付録 A

Space Weather Modeling: Status and Prospects (Journal of Geophysical Research, Vol. 113, No. A3, 2008)

"Space Weather Modeling: Status and Prospects" (Journal of Geophysical Research, Vol. 113, No. A3, 2008) は 2006 年 11 月 14 日–17 日に地球シミュレータセンター（横浜市）で開催された「CAWSES International Workshop on Space Weather Modeling」（本報告書 83 ページを参照）の集録として出版されたものである。

掲載論文

- **Introduction to special section on Space weather modeling: Status and prospects**,
  Kamide, Y. (2008),
  J. Geophys. Res., *113*, A03S01, `doi:10.1029/2007JA012972`
- **Magnetic flux emergence in the Sun**,
  Archontis, V. (2008),
  J. Geophys. Res., *113*, A03S04, `doi:10.1029/2007JA012422`
- **Radiation Belt Environment model: Application to space weather nowcasting**,
  Fok, M., R. B. Horne, N. P. Meredith, and S. A. Glauert (2008),
  J. Geophys. Res., *113*, A03S08, `doi:10.1029/2007JA012558`
- **Nonlinear force-free coronal magnetic field extrapolation scheme based on the direct boundary integral formulation**,
  He, H., and H. Wang (2008),
  J. Geophys. Res., *113*, A05S90, `doi:10.1029/2007JA012441`
- **Geomagnetically induced currents during intense storms driven by coronal mass ejections and corotating interacting regions**,
  Kataoka, R., and A. Pulkkinen (2008),
  J. Geophys. Res., *113*, A03S12, `doi:10.1029/2007JA012487`
- **Properties of AE indices derived from real-time global simulation and their implications for solar

wind-magnetosphere coupling,

Kitamura, K., H. Shimazu, S. Fujita, S. Watari, M. Kunitake, H. Shinagawa, and T. Tanaka (2008),

J. Geophys. Res., *113*, A03S10, `doi:10.1029/2007JA012514`

- **Numerical studies on three-dimensional earthward fast plasma flows in the near-Earth plasma sheet by the spontaneous fast reconnection model**,

  Kondoh, K., and M. Ugai (2008),

  J. Geophys. Res., *113*, A03S07, `doi:10.1029/2007JA012707`

- **Flux enhancement of the outer radiation belt electrons after the arrival of stream interaction regions**,

  Miyoshi, Y., and R. Kataoka (2008),

  J. Geophys. Res., *113*, A03S09, `doi:10.1029/2007JA012506`

- **A quantitative MHD study of the relation among arcade shearing, flux rope formation, and eruption due to the tearing instability**,

  Shiota, D., K. Kusano, T. Miyoshi, N. Nishikawa, and K. Shibata (2008),

  J. Geophys. Res., *113*, A03S05, `doi:10.1029/2007JA012516`

- **Parametric decay of circularly polarized Alfvén waves in the radially expanding solar wind**,

  Tanaka, S., T. Ogino, and T. Umeda (2007),

  J. Geophys. Res., *112*, A10110, `doi:10.1029/2007JA012513`

- **Nonlinear force-free modeling of the solar coronal magnetic field**,

  Wiegelmann, T. (2008),

  J. Geophys. Res., *113*, A03S02, `doi:10.1029/2007JA012432`

- **Key points model for polar region currents**,

  Xu, W., G. Chen, A. Du, Y. Wu, B. Chen, and X. Liu (2008),

  J. Geophys. Res., *113*, A03S11, `doi:10.1029/2007JA012588`

- **Some ubiquitous features of the mesospheric Fe and Na layer borders from simultaneous and common-volume Fe and Na lidar observations**,

  Yi, F., S. Zhang, X. Yue, Y. He, C. Yu, C. Huang, and W. Li (2008),

  J. Geophys. Res., *113*, A04S91, `doi:10.1029/2007JA012632`

Introduction to special section on Space weather modeling: Status and prospects

Y. Kamide[1]

[1] The importance of specifying and predicting the state of the Earth's space environment, i.e., space weather, has recently been increasingly realized, as dynamic activities of the Sun, solar wind, magnetosphere, and ionosphere can influence modern technology systems and even endanger human life. It is also widely recognized that numerical simulations driven by integrated ground-based and satellite observations are a powerful tool for quantitatively understanding the complex Sun-Earth system, and they are a vital means for predicting space weather.

[2] It is required, however, that a physical model of solar-terrestrial phenomena be constructed as a basis of space weather predictions. The model must deal with the key processes spanning from flares and coronal mass ejections on the Sun to geomagnetic storms and their effects in the upper atmosphere of Earth. To reach the goal effectively, it is acute to ask a number of fundamental questions such as, What is the triggering condition of solar flares and coronal mass ejections (CMEs)? What mechanism generates fast reconnection and leads to particle acceleration? Why particular solar flares cause superstorms on Earth? Working out these basic questions is probably equivalent to understanding multiscale coupling of various physical processes occurring in the Sun-Earth system on different scales. It is indispensable that the communications among different fields, especially between the solar and geophysical communities, be enhanced more systematically.

[3] This special section results primarily from the CAWSES International Workshop on Space Weather Modeling (CSWM), which was held on 14–17 November 2006 in Yokohama, Japan, but includes several papers that were not presented at the Workshop. It is hoped that this section will provide a timely forum for discussing all aspects of space weather modeling research, aiming to summarize the current status and future prospects of this rapidly growing field. It contains original research papers as well as topical review papers on the advanced modeling of solar and heliospheric magnetic field; solar flares, CMEs, and filament eruption; solar wind; geomagnetic storms; and radiation belts.

Y. Kamide, Research Institute for Sustainable Humanosphere, Kyoto University, Uji 611-0011, Japan. (kamide@rish.kyoto-u.ac.jp)

---

[1]Research Institute for Sustainable Humanosphere, Kyoto University, Uji, Japan.

Magnetic flux emergence in the Sun

V. Archontis[1]

[1] Space weather research is closely connected with the study of the solar magnetic activity. In past years, many solar missions (e.g., YOHKOH, SOHO, TRACE, and RHESSI) have provided outstanding observations, which have been used to improve our understanding of the structure and the dynamical evolution of solar magnetic fields. In addition, the newly launched solar missions (e.g., Hinode and STEREO) will study the interaction between the emerging magnetic field and the preexisting field in the corona (increasing our understanding of the causes of solar variability), and they will also observe the three-dimensional evolution of solar eruptions as they leave the Sun and move into the interplanetary space. One of the most important processes, responsible for many dynamical phenomena observed in the Sun, is the emergence of magnetic flux from the solar interior in active regions and the modification of the coronal magnetic field in response to the emergence. In fact, magnetic flux emergence might be responsible for the appearance of small-scale events (e.g., compact flares, plasmoids, and active-region-associated X-ray brightenings) and large-scale events (e.g., X-class flares and CMEs), which are major drivers of space weather. However, it is clear that the question of how exactly the magnetic fields rise through the convection zone of the Sun and emerge through the photosphere and chromosphere into the corona has still not been solved. It is believed that understanding the process of flux emergence is an important step toward the understanding of the initiation mechanism of eruptive events in the Sun, which is another topic of great debate. This paper provides a brief review of the theory and the numerical models, which have been used to study the process of magnetic flux emergence into the outer atmosphere of the Sun. We underline the similarities and differences between these models, and we compare the basic features of the numerical results with observations. Finally, we review the recent progress and discuss what further developments are required in the models to best describe the essential physics in the process of flux emergence.

1. Introduction

[2] Most of the solar activity is directly linked to the Sun's magnetic field. Indeed, observations have shown that the Sun's corona is highly structured, being threaded by a complex network of magnetic fields. Understanding the structure and the dynamic evolution of these fields is important because they are the building blocks of solar activity. It is believed that solar eruptive phenomena result from rapid changes in their structures and connections.

[3] The current state of the theory of the formation of the Sun's magnetic field suggests that it is produced by dynamo action in the tachocline, an interface layer separating the convection zone from the radiative zone. Then the dynamo-generated magnetic field is transported from the deep interior of the Sun to the surface by magnetic buoyancy [e.g., *Parker*, 1955], which may well be coupled with the convective motions [*Parker*, 1988]. Eventually, the buoyant magnetic fields rise through the convection zone, intersect the photosphere and create the observed sunspots and bipolar active regions [e.g., *Zwaan*, 1987]. The newly emerged bipolar active regions are called emerging flux regions (EFRs) [e.g., *Zirin*, 1970]. An extended review on the structure and dynamics of magnetic fields in the solar convection zone is given by *Fan* [2004]. The properties of active regions in terms of the dynamics of magnetic flux tubes which emerge from the solar interior to the photosphere have been reviewed by *Fisher et al.* [2000].

[4] Shortly after the appearance of flux at the photosphere, a system of bright loops appears in the EUV and X-ray detectors. Observations of large-scale magnetic fields emerging into the corona have been identified by solar satellites in the 1990s (e.g., TRACE and SOHO). The coronal fields are normally outlined by plasma emitting in EUV or X rays. Figure 1 is a high-resolution image of the Sun taken by the X-ray telescope on Hinode satellite. This image shows the detailed structure of active region loops and X-ray bright points, which are seen as concentrations of

[1]School of Mathematics and Statistics, University of St. Andrews, Fife, UK.
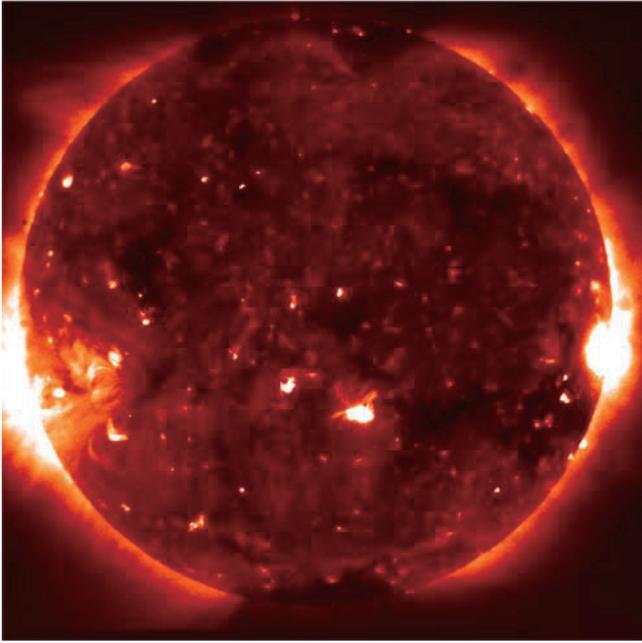
Figure 1. First-light image from the X-ray telescope on board the Hinode satellite. Credit: Japanese Aerospace Exploration Agency (JAXA).

magnetic loops. The appearance of many brightenings indicate that solar activity occurs all over the Sun.

[5] However, it is exceedingly difficult to measure coronal magnetic fields directly and, thus, we usually use the line-of-sight component of the magnetic field and vector magnetograms from photospheric measurements, to extrapolate the magnetic field in the corona. For the photosphere, direct measurements of the magnetic field, both by satellite instruments (like MDI) as well as ground-based detectors, yield a wealth of information. Vector magnetograms, in particular, permit the reconstruction of all three components of the magnetic field.

[6] *Schrijver et al.* [2006] evaluated the performance of a series of numerical models, which simulate nonlinear force-free (NLFF) magnetic fields in the solar corona. The numerical experiments were performed using vector magnetic field measurements of active regions. It was found that the solution depends strongly on the implementation of the boundary conditions. The ultimate goal of this comparison between the numerical models is to find a robust method of measuring coronal free energy in solar active regions.

[7] Another approach, which has been used to compute the large-scale magnetic field configuration of the solar corona, is the potential field source surface (PFSS) models. These models are simple to develop and implement and they can also resolve larger-scale structures than the current MHD models. On the other hand, they use idealized initial conditions and assumptions and they don't treat important processes (such as reconnection) properly. *Riley et al.* [2000] showed that PFSS models are useful tools for computing the large-scale coronal field when time-dependent changes in the photospheric flux can be neglected. Also, they found that PFSS models produce

similar results with the MHD models for configurations based on untwisted coronal fields.

[8] Observations show that most of the coronal loops seem to join the opposite polarities of the new active regions, while others establish linkages to new active regions, probably through reconnection of magnetic field lines. Space observations from YOHKOH, SOHO and TRACE have revealed evidence of magnetic reconnection [*Martens*, 2003] in emerging flux regions, in flares and it is believed that reconnection could power solar explosions called Coronal Mass Ejections (CMEs). *Antiochos et al.* [1999] presented the first model (known as breakout model), which was consistent with many observational properties and the energy requirements for CMEs. In this model, CMEs are triggered owing to reconnection between a sheared arcade and neighboring flux systems.

[9] Eruptive phenomena, such as CMEs, flares and prominences can lead to solar and geomagnetic disturbances [*Taylor et al.*, 1997] and disrupt terrestrial satellites and power systems. Recent models have increased progress on space weather prediction. The warning of an explosive event from space weather on terrestrial and space systems can be provided (a few hours ahead) on the basis of observations of such events from spacecrafts. The exact impact, however, of the potential threat in the Sun-Earth environment requires further development of space weather models [*Brun*, 2007; *Cole*, 2003] and remains a great challenge.

[10] X-ray images from YOHKOH and vector magnetograms from ground-based observatories have shown the topology of growing bipoles in areas of new magnetic flux. The analysis of the proper motions of these bipoles indicates that the flux bundles that they make up the bipoles are twisted before they emerge [*Leka et al.*, 1996]. Measurements of the twist of the emerging flux systems have shown that the change of the twist is rather small during the emergence [*Wang and Abramenko*, 1999]. In addition, photospheric measurements of the vector magnetic field have shown that the mean twist in active regions is right-handed in the southern hemisphere and left-handed in the northern hemisphere [*Pevtsov et al.*, 2001]. Soft X-ray data from observations of active regions have repeatedly shown coronal loops with a bright forward or an inverse S shape. The structures with the forward S shape appear mostly in the southern hemisphere while the inverse S shape structures usually appear in the northern hemisphere. Their shape appears to be helical because their magnetic fields are twisted. Their central part is approximately aligned with the neutral line of the normal component of the magnetic field in the photosphere. Filaments, with a forward or an inverse S shape, may also be visible along the neutral line in sigmoid active regions [*Rust and Kumar*, 1994].

[11] The above mentioned structures are called sigmoids. Using YOHKOH solar X-ray images, it has been shown that there is a relationship between structures with a sigmoidal shape and eruptions of flux ropes in the corona [*Sterling*, 2000; *Pevtsov*, 2002]. For example, CMEs appear to originate preferentially in regions of the Sun's corona that exhibit sigmoidal structures [*Canfield et al.*, 1999, 2000]. Although sigmoids play an important role in solar activity

and their evolution is closely connected to space weather forecasting [*Sterling*, 2000; *Pevtsov*, 2002], the origin and the lifetime of sigmoids are still unknown.

[12] Another interesting issue, closely connected with the study of flux emergence, is the actual structure of the solar magnetic field. Observations have shown that magnetic fields on the photosphere are intermittent and magnetic flux is predominantly concentrated in discrete areas with kilogauss field strengths [*Zwaan*, 1987; *Keller*, 1992; *Hagenaar*, 2001; *Socas Navarro and Sanchez Almeida*, 2003]. Resolution is a crucial factor on how one interprets observations since the size of these areas changes from sunspots down to very small scales, which are difficult to resolve with the available observational facilities. In the past few years, many observations have shown that the observed field adopts the form of roundish and discrete flux tubes while recent high-resolution observations (i.e., with the Swedish Solar Telescope on La Palma) show that the observed field has an intricate topology. Theoretical arguments support also the intermittent morphology of the photospheric and the subsurface magnetic fields, which are believed to be concentrated into discrete flux tubes. Numerical MHD simulations of magnetoconvection [*Galloway and Weiss*, 1981; *Nordlund et al.*, 1992] have shown that magnetic flux is concentrated in the intergranular lanes by convective motions. Eventually, the magnetic flux adopts the form of flux tubes, which are intensified because of stretching and twisting of their magnetic fieldlines by turbulent fluid motions. It is worthwhile to mention that, magnetoconvection simulations on granule and mesogranule scales in the upper part of the convection zone [*Steiner et al.*, 1998; *Weiss et al.*, 2002; *Stein and Nordlund*, 2006] have shown that the magnetic field lines of the flux tubes at the surface of the Sun are connected to various regimes, with a complex topology, below the surface and, thus, the concept of discrete flux tubes for the weak magnetic fields in the quiet Sun is not clear.

[13] Although it appears that we do not have yet a complete picture of the topology of the magnetic field on or below the photosphere, the concept of isolated magnetic flux tubes has been used extensively over the past decade or so in numerical experiments of flux emergence from the solar interior into the solar atmosphere. Most of the numerical models use, as initial configuration, a twisted flux tube or a flux sheet below the photosphere. Eventually, the initial flux system becomes unstable to perturbations or instabilities (i.e., the classical Parker buoyancy instability [*Parker*, 1978]) and makes its way up through the solar interior developing an Ω-loop shape. As the buoyant flux system rises, the top of the Ω-loop structure intersects the photosphere and creates sunspots in bipolar regions. Finally, it emerges through the photosphere and chromosphere and expands into the corona. However, even in the simplest configuration, the above phenomenon is highly time-dependent, has a complex three-dimensional geometry and the timescales of the various processes involved are remarkably different in the subphotospheric layers and in the upper atmosphere of the Sun. Hence numerical experiments are necessary to provide a first physical understanding of the flux emergence process. This paper provides a review of the results of these numerical experiments.

2. Emergence Into a Field-Free Corona

[14] The emergence of buoyant magnetic flux systems from the convectively unstable solar interior into the higher levels of the atmosphere is still a largely unexplored research domain and has been a subject of vigorous research for the past three decades. In fact, the evolution of the rising flux systems occurs on the basis of the buoyancy instability experienced by the plasma above the photosphere. There has been much interest on the literature in the buoyant instabilities in magnetized and in nonmagnetized plasmas (see the review by *Hughes and Proctor* [1998]). In the following, we first report on magnetic buoyancy instabilities in the framework of numerical MHD experiments of flux emergence, then on the issue of twisted flux tubes and finally on the dynamics of flux emergence.

2.1. Buoyant Instabilities

[15] A first series of numerical experiments were two-dimensional (2-D) and explored the excitation of magnetic instabilities of a single flux system (tube or sheet) and its subsequent emergence into a nonmagnetized corona. *Shibata et al.* [1989a, 1989b] performed 2-D magnetohydrodynamic (MHD) experiments to study the nonlinear evolution of the Parker instability, which is a kind of ideal MHD instability driven by magnetic buoyancy. More precisely, they studied the undular mode $\kappa \parallel B$ of the magnetic buoyancy instability, where κ and B are the wave number and the initial magnetic field vector, in an isolated horizontal flux sheet. The background stratified atmosphere in these experiments consisted of two unmagnetized isothermal layers as an simplified version of the Sun's photosphere/chromosphere and the ambient corona with a higher temperature. Also, small velocity perturbations were initially imposed on the magnetic flux sheet to initiate the instability. They found that as soon as the instability develops the flux sheet rises as a result of enhanced magnetic buoyancy and eventually expands into the corona. The acceleration of the rising loop shows a self-similar behavior in the low atmosphere. This self-similar solution reveals that the magnetic loop is accelerated by the magnetic pressure gradient force, which dominates the gravitational and the gas pressure gradient force. The results of these simulations were consistent with the observed small rise velocity of magnetic flux at the photospheric heights ($v \leq 1$ Km sec^{-1}) and with strong downdrafts ($v \approx 1-3$ km sec^{-1}) at the footpoints of the expanding magnetic loop, which may correspond to observed strong downdrafts near pores.

[16] Simulations by *Kaisig et al.* [1990] included a convectively unstable layer below the photosphere. Then vertical velocity fluctuations in the convection zone and horizontal shear flows at photospheric heights were considered to study the evolution of the undular mode of magnetic buoyancy of a horizontal flux sheet. The results indicated that the imposed velocity fluctuations can destabilize the initial flux sheet, generating an upward-expanding magnetic loop, as long as it is located within or just above the convection zone but not if it is originally embedded in the higher atmosphere. They also calculated eigenfunctions for

the linear and nonlinear stability problem associated with the particular initial condition that they were using.

[17] *Kusano et al.* [1998] investigated the Parker mode ($\kappa \parallel \boldsymbol{B}$) and the interchange mode ($\kappa \perp \boldsymbol{B}$) of magnetic buoyancy instabilities with the aim to understand the emerging mechanism of magnetic loops in the solar corona. They performed nonlinear, 2-D simulations in a weak two-temperature atmosphere, which consisted of the chromosphere and corona. A sheared magnetic flux was initially embedded in the bottom of the chromosphere and random velocity perturbations were added into the equilibrium state to initiate the experiment. They found that if the field is sheared, a new instability could occur through the nonlinear process of interchange instability, which leads to the formation of magnetic loops with a mushroom-like structure. Current sheets are formed in the central lower part of the magnetic loops and reconnection occurs. As a result of reconnection, magnetic bubbles are generated and eventually are detached from the original flux sheet and shortly after they rise into the upper atmosphere.

[18] *Matsumoto et al.* [1993] used the same background stratification and similar initial conditions with the previous 2-D numerical models, to study the nonlinear evolution of EFRs [e.g., *Zirin*, 1970]. On the other hand, they considered two different types of unperturbed magnetic flux system in the lower atmosphere: a horizontal magnetic flux sheet and a horizontal magnetic flux tube. They performed 3-D MHD simulations and they found that the expansion laws, derived in the previous 2-D models, are modified because in three dimensions the expansion of the rising magnetized volume occurs also perpendicular to the fieldlines in two directions. They also found that the evolution of the EFR depends on the initial structure of the magnetic flux system. As an example, the rise velocity of the expanding magnetic loops obtained in these simulations was comparable to the observed rise velocity of arch filaments (≈ 10–15 Km s^{-1}; [e.g., *Chou and Zirin*, 1988]) when the initial flux system is a flux sheet. On the other hand, the rise speed is too small when a magnetic flux tube is considered as the unperturbed magnetic flux system at the beginning of the simulation. Another interesting feature presented in this paper is the evolution of an initial flux system, which consists of a number of isolated flux tubes. In this model the flux tubes interact with each other as they rise and they finally merge into a bundle of expanding magnetic loops. Dense filaments are formed in between the expanding magnetic loops, as the plasma slides down along the outermost expanding fieldlines, with a width about 800 Km and a rise speed ≈ 10 km s^{-1}.

[19] Finally, *Archontis et al.* [2004] showed that overdense flux can be transported into the atmosphere when the following criterion is satisfied [*Newcomb*, 1961; *Acheson*, 1979]:

$$-H_p \frac{\partial}{\partial z}(logB) > -\frac{\gamma}{2}\beta\delta + \tilde{k}_\parallel^2\left(1 + \frac{\tilde{k}_\perp^2}{\tilde{k}_z^2}\right). \quad (1)$$

[20] In the above criterion, H_p is the pressure scale height, z is the height, B is the magnetic field strength, γ is the ratio of specific heats and the plasma-β is the ratio of the gas pressure over the magnetic pressure. There are also pertur-

bations with wave vector $\boldsymbol{k}$ (where $\tilde{k}_\parallel$ and $\tilde{k}_\perp$ are the horizontal components parallel and perpendicular to the magnetic field and $\tilde{k}_z$ is the vertical component). The superadiabatic excess, δ is given by $\delta = \nabla - \nabla_{ad}$, where ∇ is the actual logarithmic temperature gradient in the equilibrium stratification and ∇_{ad} is its adiabatic value.

[21] A crucial term in the above criterion is the $\beta\delta$ term. For an isothermal layer $\delta = -0.4$. Plasma-β, on the other hand, becomes small as the magnetic pressure becomes larger than the gas pressure when the uppermost layers of the tube cross the photosphere. Thus the right-hand-side term in the above criterion becomes smaller than the left-hand-side term and the instability is launched, carrying the magnetized plasma all the way up to the corona.

2.2. Twisted Flux Tubes

[22] Observationally, there is evidence that the emerging flux bundles, which rise through the solar interior and create active regions at the photosphere, are twisted and have a coherent configuration during their rise (see section 1). Also on theoretical grounds, numerical models in two dimensions have shown that an initial twist is required for a tube to retain its coherent structure as it rises through the convection zone. A nontwisted tube splits into a pair of vortex filaments rotating in opposite directions [e.g., *Schuessler*, 1979; *Longcope et al.*, 1996]. The vortex filaments separate horizontally from each other, owing to the buoyancy force on the mass elements of the vortex filaments, and eventually the rising motion of the buoyant tube turns into a horizontal expanding motion. On the other hand, if the flux tube is twisted by a sufficient amount then the magnetic tension of the twisted fieldlines can prevent the formation of vortex filaments and the tube rises as a rigid body through the convective envelope [e.g., *Moreno-Insertis and Emonet*, 1996; *Emonet and Moreno-Insertis*, 1998].

[23] The pitch angle (Ψ) of the twisted fieldlines around a horizontal magnetic flux tube has to be above a threshold, for the transverse field to be dynamically important. If the flux tube is, both in pressure balance and thermal equilibrium with its surroundings, it will be less dense than the external plasma (and therefore will rise) by a value of

$$\frac{\Delta\rho}{\rho} = \frac{(\rho - \rho_e)}{\rho} \approx -\frac{1}{\beta}, \quad (2)$$

where ρ is the density inside the tube and ρ_e is the density of the background atmosphere.

[24] The pitch angle of the twisted field lines of such a tube has a threshold of order

$$\tan\Psi \geq \left(\frac{R}{H_p}\right)^{1/2}, \quad (3)$$

where R is the radius of the tube and H_p is the local pressure scale height.

[25] All the above models were two-dimensional. *Dorch and Nordlund* [1998] performed 3-D simulations of buoyant magnetic flux tubes ascending through a solar convection zone model, and they showed that a weak random or twisting component is sufficient to make the tube rise as a coherent structure. If the initial flux tube has a nontrivial
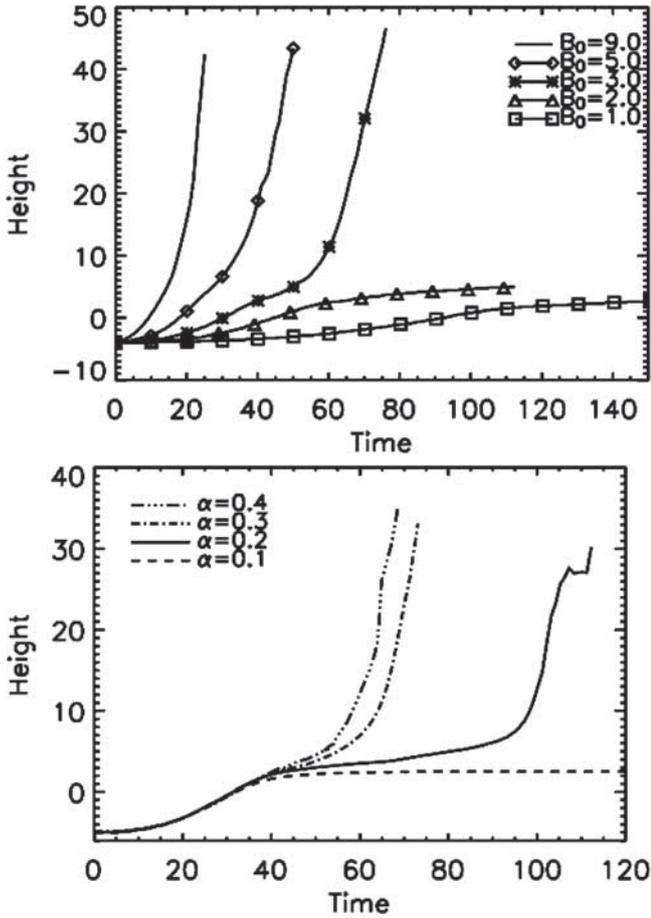
Figure 2. (top) Motion of the top of the flux tube in time for different field strengths, B_0, and fixed twist, $\alpha = 0.4$. (bottom) Fixed initial field strength and different α. For details, see *Murray et al.* [2006].

topology, the flux structure is held together for a long time and is able to keep most of its buoyancy. *Abbett et al.* [2000] found that the break-up of the flux tubes depends on the three-dimensional geometry of the problem. If the curvature of the upper part (apex) of an Ω-loop is large, the degree of fragmentation of the loop as it reaches photospheric heights is small. Also, *Abbett et al.* [2001] described how a buoyant flux tube keeps its coherent structure under the action of the Coriolis force. They also found that in the absence of forces due to convective motions, a magnetic flux tube with strong initial axial field strength will not be able to retain its cohesion. Finally, *Fan* [2001] performed 3-D simulations of arched flux tubes, which were formed from a horizontal magnetic layer due to the nonlinear growth of the undular instability. It was found that the rising arching tubes maintained their coherent structure as they moved through a significant distance inside the computational volume. Thus, although there was no net twist in the tubes, they emerged through the convection zone without significant disruption of their shape.

2.3. Dynamics of Flux Emergence

[26] The dynamical emergence of magnetic flux from the solar interior to the solar atmosphere is a big challenge in

numerical simulations. Many numerical experiments have appeared in the literature during the past few years, which yield insights into the dynamics of flux emergence and the topology of the resulting structures. Most of the numerical models have used twisted flux tubes in the lower atmosphere as an initial unperturbed configuration for the magnetic field.

[27] *Magara* [2001] investigated the emergence and expansion processes of a twisted flux tube by means of 2.5D MHD simulations. A highly stratified atmosphere was used, including a layer with increasing temperature with depth for the solar interior, an isothermal layer for the photosphere, a transition region and a high-temperature isothermal corona. Initially, the tube rises through the convection zone by magnetic buoyancy until it reaches the photosphere. However, the photosphere is strongly subadiabatic and, thus, the upper part of the tube slows down when it enters the low atmosphere. The tube is flattened and a contact surface is formed in between the rising tube and the lower atmosphere. Eventually, the Rayleigh-Taylor instability acts on this surface because dense plasma above the surface is located on top of lighter plasma, which is more magnetized. The tube emerges through the photosphere by the Rayleigh-Taylor instability and starts to expand owing to the reduction of the background atmospheric gas pressure with height and because the magnetic pressure of the emerging tube becomes larger than the outside gas pressure. However, although the tube expands into the outer atmosphere the main axis of the tube stays at the base or just below the photosphere.

[28] Recently, 3-D MHD simulations have been carried out by *Murray et al.* [2006] with the aim to understand the role of twist and magnetic field strength in shaping the emergence process. It was found that when the value of the initial field strength, B_0, and twist, α, is low the tube cannot fully emerge into the corona, but it stays in the lower atmosphere because the buoyancy instability criterion in equation (1) cannot be fulfilled. When the twist is fixed and the field strength is varied the tube experiences different magnitude of the buoyancy force, which is proportional to B_0^2. Thus, when the apex of the tube reaches the photosphere, it starts to rise into the upper atmosphere at different times (see Figure 2, top). When the field strength is fixed and the amount of twist is modified the tube rises with different configuration for each α (see Figure 2, bottom). More precisely, if $\alpha = 0.1$ the tube flattens out at photospheric heights. For $\alpha = 0.2$ the tube emerges at two side locations because the draining of plasma from the upper part of the tube is more efficient at these locations. For larger values of twist the tube rises and expands into the corona adopting a dome-like structure. In addition, it has been found [e.g., *Magara*, 2007] that the expansion of an initially high twisted flux tube below the photosphere produces coronal loops with sigmoidal structure while a tube with a weak twist produces expanded coronal loops with no clear signature of sigmoidal structures.

[29] *Fan* [2001] used a density deficit profile to initiate the rise of a flux tube from below the photosphere into the solar atmosphere. The tube is more buoyant at the middle than at the ends and evolves into an Ω-shaped loop as it rises and expands in the atmosphere. At the beginning of the emergence a bipolar region is formed with a north-south
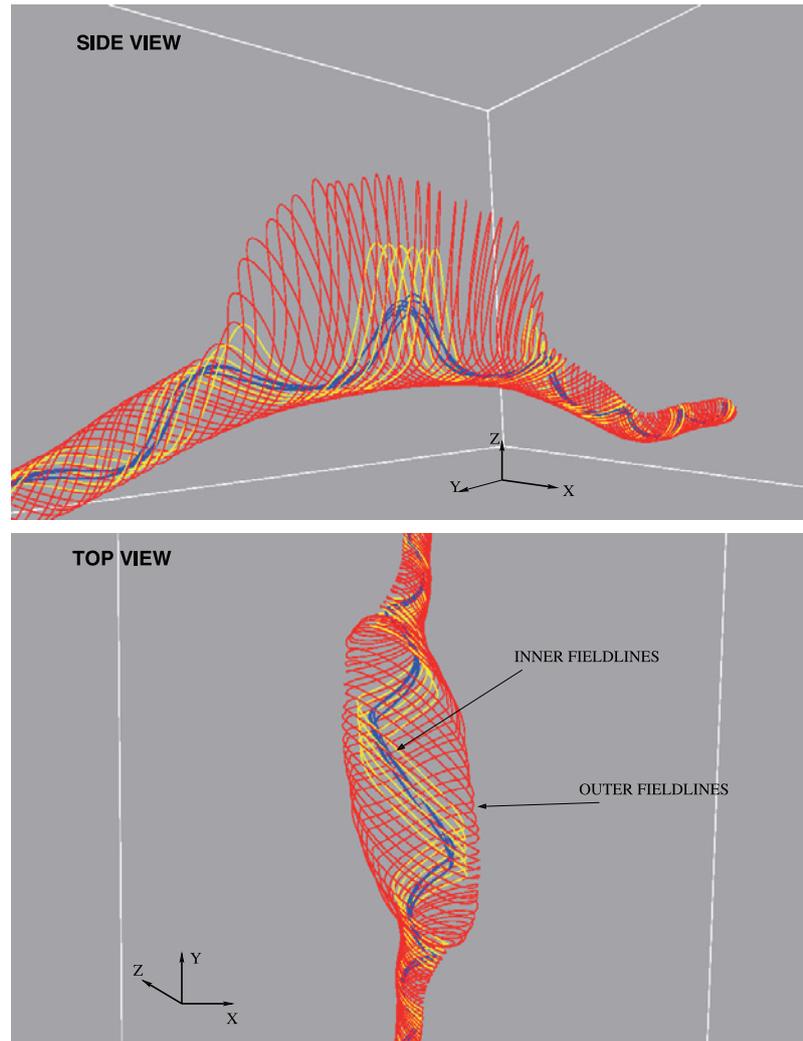
Figure 3. (top) Outermost expanding fieldlines, which adopt a fan-like shape in the corona. (bottom) Fieldlines at lower heights inside the rising flux system, which are less twisted. The innermost fieldlines have an S-shape configuration.

orientation but later on the sunspots move toward an east-west orientation, parallel to the axis of the emerging tube. In fact, there is a shear horizontal velocity flow on the two sides of the neutral line, which agrees well with the observations of emerging active regions by *Strous et al.* [1996] and was diagnosed as a shear Alfvén wave in the simulations by *Manchester* [2001].

[30] The expansion of an emerging twisted flux tube into the higher levels of the atmosphere occurs in a runaway fashion. This is because the distribution of the magnetic pressure with height above the photosphere in the expanding volume of magnetized plasma is much larger than the gas pressure. The expansion takes place in all three directions: in fact, it expands faster in the horizontal directions than in the vertical directions [e.g., *Archontis et al.*, 2004]. Because of mass conservation, the fast rise and expansion of the upcoming field cause strong downflows along the periphery of the emerging fieldlines [e.g., *Fan*, 2001; *Magara and Longcope*, 2003; *Archontis et al.*, 2004].

[31] The outermost fieldlines of the rising flux system adopts a fan-like shape as it expands into the higher levels of the atmosphere [e.g., *Fan*, 2001; *Magara*, 2001; *Archontis et al.*, 2004]. The inner fieldlines do not follow the same dynamical evolution and adopt shapes, which could appear as sigmoids in the solar corona [e.g., *Magara*, 2004]. Figure 3 shows the shape of different sets of fieldlines in a flux emergence experiment by *Archontis et al.* [2004]. It is worthwhile mentioning that, in most of the above referenced models, the fieldline on the main axis of the initial flux tube rises, but very slowly and thus does not really emerge into the corona. This is, probably, because the curvature of this fieldline is very small and the dense plasma does not drain sufficiently to make the buoyancy effects to lift up the axis of the tube. However, fieldlines that are initially located just above the main axis of the tube can emerge into the photosphere because their shape is such that enables the magnetic pressure force to lift them upward. In general, the fieldlines that emerge may be classified into two categories: expanding fieldlines and undulating fieldlines [e.g., *Magara and Longcope*, 2003]. Also, the outermost fieldlines become more twisted, than the inner fieldlines, as they expand and drive horizontal

shearing motions in the low atmosphere that contribute to the ejection of magnetic energy and helicity to the outer atmosphere [e.g., *Magara and Longcope*, 2003].

[32] *Abbett and Fisher* [2003] coupled a subsurface model of an emerging flux tube to a three dimensional model corona. They first modeled the rise of a buoyant magnetic flux system in a stable stratified convection zone. Then, they used the time-dependent vector fields and scalar variables in the upper subphotospheric boundary to drive a 3-D model corona. The simulations showed that the time-dependent flows below the surface play a crucial role on the dynamic evolution and subsequent morphology of an emerging magnetic structure. At the beginning of the simulation, the magnetic field that surrounds the emerging system is not force-free. As the emergence proceeds and the vector fields at the photosphere evolve, the overlying magnetic field relaxes to a more force-free state. The development of models where high-resolution magnetograms of active regions will be used to drive dynamic models of the solar corona is an interesting future task.

[33] Most of the MHD models described above do not include a realistic convection zone below the photosphere. A first attempt to include the cell-like convection structure in a flux emergence model was reported by *Amari et al.* [2005]. In this model, a twisted flux tube is kinematically raised by a convection cell in the convection zone and evolves in a full MHD way after its emergence into the corona. A photospheric layer is located in between the convection zone and the corona. The effective resistivity in this layer is larger than in the other two layers (resistivity layer model (RLM)). It is shown that electric current and magnetic flux emerge through the resistive photospheric layer into the corona. The emergence of flux leads to an arcade-like configuration at the beginning of simulation, while later on the emerging flux evolves more rapidly adopting the shape of a flux rope.

[34] Also, *Cheung et al.* [2007] modeled magnetic flux emergence in granular convection performing radiative MHD simulations. They showed that convection influences the evolution of the emerging fields before and after they reach photospheric heights. Subphotospheric upflows can support the rising motion of some segments of the emerging tube while downflows may suppress the emergence of other segments. It was also found that flux tubes with small longitudinal flux, are not highly buoyant and cannot rise coherently against the convective plasma motions. On the other hand, big flux tubes with large axial flux are able to rise and eventually emerge at the surface disturbing the granulation pattern owing to their dynamical horizontal expansion.

[35] *Leake and Arber* [2006] performed 2.5D MHD simulations to simulate the emergence of a twisted flux tube into the solar atmosphere taking into account two processes, which were not included in previous models: the partial ionization of certain regions of the solar atmosphere and the thermal conduction as a heat transfer mechanism. It was found that when the dense plasma rises and expands into the atmosphere is heated to its original photospheric temperature rather than being cooled (adiabatic expansion). The inclusion of partially ionized plasma in the chromosphere yields more rapid emergence and expansion of the rising field and a greater amount of flux into the

corona. An important question then is if the resulting coronal magnetic field is force-free. This is equivalent to say that the current is parallel to the magnetic field and there are no cross-field currents. It is found that when the magnetic field emerges through a partially ionized plasma, the majority of the cross-field current is destroyed, and thus the coronal magnetic field becomes force-free.

[36] *Isobe et al.* [2007] used 2-D simulations to study the emergence of a very long flux sheet from a superadiabatically stratified layer that represents the convection zone into the isothermal corona. A random velocity perturbation is given in the initial flux sheet to excite the Parker buoyancy instability. A number of loops are formed out of the current sheet. They rise, expand and eventually reconnect at the lower atmosphere. Temperature in the reconnection outflows is enhanced compared to the background plasma. This may account for the manifestation of Ellerman bombs. During the rise of the loops, dense material is accumulated at the valleys between neighboring loops. This dense plasma is squeezed because of expansion of the rising loops and creates elongated structures of cold and dense plasma, which eventually is ejected to the outer atmosphere because of reconnection between the loops. At the end of the experiment a very large loop of magnetic fieldlines is formed through successive reconnection of side-by-side emerging loops at different heights of the atmosphere (see Figure 4). This process is similar to those in the resistive emergence model proposed by *Pariat et al.* [2004].

3. Emergence Into a Magnetized Corona

[37] Detailed observations of magnetic fields in and around active regions have shown that the emergence of new magnetic flux causes noticeable changes in the topology and geometry of the magnetic flux systems in the corona: first, by creating magnetic links to preexisting magnetic fields [e.g., *Longcope et al.*, 2005] and second, by triggering the ejection of collimated, high-velocity and high-temperature outflows of plasma observed in soft X rays by the YOHKOH satellite [*Shimojo et al.*, 1996; *Shimojo and Shibata*, 2000]. These processes (change of connectivity and emission of jets) can take place through magnetic fieldline reconnection whenever an upcoming and a preexisting magnetic flux system come into contact. Reconnection is also responsible for significant localized energy release and the formation of a network of hot (with a temperature of few million degree Kelvin) plasma structures observed by the YOHKOH soft X-ray telescope. Thus it is very important to understand the dynamical interaction between emerging flux and preexisting coronal magnetic field and numerical simulations could provide a physical understanding of the afore-mentioned process.

[38] The first models, which studied the presence of a coronal magnetic field and the interaction with the emerging flux system were two-dimensional. *Shibata et al.* [1989a] included a simple horizontal field in the corona and a horizontal magnetic flux sheet below the photosphere to simulate the emerging flux. Then, they performed a series of experiments changing the direction of the overlying field [*Yokoyama and Shibata*, 1995, 1996]. In the parallel-field case (where the direction of the ambient field is parallel to the direction of the uppermost fieldlines of the rising loop
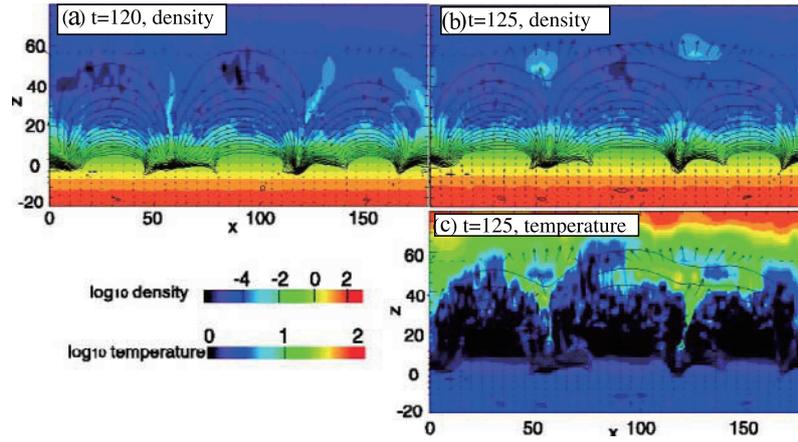
Figure 4. (a, b) Dense, cool material ejected together with (c) hot plasma from the reconnection sites in a flux emergence model by *Isobe et al.* [2007]. At the final stage of the experiment (Figure 4c), reconnected fieldlines join all the individual emerging loops and rise into the outer solar atmosphere.

upon contact) the rise velocity of the buoyant loop is drastically suppressed, while in the antiparallel-field case reconnection occurs and the rising motion of the emerging loop does not slow down. In the latter case, multiple magnetic islands are created inside the current sheet, at the interface between the two flux systems. These magnetic islands are cool and dense structures, which were originally formed in the lower atmosphere but are carried up within the current sheet. Eventually, they are ejected sideways from the current sheet and they may be observed as $H\alpha$ surges (cool jets). Also, the plasma behind the magnetic islands is heated by Joule dissipation to a few million degree Kelvin and is ejected along the reconnected fieldlines to both sides of the current sheet, creating a pair of hot jets. In the case of an oblique ambient field, one jet is ejected upward while the second one moves downward where it collides with smaller emerging loops, which have been formed from the original magnetic flux sheet below the photosphere, creating a fast MHD shock. The collision compresses the plasma and the temperature at the top of the loops increases. These hot loops may account for some observations of microflares or bright loops, which are found slightly shifted from the site of the fast and hot outflows. In the case of a vertical coronal field the jets are emitted in the vertical direction. However, the general structure and topology around the emerging region is similar to the oblique-field case. The above numerical models reproduce some of the features presented in the emerging flux model for the solar flare phenomenon suggested by *Heyvaerts et al.* [1977], the standard CSHKP model of solar flares (originally proposed by *Carmichael* [1964], *Sturrock* [1966], *Hirayama* [1974], and *Kopp and Pneuman* [1976]) that explains their observable features on the basis of magnetic reconnection and the revision of the CSHKP model based on YOHKOH observations by *Shibata et al.* [1995].

[39] The 3-D interaction between an emerging magnetic flux tube and a large-scale horizontal coronal field was studied in a series of papers by *Archontis et al.* [2004], *Galsgaard et al.* [2005], *Archontis et al.* [2005], *Archontis et al.* [2006], and *Galsgaard et al.* [2007]. Experiments were performed with the direction of the horizontal ambient field being changed from parallel to antiparallel in steps of

45 degrees. In the antiparallel case, a dome-like current concentration is formed at the interface when the two magnetic flux systems come into contact. Eventually, the current surface concentrates into a curved sheet, which is contained in a vertical midplane that is rotated almost 5 degrees away from the initial axis of the flux tube. In fact, the midplane of the current arch is rotated by an angle that increases with the relative horizontal angle between the two flux systems into contact. At the initial stage of contact of the two systems the configuration of the magnetic field vector across the current sheet is similar to the classical 2-D X-type configuration. However later on, the orientation of the field across the current sheet resembles a rotational discontinuity such that the magnetic field never goes through a null point. The above change in the profile of the magnetic field vector has important consequences for the reconnection, which occurs in a full 3-D manner not only at the top of the interface but all along the current concentration. It was also found that magnetic fieldlines reconnect in a continuous fashion while they are linked to the diffusion region and that many fieldlines that belong initially to the rising flux tube may reconnect more than once (multiple reconnection events). As a result of the reconnection process, the domain below the photosphere and the coronal domain, which were not joint at the beginning of the experiments, become linked to each other through the reconnected fieldlines. In fact, almost 75% of the emerging flux becomes reconnected to the coronal field in the experiment with antiparallel magnetic flux systems.

[40] The three-dimensional current arch is the region where the Joule dissipation has a significant impact on the heating of the plasma. It was found that the temperature in the antiparallel case could be as high as 10^7 K but becomes lower for the cases that are not so favorable for effective reconnection. In any case, it is likely that current sheets, formed between emerging and ambient magnetic fields, may constitute a source of heating for the solar corona. Similar to the 2-D models, the above 3-D experiments showed that a pair of hot and high-velocity (with the peak velocities typically reaching the local Alfvén velocity) jets are emitted sideways from the current sheet (see Figure 5). The jets do not look like horizontal thin layers (as in 2-D
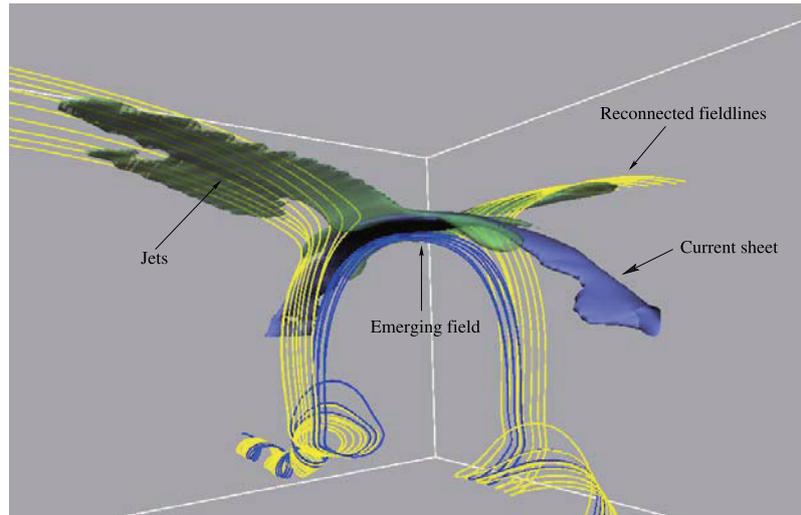
Figure 5. Curved current sheet (blue), the high-velocity outflows (green), and the reconnected fieldlines (yellow) in a 3-D flux emergence experiment [*Archontis et al.*, 2005] where the emerging field reconnects with an ambient field with antiparallel orientation.

models) but they are curved all along the sides of the current sheet.

[41] Finally, the formation and evolution of 3-D plasmoids was investigated by *Archontis et al.* [2006]. The formation of the plasmoids was possibly due to the tearing mode instability. Two phases were apparent during the evolution of the system. In the first phase, the plasmoids had the shape of solenoids lying along the current sheet with their fieldlines connecting to the subphotospheric field or with the coronal field. At this stage, they are cool (of the order of 10^4 K) and dense and their velocity upon ejection is as high as 15 Km s^{-1}. In the second phase, where the magnetic field across the current sheet undergoes through a rotational discontinuity, the fieldlines in the plasmoid become much less tightly wounded. The plasmoids now confine hot (of the order of 10^6 K) and less dense plasma while they are ejected out of the current sheet with much higher speed ($\approx$150 Km s^{-1}). The ejection of cool plasmoids may be compatible with observations of Hα or Hβ surges. It is also possible that the UV coil-like structures observed in filament eruptions may account for the appearance of helical flux strands at the second phase of the experiments. We should mention that the energy equation used in these experiments is adiabatic. Also, like other simulations, there is no radiative transfer or explicit coronal heating in the calculations. However, ohmic and viscous dissipation are included in the above mentioned experiments.

[42] *Miyagoshi and Yokoyama* [2004] studied the emergence of a flux sheet, by simulating the undular mode of the magnetic buoyancy instability, into an antiparallel ambient field by means of 2-D MHD simulations including heat conduction. They found that two different types of jets are formed around the emerging flux region: the classical reconnection jets which are formed because of reconnection between the two antiparallel flux systems and high-density evaporation jets. The basic mechanism for the formation of the evaporation jets is as follows: during reconnection, the magnetic energy is converted into thermal energy, which is transported to the chromosphere along the magnetic field lines by heat conduction. Chromospheric evaporation occurs and dense plasma rises along the reconnected magnetic field creating a secondary pair of jets. However, the temperature of these jets was found to be low because the cooling by conduction was more efficient than the heating by magnetic reconnection.

[43] *Isobe et al.* [2005] performed 3-D MHD simulations to study the interaction of an emerging flux sheet with an oblique ambient field. They found that thin current sheets and intermittent heating occurs at the interface between the two magnetic flux systems, as a result of the magnetic Rayleigh-Taylor instability. Dissipation of the filamentary current sheets leads to the heating of the plasma around dense filaments, forming a system of hot and cold loops. The Rayleigh-Taylor instability and the fast magnetic reconnection are coupled in a nonlinear way, leading to intermittent, patchy reconnection. These results may explain the intermittent nature of coronal heating and the patchy brightenings in solar flares.

[44] *Archontis et al.* [2007] performed 2.5D MHD simulations of a pair of twisted flux tubes as they emerge from the subphotospheric layers into the solar corona. A key feature in this simulation is that the first flux tube emerges, expands and creates a non uniform (in strength and direction) coronal field that the second tube emerges into. A series of dynamical phenomena is produced by this model in a self-consistent manner. A current sheet, with an arch-like shape, is formed when the two systems start to interact. Plasmoids are formed inside the sheet because of resistive instabilities and when they are ejected out of the current arch they carry cool and dense material with them. As reconnection proceeds, the magnetic topology changes dramatically: there are now four different flux systems separated by a vertical current sheet, a configuration that has been extensively used as an initial condition in many models for studying flares and eruptions of filaments. Indeed, the general characteristics of the arcade structure below the vertical current sheet may account for a compact
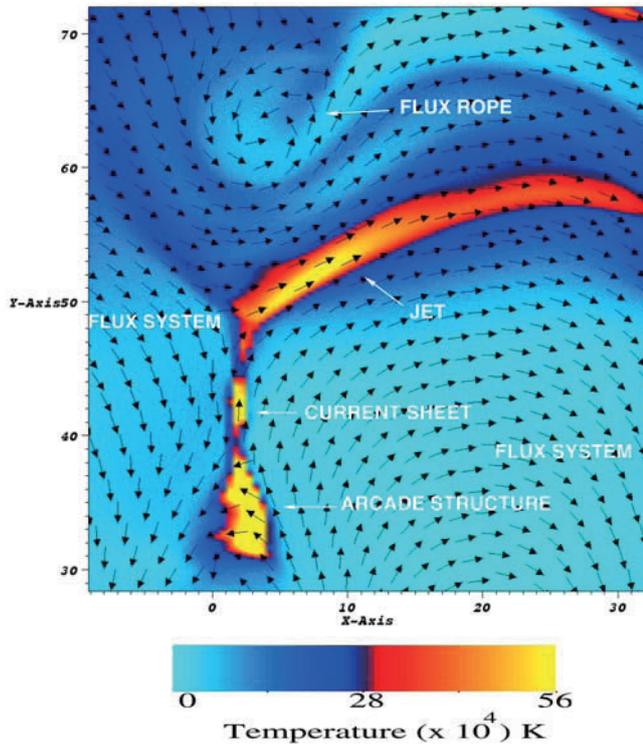
Figure 6. Two flux systems come into contact. Eventually, a vertical current sheet is formed between them, a flux rope is ejected into the outer atmosphere, and reconnection jets heat the arcade-like structure at the base of the current sheet and the top-loop of the emerging system on the right side of the current sheet.

or an arcade flare associated with flux emergence. Also, hot plasma (jet) is ejected upward from the current sheet and is moving along the reconnected fieldlines at the top of the secondary emerging system (see Figure 6). The temperature enhancement along the upper part of the secondary emerging flux system may account for a loop brightening.

4. Flux Emergence and Eruptions

[45] The emergence of magnetic flux from the solar interior to the high atmosphere of the Sun may be connected with solar eruptions, such as flares, filaments and Coronal Mass Ejections. In fact, it has been shown [*Sterling and Moore*, 2005] that even small-scale emergence of flux can change the magnetic topology of a preexisting active region and trigger an eruption.

[46] Eruptions and powerful explosions (e.g., fast CMEs) is one of the basic topics of research in space weather physics. However, details about how these eruptions are triggered and propagate in the three-dimensional space are still unknown. For details about the theory and the models of CMEs the reader is referred to the reviews by *Klimchuck* [2001], *Forbes et al.* [2006] and *Mikic and Lee* [2006]. In the following, we outline results of numerical models that incorporate flux emergence to study eruptions of magnetic flux into the outer atmosphere of the Sun.

[47] *Chen and Shibata* [2000] proposed an emerging flux trigger mechanism for CMEs using 2-D MHD simulations.

Their model consists of a quadrapolar field in a two-dimensional Cartesian plane. A detached flux rope is located above the quadrapolar field of the CME source. Two cases were studied. In the first case, magnetic flux emerges within the filament channel with direction opposite to the ambient coronal field, reconnection occurs below the detached flux rope that leads to partial magnetic cancellation and the rise of the flux rope because of loss of equilibrium. In the second case, the flux emerges at the right side outside of the filament channel. Reconnection first occurs between the emerging flux system and the outer fieldlines of the channel but it eventually proceeds in the internal layers of the filament system and, thus, it leads to the eruption of the flux rope similar to the first case. In both cases, a vertical current sheet is formed below the flux rope because of the reconnection between the emerging flux and the ambient field. The upward reconnection jet inside the filament channel pushes the flux rope toward the outer atmosphere while below the current sheet, a cusp-shaped structure with high temperature is formed. The fast ejection of the flux rope may account for a CME ejection while the cusp-shaped structure has the characteristics of LDE (long duration events) flares.

[48] *Shiota et al.* [2005] used the same initial magnetic field configuration as *Chen and Shibata* [2000] but they also included the effect of heat conduction and discussed the differences. They found that the dynamical properties (such as velocity and magnetic fields) are very similar in the two models but the thermal properties (e.g., temperature, density) are different. For example, the temperature in the reconnection region and within the reconnection outflows becomes lower. Also, the current density at the X-point in the current sheet is larger and the width of the sheet is thinner. Also, they synthesized soft X-ray images from the density and temperature in the numerical results and compared them with YOHKOH observations. They showed the Y-shape of the slow shocks associated with the reconnection, the cusp-shaped arcades below the current sheet and the dimming above them, and a bright feature at the top of the arcades that may correspond to the backbone of flare arcades observed by YOHKOH. *Dubey et al.* [2006] extended the model of *Chen and Shibata* [2000] by including the effects of gravity, spherical geometry and a stratified ambient medium in 2.5D simulations. They also studied how the rate and the total amount of emerging flux affects the velocity of the resulting CME-like structure (flux rope). It was found that the latter factor plays a more crucial role and the obtained flux rope velocities achieve higher values. However, all these models cannot reproduce the ejection of very fast CMEs.

[49] The emergence of a twisted magnetic flux tube into a preexisting potential magnetic arcade in the corona has been investigated in the work by *Fan and Gibson* [2003], *Gibson and Fan* [2006] and *Fan and Gibson* [2006]. First, it was found that a strong electric current concentration with an inverse-S shape is formed as the emerging tube develops substantial writhing as a result of the kink instability. The three-dimensional structure of the current is consistent with the shape of X-ray sigmoids. Another interesting result is that the emerging flux rope split into two parts during its eruption and interaction with the ambient field. One part is being expelled into the outer atmosphere while the other

stays behind. Such a partial expulsion is consistent with observations of CMEs. The full eruption of a kink-unstable coronal magnetic flux rope, anchored in the photosphere, was also studied by *Török and Kliem* [2005]. Their results were in good agreement with the helical shape and the rise profile of a very fast CME. They concluded that the helical kink instability of a flux rope may be the mechanism that triggers many solar eruptions.

[50] *Manchester et al.* [2004] investigated how a part of a twisted flux rope can erupt into the corona. They used similar initial conditions to *Fan* [2001] but they decreased the length of the buoyant section of the rising tube, so that the draining of the mass becomes more efficient. It was found that the magnetic fieldlines inside the expanding volume of the magnetized plasma are stretched as the tube rises and eventually reconnect forming an internal current sheet with an S-shape. Because of reconnection and to self-induced shearing motions in the lower atmosphere, the upper half of the flux rope is detached from the lower half and erupts into the corona. The general evolution of the above system is consistent, at least qualitatively, with observations of CMEs in which X-ray sigmoids may appear after the onset of the eruption.

5. Discussion and Conclusions

[51] Some of the most intense episodes of the Sun's activity are related to the dynamic process of magnetic flux emergence from the solar interior into the low-density solar corona. Over the past few years, numerical models of flux emergence have been used to study and explain some of the observed properties of solar active regions and eruptive events. In fact, some of the experiments have been quite successful in reproducing, mostly in a qualitative manner, the appearance and evolution of bipolar regions in the photosphere, the formation of sigmoidal structures strongly reminiscent of the X-ray sigmoids observed in the Sun, the morphology of the magnetic fields in solar coronal loops, the interaction between neighboring emerging flux systems, the ejection of plasmoids and cool ($H\alpha$ surges) and hot (X-ray reconnection) outflows from sites of strong current concentration, small-scale brightening events in the lower atmosphere (e.g., Ellerman bombs), the formation of flares (arcade flares, compact flares, etc.) and finally the eruption of magnetic flux ropes (e.g., CMEs) in the outer atmosphere of the Sun, which are very important components of the Space Weather system. Thus the results of these numerical experiments help us to construct a better picture of the solar magnetic activity and the agreement between numerical experiments and observations is very promising, but there is much that remains to be done and there is still much to learn about flux emergence and eruptions in the Sun.

[52] During the past few years, there has been a significant number of observations with increased resolution but there is still no complete understanding of: first, the physics behind the process of magnetic flux emergence from the deep solar interior to the outer solar atmosphere and second, how the emerging field interacts with preexisting magnetic structures leading to dynamical ejections of plasma and eruptions of flux. The evolution of such systems constitute a very intricate problem (given the large range of length and timescales, temperature, etc., involved in the process of flux

emergence) and, thus, observations are expected to provide us with information of selected aspects, which then have to fit together to build up a consistent physical picture of flux emergence in a complete manner. Progress is achieved by the recent solar mission of Hinode, which takes amazing high-resolution observations providing detailed description of the small-scale element and the large-scale magnetic field distribution and their interactions.

[53] On the other hand, there are theoretical difficulties since many numerical models are idealized. For example, the choice of the initial conditions in the majority of the experiments is not strictly in agreement with what has been shown about the evolution of the magnetic fields in realistic convection zone models. Also, there are problems if the initial entropy distribution below the photosphere is not well defined, as it influences the buoyancy and the emergence of the magnetic flux system to the photosphere. In addition, the coupling between convection and small-scale flux emergence into the corona has not been investigated so far.

[54] One of the major limitations of the 3-D simulations of flux emergence into the atmosphere is the lack of an adequate description of the thermal behavior of the system. Most of the existing simulations have assumed an adiabatic energy equation (ohmic and viscous dissipation terms are usually included) and this means that the rapid expansion into the corona is accompanied by rapid adiabatic cooling of the emerging plasma. This is not observed on the Sun and, thus, additional physics must be included to model the energy equation (e.g., radiative transfer in photosphere and chromosphere). In the corona, thermal conduction is a dominant term in the energy balance and field-aligned thermal conductivity must be also included along with optically thin radiation.

[55] Also, emergence of magnetic flux within active regions is often associated with flares and CMEs, but there is no generally accepted explanation, so far, of how the magnetic field is stressed in the corona and what the triggering processes are for the onset of the eruptions. Also, there is no definitive modeling (including realistic geometry, kinetic effects and description of the thermodynamics of the plasma) of their three-dimensional magnetic structure and evolution into the heliosphere. Thus modeling a realistic CME and producing results that may be comparable with observations has not yet been achieved.

[56] It is worthwhile to mention that a complete study of flux emergence in the solar atmosphere is a very fascinating and computationally challenging problem. It is indisputable that progress have gained through a series of numerical experiments that have investigated the various physical processes separately. We expect that the individual results will be combined and will lead to important advances about the nature and the dynamics of emerging flux systems in the near future. Finally, detailed global 3-D numerical experiments using high-performance computing and simultaneous high-resolution observations of the magnetic field structures at different heights of the solar atmosphere, are needed to advance our understanding further about the process of magnetic flux emergence and how it is associated with solar eruptions.

References

Abbett, W. P., and G. H. Fisher (2003), A coupled model for the emergence of active region magnetic flux into the solar corona, *Astrophys. J.*, *582*, 475.

Abbett, W. P., G. H. Fisher, and Y. Fan (2000), The three-dimensional evolution of rising, twisted magnetic flux tubes in a gravitationally stratified model convection zone, *Astrophys. J.*, *540*, 548.

Abbett, W. P., G. H. Fisher, and Y. Fan (2001), The effects of rotation on the evolution of rising omega loops in a stratified model convection zone, *Astrophys. J.*, *546*, 1194.

Acheson, D. J. (1979), Instability by magnetic buoyancy, *Sol. Phys.*, *62*, 23.

Amari, T., J. F. Luciani, and J. J. Aly (2005), Non-current-free coronal closure of subphotospheric MHD models, *Astron. Astrophys.*, *426*, 1047.

Antiochos, S. K., C. R. DeVore, and J. A. Klimchuk (1999), A model for solar coronal mass ejections, *Astrophys. J.*, *510*, 485.

Archontis, V., F. Moreno-Insertis, K. Galsgaard, A. W. Hood, and E. O'Shea (2004), Emergence of magnetic flux from the convection zone into the corona, *Astron. Astrophys.*, *426*, 1047.

Archontis, V., F. Moreno-Insertis, K. Galsgaard, and A. W. Hood (2005), The three-dimensional interaction between emerging magnetic flux and a large-scale coronal field: Reconnection, current sheets, and jets, *Astrophys. J.*, *635*, 1299.

Archontis, V., K. Galsgaard, K. Moreno-Insertis, and A. W. Hood (2006), Three-dimensional plasmoid evolution in the solar atmosphere, *Astrophys. J. Lett.*, *645*, 161.

Archontis, V., A. W. Hood, and C. Brady (2007), Emergence and interaction of twisted flux tubes in the Sun, *Astron. Astrophys.*, *466*, 367.

Brun, A. S. (2007), Towards using modern data assimilation and weather forecasting methods in solar physics, *Astron. Nachr.*, *328*, 329.

Canfield, R. C., H. S. Hudson, and D. E. McKenzie (1999), Sigmoidal morphology and eruptive solar activity, *Geophys. Res. Lett.*, *26*(6), 627.

Canfield, R. C., H. S. Hudson, and A. A. Pevtsov (2000), Sigmoids as precursors of solar eruptions, *IEEE Trans. Plasma Sci.*, *28*, 1786.

Carmichael, H. (1964), A process for flares, in *Proceedings of AAS-NASA Symposium on the Physics of Solar Flares*, edited by W. N. Hess, p. 451, NASA, Washington, D. C.

Chen, P. F., and K. Shibata (2000), An emerging flux trigger mechanism for coronal mass ejections, *Astrophys. J.*, *545*, 524.

Cheung, M. C. M., and M. Schüssler (2007), Magnetic flux emergence in granular convection: Radiative MHD simulations and observational signatures, *Astron. Astrophys.*, *467*, 703.

Chou, D., and H. Zirin (1988), The vertical structure of arch filament systems in solar emerging flux regions, *Astrophys. J.*, *333*, 420.

Cole, D. G. (2003), Space weather: Its effects and predictability, *Space Sci. Rev.*, *107*, 295.

Dorch, B., and A. Nordlund (1998), Numerical 3D simulations of buoyant magnetic flux tubes, *Astron. Astrophys.*, *338*, 329.

Dubey, G. B., B. van der Holst, and S. Poedts (2006), The initiation of coronal mass ejections by magnetic flux emergence, *Astron. Astrophys.*, *459*, 927.

Emonet, T., and T. Moreno-Insertis (1998), The physics of twisted magnetic tubes rising in a stratified medium: Two-dimensional results, *Astrophys. J.*, *492*, 804.

Fan, Y. (2001), The emergence of a twisted Ω-tube into the solar atmosphere, *Astrophys. J.*, *554*, 1.

Fan, Y. (2004), Magnetic fields in the solar convection zone, *Living Rev. Solar Phys.*, *1*, 1.

Fan, Y., and S. E. Gibson (2003), The emergence of a twisted magnetic flux tube into a preexisting coronal arcade, *Astrophys. J. Lett.*, *589*, 105.

Fan, Y., and S. E. Gibson (2006), On the nature of the X-ray bright core in a stable filament channel, *Astrophys. J. Lett.*, *641*, 149.

Fisher, G. H., Y. Fan, D. W. Longcope, and A. A. Pevtsov (2000), The solar dynamo and emerging flux, *Solar Phys.*, *192*, 119.

Forbes, T. G., et al. (2006), CME theory and models, *Space Sci. Rev.*, *123*, 251.

Galloway, D. J., and N. O. Weiss (1981), Convection and magnetic field in stars, *Astrophys. J.*, *243*, 945.

Galsgaard, K., F. Moreno-Insertis, V. Archontis, and A. Hood (2005), A three-dimensional study of reconnection, current sheets, and jets resulting from magnetic flux emergence in the Sun, *Astrophys. J.*, *618*, 153.

Galsgaard, K., V. Archontis, F. Moreno-Insertis, and A. Hood (2007), The effect of the relative orientation between the coronal field and new emerging flux: I. Global properties, *Astrophys. J.*, *666*, 1.

Gibson, S. E., and Y. Fan (2006), The partial expulsion of a magnetic flux rope, *Astrophys. J. Lett.*, *637*, 65.

Hagenaar, H. J. (2001), Ephemeral regions on a sequence of full-disk Michelson Doppler Imager magnetograms, *Astrophys. J.*, *555*, 448.

Heyvaerts, J., E. R. Priest, and D. M. Rust (1977), An emerging flux model for the solar flare phenomenon, *Astrophys. J.*, *216*, 123.

Hirayama, T. (1974), Solar quiescent prominence and the energy balance, *Solar Phys.*, *43*, 323.

Hughes, D. W., and M. R. E. Proctor (1998), Magnetic fields in the solar convection zone: Magnetoconvection and magnetic buoyancy, *Annu. Rev. Fluid Mech.*, *20*, 187.

Isobe, H., T. Miyagoshi, K. Shibata, and T. Yokoyama (2005), Filamentary structure on the Sun from the magnetic Rayleigh-Taylor instability, *Nature*, *434*, 478.

Isobe, H., D. Tripathi, and V. Archontis (2007), Ellerman bombs and jets associated with resistive flux emergence, *Astrophys. J. Lett.*, *657*, 53.

Kaisig, M., T. Tajima, K. Shibata, S. Nozawa, and R. Matsumoto (1990), Nonlinear excitation of magnetic undular instability by convective motion, *Astrophys. J.*, *358*, 698.

Keller, C. U. (1992), Resolution of magnetic flux tubes on the Sun, *Nature*, *359*, 307.

Klimchuk, J. A. (2001), Theory of coronal mass ejections, in *Space Weather, Geophys. Monogr. Ser.*, vol. 125, edited by P. Song, H. J. Singer, and G. Siscoe, pp. 143–157, AGU, Washington, D. C.

Kopp, R. A., and G. W. Pneuman (1976), Magnetic reconnection in the corona and the loop prominence phenomenon, *Solar Phys.*, *50*, 85.

Kusano, K., K. Moriyama, and T. Miyoshi (1998), Nonlinear dynamics of magnetic buoyancy instabilities in a sheared magnetic field, *Phys. Plasmas*, *5*, 7.

Leake, J. E., and T. D. Arber (2006), The emergence of magnetic flux through a partially ionised solar atmosphere, *Astron. Astrophys.*, *450*, 2.

Leka, K. D., R. C. Canfield, A. N. McClymont, and L. van Driel-Gesztelyi (1996), Evidence for current-carrying emerging flux, *Astrophys. J.*, *462*, 547.

Longcope, D. W., G. H. Fisher, and S. Arendt (1996), The evolution and fragmentation of rising magnetic flux tubes, *Astrophys. J.*, *464*, 999.

Longcope, D. W., D. E. McKenzie, J. Cirtain, and J. Scott (2005), Observations of separator reconnection to an emerging active region, *Astrophys. J.*, *630*, 596.

Magara, T. (2001), Dynamics of emerging flux tubes in the Sun, *Astrophys. J.*, *549*, 608.

Magara, T. (2004), A model for dynamic evolution of emerging magnetic fields in the Sun, *Astrophys. J.*, *605*, 480.

Magara, T. (2007), Dynamical and topological features of photospheric and coronal activities produced by flux emergence in the Sun, *Astrophys. J.*, *653*, 1499.

Magara, T., and D. W. Longcope (2003), Injection of magnetic energy and magnetic helicity into the solar atmosphere by an emerging magnetic flux tube, *Astrophys. J.*, *586*, 630.

Manchester, W., IV (2001), The role of nonlinear Alfvén waves in shear formation during solar magnetic flux emergence, *Astrophys. J.*, *547*, 503.

Manchester, W., IV, T. Gombosi, D. DeZeeuw, and Y. Fan (2004), Eruption of buoyantly emerging magnetic flux rope, *Astrophys. J.*, *610*, 588.

Martens, P. C. H. (2003), YOHKOH-SXT observations of reconnection, *Adv. Space Res.*, *32*, 905.

Matsumoto, R., T. Tajima, K. Shibata, and M. Kaisig (1993), Three-dimensional magnetohydrodynamics of the emerging magnetic flux in the solar atmosphere, *Astrophys. J.*, *462*, 547.

Mikic, Z., and M. A. Lee (2006), An introduction to theory and models of CMEs, shocks and solar energetic particles, *Space Sci. Rev.*, *123*, 57.

Miyagoshi, T., and T. Yokoyama (2004), Magnetohydrodynamic simulation of solar coronal chromospheric evaporation jets caused by magnetic reconnection associated with magnetic flux emergence, *Astrophys. J.*, *614*, 1042.

Moreno-Insertis, F., and T. Emonet (1996), The rise of twisted magnetic tubes in a stratified medium, *Astrophys. J.*, *472*, 53.

Murray, M. J., A. W. Hood, F. Moreno-Insertis, K. Galsgaard, and V. Archontis (2006), 3D simulations identifying the effects of varying the twist and field strength of an emerging flux tube, *Astron. Astrophys.*, *460*, 909.

Newcomb, W. (1961), Convective instability induced by gravity in a plasma with a frozen-in magnetic field, *Phys. Fluids*, *4*, 391.

Nordlund, A., A. Brandenburg, R. L. Jennings, M. Rieutord, J. Ruokolainen, R. F. Stein, and I. Tuominen (1992), Dynamo action in stratified convection with overshoot, *Astrophys. J.*, *392*, 647.

Pariat, E., G. Aulanier, B. Schmieder, M. K. Georgoulis, D. M. Rust, and P. N. Bernasconi (2004), Resistive emergence of undulatory flux tubes, *Astrophys. J.*, *614*, 1099.

Parker, E. N. (1955), Hydromagnetic dynamo models, *Astrophys. J.*, *122*, 293.

Parker, E. N. (1978), *Cosmical Magnetic Fields*, Oxford Univ. Press, New York.

Parker, E. N. (1988), Dynamical oscillation and propulsion of magnetic fields in the convective zone of a star. VI: Small flux bundles, network fields, and ephemeral active regions, *Astrophys. J.*, *326*, 407.

Pevtsov, A. (2002), Sinuous coronal loops at the Sun, in *Multi-Wavelength Observations of Coronal Structure and Dynamics*, p. 125–134, Elsevier, New York.

Pevtsov, A., C. Canfield, and S. M. Latushko (2001), Hemispheric helicity trend for solar cycle 23, *Astrophys. J.*, *549*, L261–L263.

Riley, P., J. A. Linker, Z. Mikic, R. Lionello, S. A. Ledvina, and J. G. Luhmann (2000), A comparison between global solar magnetohydrodynamic and potential field source surface model results, *Astrophys. J.*, *635*, 1510.

Rust, D. M., and A. Kumar (1994), Helical magnetic fields in filaments, *Solar Phys.*, *155*, 69.

Schrijver, C. J., M. L. Derosa, T. R. Metcalf, Y. Liu, J. McTiernan, S. Rgnier, G. Valori, M. S. Wheatland, and T. Wiegelmann (2006), Nonlinear force-free modeling of coronal magnetic fields: Part I. A quantitative comparison of methods, *Solar Phys.*, *235*, 161–190.

Schuessler, M. (1979), Magnetic buoyancy revisited: Analytical and numerical results for rising flux tubes, *Astron. Astrophys.*, *71*, 79.

Shibata, K., T. Tajima, R. S. Steinolfson, and R. Matsumoto (1989a), Two-dimensional magnetohydrodynamic model of emerging magnetic flux in the solar atmosphere, *Astrophys. J.*, *345*, 584.

Shibata, K., T. Tajima, R. Matsumoto, T. Horiuchi, T. Hanawa, R. Rosner, and Y. Uchida (1989b), Nonlinear Parker instability of isolated magnetic flux in a plasma, *Astrophys. J.*, *338*, 471.

Shibata, K., S. Masuda, M. Shimojo, H. Hara, T. Yokoyma, S. Tsuneta, T. Kosugi, and Y. Ogawara (1995), Hot-plasma ejections associated with compact-loop solar flares, *Astrophys. J. Lett.*, *451*, 83.

Shimojo, M., and K. Shibata (2000), Physical parameters of solar X-ray jets, *Astrophys. J.*, *542*, 1100.

Shimojo, M., S. Hashimoto, K. Shibata, T. Hirayama, H. S. Hudson, and L. W. Acton (1996), Statistical study of solar X-ray jets observed with the YOHKOH soft X-ray telescope, *Publ. Astron. Soc. Jpn.*, *48*, 123.

Shiota, D., H. Isobe, P. F. Chen, T. T. Yamamoto, T. Sakajiri, and K. Shibata (2005), Self-consistent magnetohydrodynamic modeling of a coronal mass ejection, coronal dimming, and a giant cusp-shaped arcade formation, *Astrophys. J.*, *634*, 663.

Socas-Navarro, H., and J. Sanchez Almeida (2003), Magnetic fields in the quiet Sun: Observational discrepancies and unresolved structure, *Astrophys. J.*, *593*, 581.

Stein, R. F., and A. Nordlund (2006), Solar small-scale magnetoconvection, *Astrophys. J.*, *642*, 1246.

Steiner, O., U. Grossmann-Doerth, M. Knoelker, and M. Schuessler (1998), Dynamical interaction of solar magnetic elements and granular convection: Results of a numerical simulation, *Astrophys. J.*, *495*, 468.

Sterling, A. C. (2000), Sigmoid CME source regions at the Sun: Some recent results, *J. Atmos. Solar Terr. Phys.*, *62*, 1427.

Sterling, A. C., and R. L. Moore (2005), Slow-rise and fast-rise phases of an erupting solar filament, and flare emission onset, *Astrophys. J.*, *630*, 1148.

Strous, L. H., G. Scharmer, T. D. Tarbell, A. M. Title, and C. Zwaan (1996), Phenomena in an emerging active region. I. Horizontal dynamics, *Astron. Astrophys.*, *306*, 947.

Sturrock, P. A. (1966), Model of the high-energy phase of solar flares, *Nature*, *211*, 695.

Taylor, J. R., M. Lester, T. K. Yeoman, B. A. Emery, D. J. Knipp, D. Orr, S. I. Solovyev, T. J. Hughes, and H. Lühr (1997), The response of the magnetosphere to the passage of a coronal mass ejection on March 20–21 1990, *Ann. Geophys.*, *15*, 671.

Török, T., and B. Kliem (2005), Confined and ejective eruptions of kink-unstable flux ropes, *Astrophys. J.*, *630*, 97.

Wang, T. J., and V. I. Abramenko (1999), Evolution of twist in an emerging flux region NOAA 7321, in *Magnetic Fields and Solar Processes: The 9th European Meeting on Solar Physics, ESA SP-448*, p. 671, Eur. Space Agency, Paris.

Weiss, N. O., M. R. E. Proctor, and D. P. Brownjohn (2002), Magnetic flux separation in photospheric convection, *Mon. Not. R. Astron. Soc.*, *337*, 293.

Yokoyama, T., and K. Shibata (1995), Magnetic reconnection as the origin of X-ray Jets and Hα surges on the Sun, *Nature*, *375*, 42.

Yokoyama, T., and K. Shibata (1996), Numerical simulation of solar coronal X-ray jets based on the magnetic reconnection model, *Publ. Astron. Soc. Jpn.*, *48*, 353.

Zirin, H. (1970), Active regions. I: The occurrence of solar flares and the development of active regions, *Solar Phys.*, *14*, 328.

Zwaan, C. (1987), Elements and patterns in the solar magnetic field, *Annu. Rev. Astron. Astrophys.*, *25*, 83.

—————————————

V. Archontis, School of Mathematics and Statistics, University of St. Andrews, North Haugh, St. Andrews, Fife KY16 9SS, UK. (vasilis@mcs.st-and.ac.uk)

Click Here
for
Full
Article

Radiation Belt Environment model: Application to space weather nowcasting

Mei-Ching Fok,[1] Richard B. Horne,[2] Nigel P. Meredith,[2] and Sarah A. Glauert[2]

[1] A data-driven physical model of the energetic electrons in the Earth's radiation belts, called the Radiation Belt Environment (RBE) model, has been developed to understand Earth's radiation belt dynamics and to predict the radiation conditions found there. This model calculates radiation belt electron fluxes from 10 keV to 6 MeV in the inner magnetosphere. It takes into account the realistic, time-varying magnetic field and considers effects of wave-particle interactions with whistler mode chorus waves. The storm on 23–27 October 2002 is simulated and the temporal evolutions of the radial and pitch angle distributions of energetic electrons are examined. The calculated electron fluxes agree very well with particle data from the low-orbit SAMPEX and LANL geosynchronous satellites, when the wave-particle interactions are taken into account during storm recovery. Flux increases begin near the plasmapause and then diffuse outward to higher L shells, consistent with previous findings from statistical studies. A simplified version of the RBE model is now running in real time to provide nowcasting of the radiation belt environment. With further improvements and refinements, this model will have important value in both scientific and space weather applications.

Citation: Fok, M.-C., R. B. Horne, N. P. Meredith, and S. A. Glauert (2008), Radiation Belt Environment model: Application to space weather nowcasting, *J. Geophys. Res.*, *113*, A03S08, doi:10.1029/2007JA012558.

1. Introduction

[2] The Earth's radiation belts consist of energetic electron (~100 keV to several MeV) and ions (~100 keV to several hundred MeV) trapped in the magnetosphere roughly from $1.2 < L < 8$. The energetic electrons reside in two distinct regions: the inner belt and the outer belt, which are usually separated by the slot region ($1.8 < L < 3$) of depleted particle populations. Pitch angle diffusion loss of electrons by interacting with whistler mode plasmaspheric hiss is believed to be the source of the slot region [*Lyons et al.*, 1972; *Albert*, 1994; *Meredith et al.*, 2007]. The inner belt is relatively stable while the outer belt is highly variable with geomagnetic activity. The fluxes of energetic electrons in the outer belt decrease during the main phase of a magnetic storm due to adiabatic effect [*Dessler and Karplus*, 1961; *Kim and Chan*, 1997]. Additional nonadiabatic processes also contribute to the flux decrease in the storm main phase [*Green et al.*, 2004]. During the recovery phase the flux of energetic electrons can change dramatically as well. While approximately half of all moderate and intense storms cause a net increase in the flux of energetic electrons by a factor of 2 or more, approximately a quarter of these storms result in a net decrease in the fluxes by more than a factor of 2

[*Reeves et al.*, 2003]. This variability is caused by an imbalance between acceleration, transport, and loss processes all of which become enhanced during geomagnetic storms [*Horne*, 2002; *Thorne et al.*, 2005; *Horne et al.*, 2006; *Summers et al.*, 2007]. The ratio of poststorm to prestorm flux is related to the solar wind speed [*Paulikas and Blake*, 1979; *Reeves et al.*, 2003] and the direction of the IMF B_z during the storm recovery phase [*Iles et al.*, 2002]. In some major storms the flux enhancement takes place close to the Earth and may even penetrate into the slot region [*Baker et al.*, 2004; *Zheng et al.*, 2006].

[3] The intensification of the radiation belts has significant impacts on the space environment. Moderate energy (~10 to 100 keV) electrons can cause surface charging effects and relativistic (~0.1 to 5 MeV) electrons can cause deep-dielectric charging on space systems [*Baker*, 2001]. Therefore understanding the physical processes that are controlling the development of the radiation belts during active periods and being able to predict their variability have both scientific and practical significance. Radial diffusion has traditionally been considered to be the leading transport and energization mechanism in the inner magnetosphere [*Schulz and Lanzerotti*, 1974]. However, it has recently been suggested that electrons can be accelerated efficiently by resonant wave particle interactions with whistler mode chorus waves [*Horne and Thorne*, 1998; *Summers et al.*, 1998]. Several studies have shown an association between relativistic electron flux enhancements and prolonged periods of enhanced chorus amplitudes lasting for the order of several days [*Meredith et al.*, 2002, 2003a; *Miyoshi et al.*, 2003]. Furthermore, theoretical

[1]Geospace Physics Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[2]Physical Sciences Division, British Antarctic Survey, Cambridge, UK.

work suggests that the timescale for acceleration is typically of the order of $1-2$ d [*Summers and Ma*, 2000; *Horne et al.*, 2005a, 2005b], consistent with observations. These studies suggest that wave acceleration by whistler mode chorus waves should be included in realistic physics-based models of the Earth's radiation environment.

