox, when the Northern hemisphere crosses from summer into winter, this effect would produce a 7 month lead in the equinox crossing, rather than a lag. Therefore, the fact that Gurnett et al. (2010) detected the 7 month lag at both the 1995 and 2009 equinoxes seems to rule out the magnetic field asymmetry as the source of the delayed equinox crossing.

An alternative explanation for the observed time lag could lie in the chemistry of the upper atmosphere and ionosphere. Consider a chemical component of the upper atmosphere whose production

rate $P$ varies seasonally. At a particular latitude we might model the variation in production rate as $P = P_{\rm eq} + \Delta P \cos \omega t$, where $\omega$ again represents the orbital angular velocity of the planet. Suppose that the loss rate is proportional to the number density $n$ with time-scale $\tau_L$, then we have

$$\frac{{\rm d}n}{{\rm d}t} = P_{\rm eq} + \Delta P \cos \omega t - \frac{n}{\tau_L}. \quad (12)$$

It is not difficult to show that provided $\tau_L \ll T$, the orbital period of the planet, then $n \simeq n_0 + \Delta n \cos \omega(t - \tau_L)$ is a solution to the equation. There is thus a time lag in the response of the chemistry to the seasonal variation that is equal to its chemical loss time-scale.

Thus, we require a chemical in the upper atmosphere whose production rate varies seasonally, with a loss time-scale of $\sim$7 months and the potential to significantly affect either the thermosphere or ionosphere. One possible candidate is external material, for example that deposited by the ablation of meteors in the upper atmosphere. Moses & Bass (2000) studied the chemistry of Mg that may have been deposited by this process. They estimated the loss time-scale of Mg due to recondensation to be of the order of $10^6$ s, or about 0.4 months, but there is a large degree of uncertainty in this value, such that they also investigated the effects of recondensation time-scales as long as $\sim$4 months. Furthermore, they found that the quantity of Mg in the upper atmosphere had a significant effect on the electron density in the lower layers of the ionosphere, with $Mg^+$ forming a secondary electron density peak at $\sim$800 km.

However, it is not obvious why the production rate of external material would vary seasonally. One possibility is that the flow of dust and meteoritic material at the orbit of Saturn is asymmetric with respect to the direction of the Sun. If there were, for example, a greater flow of material away from the Sun than towards it, then the summer hemisphere would experience a slightly greater influx of material, and this would indeed vary seasonally. At Earth the flow of meteoritic material into the atmosphere is asymmetric with respect to the Sun, but in a complex way consisting of a number of separate components of different origin (e.g. Younger et al. 2009). Without suitable observations, we can only speculate on the equivalent distribution at Saturn. However, a priori we would expect it to be asymmetric with respect to the Sun – to some extent at least – producing some degree of seasonal variation. Indeed, the complexity of such a distribution could be responsible for the non-sinusoidal nature of the variation in the SKR rotation rates, which appear to exhibit a secondary 'dip' close to solstice (Gurnett et al. 2010).

As an alternative to invoking seasonal variation in the influx of material, there could be a reasonably uniform distribution of condensed meteoric material across the upper atmosphere which exists in a cycle of evaporation by absorption of solar photons followed by recondensation over a 7 month time-scale. Another possible source of external material is the rings, which may supply water to magnetically connected latitudes (e.g. Connerney 1986). It is possible that more material may be supplied to the summer hemisphere when the corresponding side of the rings is illuminated. These considerations are largely speculative. However, the key point is that the 7 month time lag, which is much longer than any of the dynamical time-scales in the system, is easily explained if a suitable chemical loss process is important in controlling the behaviour of the upper atmosphere.

A new possibility raised by our results is that there may be multiple axial asymmetries in the thermosphere which, presumably, would generate multiple SKR rotation rates. In particular, run E1c showed a mixture of a westward drifting $m = 1$ asymmetry and an eastward drifting $m = 2$ asymmetry. The $m = 1$ asymmetry can clearly be identified with the usual planetary-period SKR emission. However, the $m = 2$ asymmetry, which is rotating slightly faster than the planetary interior, would produce SKR periods with slightly less than half the planetary period – and, crucially, the period would not be exactly half that of the westward drifting $m = 1$ asymmetry, so it would not have the expected frequency of a harmonic of the main SKR emission. It is interesting to speculate whether such a secondary SKR period may be present in the data but has been overlooked due to the focus on the main planetary-period SKR emission. Our results also showed oscillations in the rotation rate of the main $m = 1$ asymmetry, for example the 15 rotation oscillation in run E1. The conditions in our model are sufficiently different from the real conditions at Saturn for us to suggest that this specific oscillation period will probably not be present in the data, and indeed it is possibly too short to be easily measurable, however should such short time-scale oscillations be observed they would be easily explained by our results.

### *7.3.2 Magnetic field perturbations*

Smith (2011) investigated the magnetic field perturbations generated by an artificially sustained asymmetry in the thermosphere. This study found that approximately the required magnetic field morphology was produced in the equatorial plane to match the model of Southwood & Kivelson (2007), however the magnitude of the perturbations was $\sim$100 times too small. To compare our results with Smith (2011) we have carried out an identical calculation for the final output of run E1, using only the currents driven by the axial asymmetries in the Northern hemisphere.

The results are shown in Fig. 18, which shows the same information and uses identical line formats to fig. 10 of Smith (2011). The figure shows the azimuthal (thick line) and radial (thin line) magnetic field perturbations at a radial distance of $10R_S$. To match the model of Southwood & Kivelson (2007) these two components should be sinusoidal, have approximately the same amplitude, and be out of phase by 90°, thus producing an approximately uniform magnetic field lying in the equatorial plane. The dashed and dotted lines show appropriate sinusoids, with the dashed sinusoid's amplitude and phase matched by eye to the maximum of the azimuthal component. It is clear that the components predicted by our model are approximately of the correct form – each shows a single minimum and maximum and they are out of phase by $\sim$90° in the correct sense.

However, the perturbations are not a good match to the sinusoids shown. Most of the variation occurs in the 180° longitude range surrounding the peak of the azimuthal field perturbation. This is quite simply because the pseudo-vortex is too localized. This localization of the structure as it develops has already been discussed with reference to Fig. 10. A second problem with the magnetic field perturbations is that including the Southern hemisphere axial asymmetries in the calculation adds a second set of perturbation fields that, in every situation except the exact equinox conditions of E1, will rotate with a different period. Indeed, the assumed interhemispheric current system of Smith (2011) is probably inadequate to explain the presence of different perturbation periods in the Northern and Southern hemispheres (Provan et al. 2011), since interhemispheric currents driven by separate axial asymmetries in each hemisphere imply that both periods will be present in both hemispheres of the magnetosphere. It therefore seems that something similar to the partial ring current model of Provan et al. (2009) will be necessary to produce the required current closure while also isolating the magnetic field perturbations in the Northern and Southern hemispheres.

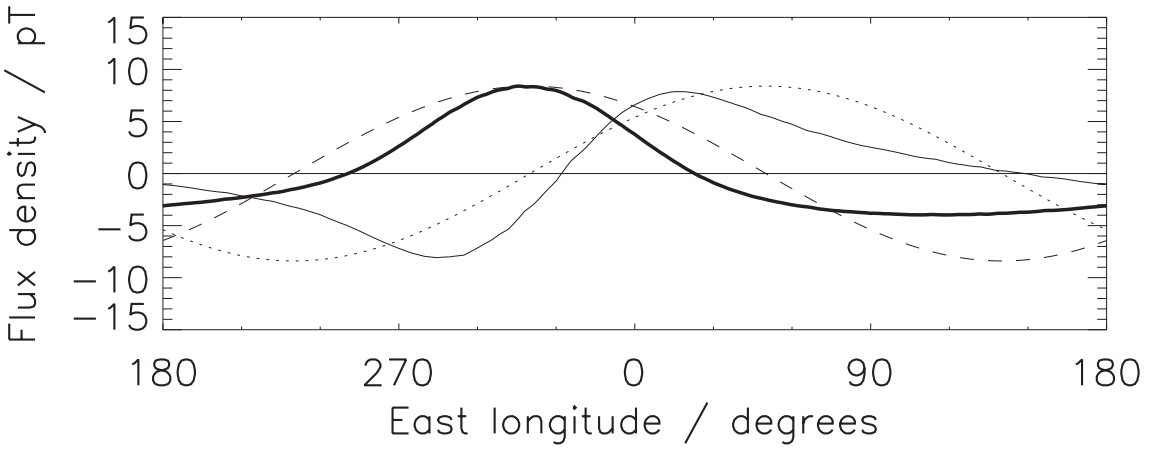


**Figure 18.** Magnetic field perturbations at a radial distance of $10R_S$ due to currents driven by the Northern hemisphere winds of run E1. This plot is equivalent to fig. 10 of Smith (2011). The perturbations are calculated assuming that divergences in the Northern hemisphere are closed by an interhemispheric current system flowing along dipolar field lines. The thick and thin solid lines show azimuthal and radial fields. The dashed line is a sinusoid fitted by eye to the maximum of the azimuthal field. The dotted line is a sinusoid of the same amplitude displaced 90° in longitude.