[4] A number of models have been established to simulate the radiation belt dynamics and to provide interpretation for observable features. Two major approaches have been taken in modeling the radiation belts: test particle and kinetic formulations. *Li et al.* [1993] and *Hudson et al.* [1996] performed guiding center simulations of the storm sudden commencement on 24 March 1991 that was induced by an interplanetary shock. They used different models for the compressional electric fields but both studies were able to reproduce the rapid formation of a "new" electron belt at $L \sim 2.5$. *Elkington et al.* [2003] investigated the effects of ULF waves on energetic electron dynamics by tracking their guiding center motion in an asymmetric, compressed dipole field. They found outer belt electrons are accelerated and diffuse radially through drift resonance with Pc-5 ULF waves. Using a test particle approach, *Ukhorskiy et al.* [2006] simulated the evolution of outer belt electrons during the magnetic storm of 7 September 2002. They found the diamagnetic effect from the storm time ring current leads to expansion of electron drift paths that intercept with the magnetopause, producing significant irreversible loss of energetic electrons at $L > 5$ during the storm main phase.

[5] Kinetic formulation is another commonly used technique to model the radiation belt. In a kinetic model the equation of the particle distribution function is solved analytically or numerically. One simple approach is based on a standard radial diffusion equation with diffusion coefficients driven by the solar wind conditions or geomagnetic activity [*Li et al.*, 2001; *Albert et al.*, 2001]. *Varotsou et al.* [2005] and *Horne et al.* [2006] combined radial diffusion with acceleration and loss due to whistler mode chorus waves and confirmed that wave acceleration by whistler mode chorus is an important acceleration mechanism in the outer radiation belt. In particular, *Horne et al.* [2006] modeled a moderate storm due to a high-speed solar wind stream, typical of the declining phase of the solar cycle. They found that the accelerated (MeV) electrons are transported both inward and outward and increase in phase space density by a factor of 10 between $3.5 < L < 7$. For models that cover a wide range of energy, drift motion must be considered since convection is an important transport mechanism for lower-energy (<50 keV) particles [*Bourdarie et al.*, 1997; *Zheng et al.*, 2003; *Miyoshi et al.*, 2006]. Using the relativistic RAM model with dipole magnetic field, *Miyoshi et al.* [2006] simulated the dynamics of energetic electrons during the October 2001 storm. They reproduced the observed local time flux asymmetry for hot electrons (30 keV). They also concluded that only convective transport and radial diffusion cannot explain the enhancement of relativistic electrons seen during storm recovery.

[6] A convection-diffusion model, namely the Radiation Belt Environment (RBE) model, has been developed to understand the radiation belt dynamics in order to predict the flux variation during active times [*Fok et al.*, 2001, 2005; *Zheng et al.*, 2003]. The model was used to simulate a

substorm injection during a dipolarization of the magnetic field [*Fok et al.*, 2001]. Observable features during substorms, such as dispersionless injection and drift echoes, are successfully reproduced. The electron flux enhancements during two magnetic storms were also studied using the RBE model [*Zheng et al.*, 2003; *Fok et al.*, 2005]. They found energization by the inductive electric field and by whistler mode waves are crucial for the flux increase during magnetic storms.

[7] The RBE model has been improved to include wave-particle interactions due to whistler mode chorus waves, to have a higher-resolution in L shell, and extended to the inner belt. In modeling the wave-particle interactions, diffusion coefficients are taken from the Pitch Angle and Energy Diffusion of Ions and Electrons (PADIE) code [*Glauert and Horne*, 2005]. In this paper, we give a brief description of the RBE formulation. We then present the RBE simulation of the storm on $23-27$ October 2002. The calculated radiation belt electron fluxes are compared with observations from the Solar Anomalous and Magnetospheric Particle Explorer (SAMPEX) low-Earth orbit satellite and Los Alamos National Laboratory (LANL) geosynchronous satellites. The dynamics of the energetic electrons are discussed and the distinct roles of transport and wave acceleration are quantitatively examined. With real-time solar wind speed, density, interplanetary magnetic field (IMF), and *Dst* data as input, a simplified version of the RBE model is running simultaneously to provide nowcasting of the radiation belt environment. Finally, we discuss the potentials of the RBE model for space weather applications.

2. Radiation Belt Environment Model

[8] The Radiation Belt Environment (RBE) model is a kinetic model that calculates the temporal variation of the phase space density of energetic electrons by solving the following bounce-averaged Boltzmann transport equation [*Fok et al.*, 2001; *Zheng et al.*, 2003]:

$$\frac{\partial f_s}{\partial t} + \langle \dot{\lambda}_i \rangle \frac{\partial f_s}{\partial \lambda_i} + \langle \dot{\varphi}_i \rangle \frac{\partial f_s}{\partial \varphi_i} = \frac{1}{\sqrt{M}} \frac{\partial}{\partial M} \left(\sqrt{M} D_{MM} \frac{\partial f_s}{\partial M} \right)$$
$$+ \frac{1}{T(y) \sin 2\alpha_o} \frac{\partial}{\partial \alpha_o} \left(T(y) \sin 2\alpha_o D_{\alpha_o \alpha_o} \frac{\partial f_s}{\partial \alpha_o} \right)$$
$$- \left(\frac{f_s}{0.5\tau_b} \right)_{\substack{\text{loss} \\ \text{cone}}} \tag{1}$$

where $f_s = f_s(t, \lambda_i, \varphi_i, M, K)$, is the average distribution function on the field line between mirror points. Here λ_i and φ_i are the magnetic latitude and local time, respectively, at the ionospheric foot point of the geomagnetic field line. M is the relativistic magnetic moment and $K = J/\sqrt{8m_oM}$, where J is the second adiabatic invariant. The motion of the particles is described by their drifts across field lines which are labeled by their ionospheric foot points. The inner boundary of λ_i is at $11.8°$, corresponding to $L = 1.06$. The outer edge of the model is bounded by field lines with λ_i at $70.2°$ and equatorial crossing at 10 Earth radius (R_E), whichever is closer. The M range is chosen to well represent the energy ranges of electrons from 10 keV to 6 MeV. The K range is chosen to cover the loss cone so that particle

precipitations can be estimated as well. Here α_o is the equatorial pitch angle and $y = \sin \alpha_o$. $T(y)$ is given by:

$$T(y) = \frac{1}{R_o} \int_0^{s_m} \frac{ds}{\cos \alpha} \qquad (2)$$

where R_o is the radial distance in R_E of the field line equatorial crossing. The integration is along the field line from the equator to the mirror point. Here τ_b is the particle bounce period.

[9] The left-hand side of (1) represents the drifts of the particle population, and the terms on the right-hand side of (1) refer to diffusion and loss. The calculation of the bounce-averaged drift velocities, $\langle \dot{\lambda}_i \rangle$ and $\langle \dot{\varphi}_i \rangle$, were described in detail by *Fok and Moore* [1997]. These drifts include gradient and curvature drift and $\mathbf{E} \times \mathbf{B}$ drift from convection and corotation electric fields. The effects of inductive electric field due to time-varying magnetic field are also taken into account implicitly in the model. For this purpose we have assumed that field lines are rooted at the ionosphere so that the inductive electric field there is zero. However, the shapes of field lines at higher altitudes vary as a function of time according to the magnetic field model. If field lines are perfect conductors, the field line motion at high altitudes, for example, at the equator, will generate an induction electric field of the form,

$$\mathbf{E}_{ind} = -\mathbf{v}_o \times \mathbf{B}_o \qquad (3)$$

where $\mathbf{v}_o$ and $\mathbf{B}_o$ are the field line velocity and magnetic field at the equator.

[10] The first term on the right-hand side of (1) represents particle diffusion in M as a result of energy diffusion due to interactions with plasma waves. The relation between the energy diffusion coefficient (D_{EE}) and the corresponding coefficient in M (D_{MM}) is given as

$$D_{MM} = D_{EE} \left(\frac{\partial M}{\partial E} \right)^2 = D_{EE} \left(\frac{E_o + E}{E_o B_m} \right)^2 \qquad (4)$$

where E_o is the electron rest energy and B_m is the magnetic field at the mirror point. The second term on the right-hand side of (1) represents pitch angle diffusion from interactions with plasma waves. For pure pitch angle diffusion (E unchanged) in the (M, K) coordinates, we first map the particle phase space density from the (M, K) to (E, α_o) coordinates, perform diffusion in α_o, and then map the updated distribution back to the (M, K) coordinates [*Fok et al.*, 1996]. The diffusion terms are followed by the loss term of the loss cone, the boundary of which is assumed to correspond to mirror height of 120 km. Particles in the loss cone are assumed to have a lifetime of one half bounce period (0.5 τ_b).

[11] In this work, the bounce-averaged pitch angle and energy diffusion coefficients, $D_{\alpha_o \alpha_o}$ and D_{EE}, are given by the PADIE code [*Glauert and Horne*, 2005]. Only resonance with lower-band whistler mode chorus (0.1 $f_{ce} < f <$ 0.5 f_{ce}) is considered. The presence of chorus waves is confined between $-15°$ and $15°$ magnetic latitude. The exclusion of high-latitude chorus may cause underestimation of the loss of electrons at energies >1 MeV [*Horne and*

Thorne, 2003]. The diffusion coefficients are calculated as a function of L shell, energy, pitch angle, and f_{pe}/f_{ce}, the ratio of plasma frequency to the cyclotron frequency, and given by

$$\frac{f_{pe}}{f_{ce}} = \frac{1}{B} \sqrt{\frac{n_e m_e}{\varepsilon_o}} \qquad (5)$$

To calculate f_{pe}/f_{ce}, which depends on the plasma density (n_e), we have embedded inside the RBE model the core plasma model of Ober and Gallagher [*Ober et al.*, 1997]. The plasma model calculates the temporal evolution of the plasma flux tube contents and equatorial plasma density distribution throughout the inner magnetosphere. The model is driven by the same magnetic and electric fields as the RBE code. The PADIE diffusion coefficients are scaled with a chorus wave intensity of 10^4 pT^2. To obtain the actual diffusion coefficients, we estimate the chorus intensity at a given location and time during the storm using the survey of CRRES plasma wave data for lower-band chorus presented by *Meredith et al.* [2001, 2003b]. For our application the wave data were binned in L shell, magnetic local time, and three levels of magnetic activity ($Kp < 2$; $2 \leq Kp < 4$; $Kp \geq 4$).

[12] Equation (1) includes multiple terms of different processes. We use the method of fractional step to decompose the equation and solve only one term at a fractional step [*Fok et al.*, 1993]. To solve (1), we have to specify the electric, magnetic fields, initial distribution, and particle distribution on the nightside boundary, which is set at 10 R_E or the last closed field line. In this work, we use the Tsyganenko 2004 model [*Tsyganenko and Sitnov*, 2005] for the magnetic field and Weimer model [*Weimer*, 2001] for electric field. The electric field is updated every time step (3 s), since updating the electric field and the recalculation of the $\mathbf{E} \times \mathbf{B}$ drift are computationally fast. Updating the magnetic field, however, requires massive calculations of field line tracing and integration along field line in order to renew the particle energy and magnetic drift. The magnetic field is thus updated every 5 min. The effect of radial diffusion is integrated in these time-varying electric and magnetic fields. The NASA trapped radiation model (AE8MAX) [*Vette*, 1991; *Fung*, 1996] is used for the initial condition in the entire RBE spatial domain. The distribution at the nightside boundary is assumed to be a kappa function with density (N_{ps}) and characteristic energy (E_{ps}) modeled by linear relations with the upstream solar wind conditions [*Zheng et al.*, 2003]:

$$N_{ps}(t) = [0.02*N_{sw}(t - 2\mathrm{hr}) + 0.316]*\sqrt{\mathrm{amu}}$$
$$E_{ps}(t) = 0.016*V_{sw}(t - 2\mathrm{hr}) - 2.4 \qquad (6)$$

where N_{ps} is in cm^{-3}, N_{sw} is the solar wind density in the same unit, amu is the atomic mass unit of the electron, E_{ps} is in keV, and V_{sw} is the solar wind velocity in km/s. Note that we assume a 2-h time lag between the plasma sheet condition and solar wind condition at the dayside magnetopause [*Borovsky et al.*, 1998].

[13] Figure 1 summarizes the RBE model architecture and logic. The only inputs to the model are the solar wind
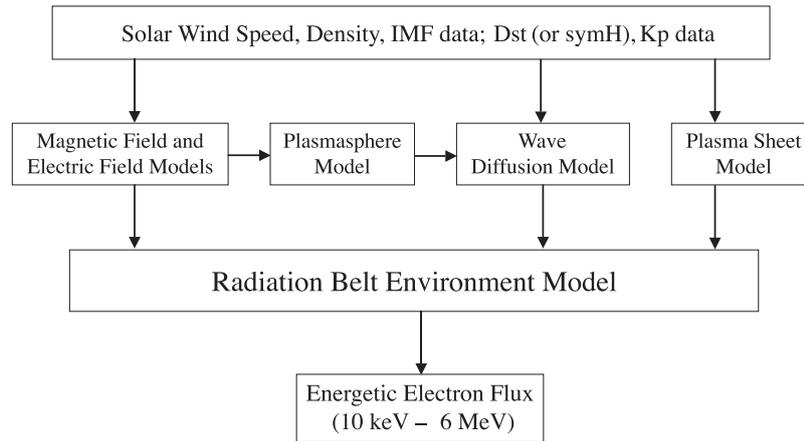
Figure 1. Model logic of the Radiation Belt Environment (RBE) model.

speed, density, and IMF and geomagnetic indices *Dst* (or *symH*) and *Kp*. The magnetic field and electric field models are driven by *Dst* or *symH*, upstream solar wind speed, density, and IMF. With the magnetic and electric fields specified, the plasmasphere model solves the core plasma density, which is used to calculate the plasma frequency needed for the PADIE code. The diffusion coefficients are then calculated for the chorus wave intensity determined by the *Kp* value. The last piece is the plasma sheet distribution, which is estimated from the solar wind density and speed (equation (6)). With all the auxiliary models in place, the RBE kinetic equation is solved to give the energetic electron

fluxes at all pitch angles in the inner and the outer radiation belts.

3. Simulation of the Storm on 23–27 October 2002

[14] The storm on 23–27 October 2002 is a moderate storm with minimum *Dst* reaching −98 nT. The storm is triggered by passing of solar wind pressure pulses and southward IMF. Figure 2 plots the *Kp*, symH, solar wind density and speed, and IMF B_z, B_y measured from the ACE satellite. The temporal resolution of the ACE data is 4 min
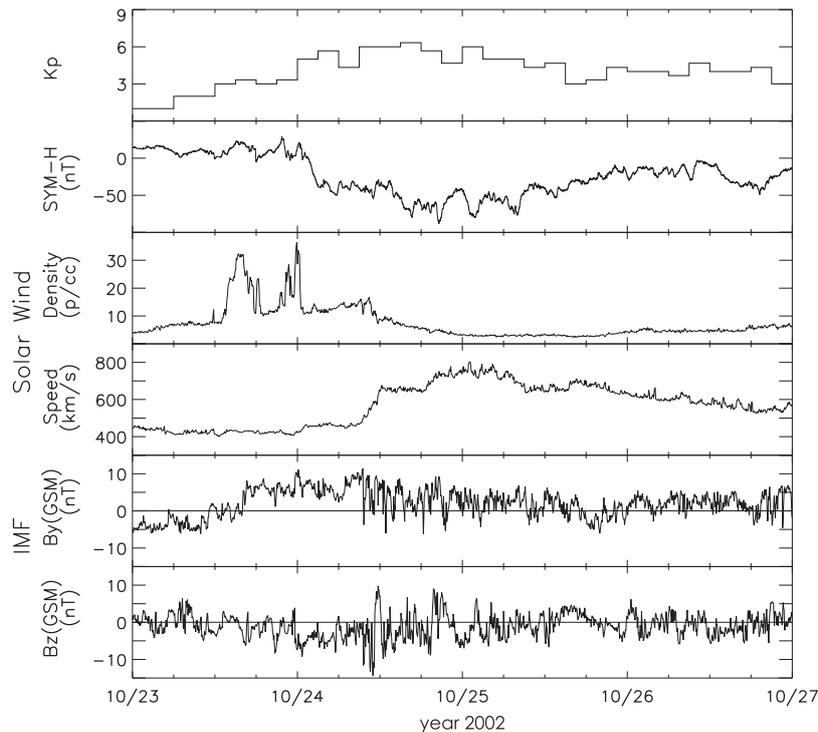
Figure 2. *Kp*, symH, solar wind density and speed, and IMF B_z, B_y measured from the ACE satellite on 23–27 October 2002. The ACE data are 44 min shifted in time.
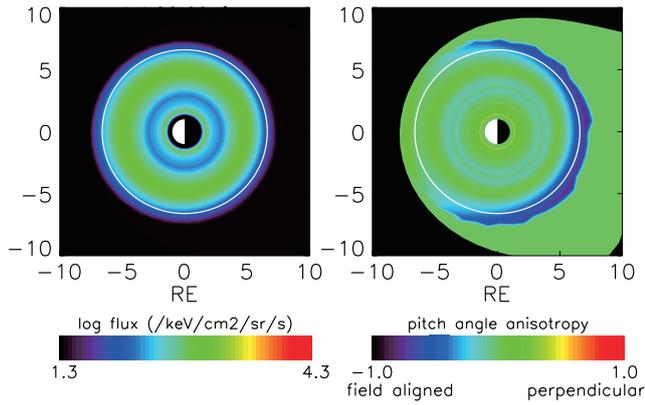
Figure 3. Simulated 0.8 MeV electron flux at the magnetic equator at quiet time, showing (left) the pitch angle averaged flux and (right) the pitch angle anisotropy. The white circles represent geosynchronous orbit.

and the data have been shifted by 44 min representing the time for solar wind to travel from the satellite to the subsolar point. The parameters plotted in Figure 2 are used to drive our magnetic (T04) and electric (Weimer 2000) field models.

[15] Figure 3 shows the simulated equatorial electron flux at 0.8 MeV during the quiet time before the storm. Figure 3 (left) is the pitch angle averaged flux and Figure 3 (right) is the pitch angle anisotropy, defined as $A = (j_\perp - j_\|)/(j_\perp + j_\|)$, where $j_\perp$ and $j_\|$ are perpendicular and parallel flux, respectively. $A = 0$ represents perpendicular and parallel fluxes are equal in magnitude, $A > 0$ perpendicular distribution, and $A < 0$ parallel distribution. The white circles in Figure 3 represent geosynchronous orbit. It can be seen from Figure 3 (left) that the inner and the outer belts are very well separated during quiet time. For this particular energy, the peak flux in the outer belt is located at ~4.5 R_E. In Figure 3 (right) a clear field-aligned pitch angle distribution (PAD) is seen on the nightside extending from dawn to dusk around the geosynchronous orbit. This field-aligned feature of energetic electrons and ions on the nightside has been observed [*Sibeck et al.*, 1987; *Garcia*, 1996; *Friedel et al.*, 2006] and is an effect of drift shell splitting in the asymmetric magnetic field. Particles with

different pitch angles injected from the nightside drift differently to the dayside. To conserve the first adiabatic invariant, perpendicular particles tend to drift farther away from the Earth on the dayside where the magnetic field is stronger for a given radial distance. Perpendicular particles originated from the nightside geosynchronous orbit may be lost when their drift paths come across the dayside magnetopause. The drift paths of parallel particles are relatively circular and have closed drift paths around the geosynchronous orbit. As a result, more particles with field-aligned pitch angles are seen on the nightside than perpendicular particles. The strong aligned region (blue partial ring in Figure 3, right) represents the difference between the last closed orbit of field-aligned electrons and that of the perpendicular particles.

[16] The energetic electron distribution is simulated throughout the storm on 23–27 October 2002 using the RBE model. Figure 4 plots the 0.8 MeV pitch angle averaged flux at 0200 UT, 25 October at the beginning of the recovery phase, 50 h into the simulation. Figure 4 (left and middle) shows the equatorial flux and pitch angle anisotropy with the same formats as Figure 3. It can be seen that the electron flux in the inner part of the outer belt has increased by an order of magnitude during the storm. We will show later in the paper that this increase is mainly coming from electrons interacting with chorus waves. The field-aligned feature in the vicinity of the nightside geosynchronous orbit persists the whole time of the event (Figure 4, middle). On the other hand, perpendicular PADs are seen on the dayside near the magnetopause. These distinct PADs at night and day can be envisaged from the meridian view in Figure 4 (right). The white curves are field lines with equatorial crossing points at 6.6 R_E. Field-aligned distributions are characterized by lower flux at the equator than at higher latitudes along the same field line. For the field line highlighted at 0000 MLT, the equatorial flux is ~3 × 10^2 cm^{-2} s^{-1} keV^{-1} sr^{-1} (cyan) and the flux at 40° magnetic latitude is ~9 × 10^2 cm^{-2} s^{-1} keV^{-1} sr^{-1} (green), three times of the value at the equator. The perpendicular distribution near the dayside magnetopause is also seen from the relatively high flux around the equator. In the outer belt where the electron flux is high, the enhancement is dominated by perpendicular particles.
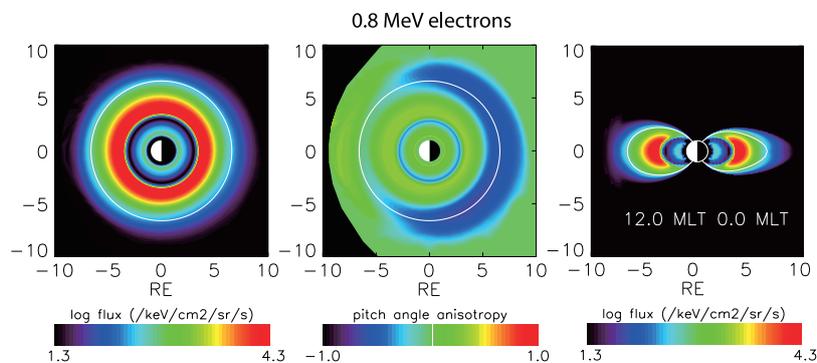
Figure 4. Simulated 0.8 MeV electron flux at the magnetic equator 50 h into the simulation at 0200 on 25 October 2002 (left, middle) in the same format as Figure 3 and (right) showing the electron flux at the noon-midnight meridian. Here, the white traces represent field lines with equatorial crossings at geosynchronous orbit.
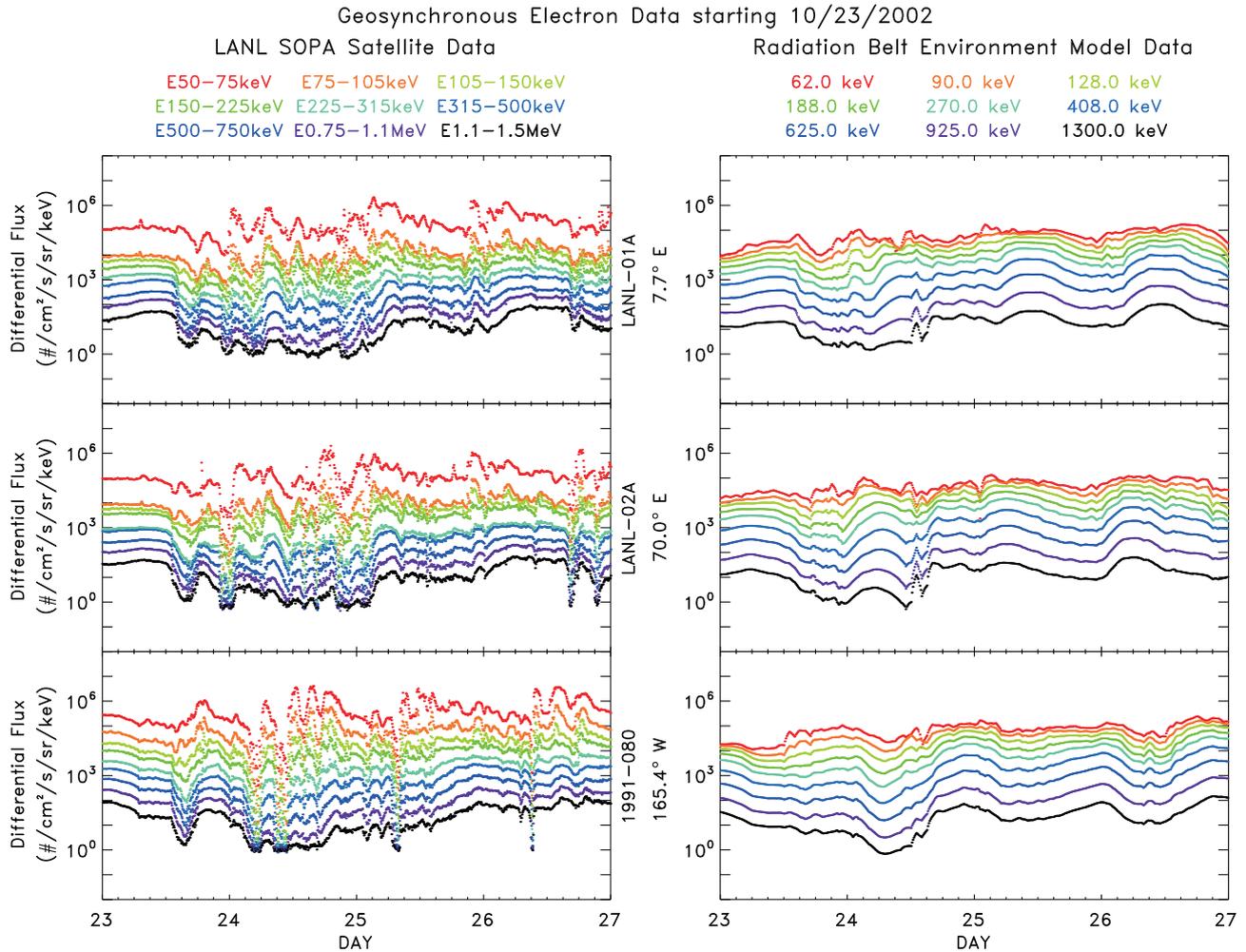
Figure 5. (left) LANL SOPA geosynchronous electron fluxes compared with (right) RBE simulated fluxes at three longitudes during the storm on 23–27 October 2002.

[17] Next we examine the storm time development of the outer belt and assess the ability of the RBE model to reproduce the observed electron flux during this event. Figure 5 (left) plots the observed electron fluxes at the geosynchronous orbits from the LANL Synchronous Orbit Particle Analyzer (SOPA) data and Figure 5 (right) plots the RBE simulated fluxes, at three geographic longitudes: 7.7°E, 70.0°E, and 165.4°W. The SOPA fluxes decrease during the storm main phase on 24 October (*Dst* effect). Electron fluxes start to recover toward the prestorm levels at late main phase. The fluctuations in the SOPA fluxes on 26 October are results of substorm activity. The simulated fluxes (Figure 5, right) exhibit similar temporal variation as those of the SOPA data, and in general the magnitudes agree very well with the data. However, the flux recovery starts earlier than the observed data, and the model fluxes lack substorm-type fluctuations.

[18] The temporal evolution of the radial profile of radiation belt particles during a storm provides important clues to identify the energization mechanisms, such as radial diffusion and wave acceleration. Figure 6 shows 2–6 MeV electron fluxes from the Proton/Electron Telescope (PET)

on SAMPEX as a function of *L* shell and time for 23–27 October 2002 (Figure 6, top). SAMPEX orbits the Earth at 520 × 670 km altitude and 82° inclination [*Baker et al.*, 1993]. It samples magnetic field lines at the ionosphere that thread the entire radiation belts every orbit period (∼90 min). PET is a zenith-pointing solid state telescope measuring electrons from 2 to 30 MeV. These electrons are either trapped particles mirroring at SAMPEX altitudes or are precipitating into the atmosphere. The latter may be in the drift loss cone and/or the local bounce loss cone depending upon the longitudinal position of SAMPEX [*Cook et al.*, 1993]. The SAMPEX fluxes shown in Figure 6 are with orbit resolution and smoothed over 15 orbits. As shown in the figure, the outer belt, slot region, and inner belt are clearly separated in this time period. The enhancement at *L* ∼ 3.4 on 23 October is a remnant from the activity taking place earlier in the month. When the storm commences on 24 October, electron fluxes start to decrease. Near the peak of the storm on late 24 October to early 25 October, electrons drift to higher *L* shells as a result of the ring current buildup. In the recovery phase, electron fluxes gradually increase. Flux enhancements are first seen at 3 <

Figure 6. (top) SAMPEX $2-6$ MeV electron L-time diagram during the storm on $23-27$ October 2002 and RBE simulated fluxes in T04 magnetic field (middle) with and (bottom) without inclusion of wave-particle interactions, respectively.

especially during geomagnetic storms. The L values in the RBE plots are calculated by:

$$L = \frac{r_i}{\cos^2 \lambda_i} \qquad (7)$$

where r_i is the ionosphere distance in R_E and λ_i is the magnetic latitude of the ionospheric foot point of the magnetic field line. The color scales of the SAMPEX and RBE data are different. The RBE fluxes are higher compared to SAMPEX since the latter observes a smaller part of the equatorial pitch angle distribution. The RBE fluxes with the inclusion of wave-particle interactions (Figure 6, middle) show similar features as the SAMPEX data. One exception is the high intensity found in the SAMPEX data on 23 October is not seen in the simulation. This is due to the fact that the RBE model uses the statistical model, AE8MAX, as initial condition and does not take into account the previous activity. When the storm hit, the flux decrease in the outer belt is also seen in the simulation. In the storm recovery, the RBE model reproduces the observed flux enhancements at about the same L shells.

[20] To identify the responsible mechanism for the flux increase, a test run is performed without the chorus wave-induced diffusions (Figure 6, bottom). Two high-flux bands are seen at around $L = 3.2$ and 4 during the storm recovery. The enhancement at $L \sim 3.2$ comes from inward radial diffusion of electrons at higher L shells. The strong flux at $L \sim 4$ is result of particle injection during the storm main phase. When wave diffusions are included (Figure 6, middle), the enhancement at the inner edge of the outer belt is diminished as a result of pitch angle diffusion loss. On the other hand, with the seed population created from particle injection, energy diffusion at $L \sim 4$ produces the flux increase in the outer belt. This enhancement cannot be explained by radial diffusion and storm injection alone. Our simulations confirm that local acceleration by wave particle interactions with whistler mode chorus can account for the flux enhancement in the center of the outer belt during the recovery phase of this storm.

[21] It is well established that strong chorus acceleration of energetic electrons takes place just outside the plasmapause where f_{pe}/f_{ce} is relatively low [*Summers et al.*, 1998; *Horne et al.*, 2006]. To demonstrate our calculation is consistent with this theory and to examine in detail how the enhancement region evolves with time, simulated $2-6$ MeV electron fluxes as a function of L and time are replotted in 1-h temporal resolution (Figure 7, top). Figure 7 (middle and bottom) shows the Ober plasma density and f_{pe}/f_{ce}, respectively. Here f_{pe}/f_{ce} is plotted only for $2.5 < L < 6.5$, since this is the L range over which the wave particle interactions are introduced into the model. It can be seen that during storm recovery the values of f_{pe}/f_{ce} are low outside the plasmapause at $3 < L < 4$. This is the region of favorable chorus-electron interactions and electron fluxes start to increase at this location and time (Figure 7, top). As wave-particle interactions continue, the high-flux region extends to higher Ls and, to a less extent, to low L shells. The expansion of the flux enhancement is a result of radial diffusion since the active wave interaction region is shrinking in late recovery phase as indicated by the increasing

$L < 4.5$ and then expand in both inward and outward direction.

[19] Figure 6 (middle and bottom) shows RBE simulated fluxes displayed in the same layout as the SAMPEX L-time plot. Note that the RBE data are equatorial fluxes rather than fluxes at the ionosphere as the SAMPEX data. However, as has been shown by *Kanekal et al.* [2001, 2005], the temporal variability of low-altitude fluxes measured by SAMPEX is nearly identical with the equatorial fluxes,

Figure 7. (top) RBE 2−6 MeV electron flux, (middle) plasmasphere density, and (bottom) $\varphi_{pe}/\varphi_{ce}$.

f_{pe}/f_{ce} ratio. *Green and Kivelson* [2004] examined the energetic electron data from POLAR/HIST instrument for acceleration event between January 1997 and January 1999. They conclude that the data is best explained by acceleration of an internal source and then outward radial diffusion to geostationary orbit. Furthermore, *Varotsou et al.* [2005] and *Horne et al.* [2006] found that local acceleration by whistler mode chorus waves followed by inward and outward radial diffusion can populate the entire radiation

belt at high energies. The RBE simulation is consistent with these findings.

4. Space Weather Application

[22] As shown in Figure 1, the only inputs to the RBE model are solar wind speed, density, and the IMF, *Dst*, and *Kp* data, which are all available near real time in public Web sites. Even with increasing complexity, the RBE CPU simulation/real time ratio is close to 1 when running on a fast PC. The RBE model thus has great potential in space weather applications in providing specification of the radiation environment in geospace. A simplified version of the RBE model, using fewer grid points in *M* and *K*, is currently running in real time to provide radiation belt nowcasting updated every 15 min. The geosynchronous fluxes at longitudes of GOES-11 and 12 are extracted from the RBE real-time run and are plotted together with real-time GOES electron (>0.6 MeV) data. The model-data comparison is continually posted at http://mcf.gsfc.nasa.gov/RB_nowcast/. Figure 8 shows the RBE prediction and the GOES data for the month of December 2006. The input data to the RBE model are plotted in the bottom. The real-time version of the RBE model does not include wave-particle interactions and thus *Kp* is not one of the input parameters. As shown in Figure 8, the RBE model agrees well with the GOES data. The simulation faithfully reproduces the diurnal variation due to the offset of geographic and magnetic equators. The RBE model also closely follows the observed flux drop-off in response to decreasing *Dst*. The model performance is less exemplary during quiet conditions, where the predicted flux is usually lower than observed. Wave-induced diffusion is being implemented in the real-time version of the RBE model. We expect these additional processes will improve the prediction accuracy not only at the geosynchronous orbit but, more importantly, at the heart of the radiation belt. It should be emphasized that the measurements presented in Figure 8 are all real-time data without any selection or noise removal.

[23] The RBE model is an effective tool to predict the response of the radiation belts to adverse solar wind conditions. However, the current design allows the model to perform near real-time nowcasting only. To enhance the space weather applications of the model, it must provide hours to days of lead time in forecasting. To accomplish this goal, the RBE code must be improved to yield higher predictive accuracy. The model is then ready to be connected to a model that is able to specify the upstream solar wind conditions with sufficient lead time. Solar wind models of this kind have been established or are under development [*Fry et al.*, 2001; *Detman et al.*, 2006]. The merger of radiation belt and solar wind models represents a very worthwhile future effort in space weather applications [*Akasofu*, 2001].

5. Discussion and Conclusions

[24] In equation (1) there is no explicit description of radial diffusion transport in the RBE model. As we mentioned above, the effects of radial transport are represented by the time-varying magnetic and electric fields. Since we only update the magnetic field every 5 min, we are missing
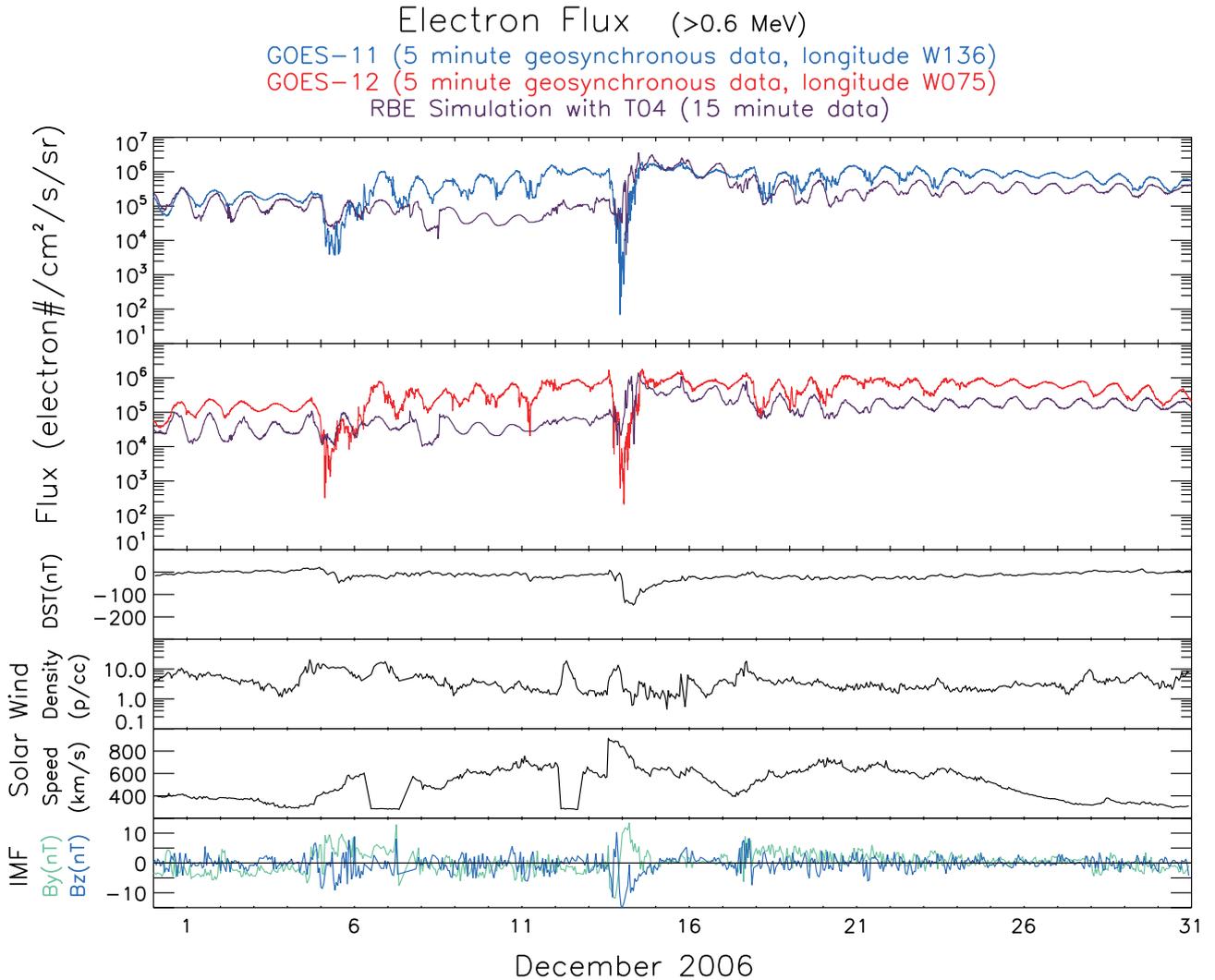
Figure 8. (top) Comparison of GOES electron fluxes (>0.6 MeV) with real-time RBE output at GOES satellite locations and (bottom) inputs to the RBE model.

the diffusive effects from the high-frequency Pc4–Pc5 fluctuations. In our previous studies [*Fok et al.*, 2001; *Zheng et al.*, 2003], we have included a radial diffusion term in the kinetic equation with diffusion coefficient given by *Cornwall* [1972]. We found, during magnetic storms, the transport due to time-varying magnetic field configuration is much larger than that from pure radial diffusion. The effect of radial diffusion may be stronger if the improved, realistic diffusion coefficients are used [i.e., *Elkington et al.*, 2003].

[25] The RBE model is sensitive to its auxiliary models: magnetic field model, electric field model, plasma sheet model, plasmasphere model, and the wave-diffusion model. Inaccuracy in these models will limit the RBE performance. A good example is the selection of the magnetic field model. It is well known that the motions of energetic ions and electrons are strongly controlled by the magnetic field configuration. To illustrate the effects of magnetic field on the RBE results, the October 2002 event is rerun using the Tsyganenko 96 (T96) model [*Tsyganenko*, 1995; *Tsyganenko and Stern*, 1996]. The T96 model does not consider the contribution from the partial ring current and the history of solar wind parameters, as all these effects are included in the

T04 model. Figure 9 shows the *L*-time plots of the simulated electron flux in T96 magnetic field with (Figure 9, top) and without (Figure 9, bottom) chorus wave interactions. We first compare the electron fluxes calculated in the two Tsyganenko models in the absence of VLF waves. It can be seen from Figure 9 (bottom) that a flux enhancement is found at $L \sim 3.2$ during the storm recovery on 25–26 October, similar to but with lower intensity than that using the T04 model (Figure 6, bottom). Furthermore, in the T96 case, there is no noticeable enhancement at $L \sim 4$ as in the T04 case. The lower flux in the RBE-T96 run indicates that the T96 model underestimates the storm time magnetic field disturbances and thus the efficiencies of radial diffusion and particle trapping. Since the T96 model fails to provide a substantial seed population in the outer belt, including chorus associated diffusion only produces a weak enhancement there at late recovery (Figure 9, top).

[26] We have shown very good agreement between the RBE simulation results and actual SAMPEX and geosynchronous particle data. Nevertheless, more work can be done to improve the model further. Cross-diffusion in energy and pitch angle, which has been found to be
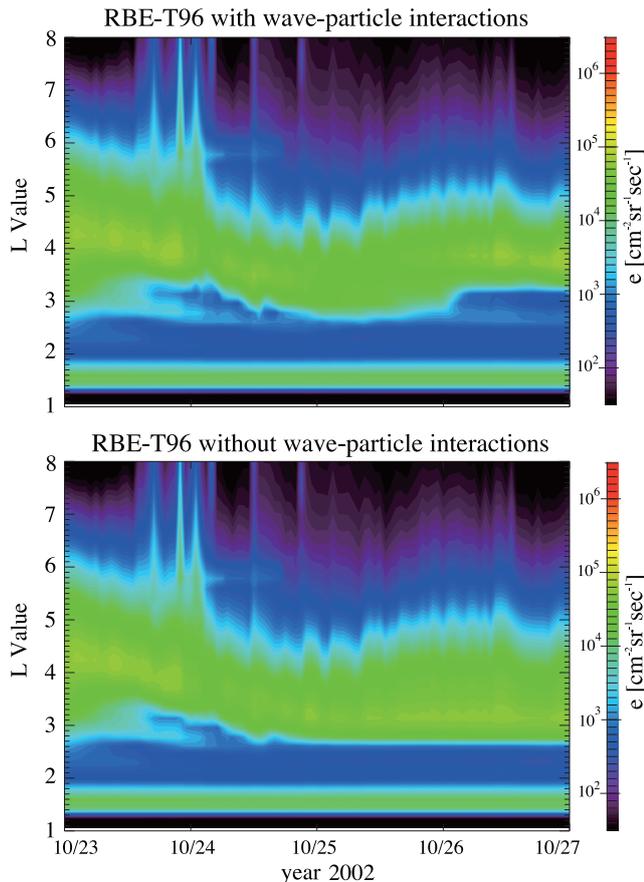
Figure 9. *L*-time diagrams of RBE simulated electron flux in T96 magnetic field with (top) and without (bottom) wave-particle interactions.

important [*Albert and Young*, 2005], has not been included in the model. Interactions with high-latitude whistler mode chorus and other waves, such as electromagnetic ion cyclotron waves and plasmaspheric hiss, should be considered as well. The diffusive effects caused by ULF waves with realistic diffusion coefficients should be reconsidered. The boundary condition of the RBE model (equation (5)) may be oversimplified and does not currently include the effects of substorm injection. All these processes will be introduced into the simulation and explored in future studies.

[27] During quiet periods, the outer belt, the slot, and the inner belt are well-defined regions. However, during intense storms, significant flux increases are seen in the slot region [e.g., *Baker et al.*, 2004] and even deep into the inner belt. The penetration distance has been found to be controlled by the strength of the storm as indicated by the *Dst* index [*Tverskaya*, 1986, 1996; *Tverskaya et al.*, 2003; *Zheng et al.*, 2006]. On the other hand, the plasmapause displaces in a similar manner. *Li et al.* [2006] found a remarkable correlation between the inner edge of the outer belt electrons and the plasmapause location. Since strong chorus acceleration of electrons takes place near the plasmapause [*Horne et al.*, 2006], the storm time erosion of the plasmasphere may place the electron acceleration zone, and thus flux enhancement, very close the Earth to the typical slot and inner belt regions. This hypothesis was tested during the Halloween storm in 2003 [*Baker et al.*, 2004]. It was shown

that the enhanced flux of energetic electrons that appeared in the slot region can be explained by wave acceleration by whistler mode chorus waves [*Horne et al.*, 2005; *Shprits et al.*, 2006].

[28] As shown in section 3, simulations from the RBE model provide an explanation of the flux enhancements seen by the SAMPEX satellite during a particular storm. The RBE model will be even more relevant to the upcoming Living With a Star Radiation Belt Storm Probe (RBSP) mission, which has a low-inclination orbit and broad particle and wave measurements (NASA/TM-2002-211613). The RBE or RBE-like models will be very useful in future RBSP data analysis and interpretation. On the other hand, data from the RBSP mission can be used to probe the importance of specific physical processes in the model.

[29] In summary, a physics-based Radiation Belt Environment (RBE) model has been developed to understand the radiation belt dynamics and provide real time predictions of the radiation belt environment that can be compared with existing and new data sets. To date the findings from this model development work include:

[30] 1. In modeling the storm on 23–27 October 2002, acceleration by chorus waves is found to be responsible for electron enhancement at the center of the outer belt.

[31] 2. The RBE model is running in real time to provide nowcasting of the radiation environment. The RBE forecasting capability will be enhanced if it is connected to a solar wind forecast model.

[32] 3. Future improvements of the RBE model include adding substorm effects, considering additional wave modes and cross-diffusion.

References

Akasofu, S.-I. (2001), Predicting geomagnetic storms as a space weather project, in *Space Weather*, Geophys. Monogr. Ser., vol. 125, edited by P. Song, et al., pp. 329–337, AGU, Washington, D. C.

Albert, J. M. (1994), Quasi-linear pitch angle diffusion coefficients: Retaining high harmonics, *J. Geophys. Res.*, 99, 23,741–23,745.

Albert, J. M., and S. L. Young (2005), Multidimensional quasi-linear diffusion of radiation belt electrons, *Geophys. Res. Lett.*, 32, L14110, doi:10.1029/2005GL023191.

Albert, J. M., D. H. Brautigam, R. V. Hilmer, and G. P. Ginet (2001), Dynamic radiation belt modeling at the Air Force Research Laboratory, in *Space Weather*, Geophys. Monogr. Ser., vol. 125, edited by P. Song, et al., pp. 281–287, AGU, Washington, D. C.

Baker, D. N. (2001), Satellite anomalies due to space storms, in *Space Storms and Space Weather Hazards*, edited by I. A. Daglis, pp. 285–311, Kluwer Acad., Boston.

Baker, D. N., G. M. Mason, O. Figueroa, G. Colon, J. G. Watzin, and R. M. Aleman (1993), An overview of the Solar, Anomalous, and Magnetospheric Particle Explorer (SAMPEX) mission, *IEEE Trans. Geosci. Remote Sens.*, 31, 531–541.

Baker, D. N., S. G. Kanekal, X. Li, S. P. Monk, J. Goldstein, and J. L. Burch (2004), An extreme distortion of the Van Allen belt arising from the Halloween solar storm in 2003, *Nature*, 432, 878–881.

Borovsky, J. E., M. F. Thomsen, and R. C. Elphic (1998), The driving of the plasma sheet by the solar wind, *J. Geophys. Res.*, *103*, 17,617–17,639.

Bourdarie, S., D. Boscher, T. Beutier, J.-A. Sauvaud, and M. Blanc (1997), Electron and proton radiation belt dynamic simulations during storm periods: A new asymmetric convection-diffusion model, *J. Geophys. Res.*, *102*, 17,541–17,552.

Cook, W. R., et al. (1993), PET: A proton/electron telescope for studies of magnetospheric, solar, and galactic particles, *IEEE Trans. Geosci. Remote Sens.*, *31*, 565–571.

Cornwall, J. M. (1972), Radial diffusion of ionized helium and protons: A probe for magnetospheric dynamics, *J. Geophys. Res.*, *77*, 1756–1770.

Dessler, A. J., and R. Karplus (1961), Some effects of diamagnetic ring currents on Van Allen radiation, *J. Geophys. Res.*, *66*, 2289–2295.

Detman, T., Z. Smith, M. Dryer, C. D. Fry, C. N. Arge, and V. Pizzo (2006), A hybrid heliospheric modeling system; Background solar wind, *J. Geophys. Res.*, *111*, A07102, doi:10.1029/2005JA011430.

Elkington, S. R., M. K. Hudson, and A. A. Chan (2003), Resonant acceleration and diffusion of outer zone electrons in an asymmetric geomagnetic field, *J. Geophys. Res.*, *108*(A3), 1116, doi:10.1029/2001JA009202.

Fok, M.-C., and T. E. Moore (1997), Ring current modeling in a realistic magnetic field configuration, *Geophys. Res. Lett.*, *24*, 1775–1778.

Fok, M.-C., J. U. Kozyra, A. F. Nagy, C. E. Rasmussen, and G. V. Khazanov (1993), Decay of equatorial ring current ions and associated aeronomical consequences, *J. Geophys. Res.*, *98*, 19,381–19,393.

Fok, M.-C., T. E. Moore, and M. E. Greenspan (1996), Ring current development during storm main phase, *J. Geophys. Res.*, *101*, 15,311–15,322.

Fok, M.-C., T. E. Moore, and W. N. Spjeldvik (2001), Rapid enhancement of radiation belt electron fluxes due to substorm dipolarization of the geomagnetic field, *J. Geophys. Res.*, *106*, 3873–3881.

Fok, M.-C., Y. Ebihara, T. E. Moore, D. M. Ober, and K. A. Keller (2005), Geospace storm processes coupling the ring current, radiation belt and plasmasphere, in *Inner Magnetosphere Interactions: New Perspectives from Imaging, Geophys. Monogr. Ser.*, vol. 159, edited by J. Burch et al., pp. 207–220, AGU, Washington, D. C.

Friedel, R. H., Y. Chen, G. D. Reeves, and T. Cayton (2006), Pitch angle evolution of energetic electrons at geosynchronous orbit during disturbed times, *Eos Trans. AGU*, *87*(52), Fall Meet. Suppl., Abstract SM41C-08.

Fry, C. D., W. Sun, C. S. Deehr, M. Dryer, Z. Smith, S.-I. Akasofu, M. Tokumaru, and M. Kojima (2001), Improvements to the HAF solar wind model for space weather predictions, *J. Geophys. Res.*, *106*, 20,985–21,001.

Fung, S. F. (1996), Recent development in the NASA trapped radiation model, in *Radiation Belts: Models and Standards, Geophys. Monogr. Ser.*, vol. 97, edited by J. F. Lemaire, D. Heynderickx, and D. N. Baker, pp. 79–91, AGU, Washington, D. C.

Garcia, H. A. (1996), Energetic electron pitch angle distribution parameters at 6.6 Re, as deduced from GOES X-ray observations, *Planet. Space Sci.*, *44*, 473–484.

Glauert, S. A., and R. B. Horne (2005), Calculation of pitch angle and energy diffusion coefficients with the PADIE code, *J. Geophys. Res.*, *110*, A04206, doi:10.1029/2004JA010851.

Green, J. C., and M. G. Kivelson (2004), Relativistic electrons in the outer radiation belt: Differentiating between acceleration mechanisms, *J. Geophys. Res.*, *109*, A03213, doi:10.1029/2003JA010153.

Green, J. C., T. G. Onsager, T. P. O'Brien, and D. N. Baker (2004), Testing loss mechanisms capable of rapidly depleting relativistic electron flux in the Earth's outer radiation belt, *J. Geophys. Res.*, *109*, A12211, doi:10.1029/2004JA010579.

Horne, R. B. (2002), The contribution of wave particle interactions to electron loss and acceleration in the Earth's radiation belts during geomagnetic storms, in *Review of Radio Science 1999–2002*, edited by W. R. Stone, chap. 33, pp. 801–828, IEEE Press, Piscataway, N. J.

Horne, R. B., and R. M. Thorne (1998), Potential waves for relativistic electron scattering and stochastic acceleration during magnetic storms, *Geophys. Res. Lett.*, *25*, 3011–3014.

Horne, R. B., and R. M. Thorne (2003), Relativistic electron acceleration and precipitation during resonant interactions with whistler-mode chorus, *Geophys. Res. Lett.*, *30*(10), 1527, doi:10.1029/2003GL016973.

Horne, R. B., R. M. Thorne, S. A. Glauert, J. M. Albert, N. P. Meredith, and R. R. Anderson (2005a), Timescale for radiation belt electron acceleration by whistler mode chorus waves, *J. Geophys. Res.*, *110*, A03225, doi:10.1029/2004JA010811.

Horne, R. B., et al. (2005b), Wave acceleration of electrons in the Van Allen radiation belts, *Nature*, *437*, 227, doi:10.1038/nature03939.

Horne, R. B., N. P. Meredith, S. A. Glauert, A. Varotsou, D. Boscher, R. M. Thorne, Y. Y. Shprits, and R. R. Anderson (2006), Mechanisms for the acceleration of radiation belt electrons, in *Recurrent Magnetic Storms: Corotating Solar Wind Streams, Geophys. Monogr. Ser.*, vol. 167, edited by B. T. Tsurutani et al., pp. 151–173, AGU, Washington, D. C.

Hudson, M. K., S. R. Elkington, J. G. Lyon, V. A. Marchenko, I. Roth, M. Temerin, and M. S. Gussenhoven (1996), MHD/particle simulations of radiation belt formation during a storm sudden commencement, in *Radiation Belts: Models and Standards, Geophys. Monogr. Ser.*, vol. 97, edited by J. F. Lemaire, D. Heynderickx, and D. N. Baker, pp. 57–62, AGU, Washington, D. C.

Iles, R. H. A., A. N. Fazakerley, A. D. Johnstone, N. P. Meredith, and P. Buhler (2002), The relativistic electron response in the outer radiation belt during magnetic storms, *Ann. Geophys.*, *20*, 957–965.

Kanekal, S. G., D. N. Baker, and J. B. Blake (2001), Multisatellite measurements of relativistic electron: Global coherence, *J. Geophys. Res.*, *106*, 29,721–29,732.

Kanekal, S. G., R. Friedel, G. D. Reeves, D. N. Baker, and J. B. Blake (2005), Relativistic electron events in 2002: Studies of pitch angle isotropization, *J. Geophys. Res.*, *110*, A12224, doi:10.1029/2004JA010974.

Kim, H.-J., and A. A. Chan (1997), Fully relativistic changes in storm time relativistic electron fluxes, *J. Geophys. Res.*, *102*, 22,107–22,116.

Li, X., I. Roth, M. Temerin, J. Wygant, M. K. Hudson, and J. B. Blake (1993), Simulation of the prompt energization and transport of radiation particles during the March 23, 1991, SSC, *Geophys. Res. Lett.*, *20*, 2423–2426.

Li, X., M. Temerin, D. N. Baker, G. D. Reeves, and D. Larson (2001), Quantitative prediction of radiation belt electrons at geostationary orbit based on solar wind measurements, *Geophys. Res. Lett.*, *28*, 1887–1890.

Li, X., D. N. Baker, T. P. O'Brien, L. Xie, and Q. G. Zong (2006), Correlation between the inner edge of outer radiation belt electrons and the innermost plasmapause location, *Geophys. Res. Lett.*, *33*, L14107, doi:10.1029/2006GL026294.

Lyons, L. R., R. M. Thorne, and C. F. Kennel (1972), Pitch-angle diffusion of radiation belt electrons with the plasmasphere, *J. Geophys. Res.*, *77*, 3455–3474.

Meredith, P. N., R. B. Horne, and R. R. Anderson (2001), Substorm dependence of chorus amplitudes: Implications for the acceleration of electrons to relativistic energy, *J. Geophys. Res.*, *106*, 13,165–13,178.

Meredith, N. P., R. B. Horne, R. H. A. Iles, R. M. Thorne, D. Heynderickx, and R. R. Anderson (2002), Outer zone relativistic electron acceleration associated with substorm-enhanced whistler-mode chorus, *J. Geophys. Res.*, *107*(A7), 1144, doi:10.1029/2001JA900146.

Meredith, N. P., M. Cain, R. B. Horne, R. M. Thorne, D. Summers, and R. R. Anderson (2003a), Evidence for chorus-driven electron acceleration to relativistic energies from a survey of geomagnetically disturbed periods, *J. Geophys. Res.*, *108*(A6), 1248, doi:10.1029/2002JA009764.

Meredith, P. N., R. B. Horne, R. M. Thorne, and R. R. Anderson (2003b), Favored regions for chorus-driven electron acceleration to relativistic energies in the Earth's outer radiation belt, *Geophys. Res. Lett.*, *30*(16), 1871, doi:10.1029/2003GL017698.

Meredith, N. P., R. B. Horne, S. A. Glauert, and R. R. Anderson (2007), Slot region electron loss timescales due to plasmaspheric hiss and lightning generated whistlers, *J. Geophys. Res.*, *112*, A08214, doi:10.1029/2007JA012413.

Miyoshi, Y., A. Morioka, T. Obara, H. Misawa, T. Nagai, and Y. Kasahara (2003), Rebuilding process of the outer radiation belt during the November 3, 1993, magnetic storm: NOAA and EXOS-D observations, *J. Geophys. Res.*, *108*(A1), 1004, doi:10.1029/2001JA007542.

Miyoshi, Y. S., V. K. Jordanova, A. Morioka, M. F. Thomsen, G. D. Reeves, D. S. Evans, and J. C. Green (2006), Observations and modeling of energetic electron dynamics during the October 2001 storm, *J. Geophys. Res.*, *111*, A11S02, doi:10.1029/2005JA011351.

Ober, D. M., J. L. Horwitz, and D. L. Gallagher (1997), Formation of density troughs embedded in the outer plasmasphere by subauroral ion drift events, *J. Geophys. Res.*, *102*, 14,595–14,602.

Paulikas, G. A., and J. B. Blake (1979), Effects of the solar wind on magnetospheric dynamics: Energetic electrons at the synchronous orbit, in *Quantitative Modeling of Magnetospheric Processes, Geophys. Monogr. Ser.*, vol. 21, edited by W.-P. Olsen, pp. 180–202, AGU, Washington, D. C.

Reeves, G. D., K. L. McAdams, and R. H. W. Friedel (2003), Acceleration and loss of relativistic electrons during geomagnetic storms, *Geophys. Res. Lett.*, *30*(10), 1529, doi:10.1029/2002GL016513.

Schulz, M., and L. J. Lanzerotti (1974), *Particle Diffusion in the Radiation Belts, Phys. and Chem. in Space*, vol. 7, Springer, New York.

Shprits, Y. Y., R. M. Thorne, R. B. Horne, S. A. Glauert, M. Cartwright, C. T. Russel, D. N. Baker, and S. G. Kanekal (2006), Acceleration mechanism responsible for the formation of the new radiation belt during the 2003 Halloween solar storms, *Geophys. Res. Lett.*, *33*, L05104, doi:10.1029/2005GL024256.

Sibeck, D. G., R. W. McEntire, A. T. Y. Lui, R. E. Lopez, and S. M. Krimigis (1987), Magnetic field drift shell splitting: Cause of unusual dayside particle pitch angle distributions during storms and substorms, *J. Geophys. Res.*, *92*, 13,485–13,497.

Summers, D., and C. Ma (2000), A model for generating relativistic electrons in the Earth's inner magnetosphere based on gyroresonant wave-particle interactions, *J. Geophys. Res.*, *105*, 2625–2640.

Summers, D., R. M. Thorne, and F. Xiao (1998), Relativistic theory of wave-particle resonant diffusion with application to electron acceleration in the magnetosphere, *J. Geophys. Res.*, *103*, 20,487–20,500.

Summers, D., B. Ni, and N. P. Meredith (2007), Timescales for radiation belt electron acceleration and loss due to resonant wave particle interactions: 2. Evaluation for VLF chorus, ELF hiss, and EMIC waves, *J. Geophys. Res.*, *112*, A04207, doi:10.1029/2006JA011993.

Thorne, R. M., R. B. Horne, S. A. Glauert, N. P. Meredith, Y. Y. Shprits, D. Summers, and R. R. Anderson (2005), The influence of wave-particle interactions on relativistic electrons during storms, in *Inner Magnetosphere Interactions: New Perspectives From Imaging, Geophys. Monogr. Ser.*, vol. 159, edited by J. Burch, M. Schulz, and H. Spence, AGU, Washington, D. C.

Tsyganenko, N. A. (1995), Modeling the Earth's magnetospheric magnetic field confined within a realistic magnetopause, *J. Geophys. Res.*, *100*, 5599–5612.

Tsyganenko, N. A., and M. I. Sitnov (2005), Modeling the dynamics of the inner magnetosphere during strong geomagnetic storms, *J. Geophys. Res.*, *110*, A03208, doi:10.1029/2004JA010798.

Tsyganenko, N. A., and D. P. Stern (1996), Modeling the global magnetic field of the large-scale Birkeland current systems, *J. Geophys. Res.*, *101*, 27,187–27,198.

Tverskaya, L. V. (1986), Electron-injection boundary in the Earth's magnetosphere, *Geomagn. Aeron.*, *26*, 736–737.

Tverskaya, L. V. (1996), The latitude position dependence of the relativistic electron maximum as a function of Dst, *Adv. Space Res.*, *18*, 135.

Tverskaya, L. V., N. N. Pavlov, J. B. Blake, R. S. Selesnick, and J. F. Fennell (2003), Predicting the L-position of the storm-injected relativistic electron belt, *Adv. Space Res.*, *31*, 1039–1044.

Ukhorskiy, A. Y., B. J. Anderson, P. C. Brandt, and N. A. Tsyganenko (2006), Storm time evolution of the outer radiation belt: Transport and losses, *J. Geophys. Res.*, *111*, A11S03, doi:10.1029/2006JA011690.

Varotsou, A., D. Boscher, S. Bourdarie, R. B. Horne, S. A. Glauert, and N. P. Meredith (2005), Simulation of the outer radiation belt electrons near geosynchronous orbit including both radial diffusion and resonant interaction with whistler mode chorus waves, *Geophys. Res. Lett.*, *32*, L19106, doi:10.1029/2005GL023282.

Vette, J. I. (1991), The AE-8 trapped electron model environment, *NSSDC/WDC-A-R&S 91-24*, NASA Goddard Space Flight Cent., Greenbelt, Md.

Weimer, D. R. (2001), An improved model of ionospheric electric potentials including substorm perturbations and applications to the Geospace Environment Modeling November 24, 1996, event, *J. Geophys. Res.*, *106*, 407–416.

Zheng, Y., M.-C. Fok, and G. V. Khazanov (2003), A radiation belt-ring current forecasting model, *Space Weather*, *1*(3), 1013, doi:10.1029/2003SW000007.

Zheng, Y., A. T. Y. Lui, X. Li, and M.-C. Fok (2006), Characteristics of 2–6 MeV electrons in the slot region and inner radiation belt, *J. Geophys. Res.*, *111*, A10204, doi:10.1029/2006JA011748.

––––––––––––––––––––

M.-C. Fok, Geospace Physics Laboratory, NASA Goddard Space Flight Center, Code 673, Greenbelt, MD 20771, USA. (mei-ching.h.fok@nasa.gov)

S. A. Glauert, R. B. Horne, and N. P. Meredith, Physical Science Division, British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK.

Nonlinear force-free coronal magnetic field extrapolation scheme based on the direct boundary integral formulation

Han He[1,2] and Huaning Wang[1]

[1]   The boundary integral equation (BIE) method was first proposed by Yan and Sakurai (2000) and used to extrapolate the nonlinear force-free magnetic field in the solar atmosphere. Recently, Yan and Li (2006) improved the BIE method and proposed the direct boundary integral equation (DBIE) formulation, which represents the nonlinear force-free magnetic field by direct integration of the magnetic field on the bottom boundary surface. On the basis of this new method, we devised a practical calculation scheme for the nonlinear force-free field extrapolation above solar active regions. The code of the scheme was tested by the analytical solutions of Low and Lou (1990) and was applied to the observed vector magnetogram of solar active region NOAA 9077. The results of the calculations show that the improvement of the new computational scheme to the scheme of Yan and Li (2006) is significant, and the force-free and divergence-free constraints are well satisfied in the extrapolated fields. The calculated field lines for NOAA 9077 present the X-shaped structure and can be helpful for understanding the magnetic configuration of the filament channel as well as the magnetic reconnection process during the Bastille Day flare on 14 July 2000.

1.   Introduction

[2]   Both observations and theoretical analyses reveal that the magnetic field plays an important role in the activity phenomena of the solar atmosphere: moving plasmas are confined to magnetic field lines, and the magnetic field also provides the energy for solar flares and other eruptive phenomena [*Tsuneta et al.*, 1992; *Masuda et al.*, 1994; *Shibata et al.*, 1995; *Wang et al.*, 1996; *Tsuneta*, 1996; *Priest and Forbes*, 2002; *Shibata*, 2004; *Lin et al.*, 2005; *Schwenn*, 2006]. To understand the physical mechanisms of these activities in the solar atmosphere, an important step is to find out the underlying structure of the magnetic field above the related active region. Currently, the direct measurement of the magnetic field in the solar chromosphere and corona is not as sophisticated as observation in the photosphere. The commonly used method for understanding the configurations of the magnetic fields above the photosphere is extrapolation: the fields can be reconstructed from a physical model in which the observed photospheric magnetic field is taken as a boundary condition. The force-free field model is often adopted for this purpose,

since it is a reasonable approximation in the solar chromosphere and corona [*Metcalf et al.*, 1995].

[3]   The nonlinear force-free magnetic field, with the field lines being everywhere aligned parallel to the electric current density, can be described by equations:

$$(\nabla \times \boldsymbol{B}) \times \boldsymbol{B} = 0, \tag{1}$$

$$\nabla \cdot \boldsymbol{B} = 0. \tag{2}$$

Equation (2) is the divergence-free constraint of the magnetic field. The force-free constraint (1) can also be written as

$$\nabla \times \boldsymbol{B} = \alpha(\boldsymbol{r})\boldsymbol{B}, \tag{3}$$

where α is a function of spatial location, usually called the force-free parameter or force-free factor. It is constant along each field line, which can be determined from the bottom boundary condition.

[4]   Currently, several methods for the nonlinear force-free field (NLFFF) extrapolation have been proposed [*Sakurai*, 1981; *Yan and Sakurai*, 2000; *Wheatland et al.*, 2000; *Wiegelmann*, 2004; *Régnier and Amari*, 2004; *Valori et al.*, 2005; *Yan*, 2005; *Wiegelmann et al.*, 2006; *Yan and Li*, 2006; *Amari et al.*, 2006; *Wheatland*, 2006; *Song et al.*, 2006; *Schrijver et al.*, 2006; *Wiegelmann*, 2007; *Song et al.*, 2007; *Valori et al.*, 2007; *Fuhrmann et al.*, 2007]. As one of

---

[1]National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China.
[2]Also at Kwasan and Hida Observatories, Graduate School of Science, Kyoto University, Kyoto, Japan.
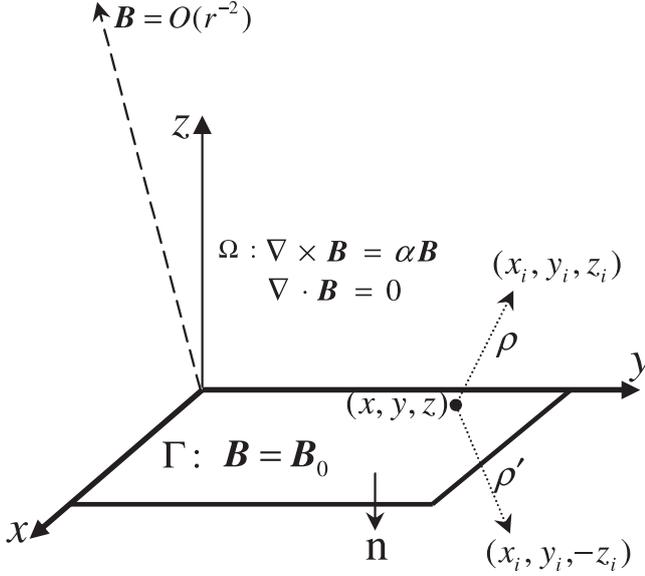
Figure 1. The geometry for application of the direct boundary integral equation (DBIE).

the techniques for nonlinear force-free field modeling, the boundary integral equation (BIE) method was first proposed by *Yan and Sakurai* [2000]. Recently, *Yan and Li* [2006] improved the BIE and proposed the direct boundary integral equation (DBIE) formulation, which represents the force-free magnetic field by direct integration of the magnetic field on the bottom boundary surface. Figure 1 shows the geometry for application of DBIE [*Yan and Li*, 2006]. On the bottom boundary surface Γ (infinite plane), the boundary condition is

$$\boldsymbol{B} = \boldsymbol{B}_0 \quad \text{on} \ \ \Gamma, \tag{4}$$

where $\boldsymbol{B}_0$ denotes the known boundary values which can be supplied from vector magnetogram measurements. At infinity, an asymptotic constraint condition is also introduced to ensure a finite energy content in the semispace Ω above Γ,

$$\boldsymbol{B} = O(r^{-2}) \quad \text{when} \ \ r \to \infty, \tag{5}$$

where r is the radial distance.

[5] According to the DBIE method, after a series of derivations using the two constraint conditions (2)–(3) and the two boundary conditions (4)–(5), the magnetic strength at the field point (x_i, y_i, z_i) in Ω can be represented by the equation [*Yan and Li*, 2006]:

$$\boldsymbol{B}_i = -\int_\Gamma \frac{\partial Y}{\partial n} \boldsymbol{B}_0 \mathrm{d}\Gamma = \int_\Gamma \frac{\partial Y}{\partial z} \boldsymbol{B}_0 \mathrm{d}\Gamma. \tag{6}$$

The reference function Y in equation (6) is chosen as

$$Y = \frac{\cos(\lambda\rho)}{4\pi\rho} - \frac{\cos(\lambda\rho')}{4\pi\rho'}, \tag{7}$$

where $\rho = [(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2]^{1/2}$ is the distance between a variable point (x, y, z) and the given field point (x_i, y_i, z_i), $\rho' = [(x - x_i)^2 + (y - y_i)^2 + (z + z_i)^2]^{1/2}$, as shown in Figure 1. The parameter λ in equation (7) has the same dimension (reciprocal of length) as the force-free factor α, which is defined by equation

$$\int_\Omega Y(\lambda^2 \boldsymbol{B} - \alpha^2 \boldsymbol{B} - \nabla\alpha \times \boldsymbol{B}) \mathrm{d}\Omega = 0. \tag{8}$$

In the case of a nonlinear force-free field, corresponding to B_{ix}, B_{iy}, and B_{iz} in equation (6), there exist three components of the reference function, Y_x, Y_y, and Y_z, together with λ_x, λ_y, and λ_z defined locally at the given field point by equation (8). The values of the three components of λ are generally different [*Li et al.*, 2004; *Yan and Li*, 2006; *He and Wang*, 2006].

[6] Once the parameter λ_x, λ_y, and λ_z at the field point (x_i, y_i, z_i) are given, the magnetic field $\boldsymbol{B}_i$ can be calculated by integration of the magnetic field on the bottom boundary surface through the DBIE formulation (6). Because it is costly to determine λ directly from the volume integral equation (8), *Yan and Li* [2006] devised an optimal approach to find the suitable λ values locally by using the DBIE (6) together with the force-free constraint condition (1). At the field point (x_i, y_i, z_i), they defined an evaluate function as:

$$f_i(\lambda_x, \lambda_y, \lambda_z) = \frac{|\boldsymbol{J} \times \boldsymbol{B}|}{|\boldsymbol{J}||\boldsymbol{B}|}, \quad \text{with} \ \ \boldsymbol{J} = \nabla \times \boldsymbol{B}, \tag{9}$$

which measures the absolute value of sine of the angle between $\boldsymbol{J}$ and $\boldsymbol{B}$. For any initial values of λ_x, λ_y, and λ_z, $\boldsymbol{B}$ at the field point and its neighborhood can be calculated by DBIE (6), then the value of $f_i(\lambda_x, \lambda_y, \lambda_z)$ can be obtained by equation (9). In an ideal situation, the suitable λ values (denoted by λ^*_x, λ^*_y, and λ^*_z) can be determined by

$$f_i\left(\lambda^*_x, \lambda^*_y, \lambda^*_z\right) = 0, \tag{10}$$

which is equivalent to the force-free field equation (1). In practical computing, one needs to find $(\lambda^*_x, \lambda^*_y, \lambda^*_z)$ that satisfies

$$f_i\left(\lambda^*_x, \lambda^*_y, \lambda^*_z\right) = \min f_i(\lambda_x, \lambda_y, \lambda_z). \tag{11}$$

A downhill simplex method can be employed to fulfil the task of equation (11) [*Nelder and Mead*, 1965; *Yan and Li*, 2006]. The computation of *Yan and Li* [2006] shows that if equation (10) is satisfied, the divergence-free constraint (2) can also be satisfied at the neighborhood of the given field point. This property indicates that the divergence-free constraint (2) has been involved during the derivation of the DBIE formulation (6).

[7] We followed the main idea of the strategy proposed by *Yan and Li* [2006] and devised the upward boundary integration scheme for the nonlinear force-free field extrapolation. That is, while we take into account the whole boundary data information through the DBIE formulation (6), the parameter λ at a given field point can be determined
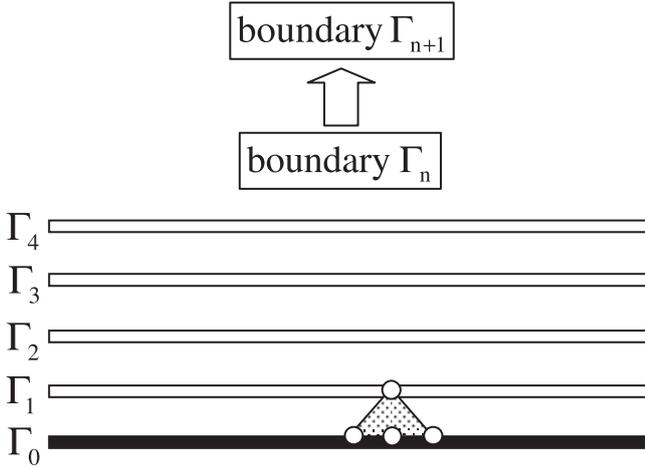
Figure 2. The diagram to illustrate the upward boundary integration scheme for nonlinear force-free field extrapolation based on the DBIE formulation (6). The computation is carried out layer by layer upwardly, the field distribution at Γ_{n+1} is calculated from the data of Γ_n.

locally through the force-free equation (1) or equation (10) with the help of only the neighboring boundary data information. Unlike the original procedure of *Yan and Li* [2006], our computation is carried out upwardly layer by layer. The procedure and techniques of the new scheme is described in section 2. The code of the scheme was tested by the analytical solutions of *Low and Lou* [1990] and was applied to the observed vector magnetogram of solar active region NOAA 9077. The results of test calculations using the analytical solutions are presented in section 3, and the results of the calculations by using the observed vector magnetogram are shown in section 4. In section 5, we give the summary and conclusion.

2. Upward Boundary Integration Scheme

[8] In the original computational procedure proposed by *Yan and Li* [2006], the equation (11) has better convergence property at the field points with lower altitude (explained in Appendix A). To take advantage of this property and to avoid the increasing errors at the field points with higher altitude, we devised a new scheme for nonlinear force-free field extrapolation based on the DBIE formulation (6), which we called the upward boundary integration scheme.

[9] In the new scheme, as shown in Figure 2, the computation is carried out upwardly, from Γ_0 (photosphere) to Γ_1, then from Γ_1 to Γ_2, and so on. In the original procedure of *Yan and Li* [2006], the bottom boundary was fixed to the photosphere, while in our scheme the bottom boundary for applying the DBIE formulation (6) is moved upwardly layer by layer. That is, we always calculate the field distribution at Γ_{n+1} from the data of the new bottom boundary Γ_n. The step distance between the two consecutive layers is the same as the space between the two consecutive grid points at Γ_n (the scale of one pixel). Since Γ_{n+1} is very close to Γ_n, the field points at Γ_{n+1} are always at very low altitude relative to the bottom boundary Γ_n. Then equation (11) can be used to achieve good convergence at

every layer. In some circumstances, especially in the case of $|\boldsymbol{J}|$ or $|\boldsymbol{B}|$ approaches to zero in equation (9), equation (11) may fail to achieve a convergent result. The problem appears in the form of isolated points where singularities appear in the raw data of Γ_{n+1}, which can be eliminated by smoothing processes with the help of the nearest grid points in the x- and y- directions.

[10] In the original procedure of *Yan and Li* [2006], the values of f_i and $\boldsymbol{J}$ in equation (9) are calculated in the infinitesimal neighborhood $\pm\delta$ of the field point in the x-, y-, and z-directions (small cubic volume surrounding the field point). The computing of the integration in equation (6) should be done seven times (six sides of the small cube plus the field point itself) to obtain the value of f_i and $\boldsymbol{J}$. To reduce the loads of computing and fully utilize the boundary data information at Γ_n in our scheme, we calculate f_i and $\boldsymbol{J}$ in a small square pyramid between the field point i and the neighboring grid points at the boundary Γ_n, as shown in Figure 3. The height of the pyramid is l (distance between Γ_{n+1} and Γ_n) and the side length of the square base is $2l$, where l denotes the space between two consecutive grid points at Γ_n (the scale of one pixel). Given an arbitrary λ at field point i, to obtain the values of $f_i(\lambda)$ and $\boldsymbol{J}(\lambda)$ in equation (9) we only need to carry out the integration of DBIE (6) one time at the field point i to calculate $\boldsymbol{B}_i(\lambda)$, thus saving computing time as compared to the original procedure of *Yan and Li* [2006]. Once $\boldsymbol{B}_i(\lambda)$ is known, J_x, J_y, and J_z can be calculated in two isosceles triangles and the square base of the pyramid, respectively, as indicated by different colors in Figure 3. $\boldsymbol{B}_o$ at the center of the square base, as shown in Figure 3, is employed to complete the calculation of $f_i(\lambda)$. Then we can use equation (11) to find the suitable value of λ^*. Once λ^* is determined, $\boldsymbol{B}_i(\lambda^*)$ will be the final result at the field point i for the nonlinear force-free field modeling.

[11] We employed the same code of the downhill simplex method as used by *Yan and Li* [2006] to perform multidimensional minimization of f_i in equation (11). The initial starting point of $(\lambda_x, \lambda_y, \lambda_z)$ for the downhill simplex
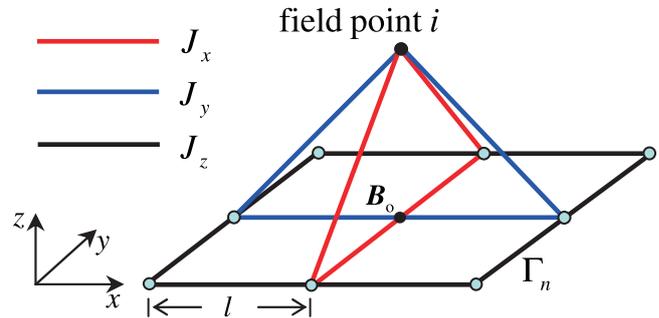
Figure 3. The diagram to illustrate the calculation of $\boldsymbol{J}(\sim\nabla\times\boldsymbol{B})$ in the small square pyramid between the field point i and the neighboring grid points at Γ_n. The height of the square pyramid is l and the side length of the square base is $2l$, where l denotes the space between two consecutive grid points at Γ_n. J_x and J_y are calculated in two isosceles triangles, respectively, and J_z is calculated in the square base of the pyramid, as indicated by different colors.
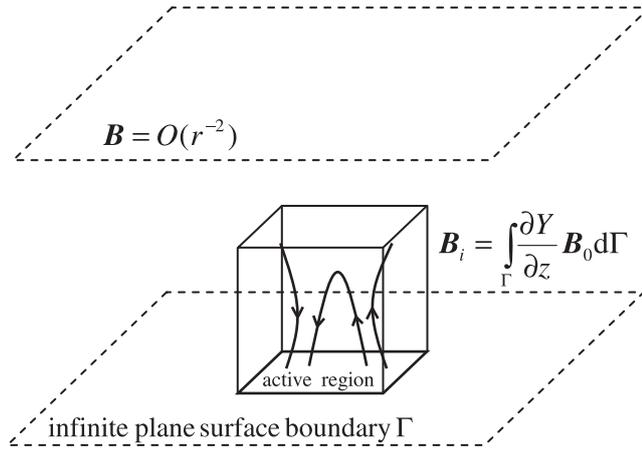
Figure 4. The diagram to illustrate the boundary conditions for application of the DBIE method to a solar active region with concentrated magnetic flux. At the bottom boundary Γ, the observed vector magnetogram is bounded in a finite square area which covers the main magnetic flux of the active region. The remained flux outside the square area is relatively very weak, and can be considered to be zero as an approximation approach. The cubic volume above the square boundary data area represents the output region of the code.

method is selected as $(0, 0, 0)$ in our code. The characteristic length scale of λ in the downhill simplex method is specified as the maximum absolute value of force-free factor α at the boundary surface Γ_0 because λ and α have the same dimension (reciprocal of length) and the same order of magnitude [*Li et al.*, 2004; *Yan and Li*, 2006; *He and Wang*, 2006]. For Case I and Case II in section 3 (analytical solutions of *Low and Lou* [1990]), the maximum absolute values of α are 11.3 and 11.1, respectively. If we adopt one pixel as the length unit (with a 64×64 grid for boundary area $x, y \in [-1, +1]$, see section 3) as in our code, the two values become 0.353 and 0.347. For the data of active regions observed in the photosphere (see section 4), we choose the characteristic length scale of λ as 10×10^{-8} m^{-1} [*Pevtsov et al.*, 1995]. Also in our code, the length unit is one pixel, for the common active regions with a field of view $300'' \times 300''$ and pixel number 64×64, the characteristic length scale of λ becomes 0.342, which approximates to the values used by Case I and Case II of the analytical solutions.

[12] As described in section 1, the DBIE formulation (6) demands that the bottom boundary Γ be an infinite plane surface [*Yan and Li*, 2006]. Figures 4 and 5 illustrates how to apply the DBIE to a solar active region with concentrated magnetic flux. At the bottom boundary Γ, the observed vector magnetogram is bounded in a finite square area which covers the main magnetic flux of the active region as shown in Figure 4. The remained flux outside the square area is relatively very weak and can be considered to be zero as an approximation approach. Then, in practical calculation, we only need to carry out the integration of DBIE (6)

over the finite square area of the active region, while the bottom boundary Γ is still an infinite plane surface.

[13] Considering that the main magnetic flux region as well as the field lines may expand at higher layers, the square area for the integration is enlarged gradually layer by layer during the calculation as illustrated in Figure 5 (in present code, from Γ_n to Γ_{n+1}, each side of the square area expanding by one pixel). Meanwhile, we keep the original pixel number of the square area at all layers by resampling the grid points to save computing time. Thus the space between the two consecutive grid points at Γ_n as well as the distance between the two consecutive layers also increases gradually with height, as shown in Figure 5. After the field distributions at a series of layers are obtained, the values of the magnetic field in the space between the layers can be calculated through the technique of interpolation.

[14] In practical calculation, the code reads the boundary data (square area, $N \times N$ array in the code), calculates the nonlinear force-free field distributions at all layers, and reforms the grid structure to a regular form through interpolation. The output of the code is the field distribution in a cubic volume ($N \times N \times N$ array in the code) which is just above the boundary data area, as illustrated in Figures 4 and 5. The current code is written in IDL programming language. The time needed for a $64 \times 64 \times 64$ output grid is about 13 h on an 1.86 GHz Intel processor.

3. Testing the Code by Using the Analytical Solutions of *Low and Lou* [1990]

[15] First, we test the code described in section 2 by using the analytical nonlinear force-free field solutions given by *Low and Lou* [1990]. The fields of the analytical solutions are basically axially symmetric. The point source of the axisymmetric fields is located at the origin of the spherical coordinate system, with the axis of symmetry pointing to the Z direction associated with the spherical coordinate system. By arbitrarily positioning the plane surface boundary Γ of DBIE in the space of the analytical fields, we
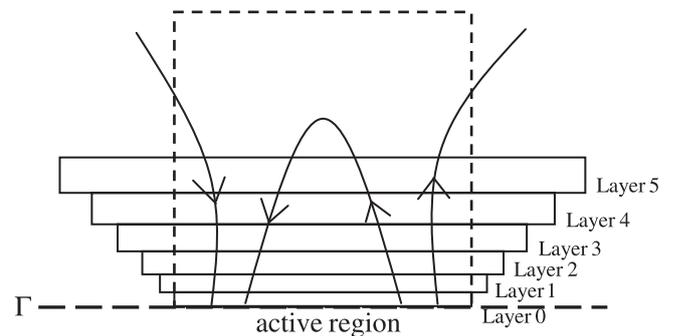
Figure 5. The diagram to illustrate how to fit expanding main magnetic flux region as well as field lines at higher layers in the new computational scheme for the DBIE method. The area for the integration of the DBIE is enlarged gradually layer by layer. Meanwhile, pixel number of the areas is fixed at all layers by resampling the grid points to save the computing time. The square dashed line indicates the final output region of the code.
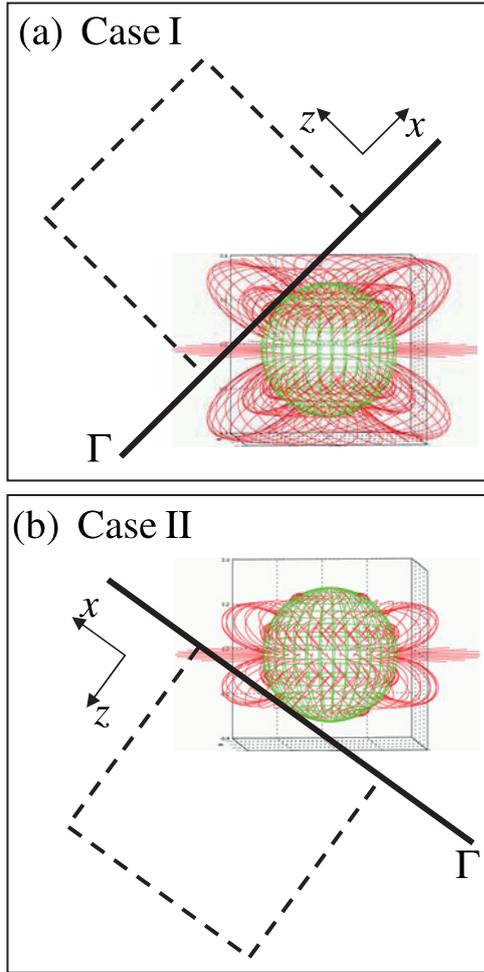
Figure 6. The global field configurations of the two analytical NLFFF solutions given by *Low and Lou* [1990], which are employed to test the validity and accuracy of the NLFFF extrapolation code (see section 3). The parameters for Case I are $n = 1$, $m = 1$, $L = 0.3$, $\Phi = \pi/4$, the parameters for Case II are $n = 3$, $m = 1$, $L = 0.3$, $\Phi = 4\pi/5$, where n and m are the eigenvalues of the solutions, L is the distance between the plane surface boundary Γ and the point source (origin of the spherical coordinate system), and Φ is the angle between the normal direction of Γ and the Z axis associated with the spherical coordinate system [*Low and Lou*, 1990; *Schrijver et al.*, 2006]. The radius of the spheres is equal to L(=0.3). The long thick lines indicate the position and scale of the bottom boundary data area ($x, y \in$ [−1, +1]). The square dashed lines represent the inner volume ($x, y \in$ [−0.5, +0.5] and $z \in$ [0, 1]) of the modeling space. x, y, and z are Cartesian coordinates defined locally on the plane surface boundary Γ.

point source (origin of the spherical coordinate system), and Φ is the angle between the normal direction of Γ and the Z axis associated with the spherical coordinate system [*Low and Lou*, 1990; *Schrijver et al.*, 2006]. In both cases, the field distributions in the modeling volume bounded by $x, y \in$ [−1, +1] and $z \in$ [0, 2] (x, y, and z are Cartesian coordinates defined locally on the boundary surface Γ) were calculated based on the bottom boundary data in the area $x, y \in$ [−1, +1], just as described in section 2 and illustrated in Figures 2−5. The pixel numbers of the bottom boundary data are 64 × 64.

3.1. Direct Comparison

[16] The extrapolated field lines for Case I and Case II are compared with the analytical solutions in Figures 7 and 8, respectively. Left columns are images of analytical solutions and right columns are images of calculated fields. It can be seen from Figures 7 and 8 that the orientations of the extrapolated field lines basically coincide with the analytical solutions. Figures 7 and 8 only show the inner volume ($x, y \in$ [−0.5, +0.5] and $z \in$ [0, 1], with 32 × 32 × 32 grid) of the modeling space as indicated in Figure 6 by square dashed lines, where the quality of agreement between the extrapolated field and the analytical solutions is better than in the margin region.

[17] A vector correlation metric C_{vec} is employed to quantify the degree of agreement between the analytical
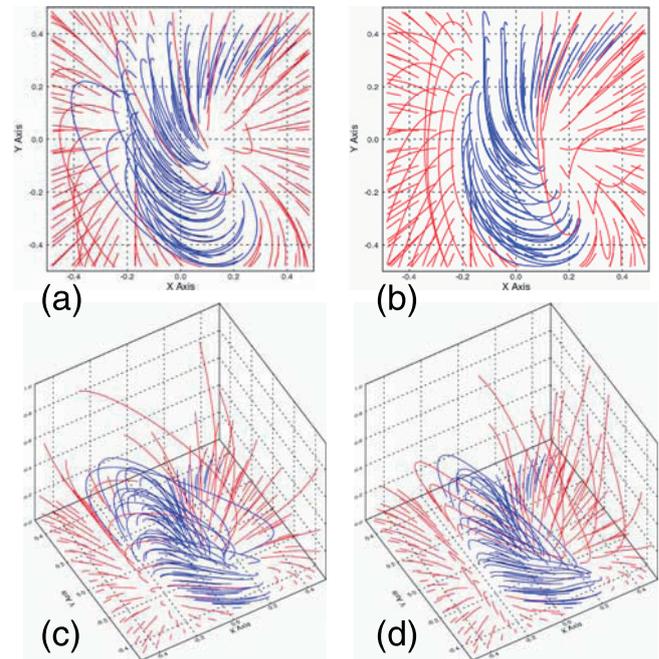
Figure 7. The extrapolated field lines in the inner volume ($x, y \in$ [−0.5, +0.5] and $z \in$ [0, 1]) of the modeling space for Case I compared with the analytical solution of *Low and Lou* [1990]. (a) and (c) Left column is images of the analytical solution and (b) and (d) right column is images of the extrapolated field. Top row are images in top view, bottom row are images in 3-D view. Closed field lines are plotted in blue, and field lines that leave the modeling box are in red.

obtain a different kind of boundary conditions prepared for extrapolation. Two cases are selected to test the validity and accuracy of our code as shown in Figure 6, which are the same cases as used by *Schrijver et al.* [2006]. The parameters for Case I are $n = 1$, $m = 1$, $L = 0.3$, $\Phi = \pi/4$, the parameters for Case II are $n = 3$, $m = 1$, $L = 0.3$, $\Phi = 4\pi/5$, where n and m are the eigenvalues of the solutions, L is the distance between the plane surface boundary Γ and the
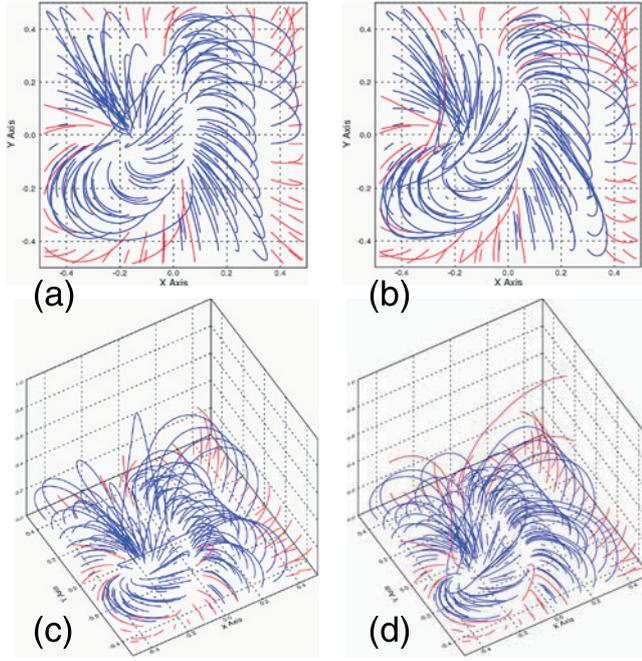
Figure 8. The extrapolated field lines in the inner volume ($x, y \in [-0.5, +0.5]$ and $z \in [0, 1]$) of the modeling space for Case II compared with the analytical solution of *Low and Lou* [1990]. (a) and (c) Left column is images of the analytical solution, and (b) and (d) right column is images of the extrapolated field. Top row are images in top view, and bottom row are images in 3-D view. Closed field lines are plotted in blue, and field lines that leave the modeling box are in red.

field B and the extrapolated field b as used by *Schrijver et al.* [2006]. To define C_{vec}, the equation

$$C_{\text{vec}} = \frac{\sum_i \boldsymbol{B}_i \cdot \boldsymbol{b}_i}{\left(\sum_i |\boldsymbol{B}_i|^2 \sum_i |\boldsymbol{b}_i|^2 \right)^{1/2}}, \qquad (12)$$

with $\boldsymbol{B}_i$ and $\boldsymbol{b}_i$ as the field vectors of the analytical field $\boldsymbol{B}$ and the extrapolated field $\boldsymbol{b}$ at each grid point i. If $\boldsymbol{B}$ and $\boldsymbol{b}$ are identical, $C_{\text{vec}} = 1$; if $\boldsymbol{B}_i \perp \boldsymbol{b}_i$, $C_{\text{vec}} = 0$. We calculated the values of C_{vec} at each layer in the central domain ($x, y \in [-0.5, +0.5]$ and $z \in [0, 2]$, with $32 \times 32 \times 64$ grid) of the modeling space, the results are shown in Figure 9. Since we only use the finite boundary data in the area $x, y \in [-1, +1]$ with 64×64 grid, it can be seen that the extrapolated fields b deviate from the analytical fields B gradually with the increasing of height. At the lower layers and in the central domain, B and b get the best agreement, as illustrated in Figures 7 and 8.

3.2. Internal Consistency of the Extrapolated Field

[18] The internal consistency of the calculated field is measured by the force-free constraint (1) and divergence-free constraint (2). To check the extent to which the extrapolated fields satisfy the force-free and divergence-free constraints, we introduced the integral measures L_f of the

Lorentz force and L_d of divergence of the fields, as used by *Schrijver et al.* [2006]. L_f and L_d are defined as:

$$L_f = \frac{1}{V} \int_V B^{-2} |(\nabla \times \boldsymbol{B}) \times \boldsymbol{B}|^2 \mathrm{d}V, \qquad (13)$$

$$L_d = \frac{1}{V} \int_V |\nabla \cdot \boldsymbol{B}|^2 \mathrm{d}V. \qquad (14)$$

We calculated L_f and L_d of the extrapolated field at each layer in the central domain ($x, y \in [-0.5, +0.5]$ and $z \in [0, 2]$, with $32 \times 32 \times 64$ grid, the unit of length in the calculation is one pixel) and plotted the curves of L_f versus height and L_d versus height. The results for Case I and Case II are shown in Figures 10 and 11, respectively. The curves for analytical solutions of *Low and Lou* [1990] are also calculated and shown in Figures 10 and 11 for reference.

[19] The curves for the analytical solutions (Figures 10a and 10c, Figures 11a and 11c) show a typical profile of L_f versus height and L_d versus height for the nonlinear force-free field. That is, the measures L_f and L_d decrease to zero rapidly with the increase of height. The relatively large values of L_f and L_d at the lower layers are the effects of the discrete grid points and the finite-difference method used in the calculations, and the values of L_f and L_d at the bottom layer (layer number 0) are in the same order of magnitude. (Exact values of L_f and L_d at the bottom layer can be found in Tables 1 and 2.)

[20] The curves for the extrapolated field (Figures 10b and 10d and Figures 11b and 11d) show similar properties of the profile as do the analytical solutions, which indicates that the force-free constraint (1) and divergence-free constraint (2) are well satisfied in the extrapolated field. Comparing the analytical solutions, the larger values of L_f and L_d for the extrapolated field are the results of the errors existing in the calculated fields, which were introduced by the numerical computation and are amplified by the $|\boldsymbol{B}|^2$ terms in equations (13) and (14).

[21] Besides the values at each layer, the integral measures L_f and L_d were also calculated over the entire volume ($x, y \in [-1, +1]$ and $z \in [0, 2]$, with $64 \times 64 \times 64$ grid) as well as the inner volume ($x, y \in [-0.5, +0.5]$ and $z \in [0, 1]$, with $32 \times 32 \times 32$ grid). All the values of L_f and L_d for Case I and Case II are listed in Tables 1 and 2, respectively. The maximum values of $|\boldsymbol{B}|$ at the bottom boundary are also included in Tables 1 and 2 for reference.

3.3. Comparing With Original Computational Scheme of *Yan and Li* [2006]

[22] As discussed in section 2 and Appendix A, in the original computational scheme of *Yan and Li* [2006], the equation (11) has better convergence property at the field points with lower altitude (near the bottom boundary) than the field points with higher altitude (away from the bottom boundary). To take advantage of this property and avoid the increasing errors at the field points with higher altitude, we devised the new upward boundary integration scheme for NLFFF extrapolation. To check the improvement of the new scheme to the original computational procedure of *Yan and Li* [2006], we produced the curves of C_{vec} versus height for Case I by using the original computational scheme proposed
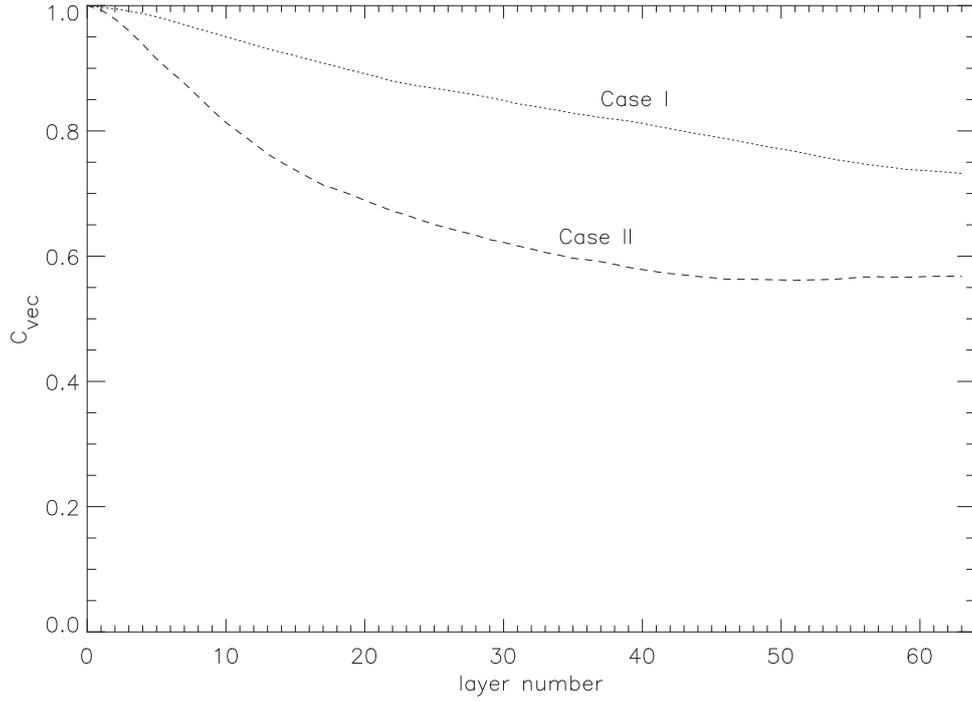
Figure 9. The curves of vector correlation metric C_{vec} (defined in equation (12)) versus height for Case I (dotted line) and Case II (dashed line). The extrapolated fields are calculated by the code of the new computational scheme. C_{vec} is calculated at each layer in the central domain $x, y \in [-0.5, +0.5]$ and $z \in [0, 2]$, with $32 \times 32 \times 64$ grid.
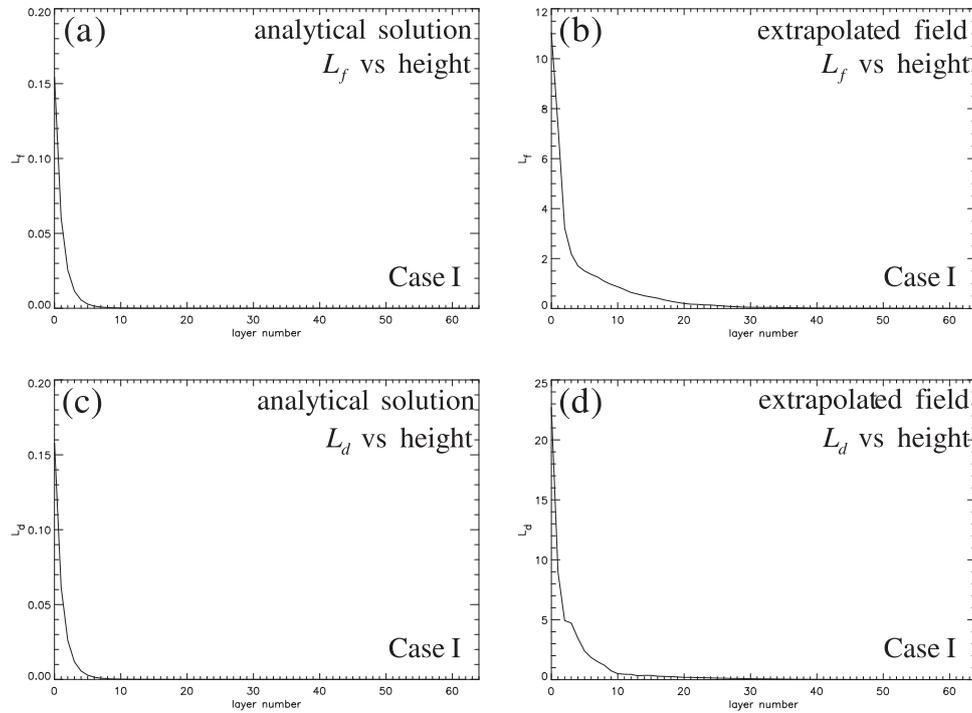
Figure 10. L_f versus height and L_d versus height curves of the extrapolated field by using the new scheme for Case I compared with the analytical solution of *Low and Lou* [1990]. (a) and (c) Left column is curves of the analytical solution, and (b) and (d) right column is curves of the extrapolated field. L_f and L_d are integral measures of the Lorentz force and divergence, as defined in equations (13) and (14). L_f and L_d are calculated at each layer in the central domain ($x, y \in [-0.5, +0.5]$ and $z \in [0, 2]$, with $32 \times 32 \times 64$ grid), the unit of length in the calculation is one pixel.
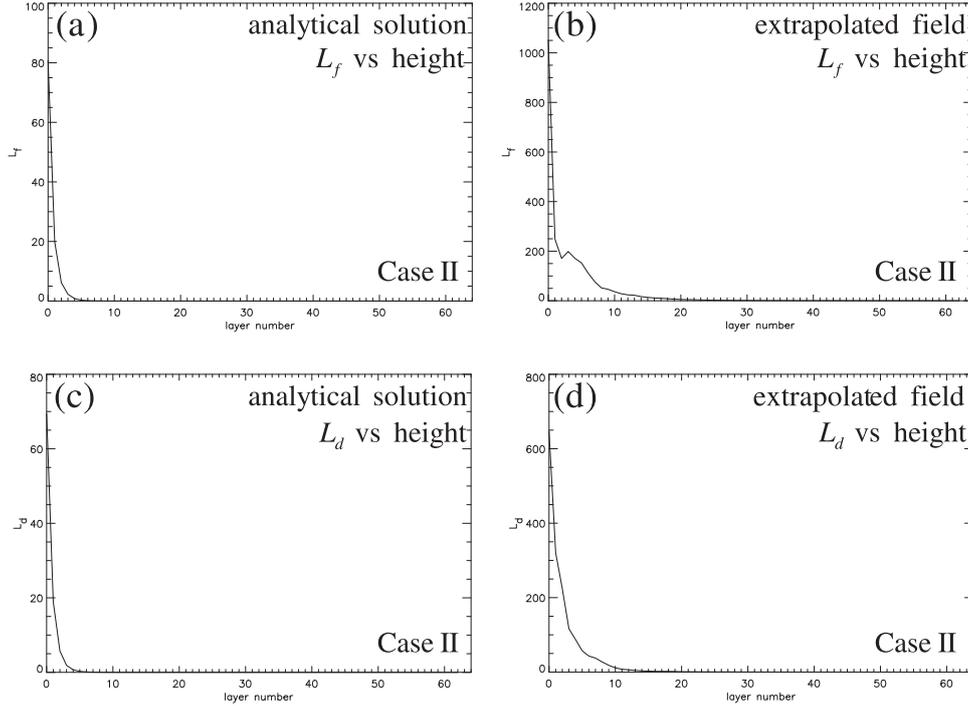
Figure 11. L_f versus height and L_d versus height curves of the extrapolated field by using the new scheme for Case II compared with the analytical solution of *Low and Lou* [1990]. (a) and (c) Left column is curves of the analytical solution, and (b) and (d) right column is curves of the extrapolated field. L_f and L_d are integral measures of the Lorentz force and divergence, as defined in equations (13) and (14). L_f and L_d are calculated at each layer in the central domain ($x, y \in [-0.5, +0.5]$ and $z \in [0, 2]$, with $32 \times 32 \times 64$ grid), the unit of length in the calculation is one pixel.

by *Yan and Li* [2006] and by using the new upward boundary integration scheme. The two curves are compared in Figure 12. The curve of C_{vec} versus height measures the agreement between the extrapolated field and the analytical solution at each layer. $C_{\mathrm{vec}} = 1$ represents the best agreement as explained in section 3.1.

[23] It can be seen from the curves in Figure 12 that the extrapolated field from the original scheme of *Yan and Li*

[2006] basically agrees with the analytical solution at lower layers, but at higher layers, the two fields are totally different. The extrapolated field by the new scheme is still similar to the analytical field at higher layers, as shown in Figure 12. The extrapolated field lines of the new scheme (Figure 7d) compared with the field lines of the original scheme (Figure A1) also illustrate this improvement.

Table 1. Integral Measures L_f and L_d of the Extrapolated Field for Case I[a]

	Analytical Solution	Extrapolated Field
L_f over the inner volume[b]	0.008500	1.228
L_d over the inner volume	0.008694	1.853
L_f over the entire volume[c]	0.001015	0.2045
L_d over the entire volume	0.001037	0.2793
L_f at the bottom layer of inner volume	0.1545	11.04
L_d at the bottom layer of inner volume	0.1580	22.71
L_f at the bottom layer of entire volume	0.03744	3.066
L_d at the bottom layer of entire volume	0.03827	5.861
Maximum of $|\boldsymbol{B}|$ at the bottom boundary	299.4	

[a]L_f and L_d are defined in equations (13) and (14), and the unit of length in the calculations is one pixel. The nonzero values of L_f and L_d for the analytical solution are due to the discrete grid points and the finite-difference method used in the calculations.
[b]Inner volume is bounded by $x, y \in [-0.5, +0.5]$ and $z \in [0, 1]$, with $32 \times 32 \times 32$ grid.
[c]Entire volume is bounded by $x, y \in [-1.0, +1.0]$ and $z \in [0, 2]$, with $64 \times 64 \times 64$ grid.

Table 2. Integral Measures L_f and L_d of the Extrapolated Field for Case II[a]

	Analytical Solution	Extrapolated Field
L_f over the inner volume[b]	3.561	80.28
L_d over the inner volume	3.134	52.48
L_f over the entire volume[c]	0.4244	9.901
L_d over the entire volume	0.3734	6.555
L_f at the bottom layer of inner volume	80.76	1050.4
L_d at the bottom layer of inner volume	69.46	646.7
L_f at the bottom layer of entire volume	19.56	255.1
L_d at the bottom layer of entire volume	16.82	160.8
Maximum of $|\boldsymbol{B}|$ at the bottom boundary	2517.7	

[a]L_f and L_d are defined in equations (13) and (14), and the unit of length in the calculations is one pixel. The nonzero values of L_f and L_d for the analytical solution are due to the discrete grid points and the finite-difference method used in the calculations.
[b]Inner volume is bounded by $x, y \in [-0.5, +0.5]$ and $z \in [0, 1]$, with $32 \times 32 \times 32$ grid.
[c]Entire volume is bounded by $x, y \in [-1.0, +1.0]$ and $z \in [0, 2]$, with $64 \times 64 \times 64$ grid.
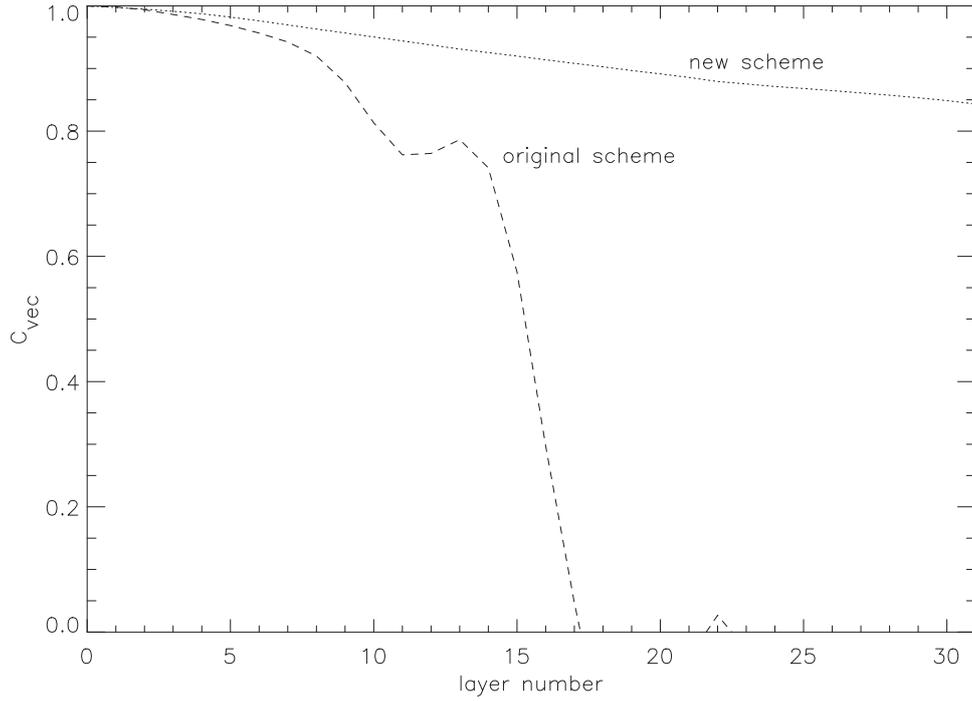
Figure 12. The curve of C_{vec} versus height for Case I by using the original computational scheme proposed by *Yan and Li* [2006] (dashed line), compared with the curve produced by using the new upward boundary integration scheme (dotted line). C_{vec} is defined in equation (12) and is calculated at each layer in the inner volume ($x, y \in [-0.5, +0.5]$ and $z \in [0, 1]$, with $32 \times 32 \times 32$ grid).

[24] Several techniques are introduced in the code of the new extrapolation scheme to save computing time as described in section 2. To calculate the NLFFF in a volume with a $64 \times 64 \times 64$ grid for Case I, the current code of the new scheme needs about 13 h on an 1.86 GHz Intel processor in IDL programming language, while the code of the original scheme of *Yan and Li* [2006] (also in IDL programming language) needs about 104 h on the same computer.

3.4. Comparing With Other Integration Scheme for NLFFF Extrapolation

[25] Five evaluation metrics C_{ves}, C_{CS}, E'_n, E'_m and ε are introduced by *Schrijver et al.* [2006] to quantify the degree

of agreement between the analytical fields $\boldsymbol{B}$ and the extrapolated field $\boldsymbol{b}$ in the modeling volume. The first metric C_{ves} are defined in equation (12), the other four metrics are defined as [*Schrijver et al.*, 2006]:

$$C_{\text{CS}} = \frac{1}{M} \sum_i \frac{\boldsymbol{B}_i \cdot \boldsymbol{b}_i}{|\boldsymbol{B}_i||\boldsymbol{b}_i|}, \tag{15}$$

$$E'_n = 1 - \frac{\sum_i |\boldsymbol{b}_i - \boldsymbol{B}_i|}{\sum_i |\boldsymbol{B}_i|}, \tag{16}$$

Table 3. Evaluation Metrics for Case I[a]

	C_{ves}	C_{CS}	E'_n	E'_m	ϵ
New scheme of DBIE (entire volume[b])	0.97	0.86	0.57	0.045	1.08
New scheme of DBIE (inner volume[c])	0.98	0.94	0.72	0.41	1.06
New scheme of DBIE (lower central domain[d])	0.99	0.97	0.81	0.70	1.04
Original scheme of DBIE (lower central domain)	0.98	0.85	0.75	0.60	0.96
Integral scheme implemented by Liu (entire volume)[e]	0.88	0.47	−0.10	−3.00	1.10
Integral scheme implemented by Liu (inner volume)	0.98	0.85	0.71	0.44	0.87
Integral scheme implemented by Liu (lower central domain)	0.98	0.94	0.77	0.69	0.86

[a]The five evaluation metrics are defined in equations (12) and (15)−(18). In the new and original computational schemes of DBIE, only bottom boundary data ($x, y \in [-1, +1]$ with 64×64 grid) were used for extrapolation, while in the scheme of Liu, data of the bottom boundary and the four side boundaries of the modeling volume were used.
[b]Entire volume is bounded by $x, y \in [-1, +1]$ and $z \in [0, 2]$, with $64 \times 64 \times 64$ grid.
[c]Inner volume is bounded by $x, y \in [-0.5, +0.5]$ and $z \in [0, 1]$, with $32 \times 32 \times 32$ grid.
[d]Lower central domain is bounded by $x, y \in [-0.5, +0.5]$ and $z \in [0, 0.5]$, with $32 \times 32 \times 16$ grid.
[e]The values of the integral scheme implemented by Liu are recalculated by courtesy of Yang Liu. The small differences between the values listed here and the values in the paper of *Schrijver et al.* [2006] are due to the different accuracies of the numerical solutions for the analytical field.

Table 4. Evaluation Metrics for Case II[a]

	C_{ves}	C_{CS}	E'_n	E'_m	ϵ
New scheme of DBIE (entire volume[b])	0.97	0.65	0.077	−12.4	1.06
New scheme of DBIE (inner volume[c])	0.97	0.81	0.48	−2.5	1.05
New scheme of DBIE (lower central domain[d])	0.98	0.90	0.65	0.044	1.04
Original scheme of DBIE (lower central domain)	0.97	0.63	0.60	−0.075	0.99
Integral scheme implemented by Liu (entire volume)[e]	0.97	0.41	−0.04	−14.1	1.04
Integral scheme implemented by Liu (inner volume)	0.97	0.54	0.47	−2.2	1.03
Integral scheme implemented by Liu (lower central domain)	0.98	0.74	0.62	−0.043	1.03

[a]The five evaluation metrics are defined in equations (12) and (15)−(18). In the new and original computational schemes of DBIE, data of the bottom boundary area $x, y \in [−1, +1]$ (64 × 64 grid) were used for extrapolation, while in the scheme of Liu, data of the bottom boundary area $x, y \in [−3, +3]$ (192 × 192 grid) were used.
[b]Entire volume is bounded by $x, y \in [−1, +1]$ and $z \in [0, 2]$, with 64 × 64 × 64 grid.
[c]Inner volume is bounded by $x, y \in [−0.5, +0.5]$ and $z \in [0, 1]$, with 32 × 32 × 32 grid.
[d]Lower central domain is bounded by $x, y \in [−0.5, +0.5]$ and $z \in [0, 0.5]$, with 32 × 32 × 16 grid.
[e]The values of the integral scheme implemented by Liu are recalculated by courtesy of Yang Liu. The small differences between the values listed here and the values in the paper of *Schrijver et al.* [2006] are due to the different accuracies of the numerical solutions for the analytical field.

Figure 13. Results of the NLFFF extrapolation for active region NOAA 9077 by using the code of the new scheme. (a) The vector magnetogram used by the code. The white contours represent the positive polarity of B_z, the black contours represent the negative polarity of B_z, the contour levels are ±50, 100, 200, 500, 1000, 1500, 2000, 3000 G. Small arrows overlying the contours represent B_t (transverse component). The data was observed at 0414 UT on 14 July 2000 by the Solar Magnetic Field Telescope (SMFT) at Huairou Solar Observing Station. The field of view is 269″ × 269″ with pixel number as 64 × 64 (4.2″/pixel). The square dashed line indicates the central domain (32 × 32 grid) of the magnetogram. (b) The top view of the extrapolated field. (c) The top view of the extrapolated field in the central domain. (d) The 3-D view of the extrapolated field in the central domain along the direction of the polarity inversion line as indicated by an arrow in the right margin of Figure 13c. Closed field lines are plotted in blue color, field lines that leave the modeling box are in red color. All field lines are plotted from layer 1 (see Figure 17) to avoid the influence of noises in the data of bottom boundary.
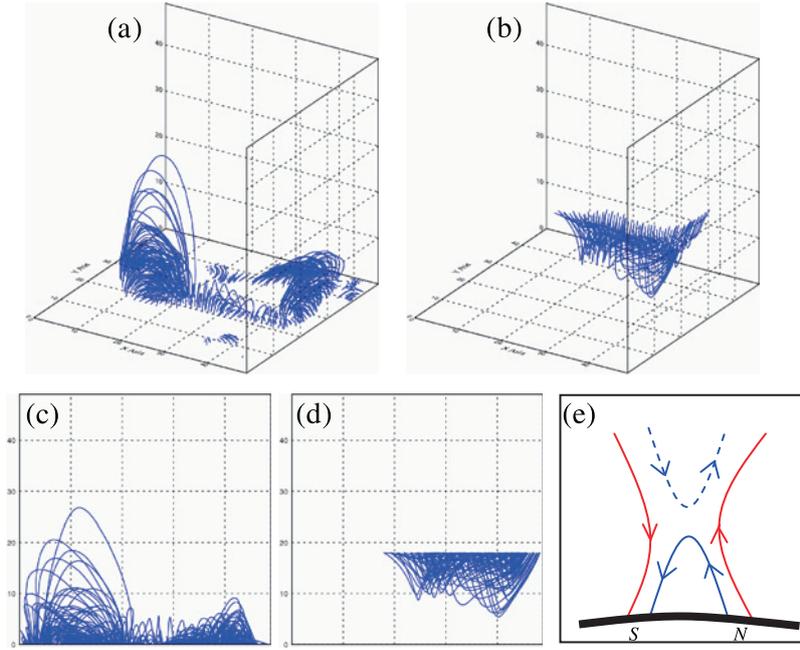
Figure 14. (a) and (b) The whole 3-D views of the closed field lines and the U-shaped field lines in the extrapolated field of NOAA 9077 (see section 4 and Figure 14e). The closed field lines are plotted from layer 1 (see Figure 17) to avoid the influence of noises in the data of bottom boundary. The U-shaped lines are plotted in the volume between layer 1 and layer 19. (c) and (d) The side views of the closed field lines and the U-shaped field lines along the direction of Y-Axis. (e) Diagram to illustrate the X-shaped structure of the field lines in the extrapolated field of NOAA 9077 as shown in Figure 13d. The dashed curve represents the U-shaped field lines above the X-point, which are not plotted and thus are displayed as a cavity in Figure 13d.

$$E'_m = 1 - \frac{1}{M} \sum_i \frac{|\boldsymbol{b}_i - \boldsymbol{B}_i|}{|\boldsymbol{B}_i|}, \tag{17}$$

$$\epsilon = \frac{\sum_i |\boldsymbol{b}_i|^2}{\sum_i |\boldsymbol{B}_i|^2}, \tag{18}$$

where $\boldsymbol{B}_i$ and $\boldsymbol{b}_i$ are the field vectors of the analytical field $\boldsymbol{B}$ and the extrapolated field $\boldsymbol{b}$ at each grid point i, M is the total number of vectors in the volume. If $\boldsymbol{B}$ and $\boldsymbol{b}$ are identical, all of the five metrics equal one.

[26] To have a quantitative comparison to other integration schemes and other NLFFF extrapolation methods, we calculated the five metrics for the extrapolated field of the new computational scheme in the entire volume ($x, y \in [-1, +1]$ and $z \in [0, 2]$, with $64 \times 64 \times 64$ grid), inner volume ($x, y \in [-0.5, +0.5]$ and $z \in [0, 1]$, with $32 \times 32 \times 32$ grid), and the lower central domain ($x, y \in [-0.5, +0.5]$ and $z \in [0, 0.5]$, with $32 \times 32 \times 16$ grid). The results for Case I and Case II are listed in Tables 3 and 4, respectively. The values of the integral scheme implemented by Liu in the paper of *Schrijver et al.* [2006] and the original computational scheme proposed by *Yan and Li* [2006] (only for lower central domain) are also included in Tables 3 and 4 for reference.

[27] It should be noted that the boundary conditions in our calculations are somewhat different to the boundary conditions used by *Schrijver et al.* [2006]. In the work of *Schrijver et al.* [2006], the data on all six boundaries of the

modeling volume are available for Case I, the data on the bottom boundary area $x, y \in [-3, +3]$ (with 192×192 grid) are provided for Case II. In our calculations, we use the bottom boundary data in the area $x, y \in [-1, +1]$ (with 64×64 grid) for both Case I and Case II. Tables 3 and 4 show that the degree of agreement of the new scheme is better than the integral scheme implemented by Liu [*Schrijver et al.*, 2006].

4. Applying the Code to the Observed Vector Magnetogram of Solar Active Region

[28] In this section, we check the validity and compatibility of the extrapolation code to deal with the observed boundary data of solar active regions. The vector magnetogram employed for testing was observed by Solar Magnetic Field Telescope (SMFT) [*Ai*, 1987], which is located at Huairou Solar Observing Station of NAOC (National Astronomical Observatories, Chinese Academy of Sciences) in Beijing. The active region associated with the magnetogram is NOAA 9077. The data were observed at 0414 UT on 14 July 2000, several hours before the Bastille Day event (X5.7 flare) at 1024 UT [*Deng et al.*, 2001; *Yan et al.*, 2001; *Liu and Zhang*, 2001; *Zhang et al.*, 2001; *Zhang*, 2002; *Tian et al.*, 2002; *Somov et al.*, 2002; *Wang et al.*, 2005].

[29] The original field of view of the magnetogram is $313'' \times 218''$ from a 512×512 pixel size CCD. We reformed the magnetogram to a square area ($269'' \times 269''$) through cropping and interpolation and reduced the
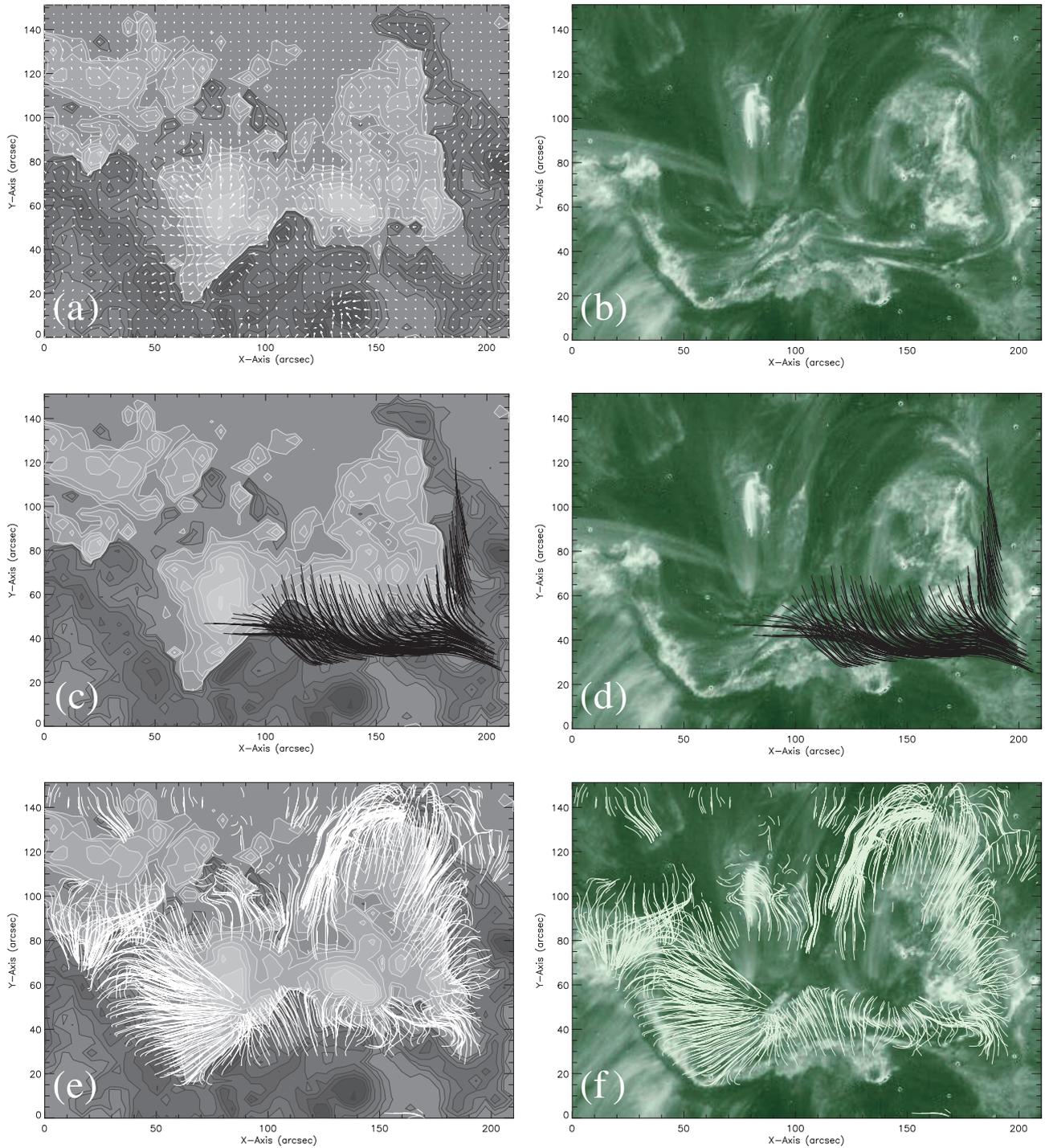
Figure 15. Comparison between the extrapolated field lines of NOAA 9077 and the TRACE 195Å image. (a) The vector magnetogram of the selected region which covers the main polarity inversion line of NOAA 9077. The data were observed at 0414 UT on 14 July 2000. (b) The TRACE 195Å image of the same region. The data were observed at 0412 UT on 14 July 2000. (c) The U-shaped field lines overlying the magnetogram of B_z. (d) The U-shaped field lines overlying the TRACE 195Å image. (e) The closed field lines overlying the magnetogram of B_z. (f) The closed field lines overlying the TRACE 195Å image. The U-shaped lines in Figures 15c and 15d are plotted in the volume between layer 0 and layer 19. The closed field lines in Figures 15e and 15f are plotted from layer 0 (see Figure 17).
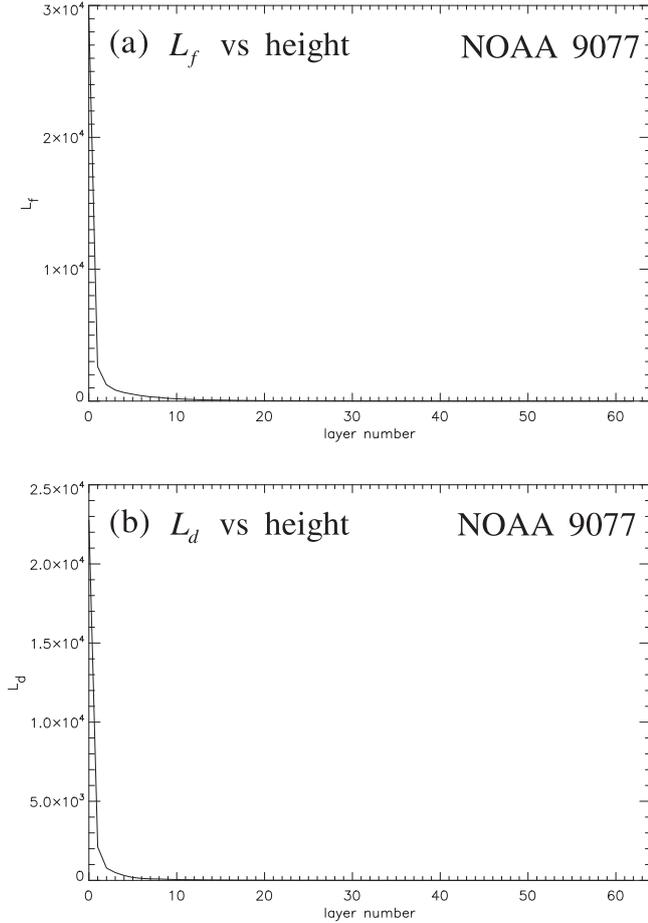
Figure 16. L_f versus height and L_d versus height curves of the extrapolated fields by using the code of the new scheme for NOAA 9077. L_f and L_d are defined in equations (13) and (14) and were calculated at each layer. The unit of magnetic field is Gauss; the unit of length in the calculation is one pixel.

pixel number to 64×64 (4.2″/pixel). The final magnetogram which is ready for the NLFFF extrapolation is shown in Figure 13a. The extrapolated field lines are shown in Figures 13b–13d. The lines in blue color are closed field lines with both footpoints being anchored at the bottom boundary, the red lines representing the field lines that leave the modeling box. Figure 13b is the top view of the field configuration in the modeling volume above the whole magnetogram. Figures 13c and 13d show the detailed structures above the central domain of the magnetogram (32×32 grid, indicated by a square dashed line in Figure 13a). Figure 13c is the top view of the field lines, Figure 13d is the 3-D view of the field lines along the direction of the polarity inversion line as indicated by an arrow in the right margin of Figure 13c. All the field lines in Figures 13b–13d are plotted from layer 1 (see Figure 17) to avoid the influence of noises in the data of the bottom boundary (the magnetogram observed in the photosphere).

[30] In Figures 13b and 13c, we can see the compact loops with different orientations aligned over the polarity inversion line of the magnetogram. The 3-D view of the

field lines in Figure 13d shows that the loops above the right half of the polarity inversion line in Figure 13c is lower than the loops above the left half of the polarity inversion line. Figure 13d also shows that in the region with lower arcade (right half of Figure 13c), the open field lines together with the underlying compact loops present an X-shaped structure. The diagram to illustrate this X-shaped structure is sketched in Figure 14e. The dashed curve in Figure 14e represents the U-shaped field lines above the X-point, which are not plotted and thus are displayed as a cavity in Figure 13d. The whole 3-D views of the closed field lines and the U-shaped field lines above the main polarity inversion line of the magnetogram are shown in Figures 14a and 14b, the side views of the field lines along the direction of Y-axis are shown in Figures 14c and 14d, respectively.

[31] To check the extent to which the extrapolated field can reflect the real distributions of the coronal magnetic fields, We compared the extrapolated field lines of NOAA 9077 with the EUV images of solar atmosphere obtained by the Transition Region and Coronal Explorer (TRACE) satellite [*Handy et al.*, 1999]. The top views of the closed field lines and the U-shaped field lines are plotted in Figure 15, and are compared with the TRACE 195Å image of the same region at almost the same time (0412 UT on 14 July 2000). It can be seen from Figures 15e and 15f that the orientations of the closed field lines above the left half of the polarity inversion line basically coincide with the coronal loops observed in the TRACE 195Å image, while in the region above the right half of the polarity inversion line, Figures 15c and 15d show that the distribution of the U-shaped field lines basically coincides with the configuration of the filament channel.

[32] To check the extent to which the extrapolated fields of the active region NOAA 9077 satisfy the force-free constraint (1) and divergence-free constraint (2), we calculated the integral measure L_f and L_d at each layer of the calculated field, and produced the curves of L_f versus height and L_d versus height, as we have done for the analytical field in section 3. The two curves are shown in Figure 16. The profiles of the curves in Figure 16 show similar properties to the analytical solutions in section 3. That is, the measures L_f and L_d decrease to zero rapidly with the increase in height, indicating that the force-free and divergence-free constraints are well satisfied in the extrapolated field.

[33] Being suitable for dealing with the noisy vector magnetogram observations is one of the great advantages of the BIE/DBIE method [*Yan and Sakurai*, 2000; *Yan*, 2005; *Yan and Li*, 2006]. The integration over the whole boundary in the DBIE formulation (6) can efficiently suppress the influence of the noises in the boundary data to the convergence property of equation (11). Thus the extrapolation code for the analytical solutions in section 3 can be applied to the real vector magnetograms directly. Figure 17 shows the magnitude distribution of the extrapolated field for NOAA 9077 at four successive layers. Figure 17a is the magnitude distribution of $|\boldsymbol{B}|$ at the bottom boundary (layer 0), Figures 17b–17d are the magnitude distribution at layers 1 to 3. Since the field are calculated
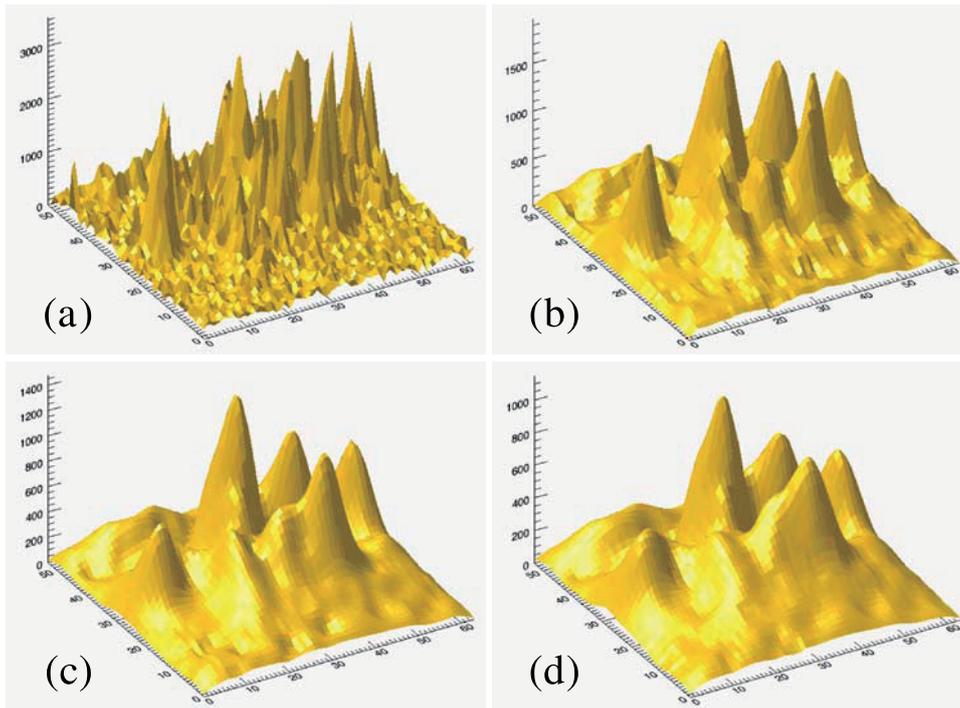
Figure 17. Magnitude distribution of the extrapolated field by using the code of the new scheme for NOAA 9077 at four successive layers. (a) The magnitude distribution of $|\boldsymbol{B}|$ at the bottom boundary (layer 0); (b)–(d) The magnitude distributions of $|\boldsymbol{B}|$ at layers 1 to 3. The unit of magnetic field (Z-Axis) is Gauss.

layer by layer, the noises in the boundary data are eliminated quickly with the increase in height.

5. Summary and Conclusion

[34] On the basis of DBIE formulation (6), we devised the upward boundary integration scheme for the nonlinear force-free field extrapolation. In this new scheme, the bottom boundary for applying the DBIE formulation (6) is moved upwardly layer by layer. That is, we always calculate the field distribution at Γ_{n+1} from the data of the new bottom boundary Γ_n, as shown in Figure 2. While we take into account the whole bottom boundary data information at Γ_n through the DBIE formulation (6), the suitable value of parameter λ at a given field point in Γ_{n+1} can be determined locally through the force-free constraint condition (11) with the help of only neighboring boundary data information at Γ_n, as shown in Figures 2 and 3.

[35] The main techniques employed in the new scheme include (1) the bottom boundary for applying the DBIE is moved upwardly layer by layer to achieve the best convergence property and accuracy, as shown in Figure 2; (2) the parameter λ at a given field point is calculated in a small square pyramid (sketched in Figure 3) to fully utilize the boundary data information at Γ_n, and thus save the computing time; (3) the square area for computing the integration of DBIE (6) is enlarged gradually layer by layer to fit the expanding field, at the same time, pixel number of the square areas is fixed at all layers by resampling the grid points to save the computing time as illustrated in Figure 5.

[36] The code of the new computational scheme was tested by the analytical solutions of *Low and Lou* [1990] and are applied to the observed vector magnetogram of solar active region. In the direct comparison between the extrapolated fields with the analytical solutions of *Low and Lou* [1990], the orientations of the extrapolated field lines are basically coincide with the analytical solutions. The quantitative comparison shows that the extrapolated fields deviate from the analytical fields gradually with the increase in height. At the lower layers and in the central domain, the best agreements are obtained. Since we only use the finite bottom boundary data in the calculation, and the analytical solutions of *Low and Lou* [1990] present global configurations [*Low and Lou*, 1990; *Wang and Sakurai*, 1998; *Li et al.*, 2004; *He and Wang*, 2006], it is natural that the extrapolated fields deviate from the analytical fields at higher layers. The force-free constraint (1) and divergence-free constraint (2) for the nonlinear force-free field are well satisfied in the extrapolated field as discussed in section 3.2.

[37] In the original computational scheme of *Yan and Li* [2006], the equation (11) has better convergence property at the field points with lower altitude (near the bottom boundary) than the field points with higher altitude (away from the bottom boundary). The new upward boundary integration scheme is proposed for taking advantage of this property and to avoid the increasing errors at higher altitude. The improvement of the new scheme to the original computational procedure of *Yan and Li* [2006] is significant as shown in section 3.3. The accuracy of the new scheme is better than the
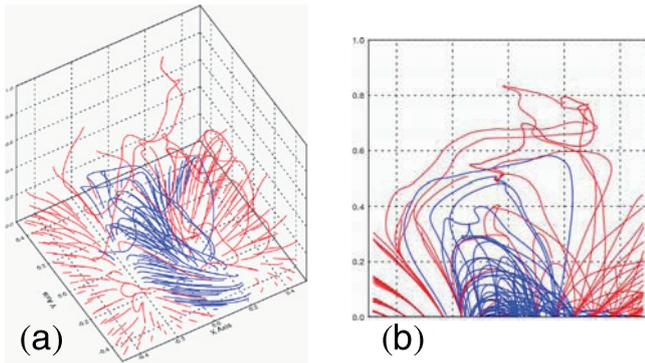
Figure A1. The extrapolated field lines in the inner volume (x, $y \in [-0.5, +0.5]$ and $z \in [0, 1]$) of the modeling space for Case I by using the original computational procedure proposed by *Yan and Li* [2006]. (a) The image in 3-D view. (b) The image in side view along the direction of Y-axis. Closed field lines are plotted in blue, and field lines that leave the modeling box are in red.

integral scheme implemented by Liu [*Schrijver et al.*, 2006] as discussed in section 3.4. If we only care about the magnetic field near the bottom boundary (at layers with very low altitude), the integral scheme of Liu can also give reasonable results as shown in Tables 3 and 4.

[38] The calculations using the observed vector magnetogram of solar active region NOAA 9077 in section 4 demonstrate that the DBIE formulation (6) and the upward boundary integration scheme can be applied to solar active regions for the nonlinear force-free field extrapolation. Only the bottom boundary data in the photosphere are needed in the calculation, and the DBIE method can suppress the noises in the observed data through the integration over the whole bottom boundary. The force-free and divergence-free constraints are well satisfied in the extrapolated field as shown in Figure 16.

[39] Since the computation is carried out point by point and layer by layer as described in section 2, fine structures can be preserved in the extrapolated field of NOAA 9077 as illustrated in Figures 13–15. In the region above the left half of the polarity inversion line of the magnetogram, the orientations of the closed field lines basically coincide with the coronal loops observed in the TRACE 195Å image as shown in Figures 15e and 15f. In the region with the dark filament above the right half of the polarity inversion line, the extrapolated field lines present the X-shaped structure as demonstrated in Figures 13–15. Beneath the X-point are the low-lying compact loops along the polarity inversion line, while above the X-point, the distribution of the U-shaped field lines coincides with the configuration of the filament channel as shown in Figures 15c and 15d, which can be helpful for understanding the magnetic structure of the dark filament as well as the magnetic reconnection process during the Bastille Day flare [*Tsuneta et al.*, 1992; *Masuda et al.*, 1994; *Shibata et al.*, 1995; *Tsuneta*, 1996; *Masuda et al.*, 2001; *Fletcher and Hudson*, 2001; *Priest and Forbes*, 2002; *Somov et al.*, 2002; *Shibata*, 2004].

[40] The application of the new DBIE extrapolation scheme to the real vector magnetograms is still in a preliminary stage. More efforts on the comparison between

the extrapolated field lines and the coronal loop observations are needed. The recent data obtained by Hinode (Solar-B) satellite are very valuable for this purpose [*Kosugi et al.*, 2007]. Topological methods and techniques will be useful for quantitatively analyzing the topological properties of the extrapolated fields [*Wang et al.*, 2000, 2001; *Longcope*, 2005; *Zhao et al.*, 2005]. Moreover, the DBIE formulation can be applied to the spherical boundary case [*Aly and Seehafer*, 1993; *Li et al.*, 2004; *Yan*, 2005; *He and Wang*, 2006]. The upward boundary integration scheme can also be adapted to the case with a spherical boundary. By using the full disk observations of vector magnetograms [*UeNo et al.*, 2004; *Zhang et al.*, 2007; *Su and Zhang*, 2007], it is possible to model the global or large-scale structures of a coronal magnetic field with a spherical boundary and study their relationships to the coronal mass ejections (CMEs) [*Chen and Shibata*, 2000; *Chen et al.*, 2002; *Zhang and Low*, 2005; *Zhou et al.*, 2006a, 2006b].

Appendix A: Convergence Property of the Original Computational Procedure Proposed by *Yan and Li* [2006]

[41] We use the analytical nonlinear force-free field solutions of *Low and Lou* [1990] to investigate the convergence property of the original computational procedure proposed by *Yan and Li* [2006]. The boundary conditions and boundary data are the same as Case I described in section 3. The field distribution in the space above the bottom boundary is calculated by using the original DBIE extrapolation code provided by Yan and Li (http://srg.bao.ac.cn/dbie-des.htm) [*Yan and Li*, 2006].

[42] The extrapolated field lines in the inner volume (x, $y \in [-0.5, +0.5]$ and $z \in [0, 1]$, with $32 \times 32 \times 32$ grid) are shown in Figure A1. The values of the extrapolated field are smoothed with the help of nearest grid points before the field lines are plotted. Figure A1 is the 3-D view, Figure A1 is the side view along the direction of Y-Axis.

[43] It can be seen from Figure A1 that the orientations of the field lines at lower altitude basically coincide with the analytical solution (Figure 7c). But the field lines at higher altitude are contorted and lose regularity. This result indicates that in the original computational procedure of *Yan and Li* [2006], the equation (11), employed to find the suitable λ values for DBIE (6), has a better convergence property at the field points with lower altitude (near the bottom boundary) than the field points with higher altitude (away from the bottom boundary). To take advantage of this property and avoid the increasing errors at the field points with higher altitude, we devised our new scheme for nonlinear force-free field extrapolation based on the DBIE formulation (6). In the new scheme, the values of magnetic field are calculated layer by layer upwardly as described in section 2.

References

Ai, G. X. (1987), Solar magnetic field telescope, *Publ. Beijing Astron. Obs.*, *9*, 27–36.

Aly, J. J., and N. Seehafer (1993), Coronal force-free magnetic field: Source-surface model, *Sol. Phys.*, *144*, 243–254.

Amari, T., T. Z. Boulmezaoud, and J. J. Aly (2006), Well posed reconstruction of the solar coronal magnetic field, *Astron. Astrophys.*, *446*, 691–705.

Chen, P. F., and K. Shibata (2000), An emerging flux trigger mechanism for coronal mass ejections, *Astrophys. J.*, *545*, 524–531.

Chen, P. F., C. Fang, K. Shibata, and Y. H. Tang (2002), Coronal mass ejections and emerging flux, *Adv. Space Res.*, *30*, 535–543.

Deng, Y., J. Wang, Y. Yan, and J. Zhang (2001), Evolution of magnetic nonpotentiality in NOAA AR 9077, *Sol. Phys.*, *204*, 11–26.

Fletcher, L., and H. Hudson (2001), The magnetic structure and generation of EUV flare ribbons, *Sol. Phys.*, *204*, 69–89.

Fuhrmann, M., N. Seehafer, and G. Valori (2007), Preprocessing of solar vector magnetograms for force-free magnetic field extrapolation, *Astron. Astrophys.*, *476*, 349–357.

Handy, B. N., et al. (1999), The transition region and coronal explorer, *Sol. Phys.*, *187*, 229–260.

He, H., and H. N. Wang (2006), The validity of the boundary integral equation for magnetic field extrapolation in open space above a spherical surface, *Mon. Not. R. Astron. Soc.*, *369*, 207–215.

Kosugi, T., et al. (2007), The Hinode (Solar-B) mission: An overview, *Sol. Phys.*, *243*, 3–17.

Li, Z., Y. Yan, and G. Song (2004), Properties of the boundary integral equation for solar non-constant-α force-free magnetic fields, *Mon. Not. R. Astron. Soc.*, *347*, 1255–1265.

Lin, J., Y.-K. Ko, L. Sui, J. C. Raymond, G. A. Stenborg, Y. Jiang, S. Zhao, and S. Mancuso (2005), Direct observations of the magnetic reconnection site of an eruption on 2003 November 18, *Astrophys. J.*, *622*, 1251–1264.

Liu, Y., and H. Zhang (2001), Relationship between magnetic field evolution and major flare event on July 14, 2000, *Astron. Astrophys.*, *372*, 1019–1029.

Longcope, D. W. (2005), Topological methods for the analysis of solar magnetic fields, *Living Rev. Sol. Phys.*, *2*, 7. (Available at http://www.livingreviews.org/lrsp-2005-7)

Low, B. C., and Y. Q. Lou (1990), Modeling solar force-free magnetic fields, *Astrophys. J.*, *352*, 343–352.

Masuda, S., T. Kosugi, H. Hara, S. Tsuneta, and Y. Ogawara (1994), A loop-top hard X-ray source in a compact solar flare as evidence for magnetic reconnection, *Nature*, *371*, 495–497.

Masuda, S., T. Kosugi, and H. S. Hudson (2001), A hard X-ray two-ribbon flare observed with Yohkoh/HXT, *Sol. Phys.*, *204*, 55–67.

Metcalf, T. R., L. Jiao, A. N. McClymont, R. C. Canfield, and H. Uitenbroek (1995), Is the solar chromospheric magnetic field force-free, *Astrophys. J.*, *439*, 474–481.

Nelder, J. A., and R. Mead (1965), A simplex method for function minimization, *Comput. J.*, *7*, 308–313.

Pevtsov, A. A., R. C. Canfield, and T. R. Metcalf (1995), Latitudinal variation of helicity of photospheric magnetic fields, *Astrophys. J.*, *440*, L109–L112.

Priest, E. R., and T. G. Forbes (2002), The magnetic nature of solar flares, *Astron. Astrophys. Rev.*, *10*, 313–377.

Régnier, S., and T. Amari (2004), 3D magnetic configuration of the Hα filament and X-ray sigmoid in NOAA AR 8151, *Astron. Astrophys.*, *425*, 345–352.

Sakurai, T. (1981), Calculation of force-free magnetic field with non-constant α, *Sol. Phys.*, *69*, 343–359.

Schrijver, C. J., M. L. Derosa, T. R. Metcalf, Y. Liu, J. McTiernan, S. Régnier, G. Valori, M. S. Wheatland, and T. Wiegelmann (2006), Nonlinear force-free modeling of coronal magnetic fields part I: A quantitative comparison of methods, *Sol. Phys.*, *235*, 161–190.

Schwenn, R. (2006), Space weather: The solar perspective, *Living Rev. Sol. Phys.*, *3*, 2. (Available at http://www.livingreviews.org/lrsp-2006-2)

Shibata, K. (2004), Magnetic reconnection in solar and astrophysical plasmas, in *Physics of Magnetic Reconnection in High-Temperature Plasmas*, edited by M. Ugai, pp. 193–237, Res. Signpost, Trivandrum, India.

Shibata, K., S. Masuda, M. Shimojo, H. Hara, T. Yokoyama, S. Tsuneta, T. Kosugi, and Y. Ogawara (1995), Hot-plasma ejections associated with compact-loop solar flares, *Astrophys. J.*, *451*, L83–L85.

Somov, B. V., T. Kosugi, H. S. Hudson, T. Sakao, and S. Masuda (2002), Magnetic reconnection scenario of the Bastille Day 2000 flare, *Astrophys. J.*, *579*, 863–873.

Song, M. T., C. Fang, Y. H. Tang, S. T. Wu, and Y. A. Zhang (2006), A new and fast way to reconstruct a nonlinear force-free field in the solar corona, *Astrophys. J.*, *649*, 1084–1092.

Song, M. T., C. Fang, H. Q. Zhang, Y. H. Tang, S. T. Wu, and Y. A. Zhang (2007), Reconstructing spherical nonlinear force-free field in the solar corona, *Astrophys. J.*, *666*, 491–500.

Su, J. T., and H. Q. Zhang (2007), The effects of polarization crosstalk and solar rotation on measuring full-disk solar photospheric vector magnetic fields, *Astrophys. J.*, *666*, 559–575.

Tian, L., J. Wang, and D. Wu (2002), Non-potentiality of the magnetic field beneath the eruptive filament in the Bastille event, *Sol. Phys.*, *209*, 375–389.

Tsuneta, S. (1996), Structure and dynamics of magnetic reconnection in a solar flare, *Astrophys. J.*, *456*, 840–849.

Tsuneta, S., H. Hara, T. Shimizu, L. W. Acton, K. T. Strong, H. S. Hudson, and Y. Ogawara (1992), Observation of a solar flare at the limb with the YOHKOH Soft X-ray Telescope, *Publ. Astron. Soc. Jpn.*, *44*, L63–L69.

UeNo, S., S. Nagata, R. Kitai, H. Kurokawa, and K. Ichimoto (2004), The development of filter vector magnetographs for the Solar Magnetic Activity Research Telescope (SMART), in *Ground-Based Instrumentation for Astronomy*, edited by A. F. M. Moorwood and M. Iye, *Proc. SPIE*, *5492*, 958–969.

Valori, G., B. Kliem, and R. Keppens (2005), Extrapolation of a nonlinear force-free field containing a highly twisted magnetic loop, *Astron. Astrophys.*, *433*, 335–347.

Valori, G., B. Kliem, and M. Fuhrmann (2007), Magnetofrictional extrapolations of Low and Lou's force-free equilibria, *Sol. Phys.*, *245*, 263–285.

Wang, H., C. Liu, Y. Deng, and H. Zhang (2005), Reevaluation of the magnetic structure and evolution associated with the Bastille day flare on 2000 July 14, *Astrophys. J.*, *627*, 1031–1039.

Wang, H. N., and T. Sakurai (1998), Cross-section variations of coronal magnetic loops, *Publ. Astron. Soc. Jpn.*, *50*, 111–123.

Wang, H. N., Y. H. Yan, T. Sakurai, and M. Zhang (2000), Topology of magnetic field and coronal heating in solar active regions, II. the role of quasi-separatrix layers, *Sol. Phys.*, *197*, 263–273.

Wang, H. N., Y. Yan, and T. Sakurai (2001), Topology of magnetic field and coronal heating in solar active regions, *Sol. Phys.*, *201*, 323–336.

Wang, J., Z. Shi, H. Wang, and Y. Lü (1996), Flares and the magnetic nonpotentiality, *Astrophys. J.*, *456*, 861–878.

Wheatland, M. S. (2006), A fast current-field iteration method for calculating nonlinear force-free fields, *Sol. Phys.*, *238*, 29–39.

Wheatland, M. S., P. A. Sturrock, and G. Roumeliotis (2000), An optimization approach to reconstructing force-free fields, *Astrophys. J.*, *540*, 1150–1155.

Wiegelmann, T. (2004), Optimization code with weighting function for the reconstruction of coronal magnetic fields, *Sol. Phys.*, *219*, 87–108.

Wiegelmann, T. (2007), Computing nonlinear force-free coronal magnetic fields in spherical geometry, *Sol. Phys.*, *240*, 227–239.

Wiegelmann, T., B. Inhester, and T. Sakurai (2006), Preprocessing of vector magnetograph data for a nonlinear force-free magnetic field reconstruction, *Sol. Phys.*, *233*, 215–232.

Yan, Y. (2005), On the coronal magnetic field configuration and solar flare/CME process, *Space Sci. Rev.*, *121*, 213–221.

Yan, Y., and Z. Li (2006), Direct boundary integral formulation for solar non-constant-α force-free magnetic fields, *Astrophys. J.*, *638*, 1162–1168.

Yan, Y., and T. Sakurai (2000), New boundary integral equation representation for finite energy force-free magnetic fields in open space above the sun, *Sol. Phys.*, *195*, 89–109.

Yan, Y., M. J. Aschwanden, S. Wang, and Y. Deng (2001), Evolution of magnetic flux rope in the active region NOAA 9077 on 14 July 2000, *Sol. Phys.*, *204*, 27–40.

Zhang, H. (2002), Magnetic field, helicity and the 2000 July 14 flare in solar active region 9077, *Mon. Not. R. Astron. Soc.*, *332*, 500–512.

Zhang, H. Q., et al. (2007), Solar magnetism and the activity telescope at HSOS, *Chin. J. Astron. Astrophys.*, *7*, 281–288.

Zhang, J., J. Wang, Y. Deng, and D. Wu (2001), Magnetic flux cancellation associated with the major solar event on 2000 July 14, *Astrophys. J.*, *548*, L99–L102.

Zhang, M., and B. C. Low (2005), The hydromagnetic nature of solar coronal mass ejections, *Annu. Rev. Astron. Astrophys.*, *43*, 103–137.

Zhao, H., J. X. Wang, J. Zhang, and C. J. Xiao (2005), A new method of identifying 3D null points in solar vector magnetic fields, *Chin. J. Astron. Astrophys.*, *5*, 443–447.

Zhou, G. P., J. X. Wang, and J. Zhang (2006a), Large-scale source regions of earth-directed coronal mass ejections, *Astron. Astrophys.*, *445*, 1133–1141.

Zhou, G. P., Y. M. Wang, and J. X. Wang (2006b), Coronal mass ejections associated with polar crown filaments, *Adv. Space Res.*, *38*, 466–469.

―――――――――――

H. He and H. Wang, National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, China. (hehan@bao.ac.cn; hnwang@bao.ac.cn)

Geomagnetically induced currents during intense storms driven by coronal mass ejections and corotating interacting regions

Ryuho Kataoka[1,2] and Antti Pulkkinen[3,4]

[1]   Geomagnetically induced currents (GIC) flowing in ground-based conductor systems during large geomagnetic storms are one of the most significant space weather phenomena that affect our ground-based technological systems. Here we show that GIC activity in subauroral latitudes depends on the storm phase and on the interplanetary drivers, such as coronal mass ejections (CMEs) and corotating interacting regions (CIRs). For example, it is shown that GIC amplitudes are relatively small during CIR storms in comparison to CME storms. However, Pc3-5 pulsation activity during CIR storms drives long-lasting GIC in the local prenoon sector. Despite of the differences between CME and CIR storms, the relationship between GIC and the time derivative of the horizontal ground magnetic field is always the same. A novel power law equation is derived to accurately characterize the relation.

1.   Introduction

[2]   Geomagnetically induced currents (GICs) flowing in long technological conductor networks on the ground are one of the manifestations of geomagnetic storms and can be thought as an end link of the chain of space weather processes from the surface of the Sun to the surface of the Earth. GIC poses a potential threat to the normal operation of technological systems such as power lines, pipelines, and railway systems [e.g., *Lanzerotti*, 1979; *Boteler et al.*, 1998].

[3]   The fundamental challenge of GIC research is to identify, understand, and model the different geophysical processes associated with large GIC events. The present understanding is that a number of different geophysical processes are capable of driving large GICs; storm sudden commencements (SSC), geomagnetic pulsations, and auroral substorms have been identified as important causes for large GICs [*Boteler*, 2001; *Lam et al.*, 2002; *Kappenman*, 2003; *Pulkkinen et al.*, 2003, 2005].

[4]   Most of the earlier studies, like those cited above, on the ionospheric and magnetospheric drivers of GICs have been more or less event based and did not provide any direct means for generalizations. Rigorous statistical analysis of GIC and the time derivative of the horizontal ground magnetic field (denoted hereafter dB/dt), a quantity closely coupled to GIC via Faraday's law of induction, have been

carried out by *Viljanen* [1997], *Viljanen et al.* [2001, 2006], *Weigel et al.* [2002], *Weigel and Baker* [2003], *Wintoft* [2005], and *Pulkkinen et al.* [2006].

[5]   However, basic characteristics of GIC during intense storms are still not well known. Recently, *Pulkkinen and Kataoka* [2006] performed a time-frequency analysis of GICs during the 10 greatest superstorms to show a clear local time and storm phase dependence of the GIC spectra. Such superstorms are driven by fast coronal mass ejections (CMEs) [e.g., *Kataoka et al.*, 2005]. *Miyoshi and Kataoka* [2005] showed evidence that there are significant differences in the magnetospheric response between intense storms driven by CMEs and corotating interaction regions (CIRs), and therefore the GIC response is also expected to be different for the CME and CIR storms. The purpose of this study is to investigate the possible differences of GIC variations between CME and CIR storms.

2.   Data and the Analysis Method

[6]   The main data to be analyzed is composed of GIC measurements carried out since November 1998 in the Finnish natural gas pipeline at the Mäntsälä pipeline section [*Pulkkinen et al.*, 2001]. The magnetic latitude of the measurement site is about 57 degrees, and the magnetic local time is about UT plus 3 h. The GIC data are obtained with a 10 s temporal resolution. The IMAGE magnetometer chain [*Lühr et al.*, 1998] is used to provide more general geomagnetic context. Nurmijärvi Geophysical Observatory (NUR), a part of the IMAGE chain, is located about 30 km southwest from the GIC measurement site. Also, the magnetometer data are obtained with a 10 s temporal resolution. Table 1 lists the coordinates of the IMAGE magnetometer stations used in this study.

[7]   Figure 1 shows a correlation between minimum *Dst* and maximum GIC for the isolated intense storms (mini-

---

[1]Computational Astrophysics Laboratory, Institute of Physics and Chemical Research (RIKEN), Saitama, Japan.
[2]Solar-Terrestrial Laboratory, Nagoya University, Nagoya, Japan.
[3]Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, Maryland, USA.
[4]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Table 1. Coordinates of IMAGE Magnetometer Stations Used in This Study

Abbreviation	Name	GEO Latitude	GEO Longitude	CGM Latitude	CGM Longitude
NAL	Ny Ålesund	78.92	11.95	75.25	112.08
LYR	Longyearbyen	78.20	15.82	75.12	113.00
HOR	Hornsund	77.00	15.60	74.13	109.59
SOR	Sørøya	70.54	22.22	67.34	106.17
KIL	Kilpisjärvi	69.02	20.79	65.88	103.79
MUO	Muonio	68.02	23.53	64.72	105.22
PEL	Pello	66.90	24.08	63.55	104.92
OUJ	Oulujärvi	64.52	27.23	60.99	106.14
HAN	Hankasalmi	62.30	26.65	58.71	104.61
NUR	Nurmijärvi	60.50	24.65	56.89	102.18
TAR	Tartu	58.26	26.46	54.47	102.89

mum $Dst < -100$ nT) during solar cycle 23 (years 1996 to 2005) as identified by *Kataoka and Miyoshi* [2006]. Continuous GIC data are available for 32 CME storms and 3 CIR storms, as indicated by open and solid circles, respectively. The superstorms investigated by *Pulkkinen and Kataoka* [2006] are included in the CME storms only if the superstorms are isolated without multiple occurrences of intense storms within 4 d. The maximum amplitude of GIC is calculated for the time interval 1 d before and after the *Dst* minima. The correlation coefficient is 0.67 for CME storms, implying that large GIC events tend to be associated with large CME storms.

[8] As is seen from Figure 1, the three CIR storms are smaller in amplitude both in terms of *Dst* and GIC. The small number of CIR storms is due to the fact that CIRs rarely produce intense storms. In fact, typical CIRs have a limited interplanetary magnetic field strength of up to about 20 nT [*Richardson et al.*, 2006]. However, since intense CIR storms can be a potential cause for large GIC events, it is worthwhile to include these events in our analysis.

[9] Generally, the waveform of GICs tends to be very noisy and nonstationary [*Pulkkinen and Kataoka*, 2006]. The S-transform [*Stockwell et al.*, 1996], capable of handing the noisy nonstationary data, is a time-frequency analysis of a time series h(t) and is defined as:

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{\frac{-f^2(\tau-t)^2}{2}} e^{-i2\pi ft} dt, \qquad (1)$$

where τ is time and f is frequency. The kernel of the S-transform is defined by a Gaussian with the window width of 1/f to give the best resolution in time-frequency space. The normalized Gaussian kernel also gives a direct mathematical connection to the Fourier transform via the integration over τ.

[10] Before calculating the S-transform spectra of GIC for all 24-h long segments centered around *Dst* minima of storm events, there are three preprocessing steps as follows: (1) We include the margins of 2 h time series just before and after the 24-h time series to reduce an artificial edge effects; (2) Hanning window is applied on the 5% of the total 28 × 360 points at the edges to further reduce the edge effect; (3) Hilbert transform is applied to obtain the analytic signal in which the spectral power is localized into positive frequencies. Finally, we apply the S-transform to the ana-

lytic signal, and the amplitude $|S(\tau, f)|$ is color coded to display the S-transform spectra.

3. Results

[11] First, we show typical examples of the GIC spectra during comparable nightside CME and CIR storm events. The *Dst* minima for the selected CME and CIR storms are -110 nT at 1930 UT (2230 MLT) on 11 May 2002 and -109 nT at 1730 UT (2030 MLT) on 11 February 2004, respectively. Using the nightside events, we can directly compare the results with the results for superstorms, as shown by *Pulkkinen and Kataoka* [2006].

[12] Figures 2 and 3 show GIC and the S-transform of GIC for the selected events. It is seen that around the storm peak taking place in the nightside, turbulent broadband spectra are seen in both CME and CIR storms. After the storm peak, the CIR storm has a stronger and longer-lasting GIC activity than that of the CME storm even though the *Dst* minima for the storms are similar to each other. Local daytime GIC in Pc3-5 pulsation ranges (0.3–10 min) are
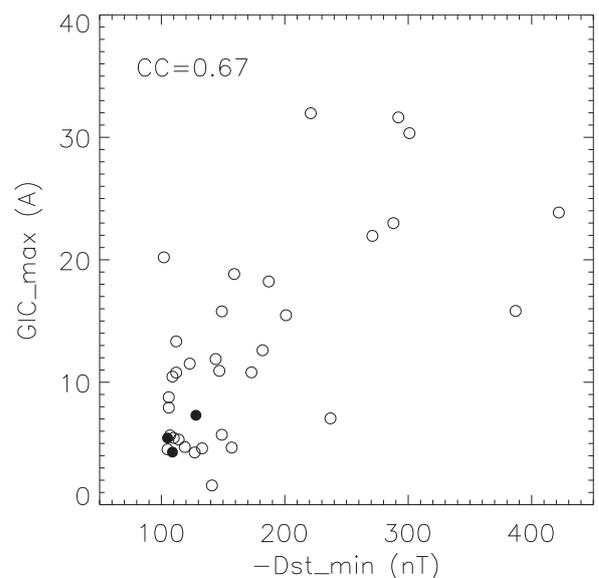
Figure 1. Scatterplot of maximum |*Dst*| and |GIC| for each of the *Dst* < −100 nT storms used in this study. Coronal mass ejection (CME) storms are indicated by open circles and corotating interaction regions (CIR) storms by solid circles.
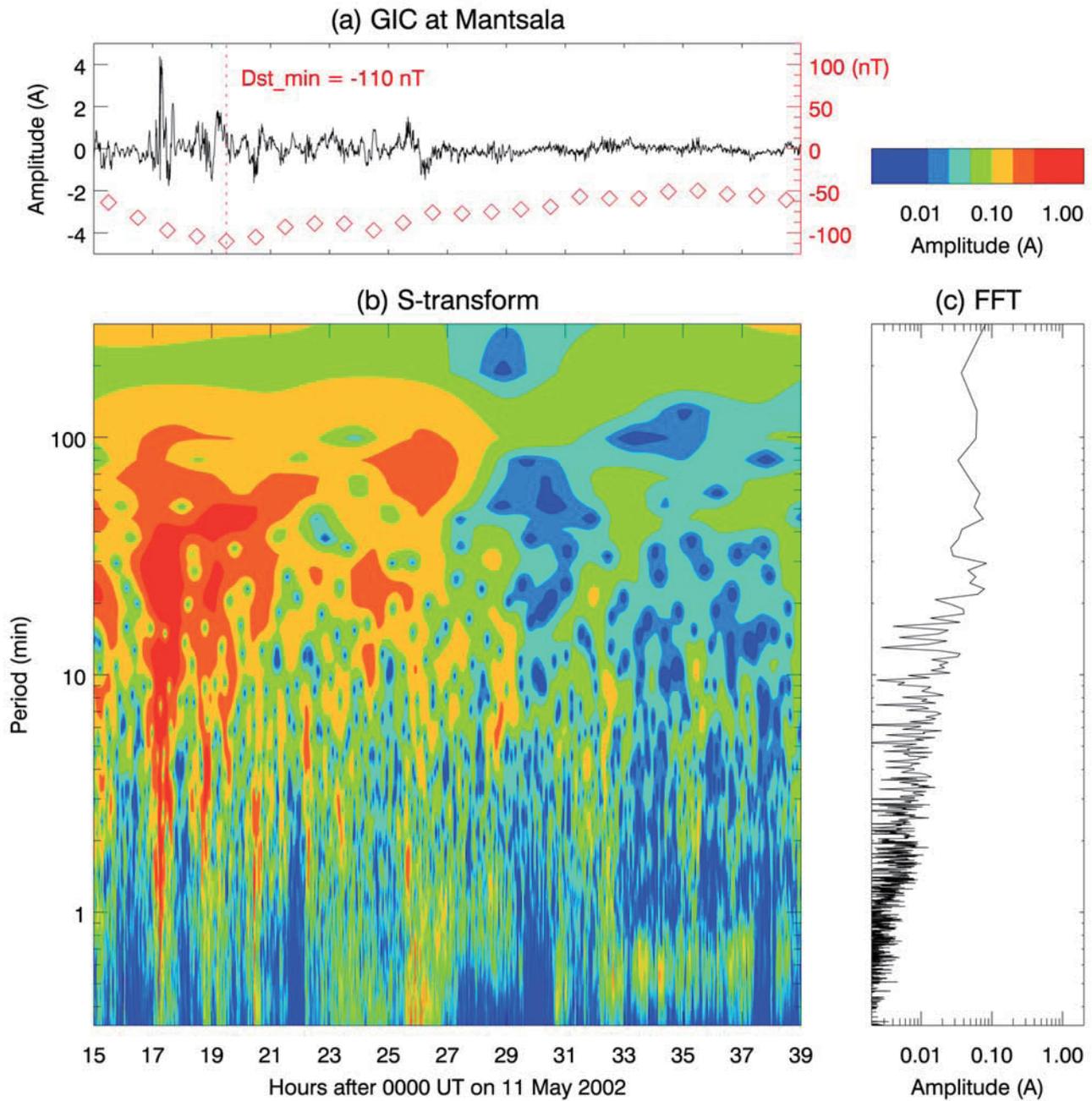
Figure 2. S-transform spectrum of a CME storm on 11 May 2002 showing (a) the waveform of geomagnetically induced currents (GIC), (b) the S-transform spectral power, and (c) the FFT power spectrum obtained by integrating the S-transform over time. The left half of Figure 2b is the local nightside (1800–0600 MLT), and the right half is the local dayside (0600–1800 MLT). Red diamonds indicate the *Dst* index. The storm peak is shown by a vertical dotted line.

especially enhanced during the recovery phase of the CIR storm. Comparing the Figures 2c and 3c, the GIC enhancement in the Pc3-5 range during the CIR storm can also be seen from the FFT spectra. Strong monochromatic GIC in the Pc5 range (2.5–10 min), as was seen for superstorms by *Pulkkinen and Kataoka* [2006], is not present during the recovery phase of the CME storm. Although not shown here, the spectral features observed above are common also for the other CME and CIR storms studied in this work.