*7.3.3 Auroral oval oscillations*

As already described, the latitudinal resolution of our model is too low to resolve the details of the currents and particle precipitation associated with the main auroral oval. However, our results are suggestive that the auroral oval may be important not just as a signature of the asymmetry but as its cause. Most suggestive of this is that the pseudo-vortex naturally develops at the shear in the plasma flow at 15° colatitude that corresponds to the location of the main auroral oval. In the model of Cowley et al. (2004a) it is the upward current sheet associated with this plasma flow shear that causes the particle precipitation that forms the main auroral oval. Indeed, it is not just the plasma flow shear that produces the current sheet, but a combination of the shear in the plasma and neutral flows at this location. Furthermore, as we have discussed, the pseudo-vortex is not a true a vortex, but a modification to the shear in the neutral flow that is partially responsible for producing the main auroral oval. It seems a reasonable possibility that there could be a more latitudinally confined vortex or pseudo-vortex that, rather than driving itself through modification of particle precipitation across the polar cap, drives itself through small latitudinal shifts in the location of the main auroral oval.

### 7.4 Implications for Jupiter

The feedback mechanism discussed here is equally applicable to Jupiter. The Jovian case is somewhat different, however, in that the axial symmetry of the thermosphere is already broken by the highly axially asymmetric and non-dipolar planetary magnetic field. Nevertheless, similar puzzles related to multiple periodicities exist at Jupiter. In addition to the System III rotation rate of the planetary magnetic field, there is an additional periodicity known as system IV (Sandel & Dessler 1988) present in observations of the Io plasma torus (a detailed review of which can be found in Steffl, Delamere & Bagenal 2006). There are also reports of periodicities at neither the System III nor System IV periods, for example in the Io torus (Steffl et al. 2006) and in radio emissions (e.g. Panchenko et al. 2010). Other more peculiar behaviours have been observed, for example Brown (1995) reported that the System IV modulation of emission intensity from the Io plasma torus underwent a sudden phase jump by 100° in longitude.

Our model may naturally account for many of these effects, as discussed previously in outline by Smith (2006). The results of this study specifically indicate that multiple periodicities may be present, given that we can not only have different axial asymmetries in the Northern and Southern hemispheres, but we can also have more than one axial asymmetry in a single hemisphere, each rotating with a different period. A phase jump could be caused if close to the relevant latitude there were multiple axially asymmetric wind systems, whose relative amplitudes varied rapidly. Such a change could possibly be initiated by an axially asymmetric wind system confined to a specific latitude interacting with the main auroral oval as it drifts in longitude.

## 8 SUMMARY AND CONCLUSIONS

We have shown that by the introduction of a feedback mechanism between thermospheric winds and particle precipitation it is possible to permanently break the axial symmetry of Saturn's thermosphere through the development of a vortex-like structure at high latitudes which we have referred to as a *pseudo-vortex*. The rotation rate of this pseudo-vortex is sensitive to the conductance of the ionosphere, producing seasonal effects that qualitatively match the seasonal variation of the SKR rotation rates. Other than at equinox, axial asymmetries in the Northern and Southern hemispheres rotate with different periods.

The feedback mechanism requires very high energy (∼5 MeV) precipitating electrons at high latitudes which must be distributed with a high degree of axial asymmetry to drive the pseudo-vortex. Neither the required energies nor the distribution match the observations. These two discrepancies are the most significant flaws in our model. It is possible that higher resolution modelling or the inclusion of further physical factors or feedback effects may reduce the required energies or alter the required distribution of precipitating electrons such that they more closely match the observations.

Although the specific physical parameters required do not fit with the actual situation at Saturn, this study does demonstrate the physical plausibility of such a feedback model. The approach that we have taken of imagining a feedback effect and then attempting to model its implications in a simplified way has proved fruitful in providing insight to how such a thermospheric asymmetry may

behave. We are able to offer partial or speculative explanations of the following.

(i) The origin of the rotating asymmetry may be due to a thermospheric vortex or pseudo-vortex driven by feedback.

(ii) The lag of the SKR rotation rate compared to the internal rotation may be due to westward drift of a thermospheric asymmetry, either due to advection or due to the locally asymmetric nature of the heating driving the axial asymmetry.