[13] We then compare the time derivative of the horizontal magnetic field (dB/dt) observed at NUR with the GIC using all of the CME and CIR storm events. The time derivative is calculated by using the three-point central difference formula. To compare the GIC and dB/dt, the data are divided into 1 h intervals, each containing 360 sample points. The maximum amplitudes of |dB/dt| and |GIC| are calculated in each interval, and 1 h values are obtained for four different time intervals: just after the storm
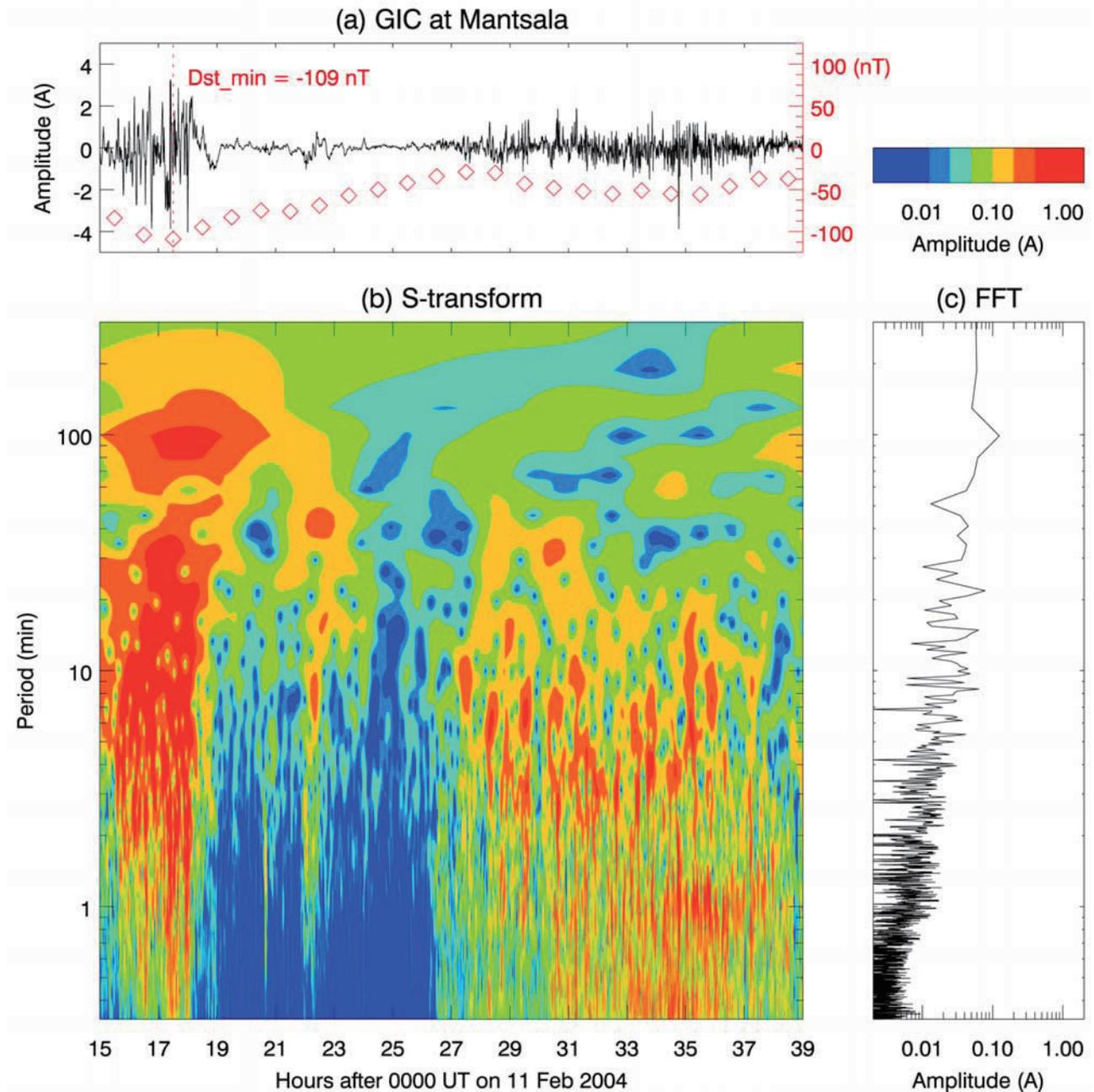
Figure 3. S-transform spectrum of a CIR storm on 11 February 2004. The format is the same as in Figure 2.

sudden commencement (SSC), main phase, recovery phase, and 24 h after the end of the recovery phase.

[14] In this paper, the beginning of the main phase is determined by the existence of minimum d(Dst)/dt or two successive points of d(Dst)/dt < −20 nT/h, where the time derivative is calculated using a two-point forward difference formula. If it is not possible to detect the beginning by these criteria, a typical time period of 6 h before the Dst minimum is taken as the main phase interval. The end of the recovery phase is determined by the recovery of the Dst index to 50% of the Dst minimum. If it is not possible to detect the end by these criteria, typical time period of 12 h after the Dst minimum is taken as the recovery phase interval.

[15] From Figure 4 showing the scatterplot and histograms of maximum |dB/dt| and |GIC| for CME storms, it is seen that from the main phase (red) to the recovery phase (blue) the amplitudes of |GIC| and |dB/dt| decrease more than a half an order of magnitude on average. Further, from the recovery phase (blue) to the 24 h after the end of the recovery phase (green), the amplitudes decrease again more than a half an order of magnitude on average. However, independent of the storm phase, all of the data points distribute around the same regression line. The correlation coefficient is better for |dBy/dt| (CC = 0.97) than that of |dBx/dt| (CC = 0.94), where the subscripts x and y denote the geographic north-south and east-west components, re-
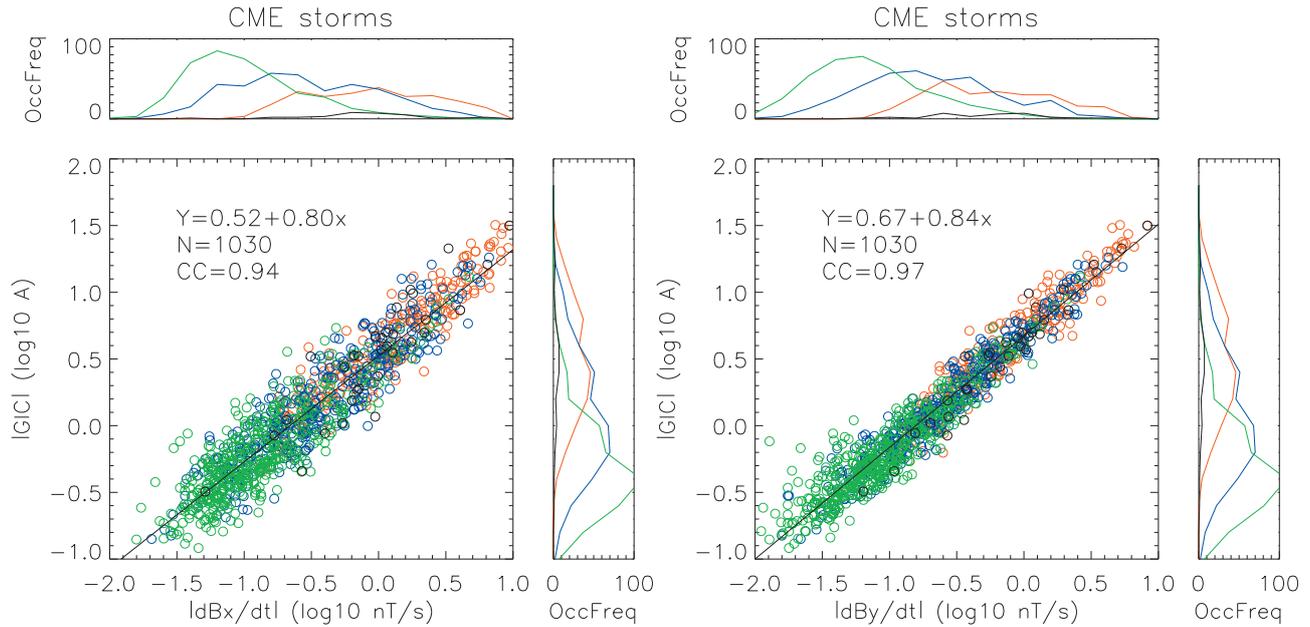
Figure 4. Scatterplot and histograms of maximum |dB/dt| showing (left) the x-component and (right) y-component and maximum |GIC| during the main phase (red), recovery phase (blue), 24 h after the end of recovery phase (green), and just after the storm sudden commencements (black) of CME storms.

spectively. The difference in correlation clearly reflects the geographical orientation of the pipeline; GIC flow in the pipeline responds differently for east-west and north-south geoelectric field in association with the temporal changes of the magnetic field [*Pulkkinen et al.*, 2001]. The relationship between |dBy/dt| and |GIC| can be approximated by power-law equation,

$$|GIC|(A) = 10^{0.67}|dB_y/dt|^{0.84}(nT/s). \qquad (2)$$

From Figure 5 showing the scatter plot and histograms of maximum |dB/dt| and |GIC| for CIR storms, it is seen that although the correlation between the two quantities is slightly worse, the basic trend of the regression line is the same as that for the CME storms. Again, the correlation coefficient is better for |dBy/dt| (CC = 0.91) than that of |dBx/dt| (CC = 0.78). The approximately log-normal histograms in Figures 4 and 5 suggest that the GIC amplitude rarely exceeds the 10 A level even in the main phase of CIR storms. From Figure 5 it is also seen that the storm phase dependence of the |dB/dt| and |GIC| amplitudes
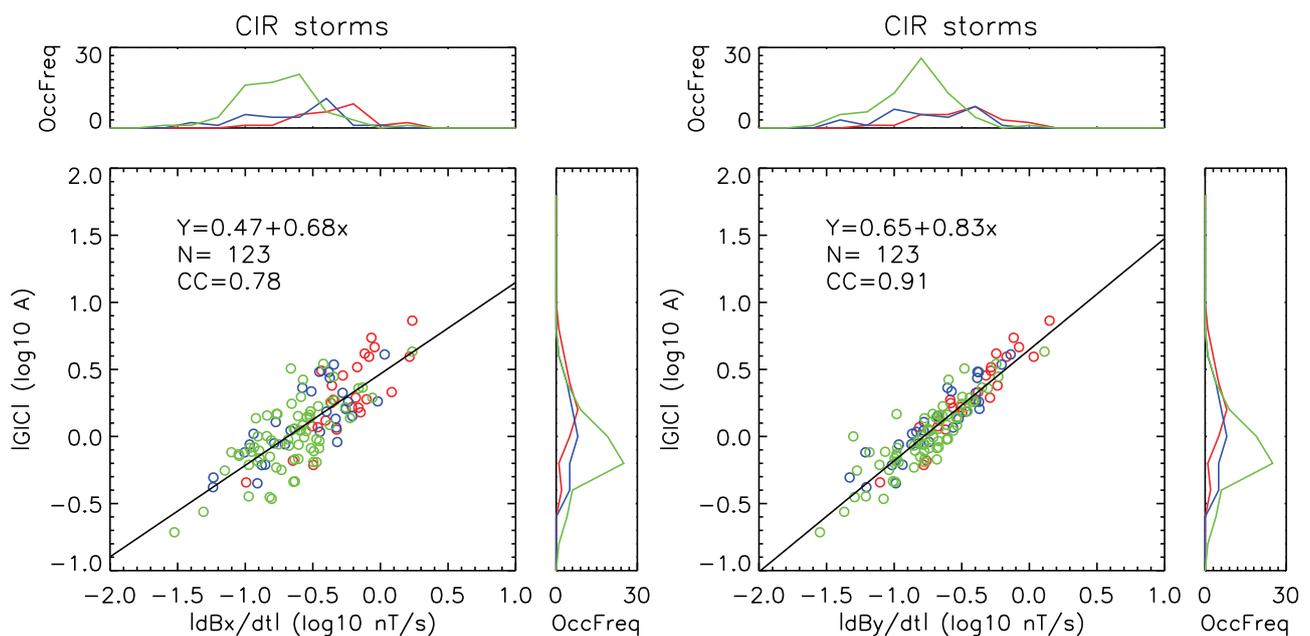
Figure 5. Same as Figure 4 but for CIR storms.

Table 2. Statistical Parameters of |GIC| Amplitudes for Different Storm Phases[a]

Interplanetary Driver	Storm Phase	Total Number of Storms	Total Number of Hours	Average (A)	Maximum (A)
CME	SSC	32	32	3.3	31.6
CME	main	32	247	3.8	32.0
CME	recovery	32	380	1.5	18.2
CME	after recovery	32	403	0.6	13.0
CIR	main	3	25	2.0	7.3
CIR	recovery	3	29	1.3	4.1
CIR	after recovery	3	69	1.0	4.3

[a]From left to right, columns show the storm driver, storm phase, the total number of the storm events, the total number of sample hours, mean values, and maximum values.

is relatively small for CIR storms, ranging well below an order of magnitude on average. Comparing the recovery phase (blue) and late recovery phase (green) in Figures 4 and 5, it is apparent, in agreement with Figures 2 and 3, that the |dB/dt| and |GIC| amplitudes during the recovery and late recovery phases are larger for CIR storms on average. Some statistical parameters of |GIC| are summarized in Table 2. The mean values are calculated after taking the logarithm. This is reasonable since the hourly values of both |GIC| and |dB/dt| follow the log-normal distribution as show in the histograms in Figures 4 and 5.

4. Discussion

[16] Figures 2 and 3 imply that partially different physical mechanisms drive GIC during CME and CIR storms. The most interesting observation is that, during CIR storms, GIC is enhanced in the Pc3-5 pulsation range, especially in the local dayside during the recovery phase. The Pc3-5 pulsations are related to the high-speed Alfvénic solar wind streams originating from large coronal holes [see *Richardson et al.*, 2006]. Thus these coronal hole streams are the origin for the long-lasting GIC activity during CIR storms. On the other hand, for the February 2004 event shown in Figure 3, the very quiet period of GIC lasting 8 h just after the storm peak was associated with unusually stable northward interplanetary magnetic fields embedded within the CIR and is not necessarily a general feature of CIR-associated storms.

[17] The well-known correlation of Pc3-5 pulsation wave power with the solar wind speed [*Singer et al.*, 1977; *Engebretson et al.*, 1998], suggests that the Kelvin-Helmholtz instabilities (KHI) at the magnetopause may act as a possible energy source for the pulsations. Actually, using multiple in situ satellite observations, *Rae et al.* [2005] investigated very strong narrowband monochromatic Pc5 pulsation, as seen in GIC during superstorms [*Pulkkinen and Kataoka*, 2006] and concluded that the discrete Pc5 pulsation is the result from the excitation of a magnetospheric waveguide mode driven by KHI at the magnetopause. Other mechanisms, however, possibly contribute to large GIC in the Pc3-5 range. On the basis of the IMAGE magnetometer data, *Howard and Menk* [2005] suggested that the daytime Pc3-4 waves as observed in this paper are generated by the upstream ion-cycrotron resonance and that there is no evidence of generation by the KHI. *Baker et al.* [2003] surveyed a 10-year data set of CANOPUS to statistically investigate the Pc5 pulsation and suggested from the property of field-line resonance that the energy

source of Pc5 pulsation is not only KHI but also that a significant part of the Pc5 pulsations should have some irregular or impulsive energy source such as traveling indentations on the magnetopause [*Mathie and Mann*, 2000] or solar wind buffeting. They also noted that transient events such as magnetic impulse events [e.g., *Kataoka et al.*, 2001] contribute to the largest power of Pc5 spectra in a limited dawn sector between 65 to 70 magnetic latitudes.

[18] Despite the spectral differences found between CME and CIR storms, hourly dB/dt is found always to be an excellent indicator of the hourly GIC activity, and the relationship between the GIC and the dB/dt is always the same in a very large dynamic range of about three orders of magnitude, as shown in Figures 4 and 5. This finding is in agreement with earlier results [see, e.g., *Viljanen et al.*, 2001, and references therein]; the novel result is that the relationship can be made very accurate by using the functional form given in equation (2). It should be noted here that there are examples [*Trichtchenko and Boteler*, 2006] where GIC behavior follows the magnetic field strength (B) rather than its time derivative. However, by making plots similar to Figures 4 and 5 for B, we have confirmed that dB/dt is a better indicator for GIC activity than B at Mäntsälä.

[19] For additional insights to |dB/dt| behavior, as shown in Figure 6, we construct a polar map showing the global distribution of average amplitude of the hourly maximum |dB/dt| and root mean square of the hourly maximum |dBx/dt| and |dBy/dt|. We use the eleven meridional magnetometers listed in Table 1. All six CIR events in the list of *Kataoka and Miyoshi* [2006] and an additional recent CIR event that occurred in the end of August 2005 are used in constructing the map. However, even with the additional CIR events the number of data points was, unfortunately, too small to construct a map for the CIR main phase. A total of 11 × 24 average points (11 points in latitude and 24 points in MLT) are linearly interpolated and color contoured. Note again that the average values are calculated after taking the logarithm.

[20] *Weigel et al.* [2002, 2003] computed similar average polar maps of |dB/dt|, but they did not make a classification of different storm phases. The results here indicate that it may be beneficial to separate predictions of |dB/dt| by different classes based on, for example, storm phase and interplanetary driver. For example, the prenoon peak of CIR storms (Figure 6c) is about twice as strong on average than that of CME storms (Figure 6b). Further, comparing the Figures 6a and 6b, the local time sector having the greatest activity may vary as a function of CME storm phase, for
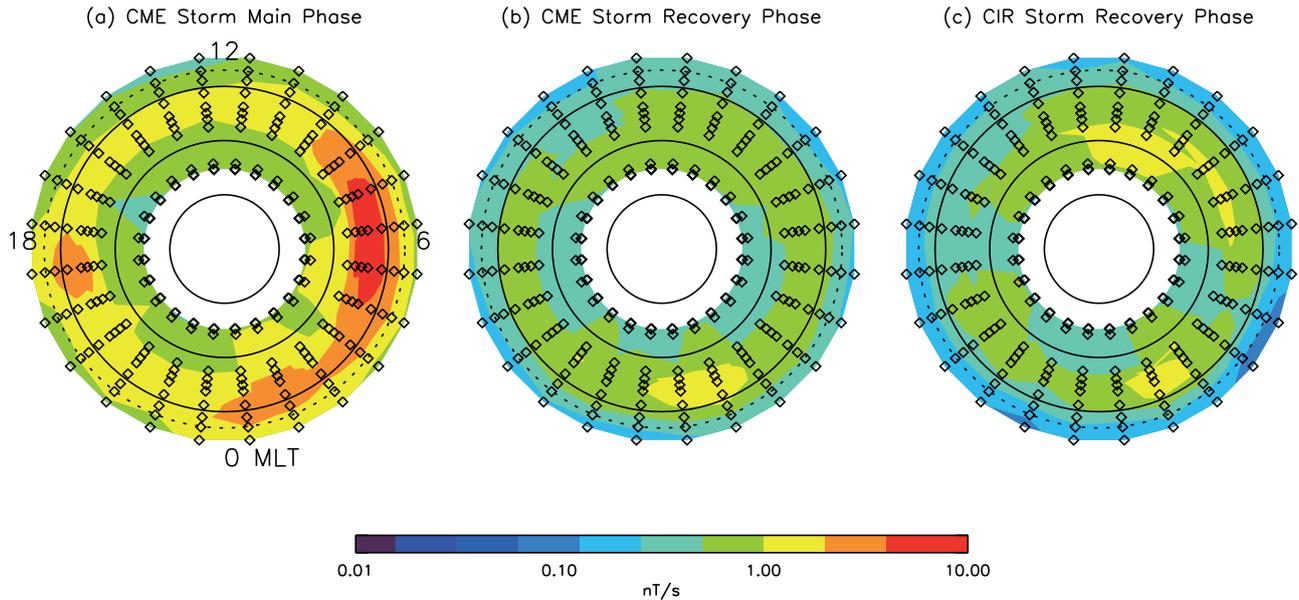
Figure 6. Polar map of the average amplitude of the hourly maximum |dB/dt| during (a) the main phase and (b) the recovery phase of CME storms and (c) during and 24 h after the recovery phase of CIR storms. Solid lines indicate the 80, 70, and 60 MLAT, and the dotted line indicates the latitude of the NUR station. Open squares indicate the meridional locations of the IMAGE chain.

example, the greatest activity is appeared in the dawn and postmidnight sector during the main and recovery phase, respectively.

5. Conclusions

[21] It was shown that subauroral GIC activity depends on the solar wind driver and phases of strong geomagnetic storms. For example, GIC amplitudes are relatively small during CIR storms in comparison to CME storms. However, Pc3-5 pulsation activity during CIR storms drives long-lasting GIC in the local prenoon sector. Despite of the differences, the maximum hourly value of |dB/dt| is always an excellent indicator for the maximum hourly amplitude of |GIC| in the Finnish pipeline for any local time and any storm phase of CME/CIR storms. A novel power law equation was derived to accurately characterize the relationship between |dB/dt| and |GIC|.

References

Baker, G. J., E. F. Donovan, and B. J. Jackel (2003), A comprehensive survey of auroral latitude Pc5 pulsation characteristics, *J. Geophys. Res.*, *108*(A10), 1384, doi:10.1029/2002JA009801.

Boteler, D. (2001), Assessment of geomagnetic hazards to power systems in Canada, *Natural Hazards*, *23*, 101–120.

Boteler, D. H., R. J. Pirjola, and H. Nevanlinna (1998), The effects of geomagnetic disturbances on electrical systems at the Earth's surface, *Adv. Space Res.*, *22*, 17–27.

Engebretson, M., K.-H. Glassmeier, M. Stellmacher, W. J. Hughes, and H. Luhr (1998), The dependence of a high latitude Pc5 wave power on solar wind velocity and on the phase of high speed solar wind streams, *J. Geophys. Res.*, *103*, 26,271.

Howard, T. A., and F. W. Menk (2005), Ground observations of high-latitude Pc3-4 ULF waves, *J. Geophys. Res.*, *110*, A04205, doi:10.1029/2004JA010417.

Kappenman, J. G. (2003), Storm sudden commencement events and the associated geomagnetically induced current risks to ground-based systems at low-latitude and midlatitude locations, *Space Weather*, *1*(3), 1016, doi:10.1029/2003SW000009.

Kataoka, R., and Y. Miyoshi (2006), Flux enhancement of radiation belt electrons during geomagnetic storms driven by coronal mass ejections and corotating interaction regions, *Space Weather*, *4*, S09004, doi:10.1029/2005SW000211.

Kataoka, R., H. Fukunishi, L. J. Lanzerotti, C. G. Maclennan, H. U. Frey, S. B. Mende, J. H. Doolittle, T. J. Rosenberg, and A. T. Weatherwax (2001), Magnetic Impulse Event: A detailed case study of extended ground and space observations, *J. Geophys. Res.*, *106*, 25,873–25,889.

Kataoka, R., S. Watari, N. Shimada, H. Shimazu, and K. Marubashi (2005), Downstream structures of interplanetary fast shocks associated with coronal mass ejections, *Geophys. Res. Lett.*, *32*, L12103, doi:10.1029/2005GL022777.

Lam, H.-L., D. H. Boteler, and L. Trichtchenko (2002), Case studies of space weather events from their launching on the Sun to their impacts on power systems on the Earth, *Ann. Geophys.*, *20*, 1073–1079.

Lanzerotti, L. J. (1979), Geomagnetic influences on man-made systems, *J. Atmos. Terr. Phys.*, *41*, 787–796.

Lühr, H., A. Aylward, S. C. Bucher, A. Pajunpaa, K. Pajunpaa, T. Holmboe, and S. M. Zalewski (1998), Westward moving dynamic substorm features observed with the IMAGE magnetometer network and other ground-based instruments, *Ann. Geophys.*, *16*, 425–440.

Mathie, R. A., and I. R. Mann (2000), Observations of Pc5 field line resonance azimuthal phase speeds: A diagnostic of their excitation mechanism, *J. Geophys. Res.*, *105*, 10,713–10,728.

Miyoshi, Y., and R. Kataoka (2005), Ring current ions and radiation belt electrons during geomagnetic storms driven by coronal mass ejections and corotating interaction regions, *Geophys. Res. Lett.*, *32*, L21105, doi:10.1029/2005GL024590.

Pulkkinen, A., and R. Kataoka (2006), S-transform view of geomagnetically induced currents during geomagnetic superstorms, *Geophys. Res. Lett.*, *33*, L12108, doi:10.1029/2006GL025822.

Pulkkinen, A., A. Viljanen, K. Pajunpää, and R. Pirjola (2001), Recordings and occurrence of geomagnetically induced currents in the Finnish natural gas pipeline network, *J. Appl. Geophys.*, *48*, 219–231.

Pulkkinen, A., A. Thomson, E. Clarke, and A. McKay (2003), April 2000 storm: ionospheric drivers of large geomagnetically induced currents, *Ann. Geophys.*, *21*, 709–717.

Pulkkinen, A., S. Lindahl, A. Viljanen, and R. Pirjola (2005), Geomagnetic storm of 29–31 October 2003: Geomagnetically induced currents and their relation to problems in the Swedish high-voltage power transmission system, *Space Weather*, *3*, S08C03, doi:10.1029/2004SW000123.

Pulkkinen, A., A. Viljanen, and R. Pirjola (2006), Estimation of geomagnetically induced current levels from different input data, *Space Weather*, *4*, S08005, doi:10.1029/2006SW000229.

Rae, I. J., et al. (2005), Evolution and characteristics of global Pc5 ULF waves during a high solar wind speed interval, *J. Geophys. Res.*, *110*, A12211, doi:10.1029/2005JA011007.

Richardson, I. G., et al. (2006), Major geomagnetic storms (Dst < −100 nT) generated by corotating interaction regions, *J. Geophys. Res.*, *111*, A07S09, doi:10.1029/2005JA011476.

Singer, H. J., C. T. Russel, M. G. Kivelson, E. Q. Greenstadt, and J. V. Olson (1977), Evidence for the control of Pc3-4 magnetic pulsations by the solar wind velocity, *Geophys. Res. Lett.*, *4*, 377–381.

Stockwell, R. G., L. Manshinha, and R. P. Lowe (1996), Localization of the complex spectrum: The S-transform, *IEEE Trans. Signal Proc.*, *44*(4), 998–1001.

Trichtchenko, L., and D. Boteler (2006), Response of power systems to the temporal characteristics of geomagnetic storms, paper presented at Canadian Conference on Electrical and Computer Engineering, IEEE, Ottawa, Ontario, Canada.

Viljanen, A. (1997), The relation between geomagnetic variations and their time derivatives and implications for estimation of induction risks, *Geophys. Res. Lett.*, *24*, 631–634.

Viljanen, A., H. Nevanlinna, K. Pajunpää, and A. Pulkkinen (2001), Time derivative of the horizontal magnetic field as an activity indicator, *Ann. Geophys.*, *19*, 1107–1118.

Viljanen, A., E. I. Taskanen, and A. Pulkkinen (2006), Relation between substorm characteristics and rapid temporal variations of the ground magnetic field, *Ann. Geophys.*, *24*, 725–733.

Weigel, R. S., and D. N. Baker (2003), Probability distribution invariance of 1-minute auroral-zone geomagnetic field fluctuations, *Geophys. Res. Lett.*, *30*(23), 2193, doi:10.1029/2003GL018470.

Weigel, R. S., D. Vassiliadis, and A. J. Klimas (2002), Coupling of the solar wind to temporal fluctuations in ground magnetic fields, *Geophys. Res. Lett.*, *29*(19), 1915, doi:10.1029/2002GL014740.

Weigel, R. S., A. J. Klimas, and D. Vassiliadis (2003), Solar wind coupling to and predictability of ground magnetic fields and their time derivatives, *J. Geophys. Res.*, *108*(A7), 1298, doi:10.1029/2002JA009627.

Wintoft, P. (2005), Study of the solar wind coupling to the time difference horizontal geomagnetic field, *Ann. Geophys.*, *23*, 1949–1957.

R. Kataoka, Computational Astrophysics Laboratory, Institute of Physics and Chemical Research (RIKEN), 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. (ryuho@riken.jp)

A. Pulkkinen, NASA Goddard Space Flight Center, Code 674, Greenbelt, MD 20771, USA. (antti.pulkkinen@gsfc.nasa.gov)

Properties of *AE* indices derived from real-time global simulation and their implications for solar wind-magnetosphere coupling

K. Kitamura,[1,2] H. Shimazu,[1,3] S. Fujita,[3,4] S. Watari,[1] M. Kunitake,[1] H. Shinagawa,[1,3] and T. Tanaka[3,5]

[1] Real-time magnetohydrodynamic (MHD) simulation of the solar wind-magnetosphere-ionosphere (S-M-I) coupling system was used to calculate auroral electrojet (*AE*) indices. This simulation reproduces the magnetic field configurations in the magnetosphere, magnetospheric convection, and field-aligned currents (FACs) using the upstream boundary conditions with the interplanetary magnetic field (IMF), solar wind speed, temperature, and proton number density measured by the ACE spacecraft. The electrical potential at 3 R_E (Earth radii) from the center of the Earth is mapped on the ionosphere. The ionospheric currents are deduced from Ohm's law to match the divergence of Pedersen and Hall currents from FACs. The *AE* indices are obtained from the magnetic field perturbation caused by the simulated ionospheric currents. We compared the simulated *AE* indices for 247 d with the *AE* indices deduced from the magnetic variations at up to 12 stations located around the auroral latitude. The results show that the simulated *AE* reproduces the observed *AE* indices well. Of the 247 d, 64% had cross-correlation coefficients of more than 0.5. We also found that the simulated *AE* indices do not correlate well with the observed *AE* indices when the standard deviations of variations in the observed *AE* indices are less than 100 nT. When variations in the *AE* indices are small, some of the short-period perturbations of the electromagnetic energy flowing from the solar wind into the magnetosphere is absorbed or filtered in the real S-M-I coupling system by some mechanism that is not included in our MHD simulation and that the resulting fluctuation in the *AE* indices is damped compared with the simulation.

1. Introduction

[2] Recent improvements in computer calculation speeds have made it possible for simulations of large-scale phenomena to be done. Global magnetohydrodynamics (MHD) simulation has played a significant role in the field of the interaction between the solar wind and the Earth's magnetosphere. It has revealed the shape of the magnetosphere, the structure of the distant magnetic tail, and how the closure of the Region-1 field-aligned currents (FACs) are reproduced [*Walker et al.*, 1993; *Fedder and Lyon*, 1995; *Raeder et al.*, 1995; *Spicer et al.*, 1996; *Elsen and Winglee*,

1997; *Janhunen and Koskinen*, 1997; *White et al.*, 1998; *Song et al.*, 1999].

[3] The ACE spacecraft's provision of real-time solar wind data makes it possible to run a global MHD model of the magnetosphere in real time. We have developed a model that simulates real-time global MHD of the solar wind interaction with the magnetosphere. The MHD code used here was developed by *Tanaka* [1994] to enable highly accurate calculations of the MHD simulation for systems with a potential magnetic field (intrinsic geomagnetic field). *Tanaka*'s [1995] MHD simulation reproduced the FACs in the magnetosphere and indicated that the distribution of the simulated FACs has signatures typical of the current patterns of regions 1 and 2, which depend on the *Z*-component of the interplanetary magnetic field (IMF). Tanaka also advanced our understanding of the structure and origin of magnetospheric convection [*Tanaka*, 1999].

[4] Another recent real-time simulation model of the global magnetosphere has being implemented at the Community Coordinated Modeling Center (CCMC) (http://ccmc.gsfc.nasa.gov) [*Narock et al.*, 2004]. This model is based on the space plasma simulation model of the Space Weather Modeling Framework (SWMF) at the University of

---

[1]National Institute of Information and Communications Technology, Tokyo, Japan.

[2]Now at Tokuyama College of Technology, Yamaguchi, Japan.

[3]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Saitama, Japan.

[4]Meteorological College, Chiba, Japan.

[5]Department of Earth and Planetary Sciences, Kyushu University, Fukuoka, Japan.

Michigan [*Tóth et al.*, 2005]. These kinds of real-time models are useful both in predicting the space weather environment and in studying physical processes in space. Another successful global magnetospheric model that can also be used to derive global geomagnetic indices is the Open Global Geospace Circulation Model (openGGCM), which was originally developed as an MHD model of Earth's magnetosphere at UCLA in the early 1990s by J. Raeder.

[5] Simulation is based on physical laws, but approximations and numerical errors are included in it. Thus all simulations must be verified against real (observed) data. Moreover, simulations in which conditions can be set freely are indispensable tools in understanding the essential elements of the observed results. Because of these characteristics, linking simulation results with observations is very important.

[6] Real-time simulation data and observed data can be complementary. With in situ observation by a satellite, even if many satellites are used, it is difficult to get a global image of the magnetosphere. In contrast, global simulation can visualize the entire image and conditions of the magnetosphere.

[7] For a quantitative comparison of observed and simulated data to be valid, the two types of data must have the same physical parameters. We use geomagnetic indices for this comparison. The most typical indices deduced from the ground magnetic variations are *Dst* and auroral electrojet (*AE*) indices. The *Dst* index represents the axially symmetric component of magnetic variations at low latitudes on the ground and is generally used as a measure of the intensity of magnetic storms. However, MHD simulation cannot accurately reproduce the *Dst* index because lack of the particle dynamics in MHD cannot reproduce the ring current that contributes strongly to the *Dst* index.

[8] The *AE* index, on the other hand, is deduced from magnetic variation at auroral latitudes on the ground, which is produced mainly by an ionospheric current called the auroral electrojet. The auroral electrojet is generated by the electric field imposed on the polar ionosphere and by confined high electric conductivities, which are mainly associated with substorms. The *AE* indices are derived from data generated by up to 12 geomagnetic observatories near the aurora oval located at roughly longitudinally equidistant intervals [*Takahashi et al.*, 2004]. Using 1-min sampling magnetic data, the *H*- (horizontal) component magnetic variation at each station whose base line is subtracted is superposed on the same time axis. In this plot the upper and lower envelopes are defined as the *AU* and *AL* indices, respectively. These provide the measure of the maximum eastward and westward electrojet currents at any time [*Davis and Sugiura*, 1966; *Kamei et al.*, 1981]. The *AE* and *AO* indices are derived from the *AL* and *AU* indices as follows:

$$AE = AU - AL$$
$$AO = (AU + AL)/2$$

Comparing the result of simulation with these indices as well as satellite observation is very valuable because the indices are derived by superposing data from a wide area.

[9] We aim to determine whether our real-time MHD simulation is practical in actual space weather forecasting.

We introduce our real-time global MHD simulation model first. Then, we derive the *AE* index from the simulation and compare it with observed indices. Finally, we discuss the implications for the characteristics of the model that concern solar wind-magnetosphere-ionosphere (S-M-I) coupling.

2. Real-Time Global MHD Simulation Model

[10] *Tanaka* [1994, 1995] developed a three-dimensional (3-D) MHD simulation model to investigate S-M-I coupling. He based his simulation code on an unstructured grid system to cope with the differences between the scales for the magnetosphere and the ionosphere and on a finite volume total variation-diminishing (TVD) scheme to capture discontinuity. The inner boundary of the grid system is set at 3 Earth radii (R_E) from the center of the Earth. The ionospheric potential can be deduced from the electric potential at 3 R_E, which is mapped on the ionosphere along the dipole field lines. The ionospheric current, which can be decomposed into the Pedersen and Hall currents, is derived from Ohm's law in the ionosphere. The current closes in the ionosphere as follows:

$$\nabla \cdot \boldsymbol{\sigma} \nabla \Phi_I = G_m(\text{rot } \mathbf{B}_1 \cdot \mathbf{n}_b) = J_{||},$$
$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_{EUV} + k_1 \boldsymbol{\sigma}_{Diff}(P, \boldsymbol{\rho}) + k_2 \boldsymbol{\sigma}_J(J_{||}),$$
$$\Phi_m = \Phi_I - k_3 f_1(J_{||})|J_{||}|,$$

where $\boldsymbol{\sigma}$ is the ionospheric conductivity tensor, Φ_I is the ionospheric potential, Φ_m is the magnetospheric potential, $\mathbf{B}_1$ is the variable component of the magnetic field, $\mathbf{n}_b$ is an outward unit vector, $J_{||}$ is the electric current parallel to the magnetic field, and G_m is a geometrical factor associated with the mapping along the field lines from $r = 3 \ R_E$ to $r = 1 \ R_E$. The current in the ionosphere closes with the FAC. Values $\boldsymbol{\sigma}_{EUV}$, $\boldsymbol{\sigma}_{Diff}$, and $\boldsymbol{\sigma}_J$ are the conductivities due to the solar EUV, the diffuse auroral precipitation, and the FAC, respectively. The diffuse auroral precipitation $\boldsymbol{\sigma}_{Diff}$ is defined as a function of the pressure P and density ρ. Here f_1 is a function being 1 at the upward FAC and 0 at the downward FAC, and $k_1 - k_3$ are scaling constants [*Tanaka*, 2000]. In this simulation the geographical north and the geomagnetic north are defined on the same axis. The resistivity (η) is defined in the present MHD simulation as:

$$\eta = k_4 f_2(x)|J|^2/|B|^2$$

where $f_2(x)$ is a time-independent fixed function (x indicates pointing to the Sun from the Earth), k_4 is a scaling constant. Here $f_2(x)$ is small in the near-Earth region at $x > -20 \ R_E$, linearly increases down tail, and saturates at $x > -60 \ R_E$ [*Tanaka*, 2000].

[11] Using the simulation code developed by *Tanaka* [1994, 1995], *Den et al.* [2006] started real-time global MHD simulation using the supercomputer at the National Institute of Information and Communications Technology (NICT). They used the density, velocity, temperature, and *Z*- and *Y*- components (in the GSM coordinate system) of the IMF of real-time solar wind data, which were recorded by the Advanced Composition Explorer (ACE) spacecraft located near the L1 point [*Zwickl et al.*, 1998]. We use
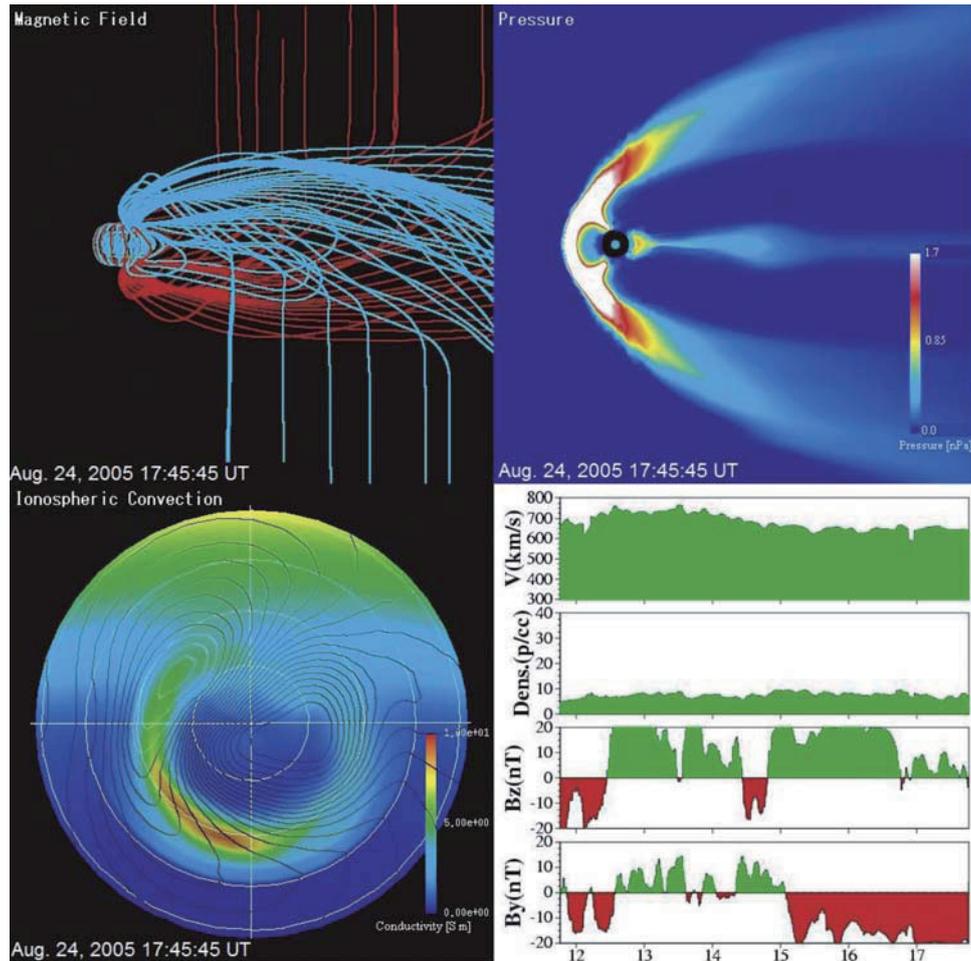
Figure 1. Types of graphics posted on Web site, showing (top left) magnetic field lines connected to the Earth. Blue and red lines are connected to northern and southern hemispheres, respectively. Some field lines with ends connected to both hemispheres are blue and some red. Also shown are (top right) plasma pressure on meridian plane, where the circle in the center represents Earth, (bottom left) contour lines represent electric potential and color code shows electrical conductivity of polar ionosphere viewed from North Pole, and (bottom right) input parameters observed by ACE spacecraft: solar wind velocity, density, and *Z*- and *Y*-components of IMF, respectively.

the modified spherical coordinates [*Tanaka*, 2000], and grid numbers 44, 56, and 60 for the r, theta, and phi directions, respectively.

[12] Processing the real-time solar wind data and calculating the MHD simulations takes only a few minutes, enabling us to obtain the 3-D distribution of the magnetic field and plasma parameters in the magnetosphere during the travel time of the solar wind. The travel time of the solar wind from the position of the ACE spacecraft to the magnetopause is approximately 1 h, though travel time basically depends on the solar wind velocity and the distance between the ACE spacecraft and the Earth. Thus real-time simulation is used to forecast the space weather of the magnetosphere for the 1-h period following the simulation.

[13] We generally simulate a period of 24 h a day for 365 d. The results are illustrated visually, and the resulting images can be posted to our Web site in almost real time (http://www2.nict.go.jp/y/y223/simulation/realtime/). The four kinds of graphics we post on the Web site are shown in Figure 1. The upper left panel shows magnetic field lines

around the Earth, and the upper right one shows the plasma pressure on the meridian plane. The circle in the center of the latter represents the Earth. Figure 1 shows that our simulation reproduces the bow shock and the tail plasma sheet. In the lower left panel, the contour lines represent the electric potential, and the color code shows the electric conductivity of the polar ionosphere viewed from the North Pole. The lower right panel shows input parameters observed by the ACE spacecraft: solar wind velocity, density, and the *Z*- and *Y*-components of the IMF. The images are archived in a movie format, and current and archived images can be accessed on the Web site.

[14] The example shown in Figure 1 is for a period after an interplanetary shock impacted the magnetopause. A pressure pulse propagates from the subsolar point to the magnetotail. When the *Z*-component of the IMF becomes negative, the current sheet thinning and plasmoid ejection can be reproduced intermittently. When the plasmoid is ejected, the conductivity of the polar ionosphere is enhanced on the nightside. When the *Y*-component of the

IMF is not oscillatory, a clear convection cell structure (a round cell and a crescent cell) can be observed [*Tanaka*, 2001]. The shape of the cells depends on the direction of the *Y*-component of the IMF.

3. Comparison of the Simulated and Observed *AE* Indices

[15] The geomagnetic variations on the ground are the consequence of the energy propagation furthest downstream
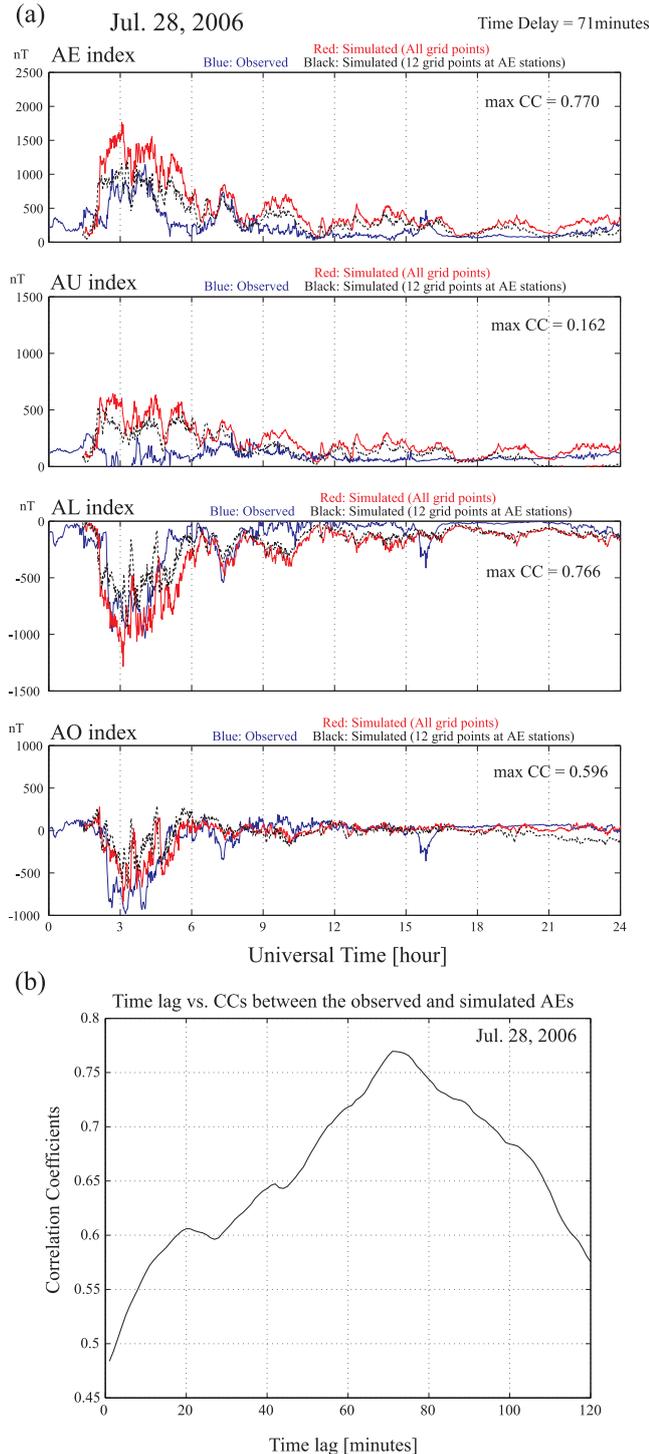
in the S-M-I coupling system and contain integrated information about the magnetosphere. Therefore the derivation of the magnetic field disturbance on the ground is suitable for verifying the availability of the actual space weather forecast using our simulation. The magnetic fields generated by the ionospheric current in the simulation are used to derive the *AE* index in real time. Although the actual *AE* index is deduced from the *H*-component variations in the magnetic field at the selected geomagnetic stations, which are located at the auroral latitudes, the simulated *AE* index here is deduced from the *H*-component magnetic variations at all grid points located at latitudes between 60° and 70°. There are seven grid points in the latitudinal direction and 40 in the longitudinal direction used to derive the *AE* index [*Den et al.*, 2006].

[16] We use digital data from the *Quick-Look (QL) AE* indices, which are calculated and distributed by the World Data Center (WDC) for Geomagnetism, Kyoto (http://swdcwww.kugi.kyoto-u.ac.jp/AEasy/index.html) as the observed *AE* indices. In general, the *QL AE* indices are based on semi-real-time magnetic data, which may include spike noises or data errors, so it is not always appropriate to use the *QL* indices for precise analysis. It is also noted that data from all 12 stations was not always available. When this happened, we used the magnetic variation data from the observatories that were operating to derive the *QL AE* indices. The real-time solar wind data measured by ACE are not processed by accurate data cleaning. Here, the *QL AE* index can be appropriately compared with the simulated *AE* index because one purpose of this study is to evaluate how well the *AE* index using real-time data predicts space weather.

[17] The simulated *AE* indices are calculated using real-time solar wind data obtained by the ACE spacecraft at a point more than 200 R_E upstream from the Earth, so the observed *AE* indices lag the simulated *AE* indices by a duration that depends on the speed and the structure of the solar wind. To obtain the time lag, we calculate the cross-correlation coefficients of the observed *AE* indices to the simulated *AE* indices whose time series are shifted from 0 to 120 min in 1-min steps. The time lag is defined as the shifted time with the maximum correlation coefficient between the observed *AE* index and time-shifted simulated *AE* index. We obtain daily correlation coefficients and lags using 1-min sampling data over 24 h. The *AU*, *AL*, and *AO* indices are calculated in the same way as the *AE* indices. In

Figure 2. (a) Comparison of simulated and observed indices for 28 July 2006. *AE*, *AU*, *AL*, and *AO* indices are shown in respective order from top to bottom. Blue and red lines represent observed and simulated indices, respectively, calculated from all the grid points located at latitudes between 60° and 70°. Black dotted lines are simulated indices calculated from 12 grid points closest to *AE* observatories. Time lag between simulated and observed *AE* indices is 71 min. MaxCCs shown in upper right of each panel are maximum correlation coefficients between observed and simulated (all stations) indices. Also shown is the (b) time lag and cross-correlation coefficients between simulated and observed *AE* indices for 28 July 2006.
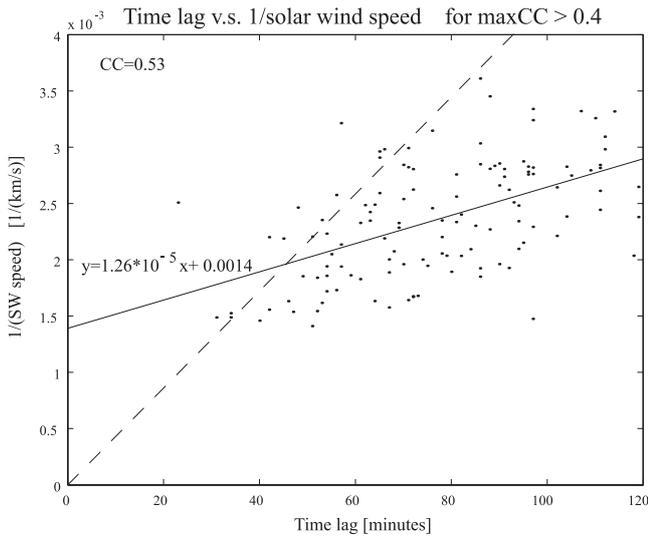
Figure 3. Comparison between time lag and $1/V$ (V: solar wind speed) for events in which maximum correlation coefficients (maxCCs) are more than 0.4. Solid line represents least-square fitting. Broken line represents expectation from propagation of solar wind from position of ACE (220 R_E) to Earth.

this procedure, if the solar wind speed changes drastically in 1 d, the results of the analysis for that day may contain significant errors. To clarify this effect, we calculated the daily mean solar wind speed ($Vmean$), the daily maximum speed (V max) and the minimum speed (V min) for 398 d from August 2005 to September 2006. The ratio $Rsw = (V$ max $- V$ min $)/Vmean$ for 398 d shows that 78% of Rsws are less than 0.3, that is, the errors caused by changes in the solar wind speed within a day are generally small.

[18] The simulated and observed indices for the interval between August 2005 and September 2006 were used in the calculations. Because of an interruption in computer operations, data for some simulated indices are missing. We also removed data that included obvious errors for either simulated or observed *AE* indices. As a result, to calculate the *AE*, *AU*, *AL*, and *AO* indices, we used data collected for 247, 238, 247, and 201 d, respectively.

[19] Figure 2a shows a comparison of simulated and observed indices for 28 July 2006. The *AE*, *AU*, *AL*, and *AO* indices are shown in respective order from the top to the bottom panel. The blue and red lines indicate observed and simulated indices, respectively, calculated from all grid points located at latitudes between 60° and 70°. The black dotted lines represent the simulated index calculated from the 12 grid points closest to the 12 *AE* observatories, which are overwritten in each panel. We used the geomagnetic coordinates for the 12 stations to derive the index.

[20] The figure shows that, as expected, the absolute values of the simulated *AL* and *AU* indices derived from the all grid-point data are generally larger than those from the 12-grid-points. Since there are two stations whose geomagnetic latitude is above 70°, the 12-grid-points index exceeds the all-grid-points index in some places. If we compare the *AL* and *AU* indices, the difference between all-grid-points indices and the 12-grid-points indices is

rather large in the *AU* index. High-latitude eveningside stations contribute to the *AU* index [*Allen and Kroehl*, 1975]. The UT variation in the *AU* index shows this tendency.

[21] However, the difference between the 12-grid-points indices and the all-grid-points indices is small compared with the difference from the observed indices. This indicates that the derivation in the *AE* index from all grid points between 60° and 70° does not cause serious discord for the statistical analysis. Hereafter, we use as the simulated indices those calculated from all the grid point data.

[22] We calculate the cross correlation coefficient between the simulated index and the observed one. The correlation coefficients of the *AE* index for the time lags are shown in Figure 2b. The correlation coefficients are distributed according to a quadratic function of the time delay. In this case, the maximum correlation coefficient (maxCC) between the simulated (red line) and observed *AE* indices (blue line) is 0.77 at a 71-min lag. This value is roughly accurate when the solar wind speed for the day, ~330 km/s, is taken into account.

[23] The observed *AE* index shows clear auroral activities from 0200 UT to 0500 UT with a maximum amplitude of ~1000 nT. The simulated *AE* index (red) also shows activities from 0200 UT to 0530 UT with a peak amplitude of ~1700 nT, which is larger than that of the observed *AE* index. In addition, a large fluctuation with an amplitude of ~700 nT appears in the simulated *AE* index, whereas no such fluctuation appears in the observed *AE* index. The amplitude of the fluctuation in the simulated *AE* index is generally larger than that in the observed *AE* index. In contrast, the simulated *AE* index for the *AE* activities from 0600 UT to 0800 UT with an amplitude of ~800 nT, corresponds well to the observed *AE* index.

[24] The *AE* index is derived by subtracting the *AL* from the *AU*, so the difference between the simulated and observed *AE* indices can be considered in terms of the difference between the *AU* and *AL* indices. We calculate the maxCC for the *AE* and for the *AU*, *AL* and *AO* indices. The maxCCs of the *AL* and *AO* indices are 0.766 and 0.596, respectively, while the maxCC of the *AU* index is 0.162, which is much less than those of the other indices. The envelopes of the simulated *AL* index correlate well with the observed *AL* index during 0200−0500 UT and 0600− 0800 UT, though the large fluctuation in the simulated *AL* index from 0200 UT to 0600 UT does not appear in the observed *AL* index. The amplitudes of the simulated *AL* activities during the above-indicated intervals are roughly comparable to that of the observed *AL* index. In contrast, the time variation of the simulated *AU* index is quite different from that of the observed *AU* index. The intensity of the simulated *AU* index is larger than that of the observed *AU* index almost all day long, especially during the above-indicated intervals. The difference between the amplitudes of the simulated and observed *AU* indices reaches 500 nT in some cases.

[25] This indicates that the westward electrojets are reproduced well by the present simulation, whereas the eastward electrojets are sometimes not reproduced well. In general, the intensity of the auroral electrojet is predominant in the westward direction during the interval of the substorm. Therefore during the substorm, the simulated *AE*

Occurrence Frequency of maximum cross-correlation coefficiency (maxCC)
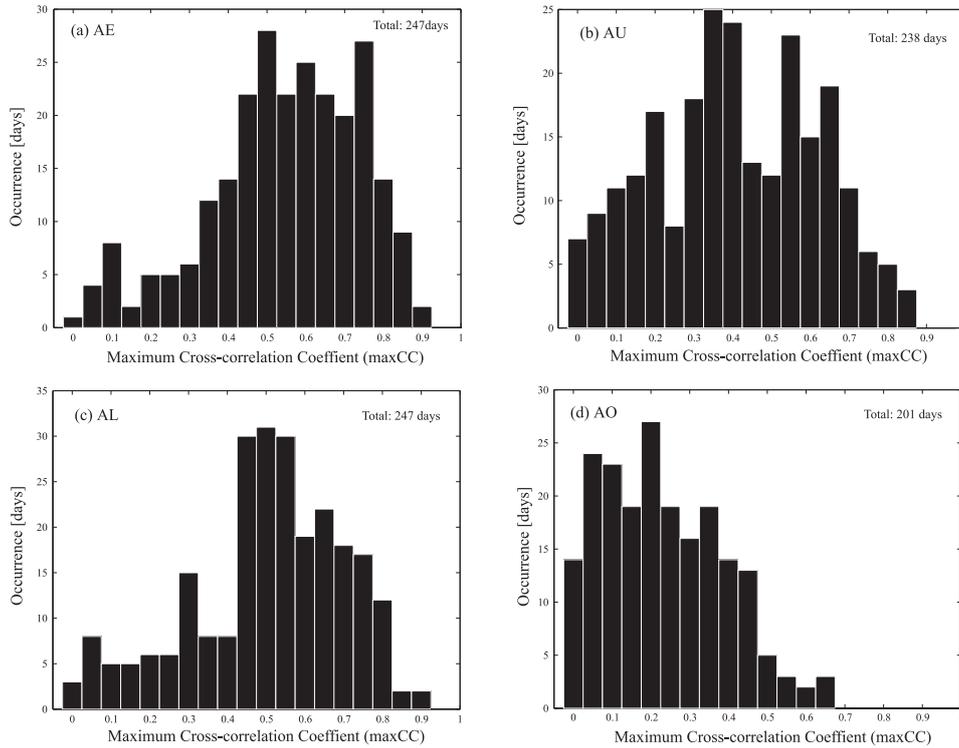
Figure 4. Occurrence frequency of maximum correlation coefficients of (a) *AE*, (b) *AU*, (c) *AL*, and (d) *AO* indices. Vertical and horizontal axes represent occurrence of days and maximum correlation coefficient (maxCC), respectively. Range of horizontal axis is between 0 and 1, and negative maxCCs (less than 1% of total) are not shown here.

index can be reproduced well. We used one event as an example of a typical substorm event and as one of the best-correlated events. Similar signatures are observed in other substorm events. To clarify what constitutes a factor limiting the reproducibility of the simulation, it is necessary to statistically investigate the correlations between the simulated and observed indices.

[26] The time lags between the observed and simulated *AE* indices are also important in evaluating the results of the simulations. The time lag mainly represents the travel time of the solar wind from the position of the ACE spacecraft to the Earth. Thus the time lag should decrease as solar wind speed increases. Figure 3 shows the relationship between the time lag and the solar wind speed observed by the ACE spacecraft. Although the data points scatter, $1/V$ (V: the solar wind speed) increases as the time lag increases, and the correlation coefficient between the two parameters is 0.53. The distribution of the data points can fit the linear function $y = 1.26 \times 10^{-5} x + 1.40 \times 10^{-3}$ (x: time lag [min], y: 1/(solar wind speed) [1/km/s]). This indicates that the time lag basically corresponds to the travel time of the solar wind speed. However, the y-intercept of 1.40×10^{-3} suggests that the time lag is not due to solar wind speed alone. The broken line in Figure 3 represents the expected relationship between $1/V$ and the time lag under the assumption that the ACE spacecraft is located 220 R_E upstream from the Earth. The figure shows that most of the events have a longer-than-expected time lag. This means that the time lag cannot be

explained by the propagation time of the solar wind from ACE to the Earth alone. The time constant of the magnetospheric response to the solar wind may significantly affect the time lag between the observed and simulated *AE* indices.

[27] Figure 4 shows the occurrence frequencies of the maxCCs for *AE*, *AU*, *AL*, and *AO* indices. In Figure 4a, most of the maxCCs for the *AE* index are distributed in the range between 0.4 and 0.8. Of the total of 247 d, 158 (64%) showed a maxCC of more than 0.5. This shows that simulation can accurately reproduce the *AE* activities for most days and is useful in predicting the *AE* index for the next hour. However, there were also 89 d (36%) for which the maxCCs were less than 0.5, which is not accurate enough to be used in prediction.

[28] The occurrence frequencies of the maxCCs for the *AU*, *AL*, and *AO* indices are shown in Figures 4b, 4c, and 4d, respectively. The distribution of the occurrence frequency of the *AL* indices is similar to that of the *AE* indices; 54% of the days analyzed had maxCCs of over 0.5. In contrast, the occurrence frequencies of the maxCCs for the *AU* and *AO* indices have rather different distributions. As can be seen in Figure 4b, the occurrence of the maxCCs for the *AU* indices was broadly distributed, with a formless peak around 0.4, and only 37% of the days had maxCCs of more than 0.5. Of the four types of indices, the *AO* indices showed the worst correlation, with 96% of the days having maxCCs of less than 0.5.
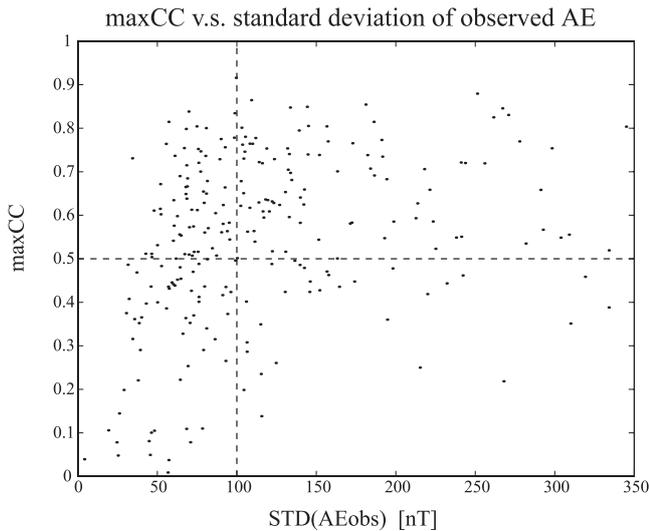
Figure 5. Relationship between maximum cross-correlation coefficients (maxCC) and standard deviation of observed *AE* (STD(*AEobs*)). Dashed horizontal line represents maxCC of 0.5, and dashed vertical line represents STD(*AEobs*) of 100 nT.

[29] A statistical analysis of the cross-correlations for the 247 d shown in Figures 3 and 4 indicates that the simulated *AE* index corresponds fairly well to the observed *AE* index, but that, for 36% of the days, the simulated *AE* indices show very little correlation with the observed *AE* indices. The factors in these lower maxCCs should be researched further.

[30] It is well known that the *AE* index is correlated with solar wind parameters and the IMF. Variations in the *AE* index and substorms occur as a consequence of the processes in the solar wind-magnetosphere coupling system. However, the triggering mechanism of substorms is still unknown. *Lyons et al.* [1997, 2005] found that a large majority of the substorms with well-defined onset times are triggered by northward turning of the IMF. In addition, the external cause of the triggering, that is, change in the dynamic pressure under a strong southward IMF, also causes auroral brightening of the substorm [*Shi et al.*, 2005, 2006]. In contrast, *Kitamura et al.* [2003, 2005] reported that quasi-periodic substorms occur during steady southward IMFs with no prominent northward turning. This means that the electromagnetic energy that flows continuously into the magnetosphere during the southward IMF is released as quasi-periodic substorms not by an external trigger but by an internal process of the magnetosphere. Even if the internal process is related to the substorm trigger, the *AE* index still correlates fairly well with the IMF and solar wind parameters.

[31] To compare the solar wind conditions and maxCCs of the *AE* indices, we used the *Z*-component of the IMF, solar wind speed, and proton number density, all of which are measured by the ACE spacecraft, to calculate the cross-correlation coefficients with the maxCC obtained as described above. The results showed no direct influence of the solar wind parameters on the maxCCs. We also calculated the cross-correlation coefficients between the maxCCs and the standard deviations of these solar wind parameters to

investigate the influence of the fluctuations of these solar wind parameters on the maxCCs. All the correlation coefficients calculated this way were less than 0.1 (data not shown here), which indicates that solar wind fluctuations have no direct influence on the maxCCs.

[32] To predict activities of the *AE* index, we should pay more attention to cases of intense *AE* activity rather than to mild *AE* activity. For this reason, we calculated the standard deviations of the observed *AE* indices (STD(*AEobs*)) for 24 h and compared them with the maxCCs. Figure 5 shows the results for this comparison. The distribution of the plots seems to be scattered, and no obvious linear relationship can be observed in Figure 5. There were 120 d (out of 247) for which the STD(*AEobs*) was over 100 nT. The maxCCs for 74% of these 120 d were over 0.5. This means that the correlations between the simulated and observed *AE* indices are better for days with observed *AE* indices with large-amplitude standard deviations.

[33] In contrast, most of the data points for the small maxCCs are distributed in the STD(*AEobs*) range between 0 and 100 nT. As indicated in the discussion of Figure 2a, the maxCCs of the *AE* indices for 89 of the 247 d were less than 0.5. For 57 of the 89 d (64%), the STD(*AEobs*) were less than 100 nT. This means that the simulated *AE* indices do not correlate well with the observed *AE* indices when the STDs are small. We discuss the reasons for this signature below.

4. Discussion and Summary

[34] In previous studies, the relationships between the variation in the *AE* index and the solar wind are analyzed using the linear prediction filter [*Bargatze et al.*, 1985]. This study implied that the solar wind-magnetosphere coupling should be described by nonlinear dynamics. *Klimas et al.* [1996] reviewed autonomous analysis and nonlinear input-output data-based methods for solar wind-magnetosphere coupling problems. They showed that unloading process should be included for realistic replications of the *AL* index. Using the Faraday Loop model, *Baker et al.* [1997] also showed that a realistic model including the loading and unloading process is necessary to describe the general substorm dynamics, as reflected in the *AL* index. These studies used mathematical models that describe the solar wind-magnetosphere coupling and demonstrated specific events. In the present study, we use the MHD simulation including the S-M-I coupling and statistically analyze a large number of events.

[35] We have shown that the *AE* indices derived from the MHD simulation using real-time solar wind data as upstream boundary conditions are accurate enough to be used in predicting the *AE* indices for the following hour. The comparison of the simulated and observed *AE* indices generally showed a good correlation. However, the correlation of these two indices is not good for 36% of the days analyzed. What is the reason for such inconsistency?

[36] We first consider this question from the simulation model side. One possible reason for the inconsistency is that the geomagnetic north and geographical north of the coordinate system of the present simulation are defined on the same axis as the spherical coordinate system [*Den et al.*, 2006]. We think this is the most significant problem in our

simulation model. The agreement of the geographic and geomagnetic coordinates causes some deviation in the current flow in the polar ionosphere and leads to errors in calculating the *AE* index. The simulation model should be modified to correct this.

[37] Here, we consider the solar wind structure to discuss the effect on the statistical results shown in Figure 5. The relationship between the time lag and the solar wind speed shown in Figure 3 indicates that the time lag increases as the solar wind speed decreases, although the correlation coefficient of 0.53 might not be valid. We used real-time solar wind data at a point L1 upstream of the Earth. Although the solar wind properties monitored at the L1 point generally represent the solar wind conditions near the magnetopause [e.g., *Paularena et al.*, 1998; *Richardson et al.*, 1998], the spatial structure of the solar wind may affect the time lag. Solar wind structure was recently investigated using multi-spacecraft observations in near-Earth interplanetary space [*Eastwood et al.*, 2002; *Lepping et al.*, 2003]. These studies compared the solar wind data measured by the ACE spacecraft with those measured by IMP-8 and Cluster satellites and showed a difference significant enough to account for the spatial structure of the heliospheric current sheet and MHD discontinuities. Such an inhomogeneity of the solar wind structure becomes more significant when there are large-amplitude variations in the solar wind speed for 24 h. Thus modifying the solar wind property during the propagation from the L1 point to the upstream boundary of the simulation may reduce the correlation between the time lag and the solar wind speed. This may also affect the signature of the maxCC shown in Figure 5. Although the solar wind properties observed at the L1 point are statistically correlated with those reported near the Earth by the studies mentioned above, *Paularena et al.* [1998] also indicated that the average solar wind flux correlation between the two spatially separated spacecraft decreases as the standard deviation of the solar wind flux decreases. This suggests that the IMF and the plasma parameters may change significantly during the propagation from ACE to the upstream boundary of the simulation, reducing the maxCCs for days when the observed *AE* indices have small standard deviations.

[38] Next, we consider the properties of the *AE*, *AU*, *AL*, and *AO* indices. In Figure 4, the simulated and observed indices show good correlations for the *AE* and *AL* indices. The *AL* index mainly represents substorm activities because the westward electrojet associated with substorm expansion is generally observed as a sudden decrease in the *AL* index. The good correlation between simulated and observed *AL* indices indicates that the present simulation model reproduces the substorm expansion and the resulting westward electrojet currents well.

[39] The occurrence frequencies of maxCCs of the *AU* and *AO* indices show quite different distributions from those of the *AE* and *AL* indices. Days with *AU* and *AO* indices with maxCCs over 0.5 represent 37% and 4% of the total days, respectively. What does this variation in *AU* indices mean? Unlike the westward electrojet, the eastward ionospheric currents do not flow in the confined region during the substorm expansion [*Kamide et al.*, 1994, 1996]. Also, the intensity of the *AU* index and the ionospheric potential in the dusk cell are generally no greater during the substorm

than under nonsubstorm conditions [*Weimer*, 1999]. It has also been reported [*Weimer*, 1999] that the *AU* index is a statistical linear function of the polar cap potential. However, the mechanism and causes of the *AU* variation seem not to have been observationally well proven. The lower correlation between the simulated and observed *AU* indices may indicate that the processes that cause the *AU* variations are not correctly included in the simulation. This should be more carefully investigated in future work.

[40] Although both *AE* and *AO* indices are deduced from the combination of the *AU* and *AL* indices, the correlation of the simulated and observed *AO* indices is statistically much lower than that of the *AE* index. This can be explained by the overestimation of the variation in the simulated indices. When the simulated *AU* index is larger (smaller) than the observed *AU* index, the simulated *AL* is also larger (smaller) than the observed *AL* index. (The *AL* is usually negative. A larger *AL* means a smaller absolute *AL* value.) In this case, the overestimated variations of the simulated *AL* and *AU* indices cancelled each other out by the subtraction used to derive the *AE* index and enhanced by the addition used to derive the *AO* index. As a result, the correlations between the simulated and observed *AE* (*AO*) indices are better (worse).

[41] Finally, we consider the reasons for the larger fluctuations in the simulated *AE* index. Figure 5 indicates that the simulated and observed *AE* indices show worse correlations when the standard deviations of the observed *AE* indices are less than 100 nT. These fluctuations may be the result of the difference of the magnetospheric response to the solar wind between the actual magnetosphere and the simulation. The other factor in the fluctuations may arise from the time resolution of the present simulation. The time step depends on the CFL (Courant-Friedrichs-Lewy) conditions, but it is usually much shorter than 1 s. We used the 1 min value for the simulated index, but it may be that the shorter time scale fluctuations appear in the simulated *AE* index rather than in the observed index.

[42] *Tanaka* [2000] showed that the MHD simulations that assumed that the magnetotail becomes more diffuse as it goes further down tail can reproduce substorms. In his simulation, the southward IMF caused the typical signatures of the substorm growth phase, which are plasma sheet thinning, increase in the size of the auroral oval, and increase in the FAC in the polar ionosphere. This means the substorm onset may be directly caused by the dipolarization of the magnetic fields, which is hastened by the northward turning of the IMF. The onset of the substorm occurs as an abrupt change in the pressure distribution in the near-Earth plasma sheet and moving of the convection flow into the inner magnetosphere. The signatures of the substorm expansion phase (dipolarization, plasma injection into the inner magnetosphere, and increase in the intensities of the FACs connected to the nightside ionosphere) were well reproduced in the simulations.

[43] These simulation results were obtained under ideal solar wind conditions to enlarge the magnetospheric convection associated with the substorm. In contrast, the *AE* indices derived from real-time MHD simulation using real-time solar wind parameters reflected actual solar wind disturbances, which consist of fluctuations and discontinuities with various ranges of temporal and spatial variations.

Thus the inconsistency in the statistics between the simulated and observed *AE* indices shown in Figures 4 and 5 may be caused by the magnetospheric responses to the realistic solar wind disturbances, which are not included in the simulation. The present result indicates that the simulation responds sensitively to small fluctuations in the IMF and dynamic pressure of the solar wind. Sometimes, small substorm-like enhancements of the *AE* indices appear in the simulation even if the observed *AE* shows no substorm signatures. *Uritsky et al.* [2001] indicated from a self-organized criticality model that the scaling features in the *AE* index are independent of the solar wind input for a time scale shorter than 3.5 h. Our results also show that short-period fluctuations in the solar wind do not appear in the observed index, but that they do appear in the simulation. Thus the real S-M-I coupling system seems to have a function that absorbs some of the short-period perturbations in the electromagnetic energy flowing from the solar wind into the magnetosphere by some mechanism that is not included in our MHD simulation and that dampens the resulting fluctuations in the *AE* indices compared with the simulation.

[44] The particle effect and phenomena smaller than the mesh size were not included in the simulation model. Moreover, bursty events whose timescale is shorter than the time step of the simulation may not be reproduced well in the present simulation. These phenomena may affect the results. If we include such effects, the simulation may reproduce the *AE* index more accurately. These problems should be resolved in the future.

[45] We have shown that simulated *AE* indices are fairly well correlated with observed *AE* indices. We have also examined when and how they did not agree. The disagreement between the simulated and observed *AE* indices has given us a significant clue about the processes of substorms and S-M-I coupling. A more detailed comparison between the simulation using the real solar wind data and the various kinds of observations in the magnetosphere should be done in the future.

References

Allen, J. H., and H. W. Kroehl (1975), Spatial and temporal distributions of magnetic effects of auroral electrojets as derived from *AE* indices, *J. Geophys. Res.*, *80*, 3667.

Baker, D. N., A. J. Klimas, D. Vassiliadis, T. I. Pulkkinen, and R. L. McPherron (1997), Reexamination of driven and unloading aspects of magnetospheric substorms, *J. Geophys. Res.*, *102*, 7169.

Bargatze, L. F., D. N. Baker, R. L. McPherron, and E. W. Hones (1985), Magnetospheric impulse response for many levels of geomagnetic activity, *J. Geophys. Res.*, *90*, 6387.

Davis, T. N., and M. Sugiura (1966), Auroral electrojet activity index *AE* and its universal time variations, *J. Geophys. Res.*, *71*, 785.

Den, M., et al. (2006), Real-time Earth magnetosphere simulator with three-dimensional magnetohydorodynamic code, *Space Weather*, *4*, S06004, doi:10.1029/2004SW000100.

Eastwood, J. P., A. Balogh, and M. W. Dunlop (2002), Cluster observations of the heliospheric current sheet and an associated magnetic flux rope and comparisons with ACE, *J. Geophys. Res.*, *107*(A11), 1365, doi:10.1029/2001JA009158.

Elsen, R. K., and R. M. Winglee (1997), The average shape of the Magnetopause: A comparison of three-dimensional global MHD and empirical models, *J. Geophys. Res.*, *102*, 4799.

Fedder, J. A., and J. G. Lyon (1995), The Earth's magnetosphere is 165 R_E long: Self-consistent currents, convection, magnetospheric structure, and processes for northward interplanetary magnetic field, *J. Geophys. Res.*, *100*, 3623.

Janhunen, P. H., and E. J. Koskinen (1997), The closure of Region-1 field-aligned current in MHD simulation, *Geophys. Res. Lett.*, *24*, 1419.

Kamei, T., M. Sugiura, and T. Araki (1981), Auroral electrojet (*AE*) indices for January-June 1978, in *Data Book 3*, World Data Cent. for Geomagn., Kyoto Univ., Japan.

Kamide, Y., et al. (1994), Ground-based studies of ionospheric convection associated with substorm expansion, *J. Geophys. Res.*, *99*, 19,451.

Kamide, Y., W. Sun, and S.-I. Akasofu (1996), The average ionospheric electrodynamics for the different substorm phases, *J. Geophys. Res.*, *101*, 99.

Kitamura, K., H. Kawano, S. Ohtani, A. Yoshikawa, K. Yumoto, and the Circum-pan Pacific Magnetometer Network Group (2003), Quasi-periodic substorms during recovery phase of magnetic storm for space weather study, paper presented at Int. Symp. on Inf. Sci. and Electr. Eng. 2003, Kyushu Univ., Fukuoka, Japan.

Kitamura, K., H. Kawano, S. Ohtani, A. Yoshikawa, and K. Yumoto (2005), Local-time Distribution of low and middle latitude ground magnetic disturbances at sawtooth injections of April 18–19, 2002, *J. Geophys. Res.*, *110*, A07208, doi:10.1029/2004JA010734.

Klimas, A. J., D. Vassiliadis, D. N. Baker, and D. A. Roberts (1996), The organized nonlinear dynamics of the magnetosphere, *J. Geophys. Res.*, *101*, 13,089.

Lepping, R. P., C.-C. Wu, and K. McClernan (2003), Two-dimensional curvature of large angle interplanetary MHD discontinuity surfaces: IMP-8 and WIND observations, *J. Geophys. Res.*, *108*(A7), 1279, doi:10.1029/2002JA009640.

Lyons, L. R., G. T. Blanchard, J. C. Samson, R. P. Lepping, T. Yamamoto, and T. Moretto (1997), Coordinated observations demonstrating external substorm triggering, *J. Geophys. Res.*, *102*, 27,039.

Lyons, L. R., D.-Y. Lee, C.-P. Wang, and S. B. Mende (2005), Global auroral responses to abrupt solar wind changes: Dynamic pressure, substorm, and null events, *J. Geophys. Res.*, *110*, A08208, doi:10.1029/2005JA011089.

Narock, A., L. Rastaetter, M. Kuznetsova, and M. Hesse (2004), Realtime magnetosphere simulation at CCMC, *Eos Trans. AGU*, *85*(47), Fall Meet. Suppl., Abstract SA43B-06.

Paularena, K. I., G. N. Zastenker, A. J. Lazarus, and P. A. Dalin (1998), Solar wind plasma correlations between IMP 8, INTERBALL-1 and WIND, *J. Geophys. Res.*, *103*, 14,601.

Raeder, J. R., J. Walker, and M. Ashour-Abdalla (1995), The structure of the distant geomagnetic tail during long periods of northwsard IMF, *Geophys. Res. Lett.*, *22*, 349.

Richardson, J. D., F. D. Dashevskiy, and K. I. Paularena (1998), Solar wind plasma correlations between L1 and Earth, *J. Geophys. Res.*, *103*, 14,619.

Shi, Y., E. Zesta, L. R. Lyons, A. Boudouridis, K. Yumoto, and K. Kitamura (2005), Effect of solar wind pressure enhancements on storm time ring current asymmetry, *J. Geophys. Res.*, *110*, A10205, doi:10.1029/2005JA011019.

Shi, Y., E. Zesta, L. R. Lyons, K. Yumoto, and K. Kitamura (2006), Statistical study of effect of solar wind dynamic pressure enhancements on dawn-to-dusk ring current asymmetry, *J. Geophys. Res.*, *111*, A10216, doi:10.1029/2005JA011532.

Song, P., D. L. De Zeeuw, T. I. Gombosi, C. P. T. Groth, and K. G. Powell (1999), numerical study of solar wind-magnetosphere interaction for northward interplanetary magnetic field, *J. Geophys. Res.*, *104*, 28,361.

Spicer, D. S., S. T. Zalesak, R. Lohner, and S. Curtis (1996), Simulation of the magnetosphere with a new three dimensional MHD code and adaptive mesh refinement: Preliminary results, *Adv. Space Res.*, *18*, 253.

Takahashi, K., C. Meng, T. Kamei, T. Kikuchi, and M. Kunitake (2004), Near-real-time auroral electrojet index: An international collaboration makes rapid delivery of auroral electrojet index, *Space Weather*, *2*, S11003, doi:10.1029/2004SW000116.

Tanaka, T. (1994), Finite volume TVD scheme on an unstructured grid system for three-dimensional MHD simulation of inhomogeneous sys-

tems including strong background potential fields, *J. Comput. Phys.*, *111*, 381.

Tanaka, T. (1995), Generation mechanisms for magnetosphere-ionosphere current systems deduced from a three-dimensional MHD simulation of the solar wind-magnetosphere-ionosphere coupling processes, *J. Geophys. Res.*, *100*, 12,057.

Tanaka, T. (1999), Configuration of the magnetosphere-ionosphere convection system under northward IMF conditions with nonzero IMF B$_y$, *J. Geophys. Res.*, *104*, 14,683.

Tanaka, T. (2000), The state transition model o the substorm onset, *J. Geophys. Res.*, *105*, 21,081.

Tanaka, T. (2001), Interplanetary magnetic field By and auroral conductance effects on high-latitude ionospheric convection patterns, *J. Geophys. Res.*, *106*, 24,505.

Tóth, G., et al. (2005), Space Weather Modeling Framework: A new tool for the space science community, *J. Geophys. Res.*, *110*, A12226, doi:10.1029/2005JA011126.

Uritsky, V. M., A. J. Klimas, and D. Vassiliadis (2001), Comparative study of dynamical critical scaling in the auroral electrojet index versus solar wind fluctuations, *Geophys. Res. Lett.*, *28*, 3809.

Walker, R. J., T. Ogino, J. Raeder, and M. Ashour-Abdalla (1993), A global magnetohydrodynamic simulation of the magnetosphere when the inter-planetary magnetic field is southward: The onset of magnetotail reconnection, *J. Geophys. Res.*, *98*, 17,235.

Weimer, D. R. (1999), Substorm influence on the ionospheric electric potentials and currents, *J. Geophys. Res.*, *104*, 185.

White, W. W., G. L. Siscoe, G. M. Erickson, Z. Kaymaz, N. C. Maynard, K. D. Siebert, B. U. O. Sonnerup, and D. R. Weimer (1998), The magnetospheric sash and the cross-tail S, *Geophys. Res. Lett.*, *25*, 1605.

Zwickl, R. D., et al. (1998), The NOAA Real-Time Solar-Wind (RTSW) system using ACE data, *Space Sci. Rev.*, *86*, 633.

————————————

S. Fujita, Meteorological College, Asahicho, Kashiwa, Chiba, 277-0852, Japan.

K. Kitamura, Tokuyama College of Technology, Gakuendai, Shunan, Yamaguchi, 745-8585, Japan.

M. Kunitake, H. Shimazu, H. Shinagawa, and S. Watari, National Institute of Information and Communications Technology, Nukuikitamachi, Koganei, Tokyo, 184-8795, Japan.

T. Tanaka, Department of Earth and Planetary Sciences, Kyushu University, Hakozaki, Higashiku, Fukuoka, 812-8581, Japan.

Numerical studies on three-dimensional earthward fast plasma flows in the near-Earth plasma sheet by the spontaneous fast reconnection model

K. Kondoh[1] and M. Ugai[1]

[1] The spontaneous fast reconnection model is applied to the earthward fast flow events observed in the near-Earth plasma sheet. Here, the earthward fast flow events include both of bursty bulk flow events and flow burst events. In order to apply it directly to actual observations, virtual probes are located in the plasma sheet region in the three-dimensional simulation domain so that we can directly observe the temporal variations of plasma quantities in accordance with the growth and proceeding of the fast reconnection. In this model, magnetic reconnection drastically evolves and Alfvénic fast plasma jet flows in the very restricted narrow channel, and a large-scale plasmoid is formed ahead of the fast plasma jet. The results of virtual observation of these evolutions are found to be in good agreement with actual satellite observations. At the same time, in the lobe region, travelling compression regions (TCRs) are observed in connection with the fast flow events. The temporal profiles of magnetic fields detected by the virtual probes are also in good agreement with actual satellite observations. It is concluded that the earthward fast flow events and earthward TCR events result from the fast reconnection mechanism.

1. Introduction

[2] It is widely accepted that magnetic reconnection plays crucial role on large dissipative events, such as geomagnetic substorms and solar flares [*Shibata*, 1999, 2004]. In space plasmas of extremely large magnetic Reynolds number, magnetic reconnection can convert magnetic energy to plasma energies most effectively by slow shocks. A possible fast reconnection configuration, involving standing slow shocks, was first proposed by *Petschek* [1964]. In this respect we have proposed the spontaneous fast reconnection model. This model predicts that the fast reconnection mechanism spontaneously develops from inside the system by the self-consistent interaction between plasma microscopic processes and macroscopic reconnection flows [*Ugai*, 1984, 1986]. The previous MHD simulations have demonstrated that the spontaneous fast reconnection model works quite effectively even in three-dimensional situations [*Ugai and Kondoh*, 2001; *Ugai et al.*, 2004, 2005]. Once the fast reconnection builds up, an Alfvénic fast flow is caused, and if the fast flow is obstructed, the reconnected field lines are piled up, and a magnetic loop is formed. A magnetic loop in our simulations is allowed to be formed because of a wall boundary assumed at one edge of a current sheet system [*Ugai et al.*, 2003].

[3] Here, we are interested in the flow enhancements ($V_x \gtrsim 300$ km/s) in the near-Earth ($X \gtrsim -30\ R_E$) plasma sheet which are called bursty bulk flows (BBFs) or flow bursts (FBs) [*Angelopoulos et al.*, 1992]. The existence of these earthward fast flow events and sharp compression of the magnetic field in the near-Earth plasma sheet during active magnetospheric conditions have been pointed out by a lot of individual case studies [*Sergeev et al.*, 2000, 2001; *Angelopoulos et al.*, 1992]. In particular, *Sergeev et al.* [2000] showed a narrow transient flow jet using five spacecraft measurements. They showed that the $\sim 3\ R_e$ wide fast plasma jet propagates from $\gtrsim 40\ R_e$ in the midtail, and is able to reach the inner magnetosphere at 6.6 R_e in about 10 min.

[4] Statistical studies (superposed epoch analyses) have also been carried out in order to understand the general characteristics of BBF [*Ohtani et al.*, 2004; *Angelopoulos et al.*, 1992; *Schodel et al.*, 2001]. To construct the superposed epoch, it is necessary to define the reference time. The physical quantity of each event is superposed about the reference time. Previous studies used some reference times based on the flow velocity [*Angelopoulos et al.*, 1992; *Ohtani et al.*, 2004] or the magnetic field [*Ohtani et al.*, 2004] or the electric field [*Schodel et al.*, 2001]. By this averaging method, any fluctuation in the parameter is preserved in the average, whereas fluctuations shifting in time are averaged out. Before constructing a superposed epoch, consideration should be given to the selection of event. Most of studies define the earthward BBF event as the interval during which (1) the ion plasma beta $\beta_i > 0.5$,

---

[1]Electrical and Electronic Engineering and Computer Science, Graduate School of Science and Engineering, Ehime University, Matsuyama, Japan.

and (2) the ion bulk velocity is continuously greater than 100 km/s and the maximum value is greater than 300 km/s. In each analysis, the other conditions are added in order to much strictly exclude suspicious events. They used the data set, which are satisfied the above conditions, observed in $X > -31$ Re from the midtail region to the near-Earth region.

[5] These previous studies have shown that earthward BBFs are generally accompanied by dipolarization of the plasma sheet, and BBFs are directed either earthward or tailward. BBFs are therefore often considered signatures of bursty and localized reconnection in the plasma sheet [*Ohtani et al.*, 2004; *Schodel et al.*, 2001]. However, there is still no agreement about their origin [*Angelopoulos et al.*, 1992; *Lui*, 1996].

[6] The main theme of the present paper is to investigate the relationship between fast magnetic reconnection and earthward fast flow using MHD simulations on the basis of the spontaneous fast reconnection model. At first, we will explain the simulation model in section 2 and show the three-dimensional evolution of magnetic reconnection and magnetic loop in section 3.1. Then, we will introduce an example of the results on superposed epoch analysis for the earthward fast flow events in section 3.2 and compare our simulation results with it using virtual satellite observation in simulation domain in section 3.3. It is well known that these earthward fast flow events are accompanied with earthward moving traveling compression regions (TCRs) [*Slavin et al.*, 2003]. Therefore we will introduce an example of the magnetic field measurements in earthward moving TCRs and compare our simulation results with it using virtual satellite measurements as well as the case of earthward fast flow events and show the connection between the earthward fast flow events and the earthward moving TCRs using the simultaneous virtual satellite observations in section 3.4. Finally, section 4 is a summary and discussion.

2. MHD Simulations

[7] In the present study, we are interested in the underlying physical mechanism of earthward fast flow evolution, so that the simulation model will be idealized and simplified in view of precise numerical computations. Also, the earthward fast flow is closely related to the formation of magnetic loop (or magnetic dipolarization), so that the present simulation model is similar to the one of the 3-D magnetic loop dynamics [*Ugai et al.*, 2003].

2.1. Simulation Modeling

[8] Simulation model in this study is the same as that in the previous simulations on the three-dimensional magnetic loop [*Ugai et al.*, 2003] except for the initial density condition. It is assumed in this study that the initial temperature is constant everywhere, while the initial density was constant in the previous study. Our previous magnetic loop simulations using the wall boundary showed that the fast magnetic reconnection produces fast reconnection jet, and it flows in plasma sheet and is suddenly braked at the boundary between the dipolar and tail-like magnetic fields because of the counterward pressure force, and the reconnection jet is decelerated [*Ugai et al.*, 2003]. This situation may be consistent with the earthward flow, and these results

are in good agreement with the situation inferred from the observation results of earthward flow [*Shiokawa et al.*, 1997]. This wall boundary is, of course, not the real one but is simply assumed in the present paper.

2.1.1. Basic Equations

[9] Our interest is directed to the evolution of fast flows, and the MHD approximations are valid for these macroscopic phenomena. The compressible MHD equations are

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{u}$$

$$\rho \frac{D\mathbf{u}}{Dt} = -\nabla P + \mathbf{J} \times \mathbf{B}$$

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{u} \times \mathbf{B}) = -\nabla \times (\eta \mathbf{J}) \qquad (1)$$

$$\rho \frac{De}{Dt} = -P \nabla \cdot \mathbf{u} + \eta \mathbf{J}^2$$

$$\mathbf{J} = \nabla \times \frac{\mathbf{B}}{\mu_0}$$

$$\nabla \cdot \mathbf{B} = 0$$

where, $D/Dt \equiv \partial/\partial t + \mathbf{u} \cdot \nabla$; the gas law, $P = (\gamma - 1)\rho e$ is assumed (e is the internal energy per unit mass, and γ is the specific heat ratio with $\gamma = 5/3$ assumed here (an adiabatic case)), as is Ohm's law, $\mathbf{E} + \mathbf{u} \times \mathbf{B} = \eta \mathbf{J}$ (η may be an effective resistivity). The basic equations (1) are transformed to a conservation-law form, and the modified Lax-Wendroff scheme is used for the numerical computation [*Ugai*, 1987].

2.1.2. Initial-Boundary-Value Problem

[10] As an initial configuration, the one-dimensional antiparallel magnetic field $\mathbf{B} = [B_x(y), 0, 0]$ is assumed as: $B_x(y) = \sin(\pi y/2)$ for $0 < y < 1$; $B_x(y) = 1$ for $1 < y < Y_1$; $B_x(y) = \cos[(y - Y_1)\pi/1.2]$ for $Y_1 < y < Y_m(= Y_1 + 0.6)$; $B_x(y) = 0$ for $Y_m < y$; also, $B_x(y) = -B_x(-y)$ for $y < 0$. The plasma pressure $P(y)$ initially satisfies the pressure-balance condition,

$$P + B_x^2 = 1 + \beta_0 \qquad (2)$$

where β_0 is the ratio of plasma pressure to the magnetic pressure in the ambient magnetic field region $1 < y < Y_1$, so that $P(y = 0) = 1 + \beta_0$ initially (in the present study, $\beta_0 = 0.15$ is taken). Initially, fluid velocity $\mathbf{u} = (0, 0, 0)$, and, constant temperature $T = P/\rho = 1 + \beta_0$ is assumed, so that the plasma density ρ initially satisfies

$$\rho(y) = P(y)/(1 + \beta_0), \qquad (3)$$

The normalization of quantities, based on the initial quantities, is self-evident; distances are normalized by the half-width of the current sheet d_0, $\mathbf{B}$ by the field strength in the magnetic field region B_{x0}, P by $B_{x0}^2/(2\mu_0)$, and ρ by $\rho_i = \rho(y = 0)$; also, $\mathbf{u}$ by $V_{Ax0}(= B_{x0}/\sqrt{\mu_0 \rho_i})$, time t by d_0/V_{Ax0}, current density $\mathbf{J}$ by $J_0 = B_{x0}/(\mu_0 d_0)$, and so forth. Note that the Alfvén velocity in the ambient magnetic field region ($1 < y < Y_1$) at initial state is given by $V_{Ae} = V_{Ax0}/\sqrt{\rho_e}$ (ρ_e is the density there).

[11] Here, the conventional symmetry boundary conditions are assumed on the (x, y), (y, z), and (z, x) planes. Initially, $B_z = 0$ because of this symmetry boundary on the (x, y) plane. Hence the computational region can be restrict-
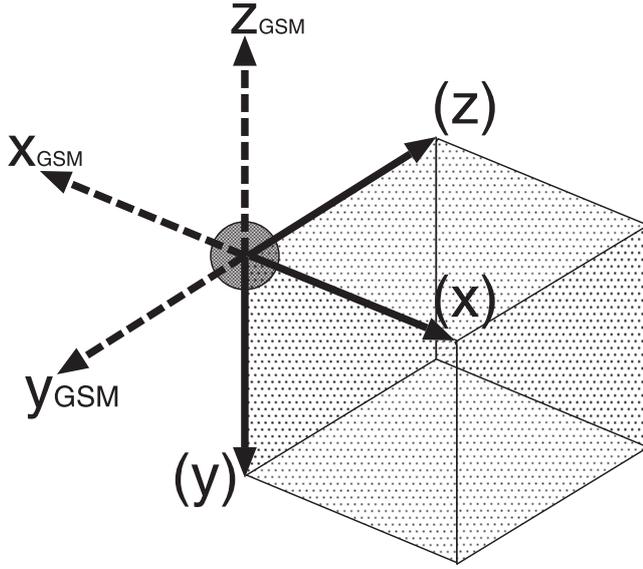
Figure 1. The relationship between coordinate systems, usual GSM one and that used in our three-dimensional simulations.

ed to the first quadrant only and taken to be a rectangular box, $0 < x < L_x$, $0 < y < L_y$, and $0 < z < L_z$; also, for simplicity, the conventional symmetry boundary condition is assumed on the outer boundary plane $x = L_x$, and on the other boundary planes($y = L_y$ and $z = L_z$) the free boundary conditions are assumed.

[12] Current-driven anomalous resistivities have been studied theoretically and experimentally [*Lui*, 2001; *Treumann*, 2001; *Ono et al.*, 2001]. Here, as in the 2-D model, the anomalous resistivity model is assumed in the form,

$$\eta(\mathbf{r}, t) = k_R[V_d(\mathbf{r}, t) - V_C] \qquad \text{for } V_d > V_C,$$
$$= 0 \qquad\qquad\qquad \text{for } V_d < V_C \qquad (4)$$

where $V_d(\mathbf{r}, t) = |\mathbf{J}(\mathbf{r}, t)/\rho(\mathbf{r}, t)|$ is the relative electron-ion drift velocity, and V_C may be a threshold for microinstabilities, and the fast reconnection evolution for different functional forms and parameters has been studied [*Ugai*, 1984, 1992; *Ugai and Kondoh*, 2001; *Ugai and Zheng*, 2005]. Here, $k_R = 0.003$ and $V_C = 12$ are taken.

[13] *Ugai* [1986] demonstrated the current-sheet thinning drastically proceeds when there is no resistivity in $t > 4$ due to the similar effect to the pinch one. Here, the ion-electron drift velocity V_d is in inverse proportional to the half-width of current sheet d, $V_d \propto 1/d$. Therefore anomalous resistivity builds up when $d \sim d_0/V_C$ from equation (4), where d_0 is the initial half-width of the current sheet. If we assume that the anomalous resistivity builds up when the half-width of the current sheet d becomes the ion inertial length λ_i, $V_C = 12$ means that initial half-width of the current sheet $d_0 \sim 10 \times \lambda_i$.

[14] In order to disturb the initial static configuration, a localized resistivity model is assumed around the point (L_x, 0, 0) in the 3-D form,

$$\eta(\mathbf{r}) = \eta_0 exp\left[-((x - L_x)/k_x)^2 - (|y|/k_y)^3 - (|z|/k_z)^3\right] \qquad (5)$$

Here, we take $k_x = k_y = 0.8$ and $\eta_0 = 0.02$ in the manner similar to the previous 2-D simulations. Also, k_z provides the 3-D effects. Previous studies showed that the k_z specifies the effective extent of diffusion region in the z-direction, and magnetic reconnection cannot effectively grow in the case of small k_z, $k_z < 3$ say, [*Kondoh et al.*, 2006], so that $k_z = 5$ is taken in the present study.

[15] The disturbance (5) is imposed only in the initial time range $0 < t < 4$, and the anomalous resistivity model (4) is assumed for $t > 4$. Hence the fast reconnection mechanism may be triggered at $x = L_x$ in this model.

[16] It should be noted that sufficiently small mesh sizes are required for precise computations of the spontaneous fast reconnection evolution, so that we assume $\Delta x = 0.04$, $\Delta y = 0.015$, and $\Delta z = 0.1$. Also, we take the magnetic field region size $Y_1 = 4$, and the whole computational region size is assumed to be $L_x = 10$, $L_y = 9.6$ and $L_z = 9.8$.

[17] Here, we should notice that the positive x, y, z-directions in this paper correspond to the tailward, southward, and eastward directions in the usual GSM coordinate system, respectively, as shown in Figure 1. The coordinate system whose axis names are parenthesized in this figure is that employed in our simulation. Hereafter, this coordinate system is used in this paper.

3. Results

3.1. General Remarks

[18] At first, let us show the general results of three-dimensional MHD simulations. Initiated by the disturbance (5) given at $x = L_x$ in the finite extent $|z| < k_z = 5$, the current sheet thinning occurs near the $x = L_x$, giving rise to distinct enhancement of the current density. Once the drift velocity V_d exceeds the given threshold (4) at time $t \sim 27$, fast reconnection drastically grows as nonlinear instability because of the positive feedback between the anomalous resistivity and the reconnection flow. Ahead of the Alfvénic reconnection jet flowing to negative x-direction, a large-scale 3-D plasmoid is formed and propagates, and the plasmoid collides with the $x = 0$ wall boundary at $t \sim 39$, giving rise to a 3-D magnetic loop. Here, we define the "plasmoid" as high pressure plasma wrapped by the reconnected magnetic filed lines.

[19] Now, we show the resulting three-dimensional configurations of some physical quantities. Figure 2 shows the magnetic field configuration (light lines) and plasma pressure configuration(isosurfaces of $P = 1.3$, 1.45, 1.6 and contour lines in the $x - z$ plane) at different times t = 36.75 (Figure 2, left), 37.5 (Figure 2, middle), and 38.25 (Figure 2, right). The grid size shown in Figure 2 is taken to be 2.0. The magnetic field lines shown in this figure are traced from the points on the $x - z$ plane; namely, all of these field lines are the reconnected magnetic field lines. The spatial interval of the start point of these field lines in the z-direction is 0.4. The grid size and the spatial interval are taken to be the same also in the similar figures that will appear. Most of the start points of magnetic field lines in this figure exist in the region $z \lesssim 1.6$; that is, the magnetic reconnection mainly occurs in $z \lesssim 1.6$ near the $x = L_x$. The 3-D plasmoid propagates to the negative x-direction as it grows. This plasmoid is pushed by the slow shock formed behind it, and the high pressure region
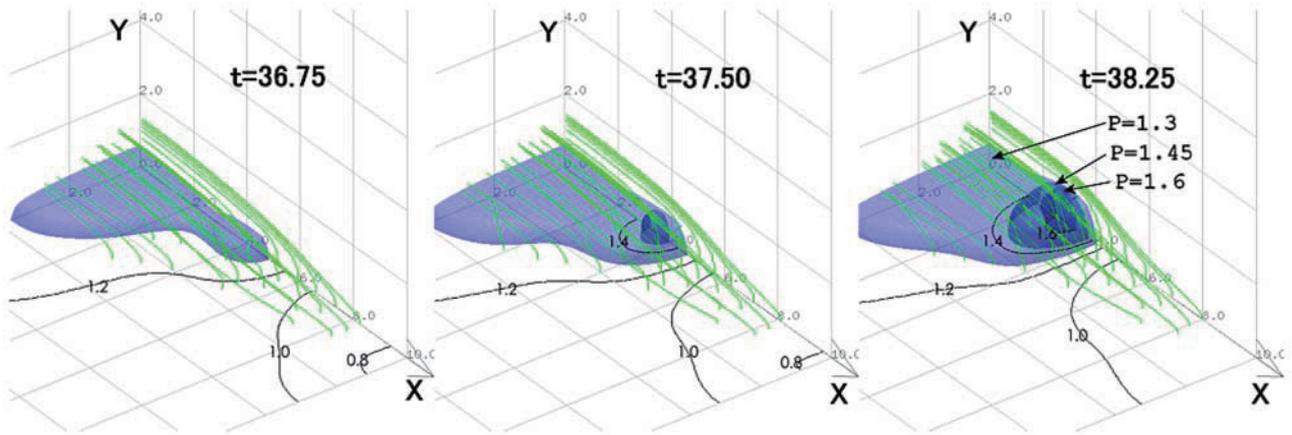
Figure 2. Magnetic field (light lines) and plasma pressure (isosurfaces (P = 1.3, 1.45, 1.6) and contour lines in the $x - z$ plane) configurations at time t = 36.75, 37.5, 38.25. All the similar figures that will appear will be shown in this way.

is produced in the plasmoid, whereas the plasma pressure decreases in the upstream region of it.

[20] We find that the 3-D configuration of plasma density is considerably different from that of plasma pressure; in particular, those configurations inside the plasmoid are quite different. Figure 3 shows the magnetic field configurations (light lines) and the plasma density configurations (isosurfaces of ρ = 1.02, 1.07 and contour lines in the $x - z$ plane) at the times same as those in Figure 2. The magnetic field lines are the same as those in Figure 2. We see that the plasma density is rather low in the plasmoid, where the plasma pressure is very high there. This means that the plasma temperature in the plasmoid is high. Dense plasma in the region ahead of the plasmoid, shown by the multi-isosurfaces, consists of the initial current sheet plasma. This region is enclosed by the magnetic field lines that have been reconnected. Therefore the high-density region does not significantly expand in the y-direction, whereas the plasmoid (high-pressure region) tends to expand in the y-direction. In the region behind the plasmoid ($x(t = 38.25) \gtrsim 4.0$) and $|z| \lesssim 2.0$, the plasma density is low near the x-axis in connection with the fast reconnection jet and strongly depends on the

z-position, whereas the plasma pressure is almost uniform in the z-direction.

[21] The plasma density configuration behind the plasmoid is closely related to the fast plasma flow configuration. Figure 4 shows the magnetic field configurations (light lines) and the plasma flow configurations (isosurfaces of $V_x = -0.4, -1.2, -2.0$ and arrows in the $x - y$ and $x - z$ plane) in the similar manner as before. It is clear that the fast plasma jet flows in the very narrow channel, and the plasma density is low in the channel where the plasma flow is fast. Very fast plasma jet $|V_x| > 1.2$ is especially localized in $|z| < 1.0$, although the magnetic reconnection occurs in $z \lesssim 1.6$ near $x = L_x$. The relatively slow plasma jet, $|V_x| \sim 0.4$, flows in the somewhat wide channel, $|z| < 2.0$. After the plasma arrives at the magnetic loop top, it flows mainly along the magnetic loop boundary.

[22] In order to see the temporal evolution of fast plasma flow, Figure 5a shows the plasma flow profile along the x-axis at different times. The horizontal dotted line indicates the Alfvén velocity V_{Ae} ($\sim$2.7) in the ambient lobe magnetic field region ($1 < y < Y_1$). As time passes, the maximum velocity of the reconnection outflow $|V_x|$ rapidly
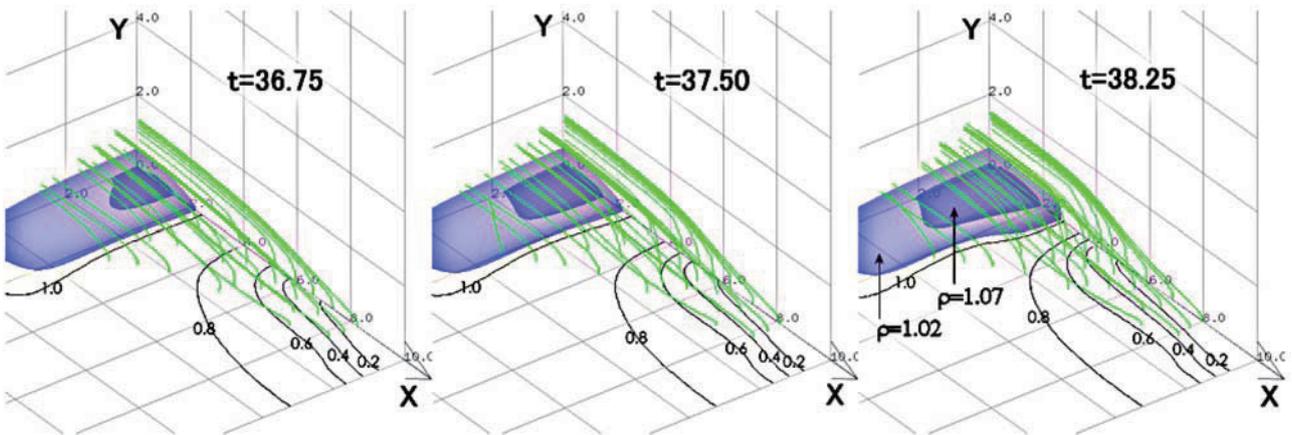
Figure 3. Magnetic field (light lines) and plasma density (isosurfaces (ρ = 1.02, 1.07) and contour lines in the $x - z$ plane) configurations at time t = 36.75, 37.5, 38.25.
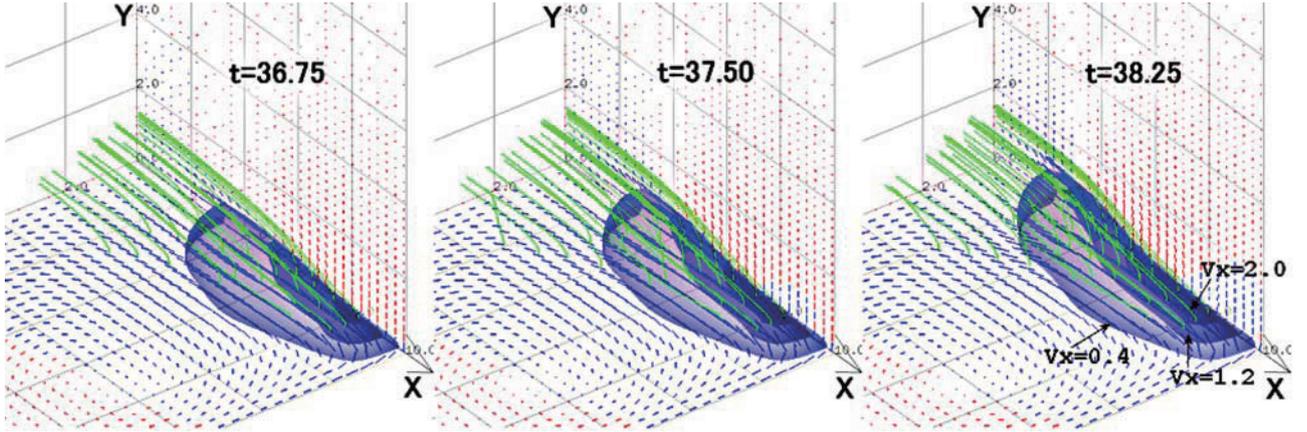
Figure 4. Magnetic field (light lines) and plasma flow (isosurfaces of x-component ($V_x = -0.4, -1.2, -2.0$) and arrows in the $x - y$, $x - z$ plane) configurations at time t = 36.75, 37.5, 38.25.

increases until $t \sim 39$. After $t \sim 39$, reconnection outflow reaches the Alfvén velocity and the profile has the flat plateau.

[23] Next, let us consider the variation of V_x at different positions in the z-direction in order to see the z-directional range of the narrow channel of the fast plasma jet. Figure 5b shows the x-directional variations of $|V_x|$ along the different lines at $y = 0$ in the positive z-direction at time t = 42, when the magnetic loop has fully developed and the maximum velocity attains the lobe Alfvén velocity. We see that the outflow velocity becomes notably smaller for the larger z-positions; only in $|z| < 1.0$, it becomes comparable with lobe Alfvén velocity V_{Ae}. In fact, in $|z| > 1.5$, the maximum value is less than $\sim 30\%$ of V_{Ae}, so that the spatial range in the z-direction of this narrow channel may be $|z| \lesssim 1.0$, and that of the region observable fast flow may be $|z| \lesssim 2.0$.

[24] As mentioned above, these four physical quantities (P, ρ, V_x, B) have different time evolutions in each other, and three-dimensional evolutions of plasma pressure and density configurations are different from those of plasma flow and reconnected magnetic field lines configurations. The extent, especially in the z-direction, of fast flow and reconnected field lines are restricted in the narrow channel differently from that of dense plasma region and high pressure region.

3.2. Comparison With Superposed Epoch Analysis

[25] Here, let us introduce the actual satellite observations. Since the fast flow events have individually different characteristics, we may compare the simulation results with the results of statistical analyses of satellite observations. Superposed epoch analysis introduced in section 1 has been conducted in order to understand the generic magnetic field and plasma parameter change in the course of the fast plasma flow in the plasma sheet.

[26] Figure 6 shows the stack plots of the various physical quantities superposed over the 818 fast earthward flow events reported by *Ohtani et al.* [2004]. These events were observed at $-31 < X_{GSMA} < -5\ R_E$ during October 1993 to July 2001 by Geotail satellite. The reference time in this analysis, which is shown by the vertical solid line ($\delta T_V = 0$) in this figure, is the start of the fast flow. The horizontal axis covers the 20-min interval centered at the reference time. We should notice that the coordinate system in this figure is GSM system differently from that used in our simulations. In order to prevent the confusion due to this difference between coordinate systems, we use the coordinate system in our simulation in text and we put the variables in parenthesis, which is corresponding one in our coordinate system, in this figure.

[27] At first, let us focus to the time profiles of the flow velocity component $|V_x|$ (Figure 6a), the three magnetic field components B_x, $|B_y|$, $|B_z|$ (Figure 6b), the plasma density ρ (Figure 6c) and the plasma pressure P (Figure 6e). The peak times, when these intensities reach peak values, are different from each other, and the order of their appearance time is ρ, P, $|B_y|$, and $|V_x|$. The peak time of
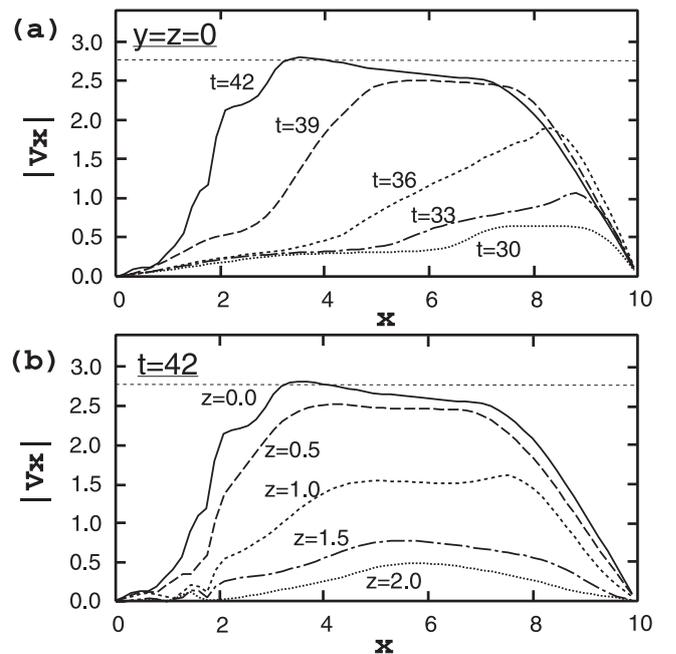
Figure 5. (a) Profiles of $|V_x|(x, y = z = 0)$ along the x-axis at different times, and (b) x-directional profiles of $|V_x|(x, y = 0, z)$ at t = 42 at different z locations.
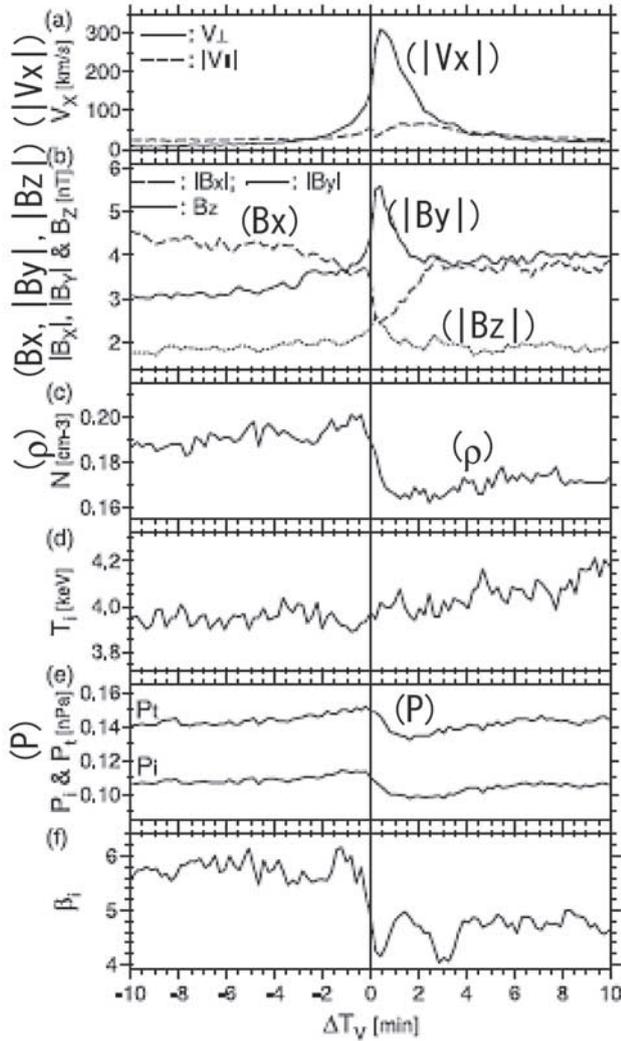
Figure 6. Various magnetic field and plasma quantities superposed over the 818 fast earthward flow events in GSM coordinate system. The start of the fast flow is used as a reference time (from *Ohtani et al.* [2004]). The variables in parenthesis indicate the corresponding ones in our coordinate system.

plasma pressure is coincident with the reference time. B_x decreases in association with the fast flow, then it falls down to the lowest point around the reference time, and then it recovers to the same level as $|B_y|$ in association with the slowing down of the fast flow. $|B_y|$ stays at levels higher than the previous levels after overshooting, namely the local magnetic field dipolarizes.

[28] The ion temperature T_i (Figure 6d) increases around the reference time. The decrease in ρ overcompensates for that increase of T_i. The value of β_i (Figure 6f) decreased in the course of the fast flow; however, we note that the minimum value of β_i is 4.0. Judging from the above, they inferred that the near-Earth reconnection is the responsible mechanism for most flow events.

[29] The magnetic field component B_x is nonzero ($2.5 \lesssim |B_x| \lesssim 4.5$) and the maximum flow velocity is only 300 km/s. From this, we can infer that the major part of these events are observed at the region off the neutral sheet in the

y-direction and outside the fast flow narrow channel in the *z*-direction.

[30] In their paper, they also examined the temporal structure of the fast flow using simulations, and compared them with the observations. In their simulations, the convergent flows generated by two X lines produces a significant enhancement of density and strong density peak in the magnetic island. Virtual satellites, which are located in the symmetry plane at the center of the current sheet, observed a strongly bipolar signature in B_y with a sudden enhancement in the density. They mentioned the discrepancies in ρ variation between the observations and the simulations. The superposed profile of ρ showed a transient enhancement followed by a sharp decrease. On the other hand, the modeled density decrease was much more gradual than the preceding density enhancement. They mentioned two possible reasons, (1) the distance from the reconnection site and (2) large density difference between the central plasma sheet and the lobes. Another discrepancy in ρ is that the peak coincides with the minimum of B_y in the superposed epoch analysis, which can also be confirmed for some individual events. Namely, the density starts to decrease before the reference time.

3.3. Virtual Observation

[31] In order to compare the simulations directly with the satellite observations, let us consider such a virtual probe that is located at a spatial point (x, y, z) in the plasma sheet. For simplicity, we assume that the probe does not move in spite of the progress of time. Then, the virtual probe can readily detect the temporal changes of plasma quantities at its location on the basis of the simulation results. In what follows, we examine what the virtual probe observes as the fast reconnection mechanism builds up and proceeds.

[32] As mentioned in section 3.2, it seems that a lot of events are observed in the region off the neutral sheet in the *y*-direction and outside the fast flow narrow channel in the *z*-direction. Therefore the virtual probe is located at the position ($x = 4.0$, $y = 0.4$. $z = 0.6$), and we compare the virtual probe observations with the actual satellite observations. Figure 7 shows the temporal variations of the plasma flow velocity (Figure 7, top), magnetic field components (Figure 7, middle), plasma pressure and plasma density (Figure 7, bottom) detected by the virtual probe located at the above position. Solid vertical line indicates the peak time T_P of plasma pressure P, and the other dotted vertical lines indicate the peak times T_ρ, T_B, T_V of ρ, $|B_y|$, and $|V_x|$, respectively. The order of these peak times are consistent with those in Figure 6. Namely, first, the plasma density ρ slightly increases until $t = T_\rho$, and then abruptly decreases. Second, the plasma pressure P largely increases until $t = T_P$, and then abruptly decreases. Third, the *y*-component of magnetic field $|B_y|$ sharply increases from $t \sim T_\rho$ until $t = T_B$, and then decreases to the higher level than the previous level until $t \sim T_V$ and keeps the level. Finally, the *x*-component of plasma flow velocity $|V_x|$ gradually increases and sharply increases at $t \sim T_P$, and then decreases after $t = T_V$.

[33] The magnetic field component B_x starts to decrease at the same time with the start of the enhancement of $|V_x|$, and then slightly recovers from T_ρ, and then falls down again at T_B, and it starts to recover at the peak time of $|V_x|$.
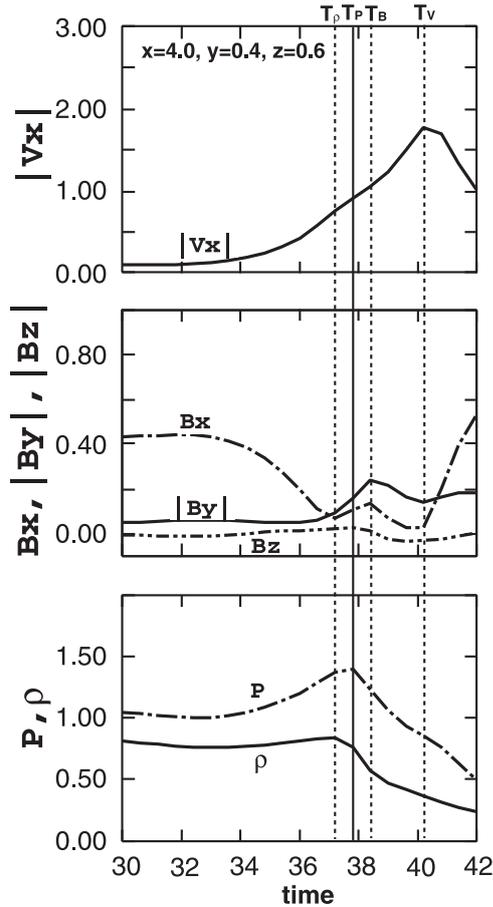
Figure 7. Time profiles of $|V_x|$, B_x, $|B_y|$, $|B_z|$, P, and ρ observed at position $x = 4.0$, $y = 0.4$, $z = 0.6$.

The former slight recovery of B_x, whose amount of enhancement is much smaller, synchronizes with the enhancement of $|B_y|$. The latter recovery of B_x means that the virtual probe goes out from the narrow channel of fast plasma jet. This profile of B_x and the relationship between B_x and V_x variations are also consistent with those in actual satellite observations. Therefore, the earthward fast plasma flow can be regarded to be the direct outcome of the fast reconnection evolution itself.

3.4. Traveling Compression Regions

[34] Earthward fast plasma flow in the plasma sheet is often accompanied by traveling compression regions (TCRs) in the magnetic lobe region [*Slavin et al.*, 1984, 1993, 2002]. *Ugai and Zheng* [2006a, 2006b] examined tailward moving TCRs in the Earth's magnetotail using the similar model with that in this paper. They demonstrated that the general features of TCR, obtained by the simulations, are in good agreement with the actual satellite observations, and the virtual satellite which is located behind the plasmoid formation observes the unipolar structure of the southward field component. Then, they concluded that the TCR event is directly caused by the fast reconnection mechanism itself. In their model, free boundary is used instead of the wall boundary adopted in our simulation, so that the plasmoid propagates freely without any obstacle, whereas it collides with ($x = 0$) wall boundary in our present

model. Also, it is important to examine simultaneously earthward moving TCR in this model. In this section, we examine particularly the connection between earthward fast plasma flow and earthward moving TCR.

[35] At first, we show an example of an earthward moving TCR in the Cluster FGM measurements on 10 September 2001 reported by *Slavin et al.* [2005] in Figure 8. In this figure, we put the variables in parenthesis, which is corresponding one in our coordinate system and use the coordinate system in our simulation in text in order to prevent the confusion in the same way of the Figure 6. The peak compression in $-B_x$ (Figure 6, top), which is indicated by the vertical dashed line, occurs near the center of the ΔB_y variations. The profile of $-B_y$ is unipolar, and it stays at level higher than the previous level after the overshooting. On the other hand, $-Bx$ stays at lower level after the overshooting. It means that the lobe magnetic field lines gently slope down tailward.

[36] In order to compare the simulation results directly to the satellite observations, let us consider such a virtual satellite that is located at a spatial point (x, y, z) in the 3-D simulation domain in the similar manner to the previous investigations on the fast flow events in the plasma sheet. Top four panels in Figure 9 show the temporal variations of the field components, $-B_x$, $-B_y$, $-B_z$ and B, observed at $x = 4.0$, $y = 1.8$, $z = 1.8$ in our simulation domain. We then find that the characteristic compression signature in B_x and draping signature in B_y are consistent with the actual satellite observations in Figure 8. The unipole variation of B_y and tailward slope down of lobe magnetic field are also consistent. Inferring from the results of *Ugai and Zheng* [2006a, 2006b], these unipole variations of B_y in actual and virtual satellite observations are attributed to the satellite location, namely, these satellites were located behind the plasmoid formation. The variation of B_z in this figure synchronizes with that of B_y. On the other hand, no clear variation is seen in B_z in Figure 8. We find, however, that such variation in B_z as seen in Figure 9 is usually found out in the other cases [*Slavin et al.*, 2005]. In order to examine the relation to earthward fast plasma flow, the temporal variations of the plasma pressure P (third and second panel from the bottom), observed at $x = 4.0$, $y = 0.4$, $z = 1.8$ (third panel from the bottom) and $z = 0.6$ (second panel from the bottom) are shown in Figure 9. The time of peak compression in B_x is coincident with the peak time of plasma pressure enhancement in both of bottom two panels. At the location just under the probe position in the magnetic lobe region (third panel from the bottom), the enhancement of plasma pressure is not so large, while the profile of plasma pressure varies much larger at the location inner the position (second panel from the bottom). This three-dimensional effect seems to be effective to the compression of the lobe magnetic field in extensive region in the z-direction. This compression of lobe magnetic field is due to the passage of plasmoid in the plasma sheet, so that this TCR observation is closely relation to the earthward fast flow event.

4. Summary and Discussion

[37] In this paper, the spontaneous fast reconnection model was applied to the earthward fast flow events
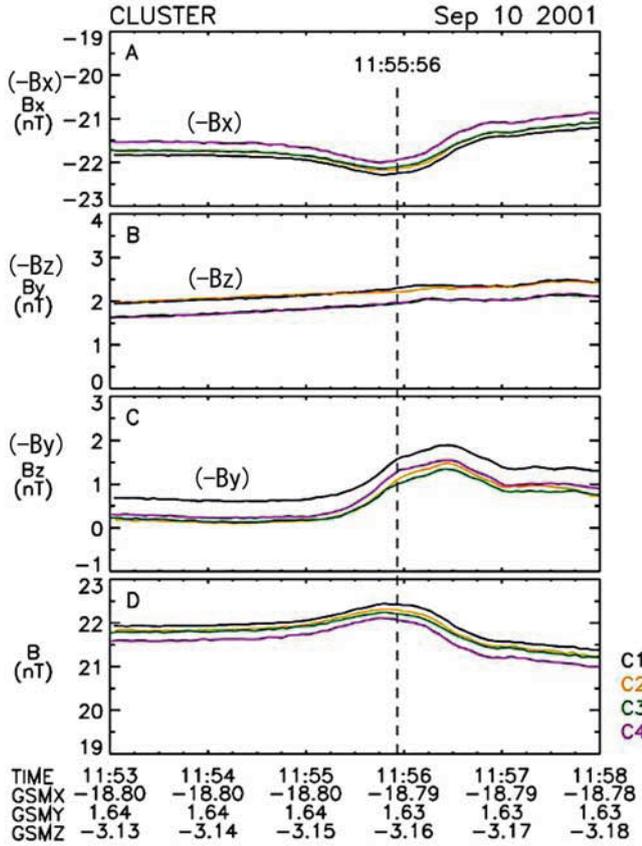
Figure 8. An example of earthward moving traveling compression region in the Cluster measurements on 10 September 2001 in GSM coordinate system (from *Slavin et al.* [2005]). The variables in parenthesis indicate the corresponding ones in our coordinate system.

observed in the near-Earth plasma sheet. The 3-D plasma configurations found in the present simulation are summarized as follows: (1) The fast reconnection jet, which attains the Alfvén velocity V_{Ae}, flows in the very narrow channel, and the decrease in the plasma density becomes much larger in the region where the flow velocity is faster. (2) The plasmoid is formed ahead of the fast reconnection jet and develops in three dimensions, and the plasma inside the plasmoid becomes high pressure and high temperature. (3) The plasma pressure in the fast flow channel tends to decrease, where the temperature (P/ρ) is remarkably enhanced. (4) The dense plasma ahead of the plasmoid, which consists of the initial plasma sheet plasma, is confined by the initially reconnected magnetic field lines. (5) The high pressure region and the high density region tend to expand in three dimensions, whereas the fast flow and the intense magnetic field, accumulated behind the plasmoid, remain to be restricted in the very narrow channel.

[38] On the basis of the simulation results, virtual probes were located in the plasma sheet and the lobe region, so that they could directly observe the temporal variations of these plasma quantities. The observations by these virtual probes were in good agreement with the actual satellite observations on both events in the very fine structure. In particular, the sequence of peak times of the plasma quantity variations

was a complete agreement. The order of this is T_ρ, T_P, T_B, and T_V. This order reflects the results of three-dimensional plasma configurations mentioned above.

[39] Now, let us discussed about the plasma density profile in the course of fast flow. We have introduced the discrepancy between the observations and simulations in the work of *Ohtani et al.* [2004], that is, the superposed profile of ρ showed a transient enhancement followed by a sharp decrease, and the density starts to decrease before the reference time. In their model, they assumed multi X line, so that the plasma density enhances in the magnetic island. As the result of it, the density does not decrease before the plasma pressure enhancement. On the other hand, the profile of plasma density in our simulation was in the good agreement with superposed profile, since the plasma density increase before the start of fast flow and decrease in the course of fast flow.

[40] *Ohtani et al.* [2004] compared the earthward flows with the tailward ones using the satellite observations. They found that the ion temperature increases for the earthward flow and decreases for the tailward flow, and inferred that this contrast may be attributed to different physical conditions surrounding the flow, that is, the associated structure gets confined in the *x*-direction for the earthward flow. In fact, the superposed profile of T_i obtained for earthward flow in Figure 6 shows a increase in the course of fast flow. On the other hand, the profile of plasma temperature T, which is observed at $x = 4.0$, $y = 0.4$, $z = 0.6$ in our simulation domain, shown in the bottom panel in Figure 9 similarly shows a increase in the course of fast flow. We note that the peak time of plasma pressure was coincident with the start of fast flow. Hence it may be reasonable to model the earthward flow by the wall boundary, which significantly confine the reconnection flow in the *x*-direction.

[41] It may be important to explicitly give the normalization units employed in the present simulation. Some laboratory experiments detected such an anomalous resistivity when the half-width of the current sheet becomes smaller than the ion inertial length λ_i [*Yamada*, 2001; *Ono et al.*, 2001]. In the present model, the current-driven anomalous resistivity builds up when the width of the current sheet becomes thin to about one-tenth of the initial width as mentioned in section 2.1.2. Hence the present unit length d_0 may be considered to be about $10\lambda_i$. If we typically take the field strength in the lobe region $B_{x0} = 20$ nT, the number density in the center of plasma sheet $n_0 = 1$ cm^3 and $\lambda_i \sim 1000$ km, we have the unit velocity $V_{Ax0} = 450$ km/s; therefore, the present unit time becomes $d_0/V_{Ax0} \sim 20$ s. Applying these quantities to Figures 7 and 9, we see that the simulation results are consistent with the actual observations in Figures 6 and 8, respectively.

[42] In summary, we have demonstrated that the earthward fast flow event and TCRs signatures are directly caused by the fast reconnection mechanism. The virtual probe observations in the simulation domain are in the good agreement with the actual satellite observations, despite the simulation model is simplified and idealized. For instance, the present simulations are done in the first-quadrant domain, so that the initial shear field in the plasma sheet is not considered; also, a realistic geomagnetic initial dipole field is not considered. These conditions may be important
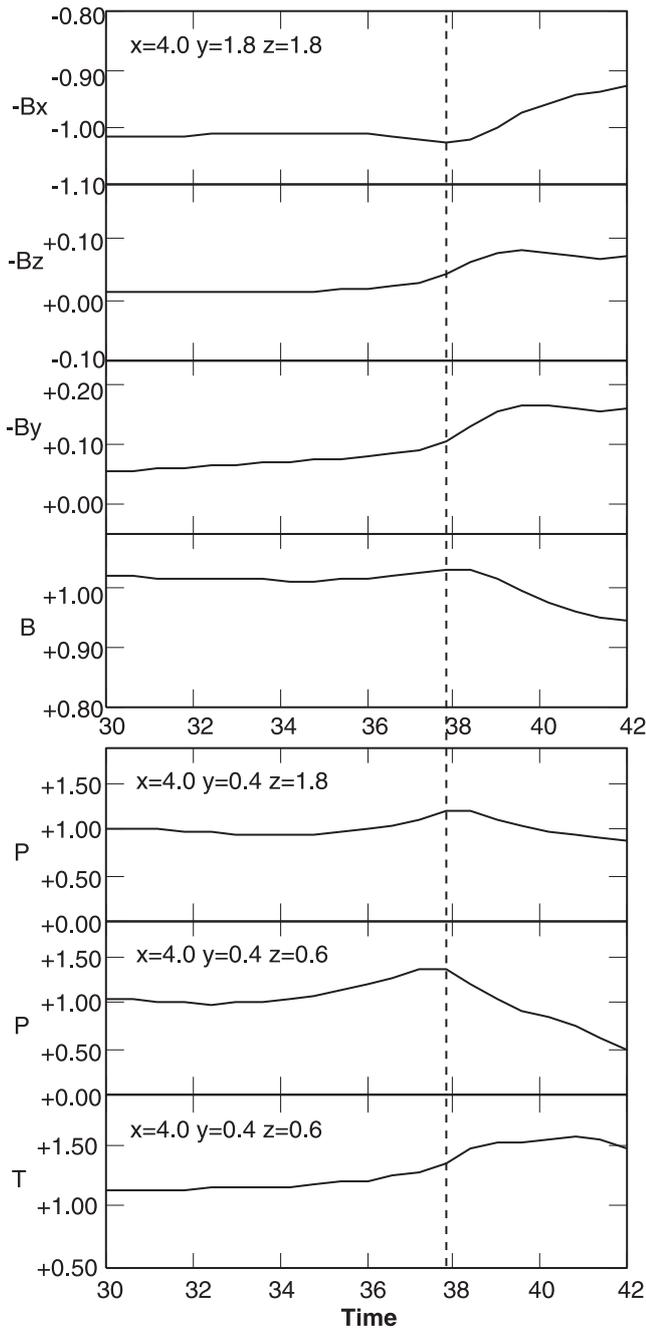
Figure 9. Traveling compression region in the virtual probe measurements in our simulation domain. Top three panels show the magnetic field variations in the lobe region, and bottom three panels show the plasma pressure and temperature variations in the plasma sheet during the same period.

in the real time predictions, which will be taken into account in the future work.

References

Angelopoulos, V., W. Baumjohn, C. F. Kennel, F. V. Voroniti, M. G. Kivelson, R. Pellat, R. J. Walker, H. Luhr, and G. Paschmann (1992), Bursty bulk flows in the inner central plasma sheet, *J. Geophys. Res.*, *97*, 4027.

Kondoh, K., M. Ugai, and T. Shimizu (2006), Three dimensional computer simulation of magnetic loop associated with magnetic reconnection, *Adv. Space Res.*, *37*, 1301.

Lui, A. T. Y. (1996), Current disruption in the Earth's magnetosphere: Observations and models, *J. Geophys. Res.*, *101*, 13,067.

Lui, A. T. Y. (2001), Multiscale model for substorms, *Space Sci. Rev.*, *95*, 325.

Ohtani, S.-I., M. A. Shay, and T. Mukai (2004), Themporal structure of the fast convective flow in the plasma sheet: Comparison between observations and two-fluid simulations, *J. Geophys. Res.*, *109*, A03210, doi:10.1029/2003JA010002.

Ono, Y., M. Inomoto, Y. Ueda, T. Matsuyama, and Y. Murata (2001), Fast compression of a current sheet during externally driven magnetic reconnection, *Earth Planets Space*, *53*, 521.

Petschek, H. E. (1964), Magnetic field annihilation, in *The Physics of Solar Flares*, edited by W. N. Hess, *NASA Spec. Publ.*, SP-50, 425.

Schodel, R., W. Baumjohann, R. Nakamura, and T. Mukai (2001), Rapid flux transport and plasma sheet reconfiguration, *J. Geophys. Res.*, *106*, 8381.

Sergeev, V. A., et al. (2000), Multiple-spacecraft observation of a narrow transient plasma jet in the Earth's plasma sheet, *Geopys. Res. Lett.*, *27*, 851.

Sergeev, V. A., et al. (2001), Substorm and convection bay compared: Auroral and magnetotail dynamics during convection bay, *J. Geophys. Res.*, *106*, 18,843.

Shibata, K. (1999), Evidence of magnetic reconnection in solar flares and a unified model of flares, *Astrophys. Space Sci.*, *264*, 129.

Shibata, K. (2004), *Physics of Magnetic Reconnection in High-Temperature Plasmas*, edited by M. Ugai, pp. 193–237, Research Signpost, India.

Shiokawa, K., et al. (1997), Braking of high-speed flows in the near-Earth tail, *Geophys. Res. Lett.*, *24*, 1179.

Slavin, J. A., R. P. Lepping, and D. N. Baker (1984), Substorm associated traveling compression regions in the distant tail-ISEE-3 geotail observations, *Geopys. Res. Lett.*, *11*, 657.

Slavin, J. A., et al. (1993), ISEE 3 observations of traveling compression regions in the Earth's magnetotail, *J. Geophys. Res.*, *98*, 15,425.

Slavin, J. A., et al. (2002), Simultaneous observations of earthward flow bursts and plasmoid ejection during magnetospheric substorms, *J. Geophys. Res.*, *107*(A7), 1106, doi:10.1029/2000JA003501.

Slavin, J. A., et al. (2003), Cluster four spacecraft measurements of small traveling compression regions in the near-tail, *Geophys. Res. Lett.*, *30*(7), 1362, doi:10.1029/2002GL016411.

Slavin, J. A., et al. (2005), Cluster observations of traveling compression regions in the near-tail, *J. Geophys. Res.*, *110*, A06207, doi:10.1029/2004JA010878.

Treumann, R. A. (2001), Origin of resistivity in reconnection, *Earth Planets Space*, *53*, 453.

Ugai, M. (1984), Self-consistent development of fast magnetic reconnection with anomalous plasma resistivity, *Plasma Phys. Controlled Fusion*, *26*, 1549.

Ugai, M. (1986), Global dynamics and rapid collapse of an isolated current-sheet system enclosed by free boundaries, *Phys. Fluids*, *29*, 3659.

Ugai, M. (1987), MHD simulations of fast reconnection spontaneously developing in a current sheet, *Comput. Phys. Commun.*, *49*, 185.

Ugai, M. (1992), Computer studies on development of the fast reconnection mechanism for different resistivity models, *Phys. Fluids*, *4*, 2953.

Ugai, M., and K. Kondoh (2001), Computer studies on the spontaneous fast reconnection evolution in various physical situations, *Phys. Plasmas*, *8*, 1545.

Ugai, M., and L. Zheng (2005), Conditions for the fast reconnection mechanism in three dimensions, *Phys. Plasmas*, *12*, 092312.

Ugai, M., and L. Zheng (2006a), Parametric studies on traveling compression regions observed in the Earth's magnetotail, *Phys. Plasmas*, *13*, 1.

Ugai, M., and L. Zheng (2006b), Modeling of traveling compression regions in the Earth's magnetotail by the spontaneous fast reconnection model, *Phys. Plasmas*, *13*, 032901.

Ugai, M., K. Kondoh, and T. Shimizu (2003), Computer simulations on three-dimensional magnetic loop dynamics by the spontaneous fast reconnection model, *Phys. Plasmas*, *10*, 357.

Ugai, M., K. Kondoh, and T. Shimizu (2004), Computer studies on the three-dimensional spontaneous fast reconnection model as a nonlinear instability, *Phys. Plasmas*, *11*, 1416.

Ugai, M., K. Kondoh, and T. Shimizu (2005), Spontaneous fast reconnection model in three dimensions, *Phys. Plasmas*, *12*, 042903.

Yamada, M. (2001), Review of the recent controlled experiments for study of local reconnection physics, *Earth Planets Space*, *53*, 509–519.

_____

K. Kondoh and M. Ugai, Electrical and Electronic Engineering and Computer Science, Graduate School of Science and Engineering, Ehime University, 790-8577, Matsuyama, Japan. (kondo@cs.ehime-u.ac.jp)

Flux enhancement of the outer radiation belt electrons after the arrival of stream interaction regions

Yoshizumi Miyoshi[1] and Ryuho Kataoka[1,2]

[1] The Earth's outer radiation belt electrons increase when the magnetosphere is surrounded by the high-speed solar wind stream, while the southward interplanetary magnetic field (IMF) is also known as an important factor for the flux enhancement. In order to distinguish the two different kinds of solar wind parameter dependence statistically, we investigate the response of the outer belt to stream interaction regions (SIRs). A total of 179 SIR events are identified for the time period from 1994 to 2005. We classify the SIR events into two groups according to the so-called "spring-toward fall-away" rule: IMF sector polarity after the stream interface is toward in spring or away in fall (group A) and vice versa (group B). According to the Russell-McPherron effect, groups A and B have a significant negative and positive offset of the IMF B_z after the stream interface, respectively. Comparing groups A and B by superposing about the stream interface, only IMF B_z dependence can be obtained because the other solar wind parameters change in the same manner. As a result, the greatest flux enhancement is found in the highest-speed streams with a southward offset of the IMF B_z, indicating that only the solar wind speed by itself is not a sufficient condition for the large flux enhancement. It is also found that the large flux enhancement tends to be associated with weak geomagnetic activities with minimum Dst of about -50 nT on average, implying that the existence of intense magnetic storms is not essential for the flux enhancement.

Citation: Miyoshi, Y., and R. Kataoka (2008), Flux enhancement of the outer radiation belt electrons after the arrival of stream interaction regions, *J. Geophys. Res.*, *113*, A03S09, doi:10.1029/2007JA012506.

1. Introduction

[2] Understanding and modeling the flux enhancement of the outer belt electrons is a central topic of the space weather research because the energetic electrons can cause anomalies of the satellites that provide weather observations, communications, and other important services to modern life [*Baker et al.*, 1987, 1998; *Kataoka and Miyoshi*, 2006]. The energetic electrons of the outer belt are accelerated up to the energy range of a few MeV in the inner magnetosphere. Classically, the inward radial diffusion and the betatron acceleration are thought to be a fundamental process to accelerate the energetic electrons [*Schulz and Lanzerotti*, 1974; *Elkington et al.*, 1999; *Li*, 2006]. Wave-particle interaction is also known as another essential mechanism to accelerate the energetic electrons [*Summers et al.*, 1998; *Meredith et al.*, 2001; *Miyoshi et al.*, 2003; *Horne et al.*, 2006]. Both processes are expected to be effective to work simultaneously for the acceleration [*O'Brien et al.*, 2003], although there is still no consensus about the relative importance.

[3] The outer radiation belt fluxes are not stable, showing a somewhat systematic change especially during magnetic storms; rapid decrease in the main phase and gradual increase in the recovery phase. *Reeves et al.* [2003] showed that about a half of magnetic storms produce the flux enhancement of the outer belt. *Miyoshi and Kataoka* [2005] found that the flux enhancement is controlled by the interplanetary driver of storms, such as coronal mass ejections (CMEs) and stream interaction regions (SIRs), suggesting that the large amplitude fluctuation of the interplanetary magnetic field (IMF) directions which has a significant southward component within the high-speed coronal hole stream following a SIR is important for the acceleration of energetic electrons. In fact, *Kataoka and Miyoshi* [2006] showed that a strong flux enhancement at geosynchronous orbit followed the 5 of 6 SIR-associated storms, while only 20 of 49 CME-associated storms have such strong enhancements. The solar wind structure dependence possibly places observational restrictions on the effective acceleration mechanism.

[4] Another important observational restriction is the solar wind parameter dependence. It has been well known that the outer belt flux enhancements are well correlated with the high-speed solar wind [*Paulikas and Blake*, 1979; *Baker et al.*, 1986; *O'Brien et al.*, 2001; *Weigel et al.*, 2003]. Several case studies have shown that the southward IMF is another important parameter for the acceleration [*Blake et al.*, 1997; *Iles et al.*, 2002; *Miyoshi et al.*, 2007].

---

[1]Solar-Terrestrial Environment Laboratory, Nagoya University, Nagoya, Japan.

[2]Now at Computational Astrophysics Laboratory, Institute of Physics and Chemical Research, Saitama, Japan.

[5] It is important that the statistical research of the solar wind parameter dependence should take into account the solar wind structure dependence at the same time, although most of the previous studies did not handle this point comprehensively. One straightforward way to do that is collecting a number of SIR events to find a pure solar wind parameter dependence on the flux enhancement without structure dependence. *McPherron and Weygand* [2006] examined a number of SIR events to clarify some statistical properties of geomagnetic activities and fluxes of the outer belt, while they have not shown parameter dependences on the flux enhancement. In this paper, we show the pure solar wind parameter dependence for the flux enhancement in the outer belt associated with the SIR statistically.

2. Data Analysis

[6] SIRs are the regions of compressed plasma formed at the leading edge of high-speed streams as they interact with the preceding slow solar wind. In this study, the stream interfaces of SIRs are detected from the OMNI-2 data (http://omniweb.gsfc.nasa.gov) using the simple criteria of (1) a large change in the azimuth angle of the solar wind velocity (>5°/hour), (2) positive slopes in the speed and temperature and negative slope in density, and (3) the magnetic field strength higher than the weekly average. We use these parameters because stream interfaces are usually characterized by these parameters [*Gosling et al.*, 1978] and using these parameters work successfully to detect a number of stream interfaces [*McPherron and Weygand*, 2006]. We further restrict our selection of the events with relatively stable magnetic fields for 72 hours after the stream interface, where 75% of the magnetic fields directing in the away quadrant (the IMF azimuthal angle from the x axis ranges from 90° to 180°) or 75% of the magnetic fields directed in the toward quadrant (the IMF azimuthal angle from the x axis ranges from 270° to 360°). As a result, a total of 179 SIR events are identified for the time interval from 1994 to 2005. It is worthwhile to note that in this paper we identify the SIR events regardless of the storm amplitude.

[7] In order to categorize the SIR events into different IMF B_z conditions, we consider the Russell-McPherron effect [*Russell and McPherron*, 1973]. The Russell-McPherron effect controls the IMF B_z of GSM coordinate through the projection of the IMF B_y of GSE coordinate. According to the Russell-McPherron effect, geomagnetic activities are enhanced during the time intervals of so-called "spring-toward fall-away" (STFA) because the IMF polarities of toward and away have a large projection component of the southward IMF around the spring and fall equinox, respectively. The southward component of IMF in a high-speed stream is always significantly suppressed or enhanced because of the STFA rule, because of comparable amplitudes of both the Russell-McPherron effect and the Alfvénic IMF fluctuation of a high-speed stream.

[8] According to the STFA rule, we classify the SIR events into two groups: IMF sector after the stream interface is toward in spring (February, March, April, May) or away in fall (August, September, October, November) (group A) and vice versa (group B). As a result of this categorization, two sets of SIR events with the same solar wind structure of different IMF offset after the stream interface can be

obtained. Hereafter we call the group A as SBZ events, and the group B as NBZ events. The total numbers of the SBZ and NBZ events are 75 and 43, respectively.

[9] Figure 1 shows the result of superposed epoch analysis about the solar wind parameters and geomagnetic activities. The reference time ($t = 0.0$; time unit is shown in days) of the superposed epoch analysis is chosen at the time of a stream interface. The fluctuation of the magnetic field dB is calculated as $\sqrt{\sigma_{Bx}^2 + \sigma_{By}^2 + \sigma_{Bz}^2}$, where σ_{Bx}, σ_{By}, and σ_{Bz} are the standard deviations of the x, y, and z components, respectively. The typical solar wind parameter changes associated with a SIR can be found, i.e., speed and temperature increase accompanied with density drop, following high-speed stream [e.g., *Gosling et al.*, 1978; *Richardson et al.*, 2006; *McPherron and Weygand*, 2006]. Following the STFA rule, the average IMF B_z is systematically shifted to southward after $t = 0.0$ day in SBZ events, while the IMF B_z is shifted to northward in NBZ events. Modest Kp activity continues for a few days during SBZ events, while the Kp index is significantly less enhanced in NBZ events. Although the AE index has not been available during the whole period, continuous Kp activities indicate the continuous substorms and enhanced convections, i.e., High Intensity Long Duration Continuous AE Activity (HILDCAA)-like activities [*Tsurutani and Gonzalez*, 1987; *Tsurutani et al.*, 2006]. The solar wind speed is slightly higher in NBZ events, causing slightly hot and tenuous plasma parameters, although the slightly different plasma parameters do not affect the essential results of this paper.

[10] As a direct measurement of the outer belt variation, we use the flux of >2 MeV electrons observed in situ by the GOES satellites at geosynchronous orbit. In addition, the radiation belt electrons over a wide range of L value are continuously monitored by the low-altitude NOAA/POES satellites. Note that L value in this study is McIlwain's L derived from IGRF. We use NOAA 12 and 15 satellites, orbiting 0730–1930 local time meridians, before and after 1 July 1998, respectively. We use the 90° detectors mostly measuring the local mirroring >300 keV electrons. Though >300 keV electron detectors are also sensitive to >440 keV protons [see *Evans and Greer*, 2000], we confirmed that the contamination is not a severe problem by examination of the 240–800 keV ion measurement.

3. Results

[11] First, we show that the flux enhancement in SBZ events is faster and stronger than that in NBZ events. Figures 2a and 2b show the superposed flux of >2.0 MeV electrons for SBZ and NBZ events as observed by the GOES satellites with the superposed Dst index. About 0.5 day before the stream interface, the electron flux begins to decrease in both groups. The flux decrease in SBZ events is relatively faster than that in NBZ events. Figure 2a shows that the average flux begins to increase just after the stream interface during SBZ events. The flux recovers to the pre-SIR level at $t = 1.0$ day and then increases above the pre-SIR level. Figure 2b shows that the flux continues to decrease till one day after the stream interface, and then gradually recovers to the pre-SIR level at $t = 2.0$ day during NBZ events. After $t = 3.0$ day, although the flux is over the pre-
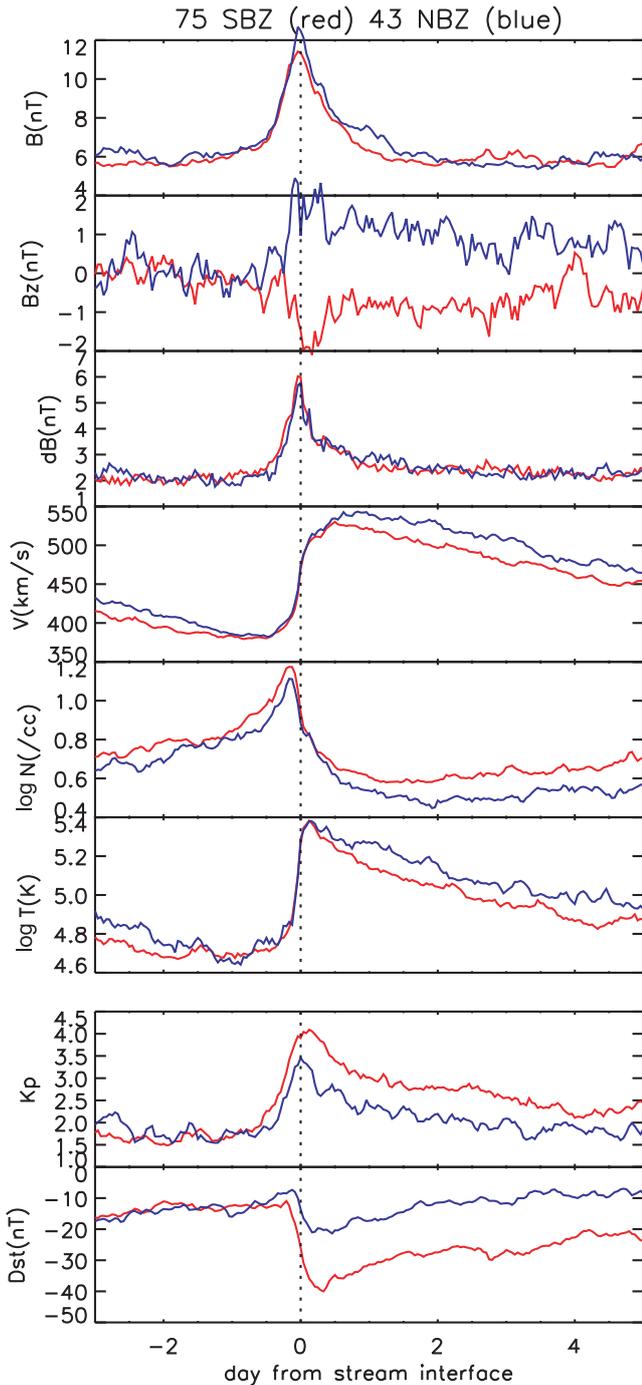
Figure 1. Superposed time series of solar wind parameters and geomagnetic activities. SBZ and NBZ events are shown in red and blue, respectively (see text for the categorization in detail). Shown are (top to bottom) magnetic field strength *B*, southward magnetic field B_z, fluctuation of the magnetic field *dB*, solar wind speed *V*, the logarithm of density *N*, the logarithm of proton temperature *T*, and the *Kp* and *Dst* indices.

SIR level, the amount of the flux enhancement is relatively weak.

[12] Secondly, we examine the location of the outer belt using the data from NOAA/POES satellites. Figures 3a

and 3b are the superposed $L - t$ diagrams of the radiation belt electrons for SBZ and NBZ events, respectively. Similar to the GOES observations, the flux of the outer belt electron begins to decrease just before the stream interface. Note that the flux decrease occurs only at the outer belt at $L > 4.0$ and the flux does not change at the inner portion. Figure 3a shows that flux recovers quickly to the pre-SIR level and then increases over the pre-SIR level during SBZ events. The flux enhancement after the stream interface at the lower *L* shell is faster than that at the higher *L* shell. Figure 3b shows that the flux gradually recovers to the pre-SIR level after the stream interface during NBZ events. As same as Figure 3a, the flux recovers to the pre-SIR level till $t = 1.0$ day at $L < 4.5$, while the recovery takes much time at the outer portion. Similar to the GOES observations in NBZ events (Figure 2b), the flux recovers to the pre-SIR level and the amount of the flux enhancement is relatively weak.

[13] The above results showed that the flux enhancement in the outer radiation belt is significantly controlled by the STFA rule. As a next examination, in order to clarify the effect of the solar wind speed for the flux enhancement, we further categorize SBZ events into two different sets in terms of the solar wind speed. We calculate the average solar wind speed *Vsw* for 72 hours after the stream interface. In SBZ events, 37 events are identified as the solar wind speed more than 500 km/s (hereafter, SBZ-fast), while 38 events are identified as the solar wind speed less than 500 km/s (SBZ-slow). Comparing SBZ-fast and SBZ-slow events, we can investigate the effect of the solar wind speed for the evolution of the outer belt with the same southward IMF offset.

[14] The solar wind parameters and geomagnetic activities for SBZ-fast and SBZ-slow events are shown in Figure 4. The offsets of the southward IMF in both groups are almost the same, while the solar wind speed after the stream interface is clearly different from each other. In SBZ-fast events, the solar wind speed increases up to 600 km/s on average and remains above 500 km/s till $t = 2.0$ day. The temperature of the solar wind in SBZ-fast events is higher than that in SBZ-slow events, which is consistent with statistical results that the solar wind temperature is well correlated with the solar wind speed [e.g., *Borovsky et al.*, 1998; *Borovsky and Steinberg*, 2006a]. Although the *Kp* index of both groups shows continuous activities after the stream interface, the amplitude of the *Kp* index in SBZ-fast events is larger than that in SBZ-slow events. As same as the variation of the *Kp* index, the *Dst* index in SBZ-fast events is smaller than that in SBZ-slow events and keeps small values for several days after the stream interface.

[15] The solar wind speed difference can be clearly seen in Figures 5 and 6 where the format is the same as Figures 2 and 3, respectively. It is interesting to note that the flux recovery and enhancement are very different between SBZ-fast and SBZ-slow events. The flux recovery starts at $t = 0.0$ day in SBZ-fast events, while the recovery start at $t = 0.5$ day in SBZ-slow events as shown in Figure 5. The flux enhancement is faster and stronger in SBZ-fast events, which is similar variation to the SIR-driven storms of $Dst < -100$ nT [*Miyoshi and Kataoka*, 2005]. During SBZ-fast events, the flux recovers to the pre-SIR level within one day after the stream interface, and the enhanced
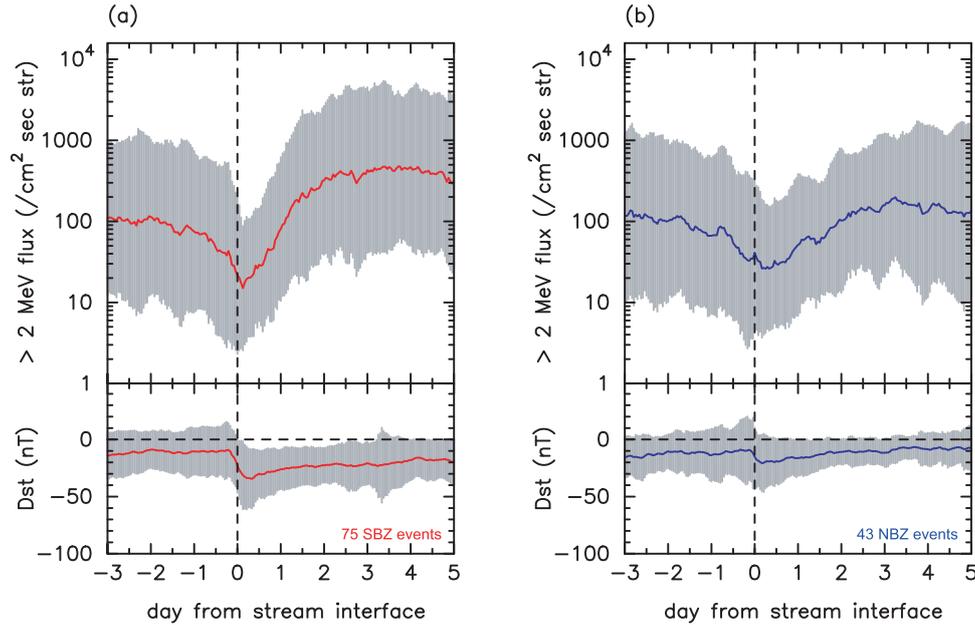
Figure 2. Superposed time series of the logarithm of the flux of >2.0 MeV electrons as observed by the GOES satellites for (a) SBZ and (b) NBZ events. The bottom plots show the superposed *Dst* index. The gray region indicates the standard deviation.

flux is about ten times of the pre-SIR level at $t = 1.0$ day. In contrast, the flux recovery takes a longer time in SBZ-slow events, where the flux recovers to the pre-SIR level at $t = 2.0$ day, and the increase after $t = 2.0$ day is relatively weak. In SBZ-fast event, the flux increase takes place so quickly at $L = 4.5$ and then the region of the flux increase spreads over the outer belt as shown in Figure 6a. The flux increases more than the pre-SIR level within one day after the stream interface. Again, there exists the L shell dependence for the period of the flux recovery and enhancement: the flux enhancement at the lower L shell is

faster than that at the higher L shell. In SBZ-fast events, after the recovery to the pre-SIR level, the flux continues to increase and the amount of enhancements is large. On the other hand, in SBZ-slow events, again the flux recovers to the pre-SIR level at $t = 1.0$ day but the flux enhancement is significantly weak.

[16] The above results are summarized in Figure 7, showing a scatter diagram of the daily maximum flux of geosynchronous orbit at $t = 3.0$ day for all events. Figure 7a shows the SIR events that *Vsw* is faster than 500 km/s, and Figure 7b shows the SIR events that *Vsw* is slower than
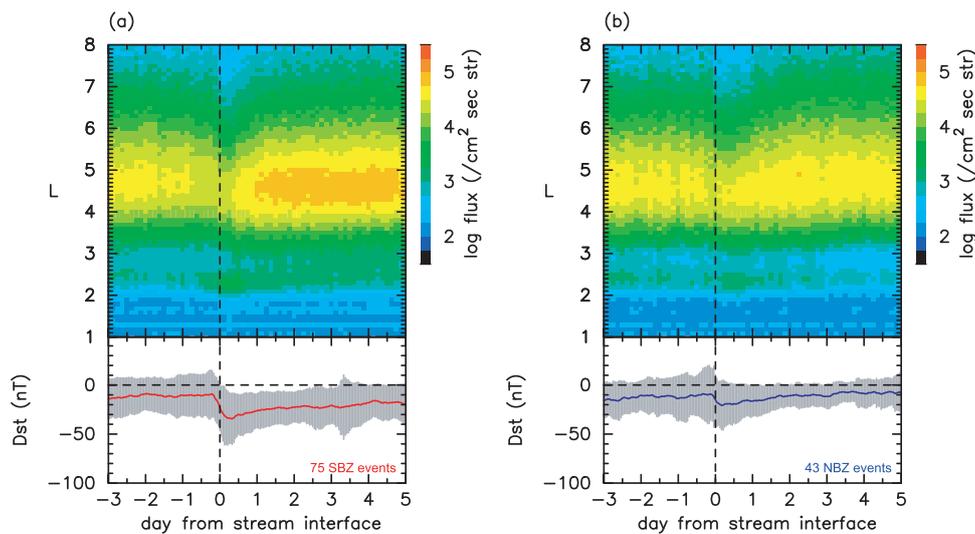
Figure 3. Superposed *L-t* diagram of >300 keV electrons as observed by the NOAA/POES satellites for (a) SBZ and (b) NBZ events. Average values after taking the logarithm of the flux are color coded. The bottom plots show the superposed *Dst* index. The gray region indicates the standard deviation.
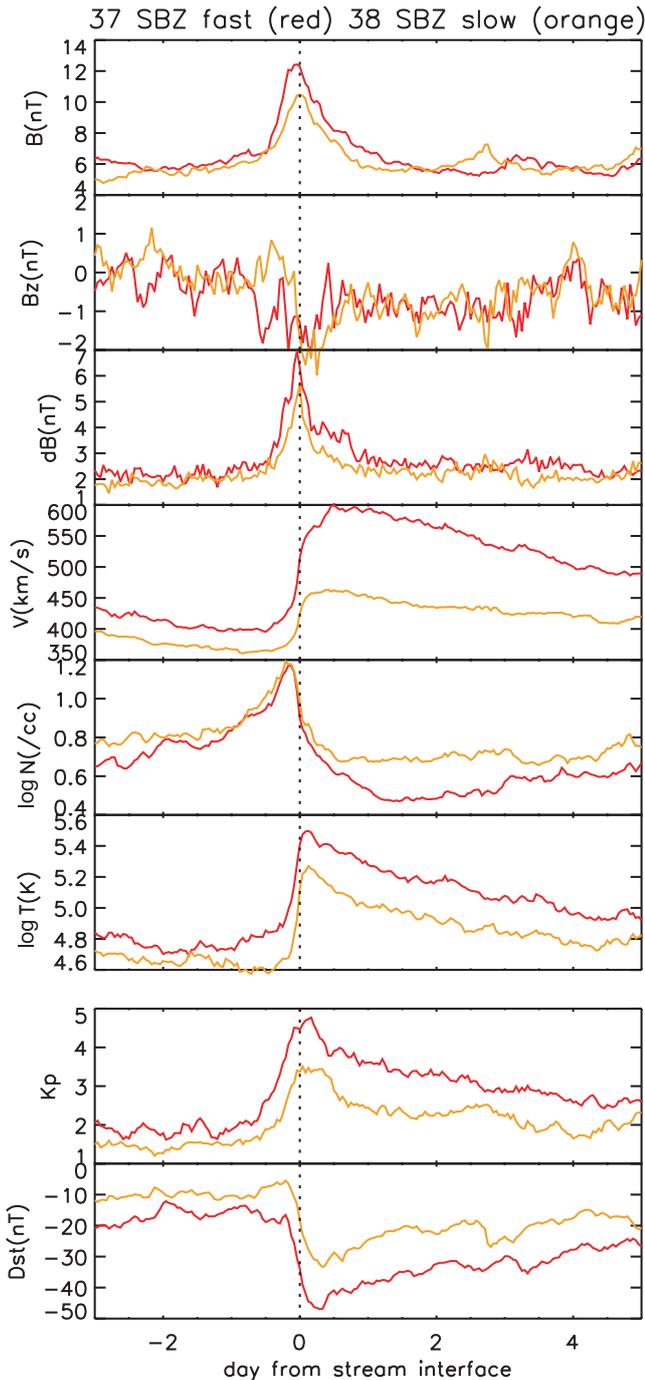
Figure 4. Superposed time series of solar wind parameters and geomagnetic activities. SBZ-fast and SBZ-slow events are shown in red and orange, respectively (see text for the categorization in detail). The format is the same as Figure 1.

500 km/s. It is clearly demonstrated that the flux enhancement in the outer belt depends on the STFA rule. Comparing Figures 7a and 7b, the dependence of solar wind speed is also apparent. The greatest flux enhancement is achieved in a high-speed coronal hole stream with a southward IMF offset of a few nT.

[17] Before finishing to show the results, from a different point of view, it may also be interesting to see the correla-

tion between fluxes of pre-SIR and post-SIR considering the STFA rule and solar wind speed because it has been reported that prestorm and poststorm fluxes are highly uncorrelated [*Reeves et al.*, 2003]. Figure 8 shows the maximum >2 MeV electron flux at geosynchronous orbit for 72 hours after the stream interface against the maximum flux for 72 hours before the stream interface. Figure 8a corresponds to events that *Vsw* is faster than 500 km/s, and Figure 8b is for events with *Vsw* below 500 km/s. Colors are the same as Figure 7. It is found that the correlation between the fluxes of pre-SIR and post-SIR depends on the solar wind speed and the STFA rule. Correlation coefficients for SBZ, NBZ, and other SIR events are 0.36, 0.77, 0.42 for fast stream events as shown in Figure 8a, and 0.44, 0.82, and 0.86 for slow stream events as shown in Figure 8b, respectively. The correlations of fast stream events are basically smaller than that of slow stream events. The smallest correlation (0.36) is seen in the SBZ-fast events; the increased fluxes are almost independent of the pre-SIR fluxes. The highest correlation (0.77) in the fast stream events is seen in NBZ events, indicating that the flux after the SIR does not largely increase or decrease relative to the pre-SIR level. In the slow stream events, the high-correlation coefficients (more than 0.8) are found in NBZ events and other SIR events; flux variations are small in the events between before and after the SIR. This may suggest that the slow stream without the southward IMF offset do not work for the net flux variation at geosynchronous orbit. The correlations of other SIR events for fast and slow stream events are natural, regarding the events as somewhere between SBZ and NBZ events.

4. Discussion

[18] The main result of this paper is that the southward IMF as well as the solar wind speed controls the flux enhancement of the outer belt in the SIR events. The most effective solar wind stream to produce the large flux enhancement is the highest-speed stream with the southward offset of IMF because of the STFA rule. The dependence on the STFA rule itself is not a new finding. Similar results about the outer belt flux variation caused by the STFA rule were reported by *McPherron* [2006], but they did not investigate the solar wind speed dependence. Although the excellent correlation for the solar wind speed is consistent with past studies [*Paulikas and Blake*, 1979; *Baker et al.*, 1986; *O'Brien et al.*, 2001; *Weigel et al.*, 2003], the result of this study indicates for the first time that only the solar wind speed by itself is not a sufficient condition for the large acceleration.

[19] The significant role of the STFA rule for the flux enhancement would be reasonable because the small amplitude southward IMF within a coronal hole stream, which cannot be a driver for the storm main phase, can be a driver for continuous substorms and enhanced convections, so-called HILDCAAs [*Tsurutani and Gonzalez*, 1987; *Tsurutani et al.*, 2006]. The southward IMF offset is a preferable condition for substorms and enhanced convections, while the northward IMF offset suppresses such activities in the magnetosphere. There have been several reports on the effect of substorms and enhanced convections for the flux enhancement in the outer belt through the nonadiabatic acceleration
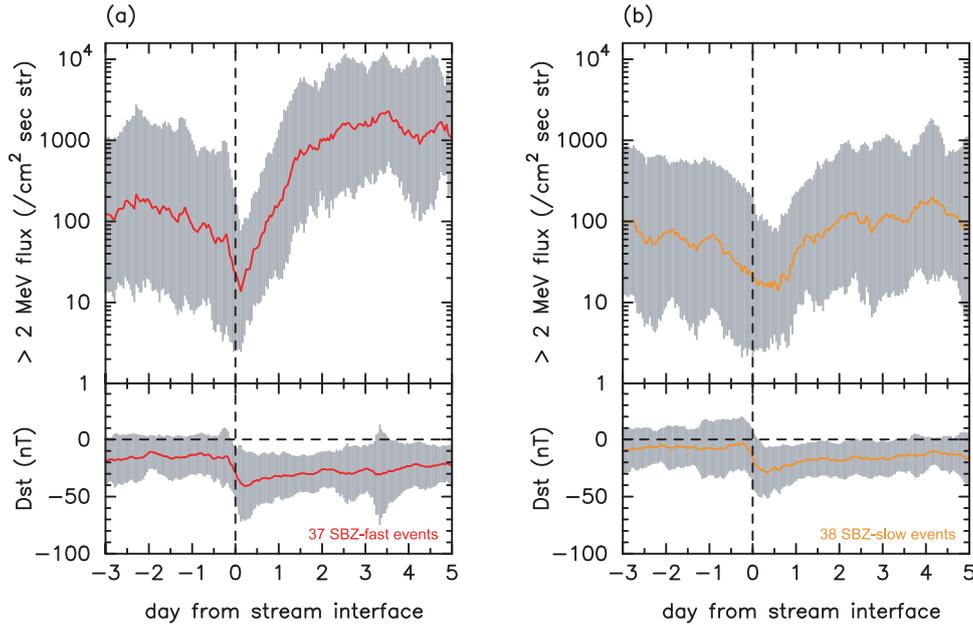
Figure 5. Superposed time series of the logarithm of the flux of >2.0 MeV electrons as observed by the GOES satellites for (a) SBZ-fast and (b) SBZ-slow events. The format is the same as Figure 2.