(iii) The seasonal variation of the SKR rotation rate may be explained in terms of seasonally variable conductances altering the magnitude of the westward drift.

(iv) The 7 month time lag observed in the crossing of the northern and southern rotation rates may be explained in terms of long chemical time-scales producing a 7 month lag in the seasonal variation of conductance.

(v) Magnetic field perturbations in the magnetosphere driven by a thermospheric asymmetry are of approximately the correct form to explain some of the observations (see also Smith 2011).

We can also predict the possible existence of the following effects which are implied by our results and would constitute indirect evidence for the existence of axially asymmetric thermospheric wind systems.

(i) Multiple periodicities in a single hemisphere, for example a weak $m = 2$ asymmetry coexisting with an $m = 1$ asymmetry and rotating at a slightly different rate.

(ii) Oscillations in the periodicities on time-scales much shorter than the time-scale for seasonal variation, of a similar nature to the $\sim$15 rotation period oscillations predicted by our model.

(iii) Rotating longitudinal brightness asymmetries in the auroral emissions, correlated with the SKR emission period.

Finally, the following are problems with the model that remain to be resolved.

(i) It has not been shown whether a thermospheric asymmetry can actually be driven by plausible distributions of precipitating electrons, or by some other mechanism, without contradicting existing observations.

(ii) The magnitude of the magnetic field perturbations driven by a thermospheric asymmetry remains $\sim$100 times too small to explain the observations (see also Smith 2011).

(iii) It has not been demonstrated how the current systems driven by a thermospheric asymmetry can close in such a way that the northern and southern magnetospheres exhibit magnetic field perturbations with distinct rotation rates.

(iv) It has not been demonstrated how a thermospheric asymmetry could either trigger SKR emission, shift the location of the main auroral oval, or cause any of the other observed periodic effects.

## ACKNOWLEDGMENTS

The authors thank Luke Moore for providing results from his ionosphere model.