[*Meredith et al.*, 2002; *Miyoshi et al.*, 2003, 2007; *Lyons et al.*, 2005]. *Meredith et al.* [2002] suggested that the gradual acceleration of electrons during storms can be effective when there are periods of prolonged substorm activity during the storm recovery phase. *Lyons et al.* [2005] claimed that the enhanced convection leads to the enhanced dawnside chorus waves which accelerate the relativistic electrons. Further, they suggested that the large-amplitude IMF B_z fluctuation which has a large southward component and large IMF $|B_y|$ with a high solar wind speed are expected to lead to enhanced convection. *Miyoshi et al.* [2007] examined two magnetic storms driven by typical SIRs occurred in November 1993, in which the IMF sector polarities were different from each other: the outer belt electrons strongly increased in one storm,

but did not increase in another storm. Dynamics of seed electron population and chorus waves were quite different between two storms because of different activities of substorms and enhanced convection, and these differences produced the different evolution of the outer belt.

[20] The results of superposed epoch analysis of the NOAA data (Figures 3 and 6) showed that flux enhancements during the events take place at $L > 4$. That is, the SIR generally causes the flux enhancement at the outer portion and is less effective in the flux variation at the inner portion such as the slot region. Since the *Dst* index during the SIR events is greater than -50 nT on average, this is consistent with the statistical result between the peak L shell and the *Dst* index [*O'Brien et al.*, 2003]. It is worthwhile to note
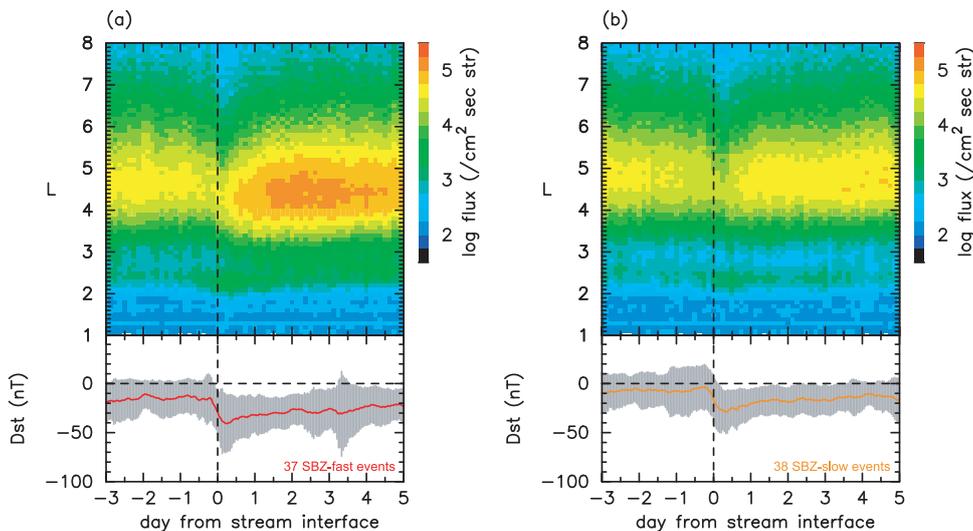
Figure 6. Superposed L-t diagram of >300 keV electrons as observed by the NOAA/POES satellites for (a) SBZ-fast and (b) SBZ-slow events. The format is the same as Figure 3.
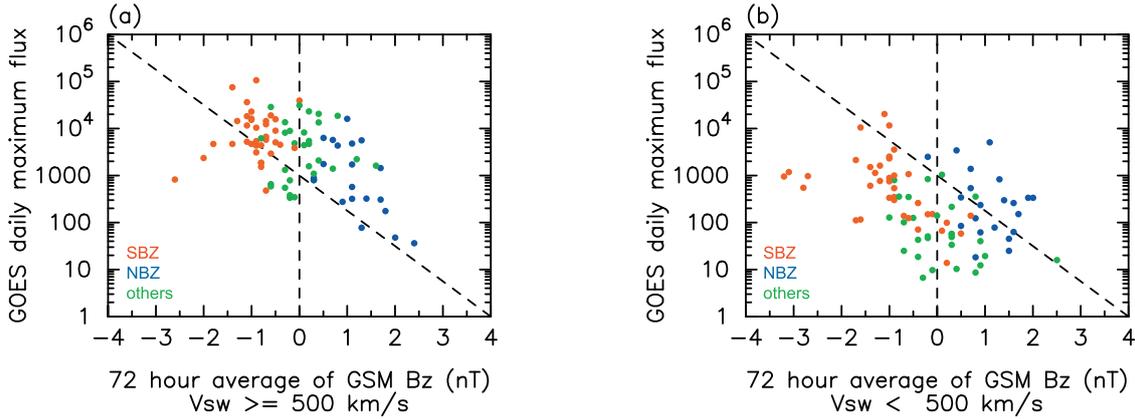
Figure 7. Scatter diagram of the daily maximum flux of geosynchronous orbit at $t = 3.0$ days for SIR events from 1994 to 2005. Horizontal axis is the 72 hour average of GSM B_z after the stream interface. Shown are (a) fast and (b) slow SIR events (see text for more detail). Red and blue circles correspond to SBZ and NBZ events, and green circles show other SIR events.

that the solar wind driver dependences of peak L shell control the solar cycle variation of the outer belt. Since the number of SIRs with fast coronal hole streams increases during the solar declining phase, it is expected that the flux at the outer portion of the outer belt increases largely during the declining phase. Moreover, CME-driven great storms with average minimum Dst of -200 nT cause the flux enhancement at the inner portion [*Miyoshi and Kataoka*, 2005] and the number of CME-driven great storms increases during the solar active period [e.g., *Kataoka and Miyoshi*, 2006], so it is also expected that the flux at the inner portion increases during the solar maximum. This is

consistent with the observational results: the outer portion of the outer belt develops during the solar declining phase, while the inner portion develops during the solar active period [*Miyoshi et al.*, 2004].

[21] Here we discuss the relationship between relativistic electron flux enhancements and magnetic storm amplitudes. The statistical examination showed that the flux enhancement of the outer belt does not depend on the storm amplitude [*Reeves et al.*, 2003]. Recently, *Kim et al.* [2006] showed that the average of minimum Dst indices for which the relativistic electron flux increased is above -50 nT, and they are normally classified as weak magnetic storms or
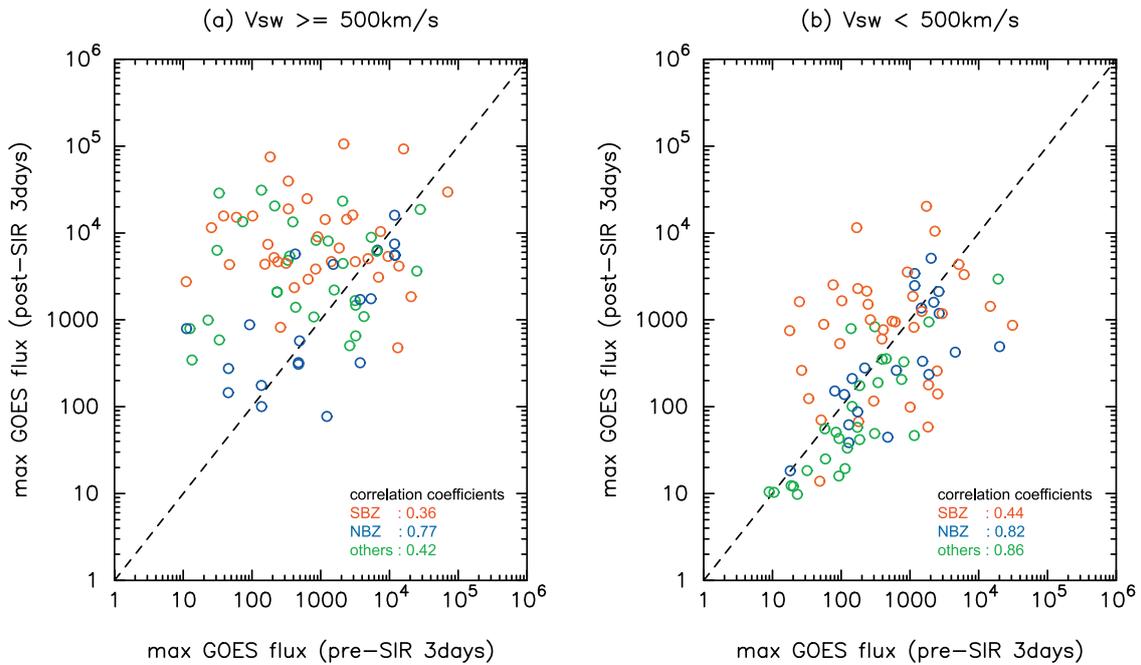
Figure 8. Flux of geosynchronous orbit for 72 hours before and after the stream interface. Horizontal axis is the pre-SIR flux, and vertical axis is the post-SIR flux. Shown are (a) fast and (b) slow SIR events (see text for more detail). Red and blue circles correspond to SBZ and NBZ events, and green circles show other SIR events.

absence of magnetic storms. This is consistent with our results as shown in Figure 5a: the superposed *Dst* index for SBZ-fast events, in which the greatest flux enhancement in the outer belt was observed, is greater than −50 nT. Since the period of a high-speed stream that produces the flux enhancement typically corresponds to the storm recovery phase, the flux enhancement is independent of the storm main phase which is driven by different structure such as CMEs or SIRs. *McPherron and Weygand* [2006] also pointed out that different solar wind structures relative to a stream interface of SIRs are important for the flux enhancements. Thus intense magnetic storms are not essential for the flux enhancement in the outer belt. This is consistent with discussions of *Meredith et al.* [2002]. Moreover, time delay between the storm main phase and the flux enhancement of the outer belt is readily understood from this consideration.

[22] The *Dst* effect should be important for the loss of electrons during the main phase [*Kim and Chan*, 1997]. The adiabatic *Dst* effect, however, can be negligible because in the superposed epoch analysis in this study (e.g., Figure 2), the loss of the outer belt were observed around the SIR when the *Dst* index did not show significant evolution. As shown by *Green et al.* [2004] and *Kim et al.* [2006], when relativistic electrons decrease at geosynchronous orbit, the ambient magnetic field is less dipolar, implying that a distortion of the ambient magnetic field causes a net reduction in the outer belt. The result in this study suggests that the SIR is one of the interplanetary loss drivers, probably leading to nonadiabatic loss due to the distortion of the magnetic field. This is consistent with the report of *Onsager et al.* [2002] showing that several nonadiabatic flux decreases at geosynchronous orbit are correlated with discontinuities in the solar wind. Note that the loss associated with the SIR during weak magnetic storms or absence of magnetic storms takes place only around the outer portion at the outer belt, and the inner portion is less sensitive. This suggests that the field distortion associated with the SIR would be typically limited at the outer portion.

[23] The small correlation between pre-SIR and post-SIR fluxes in SBZ events in Figure 8 results from resetting the outer belt by rapid loss during the storm main phase and subsequent large enhancements during the recovery phase. One of the mechanisms for the rapid loss of the outer belt in SBZ events (Figure 5a) may be caused by enhanced pitch angle scattering within plasmapause drainage [*Borovsky and Steinberg*, 2006b]. According to the paper by *Borovsky and Steinberg* [2006b], SBZ events are likely to produce so-called "calm" just prior to the SIR and "calm" stimulates the formation of plasmapause drainage by the preconditioning effect. The high correlation found in NBZ events is also interesting. During NBZ events, the outer belt electrons show slow decrease and recovery to the pre-SIR level, suggesting that the loss balances with the acceleration. Although this high correlation in NBZ events may not be consistent with *Reeves et al.* [2003], we have to note that their statistical correlation analysis does not involve most of our NBZ events because the *Dst* minima in NBZ events are somewhat large, and hard to detect the events as storms.

[24] The obtained fact that the flux enhancement of the outer belt is controlled by both the solar wind speed and the STFA rule would be important for application of the space weather forecast. One can predict which coronal hole stream produces the large flux enhancement by looking at the solar wind speed and the IMF sector polarity. In fact, the diagrams for the probabilistic forecast are presented by *Miyoshi and Kataoka* [2008], which predict probability of the flux enhancement associated with the SIR events on the basis of the STFA rule and the solar wind speed. Furthermore, the results shown in this paper can give a possible observational restriction to determine the acceleration mechanism of the energetic electrons in the outer belt, and can contribute as a basic research to future inner magetospheric satellite missions such as RBSP (US), ORBITALS (Canada), and ERG (Japan).

5. Conclusion

[25] We evaluated the solar wind parameter dependence of the outer belt to find that both the IMF B_z and solar wind speed are important for the large flux enhancements in SIR events. The STFA rule plays a significant role to control the flux enhancement via the IMF B_z offset of large-amplitude IMF fluctuations in a high-speed stream, that is, only the solar wind speed by itself is not a sufficient condition for the large acceleration. The largest flux enhancement is found in the highest-speed streams with a southward IMF offset. The strong flux enhancements tend to be associated with weak geomagnetic storms with minimum *Dst* of about −50 nT on average.

References

Baker, D. N., R. W. Klebesadel, P. R. Higbie, and J. B. Blake (1986), Highly relativistic electrons in the Earth's outer magnetosphere: 1. Lifetimes and temporal history 1979–1984, *J. Geophys. Res.*, *91*, 4265–4276.

Baker, D. N., R. D. Belian, P. R. Higbie, R. W. Kebesadel, and J. B. Blake (1987), Deep dielectric charging effects due to high energy electrons in the Earth's outer magnetosphere, *J. Electrost.*, *20*, 3–19.

Baker, D. N., J. H. Allen, S. G. Kanekal, and G. D. Reeves (1998), Disturbed space environment may have been related to pager satellite failure, *Eos Trans. AGU*, *79*, 477.

Blake, J. B., D. N. Baker, N. Turner, K. W. Ogilvie, and R. P. Lepping, (1997), Correlation of changes in the outer-zone relativistic-electron population with upstream solar wind and magnetic field measurements, *Geophys. Res. Lett.*, *24*, 927–929.

Borovsky, J. E., and J. T. Steinberg (2006a), The freestream turbulence effect in solar-wind/magnetosphere coupling: Analysis through the solar cycle and for various types of solar wind, in *Recurrent Magnetic Storms: Corotating Solar Wind Streams*, *Geophys. Monogr. Ser.*, vol. 167, edited by B. Tsurutani et al., pp. 59–76, AGU, Washington, D. C.

Borovsky, J. E., and J. T. Steinberg (2006b), The "calm before the storm" in CIR/magnetosphere interactions: Occurrence statistics, solar wind statistics, and magnetospheric preconditioning, *J. Geophys. Res.*, *111*, A07S10, doi:10.1029/2005JA011397.

Borovsky, J. E., M. F. Thomsen, and R. C. Elphic (1998), The driving of the plasma sheet by the solar wind, *J. Geophys. Res.*, *103*, 17,617–17,640.

Elkington, S. R., M. K. Hudson, and A. A. Chan (1999), Acceleration of relativistic electrons via drift resonant interactions with toroidal-mode Pc5 ULF oscillation, *Geophys. Res. Lett.*, *26*, 3273–3276.

Evans, D. S., and M. S. Greer (2000), Polar orbiting environmental satellite space environment monitor: 2. Instrument description and archive data documentation, *NOAA Tech. Memo. OAR SEC-93*, NOAA, Boulder, Colo.

Gosling, J. T., J. R. Asbridge, S. J. Bame, and W. C. Feldman (1978), Solar wind stream interface, *J. Geophys. Res.*, *83*, 1401–1412.

Green, J. C., T. G. Onsager, T. P. O'Brien, and D. N. Baker (2004), Testing loss mechanisms capable of rapidly depleting relativistic electron flux in the Earth's outer radiation belt, *J. Geophys. Res.*, *109*, A12211, doi:10.1029/2004JA010579.

Horne, R. B., N. P. Meredith, S. A. Glauert, A. Varotsou, D. Boscher, R. M. Thorne, Y. Y. Shprits, and R. R. Anderson (2006), Mechanisms for the acceleration of radiation belt electrons, in *Recurrent Magnetic Storms: Corotating Solar Wind Streams*, *Geophys. Monogr. Ser.*, vol. 167, edited by B. Tsurutani et al., pp. 151–174, AGU, Washington, D. C.

Iles, R. H. A., A. N. Fazakerley, A. D. Johnstone, N. P. Meredith, and P. Buhler (2002), The relativistic electron response in the outer radiation belt during magnetic storms, *Ann. Geophys.*, *20*, 957–965.

Kataoka, R., and Y. Miyoshi (2006), Flux enhancement of radiation belt electrons during geomagnetic storms driven by coronal mass ejections and corotating interaction regions, *Space Weather*, *4*, S09004, doi:10.1029/2005SW000211.

Kim, H.-J., and A. A. Chan (1997), Fully adiabatic changes in storm time relativistic electron fluxes, *J. Geophys. Res.*, *102*, 22,107–22,116.

Kim, H.-J., K. C. Kim, D.-Y. Lee, and G. Rostoker (2006), Origin of geosynchronous relativistic electron events, *J. Geophys. Res.*, *111*, A03208, doi:10.1029/2005JA011469.

Li, X. (2006), The role of radial transport in accelerating radiation belt electrons, in *Recurrent Magnetic Storms: Corotating Solar Wind Streams*, *Geophys. Monogr. Ser.*, vol. 167, edited by B. Tsurutani et al., pp. 139–174, AGU, Washington, D. C.

Lyons, L. R., D.-Y. Lee, R. M. Thorne, R. B. Horne, and A. J. Smith (2005), Solar wind–magnetosphere coupling leading to relativistic electron energization during high-speed streams, *J. Geophys. Res.*, *110*, A11202, doi:10.1029/2005JA011254.

McPherron, R. L. (2006), Statistical properties of the solar wind and IMF at 1 AU, *Eos Trans., AGU*, *87*(52), Fall Meet. Suppl., Abstract SH43C-01.

McPherron, R. L., and J. Weygand (2006), The solar wind and geomagnetic activity as a function of time relative to corotating interaction regions, in *Recurrent Magnetic Storms: Corotating Solar Wind Streams*, *Geophys. Monogr. Ser.*, vol. 167, edited by B. Tsurutani et al., pp. 125–138, AGU, Washington, D. C.

Meredith, N. P., R. B. Horne, and R. R. Anderson (2001), Substorm dependence of chorus amplitudes: Implications for the acceleration of electrons to relativistic energies, *J. Geophys. Res.*, *106*, 13,165–13,178.

Meredith, N. P., R. B. Horne, R. H. A. Iles, R. M. Thorne, D. Heynderickx, and R. R. Anderson (2002), Outer zone relativistic electron acceleration associated with substorm-enhanced whistler mode chorus, *J. Geophys. Res.*, *107*(A7), 1144, doi:10.1029/2001JA900146.

Miyoshi, Y., and R. Kataoka (2005), Ring current ions and radiation belt electrons during geomagnetic storms driven by coronal mass ejections and corotating interaction regions, *Geophys. Res. Lett.*, *32*, L21105, doi:10.1029/2005GL024590.

Miyoshi, Y., and R. Kataoka (2008), Probabilistic space weather forecast of the relativistic electron flux enhancement at geosynchronous orbit, *J. Atmos. Sol. Terr. Phys.*, in press.

Miyoshi, Y., A. Morioka, H. Misawa, T. Obara, T. Nagai, and Y. Kasahara (2003), Rebuilding process of the outer radiation belt during the 3 November 1993 magnetic storm: NOAA and Exos-D observations, *J. Geophys. Res.*, *108*(A1), 1004, doi:10.1029/2001JA007542.

Miyoshi, Y. S., V. K. Jordanova, A. Morioka, and D. S. Evans (2004), Solar cycle variations of the electron radiation belts: Observations and radial diffusion simulation, *Space Weather*, *2*, S10S02, doi:10.1029/2004SW000070.

Miyoshi, Y., A. Morioka, R. Kataoka, Y. Kasahara, and T. Mukai (2007), Evolution of the outer radiation belt during the November 1993 storms driven by corotating interaction regions, *J. Geophys. Res.*, *112*, A05210, doi:10.1029/2006JA012148.

O'Brien, T. P., R. L. McPherron, D. Sornette, G. D. Reeves, R. Friedel, and H. J. Singer (2001), Which magnetic storms produce relativistic electrons at geosynchronous orbit?, *J. Geophys. Res.*, *106*, 15,533–15,544.

O'Brien, T. P., K. R. Lorentzen, I. R. Mann, N. P. Meredith, J. B. Blake, J. F. Fennell, M. D. Looper, D. K. Milling, and R. R. Anderson (2003), Energization of relativistic electrons in the presence of ULF power and MeV microbursts: Evidence for dual ULF and VLF acceleration, *J. Geophys. Res.*, *108*(A8), 1329, doi:10.1029/2002JA009784.

Onsager, T. G., G. Rostoker, H.-J. Kim, G. D. Reeves, T. Obara, H. J. Singer, and C. Smithtro (2002), Radiation belt electron flux dropouts: Local time, radial, and particle-energy dependence, *J. Geophys. Res.*, *107*(A11), 1382, doi:10.1029/2001JA000187.

Paulikas, G. A., and J. B. Blake (1979), Effects of the solar wind on magnetospheric dynamics: Energetic electrons at the synchronous orbit, in *Quantitative Modeling of Magnetospheric Processes*, *Geophys. Monogr. Ser.*, vol. 21, edited by W. P. Olson, pp. 180–202, AGU, Washington D. C.

Reeves, G. D., K. L. McAdams, R. H. W. Friedel, and T. P. O'Brien (2003), Acceleration and loss of relativistic electrons during geomagnetic storms, *Geophys. Res. Lett.*, *30*(10), 1529, doi:10.1029/2002GL016513.

Richardson, I. G., et al. (2006), Major geomagnetic storms ($Dst \leq -100$ nT) generated by corotating interaction regions, *J. Geophys. Res.*, *111*, A07S09, doi:10.1029/2005JA011476.

Russell, C. T., and R. L. McPherron (1973), Semiannual variation of geomagnetic activity, *J. Geophys. Res.*, *78*, 92–108.

Schulz, M., and L. Lanzerotti (1974), *Particle Diffusion in the Radiation Belts*, Springer, New York.

Summers, D., R. M. Thorne, and F. Xiao (1998), Relativistic theory of wave-particle resonant diffusion with application to electron acceleration in the magnetosphere, *J. Geophys. Res.*, *103*, 20,487–20,500.

Tsurutani, B. T., and W. D. Gonzalez (1987), The cause of high-intensity long-duration continuous AE activity (HILDCAAs): Interplanetary Alfvén wave trains, *Planet. Space Sci.*, *35*(4), 305–412.

Tsurutani, B. T., et al. (2006), Corotating solar wind streams and recurrent geomagnetic activity: A review, *J. Geophys. Res.*, *111*, A07S01, doi:10.1029/2005JA011273.

Weigel, R. S., A. J. Klimas, and D. Vassiliadis (2003), Precursor analysis and prediction of large-amplitude relativistic electron fluxes, *Space Weather*, *1*(3), 1014, doi:10.1029/2003SW000023.

————————

R. Kataoka, Computational Astrophysics Laboratory, Institute of Physics and Chemical Research (RIKEN), Saitama 351-0198, Japan. (ryuho@riken.jp)

Y. Miyoshi, Solar-Terrestrial Environment Laboratory, Nagoya University, Nagoya 464-8601, Japan. (miyoshi@stelab.nagoya-u.ac.jp)

A quantitative MHD study of the relation among arcade shearing, flux rope formation, and eruption due to the tearing instability

D. Shiota,[1] K. Kusano,[2] T. Miyoshi,[3] N. Nishikawa,[4] and K. Shibata[5]

[1] The quantitative relationship between the magnetohydrodynamic (MHD) activity of solar coronal arcade and the magnetic helicity injection, which is caused by shearing motion, has been investigated, using azimuthally symmetric model of MHD simulation. We have calculated several cases in which the width of the shearing region is varied and examined the relationship between the magnetic arcade dynamics and magnetic helicity evolution. As a result, it is found that as the shearing motion is imposed on narrower regions along each side of the magnetic inversion line, the magnetic arcade can be easily destabilized by the resistive tearing mode. However, in this case, even though reconnection driven by the tearing mode produces plasmoids, the plasmoid elevation is almost in proportion to the total amount of magnetic helicity contained in the arcade, and it is too slow to explain the trigger process of coronal mass ejections (CMEs). On the other hand, in the case where the shearing motion is imposed on the entire region, much larger magnetic helicity injection is required to injected arcade in order to destabilize the system, compared to practical helicity injection measured in the solar corona. The results suggest that it may be difficult to trigger a CME just by the axisymmetric shearing motion and that some other mechanisms should be involved in the triggering process of a CME. The results also imply that the relation between the magnetic helicity and the overlying magnetic flux can be a key parameter for the CME occurrence.

Citation: Shiota, D., K. Kusano, T. Miyoshi, N. Nishikawa, and K. Shibata (2008), A quantitative MHD study of the relation among arcade shearing, flux rope formation, and eruption due to the tearing instability, *J. Geophys. Res.*, *113*, A03S05, doi:10.1029/2007JA012516.

1. Introduction

[2] Coronal mass ejections (CMEs) are the most explosive phenomena in the solar corona, whose triggering mechanism is one of the major problems in the solar physics. In addition, it is also important for the understanding of CMEs and space weather forecasting to make clear a kind of critical condition to determine whether CMEs can be launched as the result of eruptions of coronal fields.

[3] CMEs are defined as ejections of a large amount of mass within magnetic flux, which are produced by eruptions of the coronal magnetic field. In many observations with coronagraphs, CMEs have been observed to have fine complicated density structures which are thought to be traces of magnetic field lines due to freezing of the plasma to the magnetic field. Usual structures in CMEs are helical structures, i.e., twisted magnetic flux ropes, and in some of them there are dense cores corresponding to erupted chromospheric prominences (see *Hundhausen* [1999] for review). This results suggested that the magnetic structures in prominences are closely related to the CME initial processes. Furthermore, because twisted magnetic flux ropes represent current carrying magnetic flux, they are thought to be the most essential features of CMEs.

[4] Since the discovery of CMEs in the early 1970s, many attempts to understand the mechanism of CMEs theoretically or numerically have been made by many solar physicists. There are some reviews for such historical works [*Forbes*, 2000; *Priest and Forbes*, 2002; *Lin et al.*, 2003; *Forbes et al.*, 2006], and we do not duplicate them here. Most of theoretical works about CMEs since the 1990s can be roughly classified into two major categories: flux rope models and sheared arcade models.

[5] Studies with the former models focused only on the energy release process assuming initial conditions including a twisted magnetic flux rope [*Forbes and Priest*, 1995; *Lin and Forbes*, 2000; *Chen and Shibata*, 2000; *Kliem and Torok*, 2005]. These works succeeded in explaining the rapid energy release and acceleration in the initial phase of CMEs due to loss of equilibrium or stability of the initial

---

[1]Center for Computational Astrophysics, National Astronomical Observatory of Japan, Tokyo, Japan.
[2]Earth Simulator Center, Japan Agency for Marine-Earth Science and Technology, Kanagawa, Japan.
[3]Facility of Science, Hiroshima University, Hiroshima, Japan.
[4]Super Computer System Planning and Operations Department, Japan Agency for Marine-Earth Science and Technology, Kanagawa, Japan.
[5]Kwasan and Hida Observatories, Kyoto University, Kyoto, Japan.

conditions containing high magnetic energy but do not discuss how such the initial conditions can be formed.

[6] In contrast, the other group studies the shearing of the coronal potential magnetic arcades whose magnetic energy is minimum [*Mikic et al.*, 1988; *Steinolfson*, 1991; *Inhester et al.*, 1992; *Mikic and Linker*, 1994; *Linker and Mikic*, 1995; *Kusano et al.*, 1995; *Choe and Lee*, 1996; *Antiochos et al.*, 1999; *Amari et al.*, 2003; *MacNeice et al.*, 2004]. These works succeed in explaining the energy buildup and release processes with slow velocity on the lower boundary, starting from initial minimum energy conditions. In these models, twisted flux ropes are produced as a result of magnetic reconnection in the sheared arcades, and the flux ropes are ejected without remaining in equilibrium states as with the initial conditions of the flux rope models.

[7] Because the magnetic configuration of each model is too simple, these results show no continuity between these scenarios. To avoid the incompleteness due to the simplification, recent studies tend to simulate as realistically as possible. *Manchester et al.* [2004] performed MHD simulation of magnetic buoyant processes and the initiation of eruptions starting from the initial condition where the twisted magnetic flux rope embedded below the photosphere. Their results show that the emergence of the twisted magnetic flux ropes achieves shearing motion of the initial coronal arcades on the photosphere. Because of the stratification in the solar interior, the buoyant process proceeds with nonuniform speed due to mass accumulated in the concave part inside the flux rope. As a result, most of the poloidal flux has footpoints anchored to the surface and appears to bind the strongly sheared arcade corresponding to the axis of the original large-scale flux rope (also investigated by *Magara* [2006]). These result suggests that a strong sheared arcade embedded in a weakly or unsheared arcade, i.e., a magnetic configuration where its shear component is localized, could be a common magnetic configuration in an emerging flux region.

[8] On the other hands, the formation process of a coronal twisted flux rope has been studied with the model of magnetic field evolutions [*van Ballegooijen and Martens*, 1989; *MacKay and van Ballegooijen*, 2005, 2006a, 2006b]. Those studies suggests that the important process for the formation is "magnetic cancellation" of a sheared arcade which is achieved by the motion of opposite magnetic elements on the photosphere, toward the polarity inversion line. The cancellation process leads to formation of a current sheet and then reconnection, which results in the formation the flux rope. However, these authors did not include the dynamical effect of current sheet formation (inhibition of reconnection), and the effect on CME occurrence of the impulsive energy release by reconnection On the basis of these results, in this paper, we developed a classical arcade shearing model in a simple axisymmetric configuration to investigate the questions: how coronal twisted flux rope can be formed via magnetic reconnection and what is an appropriate condition for CME occurrence. Specifically, we use a single potential arcade across the equator and impose shearing motion on its bottom boundary.

[9] In this study, we focus on the relation between the condition for a CME and the following quantities: magnetic

helicity, magnetic flux, and electric resistivity. Magnetic helicity [*Berger and Field*, 1984] defined as

$$H = \int \mathbf{A} \cdot \mathbf{B} dV. \tag{1}$$

The magnetic helicity is an important quantity related to the topology of magnetic field and the free magnetic energy where **B** and **A** are magnetic field and the corresponding vector potential. In order to exclude their gauge arbitrary, we use relative helicity uniquely defined as

$$H_{rel} = \int \left(\mathbf{A} + \mathbf{A}_p \right) \cdot \left(\mathbf{B} - \mathbf{B}_p \right) dV, \tag{2}$$

where $\mathbf{B}_p$ and $\mathbf{A}_p$ are the potential magnetic field and the corresponding vector potential, whose details are described in section 2.2. It is widely accepted that a CME is the process in which magnetic helicity in the solar corona is expelled into interplanetary space, and therefore the existence of magnetic helicity is the necessary condition for a CME occurrence. However, how the minimum amount of helicity leads to a CME is not clear. Magnetic flux is another key quantity for determining whether an eruption in a sheared arcade can evolve into a CME or not. That is, if enough magnetic flux exists outside of the reconnecting arcade, it can confine and stabilize the ejected plasmoid (flux rope) by magnetic tension force. Therefore the amount of overlying flux is thought to be an essential quantity for a CME occurrence. Finally, resistivity is known as the essential quantity to determine the magnetic reconnection process and the growth rate of a tearing-mode instability. In this study we try to make clear the quantitative relations among these quantities and the condition for a CME occurrence.

[10] Here we briefly summarize the contexts of this paper. In the next section, we describe the detailed methodology of the numerical simulation. In section 3, the numerical results of different shear and different resistivity cases are shown. We discuss the quantitative relations suggested by the numerical results in section 4. Finally, we summarize this paper in section 5.

2. Numerical Model
2.1. Numerical Scheme

[11] In a numerical simulation performed in this work, we solve the following MHD equations using a finite volume method with the HLLD nonlinear Riemann solver [*Miyoshi and Kusano*, 2005], the third-order TVD MUSCL, and the Runge-Kutta time integration:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \tag{3}$$

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v} \mathbf{v} - \mathbf{B} \mathbf{B} + p_T \hat{\mathbf{e}}) = \rho \mathbf{g}, \tag{4}$$

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \cdot (\mathbf{B} \mathbf{v} - \mathbf{v} \mathbf{B} + \psi \hat{\mathbf{e}}) + \nabla \times (\eta \mathbf{J}) = 0, \tag{5}$$

$$\frac{\partial e}{\partial t} + \nabla \cdot [(e + p_T)\mathbf{v} - (\mathbf{v} \cdot \mathbf{B})\mathbf{B}] = 0, \tag{6}$$

$$\frac{\partial \psi}{\partial t} + c_p^2 \nabla \cdot \mathbf{B} + c_d \psi = 0, \tag{7}$$

where

$$\mathbf{J} = \nabla \times \mathbf{B}, \tag{8}$$

$$e = \frac{\rho \mathbf{v}^2}{2} + \frac{p}{\gamma - 1} + \frac{\mathbf{B}^2}{2} - \rho \Psi, \tag{9}$$

$$p_T = p + \frac{\mathbf{B}^2}{2}, \tag{10}$$

$$\hat{\mathbf{e}} = \begin{pmatrix} e_{x,r} & e_{y,r} & e_{z,r} \\ e_{x,\theta} & e_{y,\theta} & e_{z,\theta} \\ e_{x,\phi} & e_{y,\phi} & e_{z,\phi} \end{pmatrix} \tag{11}$$

$$G_0 = \frac{GM_\odot \rho_0}{R_\odot p_0}, \tag{12}$$

$$\mathbf{g} = -G_0 \mathbf{r}/r^3, \tag{13}$$

$$\Psi = -\frac{GM_\odot}{r}. \tag{14}$$

Note that we introduce a new variable ψ and the corresponding equation (7) in order to remove numerical $\nabla \cdot \mathbf{B}$ from the domain of interest. Detail of this method is described in the next paragraph. All other variables are defined as standard usages. The vector potential, which is used for the calculation of magnetic helicity, is also time integrated with the following equation:

$$\frac{\partial \mathbf{A}}{\partial t} = -\mathbf{E}, \tag{15}$$

where

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B} + \eta \mathbf{J}. \tag{16}$$

[12] Here we briefly introduce the divergence-free correction method [*Dedner et al.*, 2002]. Equations (4), (5), and (6) are derived assuming the Solenoidal condition of the magnetic field; i.e., these equations have the following neglected right-hand terms: $-(\nabla \cdot \mathbf{B})\mathbf{B}$, $-(\nabla \cdot \mathbf{B})\mathbf{v}$, and $-(\nabla \cdot \mathbf{B})(\mathbf{v} \cdot \mathbf{B})$, respectively. Generally, in the result of the time integration of equation (5) in a three-dimensional simulation, $\nabla \cdot \mathbf{B}$ cannot be kept to zero, and the resulting finite $\nabla \cdot \mathbf{B}$ generates numerical errors through the neglected terms above. Because an accumulation of the errors causes wrong solutions, it is very important to keep the magnitude of

numerical $\nabla \cdot \mathbf{B}$ small so it does not change the solution significantly. The essential point of the method of *Dedner et al.* [2002] is that introducing a new variable ψ and equation (7) with the divergence of equation (5)

$$\frac{\partial (\nabla \cdot \mathbf{B})}{\partial t} + \nabla^2 \psi = 0, \tag{17}$$

we can establish a hyperbolic and parabolic system about $\nabla \cdot \mathbf{B}$ and ψ. The coefficient c_p in equation (7) is the propagation speed of numerical $\nabla \cdot \mathbf{B}$, and the coefficient c_d is the diffusion coefficient of ψ which has finite value only near the outer boundary ($r > 3$). When nonzero values of $\nabla \cdot \mathbf{B}$ are generated somewhere in the numerical domain, this method distributes that value to the all direction as a wave, thus reducing its local magnitude. Furthermore, when the wave approaches the outer boundary, the amplitudes of both $\nabla \cdot \mathbf{B}$ and ψ reduces rapidly. On the other hand, the curl procedure of equation (5) vanishes the second term of the right-hand side, i.e., the term due to the correction method. Hence this method does not directly affect the evolution of the current density, which is significant in resistive MHD processes. This means that the correction method developed for ideal MHD process can be adopted to resistive MHD simulations.

[13] Parameters are chosen to realistic value: the solar radius $R_\odot = 6.998 \times 10^{10}$ cm for length, sound speed c_s for velocity, free-fall time $\tau_0 = R_\odot/c_s$ for time, and $B_0 = \sqrt{8\pi p_0}$ for magnetic field. When these values are chosen to be characteristic values of the solar corona, such as $n_0 = 1 \times 10^9$ cm^{-3} and $T_0 = 2$ MK with a specific heat ratio $\gamma = 1.02$, the normalization parameters become $c_s = 1.83 \times 10^2$ km s^{-1}, $\tau_0 = 3.81 \times 10^3$ s, and $B_0 = 3.72$ Gauss.

[14] Thermodynamic effects, such as coronal heating and radiative cooling, are neglected in this study, and we assume the specific heat ratio γ is set to be 1.02 for simplicity. The electric resistivity η is set to be uniform: zero from cases A to E, 10^{-4} for case F, 10^{-5} for case G, and 10^{-6} for case H. However, when the width of the current sheet becomes as small as the grid size, TVD limiter, installed in order to avoid numerical oscillations at discontinuities, behaves as anomalous resistivity which could lead to fast reconnection.

[15] The numerical domain is set to a sector in the spherical coordinate, i.e., ($R_{\min} < r < R_{\max}$, $0 < \theta < \pi$, $0 < \phi < 2\pi/N_d$) where $R_{\min} = 1$, $R_{\max} = 5$, and $N_d = 256$. The domain is discretized by (N_r, N_θ, N_ϕ) = (510, 255, 1) grid points which are distributed nonuniformly to enhance the resolution in the equatorial and lower regions. The radial and latitudinal location of grid points are determined so that its first and second derivatives of the grid coordinate become continuous between uniform and nonuniform distributions, as follows:

$$r(i) = \begin{cases} R_{\min} + dr_1 i & i \le i_c \\ R_{\min} + a_r i^3 + b_r i^2 + (c_r + dr_1)i + d_r & i > i_c \end{cases} \tag{18}$$

where $dr_1 = (R_{\max} - R_{\min})\varepsilon_r/N_r$, $0 < \varepsilon_r \le 1$, and

$$\begin{pmatrix} a_r \\ b_r \\ c_r \\ d_r \end{pmatrix} = \frac{R_{\max} - R_{\min} - dr_1 N_r}{(N_r - i_c)^3} \begin{pmatrix} 1 \\ -3i_c \\ 3i_c^2 \\ -i_c^3 \end{pmatrix}. \tag{19}$$

$$
\theta(j) = \begin{cases} \pi/2 + d\theta_1 j' - a_\theta\{\cos[b_\theta(j'+j_c)] - 1\} \\ \quad -c_\theta(j'+j_c)^2 & (-j_h \leq j' < -j_c) \\ \pi/2 + d\theta_1 j' & (-j_c \leq j' \leq j_c) \\ \pi/2 + d\theta_1 j' + a_\theta\{\cos[b_\theta(j'-j_c)] - 1\} \\ \quad +c_\theta(j'-j_c)^2 & (j_c < j' \leq j_h), \end{cases}
$$

(20)

where $j' = j - j_h$, $j_h = (N_\theta + 1)/2$, $d\theta_1 = \pi\varepsilon_\theta/(N_\theta + 1)$, $0 < \varepsilon_\theta \leq 1$, and

$$
\begin{pmatrix} a_\theta \\ b_\theta \\ c_\theta \end{pmatrix} = \begin{pmatrix} (\pi/2 - d\theta_1 j_h)/2\pi^2 \\ 2\pi/(N_\theta - j_c - j_h) \\ (\pi/2 - d\theta_1 j_h)/(N_\theta - j_c - j_h)^2 \end{pmatrix}.
$$

(21)

Parameters $(i_c, j_c, \varepsilon_r, \varepsilon_\theta)$ are set to be $(0.4\,N_r, 0.0625\,N_\theta, 0.125, 0.0625)$, so that minimum and maximum grid size are $(dr_{min}, dr_{max}, d\theta_{min}, d\theta_{max}) = (9.8 \times 10^{-4}, 3.5 \times 10^{-2}, \pi/4096, 2.5 \times 10^{-2})$.

2.2. Initial Condition

[16] Pressure, velocity, and density throughout the numerical domain are determined by assuming hydrostatic equilibrium at uniform temperature. Potential magnetic field is obtained from a boundary value problem of Laplace equation for Φ [*Altschuler and Newkirk*, 1969], defined as follows:

$$
\nabla \cdot \mathbf{B_p} = \nabla \cdot (-\nabla\Phi) = -\nabla^2\Phi = 0.
$$

(22)

With the outer boundary at infinity, the solution is determined by the boundary condition at the solar surface ($r = 1$). When the normal component on the surface $B_{p,r0}$ is given as Neumann condition of Φ, we get the solution

$$
\Phi = \sum_{n=1}^{N} \sum_{m=0}^{n} \left\{ \left(\frac{1}{r}\right)^{n+1} (g_n^m \cos\phi + h_n^m \sin\phi)P_n^m(\cos\theta) \right\},
$$

(23)

where $P_n^m(\cos\theta)$ are associated Legendre polynomials, and g_n^m and h_n^m are coefficients obtained by spherical harmonics analysis. Then we get the potential magnetic field as,

$$
B_r = -\sum_{n=1}^{N} \sum_{m=0}^{n}
$$
$$
\cdot \left\{ (n+1)\left(\frac{1}{r}\right)^{n+2} (g_n^m \cos\phi + h_n^m \sin\phi)P_n^m(\cos\theta) \right\},
$$

(24)

$$
B_\theta = \sum_{n=1}^{N} \sum_{m=0}^{n} \left\{ \left(\frac{1}{r}\right)^{n+2} (g_n^m \cos\phi + h_n^m \sin\phi)\frac{dP_n^m(\cos\theta)}{d\theta} \right\},
$$

(25)

$$
B_\phi = -\sum_{n=1}^{N} \sum_{m=0}^{n} \left\{ m\left(\frac{1}{r}\right)^{n+2} (g_n^m \sin\phi - h_n^m \cos\phi)\frac{P_n^m(\cos\theta)}{\sin\theta} \right\}.
$$

(26)

[17] Once $\mathbf{B}_p$ is obtained, using the gauge as $A_{p,r} = 0$, we can get the vector potential $\mathbf{A}_p$ by integration of $\mathbf{B}_p = \nabla \times \mathbf{A}_p$. Assuming $\lim_{r\to\infty}\mathbf{A}_p = 0$, we get a simple solution for the vector potential:

$$
A_{p,\theta} = \sum_{n=1}^{N} \sum_{m=0}^{n} \left\{ \frac{m}{n}\left(\frac{1}{r}\right)^{n+1} (g_n^m \sin\phi - h_n^m \cos\phi)\frac{P_n^m(\cos\theta)}{\sin\theta} \right\}
$$

(27)

$$
A_{p,\phi} = -\sum_{n=1}^{N} \sum_{m=0}^{n} \left\{ \frac{1}{n}\left(\frac{1}{r}\right)^{n+1} (g_n^m \cos\phi + h_n^m \sin\phi)\frac{dP_n^m(\cos\theta)}{d\theta} \right\}.
$$

(28)

[18] In this study, the mode limit N is set to 64 and only the axisymmetric component ($m = 0$) is used. Thus B_ϕ and A_θ vanish. The coefficients g and h are determined from the boundary condition

$$
B_{p,r0}(\theta) = B_0 c \exp\left[-\left(\frac{\theta - \theta_c}{\theta_r}\right)^2\right]\tanh\left(\frac{\theta - \theta_c}{\theta_d}\right),
$$

(29)

where the normalized coefficient c is determined as max $(B_{p,r0}(\theta)) = B_0 = 3.0$, and $(\theta_c, \theta_r, \theta_d) = (0.5\pi, 0.1\pi, 0.05\pi)$.

2.3. Boundary Conditions

[19] The bottom boundary ($r = 1$) is a line-tied boundary on which the normal component of the magnetic field (29) is conserved. If we assume that the photosphere can be treated as a no-slip wall, then the normal and meridional components of velocity v_r and v_θ are set to be zero. In order to energize the magnetic field, we impose only azimuthal (toroidal) shearing motion $v_{\psi,sh}(\theta, t)$ to the footpoints of the arcade. We carried out five different simulations, cases A through E, with the five different patterns of shear motion, as shown in Figure 1. Only in case A, is the shearing motion imposed on the whole area between the magnetic poles, that is,

$$
v_{\psi,sh}(\theta, t) = v_0 R(t)\frac{B_{p,r0}(\theta)}{B_0},
$$

(30)

where v_0 is maximum speed and $R(t) = \min(t, 1)$ is a ramp function. The function is useful for reducing the fluctuation due to a sudden velocity injection. The maximum velocity v_0 is set to be 0.03; that is, about 1% of Alfvén wave in the arcade region. In cases B through E, only a limited region near the magnetic neutral line between the latitudes $\theta_{p,max}$ and $\theta_{p,min}$, where max(B_r) and min(B_r), respectively, are sheared by means of the following forms:

$$
v_{\psi,sh}(\theta, t) = v_0 R(t)\begin{cases} 0 & \text{for } \theta < \theta_{p,min} \text{ or } \theta_{p,max} < \theta \\ \frac{\partial f(\theta)}{\partial \theta} & \text{for } \theta_{p,max} \leq \theta \leq \theta_{p,min} \end{cases} ;
$$

(31)

in case B,

$$
f(\theta) = \left\{ 1 - \left(1 - \frac{|B_{p,r0}(\theta)|}{B_0}\right)^2 \right\}^2 \{B_{p,r0}(\theta)\}^2 \mathrm{sgn}\left[B_{p,r0}(\theta)\right].
$$

(32)
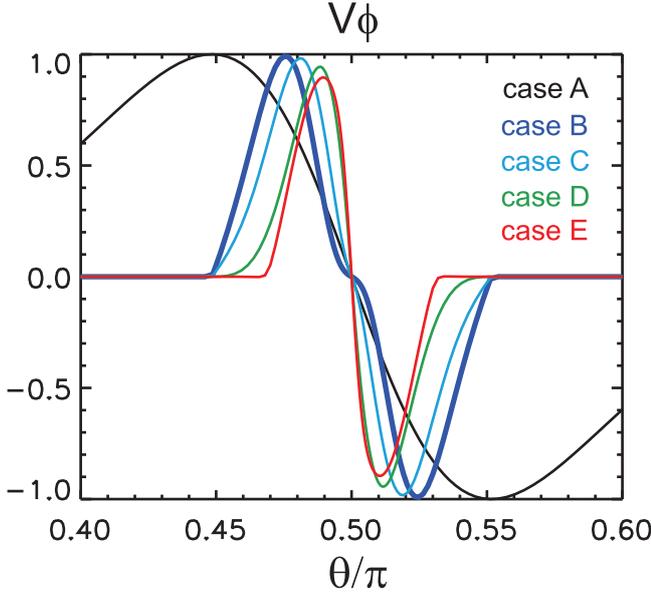
Figure 1. Velocity profiles of the boundary shearing motion.

In case C,

$$f(\theta) = \left\{ 1 - \left(1 - \frac{|B_{p,r0}(\theta)|}{B_0} \right)^2 \right\}^2 B_{p,r0}(\theta). \qquad (33)$$

In case D,

$$f(\theta) = \left\{ 1 - \left(1 - \frac{|B_{p,r0}(\theta)|}{B_0} \right)^2 \right\}^2 \mathrm{sgn}\left[B_{p,r0}(\theta) \right]. \qquad (34)$$

Finally, in case E,

$$f(\theta) = \left\{ 1 - \left(1 - \min\left[2\frac{|B_{p,r0}(\theta)|}{B_0}, 1 \right] \right)^2 \right\}^2 \mathrm{sgn}\left[B_{p,r0}(\theta) \right]. \qquad (35)$$

We have also carried out other simulations with different finite resistivities in cases F through H, with the same shear pattern used for case D.

3. Results

3.1. No Overlying Flux Case

[20] In case A the imposed shearing motion has a distribution that is proportional to the radial component of the magnetic field shown in equation (30). This case is similar to that of *Choe and Lee* [1996]. All the magnetic flux in the arcade is sheared, and there is no overlying arcade which confines the nonpotential flux injected by the shearing motion.

[21] The time evolution of the magnetic field and the toroidal current density is shown in Figure 2 in which the magnetic field lines are illustrated as the contours of the toroidal flux function defined as rA_ϕ. The evolution of the arcade is almost the same as that of rectangular case studied by *Choe and Lee* [1996]. As in their simulation, the

arcade quasi-statically evolves by three steps; in the first step the toroidal component of the magnetic field increases without expanding, in the second step the arcade begins to expand vertically, and in the third step the vertically expanded arcade begins to thin and form a vertical current sheet. Note that the "vertical" direction means to be parallel to the opposite direction of the gravity, and hence it corresponds to the radial direction. As a result, the initial potential arcade evolves into a vertically elongated sheared arcade with an inverse Y-shape current structure (see Figure 2). Further shearing motion leads to continual thinning of the current sheet. When the width of the current sheet approaches the grid scale, numerical diffusion becomes significant and a tearing mode instability ensues. The reconnection which results from the instability forms a magnetic island (a helical flux rope in three dimensions (3-D)), as shown in Figure 2d; this flux rope subsequently ejected into interplanetary space.

[22] Figure 3 shows the time evolution of the toroidal flux function and the toroidal current density. In the axisymmetric spherical coordinates system, the flux function evolving in time is $rA_\phi(r, \theta, t)$. Figure 3 plots the contour of the flux function at the equator determined $rA_\phi(r, \pi/2, t) = $ const with the solid lines and indicates the toroidal current density with grey. The contours of the flux function correspond to the trajectories of plasma frozen in the same magnetic field lines across the equator, and the vertical distance from a contour to the next contour are inversely proportional to the strength of B_θ.

3.2. Partially Sheared Cases

[23] In cases B, C, D, and E, shearing motions are imposed only on the footpoints of the inner region of the initial potential arcade. From case B to case E, the area where shear motion is imposed becomes narrower (see Figure 1). The behavior in these cases is similar to that found by *Choe and Cheng* [2000]. In contrast to case A described in the previous subsection, there are two distinct flux systems; i.e., one corresponds to the overlying unsheared arcade and the other to the inner sheared arcade. Figure 4 shows the current density distribution and contours of the flux function displaying magnetic field lines. In Figure 4, the overlying flux region can be recognized as the region where B_θ is dominant and field lines are concentrated, while the sheared region is where B_ϕ is dominant and the distances among the field lines are large. Between these two regions there is a separatrix surface where the current is enhanced slightly as shown in Figure 4.

[24] A vertical current sheet similar to case A is also formed above the inversion line inside the sheared arcade (see Figure 4b). An inverse Y-shape current structure also forms, and the upper sheet of this structure eventually becomes tearing unstable. When this happens an X point forms at the center of the current sheet. The reconnection at the X point produces a magnetic island corresponding to a helical flux rope in 3-D. However, the speed of the helical flux rope soon decreases because the system approaches a new equilibrium state. Indeed, the newly formed helical flux rope overruns the new equilibrium state and bounces around it (see Figure 3). In the downward phase of the bounce, a horizontal current sheet between the flux rope surface and
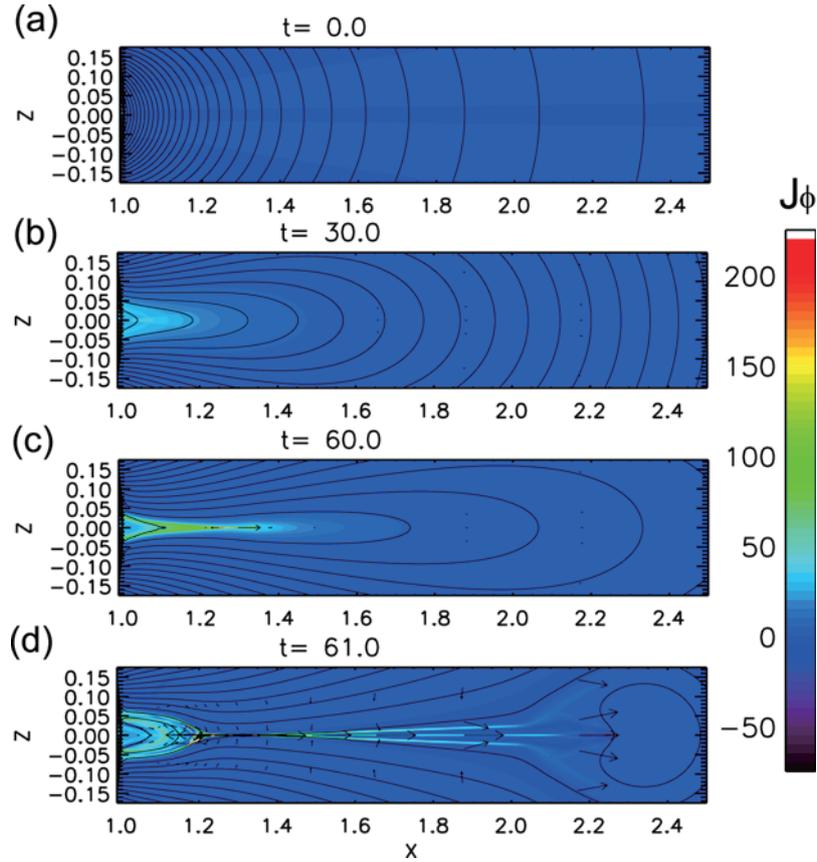
Figure 2. Meridional distribution of the toroidal current density (J_ϕ) in case A. Arrows show the velocity field meridional component. Solid lines projected the magnetic field lines.

the sheared arcade is formed in which reconnection repetitively occurs.

[25] In the other cases (B, C, and E), the evolution is similar to those described above, except for the start times and intervals of the repeated disruption of the sheared arcade. Figure 3 shows the time evolution of the current density distribution above the equator in cases B, C, D, and E.

3.3. Partially Sheared Case With Finite Resistivity

[26] In cases F, G, and H, a shearing motion whose distribution is the same as that in case D is imposed. From case F to case H, the resistivity becomes smaller raging from 10^{-4} to 10^{-6}.

[27] In the case H ($\eta = 10^{-6}$), the evolution is almost the same as that in case D. This suggests that effective numerical diffusion is of the order of 10^{-6}.

[28] In case F, a large resistivity case, the evolution of the magnetic field is much different from the cases described above (Figure 5). The evolution of the field in this case is similar to that in case D during the initial phase when current distribution is smooth (first and second stages discussed in section 3.1). However, as the sheared arcade starts to develop a vertical current sheet (in the third stage), the evolution differs significantly from that in case D. As shown in Figure 5, an extended vertical current sheet is never formed. Instead an elongated sheared arcade (like an inverse V-shape) is formed during this simulation. This result suggests that the timescale of the current dissipation

is shorter than the Alfven timescale in the sheared arcade in this case.

[29] The evolution in case G is the intermediate case between cases D and F. As shown in Figure 3, the start time of the tearing instability is much later than in case D. The reason seems to be the same as for case F and that the inverse V-shape arcade forms before the tearing start time. In contrast to case F, however, reconnection due to the tearing instability is sufficient to form a flux rope. The interval of the repetitive reconnection events are shorter than those in case D.

4. Discussion

4.1. Evolution of Energy, Helicity, and Flux

[30] The change of helicity can be written as follows [*Priest and Forbes*, 2000, chap. 8] with the use of the vector potentials $\mathbf{A_p}$ of equations (27) and (28):

$$\frac{dH}{dt} = -2 \int_V \mathbf{E} \cdot \mathbf{B} dV + \int_S \mathbf{A_p} \times \mathbf{E} \cdot \mathbf{n} dS \tag{36}$$

then

$$\frac{dH}{dt} = -2 \int_V \eta \mathbf{J} \cdot \mathbf{B} dV + \int_S (\mathbf{B} \cdot \mathbf{A_p})(\mathbf{v} \cdot \mathbf{n}) - (\mathbf{v} \cdot \mathbf{A_p})(\mathbf{B} \cdot \mathbf{n}) dS. \tag{37}$$
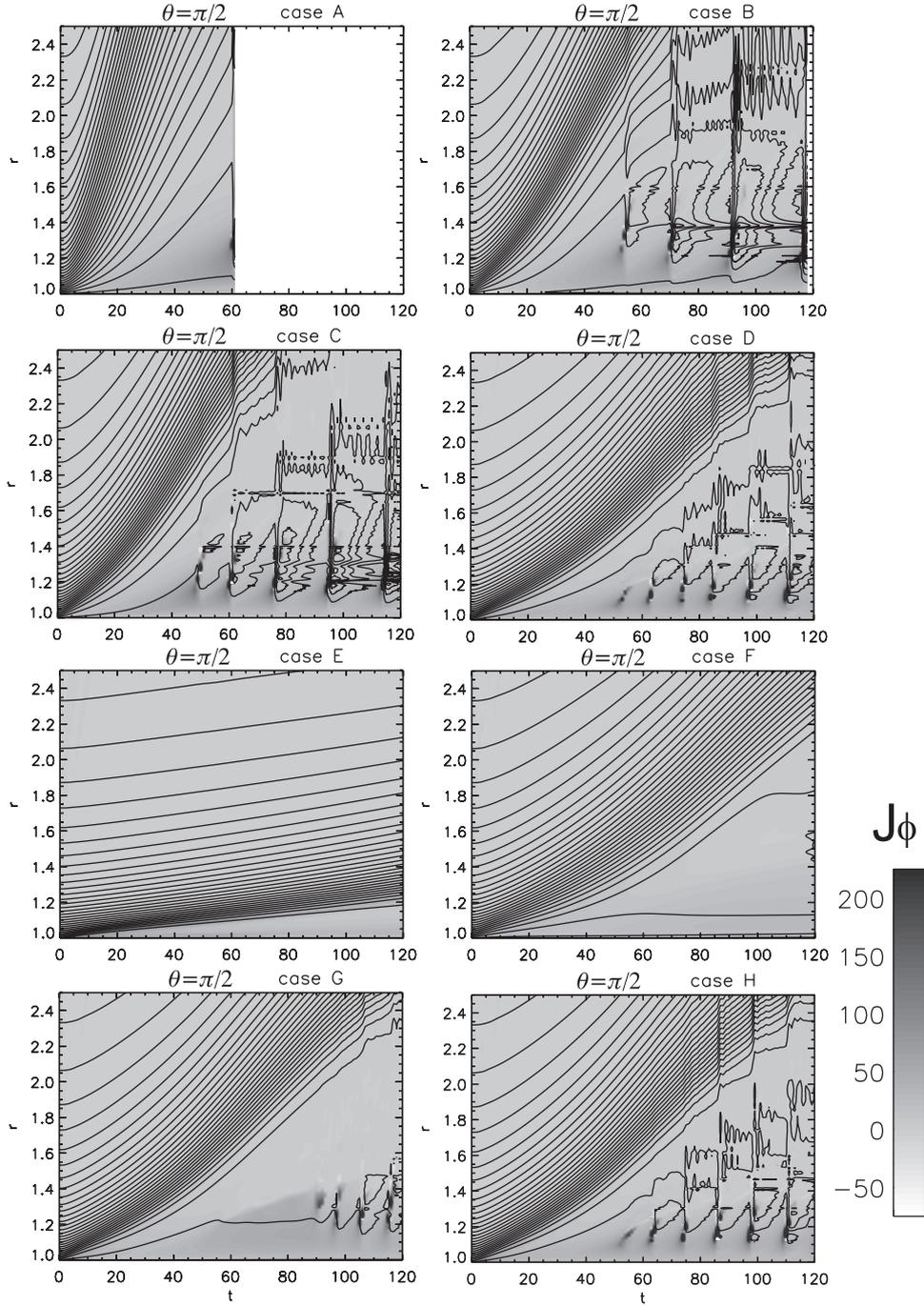
Figure 3. Time evolution of the toroidal current density (J_ϕ) above the equator ($\theta = \pi/2$) in each case. Solid lines are trajectories of the magnetic field.

In this study, $\mathbf{A_p} = A_{p,\phi}\,\mathbf{e}_\phi$ and $\mathbf{v} = v_{sh,\phi}\,\mathbf{e}_\phi$ on the bottom boundary. Therefore equation (37) reduces to

$$\frac{dH}{dt} = - \int_S v_{sh,\phi} A_{p,\phi} B_{p,r0}\,\mathrm{d}S. \qquad (38)$$

which is constant as long as the shearing motion is constant. This relation is confirmed in Figure 6a which shows the time evolution of the total magnetic helicity in the entire numerical domain in each case.

[31] On the other hand, the time evolution of the magnetic energy (Figure 6b) shows three stages in the evolution. In the first stage, the magnetic energy increases until it is about twice that of the potential magnetic energy; in the second stage, the rate of energy increase gradually slows; and in the third stage, the rate of energy increase becomes almost constant. In the third stage, we find many fluctuations of both the magnetic and kinetic energies. These correspond to the disruption of the arcade due to the tearing-mode instability in the central current sheet. These three evolution stages correspond to those of the topological evolution
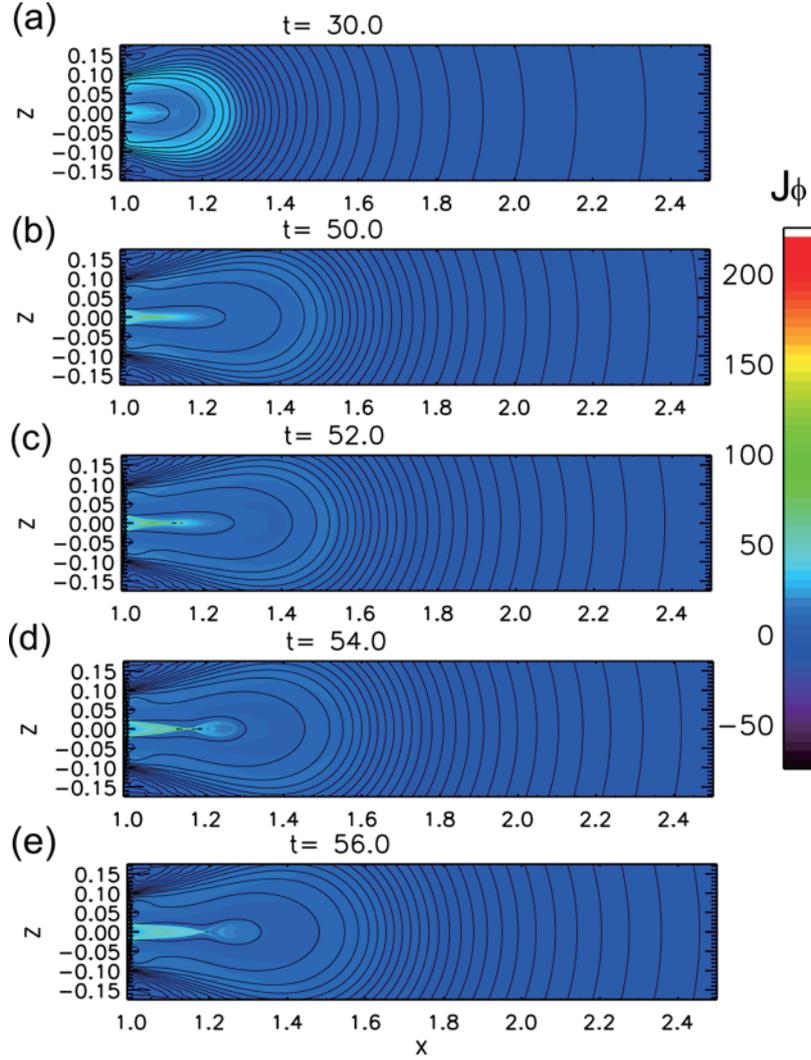
Figure 4. Current density distribution and its time evolution in case D. The meaning of each panel is the same as that of Figure 2.

described in the previous section. This behavior is also similar to that of *Choe and Cheng* [2000].

[32] There are a few observations which succeed in detecting features that result from tearing-mode instability such as plasmoid in the current sheet [e.g., *Ko et al.*, 2003; *Lin et al.*, 2005; *Lin et al.*, 2007] and numerically illustrated by *Riley et al.* [2007]. However, these features were found in the current sheet behind some CMEs where reconnection might continue and accelerate the CMEs. Therefore these features were thought to be not the results of instability in a sheared arcade which discussed in this study, but the results of some instability of a reconnecting current sheet, such as secondary tearing [*Shibata and Tanuma*, 2001].

[33] The relation of kinetic energy to helicity (Figure 6e) shows that the maximum kinetic energy produced by the tearing instability appears to have an upper limit (pink line) which is a function of helicity. This result suggests that the upper limit of released magnetic energy is closely related to the total helicity or energy.

4.2. Relation of Evolution to Energy, Helicity, and Flux

[34] As discussed in section 1, it is thought that the magnetic fluxes in various domains are an important factor in determining whether a CME occurs or not. We first examine how much initial potential magnetic flux is involved in the sheared arcade. The initial potential magnetic flux is calculated by the following surface integral as

$$\Psi_p = \frac{1}{2} \int_S |B_{p,r0}| dS, \tag{39}$$

where S is the bottom boundary. Considering the intersection between the surface and the separatrix, the bottom boundary is divided into two areas S_{ov} and S_{sh}. The total poloidal magnetic flux involved in sheared arcade and overlying arcade as follows:

$$\Psi_{sh} = \frac{1}{2} \int_{S_{sh}} |B_{p,r0}| dS, \tag{40}$$
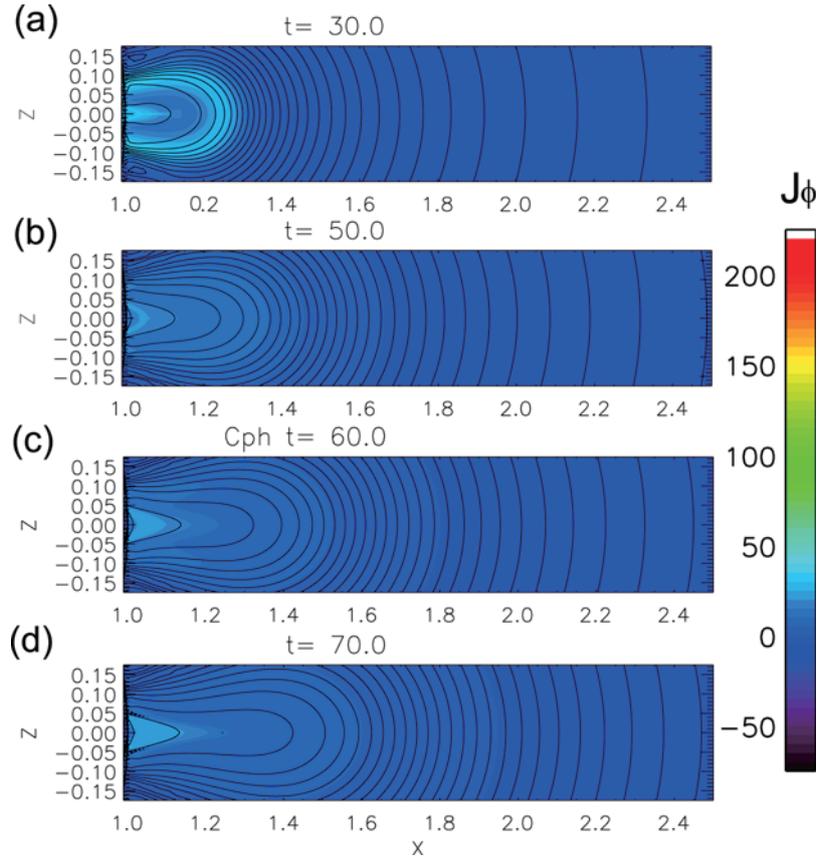
Figure 5. Snapshots of the current density toroidal (J_ϕ) component distribution in the meridional plane ($\phi = 0$) in case F. Solid lines are projected magnetic field lines.

$$\Psi_{ov} = \frac{1}{2} \int_{S_{ov}} |B_{p,r0}| dS. \qquad (41)$$

The results are tabulated in Table 1. It is found that a largest part of the magnetic flux is involved in the overlying arcade. This is the reason why the newly formed flux rope did not eject like a CME in cases B-E. If the condition for CME ejection depends on both the magnetic helicity and the overlying magnetic flux, then there may be some relation between them. The dimensionless parameter

$$\alpha_{CME} = \frac{\Psi_{ov}^2}{H} \qquad (42)$$

could be a key parameter for CME triggering. In case A the flux rope appears to ejected to the outer boundary as a CME (before it reached the boundary the simulation was halted), while in other cases it clearly increases and bounces around an equilibrium point. From Table 1 we get

$$0(\text{case A}) < \alpha_{CME} < 1.7(\text{case B}). \qquad (43)$$

Because this limitation is too broad, more investigation is necessary.

[35] In order to examine how the total helicity is related to the evolution of the field, we compare the evolution in cases A to E based on the total helicity amount. Figure 7 plots the heights of the separatrix and the center of the vertical current sheet versus the total helicity. It is found that the heights of the separatrix in each case, are almost the same, which suggests that the vertical extension of the sheared arcade is closely correlated to the total amount of helicity; i.e., the rise velocity of the separatrix is correlated to the helicity injection rate at the bottom boundary, as long as the magnetic helicity is conserved. As described in section 3.2, a plasmoid (a twisted flux rope) are generated as the result of the nonlinear evolution of tearing-mode instability, and the repeating eruption due to the tearing instability feeds the magnetic flux to the plasmoid in each of cases B through E. Because the flux ropes are confined within the separatrix, which is closely related to the total magnetic helicity, the elevation of the flux ropes is also related to the total magnetic helicity. Figure 8

[36] The growth rate of the tearing-mode instability depends on the aspect ratio of the current sheet, i.e., the height and width of the sheet. As discussed above, the height of the sheared arcade is closely related to the total helicity. Comparing two fields with the same helicity but with different sheared region widths, the growth rate in the field whose sheared region is the most narrow should be the largest. In fact, Figure 3 shows that the interval of the eruption becomes shorter from case B to case E. This result is consistent with that of *Kusano et al.* [1995].
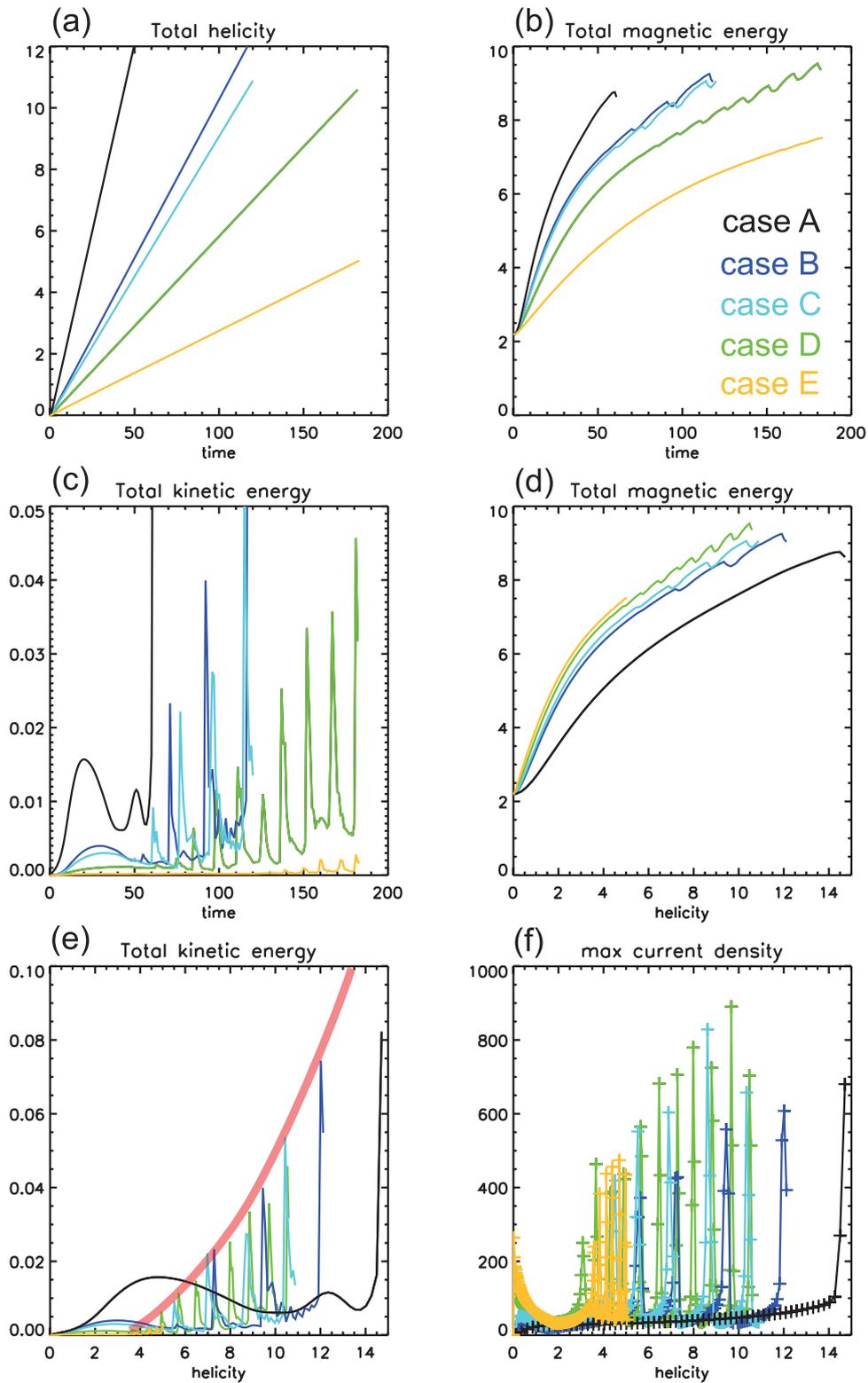
Figure 6. (a) Time profile of total magnetic helicity, (b) total magnetic energy, and (c) total kinetic energy in each case. Also shown is (d,e) the total magnetic energy and the kinetic energy relative to the total helicity. The unit of energy is 1.88×10^{32} erg. Also shown is (f) the maximum current density in the vertical current sheet relative to the total helicity.

Table 1. Fraction of the Magnetic Flux of the Sheared Arcade, the Overlying Arcade, its Square, and the Total Helicity When the Simulation is Stopped

Case	Sheared Arcade	Overlying Arcade	Overlying Arcade^2	Final Helicity
A	5.59	0	0	14.72
B	1.06	4.53	20.52	12.11
C	0.76	4.83	23.33	10.88
D	0.60	4.99	24.90	10.59
E	0.34	5.25	27.56	5.03

[37] However, when the area of the sheared region is small, the amount of the overlying flux is large. As discussed above, the amount of overlying flux contributes to confine the eruption and stabilize the newly formed flux rope as the result of reconnection. Therefore the flux rope does not erupt but finds a new equilibrium state in the lower corona. The necessary condition for this to occur can be determined by the relation between the magnetic helicity and the square of the overlying flux. Hence there is a kind of trade-off between the condition for the tearing instability and that for the eruption of a CME. Consequently, a vast amount of helicity is needed for eruption in an axisymmetric system. In the real solar corona the amount of magnetic helicity injected into the corona is much smaller than that used in this simulation [*Kusano et al.*, 2002; *Maeshiro et al.*, 2005; *Yamamoto et al.*, 2005]. This suggests that it may be difficult to trigger a CME just by the axisymmetric shearing motion and that some other mechanisms should be involved in the trigger process of CME, such as helicity annihilation [*Kusano et al.*, 2004], flux cancellation, or emergence [*Chen and Shibata*, 2000].

[38] From the point of view of magnetic flux, we compare our results with those of other works. Note that their concepts are different from this study because these studies concentrated on only the way to cause CMEs. However, the comparisons give us important suggestions.
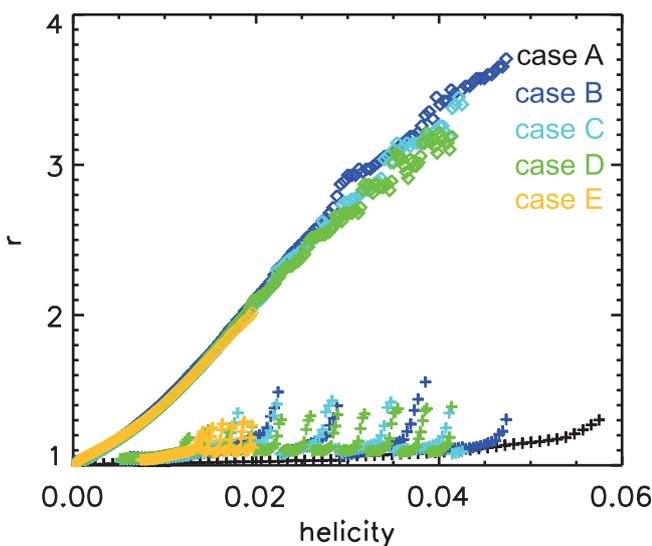
Figure 7. Variations of height of the separatrix between the sheared and unsheared portions of the arcade (shown by diamond symbol) and the center of the central current sheet (shown by cross symbol) versus the total magnetic helicity.

[39] *Mikic and Linker* [1994] performed an early study of arcade shearing in the spherical geometry. In the study, the initial magnetic field is a dipole field and shearing motion is imposed on a localized region beside the equator. It is common to our studies that the shearing motion is imposed only in the inner loops of the initial magnetic field, but a significant difference is that the initial magnetic field is the dipole field. Our results suggest that the confinement of sheared flux by an overlying magnetic flux is an important process for determining whether a CME can occur. However, in the case of the dipole field, the spatial scale gap between the overlying unsheared magnetic flux and the sheared flux is so large (comparable to the solar radius) that magnetic tension force of the unsheared flux is much smaller than our cases. Thus the spatial scale of the region involved in the CME may be one of the key factors for determining the CME occurrence condition.

[40] The breakout model [*Antiochos et al.*, 1999; *MacNeice et al.*, 2004] is also an important model of CME initiation. In these studies, the initial magnetic field is a quadrupole field and a shearing motion is imposed on the inner region beside the equatorial inversion line where the magnetic field radial component becomes 0. As shear is injected to the inner arcade, it swells upward, similar to our results and others. The essential point of the model is that the direction of the magnetic field further above the inner arcade is opposite so that reconnection can strip the magnetic flux which confines the innermost sheared flux. This result and our results suggest that the ambient magnetic field structure may be also one of key factors for determining CME occurrence condition. By the same reason, the effect of open field due to the solar wind should be taken into account.

4.3. Effect of Resistivity

[41] Electric resistivity is another factor governing the tearing instability. The fastest growth rate is

$$\omega = \left(\frac{1}{\tau_d \tau_A} \right)^{1/2}, \tag{44}$$

where $\tau_d = l^2/\eta$ and $\tau_A = l/V_A$ are the diffusion and Alfvén timescales, and l is the width of the current sheet. Therefore in a case with a larger resistivity, the tearing instability will grow faster. This finding is confirmed by the fact the interval for the reconnection events in case G is shorter than in case D.

[42] On the other hand, the larger resistivity can prevent the formation of a current sheet. A current sheet is formed by the thinning of a current containing a sheared arcade. However, if the current diffusion due to resistivity is faster
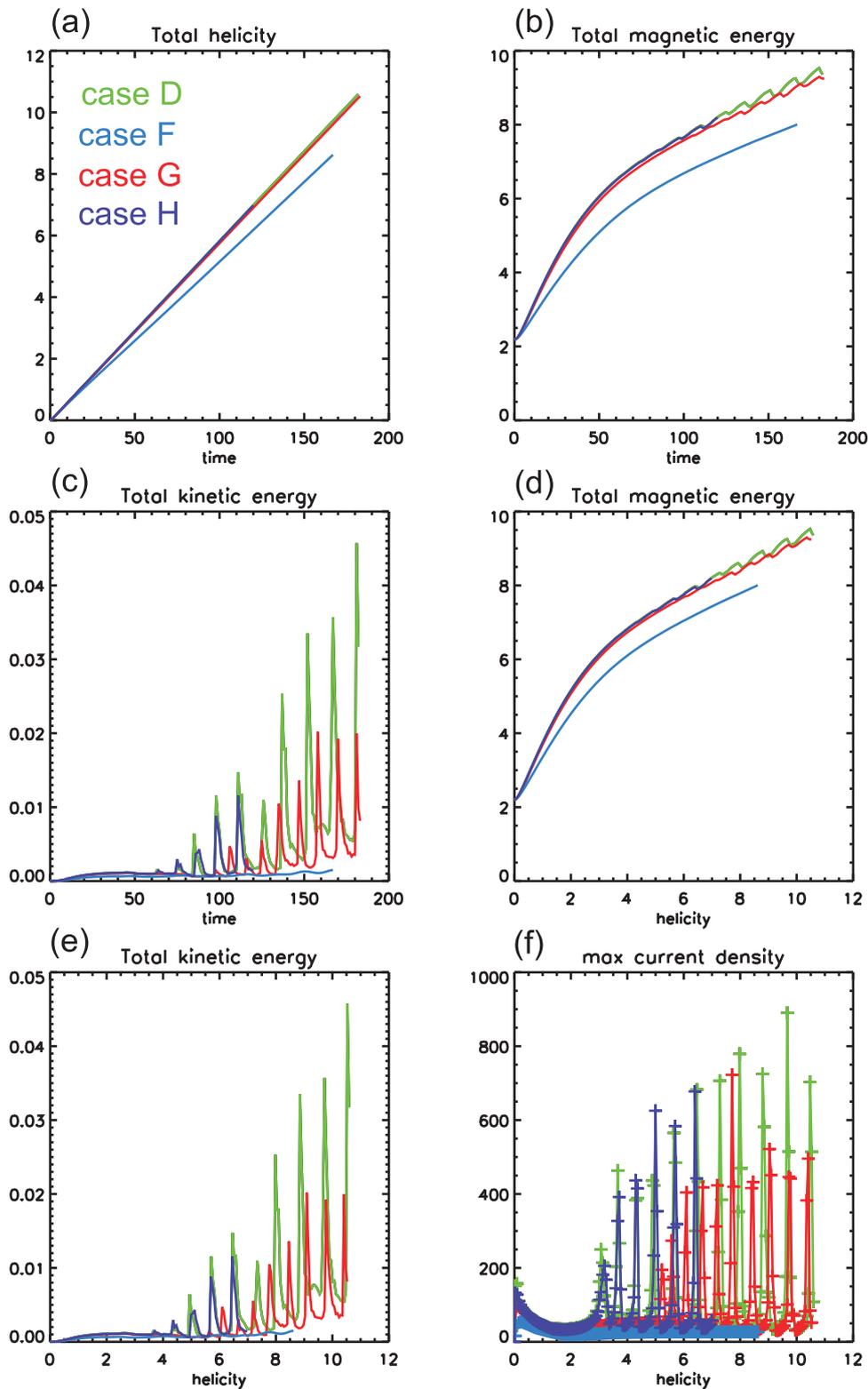
Figure 8. (a) Time profile of total magnetic helicity, (b) total magnetic energy, and (c) total kinetic energy in each case. Also shown are (d,e) the total magnetic energy and the kinetic energy relative to the total helicity. The unit of energy is 1.88×10^{32} erg. Also shown is (f) the maximum current density in the vertical current sheet relative to the total helicity.

than the current enhancement due to thinning, the current sheet is never formed, as in case F.

[43] In cases A to E we set explicit resistivity to be 0, and therefore the evolutions are caused due only to the effective resistivity, numerical diffusion. In cases F, G, and H the resistivity values are set to be 10^{-4}, 10^{-5}, and 10^{-6}, respectively. The evolution in case D is almost the same as that in case H, while it is a little different from that in case G. This result suggests that the effective resistivity due to numerical diffusion is larger than 10^{-6} and less than 10^{-5}.

[44] In our simulation (case G for example) the magnetic Reynolds number is $R_m = V_A L/\eta \sim 1/\eta \sim 10^5$ for the spatial scale of the active region $\sim 0.3 \; R_\odot$. In the solar corona, considering just the Spitzer type resistivity, the magnetic Reynolds number is 10^{14} for an active region. Furthermore, *Lin et al.* [2007] derived the electric resistivity from the recent observations of reconnecting current sheets [*Ko et al.*, 2003; *Lin et al.*, 2005]. The obtained magnetic diffusivities of quiet corona and anomalous resistivity in the cgs units $\eta_{dif} = 4 \; \pi\eta_e/c$ are

$$\eta_{dif,c} = (1.7 \sim 4.6) \times 10^{-6} \text{cm}^2\text{s}^{-1} \qquad (45)$$

and $\eta_{dif,a} = (1.1 \sim 4.2) \times 10^2$ cm^{2}s^{-1}. The corresponding magnetic Reynolds numbers $R_m = V_A L/\eta_{dif} \sim c_s R_\odot/\eta$ are $R_{m,c} = (7.6 \sim 2.8) \times 10^{23}$ cm^{2}s^{-1} and $R_{m,a} = (12. \sim 3.1) \times 10^{15}$ cm^{2}s^{-1} for the spatial scale of the active region. Hence the resistivity in this study is much larger than that in the solar corona because of the following limitation of the numerical simulation. In the numerical simulation the width of the current sheet is limited to the minimum grid scale. This limitation makes the numerical diffusion large, so the tearing instability grows easily. Future investigations with different resolutions will be necessary to extrapolate our results to real coronal conditions. We need more understanding about the anomalous resistivity which might be caused by a plasma microinstability.

5. Conclusion

[45] We have investigated the quantitative relations between arcade shearing, flux rope formation, and eruption with a 2.5-dimensional axisymmetric MHD simulation. We studied several cases with different velocity profiles of the shearing motion and with different resistivity value cases. In the case in which all magnetic flux is sheared, a vertical current sheet is formed, and then reconnection due to tearing-mode instability in the sheet forms helical flux rope and leads to an eruption.

[46] On the other hand, in the cases in which only the inner magnetic flux is sheared, a different structure is formed. This structure consists of both sheared and unsheared flux, and it has a vertical current sheet. As the results of reconnection in the current sheet, a helical flux rope is formed, but this flux rope remains in an equilibrium state and is not ejected. Further injection of shear leads to repeating eruptions of the sheared arcade, which feed toroidal flux to the flux rope. As the results, although the flux rope was elevated gradually, the flux rope has not been ejected as a CME until the end of the simulation. The results suggest that it may be difficult to trigger a CME just by the axisymmetric shearing motion and that some other mechanisms could be involved in the trigger process of a CME.

[47] Resistive cases have also been studied. The results suggest that a sufficiently large resistivity can prevent the current sheet from forming. However, somewhat smaller but still large values enhance the growth rate of the tearing instability.

[48] We examined the relations among evolution, magnetic helicity magnetic energy, and fraction of magnetic flux. It is found that the amount of magnetic helicity is closely related to the topology of the shear configuration and the magnitude of kinetic energy released in a confined eruption.

[49] The parameter ranges in this study are insufficient to make clear the quantitative relations. In order to extrapolate our results to real coronal conditions, we will investigate the issue with wider parameter range and different resolutions in our future studies.

References

Altschuler, M. D., and G. Newkirk (1969), Magnetic fields and the structure of the solar corona. I: Methods of calculating coronal fields, *Sol. Phys.*, *9*, 131–149, doi:10.1007/BF00145734.

Amari, T., J. F. Luciani, J. J. Aly, Z. Mikic, and J. Linker (2003), Coronal mass ejection: Initiation, magnetic helicity, and flux ropes. I. Boundary motion-driven evolution, *Astrophys. J.*, *585*, 1073–1086, doi:10.1086/345501.

Antiochos, S. K., C. R. DeVore, and J. A. Klimchuk (1999), A model for solar coronal mass ejections, *Astrophys. J.*, *510*, 485–493, doi:10.1086/306563.

Berger, M. A., and G. B. Field (1984), The topological properties of magnetic helicity, *J. Fluid Mech.*, *147*, 133–148.

Chen, P. F., and K. Shibata (2000), An emerging flux trigger mechanism for coronal mass ejections, *Astrophys. J.*, *545*, 524–531, doi:10.1086/317803.

Choe, G. S., and C. Z. Cheng (2000), A model of solar flares and their homologous behavior, *Astrophys. J.*, *541*, 449–467, doi:10.1086/309415.

Choe, G. S., and L. C. Lee (1996), Evolution of solar magnetic arcades. II. Effect of resistivity and solar eruptive processes, *Astrophys. J.*, *472*, 372, doi:10.1086/178070.

Dedner, A., F. Kemm, D. Kröner, C.-D. Munz, T. Schnitzer, and M. Wesenberg (2002), Hyperbolic divergence cleaning for the MHD equations, *J. Comput. Phys.*, *175*, 645–673.

Forbes, T. G. (2000), A review on the genesis of coronal mass ejections, *J. Geophys. Res.*, *105*, 23,153–23,166.

Forbes, T. G., and E. R. Priest (1995), Photospheric magnetic field evolution and eruptive flares, *Astrophys. J.*, *446*, 377, doi:10.1086/175797.

Forbes, T. G., et al. (2006), CME theory and models, *Space Sci. Rev.*, *123*, 251–302, doi:10.1007/s11214-006-9019-8.

Hundhausen, A. (1999), Coronal mass ejections, in *The Many Faces of the Sun: A Summary of the Results From NASA's Solar Maximum Mission*, edited by K. T. Strong et al., p. 143, Springer, New York.

Inhester, B., J. Birn, and M. Hesse (1992), The evolution of line-tied coronal arcades including a converging footpoint motion, *Sol. Phys.*, *138*, 257–281.

Kliem, B., and T. Török (2006), Torus instability, *Phys. Rev. Lett.*, *96*(25), 255,002, doi:10.1103/PhysRevLett.96.255002.

Ko, Y.-K., J. C. Raymond, J. Lin, G. Lawrence, J. Li, and A. Fludra (2003), Dynamical and physical properties of a post-coronal mass ejection current sheet, *Astrophys. J.*, *594*, 1068–1084, doi:10.1086/376982.

Kusano, K., Y. Suzuki, and K. Nishikawa (1995), A solar flare triggering mechanism based on the Woltjer-Taylor minimum energy principle, *Astrophys. J.*, *441*, 942–951, doi:10.1086/175413.

Kusano, K., T. Maeshiro, T. Yokoyama, and T. Sakurai (2002), Measurement of magnetic helicity injection and free energy loading into the solar corona, *Astrophys. J.*, *577*, 501–512, doi:10.1086/342171.

Kusano, K., T. Maeshiro, T. Yokoyama, and T. Sakurai (2004), The trigger mechanism of solar flares in a coronal arcade with reversed magnetic shear, *Astrophys. J.*, *610*, 537–549, doi:10.1086/421547.

Lin, J., and T. G. Forbes (2000), Effects of reconnection on the coronal mass ejection process, *J. Geophys. Res.*, *105*, 2375–2392.

Lin, J., W. Soon, and S. L. Baliunas (2003), Theories of solar eruptions: A review, *New Astron. Rev.*, *47*, 53–84.

Lin, J., Y.-K. Ko, L. Sui, J. C. Raymond, G. A. Stenborg, Y. Jiang, S. Zhao, and S. Mancuso (2005), Direct observations of the magnetic reconnection site of an eruption on 2003 November 18, *Astrophys. J.*, *622*, 1251–1264, doi:10.1086/428110.

Lin, J., T. G. Forbes, Y.-K. Ko, J. C. Raymond, and A. Vourlidas (2007), Features and properties of coronal mass ejection/flare current sheets, *Astrophys. J.*, *658*, L123–L126, doi:10.1086/515568.

Linker, J. A., and Z. Mikic (1995), Disruption of a helmet streamer by photospheric shear, *Astrophys. J.*, *438*, L45–L48, doi:10.1086/187711.

Mackay, D. H., and A. A. van Ballegooijen (2005), New results in modeling the hemispheric pattern of solar filaments, *Astrophys. J.*, *621*, L77–L80, doi:10.1086/428904.

Mackay, D. H., and A. A. van Ballegooijen (2006a), Models of the large-scale corona. I. Formation, evolution, and liftoff of magnetic flux ropes, *Astrophys. J.*, *641*, 577–589, doi:10.1086/500425.

Mackay, D. H., and A. A. van Ballegooijen (2006b), Models of the large-scale corona. II. Magnetic connectivity and open flux variation, *Astrophys. J.*, *642*, 1193–1204, doi:10.1086/501043.

MacNeice, P., S. K. Antiochos, A. Phillips, D. S. Spicer, C. R. DeVore, and K. Olson (2004), A numerical study of the breakout model for coronal mass ejection initiation, *Astrophys. J.*, *614*, 1028–1041, doi:10.1086/423887.

Maeshiro, T., K. Kusano, T. Yokoyama, and T. Sakurai (2005), A statistical study of the correlation between magnetic helicity injection and soft X-ray activity in solar active regions, *Astrophys. J.*, *620*, 1069–1084, doi:10.1086/426806.

Magara, T. (2006), Dynamic and topological features of photospheric and coronal activities produced by flux emergence in the Sun, *Astrophys. J.*, *653*, 1499–1509, doi:10.1086/508926.

Manchester, W., IV, T. Gombosi, D. DeZeeuw, and Y. Fan (2004), Eruption of a buoyantly emerging magnetic flux rope, *Astrophys. J.*, *610*, 588–596, doi:10.1086/421516.

Mikic, Z., and J. A. Linker (1994), Disruption of coronal magnetic field arcades, *Astrophys. J.*, *430*, 898–912, doi:10.1086/174460.

Mikic, Z., D. C. Barnes, and D. D. Schnack (1988), Dynamical evolution of a solar coronal magnetic field arcade, *Astrophys. J.*, *328*, 830–847, doi:10.1086/166341.

Miyoshi, T., and K. Kusano (2005), A multi-state HLL approximate Riemann solver for ideal magnetohydrodynamics, *J. Comput. Phys.*, *208*, 315–344, doi:10.1016/j.jcp.2005.02.017.

Priest, E., and T. Forbes (2000), *Magnetic Reconnection*, 612 pp., Cambridge Univ. Press, Cambridge, UK.

Priest, E. R., and T. G. Forbes (2002), The magnetic nature of solar flares, *Astron. Astrophys. Rev.*, *10*, 313–377, doi:10.1007/s001590100013.

Riley, P., R. Lionello, Z. Mikiæ, J. Linker, E. Clark, J. Lin, and Y.-K. Ko (2007), "Bursty" reconnection following solar eruptions: MHD simulations and comparison with observations, *Astrophys. J.*, *655*, 591–597, doi:10.1086/509913.

Shibata, K., and S. Tanuma (2001), Plasmoid-induced-reconnection and fractal reconnection, *Earth Planets Space*, *53*, 473–482.

Steinolfson, R. S. (1991), Coronal evolution due to shear motion, *Astrophys. J.*, *382*, 677–687, doi:10.1086/170755.

van Ballegooijen, A. A., and P. C. H. Martens (1989), Formation and eruption of solar prominences, *Astrophys. J.*, *343*, 971–984, doi:10.1086/167766.

Yamamoto, T. T., K. Kusano, T. Maeshiro, T. Yokoyama, and T. Sakurai (2005), Magnetic helicity injection and sigmoidal coronal loops, *Astrophys. J.*, *624*, 1072–1079, doi:10.1086/429363.

_____

K. Kusano, Earth Simulator Center, Japan Agency for Marine-Earth Science and Technology, Yokohama, Kanagawa, Japan.

T. Miyoshi, Facility of Science, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan.

N. Nishikawa, Super Computer System Planning and Operations Department, Japan Agency for Marine-Earth Science and Technology, Yokohama, Kanagawa, Japan.

K. Shibata, Kwasan and Hida Observatories, Kyoto University, Yamashina, Kyoto, Japan.

D. Shiota, Center for Computational Astrophysics (CfCA), National Astronomical Observatory of Japan, 2-21-1 Ohsawa, Mitaka, Tokyo, Japan 181-8588. (shiota@cfca.jp)

Parametric decay of circularly polarized Alfvén waves in the radially expanding solar wind

Shin Tanaka,[1] Tatsuki Ogino,[1] and Takayuki Umeda[1]

[1]   We present a magnetohydrodynamic (MHD) simulation study of the parametric decay of the circularly polarized Alfvén waves propagating in the radial outflow of the solar wind. Assuming the transonic wind solution as an initial condition, we continuously injected monochromatic circularly polarized Alfvén waves from the inner boundary at the lower corona, and simulated the wave propagation. The injected Alfvén waves are subject to the parametric decay, and density fluctuations in the solar wind plasma grow rapidly at a specific region. The location of the most unstable region depends on the amplitude and frequency of injected Alfvén waves. We found that the unstable region of simulation results can be well estimated by considering a localized dispersion relation in the frame of reference moving with the background solar wind.

1.   Introduction

[2]   Alfvén waves are ubiquitous phenomena in space plasmas, and are commonly observed in various situations such as the solar wind [*Belcher and Davis*, 1971] and the interplanetary shocks [*Agim et al.*, 1995]. Since Alfvén waves are non-compressional transverse magnetohydrodynamic (MHD) modes, they can propagate a long distance compared to compressional MHD modes which can easily damp. Because of this nature, Alfvén waves are thought to propagate along the open magnetic flux tube in a polar coronal hole and to play an important role in the acceleration of the high-speed solar wind.

[3]   It is well known that circularly polarized Alfvén waves with a finite amplitude are an exact solution to the nonlinear MHD equations, and are unstable to a decay process in the presence of the density fluctuation [*Sagdeev and Galeev*, 1969]. The pump Alfvén waves decay into the forward-propagating sound waves and the backward-propagating Alfvén waves due to a nonlinear three-wave coupling. *Goldstein* [1978] and *Derby* [1978] derived the general dispersion relation for fundamental modes from ideal MHD equations with finite β. Recent observation of the solar wind suggests the presence of outward- and inward-propagating Alfvén waves [*Bavassano et al.*, 2000]. The parametric decay was thought to be one of triggering mechanisms of inward-traveling Alfvénic modes and an MHD turbulence in the solar wind.

[4]   Recently, several one- and multidimensional MHD simulations have been performed to examine the parametric decay of Alfvén waves in a homogeneous medium with both the periodic and open boundaries [*Del Zanna et al.*, 2001; *Malara et al.*, 2000; *Ghosh and Goldstein*, 1994; *Ghosh et al.*, 1994]. *Pruneti and Velli* [1997] and *Turkmani and Torkelsson* [2004] consider the propagation of Alfvén waves in a gravitationally stratified medium in a slab geometry. In their simulations, Alfvén waves are injected from the one side of boundaries in open systems. However, they assumed a static background medium as an initial state.

[5]   The propagation of Alfvén waves in the solar wind has been extensively studied by many authors through numerical simulations in association with the coronal heating and wind acceleration [*Ofman*, 2004; *Ofman and Davila*, 1998; *Lau and Siregar*, 1996; *Buti et al.*, 1998; *Buti et al.*, 1999]. *Suzuki and Inutsuka* [2006] performed an MHD simulation from the photosphere toward the solar wind, and concluded that low-frequency Alfvén waves can establish both the coronal high temperature and the high-speed solar wind. Most of these studies focused on a quasi-steady state after injected Alfvén waves and compressive waves were saturated through linear and nonlinear processes such as wave steepening, reflection, and possibly parametric decay.

[6]   In the present study, we focus on a growing stage of the parametric decay of Alfvén waves in the radially expanding medium before the quasi-steady state. Circularly polarized monochromatic Alfvén waves are injected into a spherically symmetric transonic wind [*Parker*, 1963]. The injected Alfvén waves propagate outward along the radially expanding flow tube of the solar wind. When the amplitude is sufficiently large, Alfvén waves are subject to the parametric decay, generating outward-propagating sound waves and inward-propagating Alfvén waves. We analyze the dispersion relation to determine the heliocentric distance at which the compressive fluctuation grows most rapidly by the parametric decay.

---

[1]Solar-Terrestrial Environment Laboratory, Nagoya University, Aichi, Japan.

[7] The paper is organized as follows. In Section 2, we introduce the model equation of Alfvén waves in the solar wind, the initial and boundary condition. In Section 3, we show the results of MHD numerical simulations for a few parameters. In Section 4, we discuss the parametric decay for numerical results and compare with the linear theory.

2. Simulation Model

[8] In order to model the radial expansion of the solar wind, we solve the following ideal MHD equations in a one dimensional spherical coordinate in a heliocentric distance r:

$$\frac{\partial \rho'}{\partial t} + \frac{\partial}{\partial r}(\rho' v_r) = 0 \tag{1}$$

$$\frac{\partial}{\partial t}(\rho' v_r) + \frac{\partial}{\partial r}(\rho' v_r^2) = -\frac{\partial}{\partial r}\left(p' + \frac{|\mathbf{B}'_\perp|^2}{2\mu_0}\right) + \rho' \frac{|\mathbf{v}'_\perp|^2}{r^3} + \frac{2p'}{r} + \rho' g \tag{2}$$

$$\frac{\partial}{\partial t}(\rho' \mathbf{v}'_\perp) + \frac{\partial}{\partial r}(v_r \rho' \mathbf{v}'_\perp) = \frac{\partial}{\partial r}(C\mathbf{B}'_\perp) \tag{3}$$

$$\frac{\partial e'}{\partial t} + \frac{\partial}{\partial r}\left[v_r\left(e' + p' + \frac{|\mathbf{B}'_\perp|^2}{2\mu_0}\right)\right] = \frac{\partial}{\partial r}\left[B_r(\mathbf{v}'_\perp \cdot \mathbf{B}'_\perp)\right] + \rho' v_r g \tag{4}$$

$$\frac{\partial \mathbf{B}'_\perp}{\partial t} = -\frac{\partial}{\partial r}\left(v_r \mathbf{B}'_\perp - B_r \mathbf{v}'_\perp\right) \tag{5}$$

where $\rho' \equiv r^2\rho$, $p' \equiv r^2 p$, $e' \equiv r^2 e$, $\mathbf{v}'_\perp \equiv r\mathbf{v}_\perp$ and $\mathbf{B}'_\perp \equiv r\mathbf{B}_\perp$ with ρ, p, v_r, $\mathbf{v}_\perp$, $\mathbf{B}_\perp$ and μ_0 being the density, pressure, radial component of velocity, tangential component of velocity, tangential component of magnetic field, and magnetic permeability in vacuum, respectively. Here $g = GM_S/r^2$ is a gravitational acceleration where G and M_S are a gravitational constant and a solar mass, respectively, and

$$e = \frac{\rho}{2}\left(v_r^2 + |\mathbf{v}_\perp|^2\right) + \frac{p}{\gamma - 1} + \frac{|\mathbf{B}_\perp|^2}{2\mu_0}, \tag{6}$$

is the total energy of the fluid, where γ is the polytropic index. The radial component of the magnetic field B_r is given by $B_r = Cr^{-2}$ where C is a constant.

[9] We solve Equations (1)–(5) with the PIC-MOCCT scheme [S. Tanaka et al., submitted to *Comput. Phys. Commun.*], which consists of the PIC scheme [T. Umeda, submitted to *Comput. Phys. Commun.*] and the MOCCT [*Hawley and Stone*, 1995]. The PIC scheme is a non-oscillatory, positivity preserving and conservative numerical interpolation scheme for solving one-dimensional advection equations, which is used to solve advection phases of the MHD equations in the left hand side of Equations (1)–(4). The MOCCT solves the characteristics of Alfvén waves, maintaining $\nabla \cdot \mathbf{B} = 0$, which is applied to the induction

Equation (5) and terms for the magnetic stress, i.e., the right hand side of Equation (3) and the first term in the right hand side of Equation (4). In the present study, we adopted the PIC-MOCCT scheme is because this scheme does not need to introduce an additional diffusion term to suppress numerical oscillations. This scheme can treat Alfvén waves more accurately than previous non-oscillatory numerical schemes for MHD equations such as TVD schemes [e.g., *Fukuda and Hanawa*, 1999; *Tanaka*, 1994].

[10] The simulation domain is taken from $r = 1\ R_S$ to $r = 40\ R_S$ on an uniform mesh with 16384 grid points. As an initial condition, we used a steady transonic wind solution [*Parker*, 1963] for ρ, p and v_r. We used the following solar wind parameters: $T_0 = 1.8 \times 10^6$ K, $n_0 = \rho_0/m = 1.0 \times 10^8$ cm^{-3} and $B_{r0} = 5.0$ G, where T_0, n_0, B_{r0}, ρ_0 and m are the temperature, number density, radial component of the magnetic field, density at the lower boundary ($r = 1\ R_S$) and a proton mass, respectively. We imposed a very small random initial noise ($\sim 10^{-5}$ to the initial background profiles) to the initial density, pressure and radial velocity to give compressive fluctuations.

[11] In the present study we set $\gamma = 1.1$. Note that the properties of the wind solution for the initial condition depend on γ. With $\gamma = 1.1$ we have the radial velocity $v_r \sim$ 300 km/s at the upper boundary $r = 40\ R_S$, while we have $v_r \sim 400$ km/s with $\gamma = 1.05$. The previous coronal models used, for example, $\gamma = 1.05$ [e.g., *Linker et al.*, 1990; *Linker et al.*, 1999]. We have performed several test runs with different γ, and found that physical processes were essentially the same.

[12] Figure 1 shows the initial profiles of the bulk velocity of the solar wind v_r, sound speed $v_S = \sqrt{\gamma p/\rho}$, and Alfvén speed $v_A = B_r/\sqrt{\mu_0\rho}$, respectively.

[13] We impose the following boundary conditions for the tangential component of the velocity and magnetic field to inject the monochromatic circularly polarized Alfvén waves at the lower boundary ($r = 1\ R_S$):

$$v_\perp = \delta v_0\left(\sin \omega_0 t \cdot \mathbf{e}_\theta + \cos \omega_0 t \cdot \mathbf{e}_\phi\right) \tag{7}$$

$$\mathbf{B}_\perp = -\mathbf{v}_\perp\sqrt{\mu_0\rho_0} \tag{8}$$

where δv_0, ω_0, and ρ_0 are the velocity amplitude, frequency of injected Alfvén waves, and the density at the lower boundary, respectively. Here, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ are the unit vectors of tangential components. Equations (7) and (8) give the relationship between the velocity amplitude and magnetic field amplitude of the outward-propagating Alfvén waves. Note that since we use ideal MHD equations the circular polarization does not affect the present simulation results.

[14] The generation mechanisms of Alfvén waves are expected to be due to solar activities near the solar surface. *Cranmer and Ballegooijen* [2005] consider the turbulent motion of the foot point of the solar magnetic field as a source of waves. *Parhi et al.* [1997] simulated a propagation of MHD waves which are constantly driven by photospheric foot point motions in the context of the coronal heating. *Sturrock* [1999] suggests that magnetic reconnections in the chromosphere can generate MHD waves.
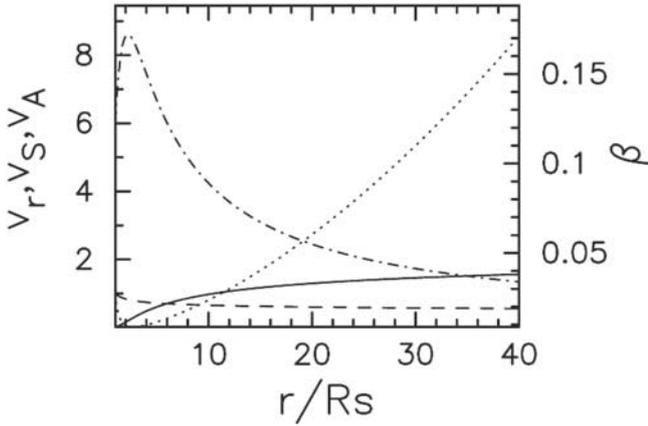
Figure 1. Initial profiles of the radial velocity of the solar wind v_r (solid line), sound speed v_S (dashed line), Alfvén speed v_A (dashed-dotted line), and plasma β (dotted line). The velocities v_r, v_S and v_A are normalized by the sound speed at the lower boundary, $r = 1\ R_S$.

[15] The lower boundary condition for ρ, e, and v_r are fixed at the initial values. Note that ρ, e and v_r should be determined from the characteristic equations to avoid an unphysical wave reflection [*Steinolfson and Nakagawa*, 1976]. However, we confirmed that the wave reflection at the lower boundary is less than 3% for the present parameters. At the upper boundary ($r = 40\ R_S$), we impose the open boundary condition where the spatial derivatives of all the variables are set to be zero. Since the bulk velocity of the solar wind exceeds both the sound speed and the Alfvén speed near the upper boundary as shown in Figure 1, unphysical reflections, which might occur at the upper boundary, is considered to be negligibly small.

3. Simulation Results

[16] In the present study, the amplitude and frequency of injected Alfvén waves are varied as different simulation runs. Detailed parameters are shown in Table 1.

[17] Figure 2 shows the profiles of the number density n (a), the radial velocity components v_r (b), and the ratio of the tangential magnetic field component to the ambient magnetic field $|B_\perp|/B_r$ (c) at time $t = 430$ min. for Run A, respectively. As injected Alfvén waves propagate outward, v_r and ρ become larger than their initial values. At $r \sim 30\ R_S$, we found discontinuities in v_r and n. These are generated by the outward-directed magnetic pressure due to the discontinuous variation of the magnetic field amplitude at the wavefront shown in Figure 2c. Around $r \sim 18\ R_S$, there appear fluctuations in the velocity, density, and the tangential magnetic field. These fluctuations propagate both outward and inward as they grow. This indicates the parametric decay, in which outward-propagating sound waves and inward-propagating Alfvén waves are excited by the large-amplitude Alfvén waves.

[18] To confirm the inward-propagating Alfvén waves, we plot the θ component of the Elsässer vectors, $Z_\theta^\pm = v_\theta \mp B_\theta/\sqrt{\mu_0 \rho}$ in Figures 2d and 2e. The Elsässer vectors Z_θ^+ and Z_θ^- correspond to amplitude of the outward- and inward-

propagating Alfvén waves, respectively. The profile of Z_θ^- shows the apparent generation of inward-propagating Alfvén waves in the unstable region.

[19] There are two possible mechanisms to generate the inward-propagating Alfvén waves. One is the parametric decay and the other is the wave reflection due to the inhomogeneity of background medium. The wave reflection becomes important when the frequency of Alfvén waves ω_0 is smaller than the spatial gradient of the Alfvén speed dv_A/dr [*An et al.*, 1990]. Figure 2e shows that there also exist inward-propagating components outside the unstable region. However, the frequency of injected Alfvén waves are $\omega_0 \geq 10$ mHz in Runs A-C, while the maximum value of dv_A/dr is ~ 1.4 mHz. Thus the amplitude of inward-propagating waves generated by the reflection is considerably small compared to that by the parametric decay. It is also noted that when Alfvén waves are linearly polarized, the variation of the magnetic pressure can generate secondary compressive modes [e.g., *Hollweg*, 1971] which leads to a reflection of Alfvén waves. However, since we inject circularly polarized Alfvén waves, the compressive motions are negligibly small compared to the transverse motion until the parametric decay evolves.

[20] Note that the beat instability is also a possible mechanism to generate inward-propagating Alfvén waves in an ideal MHD system for $\beta > 1$ [*Jayanti and Hollweg*, 1993]. The beat instability involves the interaction between the forward-propagating pump Alfvén waves and backward-propagating lower sideband Alfvén waves. The beat instability does not occur in the present simulations because $\beta < 1$ as shown in Figure 1.

[21] To confirm the parametric decay, in Figure 3, we plot the $\omega - k$ spectrum with a gray scale obtained by Fourier transformation of simulation data in the range of $15.56\ R_S < r < 20.44\ R_S$ and 319 min $< t < 474$ min for (B_θ, B_ϕ) and v_r. This range of r covers the most unstable region as seen in Figure 2, and the range of t covers from an early stage with a steady flow to the saturation stage in which the growing compressive modes have sufficiently large amplitude. We found the injected Alfvén waves at $K \sim 10\ R_S^{-1}$, $\Omega \sim 11$ mHz in Figure 3a, outward-propagating sound waves at $k_S \sim 15\ R_S^{-1}$, $\omega_S \sim 8$ mHz in Figure 3b, and inward-propagating Alfvén waves at $k_A \sim -5\ R_S^{-1}$, $\omega_A = 2$ mHz in Figure 3a. Note that the frequency in the $\omega - k$ spectra is Doppler-shifted, i.e., $\omega = \omega' + kv_r$, where ω and ω' represent the frequency of waves observed from the rest frame and from the moving reference frame with the solar wind velocity v_r, respectively. We found that observed three waves satisfied the resonant condition $K = k_S + k_A$, $\Omega = \omega_S + \omega_A$ for the parametric decay [*Sagdeev and Galeev*, 1969] by taking into account the Doppler shift with the radial velocity $v_r \sim 240$ km/s in the range of the spectrum data ($15.56\ R_S < r < 20.44\ R_S$). We also found another wave mode at $k_A \sim 25\ R_S^{-1}$ and $\omega_A \sim 19$ mHz in Figure 3a. This wave mode satisfies

Table 1. Parameters of Injected Alfvén Waves for Different Simulation Runs

Run	δv_0	ω_0
A	30 km/s	10 mHz
B	20 km/s	10 mHz
C	30 km/s	20 mHz

another resonant condition $k_A = k_S + K$ and $\omega_A = \omega_S + \Omega$, which is associated with the excitation of the outward-propagating magnetic fluctuation. Since this wave is not a normal mode, the power of this wave is much smaller than that of inward-propagating waves [*Goldstein*, 1978].
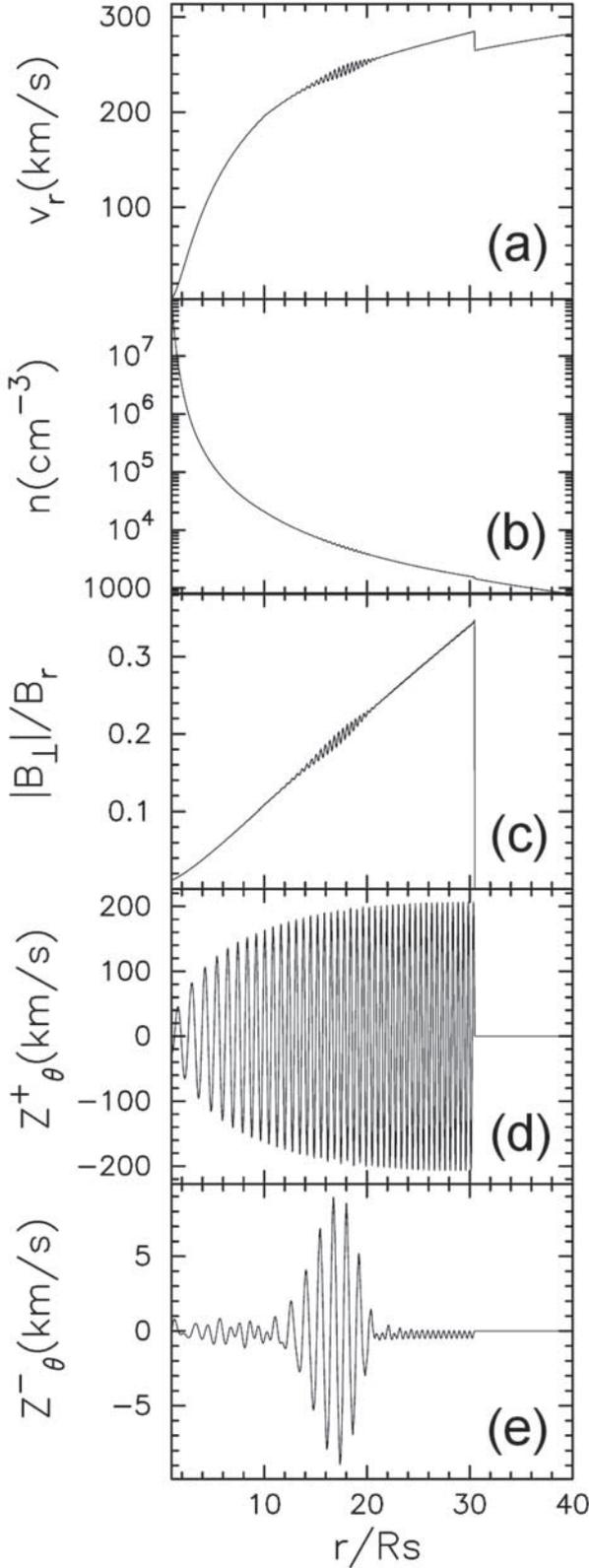
[22] Figure 4 shows the profile of v_r for Runs B and C. In both cases, the injected Alfvén waves generate the fluctuations in a quite similar manner to Run A, and we confirmed the parametric decay (not shown). However, the positions of the unstable region are different from each other. In Run B the amplitude of injected Alfvén waves is smaller than in Run A. Then the position of unstable region becomes farther from the lower boundary. In Run C, on the other hand, the frequency of injected Alfvén waves is higher than in Run A, Then the position of unstable region becomes closer to the lower boundary. From these results, it is expected that the position of the unstable region becomes closer to the lower boundary as the amplitude or frequency of injected Alfvén waves becomes larger.

4. Discussion

[23] The present simulation results suggest that compressive fluctuation grows most rapidly at a localized region by the parametric decay. It is known that a larger amplitude of Alfvén waves or a lower β lead to higher growth rate of the parametric decay [*Goldstein*, 1978; *Derby*, 1978]. In the present case the profile of the normalized amplitude of Alfvén waves $b \equiv B_\perp 0/B_r$ increases as the injected Alfvén waves propagate outward while the local beta value $\beta \equiv v_S^2/v_A^2$ decreases (see Figures 1 and 2c). These profiles imply that the growth rate of the parametric decay has a maximum value at a certain position.

[24] We discuss the location and wave number of the most rapidly growing compressive mode due to the parametric decay instability. The dispersion relation for the parametric decay instability in a homogeneous plasma was first derived by *Goldstein* [1978] and *Derby* [1978]. To consider the dispersion equation in a inhomogeneous background as in a present simulation, we assume so-called Wentzel-Kramers-Brillouin (WKB) approximation in which the wavelength of Alfvén waves is small compared to the scale height of the background. We computed the pressure scale height $H = p(\partial p/\partial r)^{-1}$ and the wavelength of injected Alfvén waves $\lambda = 2\pi (v_A + v_r)/\omega_0$. We found that the wavelength is shorter than the pressure scale height for $r > 4 R_S$, and is comparable for $r < 4 R_S$, indicating that the condition of the WKB approximation is well satisfied for $r > 4 R_S$.

[25] In the WKB approximation, we can consider the dispersion equation for a homogeneous medium as a local dispersion equation at each positions. Let Ω' and K be the frequency and wave number of injected Alfvén waves, which is regarded as the pump wave. Then, we have the local dispersion relation for fundamental modes induced by the injected circularly polarized Alfvén waves with the

Figure 2. Results of Run A at $t = 430$ min. The radial dependence of (a) radial velocity v_r (km/s), (b) number density n (cm^{-3}), (c) magnitude of the tangential magnetic field $|B_\perp|$ normalized by local radial magnetic field B_r, and (d)(e) θ component of Elsässer vectors $Z_\theta^\pm = v_\theta \mp B_\theta/\sqrt{\mu_0\rho}$ (km/s) are shown. Note the compressive fluctuations and inward-propagating Alfvén waves growing around $r \sim 18 R_S$.
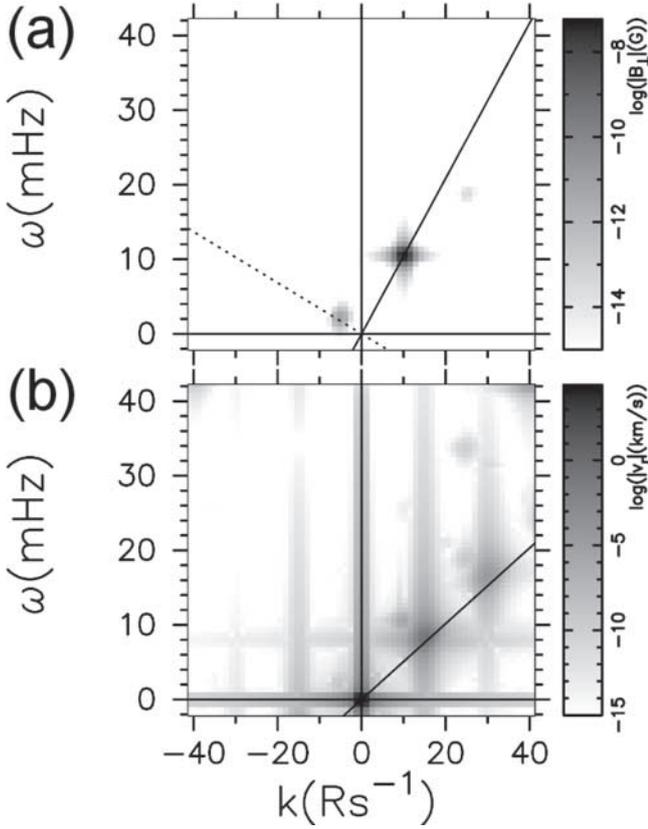
Figure 3. The ω-k diagram of simulation data in the range of $15.56\,R_S < r < 20.44\,R_S$ and $319\,\text{min} < t < 474\,\text{min}$. The horizontal and vertical axes indicate $k(R_S^{-1})$ and $\omega(\text{mHz})$, respectively. Positive (negative) k correspond to outward- (inward-) propagating waves. (a) Fourier amplitude $|B_\perp(k, \omega)|$ plotted with a gray scale. Solid line and dotted line indicate the dispersion relation of outward- and inward-propagating Alfvén waves observed in the rest frame $\omega = (v_r + v_A)k$ and $\omega = (v_r - v_A)k$, respectively, where the solar wind velocity and Alfvén speed are $v_r = 240$ km/s and $v_A = 500$ km/s, respectively. (b) Fourier amplitude $|v_r(k, \omega)|$ plotted with a gray scale. Solid line indicates the dispersion relation of outward-propagating sound waves in the rest frame $\omega = (v_r + v_S)k$ where the sound speed is $v_S = 100$ km/s. These results show that excited modes satisfy the resonant condition of the parametric decay(see text).

finite amplitude in the solar wind frame [*Goldstein*, 1978; *Derby*, 1978],

$$(\omega'^2 - v_S^2 k^2)(\omega' - v_A k)(\omega' + v_A k + 2\Omega')(\omega' + v_A k - 2\Omega')$$
$$= b^2 v_A^2 k^2 (\omega'^3 + \omega'^2 v_A k - 3\Omega'^2 \omega' + \Omega'^2 v_A k) \qquad (9)$$

where ω' and k are the frequency and wave number of compressive fluctuations, respectively. Here, v_S and v_A are the sound speed and Alfvén speed of the solar wind plasma, respectively, and b is a normalized magnetic field amplitude defined as $b \equiv B_\perp/B_r$ where $B_\perp$ and B_r are the magnetic field amplitude of injected Alfvén waves and the background radial magnetic field, respectively. The frequency ω' and Ω' are observed in the solar wind frame moving at the velocity v_r, and are expressed by using the relation of the

Doppler shift as, $\omega' = \omega - k v_r$ and $\Omega' = \Omega - K v_r$ where ω and Ω are the frequency of compressive fluctuations and injected Alfvén waves observed in the rest frame, respectively. Note that v_S, v_A, v_r, b, K, and Ω' are functions of r.

[26] We solved Equation (8) numerically for a complex solution $\omega = \omega_r + i\gamma$ which exhibits an unstable compressive mode, and obtained the growth rate γ at each wave number k and each position r. The values of v_S, v_A, v_r, and b, in Equation (8) are given as a function of r. Note that Ω is assumed to be equal to ω_0 which is a frequency of the injection of Alfvén waves at the lower boundary and is constant anywhere in a simulation domain while K varies with r. Then the wave number K for injected Alfvén waves in the rest frame is given as $K = \Omega/(v_A + v_r)$. By solving Equation (8) as described above, we have the locally defined growth rate $\gamma(r, k)$ for the parametric decay instability.

[27] In the present simulation the pump Alfvén waves are injected into the background medium which is initially a steady solar wind outflow with the bulk velocity $v_r(r)$. It is expected that the fluid element of the solar wind at a certain distance undergoes injected waves at a certain time, and then the density fluctuation in the fluid element will begin to grow due to the instability. We assume that this density
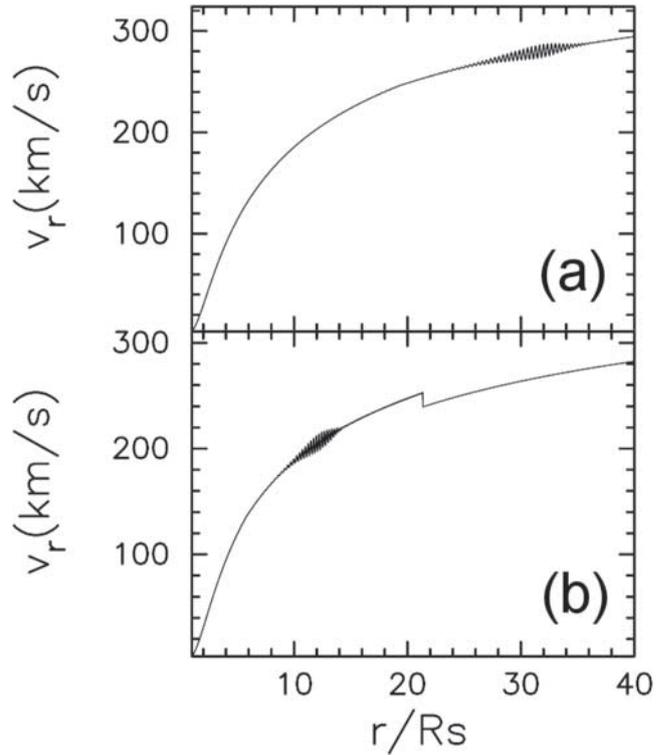
Figure 4. The radial dependence of the radial velocity v_r. (a) Result of Run B (a case of a smaller δv_0 than Run A) at $t = 758$ min. The discontinuity due to the magnetic pressure as shown in Figure 2 does not exist because it propagated out of the simulation domain. (b) Result of Run C (a case of a larger ω_0 than Run A) at $t = 257$ min. These results indicate that a larger amplitude and frequency of injected Alfvén waves make the most unstable region closer to the lower boundary.
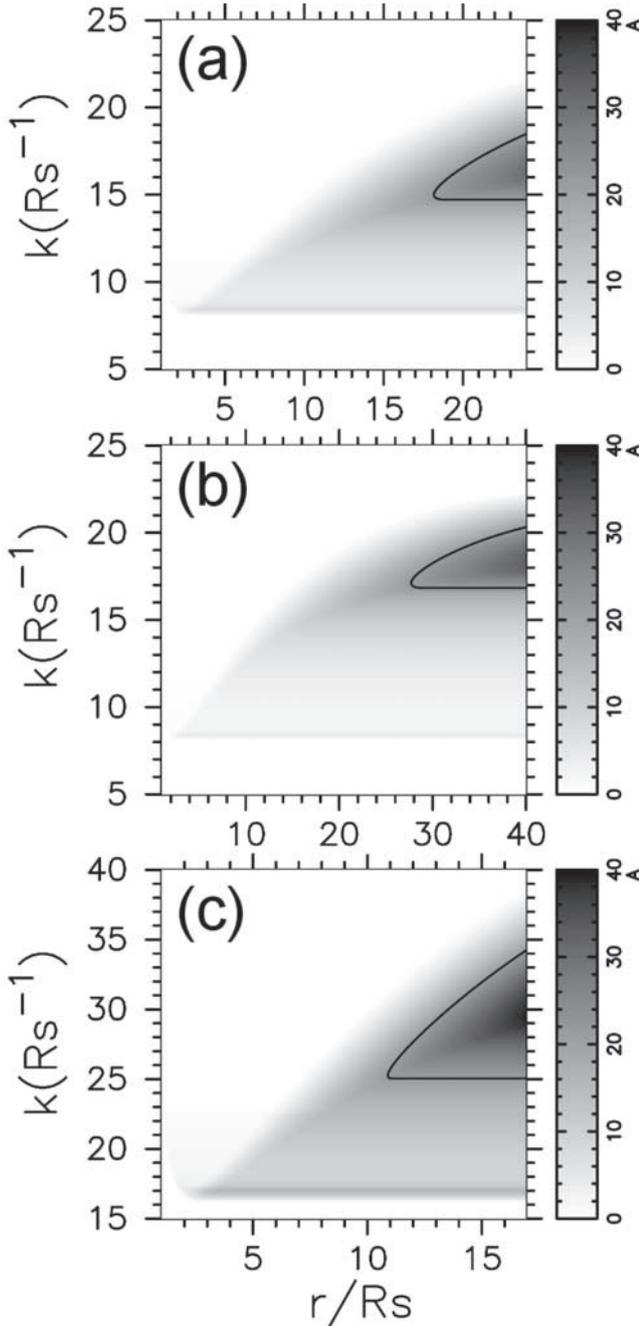
Figure 5. The plot of the amplification factor $A(r, k)$ defined by Equation (10) for (a) Run A, (b) Run B, and (c) Run C. The solid lines indicate the contour line for $A = 21$ which corresponds to the amplitude of the compressive fluctuations seen in Figure 2a.

fluctuation propagates outward at the bulk velocity $v_r(r)$. Here we evaluate the spatial growth of the compressive fluctuation in the solar wind frame by integrating the local growth rate as the fluid element of the solar wind moves from a position r_0 to r with the bulk velocity v_r,

$$A(r,k) \equiv \int_{r_0}^{r} \frac{\gamma(r',k)}{v_r(r')} dr'. \qquad (10)$$

Note that $A(r, k)$ is a non-dimensional value corresponding to an effective amplification factor of the compressive fluctuation for a specific wave number k.

[28] We plot $A(r, k)$ with the gray scale for Runs A-C in Figures 5a–5c. These figures correspond to the spatial evolution of the amplitude of the compressive fluctuation for various wave number in the solar wind frame. We compare the theoretical calculations with the simulation results. The compressive fluctuation has the largest amplitude at $r \sim 18 \; R_S$ as shown in Figure 2a, and this fluctuation has the wave number $k \sim 15 \; R_S^{-1}$ as shown in Figure 3b. The solid line in Figure 5 corresponds to the contour line of $A(r, k) = 21$, which pass through the point $r \sim 18 \; R_S$, $k \sim 15 \; R_S^{-1}$. We expect that a wave mode with and amplification factor $A > 21$ will be dominant due to the spatial growth in the solar wind frame.

[29] Next we compared Figure 4a and Figure 5b for Run B. The solid line in Figure 5b corresponds to the contour line of $A = 21$. This contour line indicates that the unstable mode with the wave number $k \sim 17 \; R_S^{-1}$ grows most rapidly at $r \sim 28 \; R_S$. Figure 4a shows that the compressive fluctuations appear at $r = 28 \sim 34 \; R_S$, which is in agreement with our theoretical analysis. We also compared Figure 4b and Figure 5c for Run C. The solid line in Figure 5c corresponds the contour line of $A = 21$. This contour line indicates that the unstable mode with the wave number $k \sim 25 \; R_S$ grows most rapidly at $r \sim 11 \; R_S$. Figure 4b shows that the compressive fluctuations appear at $r = 10 \sim 14 \; R_S$, which is again in agreement with our theoretical analysis.

[30] From these comparison we found that the profile of $A(k, r)$ can well estimate the wave number of the most rapidly growing compressive mode, where the parametric instability occurs. Note that it is difficult to evaluate the exact amplification factor of the compressive mode because of the spatial variation of the growth rate and the background bulk velocity. However, the simulation results are qualitatively explained by the effective amplification factor obtained by the local linear dispersion analysis with the WKB approximation.

5. Conclusion

[31] We have performed MHD simulations on the propagation of Alfvén waves in the solar wind. Circularly polarized monochromatic Alfvén waves have been injected from the lower boundary into the transonic wind solution. Simulation results have shows that the injected Alfvén waves are subject to the parametric decay, exciting outward-propagating sound waves and inward-propagating Alfvén waves. The compressive fluctuations grow most rapidly at a certain position in the simulation domain. To estimate the most unstable position, we analyzed the local dispersion relation for the parametric decay instability with the WKB approximation. We computed the effective amplification factor with the variation of the local growth rate due to the background bulk velocity, which is in good agreement with the simulation results.

[32] In the present study we have adopted a simplified model in which circularly polarized monochromatic Alfvén waves propagate in the radially expanding transonic flow in a one-dimensional system. However, the present result gives

a quantitative description of nonlinear evolution of Alfvén waves in the solar wind. Better understanding on the propagation of Alfvén waves in the solar wind would contribute the space weather modeling research in associated with the acceleration and heating processes of the solar wind. For a more realistic description of Alfvén waves propagation in the solar wind, we need to introduce a realistic energy equation and multidimensional effects. These are left as future works.

[33] Kinetic effects are also important. They can change the growth rate of unstable modes [*Gomberoff*, 2000; *Gomberoff et al.*, 2001], which would modify the location and wave number of the most rapidly growing compressive mode. Even with the damping effects of ion acoustic waves, however, we think that the decay instability can occur and that the compressive mode can grow at a certain position as discussed above. Full kinetic effects for parametric decay instability are important issues.

[35] Amitava Bhattacharjee thanks Luis Gomberoff and G. Lakhina for their assistance in evaluating this paper.

References

Agim, Y. Z., A. F. Viñas, and M. L. Goldstein (1995), Magnetohydrodynamic and hybrid simulations of broadband fluctuations near interplanetary shocks, *J. Geophys. Res.*, *100*, 17,081–17,106, doi:10.1029/95JA00527.

An, C.-H., S. T. Suess, R. L. Moore, and Z. E. Musielak (1990), Reflection and trapping of Alfven waves in a spherically symmetric stellar atmosphere, *Astrophys. J.*, *350*, 309–323, doi:10.1086/168384.

Bavassano, B., E. Pietropaolo, and R. Bruno (2000), On the evolution of outward and inward Alfvénic fluctuations in the polar wind, *J. Geophys. Res.*, *105*, 15,959–15,964, doi:10.1029/1999JA000276.

Belcher, J. W., and L. Davis (1971), Large amplitude Alfvén waves in the interplanetary medium, *J. Geophys. Res.*, *76*, 3534–3563.

Buti, B., V. Jayanti, A. F. Viñas, S. Ghosh, M. L. Goldstein, D. A. Roberts, G. S. Lakhina, and B. T. Tsurutani (1998), Nonlinear evolution of Alfvénic wave packets, *Geophys. Res. Lett.*, *25*, 2377–2380, doi:10.1029/98GL01688.

Buti, B., V. L. Galinski, V. I. Shevchenko, G. S. Lakhina, B. T. Tsurutani, B. E. Goldstein, P. Diamond, and M. V. Medvedev (1999), Evolution of nonlinear Alfvén waves in streaming inhomogeneous plasmas, *Astrophys. J.*, *523*, 849–854, doi:10.1086/307743.

Cranmer, S. R., and A. A. Ballegooijen (2005), On the generation, propagation, and reflection of Alfvén waves from the solar photosphere to the distant heliosphere, *Astrophys. J.*, *156*, 265–293, doi:10.1086/426507.

Del Zanna, L., M. Velli, and P. Londrillo (2001), Parametric decay of circularly polarized Alfvén waves: Multidimensional simulations in periodic and open domains, *Astron. Astrophys.*, *367*, 705–718, doi:10.1051/0004-6361:20000455.

Derby, N. F., Jr. (1978), Modulational instability of finite-amplitude, circularly polarized Alfven waves, *Astrophys. J.*, *224*, 1013–1016, doi:10.1086/156451.

Fukuda, N., and T. Hanawa (1999), Gravitational and parametric instabilities of the interstellar medium in which the Alfvén wave travels, *Astrophys. J.*, *517*, 226–241, doi:10.1086/307169.

Ghosh, S., and M. L. Goldstein (1994), Nonlinear evolution of a large-amplitude circularly polarized Alfven wave: Low beta, *J. Geophys. Res.*, *99*, 13,351–13,362.

Ghosh, S., A. F. Vinas, and M. L. Goldstein (1994), Nonlinear evolution of a large-amplitude circularly polarized Alfven wave: High beta, *J. Geophys. Res.*, *99*, 19,289–19,300.

Goldstein, M. L. (1978), An instability of finite amplitude circularly polarized Alfven waves, *Astrophys. J.*, *219*, 700–704, doi:10.1086/155829.

Gomberoff, L. (2000), Ion acoustic damping effects on parametric decays of Alfvén waves, *J. Geophys. Res.*, *105*, 10,509–10,518, doi:10.1029/1999JA000281.

Gomberoff, L., K. Gomberoff, and A. L. Brinca (2001), Ion acoustic damping effects on parametric decays of Alfvén waves: Right-hand polarization, *J. Geophys. Res.*, *106*, 18,713–18,720, doi:10.1029/2000JA000384.

Hawley, J. F., and J. M. Stone (1995), MOCCT: A numerical technique for astrophysical MHD, *Comput. Phys. Commun.*, *89*, 127–148, doi:10.1016/0010-4655 (95)00190-Q.

Hollweg, J. V. (1971), Density fluctuations driven by Alfvén waves, *J. Geophys. Res.*, *76*, 5155–5161.

Jayanti, V., and J. V. Hollweg (1993), On the dispersion relations for parametric instabilities of parallel-propagating Alfvén waves, *J. Geophys. Res.*, *98*, 13,247–13,252, doi:10.1029/93JA00920.

Lau, Y.-T., and E. Siregar (1996), Nonlinear Alfven wave propagation in the solar wind, *Astrophys. J.*, *465*, 451–461, doi:10.1086/177432.

Linker, J. A., G. van Hoven, and D. D. Schnack (1990), A three-dimensional simulation of a coronal streamer, *Geophys Res. Lett.*, *17*, 2281–2284.

Linker, J. A., et al. (1999), Magnetohydrodynamic modeling of the solar corona during Whole Sun Month, *J. Geophys. Res.*, *104*, 9809–9830, doi:10.1029/1998JA900159.

Malara, F., L. Primavera, and P. Veltri (2000), Nonlinear evolution of parametric instability of a large-amplitude nonmonochromatic Alfvén wave, *Phys. Plasmas*, *7*, 2866–2877, doi:10.1063/1.874136.

Ofman, L. (2004), Three-fluid model of the heating and acceleration of the fast solar wind, *J. Geophys. Res.*, *109*, A07102, doi:10.1029/2003JA010221.

Ofman, L., and J. M. Davila (1998), Solar wind acceleration by large-amplitude nonlinear waves: Parametric study, *J. Geophys. Res.*, *103*, 23,677–23,690, doi:10.1029/98JA01996.

Parhi, S., B. P. Pandey, M. Goossens, G. S. Lakhina, and P. de Bruyne (1997), MHD study of coronal waves: A numerical approach, *Astrophys. Space Sci.*, *250*, 147–162, doi:10.1023/A:1000444410913.

Parker, E. N. (1963), *Interplanetary Dynamical Processes*, Wiley Intersci., New York.

Pruneti, F., and M. Velli (1997), Parametric Decay of Large Amplitude Alfven Waves in the Solar Atmosphere, in *Eur. Space Agency, Spec. Publ.-404: Fifth SOHO Workshop: The Corona and Solar Wind Near Minimum Activity*, edited by A. Wilson, pp. 623–627.

Sagdeev, R. Z., and A. A. Galeev (1969), *Nonlinear Plasma Theory*, Nonlinear Plasma Theory, Benjamin, New York.

Steinolfson, R. S., and Y. Nakagawa (1976), Dynamical response of the solar corona. II. Numerical simulations near the sun, *Astrophys. J.*, *207*, 300–307, doi:10.1086/154494.

Sturrock, P. A. (1999), Chromospheric magnetic reconnection and its possible relationship to coronal heating, *Astrophys. J.*, *521*, 451–459, doi:10.1086/307544.

Suzuki, T. K., and S.-I. Inutsuka (2006), Solar winds driven by nonlinear low-frequency Alfvén waves from the photosphere: Parametric study for fast/slow winds and disappearance of solar winds, *J. Geophys. Res.*, *111*, A06101, doi:10.1029/2005JA011502.

Tanaka, T. (1994), Finite volume TVD scheme on an unstructured grid system for three-dimensional MHD simulation of inhomogeneous systems including strong background potential fields, *J. Comput. Phys.*, *111*, 381–390.

Turkmani, R., and U. Torkelsson (2004), Dynamics of circularly polarised Alfvén waves in a stratified medium, *Astron. Astrophys.*, *428*, 227–233, doi:10.1051/0004-6361:20041071.

―――――――――――

T. Ogino, S. Tanaka, and T. Umeda, Solar-Terrestrial Environment Laboratory, Nagoya University, Nagoya, Aichi 464-8601, Japan. (ogino@stelab.nagoya-u.ac.jp; stanaka@stelab.nagoya-u.ac.jp; umeda@stelab.nagoya-u.ac.jp)

Nonlinear force-free modeling of the solar coronal magnetic field

T. Wiegelmann[1]

[1]   The coronal magnetic field is an important quantity because the magnetic field dominates the structure of the solar corona. Unfortunately, direct measurements of coronal magnetic fields are usually not available. The photospheric magnetic field is measured routinely with vector magnetographs. These photospheric measurements are extrapolated into the solar corona. The extrapolated coronal magnetic field depends on assumptions regarding the coronal plasma, for example, force-freeness. Force-free means that all nonmagnetic forces like pressure gradients and gravity are neglected. This approach is well justified in the solar corona owing to the low plasma beta. One has to take care, however, about ambiguities, noise and nonmagnetic forces in the photosphere, where the magnetic field vector is measured. Here we review different numerical methods for a nonlinear force-free coronal magnetic field extrapolation: Grad-Rubin codes, upward integration method, MHD relaxation, optimization, and the boundary element approach. We briefly discuss the main features of the different methods and concentrate mainly on recently developed new codes.

Citation: Wiegelmann, T. (2008), Nonlinear force-free modeling of the solar coronal magnetic field, *J. Geophys. Res.*, *113*, A03S02, doi:10.1029/2007JA012432.

1.   Introduction

1.1.   How to Obtain the Coronal Magnetic Field?

[2]   Information regarding the coronal magnetic field is important for space weather application like the onset of flares and coronal mass ejections (CMEs). Unfortunately, we usually cannot measure the coronal magnetic field directly, although recently some progress has been made [see, e.g., *Judge*, 1998; *Solanki et al.*, 2003; *Lin et al.*, 2004]. Because of the optically thin coronal plasma, direct measurements of the coronal magnetic field have a line-of-sight integrated character and to derive the accurate 3D structure of the coronal magnetic field a vector tomographic inversion is required. Corresponding feasibility studies based on coronal Zeeman and Hanle effect measurements have recently been done by *Kramar et al.* [2006] and *Kramar and Inhester* [2006]. These direct measurements are only available for a few individual cases and usually one has to extrapolate the coronal magnetic field from photospheric magnetic measurements. To do so, one has to make assumptions regarding the coronal plasma. It is helpful that the low solar corona is strongly dominated by the coronal magnetic field and the magnetic pressure is orders of magnitude higher than the plasma pressure. The quotient of plasma pressure p and magnetic pressure, $B^2/(2\,\mu_0)$ is small compared to unity ($\beta = 2\,\mu_0\,p/B^2 \ll 1$). In lowest-order

nonmagnetic forces like pressure gradient and gravity can be neglected which leads to the force-free assumption. Force-free fields are characterized by the equations

$$\mathbf{j} \times \mathbf{B} = 0, \tag{1}$$

$$\nabla \times \mathbf{B} = \mu_0\mathbf{j}, \tag{2}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{3}$$

where $\mathbf{B}$ is the magnetic field, $\mathbf{j}$ the electric current density and μ_0 the permeability of vacuum. Equation (1) implies that for force-free fields the current density and the magnetic field are parallel, i.e.,

$$\mu_0\mathbf{j} = \alpha\mathbf{B}, \tag{4}$$

or by replacing $\mathbf{j}$ with equation (2),

$$\nabla \times \mathbf{B} = \alpha\mathbf{B}, \tag{5}$$

where α is called the force-free function. To get some insights in the structure of the space-dependent function α, we take the divergence of equation (4) and make use of equations (2) and (3),

$$\mathbf{B} \cdot \nabla\alpha = 0, \tag{6}$$

which tells us that the force-free function α is constant on every field line, but will usually change from one field line

---

[1]Max-Planck-Institut für Sonnensystemforschung, Katlenburg-Lindau, Germany.

to another. This generic case is called nonlinear force-free approach.

[3] Popular simplifications are $\alpha = 0$ (current free potential fields [see, e.g., *Schmidt*, 1964; *Semel*, 1967; *Schatten et al.*, 1969; *Sakurai*, 1982]) and $\alpha = $ constant (linear force-free approach [see, e.g., *Nakagawa and Raadu*, 1972; *Chiu and Hilton*, 1977; *Seehafer*, 1978; *Alissandrakis*, 1981; *Seehafer*, 1982; *Semel*, 1988]). These simplified models have been in particular popular owing to their relative mathematical simplicity and because only line-of-sight photospheric magnetic field measurements are required. Linear force-free fields still contain one free global parameter α, which can be derived by comparing coronal images with projections of magnetic field lines [e.g., *Carcedo et al.*, 2003]. It is also possible to derive an averaged value of α from transverse photospheric magnetic field measurements [e.g., *Pevtsov et al.*, 1994; *Wheatland*, 1999; *Leka and Skumanich*, 1999]. Despite the popularity and frequent use of these simplified models in the past, there are several limitations in these models (see below) which ask for considering the more sophisticated nonlinear force-free approach.

[4] Our aim is to review recent developments of the extrapolation of nonlinear force-free fields (NLFFF). For earlier reviews on force-free fields we refer to *Sakurai* [1989], *Aly* [1989], *Amari et al.* [1997], *McClymont et al.* [1997], and chapter 5 of *Aschwanden* [2005]. Here we will concentrate mainly on new developments which took place after these earlier reviews. Our main emphasis is to study methods which extrapolate the coronal magnetic field from photospheric vector magnetograms. Several vector magnetographs are currently operating or planed for the nearest future, for example, ground based: the solar flare telescope/ NAOJ [*Sakurai et al.*, 1995], the imaging vector magnetograph/MEES Observatory [*Mickey et al.*, 1996], Big Bear Solar Observatory, Infrared Polarimeter VTT, SOLIS/NSO [*Henney et al.*, 2006] and space born: Hinode/SOT [*Shimizu*, 2004], SDO/HMI [*Borrero et al.*, 2006]. Measurements from these vector magnetograms will provide us eventually with the magnetic field vector on the photosphere, say B_{z0} for the normal and B_{x0} and B_{y0} for the transverse field. Deriving these quantities from the measurements is an involved physical process, which includes measurements based on the Zeeman and Hanle effect, the inversion of Stokes profiles [e.g., *LaBonte et al.*, 1999] and removing the 180 ambiguity [e.g., *Metcalf*, 1994; *Metcalf et al.*, 2006] of the horizontal magnetic field component. Special care has to be taken for vector magnetograph measurements which are not close to the solar disk, when the line-of-sight and normal magnetic field component are far apart [e.g., *Gary and Hagyard*, 1990]. For the purpose of this paper we do not address the observational methods and recent developments and problems related to deriving the photospheric magnetic field vector. We rather will concentrate on how to use the photospheric B_{x0}, B_{y0} and B_{z0} to derive the coronal magnetic field.

[5] The transverse photospheric magnetic field (B_{x0}, B_{y0}) can be used to approximate the normal electric current distribution by

$$\mu_0 j_{z0} = \frac{\partial B_{y0}}{\partial x} - \frac{\partial B_{x0}}{\partial y}, \tag{7}$$

and from this one gets the distribution of α on the photosphere by

$$\alpha(x, y) = \mu_0 \frac{j_{z0}}{B_{z0}}. \tag{8}$$

[6] By using equation (8) one has to keep in mind that rather large uncertainties in the transverse field component and numerical derivations used in (7) can cumulate in significant errors for the current density. The problem becomes even more severe by using (8) to compute α in regions with a low normal magnetic field strength B_{z0}. Special care has to be taken at photospheric polarity inversion lines, i.e., lines along which $B_{z0} = 0$ [see, e.g., *Cuperman et al.*, 1991]. The nonlinear force-free coronal magnetic field extrapolation is a boundary value problem. As we will see later some of the NLFFF codes make use of (8) to specify the boundary conditions while other methods use the photospheric magnetic field vector more directly to extrapolate the field into the corona.

[7] Pure mathematical investigations of the nonlinear force-free equations [see, e.g., *Aly*, 1984; *Boulmezaoud and Amari*, 2000; *Aly*, 2005] and modeling approaches not based on vector magnetograms are important and occasionally mentioned in this paper. A detailed review of these topics is well outside the scope of this paper, however. Some of the model approaches not based on vector magnetograms are occasionally used to test the nonlinear force-free extrapolation codes described here.

1.2. Why Do We Need Nonlinear Force-Free Fields?

[8] 1. A comparison of global potential magnetic field models with TRACE images by *Schrijver et al.* [2005] revealed that significant nonpotentially occurs regularly in active regions, in particular when new flux has emerged in or close to the regions.

[9] 2. Usually α changes in space, even inside one active region. This can be seen, if we try to fit for the optimal linear force-free parameter α by comparing field lines with coronal plasma structures. An example is given by *Wiegelmann and Neukirch* [2002] where stereoscopic reconstructed loops by *Aschwanden et al.* [1999] have been compared with a linear force-free field model. The optimal value of α changes even sign within the investigated active regions, which is a contradiction to the $\alpha = $ constant linear force-free approach (see Figure 1).

[10] 3. Photospheric α distributions derived from vector magnetic field measurements by equation (8) show as well that α usually changes within an active region [see, e.g., *Régnier et al.*, 2002].

[11] 4. Potential and linear force-free fields are too simple to estimate the free magnetic energy and magnetic topology accurately. The magnetic energy of linear force-free fields is unbounded in a halfspace [*Seehafer*, 1978] which makes this approach unsuitable for energy approximations of the coronal magnetic field. Potential fields have a minimum energy for an observed line-of-sight photospheric magnetic field. An estimate of the excess of energy a configuration has above that of a potential field is an important quantity which might help to understand the onset of flares and coronal mass ejections.
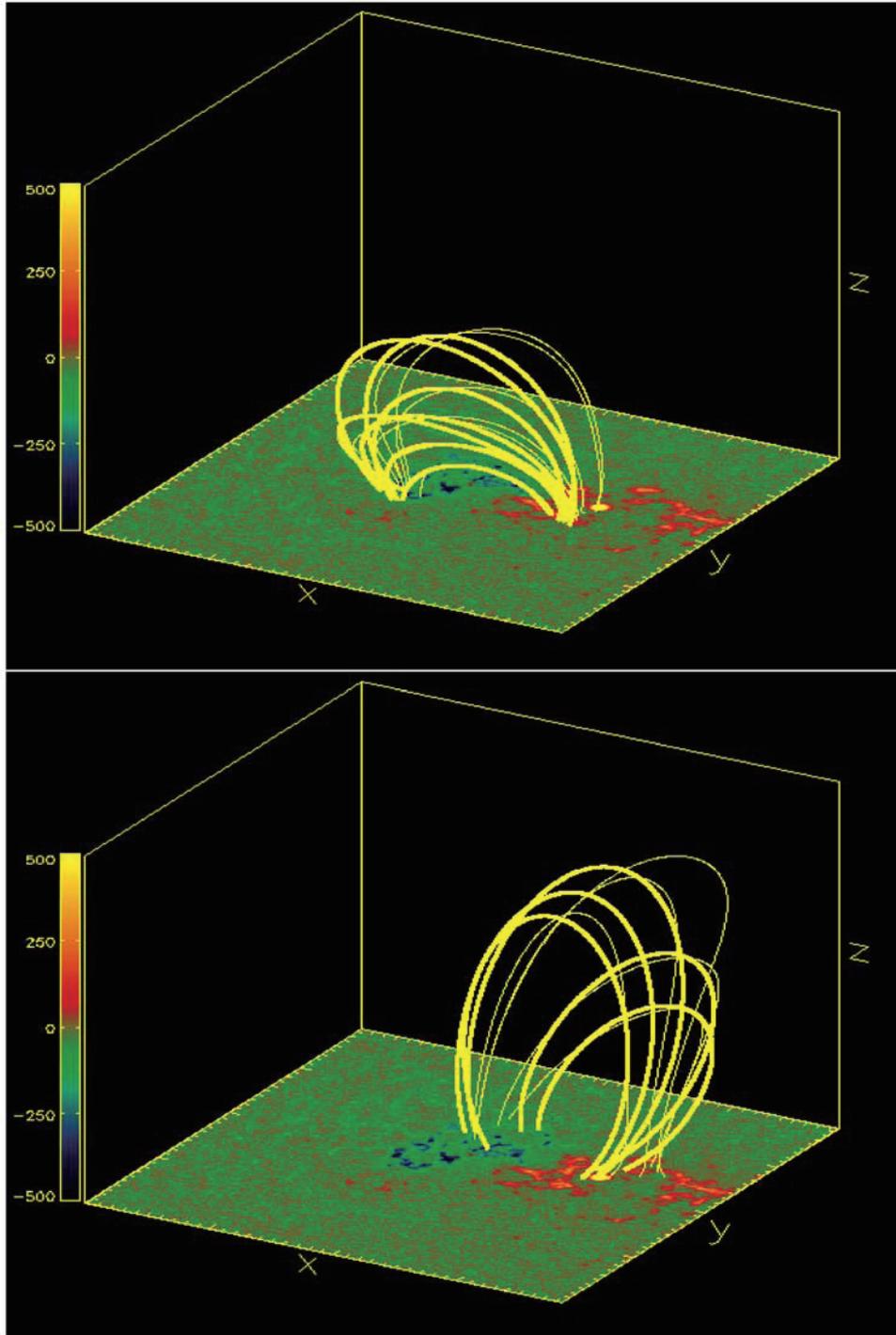
Figure 1. Linear force−free field model for NOAA 7986 with the best fit for α. (top) A group of loops with $\alpha = 2.5$ and (bottom) another group of loops with $\alpha = -2.0$. The different optimal values of the linear force-free parameter within one active region are a contradiction to the linear assumption (α constant) and tell us that a consistent modeling of this active regions requires nonlinear force-free approach. (This figure was originally published as *Wiegelmann and Neukirch*'s [2002] Figure 7). Used with permission of Springer.).

[12] 5. A direct comparison of measured fields in a newly developed active region by *Solanki et al.* [2003] with extrapolations from the photosphere with a potential, linear and nonlinear force-free model by *Wiegelmann et al.* [2005b] showed that nonlinear fields are more accurate than simpler models. Figure 2 shows some selected magnetic field lines for the original measured field and extrapolations from the photosphere with the help of a potential, linear and nonlinear force-free model.
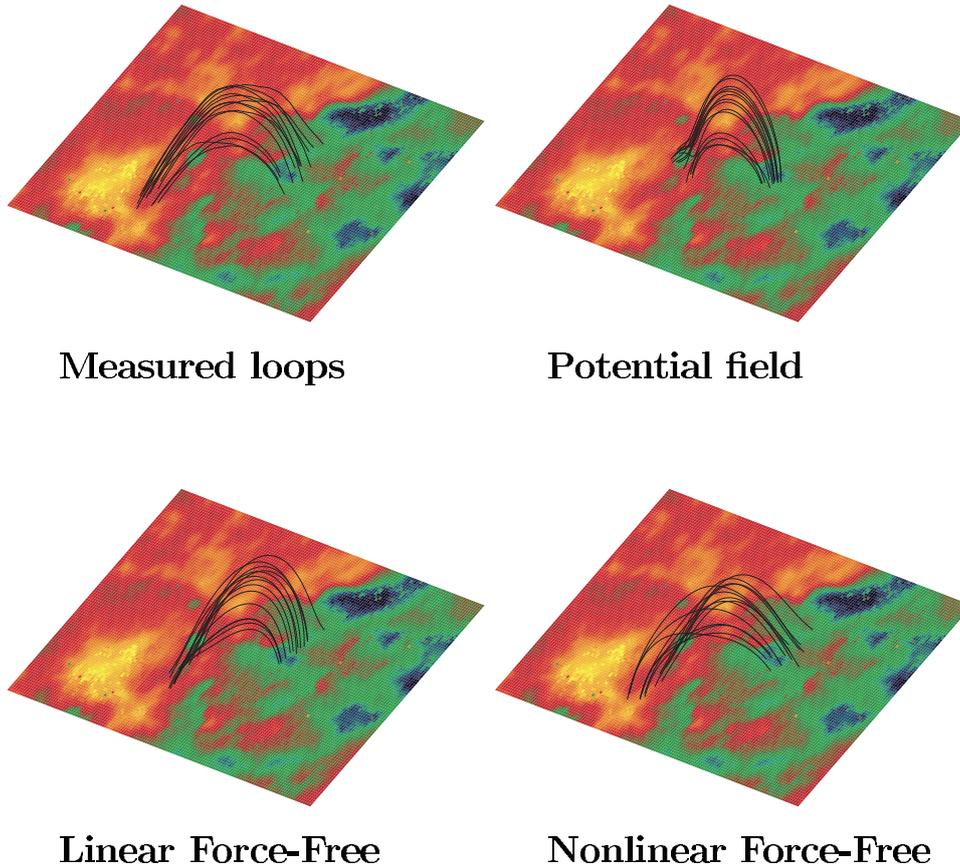
Figure 2. Magnetic field structure of the newly developed active region NOAA 9451. Direct measurements of the field have been compared with potential, linear, and nonlinear force-free models. Best agreement has been found for the nonlinear model. (This figure was originally published as part of *Wiegelmann et al.*'s [2005b] Figure 1. Used with permission of *Astronomy and Astrophysics*).

[13] These points tell us that nonlinear force-free modeling is required for an accurate reconstruction of the coronal magnetic field. Simpler models have been used frequently in the past. Global potential fields provide some information of the coronal magnetic field structure already, for example, the location of coronal holes. The generic case of force-free coronal magnetic field models are nonlinear force-free fields, however. Under generic we understand that α can (and usually will) change in space, but this approach also includes the special cases α = constant and α = 0. Some active regions just happen to be more potential (or linear force-free) and if this is the case they can be described with simpler models. Linear force-free models might provide a rough estimate of the true 3D magnetic field structure if the nonlinearity is weak. The use of simpler models was often justified owing to limited observational data, in particular if only the line-of-sight photospheric magnetic field has been measured.

[14] While the assumption of nonlinear force-free fields is well accepted for the coronal magnetic fields in active regions, this is not true for the photosphere. The photospheric plasma is a finite β plasma and nonmagnetic forces like pressure gradient and gravity cannot be neglected here. As a result electric currents have a component perpendicular to the magnetic field, which contradicts the force-free assumption. We will discuss later how these difficulties can be overcome.

2. Nonlinear Force-Free Codes

[15] Different methods have been proposed to extrapolate nonlinear force-free fields from photospheric vector magnetic field measurements. (1) The Grad-Rubin method was proposed for fusion plasmas by *Grad and Rubin* [1958] and first applied to coronal magnetic fields by *Sakurai* [1981]. (2) The upward integration method was proposed by *Nakagawa* [1974] and encoded by *Wu et al.* [1985]. (3) The MHD relaxation method was proposed for general MHD equilibria by *Chodura and Schlueter* [1981] and applied to force-free coronal magnetic fields by *Mikic and McClymont* [1994]. (4) The optimization approach was developed by *Wheatland et al.* [2000]. (5) The boundary element (or Greens function like) method was developed by *Yan and Sakurai* [2000].

2.1. Grad-Rubin Method

[16] The Grad-Rubin method reformulates the nonlinear force-free equations in such a way, that one has to solve a well posed boundary value problem. This makes this approach also interesting for a mathematical investigation of the structure of the nonlinear force-free equations. *Bineau*

[1972] demonstrated that the used boundary conditions (vertical magnetic field on the photosphere and α distribution at one polarity) ensure, at least for small values of α and weak nonlinearities the existence of a unique nonlinear force-free solution. A detailed analysis of the mathematical problem of existence and uniqueness of nonlinear force-free fields is outside the scope of this review and is given, for example, by *Amari et al.* [1997, 2006].

[17] The method first computes a potential field, which can be obtained from the observed line-of-sight photospheric magnetic field (say B_z in Cartesian geometry) by different methods, for example, a Greens function method as described by *Aly* [1989]. It is also popular to use linear force-free solvers, for example, as implemented by *Seehafer* [1978] and *Alissandrakis* [1981] with the linear force-free parameter $\alpha = 0$ to compute the initial potential field. The transverse component of the measured magnetic field is then used to compute the distribution of α on the photosphere by equation (8). While α is described this way on the entire photosphere, for both polarities, a well posed boundary value problem requires that the α distribution becomes only described for one polarity. The basic idea is to iteratively calculate α for a given **B** field from (6), then calculate the current via (4) and finally update **B** from the Biot-Savart problem (5). These processes are repeated until the full current as prescribed by the α distribution has been injected into the magnetic field and the updated magnetic field configuration becomes stationary in the sense that eventually the recalculation of the magnetic field with Amperes law does not change the configuration anymore. To our knowledge the Grad-Rubin approach has been first implemented by *Sakurai* [1981]. Here α has been prescribed on several nodal points along a number of magnetic field lines of the initial potential field. The method used a finite-element-like discretization of current tubes associated with magnetic field lines. Each current tube was divided into elementary current tubes of cylindrical shape. The magnetic field is updated with Ampere's law using a superposition of the elementary current tubes. The method was limited by the number of current-carrying field lines, nodal points and the corresponding number of nonlinear equations (N^9) to solve with the available computer resources more than a quarter century ago.

[18] Computer resources have increased rapidly since the first NLFFF implementation by *Sakurai* [1981] and about a decade ago the Grad-Rubin method has been implemented on a finite difference grid by *Amari et al.* [1997, 1999]. This approach decomposes equations (1)–(3) into a hyperbolic part for evolving α along the magnetic field lines and an elliptic one to iterate the updated magnetic field from Amperes law. For every iteration step k one has to solve iteratively for

$$\mathbf{B}^{(k)} \cdot \nabla \alpha^{(k)} = 0 \tag{9}$$

$$\alpha^{(k)} \mid_{S\pm} = \alpha_{0\pm}, \tag{10}$$

which evolves α in the volume and

$$\nabla \times \mathbf{B}^{(k+1)} = \alpha^{(k)} \mathbf{B}^{(k)}, \tag{11}$$

$$\nabla \cdot \mathbf{B}^{(k+1)} = 0, \tag{12}$$

$$B_z^{(k+1)} \mid_{S\pm} = B_{z0}, \tag{13}$$

$$\lim_{|r| \to \infty} \mid \mathbf{B}^{(k+1)} \mid = 0, \tag{14}$$

where $\alpha_{0\pm}$ corresponds to the photospheric distribution of α for either on the positive or the negative polarity. The Grad-Rubin method as described by *Amari et al.* [1997, 1999] has been applied to investigate particular active regions by *Bleybel et al.* [2002] and a comparison of the extrapolated field with 2D projections of plasma structures as seen in Hα, EUV and X ray has been done by *Régnier et al.* [2002] and *Régnier and Amari* [2004]. The code has also been used to investigate mutual and self helicity in active regions by *Régnier et al.* [2005] and to flaring active regions by *Régnier and Canfield* [2006].

[19] A similar approach as done by *Sakurai* [1981] has been implemented by *Wheatland* [2004]. The implemented method computes the magnetic field directly on the numerical grid from Ampere's law. This is somewhat simpler and faster as Sakurai's approach which required solving a large system of nonlinear equations for this aim. The implementation by *Wheatland* [2004] has, in particular, been developed with the aim of parallelization. The parallelization approach seems to be effective owing to a limited number of interprocess communications. This is possible because as the result of the linearity of Ampere's law the contributions of the different current carrying field lines are basically independent from each other. In the original paper Wheatland reported problems for large currents on the field lines. These problems have been related to an error in current representation of the code and the corrected code worked significantly better [see also *Schrijver et al.*, 2006]. The method has been further developed by *Wheatland* [2006]. This newest Wheatland-implementation scales with the number of grid points N^4 for a N^3 volume, rather than N^6 for the earlier [*Wheatland*, 2004] implementation. The main new development is a faster implementation of the current-field iteration. To do so the magnetic field has been separated into a current-free and a current carrying part at each iteration step. Both parts are solved using a discrete Fast Fourier Transformation, which imposes the required boundary conditions implicitly. The code has been parallelized on shared memory distributions with OpenMP.

[20] *Amari et al.* [2006] developed two new versions of their Grad-Rubin code. The first version is a finite difference method and the code was called 'XTRAPOL.' This code prescribe the coronal magnetic field with the help of a vector potential **A**. The code has obviously it's heritage from the earlier implementation of *Amari et al.* [1999], but with several remarkable differences:

[21] 1. The code includes a divergence cleaning routine, which takes care about $\nabla \cdot \mathbf{A} = 0$. The condition $\nabla \cdot \mathbf{A} = 0$ is fulfilled with high accuracy in the new code 10^{-9} compared to 10^{-2} in the earlier implementation.

[22] 2. The lateral and top boundaries are more flexible compared to the earlier implementation and allow a finite B_n

and non zero α values for one polarity on all boundaries. This treats the whole boundary (all six faces) as a whole.

[23] 3. The slow current input as reported for the earlier implementation, which lead to a two level iteration, has been replaced. Now the whole current is injected at once and only the inner iteration loop of the earlier code remained in the new version.

[24] 4. The computation of the α characteristics has been improved with an adaptive Adams-Bashforth integration scheme (see *Press* [2002] for details).

[25] 5. The fixed number of iteration loops have been replaced by a quantitative convergence criterium.

[26] In the same paper [*Amari et al.*, 2006] introduced another Grad-Rubin approach based on finite elements, which they called 'FEMQ'. Different from alternative implementation this code does not use a vector potential but iterates the coupled divergence and curl system, which is solved with the help of a finite element discretization. The method transforms the nonlinear force-free equations into a global linear algebraic system.

[27] *Inhester and Wiegelmann* [2006] implemented a Grad-Rubin code on a finite element grid with staggered field components [see *Yee*, 1966] which uses discrete Whitney forms [*Bossavit*, 1988]. Whitney forms allow to transform standard vector analysis (as the differential operators gradient, curl and divergence) consistently into the discrete space used for numerical computations. Whitney forms contain four types of finite elements (form $0-3$). They can be considered as a discrete approximation of differential forms. The finite element base may consist of polynomials of any order. In its simplest form, the 0-forms have as parameters the function values at the vertices of the cells and are linearly interpolated within each cell. The 1-forms are a discrete representation of a vector field defined on the cell edges. The 2-forms are defined as the field component normal to the surfaces of the cells. The 3-forms are finite volume elements for a scalar function approximation, which represents the average of a scalar over the entire cell. The 4-forms are related to each other by GRAD (0 to 1 form), CURL (1 to 2 form) and DIV (2 to 3 form). As for continuous differential forms, double differentiation (CURL of GRAD, DIV of CURL) give exactly zero, independent of the numerical precision. A dual grid, shifted by half a grid size in each axis, was introduced in order to allow for Laplacians. Whitney forms on the dual grid are related to forms on the primary grid in a consistent way.