## REFERENCES

Achilleos N., Miller S., Tennyson J., Aylward A. D., Mueller-Wodarg I., Rees D., 1998, J. Geophys. Res., 103, 20089
Achilleos N., Guio P., Arridge C. S., Sergis N., Wilson R. J., Thomsen M. F., Coates A. J., 2010, Geophys. Res. Lett., 37, 20201
Anderson J. D., Schubert G., 2007, Sci, 317, 1384
Andrews D. J., Bunce E. J., Cowley S. W. H., Dougherty M. K., Provan G., Southwood D. J., 2008, J. Geophys. Res., 113, 9205
Andrews D. J., Coates A. J., Cowley S. W. H., Dougherty M. K., Lamy L., Provan G., Zarka P., 2010, J. Geophys. Res., 115, A12252
Andrews D. J., Cecconi B., Cowley S. W. H., Dougherty M. K., Lamy L., Provan G., Zarka P., 2011, J. Geophys. Res., 116, 9206
Bougher S. W., Waite J. H., Majeed T., Gladstone G. R., 2005, J. Geophy. Res., 110, 4008
Brown M. E., 1995, J. Geophys. Res., 100, 21683
Burch J. L., DeJong A. D., Goldstein J., Young D. T., 2009, Geophys. Res. Lett., 36, L14203
Clarke J. T. et al., 2005, Nat, 433, 717
Connerney J. E. P., 1986, Geophys. Res. Lett., 13, 773
Cowley S. W. H., Bunce E. J., O'Rourke J. M., 2004a, J. Geophys. Res., A18, 5212
Cowley S., Bunce E., Prangé R., 2004b, Ann. Geophys., 22, 1379
Davis L. J., Smith E. J., 1990, J. Geophys. Res., 95, 15257
Desch M. D., Kaiser M. L., 1981, Geophys. Res. Lett., 8, 253
Dessler A. J., 1985, Geophys. Res. Lett., 12, 299
Espinosa S. A., Southwood D. J., Dougherty M. K., 2003a, J. Geophys. Res., 108
Espinosa S. A., Southwood D. J., Dougherty M. K., 2003b, J. Geophys. Res., 108
Galand M., Moore L., Mueller-Wodarg I., Mendillo M., Miller S., 2011, J. Geophys. Res., 116, 9306
Galopeau P. H. M., Lecacheux A., 2000, J. Geophys. Res., 105, 13089
Gérard J., Grodent D., Gustin J., Saglam A., Clarke J. T., Trauger J. T., 2004, J. Geophys. Res., 109, 9207
Goldreich P., Farmer A. J., 2007, J. Geophys. Res., 112, A05225
Grodent D., Gustin J., Gérard J.-C., Radioti A., Bonfond B., Pryor W. R., 2011, J. Geophys. Res., 116, 9225
Gurnett D. A., Persoon A. M., Kurth W. S., Groene J. B., Averkamp T. F., Dougherty M. K., Southwood D. J., 2007, Sci, 316, 442
Gurnett D. A., Lecacheux A., Kurth W. S., Persoon A. M., Groene J. B., Lamy L., Zarka P., Carbary J. F., 2009, Geophys. Res. Lett., 36, 16102
Gurnett D. A., Groene J. B., Persoon A. M., Menietti J. D., Ye S., Kurth W. S., MacDowall R. J., Lecacheux A., 2010, Geophys. Res. Lett., 37, L24101
Holton J. R., 1992, An Introduction to Dynamic Meteorology. Academic Press, New York
Hubbard W. B. et al., 1997, Icarus, 130, 404
Matcheva K. I., Strobel D. F., 1999, Icarus, 140, 328
Moore L. E., Mendillo M., Müller-Wodarg I. C. F., Murr D. L., 2004, Icarus, 172, 503
Moore L., Mueller-Wodarg I., Galand M., Kliore A., Mendillo M., 2010, J. Geophys. Res., 115, A11317
Moses J. I., Bass S. F., 2000, J. Geophys. Res., 105, 7013
Moses J. I., Bézard B., Lellouch E., Gladstone G. R., Feuchtgruber H., Allen M., 2000, Icarus, 143, 244
Müller-Wodarg I. C. F., Mendillo M., Yelle R. V., Aylward A. D., 2006, Icarus, 180, 147
Nichols J. D., Clarke J. T., Cowley S. W. H., Duval J., Farmer A. J., Gérard J.-C., Grodent D., Wannawichian S., 2008, J. Geophys. Res., 113, A11205
Nichols J. D. et al., 2010a, Geophys. Res. Lett., 37, L15102
Nichols J. D., Cowley S. W. H., Lamy L., 2010b, Geophys. Res. Lett., 37, L24102
Panchenko M., Rucker H. O., Kaiser M. L. St., Cyr O. C., Bougeret J.-L., Goetz K., Bale S. D., 2010, Geophys. Res. Lett., 37, L05106
Provan G., Andrews D. J., Arridge C. S., Coates A. J., Cowley S. W. H., Milan S. E., Dougherty M. K., Wright D. M., 2009, J. Geophys. Res., 114, 2225
Provan G., Andrews D. J., Cecconi B., Cowley S. W. H., Dougherty M. K., Lamy L., Zarka P. M., 2011, J. Geophys. Res., 116, A04225
Radioti A., Grodent D., Gérard J.-C., Milan S. E., Bonfond B., Gustin J., Pryor W., 2011, J. Geophys. Res., 116, 11209
Read P. L., Dowling T. E., Schubert G., 2009, Nat, 460, 608
Sandel B. R., Dessler A. J., 1988, J. Geophys. Res., 93, 5487

Smith C. G. A., 2006, Ann. Geophys., 24, 2709
Smith C. G. A., 2011, MNRAS, 410, 2315
Smith C. G. A., Aylward A. D., 2008, Ann. Geophys., 26, 1007
Smith G. R., Shemansky D. E., Holberg J. B., Broadfoot A. L., Sandel B. R., McConnell J. C., 1983, J. Geophys. Res., 88, 8667
Smith C. G. A., Aylward A. D., Miller S., Müller-Wodarg I. C. F., 2005, Ann. Geophys., 23, 2465
Smith C. G. A., Aylward A. D., Millward G., Miller S., Moore L. E., 2007, Nat, 445, 399
Southwood D., 2011, J. Geophys. Res., 116, A01201
Southwood D. J., Kivelson M. G., 2007, J. Geophys. Res., 112, 12222
Southwood D. J., Kivelson M. G., 2009, J. Geophys. Res., 114, A09201
Stallard T. S., Miller S., Trafton L. M., Geballe T. R., Joseph R. D., 2004, Icarus, 167, 204
Stallard T., Miller S., Melin H., Lystrup M., Cowley S. W. H., Bunce E. J., Achilleos N., Dougherty M., 2008, Nat, 453, 1083
Steffl A. J., Delamere P. A., Bagenal F., 2006, Icarus, 180, 124
Younger P. T., Astin I., Sandford D. J., Mitchell N. J., 2009, Ann. Geophys., 27, 2831

This paper has been typeset from a TEX/LATEX file prepared by the author.