[28] The Grad-Rubin implementation uses a vector potential representation of the magnetic field, where the vector potential is updated with a Poisson equation in each iteration step. The Poisson equation is effectively solved with the help of a multigrid solver. The main computing time is spend to distribute α along the field lines with (6). This seems to be a general property of Grad-Rubin implementations. One can estimate the scaling of (6) by $\propto N^4$, where the number of field lines to compute is $\propto N^3$ and the length of a field line $\propto N$. The Biot-Savart step (5) solved with FFT or multigrid methods scales only with $\propto N \log N$. Empirical tests show that the number of iteration steps until a stationary state is reached does not depend on the number of grid points N for Grad-Rubin solvers. We have explained before, that the Grad-Rubin implementation requires the prescription of α only for one polarity to have a well posed

mathematical problem. The *Inhester and Wiegelmann* [2006] implementation allows these choice of boundary conditions as a special case. In general one does not need to make the distinction between $(\partial V)^+$ and $(\partial V)^-$ in the new implementation. A well posed mathematical problem is still ensured, however, in the following way. Each boundary value of α is attached with a weight. The final version of α on each field line is then determined by a weighted average of the α values on both endpoints of a field lines. By this way the influence of uncertain boundary values, for example, on the side walls and imprecise photospheric measurements can be suppressed.

2.2. Upward Integration Method

[29] The basic equations for the upward integration method (or progressive extension method) have been published already by *Nakagawa* [1974] and a corresponding code has been developed by *Wu et al.* [1985, 1990a]. The upward integration method is a straight forward approach to use the nonlinear force-free equations directly to extrapolate the photospheric magnetic field into the corona. To do so one reformulates the force-free equations (1)–(3) in order to extrapolate the measured photospheric magnetic field vector into the solar corona.

[30] As a first step the magnetic field vector on the lower boundary $\mathbf{B}_0(x,y,0)$ is used to compute the z component of the electric current $\mu_0 j_{z0}$ with equation (7) and the photospheric α distribution (say α_0) by equation (8). With the help of equation (4) we calculate the x and y component of the current density

$$\mu_0 j_{x0} = \alpha_0 B_{x0} \tag{15}$$

$$\mu_0 j_{y0} = \alpha_0 B_{y0}. \tag{16}$$

[31] We now use equation (3) and the x and y component of equation (2) to obtain expressions for the z derivatives of all three magnetic field components in the form

$$\frac{\partial B_{x0}}{\partial z} = \mu_0 j_{y0} + \frac{\partial B_{z0}}{\partial x}, \tag{17}$$

$$\frac{\partial B_{y0}}{\partial z} = \frac{\partial B_{z0}}{\partial y} - \mu_0 j_{x0}, \tag{18}$$

$$\frac{\partial B_{z0}}{\partial z} = -\frac{\partial B_{x0}}{\partial x} - \frac{\partial B_{y0}}{\partial y}. \tag{19}$$

The idea is to integrate this set of equations numerically upward in z by repeating the previous steps at each height. As a result we get in principle the 3D magnetic field vector in the corona. While this approach is straight forward, easy to implement and computational fast (no iteration is required), a serious drawback is that it is unstable. Several authors [e.g., *Cuperman et al.*, 1990; *Amari et al.*, 1997] pointed out that the formulation of the force-free equations in this way is unstable because it is based on an ill-posed mathematical problem. In particular one finds that expo-

nential growth of the magnetic field with increasing height is a typical behavior. What makes this boundary value problem ill-posed is that the solution does not depend continuously on the boundary data. Small changes or inaccuracies in the measured boundary data lead to a divergent extrapolated field (see *Low and Lou* [1990] for a more detailed discussion). As pointed out by Low and Lou meaningful boundary conditions are required also on the outer boundaries of the computational domain. It is also possible to prescribe open boundaries in the sense that the magnetic field vanishes at infinity. This causes an additional problem for the upward integration method, because the method transports information only from the photosphere upward and does not incorporate boundary information on other boundaries or at infinity. Attempts have been made to regularize the method [e.g., *Cuperman et al.*, 1991; *Demoulin and Priest*, 1992], but cannot be considered as fully successful.

[32] *Wu et al.* [1990b] compared the Grad-Rubin method in the implementation of *Sakurai* [1981] with the upward integration method in the implementation of *Wu et al.* [1990a]. (The authors used a somewhat different nomenclature: The upward integration method was called 'progressive extension method' and the Grad-Rubin method 'iterative method'. That time the term 'iterative method' was reasonable because Grad-Rubin was the only iterative approach available, but now, 17 years later, several other iterative methods are available to compute nonlinear force-free fields.) The comparison showed qualitatively similar results for extrapolations from an observed magnetogram, but quantitatively differences. The NLFFF computations have been very similar to potential field extrapolations, however, too. One reason for this behavior was, that the method of *Sakurai* [1981] is limited to small values of α and a 'by eye' comparison shows that the corresponding NLFFF field is very close to a potential field configuration. The field computed with the upward integration method deteriorated if the height of the extrapolation exceeded a typical horizontal scale length.

[33] The upward integration method has been recently reexamined by *Song et al.* [2006] who developed a new formulation of this approach. The new implementation uses smooth continuous functions and the equations are solved in asymptotic manner iteratively. The original upward integration equations are reformulated into a set of ordinary differential equations and uniqueness of the solution seems to be guaranteed at least locally. While *Demoulin and Priest* [1992] stated that 'no further improvement has been obtained with other types of smoothing functions' *Song et al.* [2006] point out that the transformation of the original partial differential equations into ordinary ones eliminates the growing modes in the upward integration method, which have been reported before by *Wu et al.* [1990a] and subsequent papers. The problem that all three components of the photospheric magnetic field and the photospheric α distribution has to be prescribed in a consistent way remains in principle, but some compatibility conditions to compute a slowly varying α have been provided by *Song et al.* [2006]. These compatibility conditions are slightly different for real photospheric observations and tests with smooth boundaries extracted from semianalytic equilibria. For the latter kind of problems the new formulation provided reasonable results

with the standard test equilibrium found by *Low and Lou* [1990]. The method seems to be also reasonably fast. Of course, further tests with more sophisticated equilibria and real data are necessary to evaluate this approach in more detail.

2.3. MHD Relaxation

[34] MHD relaxation codes [e.g., *Chodura and Schlueter*, 1981] can be applied to solve nonlinear force-free fields as well. The idea is to start with a suitable magnetic field which is not in equilibrium and to relax it into a force-free state. This is done by using the MHD equations in the following form:

$$\upsilon\mathbf{v} = (\nabla \times \mathbf{B}) \times \mathbf{B}, \tag{20}$$

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0, \tag{21}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \tag{22}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{23}$$

where υ is a viscosity and $\mathbf{E}$ the electric field. As the MHD relaxation aims for a quasiphysical temporal evolution of the magnetic field from a nonequilibrium toward a (nonlinear force-free) equilibrium this method is also called 'evolutionary method' or 'magneto-frictional method'. The basic idea is that the velocity field in the equation of motion (21) is reduced during the relaxation process. Ideal Ohm's law (22) ensures that the magnetic connectivity remains unchanged during the relaxation. The artificial viscosity v plays the role of a relaxation coefficient which can be chosen in such way that it accelerates the approach to the equilibrium state. A typical choice is

$$\upsilon = \frac{1}{\mu} \mid \mathbf{B} \mid^2, \tag{24}$$

with μ = constant. Combining equations (20), (21), (22) and (24) we get an equation for the evolution of the magnetic field during the relaxation process,

$$\frac{\partial \mathbf{B}}{\partial t} = \mu \mathbf{F}_{MHD}, \tag{25}$$

with

$$\mathbf{F}_{MHD} = \nabla \times \left(\frac{[(\nabla \times \mathbf{B}) \times \mathbf{B}] \times \mathbf{B}}{B^2} \right). \tag{26}$$

This equation is then solved numerically starting with a given initial condition for $\mathbf{B}$, usually a potential field. Equation (25) ensures that equation (23) is satisfied during the relaxation if the initial magnetic field satisfies it. (As we will see below the 'optimization' approach leads to a similar iteration equations for the magnetic field, but a different artificial driving force $\mathbf{F}$.) The difficulty with this method is that it cannot be guaranteed that for given boundary

conditions and initial magnetic field (i.e., given connectivity), a smooth force-free equilibrium exists to which the system can relax. If such a smooth equilibrium does not exist the formation of current sheets is to be expected which will lead to numerical difficulties. Therefore care has to be taken when choosing an initial magnetic field.

[35] *Yang et al.* [1986] developed a magneto frictional method which represent the magnetic field with the help of Euler (or Clebsch) potentials.

$$\mathbf{B} = \nabla g \times \nabla h, \qquad (27)$$

where the potentials g and h are scalar functions. The general method has been developed for three dimensional fields and iterative equations for $g(x, y, z)$ and $h(x, y, z)$ have been derived. The Clebsch representation automatically ensures $\nabla \cdot \mathbf{B} = 0$. The method has been explicitly tested in the paper by [*Yang et al.*, 1986] with the help of an equilibrium with one invariant coordinate. In principle it should be possible to use this representation for the extrapolation of nonlinear force-free fields, but we are not aware of a corresponding implementation. On the basis of the discussion by *Yang et al.* [1986] a difficulty seems to be that one needs to specify boundary conditions for the potentials, rather than for the magnetic fields. It seems in particular to be difficult to find boundaries conditions for potentials which correspond to the transverse component of the photospheric magnetic field vector. One problem is that boundary conditions for g and h prescribe the connectivity. Every field line can be labeled by its (g,h) values. Hence boundary values for g and h establish foot point relations although the field is not known yet.

[36] The MHD relaxation (or evolutionary) method has been implemented by *Mikic and McClymont* [1994] and *McClymont et al.* [1997] on the basis of the time-dependent MHD code by *Mikic et al.* [1988]. The code uses a nonuniform mesh and the region of interested is embedded in a large computational domain to reduce the influence of the lateral boundaries. The method has been applied to extrapolate the magnetic field above an active region by *Jiao et al.* [1997]. The computations have been carried out with a resolution of the order of 100^3 points. A supercomputer was required for these computations that time (10 years ago), but because of the rapid increase of computer speed and memory within the last decade this restriction is very probably not valid anymore.

[37] *Roumeliotis* [1996] developed the so-called stress and relax method. In this approach the initial potential field becomes disturbed by the observed transverse field component on the photosphere. The boundary conditions are replaced in subsequently in several small steps and always relaxed with a similar MHD relaxation scheme as described above toward a force-free equilibrium. The code by *Roumeliotis* [1996] has implemented a function $w(x,y)$ which allows to give a lower weight to regions where the transverse photospheric field has been measured with lower accuracy. Additional to the iterative equations as discussed above, the method includes a resistivity η (or diffusivity) by adding a term $\eta \mathbf{j}$ on the right hand site of Ohms law (21). This relaxes somewhat the topological constrains of ideal MHD relaxation, because a finite resistivity allows a kind of

artificial reconnection and corresponding changes of the initial potential field topology. The method has been tested with a force-free equilibrium found by *Klimchuk and Sturrock* [1992] and applied to an active region measured with the MSFC vector-magnetograph.

[38] The stress and relax method has been revisited by *Valori et al.* [2005]. Different from the earlier implementation by *Roumeliotis* [1996] the new implementation uses directly the magnetic field, rather than the vector potential in order to keep errors from taking numerical deviations from noisy magnetograms minimal. The solenoidal condition is controlled by a diffusive approach by *Dedner et al.* [2002] which removes effectively a numerically created finite divergence of the relaxed magnetic field. The new implementation uses a single stress step, rather than the multiple small stress used by *Roumeliotis* [1996] to speed up the computation. The single step stress and relax method is connected with a suitable control of artificial plasma flows by the Courant criterium. The authors reported that a multistep and single-step implementation do not reveal significant differences. The numerical implementation is based on the time-dependent full MHD code 'AMRVAC' by *Keppens et al.* [2003]. *Valori et al.* [2005] tested their nonlinear force-free implementation with a numerically constructed nonlinear force-free twisted loop computed by *Török and Kliem* [2003].

2.4. Optimization Approach

[39] The optimization approach has been developed by *Wheatland et al.* [2000]. The solution is found by minimizing the functional

$$L = \int_V [B^{-2} \mid (\nabla \times \mathbf{B}) \times \mathbf{B} \mid^2 + \mid \nabla \cdot \mathbf{B} \mid^2] d^3 V. \qquad (28)$$

Obviously, L is bound from below by 0. This bound is attained if the magnetic field satisfies the force-free equations (1)–(3).

[40] By taking the functional derivatives with respect to some iteration parameter t we get

$$\Rightarrow \frac{1}{2} \frac{dL}{dt} = -\int_V \frac{\partial \mathbf{B}}{\partial t} \cdot \tilde{\mathbf{F}} d^3 x - \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \tilde{\mathbf{G}} d^2 x, \qquad (29)$$

with

$$\mathbf{F} = \nabla \times \left(\frac{[(\nabla \times \mathbf{B}) \times \mathbf{B}] \times \mathbf{B}}{B^2} \right) + \left\{ -\nabla \times \left(\frac{((\nabla \cdot \mathbf{B})\mathbf{B}) \times \mathbf{B}}{B^2} \right) \right.$$
$$\left. - \Omega \times (\nabla \times \mathbf{B}) - \nabla(\Omega \cdot \mathbf{B}) + \Omega(\nabla \cdot \mathbf{B}) + \Omega^2 \mathbf{B} \right\} \qquad (30)$$

$$\Omega = \mathbf{B}^{-2}[(\nabla \times \mathbf{B}) \times \mathbf{B} - (\nabla \cdot \mathbf{B})\mathbf{B}]. \qquad (31)$$

The surface term vanishes if the magnetic field vector is kept constant on the surface, for example, prescribed from photospheric measurements. In this case L decreases monotonically if the magnetic field is iterated by

$$\frac{\partial \mathbf{B}}{\partial t} = \mu \mathbf{F}. \qquad (32)$$

[41] Let us remark that $\mathbf{F}_{\text{MHS}}$ as defined in equation (26) and used for MHD relaxation is identical with the first term on the right-hand-side of equation (30), but equation (30) contains additional terms.

[42] For this method the vector field $\mathbf{B}$ is not necessarily solenoidal during the computation, but will be divergence-free if the optimal state with $L = 0$ is reached. A disadvantage of the method is that it cannot be guaranteed that this optimal state is indeed reached for a given initial field and boundary conditions. If this is not the case then the resulting $\mathbf{B}$ will either be not force-free or not solenoidal or both.

[43] McTiernan has implemented the optimization approach basically as described by *Wheatland et al.* [2000] in IDL (see *Schrijver et al.* [2006] for a brief description of the McTiernan implementation). This code allows the use of a nonuniform computational grid. In a code intercomparison by *Schrijver et al.* [2006] the IDL optimization code by McTiernan was about a factor of 50 slower compared to an implementation in parallelized C by *Wiegelmann* [2004]. To our knowledge McTiernan has translated his IDL code into FORTRAN in the meantime for faster computation (J. M. McTiernan, personal communication on the NLFFF workshop Palo Alto, June 2006 [see also *Metcalf et al.*, 2007]).

[44] Several tests have been performed with the optimization approach of *Wiegelmann and Neukirch* [2003]. It has been investigated how the unknown lateral and top boundary influence the solution. The original optimization approach by *Wheatland et al.* [2000] has been extended toward more flexible boundary conditions, which allow $\frac{\partial B}{\partial t} \neq 0$ on the lateral and top boundaries. This has been made with the help of the surface integral term in (29) and led to an additional term $\frac{\partial B}{\partial t} = \mu \mathbf{G}$ on the boundaries. This approached improved the performance of the code for cases where only the bottom boundary was prescribed. No improvement was found for a slow multistep replacement of the boundary and this possibility has been abandoned in favor of a single step method. It has been also investigated how noise influences the optimization code and this study revealed that noise in the vector magnetograms leads to less accurate nonlinear force-free fields.

[45] *Wiegelmann* [2004] has reformulated the optimization principle by introducing weighting functions One defines the functional

$$L = \int_V [wB^{-2} \mid (\nabla \times \mathbf{B}) \times \mathbf{B} \mid^2 + w \mid \nabla \cdot \mathbf{B} \mid^2] d^3x, \qquad (33)$$

where $w(x, y, z)$ is a weighting function. It is obvious that (for $w > 0$) the force-free equations $(1-3)$ are fulfilled when L is equal zero. Minimization of the functional (34) lead to

$$\frac{\partial \mathbf{B}}{\partial t} = \mu \tilde{\mathbf{F}}, \qquad (34)$$

$$\tilde{\mathbf{F}} = w\mathbf{F} + (\Omega_a \times \mathbf{B}) \times \nabla w + (\Omega_b \cdot \mathbf{B})\nabla w, \qquad (35)$$

$$\Omega_a = B^{-2}[(\nabla \times \mathbf{B}) \times \mathbf{B}], \qquad (36)$$

$$\Omega_b = B^{-2}[(\nabla \cdot \mathbf{B})\mathbf{B}], \qquad (37)$$

with $\mathbf{F}$ as defined in (30). With $w(x, y, z) = 1$ this approach reduces to the *Wheatland et al.* [2000] method as described above. The weighting function is useful if only the bottom boundary data are known. In this case we a buffer boundary of several grid points toward the lateral and top boundary of the computational box is introduced. The weighting function is chosen constant in the inner, physical domain and drop to 0 with a cosine profile in the buffer boundary toward the lateral and top boundary of the computational box. In the work of *Schrijver et al.* [2006] some tests have been made with different weighting functions for the force-free and solenoidal part of the functional (34), but the best results have been obtained if both terms got the same weight. The computational implementation involves the following steps.

[46] 1. Compute start equilibrium (e.g., a potential field) in the computational box.

[47] 2. Replace the bottom boundary with the vector magnetogram.

[48] 3. Minimize the functional (33) with the help of equation (34). The continuous form of (34) guaranties a monotonically decreasing L. This is as well ensured in the discretized form if the iteration step dt is sufficiently small. The code checks if $L(t+dt) < L(t)$ after each time step. If the condition is not fulfilled, the iteration step is repeated with dt reduced by a factor of 2. After each successful iteration step we increase dt slowly by a factor of 1.01 to allow the time step to become as large as possible with respect to the stability condition.

[49] 4. The iteration stops if L becomes stationary. Stationarity is assumed if $\frac{\partial L}{\partial t}/L < 1.0 \ 10^{-4}$ for 100 consecutive iteration steps.

[50] The program has been tested with the semianalytic nonlinear force-free configuration of *Low and Lou* [1990] and *Titov and Démoulin* [1999] by *Wiegelmann et al.* [2006a]. The code has been applied to extrapolate the coronal magnetic field in active regions by *Wiegelmann et al.* [2005b, 2005a].

[51] A finite element optimization approach has been implemented by *Inhester and Wiegelmann* [2006] using the Whitney elements as for the Grad-Rubin code (which has been described above). The optimization method uses exactly the same staggered finite element grid as described above, which is different from the finite difference grids used in the earlier implementations by *Wheatland et al.* [2000], *Wiegelmann and Neukirch* [2003], and *Wiegelmann* [2004]. Another difference is that earlier implementations discretized the analytical derivative of the functional L (28), while the new code takes the numerical more consistent derivative of the discretized function L. All other implementations used a simple Landweber scheme for updating the magnetic field, which is replaced here by an unpreconditioned conjugate gradient iteration, which at every time step performs an exact line search to the minimum of L in the current search direction and additional selects an improve search direction instead of the gradient of the functional L. To do so the Hessian matrix of the functional L is computed during every iteration step. An effective computation of the Hessian matrix is possible, because the reformulated function $L(s)$ is a fourth-order polynomial in $\mathbf{B}$ and all five polynomial coefficients can be computed in one go. The code has been tested by *Low and Lou* [1990] and the

result of twisted loop computations of the Grad-Rubin implementation on the same grid.

[52] The optimization code in the implementation of *Wiegelmann* [2004] has recently been extended toward using a multiscale implementation. The main difference from the original code are [see also *Metcalf et al.*, 2007] as follows. (1) The method is not full multigrid, but computes the solution on different grids only once, for example, something like 50^3, 100^3, 200^3. (2) The main idea is to get a better (than potential field) start equilibrium on the full resolution box. (3) Solution of smaller grids are interpolated onto larger grids as initial state for the magnetic field in the computational domain of the next larger box. The multiscale implementation has been tested as part of a code-intercomparison test by *Metcalf et al.* [2007] with the help of solar-like reference model computed by *van Ballegooijen* [2004] and *van Ballegooijen et al.* [2007].

[53] The optimization approach has recently been implemented in spherical geometry by *Wiegelmann* [2007] and tested by *Low and Lou* [1990]. The original longitudinal symmetric Low and Lou solution has been shifted by 1/4 of a solar radius to test the code without any symmetry with respect to the Suns surface. The numerical implementation is very similar as the Cartesian implementation described by *Wiegelmann* [2004]. The spherical implementation converged fast for low-latitude regions, but the computing time increased significantly if polar regions have been included. It has been suggested to implement the code on a so called 'Yin and Yang' grid as developed by *Kageyama and Sato* [2004] to reduce the computing time. The 'Yin and Yang' grid is suitable for massive parallelization, which is necessary for full-sphere high-resolution NLFFF computations.

2.5. Boundary Element or Greens-Function-Like Method

[54] The boundary integral method has been developed by *Yan and Sakurai* [2000]. The method relates the measured boundary values with the nonlinear force-free field in the entire volume by

$$c_i \mathbf{B_i} = \oint_S \left(\bar{\mathbf{Y}} \frac{\partial \mathbf{B}}{\partial n} - \frac{\partial \bar{\mathbf{Y}}}{\partial n} \mathbf{B}_0 \right) \mathbf{dS}, \qquad (38)$$

where $c_i = 1$ for points in the volume and $c_i = 1/2$ for boundary points and $\mathbf{B}_0$ is the measured vector magnetic field on the photosphere. The auxiliary vector function is defined as

$$\bar{\mathbf{Y}} = \mathrm{diag} \left(\frac{\cos(\lambda_x r)}{4\pi r}, \frac{\cos(\lambda_y r)}{4\pi r}, \frac{\cos(\lambda_z r)}{4\pi r} \right), \qquad (39)$$

and the λ_i, ($i = x, y, z$) are computed in the original approach by *Yan and Sakurai* [2000] with integrals over the whole volume, which define the λ_i implicitly,

$$\int_V Y_i [\lambda_i^2 B_i - \alpha^2 B_i - (\nabla \alpha \times \mathbf{B}_i)] dV = 0. \qquad (40)$$

This volume integration, which has to be carried out for every point in the volume is certainly very time consuming (a sixth-order process). The λ_i have the same dimension as

the magnetic field. The existence of the λ_i has been confirmed for the semianalytic field of *Low and Lou* [1990] by *Li et al.* [2004]. While the work of *Li et al.* [2004] showed that one can find the auxiliary function $\bar{\mathbf{Y}}$ for a given force-free field in 3D, the difficulty is that $\bar{\mathbf{Y}}$ is a priori unknown if only the photospheric magnetic field vector is given. *Yan and Sakurai* [2000] proposed an iterative scheme to compute the auxiliary functions and the nonlinear force-free magnetic field self-consistently. They use the approximate solution k on the right-hand side of equation (38) to compute a better solution $k + 1$ by

$$c_i \mathbf{B_i}^{(k+1)} = \oint_S \left(\bar{\mathbf{Y}}^{(\mathbf{k})} \frac{\partial \mathbf{B}^{(\mathbf{k})}}{\partial n} - \frac{\partial \bar{\mathbf{Y}}^{(\mathbf{k})}}{\partial n} \mathbf{B}_0 \right) \mathbf{dS}, \qquad (41)$$

where the initial guess for the magnetic field in the volume is $\mathbf{B} = 0$ and also the initial $\frac{\partial \bar{\mathbf{Y}}}{\partial n} = 0$. In principle it would be also possible to compute a potential field first and derive the auxiliary functions for this field as done by *Li et al.* [2004] and iterate subsequently for the nonlinear force-free fields and the associated auxiliary functions with equation (41). This possibility has not been tried out to our knowledge until now, however. The method iterates the magnetic field until $\mathbf{B}$ and $\frac{\partial \mathbf{B}}{\partial n}$ converge. In an inter code comparison by *Schrijver et al.* [2006] one iteration step of (41) took about 80 hours for this method and only this one step was carried out without further iteration. This seems, however, not to be sufficient to derive an accurate nonlinear force-free solution. The method has been applied for the comparison with soft X-ray loops observed with YOHKOH by *Wang et al.* [2000] and *Liu et al.* [2002] and to model a magnetic flux robe by *Yan et al.* [2001a, 2001b].

[55] In a new implementation of the boundary element method by *Yan and Li* [2006] the auxiliary functions are computed iteratively with the help of a simplex method. This avoids the numerical expensive computation of the volume integral (40). The boundary element method is still rather slow if a magnetic field has to be computed in an entire 3D domain. Different from other method, it allows, however, to evaluate the NLFFF field at every arbitrary point within the domain from the boundary data, without the requirement to compute the field in an entire domain. This is in particular useful if one is interested to compute the NLFFF field only along a given loop.

[56] *He and Wang* [2006] investigated the validity of the boundary integral representation for a spherical implementation. The method has been tested with the longitudinal invariant [*Low and Lou*, 1990] solution. The spherical implementation method of this method revealed reasonable results for smooth modestly nonlinear fields, but a poor convergence for complex magnetic field structures and large values of α.

3. How to Deal With Non-Force-Free Boundaries and Noise?

[57] Given arbitrary boundary conditions of the magnetic field vector on the photosphere, the solution to the force-free equations in 3D may not exist. Nonlinear force-free coronal magnetic field models assume, however, that the solution exists. It is certainly possible and necessary to

check after or during the computation if a solution has been found. In the following we will discuss what we can do if the measured photospheric data are incompatible with the assumption of a force-free coronal magnetic field.

3.1. Consistency Check of Vector Magnetograms

[58] We reexamine some necessary conditions with the photospheric field (or bottom boundary of a computational box). These conditions have to be fulfilled in order to be suitable boundary conditions for a nonlinear force-free coronal magnetic field extrapolation. An a priori assumption about the photospheric data is that the magnetic flux from the photosphere is sufficiently distant from the lateral boundaries of the observational domain and the net flux is in balance, i.e.,

$$\int_S B_z(x,y,0)dxdy = 0. \tag{42}$$

Molodensky [1969, 1974], *Aly* [1989], and *Sakurai* [1989] used the viral theorem to define which conditions a vector magnetogram has to fulfill to be consistent with the assumption of a force-free field in the corona above the boundary. These conditions are as follows.

[59] 1. The total force on the boundary vanishes

$$\int_S B_x\,B_z dxdy = \int_S B_y\,B_z dxdy = 0 \tag{43}$$

$$\int_S (B_x^2 + B_y^2)dxdy = \int_S B_z^2 dxdy. \tag{44}$$

[60] 2. The total torque on the boundary vanishes

$$\int_S x(B_x^2 + B_y^2)dxdy = \int_S xB_z^2 dxdy, \tag{45}$$

$$\int_S y(B_x^2 + B_y^2)dxdy = \int_S yB_z^2 dxdy, \tag{46}$$

$$\int_S yB_x B_z dxdy = \int_S xB_y B_z dxdy. \tag{47}$$

[61] In an earlier review, *Aly* [1989] has mentioned already that the magnetic field is probably not force-free in the photosphere, where **B** is measured because the plasma β in the photosphere is of the order of 1 and pressure and gravity forces are not negligible. The integral relations (43)–(47) are not satisfied in this case in the photosphere and the measured photospheric field is not a suitable boundary condition for a force-free extrapolation. *Metcalf et al.* [1995] concluded that the solar magnetic field is not force-free in the photosphere, but becomes force-free only at about 400 km above the photosphere. *Gary* [2001] pointed out that care has to be taken when extrapolating the coronal magnetic field as a force-free field from photospheric measurements, because the force-free low corona is sandwiched between two regions (photosphere and higher

corona) with a plasma $\beta \approx 1$, where the force-free assumption might break down. An additional problem is that measurements of the photospheric magnetic vector field contain inconsistencies and noise. In particular the transverse components (say B_x and B_y) of current vector magnetographs include uncertainties.

[62] The force-free field in a domain requires the Maxwell stress (43)–(47) to sum to zero over the boundary. If these conditions are not fulfilled a force-free field cannot be found in the volume. A faithful algorithm should therefore have the capability of rejecting a prescription of the vector field at the boundary that fails to produce zero net Maxwell stress. A simple way to incorporate these conditions would be to evaluate the integrals (43)–(47) within or prior to the NLFFF computation and to refuse the vector field if the conditions are not fulfilled with sufficient accuracy. Current codes do run, however, although if fed with inconsistent boundary data, but they certainly cannot find a force-free solution in this case (because it does not exist). This property of current codes does, however, not challenge the trustworthiness of the algorithms, because the force-free and solenoidal conditions are checked in 3D, for example, with the help of the functional L as defined in (28). A non zero value of L (within numerical accuracy) tells the user that a force-free state has not been reached. In principle it would be possible that the codes do refuse to output the magnetic field in this case. For current codes this is not automatically controlled but responsibility of the user.

[63] Unfortunately current measurements of the magnetic field vector are only available routinely in the photosphere, where we have a finite β plasma and nonmagnetic forces might become important. The force-free compatibility conditions (43)–(47) are not fulfilled in the photosphere, but they should be fulfilled in the low β chromospheric and coronal plasma above. The question is if we still can use the photospheric measurements to find suitable consistent boundary conditions for a nonlinear force-free modeling. Such an approach has been called preprocessing of vector magnetograms.

3.2. Preprocessing

[64] The preprocessing routine has been developed by *Wiegelmann et al.* [2006b]. The integral relations (43)–(47) have been used to define a 2D functional of quadratic forms:

$$L_{\text{prep}} = \mu_1 L_1 + \mu_2 L_2 + \mu_3 L_3 + \mu_4 L_4, \tag{48}$$

where

$$L_1 = \left[\left(\sum_p B_x\,B_z \right)^2 + \left(\sum_p B_y\,B_z \right)^2 + \left(\sum_p B_z^2 - B_x^2 - B_y^2 \right)^2 \right], \tag{49}$$

$$L_2 = \left[\left(\sum_p x(B_z^2 - B_x^2 - B_y^2) \right)^2 + \left(\sum_p y(B_z^2 - B_x^2 - B_y^2) \right)^2 + \left(\sum_p yB_x\,B_z - xB_y\,B_z \right)^2 \right], \tag{50}$$

$$L_3 = \left[\sum_p (B_x - B_{xobs})^2 + \sum_p (B_y - B_{yobs})^2 + \sum_p (B_z - B_{zobs})^2 \right],$$

(51)

$$L_4 = \left[\sum_p (\Delta B_x)^2 + (\Delta B_y)^2 + (\Delta B_z)^2 \right].$$

(52)

The surface integrals are here replaced by a summation $\sum_p$ over all grid nodes p of the bottom surface grid and the differentiation in the smoothing term is achieved by the usual five-point stencil for the 2D-Laplace operator. Each constraint L_n is weighted by a yet undetermined factor μ_n. The first term ($n = 1$) corresponds to the force-balance conditions (43)–(44), the next ($n = 2$) to the torque-free condition (45)–(47). The following term ($n = 3$) ensures that the optimized boundary condition agrees with the measured photospheric data and the last terms ($n = 4$) controls the smoothing. The 2D-Laplace operator is designated by Δ. The aim of the preprocessing procedure is to minimize L_{prep} so that all terms L_n if possible are made small simultaneously. A strategy on how to find the optimal yet undefined parameters μ_n is described by *Wiegelmann et al.* [2006b]. As result of the preprocessing we get a data set which is consistent with the assumption of a force-free magnetic field in the corona but also as close as possible to the measured data within the noise level.

4. Code Testing and Code Comparisons

4.1. Code Testing

[65] Newly developed codes for the extrapolation of nonlinear force-free fields from boundary data have to be tested before they are applied to measurements. In principle any analytical or numerically created solution of the force-free equations (1)–(3) can be used as a reference case. One cuts a plane (artificial photosphere, bottom boundary) out of the 3D reference solution (for the pure task of code testing it is also acceptable to use all six boundaries of the reference solution; these kind of data are not available for real solar cases of course) and uses the above described extrapolation codes to reconstruct the magnetic field. The result of this extrapolation is then compared with the reference to rate the quality of the reconstruction. Unfortunately, it is very hard to find a truly nonlinear 3D solution of (1)–(3) analytically and very few solutions are known. *Low and Lou* [1990] (LL) found a class of solutions which have become a standard reference for testing NLFFF extrapolation codes. LL found axisymmetric equilibria which are separable in spherical coordinates. They are self-similar in the radial coordinate, and the polar angle dependence is determined from a nonlinear eigenvalue equation. The symmetry is broken by cutting out a rectangular chunk of the solution by using a Cartesian coordinate system which is shifted and rotated with respect to the original coordinate system in which the LL equilibria are calculated. The parameters of the LL solutions and the parameters of the new Cartesian coordinate system allow for a large number of different situations which can be used for tests. The original axisymmetric spherical LL solution has also been

used (with and without symmetry breaking by shifting the origin of the coordinate system) to test spherical NLFFF programs. To our knowledge all recent implementations of the described NLFFF approaches have been tested with LL, either immediately in the original code-describing papers or in subsequent works, for example, in a blind-algorithm test within the NLFFF consortium, as described below.

[66] The MHD relaxation method and the optimization approach have been compared by *Wiegelmann and Neukirch* [2003]. Both methods have been applied to the *Low and Lou* [1990] equilibrium with exactly the same finite difference grid. The iterative equations for MHD relaxation and optimization have both the form $\frac{\partial \mathbf{B}}{\partial t} = \mu \mathbf{F}$ but the structure of $\mathbf{F}$ is more complicated for optimization than for MHD relaxation. The MHD relaxation term is indeed identical with the first term of the optimization approach. While MHD relaxation minimizes only the Lorentz force, the optimization does additional minimize $\nabla \cdot \mathbf{B}$, while a decreasing magnetic field divergence during MHD relaxation [as shown by *Wiegelmann and Neukirch*, 2003] is the result of numerical diffusion. Despite the numerical overhead in computing $\mathbf{F}$ for the optimization code, optimization provided more accurate results and faster convergence.

[67] A practical advantage of the MHD approach is that several time-dependent MHD codes are well known and established and can be used for the force-free relaxation discussed here. The inclusion of nonmagnetic forces like pressure gradients and gravity looks straight forward for the MHD approach. Other methods are usually developed with the only task of computing nonlinear force-free coronal magnetic fields, also a generalization toward magnetohydrostatic and stationary MHD equilibria is possible and has been done for the optimization approach [see *Wiegelmann and Inhester*, 2003; *Wiegelmann and Neukirch*, 2006]. Another advantage of using time-dependent MHD codes for relaxation is that the computed force-free equilibrium can be used on the same grid and with the same code as initial state for time-dependent MHD simulations. One can, in principle, use the force-free equilibria computed with any of the described method as initial state for time-dependent MHD simulation, but having the initial equilibrium state already directly on the MHD grid might be very handy, because no further adjustments are needed.

4.2. NLFFF Consortium

[68] Since the year 2004 activities are ongoing to bring NLFFF modelers together and to compare the different existing codes. A workshop series has been organized for this aim by Karel Schrijver and three workshops took place so far from 2004–2006. The next workshop is planed for June 2007. As we have been asked to summarize the workshop results on the CSWM meeting, we give also a very brief overview in the corresponding special issue paper here. The main results of the first two workshops have been published by *Schrijver et al.* [2006]. In this paper six different NLFFF implementations (Grad-Rubin codes of *Amari et al.* [1999] and *Wheatland* [2004], MHD relaxation code of *Valori et al.* [2005], optimization codes by McTiernan and by *Wiegelmann* [2004], boundary element method by *Yan and Sakurai* [2000]) have been compared. The codes have been tested in a blind algorithm test with the help of the semianalytic equilibrium by *Low and Lou* [1990]

in two cases. In case I all six boundaries of a computational box have been described, and in case II only the bottom boundary has been described. The comparison of the extrapolation results with the reference solution has been done qualitatively by magnetic field line plots (shown here in Figure 3 for the central region of case II) and quantitatively by a number of sophisticated comparison matrices. All NLFFF fields agreed best with the reference field for the low-lying central magnetic field region, where the magnetic field and electric currents are strongest and the influence of the boundaries lowest. The code converged with speeds that differed by a factor of one million per iteration steps. (The codes run on different machines, have been written in different programming languages and used different compilers. A real test of the exact computing time would comprise a proper operation count, for example, the number of fixed point additions and multiplications per iteration step.) The fastest-converging and best-performing code was the *Wheatland et al.* [2000] optimization code as implemented by *Wiegelmann* [2004]. Recent implementations of the Grad-Rubin code by *Amari et al.* [2006] and *Inhester and Wiegelmann* [2006] and a new implementation of the upward integration method by *Song et al.* [2006] did not participate in the blind-algorithm intercomparison by *Schrijver et al.* [2006], but these three new codes have been tested by the authors with similar measures and revealed similar accuracy as the best performing codes in the blind algorithm test. It seems that the somewhat more flexible boundary conditions used in the Grad-Rubin approaches of *Amari et al.* [2006] and *Inhester and Wiegelmann* [2006] are responsible for the better performance compared to the earlier implementation by *Amari et al.* [1999], which has been used in the blind algorithm test.

[69] The widely used LL equilibrium contains a very smooth photospheric magnetic field and an extended current distribution. It is therefore also desirable to test NLFFF codes also with other, more challenging boundary fields, which are less smooth, have localized current distribution and to investigate also the effects of noise and effects from non-force-free boundaries. A somewhat more challenging reference case is the equilibrium found by *Titov and Démoulin* [1999] (TD). Similar to LL, the TD equilibrium is an axisymmetric equilibrium. The TD model contains a potential field which is disturbed by a toroidal nonlinear force-free current. This equilibrium has been used for testing the MHD relaxation code (G. Valori and B. Kliem, personal communication, 2006) and the optimization code from *Wiegelmann et al.* [2006a].

[70] Any numerically created NLFFF model might be suitable for code testing, too. It is in particular interesting to use models, which are partly related on observational data. Very recently, *van Ballegooijen et al.* [2007] used line-of-sight photospheric measurements from SOHO/MDI to compute a potential field, which was then disturbed by inserting a twisted flux robe and relaxed toward a nonlinear force-free state with a magnetofrictional method as described by *van Ballegooijen* [2004]. The *van Ballegooijen et al.* [2007] model is not force-free in the entire computational domain, but only above a certain height above the bottom boundary (artificial chromosphere). On the lowest boundary (photosphere) the model contains significant nonmagnetic forces. Both the chromospheric as well as the photospheric mag-

netic field vector from the *van Ballegooijen et al.* [2007] model have been used to test four of the recently developed extrapolation codes (one Grad-Rubin method, one MHD relaxation code and two optimization approaches) in a second blind algorithm test by *Metcalf et al.* [2007]. While the NLFFF consortium paper, part I [*Schrijver et al.*, 2006], used a domain of just 64^3 pixel, the part II paper used a computational domain of $320 \times 320 \times 258$ pixel and modern NLFFF codes where able to compute the nonlinear force-free field in such relatively large boxes within a few hours for a moderate parallelization on only $1-4$ processors and a memory requirement of $2.5-4$ GB of RAM. This very recent code comparison shows a major improvement regarding computing time and suitable grid sizes within less than 3 years. On the first NLFFF consortium meeting in 2004, box sizes of some 64^3 have been a kind of standard or computing times of some 2 weeks have been reported for 150^3 boxes. We briefly summarize the results of *Metcalf et al.* [2007] as: (1) NLFFF extrapolations from chromospheric data recover the original reference field with high accuracy; (2) when the extrapolations are applied to the photospheric data, the reference field is not well recovered; and (3) preprocessing of the photospheric data improve the result, but the accuracy is still lower as for extrapolations from the chromosphere.

5. Conclusions and Outlook

[71] Within the last few years the scientific community showed a growing interest into coronal magnetic fields. (Publications containing the phrase 'coronal magnetic fields' in title or abstract have been cited less than about 50 times per year until the early 1990s and this number increased to about 150 citations per year in 2004. A peak year was 2006 (last year) with more than 300 citations (source: ISI Web of Knowledge, March 2007)). The development of new ground based and space born vector magnetographs provide us measurements of the magnetic field vector on the suns photosphere. Accompanied from these hardware development, software has been developed to extrapolate the photospheric measurements into the corona. Special attention has recently been given to nonlinear force-free codes. Five different numerical approaches (Grad-Rubin, upward integration, MHD relaxation, optimization, boundary elements) have been developed for this aim. It is remarkable that new codes or major updates of existing codes have been published for all five methods within the last two years, mainly in the last year (2006). A workshop series (NLFFF consortium) since 2004 on nonlinear force-free fields has recently released synergy effects, by bringing modelers of the different numerical implementations together to compare, evaluate and improve the programs. Several of the most recent new codes and utility programs (e.g., preprocessing) have at least been partly inspired by these workshops. The new implementations have been tested with the smooth semianalytic Low-Lou equilibrium and showed reasonable agreement with this reference field. While all methods aim for a reconstruction of the coronal magnetic field from the photospheric magnetic field vector, the way how these measurements are used to prescribe the boundaries of the codes is different.
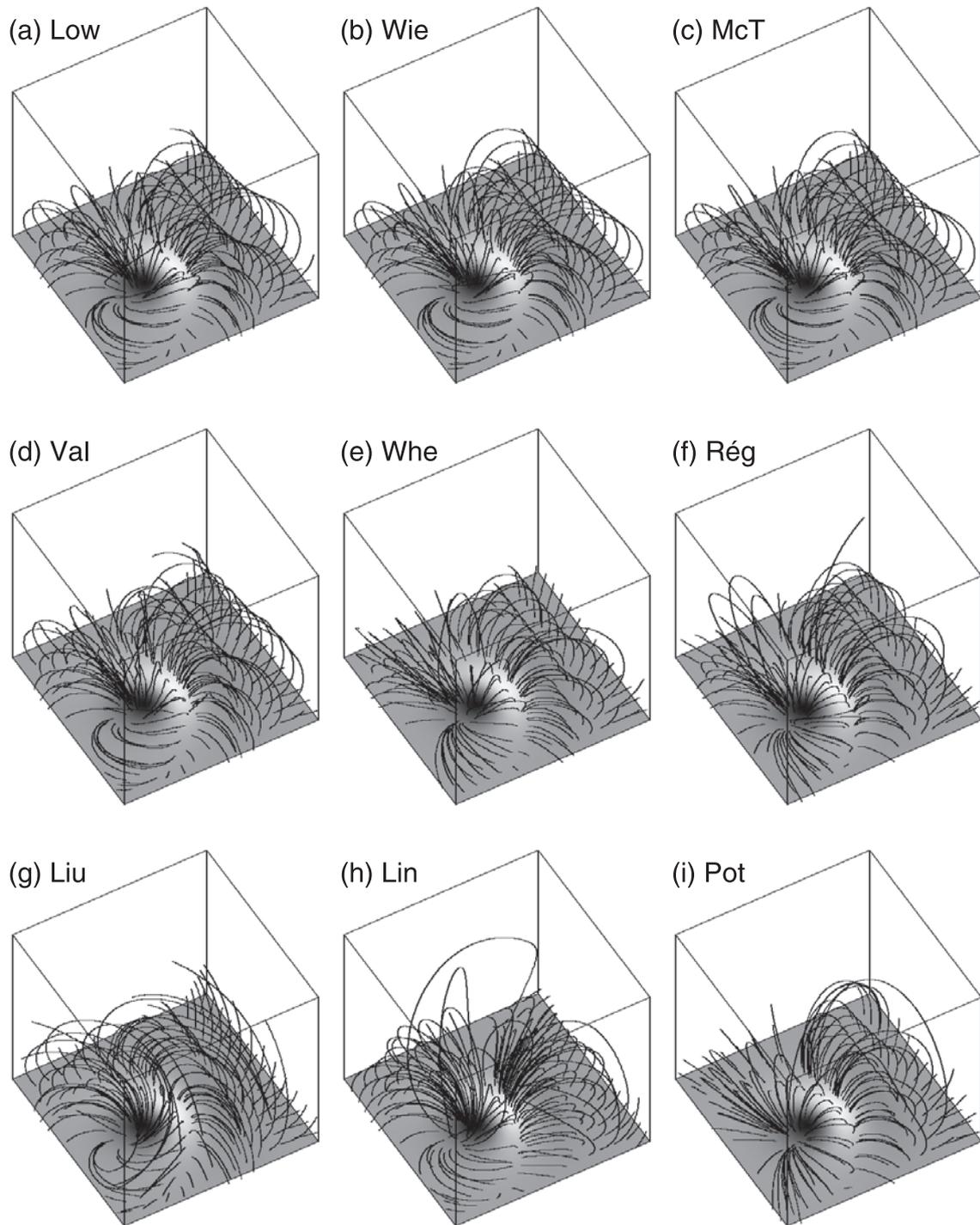
Figure 3. Evaluation of six nonlinear force–free codes. The (a) reference solution has been compared with extrapolations with (b, c) optimation, (d) MHD relaxation, (e, f) Grand–Rubin, and (g) boundary element. For comparison, (h) linear-force-free and (i) potential fields are shown, too. The images show the central domain of the model. Only the bottom boundary has been provided for the extrapolation. (This figure was originally published as *Schrijver et al.*'s [2006] Figure 4. Used with permission of Springer.)

[72]   1. MHD relaxation and optimization use B_{x0}, B_{y0}, B_{z0} on the bottom boundary. This overdetermines the boundary value problem. Both methods are closely related and compute the magnetic field in a computational box with

$$\frac{\partial \mathbf{B}}{\partial t} = \mu \mathbf{F}, \qquad (53)$$

where the structure of $\mathbf{F}$ is somewhat different (the optimization approach has more terms) for both methods. Usually a potential field is used as initial state for both approaches, also the use of a linear force-free initial state is possible. Recently a multiscale version of optimization has been installed, which uses a low-resolution NLFFF field as input for higher-resolution computations. Specifying the entire magnetic field vector on the bottom boundary is an overimposed problem and a unique NLFFF field (or a solution at all) requires that the boundary data fulfill certain consistency criteria. A recently developed preprocessing routine helps to find suitable consistent boundary data from inconsistent photospheric measurements. Earlier and current comparisons showed a somewhat higher accuracy for the optimization approach. A practical advantage of the MHD approach is that in principle any available time-dependent MHD code can be adjusted to compute the NLFFF field.

[73]   2. The Grad-Rubin approach uses B_{z0} and the distribution of α computed with equation (8) for one polarity, which corresponds to well posed mathematical problem. A practical problem is that the computation of α requires numerical differences of the noisy and forced transverse photospheric field B_{x0}, B_{y0} with (7) leading to inaccuracies in the normal electric current distribution and in α. For smooth semianalytic test cases this is certainly not a problem, but real data require special attention (smoothing, preprocessing, limiting $\alpha \neq 0$ to regions where B_{z0} is above a certain limit) to derive a meaningful distribution of α. While the method requires only α for one polarity, the computation from photospheric data provide α for both polarities. We are not aware of any tests on how well NLFFF solutions computed from α prescribed on the positive and negative polarity coincide. It is also unclear how well the computed transverse field components on the bottom boundary agree with the measured values of B_{x0}, B_{y0}. (In principle B_{x0}, B_{y0} may have an additional field (B_{x0}, B_{y0}) + $(\partial_x \partial_y) \varphi$ without making a difference for α and hence for the Grad-Rubin result.) More tests on this topics are necessary, including the recently installed possibility to prescribe α for both polarities and adjust the boundary by a weighed average of α on both polarities to fulfill equation (6). As initial state the Grad-Rubin method uses a potential field, which is also true for MHD relaxation and optimization.

[74]   3. The upward integration and the boundary element method prescribe both all components of the bottom boundary magnetic field vector and the α distribution computed with equation (8). This approach over imposes the boundary and B_{x0}, B_{y0}, B_{z0} and α have to be consistent which each other and the force-free assumption. This is certainly not a problem at all for smooth semianalytic test equilibria and strategies to derive consistent boundary data from measured data have been developed recently. Different from the three approaches discussed above, upward integration and boundary element methods do not require to compute first an initial potential field in the computational domain. It is well known that the upward integration method is based on an ill-posed problem and the method has not been considered for several years, but a recent implementation with smooth analytic functions might help to regularize this method. First tests showed a reasonable results for computations with the smooth semianalytic Low-Lou solution.

[75]   The boundary element method has the problem to be very slow and an earlier implementation of this method could not reach a converged state for a 64^3 boxed used in the NLFFF consortium paper, part I, owing to this problem. A new 'direct boundary method' has been developed, which seems to be faster than the original 'boundary element method', but still slower compared with the four other NLFFF approaches if the task is to compute a 3D magnetic field in an entire 3D domain. Different from all other described methods the boundary element approach allows to compute the nonlinear force-free field vector at any arbitrary point above the boundary and it is not necessary to compute the entire 3D field above the photosphere. This might be a very useful feature if one is interested in computing the magnetic field only along a single loop and not interested in an entire active region.

[76]   The new implementations of upward integration and boundary element method show both reasonable results for first tests with the smooth semianalytic Low and Lou equilibrium. Further tests with more sophisticated equilibria, for example, a solar-like test case as used in the NLFFF consortium paper, part II, would be useful to come to more sound conclusions regarding the feasibility of these methods.

[77]   Most of the efforts done in nonlinear force-free modeling until now concentrated mainly on developing these models and testing their accuracy and speed with the help of well known test configuration. Not too many applications of nonlinear force-free models to real data are currently available, from which we learned new physics. One reason was the insufficient access to high-accuracy photospheric vector magnetograms and a second one were limitations of the models. Force-free field extrapolation is a mere tool, if properly employed on vector magnetograms, it can help to understand physical, magnetic field dominated processes in the corona. Both the computational methods as well as the accuracy of required measurements (e.g., with Hinode, SDO) are rapidly improving. Within the NLFFF consortium we just started (since April 2007) to apply the different codes to compute nonlinear force-free coronal magnetic fields from Hinode vectormagnetograms. This project might provide us already some new insights about coronal physics.

[78]   To conclude, we can say that the capability of Cartesian nonlinear force-free extrapolation codes has rapidly increased in recent years. Only 3 years ago most codes run usually on grids of about 64^3 pixel. Recently developed or updated codes (Grad-Rubin by Wheatland, MHD relaxation by Valori, optimization by Wiegelmann, optimization by McTiernan) have been applied to grids of about 300^3 pixel. Although this increase of traceable grid sizes is certainly encouraging, the resolution of current and near-

future vector magnetographs (which of course measure only data in 2D!) is significantly higher. We should keep in mind, however, that the currently implemented NLFFF codes have been only moderately parallelized using only a few processors. The CSWM conference, where this paper has been presented, took place at the 'Earth simulator' in Yokohama, which contains several thousands of processors used for Earth-science computer simulations. An installation of NLFFF codes on such massive parallel computers (which has been briefly addressed on NLFFF consortium meetings) combined with adaptive mesh refinements might enable drastically improved grid sizes. One should not underestimate the time and effort necessary to program and install such massive parallelized versions of existing codes. As full disk vectormagnetograms will become available soon (SOLIS, SDO/HMI) it is also an important task to take a spherical geometry into account. First steps in this direction have been carried out with the optimization and boundary element methods. Spherical NLFFF geometries are currently still in it's infancy and have been tested until now only with smooth semianalytic Low and Lou equilibria and require further developments.

[79] Attention has also recently been drawn to the problem that the coronal magnetic field is force-free, but the photospheric one is not. Tests with extrapolations from solar-like artificial photospheric and chromospheric measurements within the NLFFF consortium paper, part II, revealed that extrapolations from the (force-free) chromospheric field provide significantly better results as extrapolations using directly the (forced) photospheric field. Applying a preprocessing program on the photospheric data, which effectively removes the nonmagnetic forces, leads to significantly better results, but they are not as good as by using the chromospheric magnetic field vector as boundary condition. An area of current research is the possibility to use chromospheric images to improve the preprocessing of photospheric magnetic field measurements. Improvements in measuring the chromospheric magnetic field directly [e.g., *Lagg et al.*, 2004] might further improve to find suitable boundary conditions for NLFFF extrapolations. Force-free extrapolations are not suitable, however, to understand the details of physical processes on how the magnetic field evolves from the forced photosphere into the chromosphere, because nonmagnetic forces are important in the photosphere. For a better understanding of these phenomena more sophisticated models which take pressure gradients and gravity (and maybe also plasma flow) into account are required. Some first steps have been done with a generalization of the optimization method by *Wiegelmann and Neukirch* [2006], but such approaches are still in their infancy and have been tested so far only with smooth MHD equilibria. It is also not entirely clear how well necessary information regarding the plasma (density, pressure, temperature, flow) can be derived from measurements. Nonmagnetic forces become important also in quiet sun regions [*Schrijver and van Ballegooijen*, 2005] and in the higher layers of the corona, where the plasma β is of the order of unity. Coronagraph measurements, preferably from two viewpoints as provided by the STEREO mission, combined with a tomographic inversion might help here to get insights in the required 3D structure of the plasma density. One should also pay attention to the combination of extrapolation methods, as described here, with measurements of the Hanle and Zeeman effects in coronal lines which allows the reconstruction of the coronal magnetic field as proposed in feasibility studies of vector tomography by *Kramar et al.* [2006] and *Kramar and Inhester* [2006]. Other measurements of coronal features, for example, coronal plasma images from two STEREO viewpoints, can be used for observational tests of coronal magnetic field models. Using two viewpoints provide a much more restrictive test of models as images from only one view direction. While a nonlinear force-free coronal magnetic field model helps us to derive the topology, magnetic field and electric current strength in coronal loops, they do not provide plasma parameters. One way to get insights regarding the coronal plasma is the use of scaling laws to model the plasma along the reconstructed 3D field lines and compare correspondent artificial plasma images with real coronal images. *Schrijver et al.* [2004] applied such an approach to global potential coronal magnetic fields and compared simulated and real coronal images from one viewpoint. A generalization of such methods toward the use of more sophisticated magnetic field models and coronal images from two STEREO viewpoints will probably provide many insights regarding the structure and physics of the coronal plasma. An important challenge is for example the coronal heating problem. The dominating coronal magnetic field is assumed to play an important role here, because magnetic field configuration containing free energy can under certain circumstances reconnect *Priest* [1996, 1999] and supply energy for coronal heating. *Priest et al.* [2005] pointed out that magnetic reconnection at separators and separatrices plays an important role for coronal heating. Nonlinear force-free models can help here to identify the magnetic field topology, magnetic null points, separatrices and localized strong current concentration. While magnetic reconnection [see, e.g., *Priest and Schrijver*, 1999] is a dynamical phenomenon, the static magnetic field models discussed here can help to identify the locations favorable for reconnection. Time sequences of nonlinear force-free models computed from corresponding vector magnetograms will also tell wether the topology of the coronal magnetic field has changed due to reconnection, even if the physics of reconnection is not described by force-free models. Sophisticated 3D coronal magnetic field models and plasma images from two viewpoints might help to constrain the coronal heating function further, which has been done so far with plasma images from one viewpoint [*Aschwanden*, 2001a, 2001b] (by using data from Yokoh, Soho and Trace).

References

Alissandrakis, C. E. (1981), On the computation of constant alpha force-free magnetic field, *Astron. Astrophys.*, *100*, 197–200.

Aly, J. J. (1984), On some properties of force-free magnetic fields in infinite regions of space, *Astrophys. J.*, *283*, 349–362, doi:10.1086/162313.

Aly, J. J. (1989), On the reconstruction of the nonlinear force-free coronal magnetic field from boundary data, *Sol. Phys.*, *120*, 19–48.

Aly, J. J. (2005), A uniqueness result for a simple force-free magnetic field submitted to a topological constraint, *Astron. Astrophys.*, *429*, 15–18, doi:10.1051/0004-6361:20041547.

Amari, T., J. J. Aly, J. F. Luciani, T. Z. Boulmezaoud, and Z. Mikic (1997), Reconstructing the solar coronal magnetic field as a force-free magnetic field, *Sol. Phys.*, *174*, 129–149.

Amari, T., T. Z. Boulmezaoud, and Z. Mikic (1999), An iterative method for the reconstructionbreak of the solar coronal magnetic field. I. Method for regular solutions, *Astron. Astrophys.*, *350*, 1051–1059.

Amari, T., T. Z. Boulmezaoud, and J. J. Aly (2006), Well posed reconstruction of the solar coronal magnetic field, Astron, *Astrophys.*, *446*, 691–705, doi:10.1051/0004-6361:20054076.

Aschwanden, M. J. (2001a), Revisiting the determination of the coronal heating function from Yohkoh data, *Astrophys. J.*, *559*, L171–L174, doi:10.1086/323788.

Aschwanden, M. J. (2001b), An evaluation of coronal heating models for active regions based on Yohkoh, SOHO, and TRACE observations, *Astrophys. J.*, *560*, 1035–1044, doi10.1086/323064.

Aschwanden, M. J. (2005), *Physics of the Solar Corona: An Introduction with Problems and Solutions*, 2nd ed., Springer, New York.

Aschwanden, M. J., J. S. Newmark, J.-P. Delaboudinière, W. M. Neupert, J. A. Klimchuk, G. A. Gary, F. Portier-Fozzani, and A. Zucker (1999), Three-dimensional stereoscopic analysis of solar active region loops. I. SOHO/EIT observations at temperatures of $(1.0–1.5) \times 10^6$ K, *Astrophys. J.*, *515*, 842–867, doi:10.1086/307036.

Bineau, M. (1972), Existence of force-free magnetic-fields, *Commun. Pure Appl. Math.*, *25*, 77–84.

Bleybel, A., T. Amari, L. van Driel-Gesztelyi, and K. D. Leka (2002), Global budget for an eruptive active region. I. Equilibrium reconstruction approach, *Astron. Astrophys.*, *395*, 685–695, doi:10.1051/0004-6361:20021332.

Borrero, J. M., S. Tomczyk, A. Norton, T. Darnell, J. Schou, P. Scherrer, R. Bush, and Y. Lui (2006), Magnetic field vector retrieval with HMI, in *Solar Polarization 4, ASP Conf. Ser.*, vol. 358, edited by R. Casini and B. W. Lites, p. 144, Astron. Soc. of the Pac., San Francisco, Calif. (Available at http://arxiv.org/abs/astro-ph/0602130)

Bossavit, A. (1988), Mixed finite elements and the complex of Whitney forms, in *The Mathemetics of Finite Elements and Applications VI*, edited by J. Whiteman, pp. 137–144, Academic Press, London.

Boulmezaoud, T. Z., and T. Amari (2000), On the existence of non-linear force-free fields in three-dimensional domains, *ZAMP*, *51*, 942–967.

Carcedo, L., D. S. Brown, A. W. Hood, T. Neukirch, and T. Wiegelmann (2003), A quantitative method to optimise magnetic field line fitting of observed coronal loops, *Sol. Phys.*, *218*, 29–40.

Chiu, Y. T., and H. H. Hilton (1977), Exact Green's function method of solar force-free magnetic-field computations with constant alpha. I. Theory and basic test cases, *Astrophys. J.*, *212*, 873–885.

Chodura, R., and A. Schlueter (1981), A 3D code for MHD equilibrium and stability, *J. Comput. Phys.*, *41*, 68–88.

Cuperman, S., L. Ofman, and M. Semel (1990), Extrapolation of photospheric potential magnetic fields using oblique boundary values—A simplified approach, *Astron. Astrophys.*, *227*, 583–590.

Cuperman, S., P. Demoulin, and M. Semel (1991), Removal of singularities in the Cauchy problem for the extrapolation of solar force-free magnetic fields, *Astron. Astrophys.*, *245*, 285–288.

Dedner, A., F. Kemm, D. Kröner, C.-D. Munz, T. Schnitzer, and M. Wesenberg (2002), Hyperbolic divergence cleaning for the MHD equations, *J. Comput. Phys.*, *175*, 645–673.

Demoulin, P., and E. R. Priest (1992), The properties of sources and sinks of a linear force-free field, *Astron. Astrophys.*, *258*, 535–541.

Gary, G. A. (2001), Plasma Beta above a solar active region: Rethinking the paradigm, *Sol. Phys.*, *203*, 71–86.

Gary, G. A., and M. J. Hagyard (1990), Transformation of vector magnetograms and the problems associated with the effects of perspective and the azimuthal ambiguity, *Sol. Phys.*, *126*, 21–36.

Grad, H., and H. Rubin (1958), Hydromagnetic equilibria and force-free fields, paper presented at 2nd International Conference on Peaceful Uses of Atomic Energy, Int. At. Energy Agency, Geneva.

He, H., and H. Wang (2006), The validity of the boundary integral equation for magnetic field extrapolation in open space above a spherical surface, *Mon. Not. R. Astron. Soc.*, *369*, 207–215, doi:10.1111/j.1365-2966.2006.10288.x.

Henney, C. J., C. U. Keller, and J. W. Harvey (2006), SOLIS-VSM solar vector magnetograms, in *Solar Polarization 4, ASP Conf. Ser.*, vol. 358, edited by R. Casini and B. W. Lites, p. 92, Astron. Soc. of the Pac., San Francisco, Calif. (Available at http://arxiv.org/abs/astro-ph/0701122)

Inhester, B., and T. Wiegelmann (2006), Nonlinear force-free magnetic field extrapolations: Comparison of the Grad Rubin and Wheatland Sturrock Roumeliotis algorithm, *Sol. Phys.*, *235*, 201–221, doi:10.1007/s11207-006-0065-x.

Jiao, L., A. N. McClymont, and Z. Mikic (1997), Reconstruction of the three-dimensional coronal magnetic field, *Sol. Phys.*, *174*, 311–327.

Judge, P. G. (1998), Spectral lines for polarization measurements of the coronal magnetic field. I. Theoretical intensities, *Astrophys. J.*, *500*, 1009–1022.

Kageyama, A., and T. Sato (2004), "Yin-Yang grid": An overset grid in spherical geometry, *Geochem. Geophys. Geosyst.*, *5*, Q09005, doi:10.1029/2004GC000734.

Keppens, R., M. Nool, G. Tóth, and J. P. Goedbloed (2003), Adaptive mesh refinement for conservative systems: Multi-dimensional efficiency evaluation, *Comput. Phys. Commun.*, *153*, 317–339.

Klimchuk, J. A., and P. A. Sturrock (1992), Three-dimensional force-free magnetic fields and flare energy buildup, *Astrophys. J.*, *385*, 344–353, doi:10.1086/170943.

Kramar, M., and B. Inhester (2006), Inversion of coronal Zeeman and Hanle observations to reconstruct the coronal magnetic field, *Mem. Soc. Astron. Ital.*, *78*, 120. (Available at http://arxiv.org/abs/astro-ph/0612639)

Kramar, M., B. Inhester, and S. K. Solanki (2006), Vector tomography for the coronal magnetic field. I. Longitudinal Zeeman effect measurements, *Astron. Astrophys.*, *456*, 665–673, doi:10.1051/0004-6361:20064865.

LaBonte, B. J., D. L. Mickey, and K. D. Leka (1999), The imaging vector magnetograph at Haleakalā: II. Reconstruction of Stokes spectra, *Sol. Phys.*, *189*, 1–24.

Lagg, A., J. Woch, N. Krupp, and S. K. Solanki (2004), Retrieval of the full magnetic vector with the He I multiplet at 1083 nm: Maps of an emerging flux region, *Astron. Astrophys.*, *414*, 1109–1120.

Leka, K. D., and A. Skumanich (1999), On the value of 'αAR' from vector magnetograph data: I. Methods and caveats, *Sol. Phys.*, *188*, 3–19.

Li, Z., Y. Yan, and G. Song (2004), Properties of the boundary integral equation for solar non-constant-α force-free magnetic fields, *Mon. Not. R. Astron. Soc.*, *347*, 1255–1265, doi:10.1111/j.1365-2966.2004.07309.x.

Lin, H., J. R. Kuhn, and R. Coulter (2004), Coronal magnetic field measurements, *Astrophys. J. Lett.*, *613*, L177–L180, doi:10.1086/425217.

Liu, Y., X. P. Zhao, J. T. Hoeksema, P. H. Scherrer, J. Wang, and Y. Yan (2002), On formation of the sigmoidal structure in solar active region NOAA 8100, *Sol. Phys.*, *206*, 333–346.

Low, B. C., and Y. Q. Lou (1990), Modeling solar force-free magnetic fields, *Astrophys. J.*, *352*, 343–352.

McClymont, A. N., L. Jiao, and Z. Mikic (1997), Problems and progress in computing three-dimensional coronal active region magnetic fields from boundary data, *Sol. Phys.*, *174*, 191–218.

Metcalf, T. R. (1994), Resolving the 180-degree ambiguity in vector magnetic field measurements: The 'minimum' energy solution, *Sol. Phys.*, *155*, 235–242.

Metcalf, T. R., L. Jiao, A. N. McClymont, R. C. Canfield, and H. Uitenbroek (1995), Is the solar chromospheric magnetic field force-free?, *Astrophys. J.*, *439*, 474–481, doi:10.1086/175188.

Metcalf, T. R., et al. (2006), An overview of existing algorithms for resolving the 180-degree ambiguity in vector magnetic fields: Quantitative tests with synthetic data, *Sol. Phys.*, *237*, 267–296, doi:10.1007/s11207-006-0170-x.

Metcalf, T. R., M. L. DeRosa, C. J. Schrijver, G. Barnes, A. VanBallegooijen, T. Wiegelmann, M. S. Wheatland, G. Valori, and J. M. McTiernan (2007), Non-linear force-free modeling of coronal magnetic fields. II. Modeling a filament arcade from simulated chromospheric and photospheric vector fields, *Sol. Phys.*, in press.

Mickey, D. L., R. C. Canfield, B. J. Labonte, K. D. Leka, M. F. Waterson, and H. M. Weber (1996), The imaging vector magnetograph at Haleakala, *Sol. Phys.*, *168*, 229–250.

Mikic, Z., and A. N. McClymont (1994), Deducing coronal magnetic fields from vector magnetograms, in *Solar Active Region Evolution: Comparing Models with Observations, ASP Conf. Ser.*, vol. 68, edited by K. S. Balasubramaniam and G. W. Simon, pp. 225-232, Astron. Soc. of the Pac., San Francisco, Calif.

Mikic, Z., D. C. Barnes, and D. D. Schnack (1988), Dynamical evolution of a solar coronal magnetic field arcade, *Astrophys. J.*, *328*, 830–847, doi:10.1086/166341.

Molodensky, M. M. (1969), Integral properties of force-free fields, *Soviet Astron., Engl. Transl.*, *12*, 585–588.

Molodensky, M. M. (1974), Equilibrium and stability of force-free magnetic field, *Sol. Phys.*, *39*, 393–404.

Nakagawa, Y. (1974), Dynamics of the solar magnetic field. I. Method of examination of force-free magnetic fields, *Astrophys. J.*, *190*, 437–440.

Nakagawa, Y., and M. A. Raadu (1972), On practical representation of magnetic field, *Sol. Phys.*, *25*, 127–135.

Pevtsov, A. A., R. C. Canfield, and T. R. Metcalf (1994), Patterns of helicity in solar active regions, *Astrophys. J. Lett.*, *425*, L117–L119, doi:10.1086/187324.

Press, W. H. (2002), *Numerical Recipes in C++: The Art of Scientific Computing*, 1002 pp., Cambridge Univ. Press, New York.

Priest, E. (1996), Coronal heating by magnetic reconnection, *Astrophys. Space Sci.*, *237*, 49–73.

Priest, E. R. (1999), Heating the solar corona by magnetic reconnection, *Astrophys. Space Sci.*, *264*, 77–100.

Priest, E. R., and C. J. Schrijver (1999), Aspects of three-dimensional magnetic reconnection (invited review), *Sol. Phys.*, *190*, 1–24, doi:10.1023/A:1005248007615.

Priest, E. R., D. W. Longcope, and J. Heyvaerts (2005), Coronal heating at separators and separatrices, *Astrophys. J.*, *624*, 1057–1071, doi:10.1086/429312.

Régnier, S., and T. Amari (2004), 3D magnetic configuration of the Hα filament and X-ray sigmoid in NOAA AR 8151, *Astron. Astrophys.*, *425*, 345–352.

Régnier, S., and R. C. Canfield (2006), Evolution of magnetic fields and energetics of flares in active region 8210, *Astron. Astrophys.*, *451*, 319–330, doi:10.1051/0004-6361:20054171.

Régnier, S., T. Amari, and E. Kersalé (2002), 3D coronal magnetic field from vector magnetograms: Non-constant-alpha force-free configuration of the active region NOAA 8151, *Astron. Astrophys.*, *392*, 1119–1127.

Régnier, S., T. Amari, and R. C. Canfield (2005), Self and mutual magnetic helicities in coronal magnetic configurations, *Astron. Astrophys.*, *442*, 345–349, doi:10.1051/0004-6361:20053509.

Roumeliotis, G. (1996), The "stress-and-relax" method for reconstructing the coronal magnetic field from vector magnetograph data, *Astrophys. J.*, *473*, 1095–1103.

Sakurai, T. (1981), Calculation of force-free magnetic field with non constant alpha, *Sol. Phys.*, *69*, 343–359.

Sakurai, T. (1982), Green's function methods for potential magnetic fields, *Sol. Phys.*, *76*, 301–321.

Sakurai, T. (1989), Computational modeling of magnetic fields in solar active regions, *Space Sci. Rev.*, *51*, 11–48.

Sakurai, T., et al. (1995), Solar flare telescope at Mitaka, *Publ. Astron. Soc. Jpn.*, *47*, 81–92.

Schatten, K. H., J. M. Wilcox, and N. F. Ness (1969), A model of interplanetary and coronal magnetic fields, *Sol. Phys.*, *6*, 442–455.

Schmidt, H. U. (1964), On the observable effects of magnetic energy storage and release connected with solar flares, in *The Physics of Solar Flares*, edited by W. N. Hess, pp. 107–114, NASA, Washington, D. C.

Schrijver, C. J., and A. A. van Ballegooijen (2005), Is the quiet-sun corona a quasi-steady, force-free environment?, *Astrophys. J.*, *630*, 552–560, doi:10.1086/431754.

Schrijver, C. J., A. W. Sandman, M. J. Aschwanden, and M. L. DeRosa (2004), The coronal heating mechanism as identified by full-sun visualizations, *Astrophys. J.*, *615*, 512–525, doi:10.1086/424028.

Schrijver, C. J., M. L. DeRosa, A. M. Title, and T. R. Metcalf (2005), The nonpotentiality of active-region coronae and the dynamics of the photospheric magnetic field, *Astrophys. J.*, *628*, 501–513, doi:10.1086/430733.

Schrijver, C. J., M. L. Derosa, T. R. Metcalf, Y. Liu, J. McTiernan, S. Régnier, G. Valori, M. S. Wheatland, and T. Wiegelmann (2006), Nonlinear force-free modeling of coronal magnetic fields part i: A quantitative comparison of methods, *Sol. Phys.*, *235*, 161–190, doi:10.1007/s11207-006-0068-7.

Seehafer, N. (1978), Determination of constant alpha force-free solar magnetic fields from magnetograph data, *Sol. Phys.*, *58*, 215–223.

Seehafer, N. (1982), A comparison of different solar magnetic field extrapolation procedures, *Sol. Phys.*, *81*, 69–80.

Semel, M. (1967), Contribution à l'étude des champs magn\' etiques dans les régions actives solaires, *Ann. Astrophys.*, *30*, 513.

Semel, M. (1988), Extrapolation functions for constant-alpha force-free fields—Green's method for the oblique boundary value, *Astron. Astrophys.*, *198*, 293–299.

Shimizu, T. (2004), SolarB Solar Optical Telescope (SOT), in *The Solar-B Mission and the Forefront of Solar Physics, ASP Conf. Ser.*, vol. 325, edited by T. Sakurai and T. Sekii, p. 3, Astron. Soc. of the Pac., San Francisco, Calif.

Solanki, S. K., A. Lagg, J. Woch, N. Krupp, and M. Collados (2003), Three-dimensional magnetic field topology in a region of solar coronal heating, *Nature*, *425*, 692–695.

Song, M. T., C. Fang, Y. H. Tang, S. T. Wu, and Y. A. Zhang (2006), A new and fast way to reconstruct a nonlinear force-free field in the solar corona, *Astrophys. J.*, *649*, 1084–1092, doi:10.1086/506249.

Titov, V. S., and P. Démoulin (1999), Basic topology of twisted magnetic configurations in solar flares, *Astron. Astrophys.*, *351*, 707–720.

Török, T., and B. Kliem (2003), The evolution of twisting coronal magnetic flux tubes, *Astron. Astrophys.*, *406*, 1043–1059, doi:10.1051/0004-6361:20030692.

Valori, G., B. Kliem, and R. Keppens (2005), Extrapolation of a nonlinear force-free field containing a highly twisted magnetic loop, *Astron. Astrophys.*, *433*, 335–347.

van Ballegooijen, A. A. (2004), Observations and modeling of a filament on the Sun, *Astrophys. J.*, *612*, 519–529, doi:10.1086/422512.

van Ballegooijen, A. A., E. E. Deluca, K. Squires, and D. H. Mackay (2007), Modeling magnetic flux ropes in the solar atmosphere, *J. Atmos. Terr. Phys.*, *69*, 24–31, doi:10.1016/j.jastp.2006.06.007.

Wang, H., Y. Yan, T. Sakurai, and M. Zhang (2000), Topology of magnetic field and coronal heating in solar active regions. II. The role of quasi-separatrix layers, *Sol. Phys.*, *197*, 263–273.

Wheatland, M. S. (1999), A better linear force-free field, *Astrophys. J.*, *518*, 948–953, doi:10.1086/307301.

Wheatland, M. S. (2004), Parallel construction of nonlinear force-free fields, *Sol. Phys.*, *222*, 247–264.

Wheatland, M. S. (2006), A fast current-field iteration method for calculating nonlinear force-free fields, *Sol. Phys.*, *238*, 29–39, doi:10.1007/s11207-006-0232-0.

Wheatland, M. S., P. A. Sturrock, and G. Roumeliotis (2000), An optimization approach to reconstructing force-free fields, *Astrophys. J.*, *540*, 1150–1155.

Wiegelmann, T. (2004), Optimization code with weighting function for the reconstruction of coronal magnetic fields, *Sol. Phys.*, *219*, 87–108.

Wiegelmann, T. (2007), Computing nonlinear force-free coronal magnetic fields in spherical geometry, *Sol. Phys.*, *240*, 227–239, doi:10.1007/s11207-006-0266-3.

Wiegelmann, T., and B. Inhester (2003), Magnetic modeling and tomography: First steps towards a consistent reconstruction of the solar corona, *Sol. Phys.*, *214*, 287–312.

Wiegelmann, T., and T. Neukirch (2002), Including stereoscopic information in the reconstruction of coronal magnetic fields, *Sol. Phys.*, *208*, 233–251.

Wiegelmann, T., and T. Neukirch (2003), Computing nonlinear force free coronal magnetic fields, *Nonlinear Processes Geophys.*, *10*, 313–322.

Wiegelmann, T., and T. Neukirch (2006), An optimization principle for the computation of MHD equilibria in the solar corona, *Astron. Astrophys.*, *457*, 1053–1058, doi:10.1051/0004-6361:20065281.

Wiegelmann, T., B. Inhester, A. Lagg, and S. K. Solanki (2005a), How to use magnetic field information for coronal loop identification, *Sol. Phys.*, *228*, 67–78, doi:10.1007/s11207-005-2511-6.

Wiegelmann, T., A. Lagg, S. K. Solanki, B. Inhester, and J. Woch (2005b), Comparing magnetic field extrapolations with measurements of magnetic loops, *Astron. Astrophys.*, *433*, 701–705.

Wiegelmann, T., B. Inhester, B. Kliem, G. Valori, and T. Neukirch (2006a), Testing non-linear force-free coronal magnetic field extrapolations with the Titov-Démoulin equilibrium, *Astron. Astrophys.*, *453*, 737–741, doi:10.1051/0004-6361:20054751.

Wiegelmann, T., B. Inhester, and T. Sakurai (2006b), Preprocessing of vector magnetograph data for a nonlinear force-free magnetic field reconstruction, *Sol. Phys.*, *233*, 215–232.

Wu, S. T., H. M. Chang, and M. J. Hagyard (1985), On the numerical computation of nonlinear force-free magnetic fields, in *Measurements of Solar Vector Magnetic Fields*, NASA Conf. Publ., 2374, 17–40.

Wu, S. T., M. T. Sun, H. M. Chang, M. J. Hagyard, and G. A. Gary (1990a), On the numerical computation of nonlinear force-free magnetic fields, *Astrophys. J.*, *362*, 698–708.

Wu, S. T., M. T. Sun, and T. Sakurai (1990b), A comparison between progressive extension method (PEM) and iterative method (IM) for magnetic field extrapolations in the solar atmosphere, *Mem. Soc. Astron. Ital.*, *61*, 477–484.

Yan, Y., and Z. Li (2006), Direct boundary integral formulation for solar non-constant-α force-free magnetic fields, *Astrophys. J.*, *638*, 1162–1168, doi:10.1086/499064.

Yan, Y., and T. Sakurai (2000), New boundary integral equation representation for finite energy force-free magnetic fields in open space above the Sun, *Sol. Phys.*, *195*, 89–109.

Yan, Y., M. J. Aschwanden, S. Wang, and Y. Deng (2001a), Evolution of magnetic flux rope in the active region NOAA 9077 on 14 July 2000, *Sol. Phys.*, *204*, 27–40, doi:10.1023/A:1014265123601.

Yan, Y., Y. Deng, M. Karlický, Q. Fu, S. Wang, and Y. Liu (2001b), The magnetic rope structure and associated energetic processes in the 2000 July 14 solar flare, *Astrophys. J. Lett.*, *551*, L115–L119, doi:10.1086/319829.

Yang, W. H., P. A. Sturrock, and S. K. Antiochos (1986), Force-free magnetic fields: The magneto-frictional method, *Astrophys. J.*, *309*, 383–391, doi:10.1086/164610.

Yee, K. S. (1966), Numerical solition of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas Propag.*, *AP-14*(4), 302–307.

————————————

T. Wiegelmann, Max-Planck-Institut für Sonnensystemforschung, Max-Planck-Strasse 2, D-37191 Katlenburg-Lindau, Germany. (wiegelmann@mps.mpg.de)

Key points model for polar region currents

Wen-Yao Xu,[1] Geng-Xiong Chen,[1] Ai-Min Du,[1] Ying-Yan Wu,[1] Bo Chen,[1]
and Xiao-Can Liu[1]

[1] The equivalent ionospheric electric currents in the polar region mainly consist of the DP1 and DP2 systems. The former involves a westward electrojet around midnight, while the later involves a two-cell system with foci on the morningside and eveningside, respectively. In space weather prediction and nowcasting, sophisticated models of the polar currents are needed, but concise and convenient models are also useful to predict or nowcast the principal characteristics of the current systems, such as intensity and position. In this paper, we propose a "key points model" for outlining the basic features of the polar region current system for different disturbance levels. The "key points model" (or KP model) includes six key points of the current system: the centers of two DP2 cells, the maximum densities of the eastward and westward electrojets, and the maximum densities of the northward and southward currents. Each of six key points is described by three parameters: intensity, local time, and latitude. The AE-dependences of the 18 parameters are deduced from the equivalent current systems for every 5 min during a 2-d period (18–19 March 1978). The KP model reveals systematic variations of the current systems. When AE increases, the currents and the current densities are simultaneously enhanced linearly, and most of the key points concentrate towards midnight. In addition, when AE increases, the key points K2 and K4 for the evening current cell move equatorward, while the key points K1 and K3 for morning cell move poleward.

1. Introduction

[2] The ionospheric electric currents in the polar region and associated magnetic disturbances are essential manifestation of solar wind-magnetosphere-ionosphere coupling and are what space weather prediction concerns with [*Akasofu*, 1979; *Chen et al.*, 2000]. The equivalent currents in the polar region mainly consist of the DP1 and DP2 systems [*Clauer and Kamide*, 1985; *Kamide and Kokubun*, 1996]. The former involves a westward electrojet around midnight, associated with the substorm current wedge that reflects the unloading process in the magnetic tail. The latter is a two-cell system with the foci of the cells on the morningside and eveningside and which is driven by large-scale magnetospheric convection [*Akasofu*, 1964; *Rostoker*, 1969, 1996; *Rostoker et al.*, 1987]. During quiet periods, the DP2 currents dominate the whole polar current pattern. During disturbed conditions, especially in the expansion phase of substorms, the DP1 currents are drastically intensified, while at the same time, a steady enhancement of the DP2 currents is observed [*Sun et al.*, 1998, 2000].

[3] Magnetic data from observatories distributed in the polar region have been widely used to deduce the ionospheric equivalent current system, from which the three dimensional current system with field-aligned currents can be inferred [*Kamide et al.*, 1981]. Using the ground-based magnetometer data from 71 stations distributed along six meridian chains at high latitudes during the period of the International Magnetosphere Study (IMS), *Kamide et al.* [1982] calculated equivalent ionospheric current systems in the polar region for every 5 min of 17–19 March 1978. The current systems demonstrate the dominant DP1 and DP2 components [*Kamide and Kokubun*, 1996], although they undergo very complicated variations in both pattern and intensity (Figure 1).

[4] The concept of two-component process of magnetospheric substorm was proposed by *Akasofu* [1979], and has been extensively studied over the past few decades [*Rostoker et al.*, 1987; *Kamide and Kokubun*, 1996; *Sun et al.*, 1998, 2000]. The first component involves the directly driven process, resulting in a two-cell convection and current system in the polar regions. The second component is unloading process, producing a current wedge and an ionospheric westward electrojet around midnight. *Kamide and Kokubun* [1996] illustrated two pairs of the electric potential cells in a schematic diagram, representing the effects of enhanced plasma convection and substorm expansion.

[1]Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China.
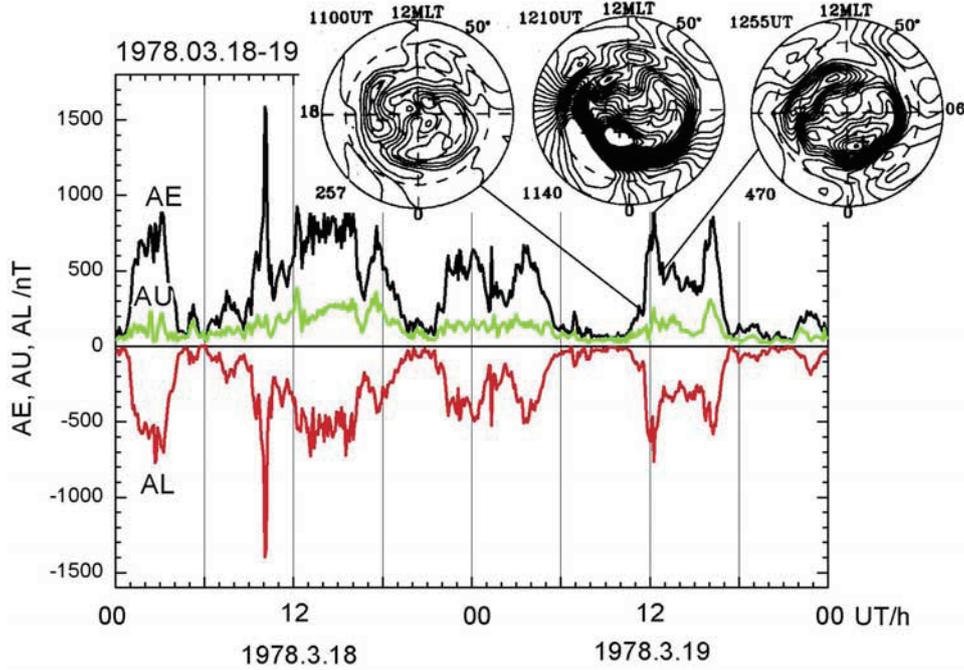
Figure 1. Magnetic indices *AE*, *AL*, and *AU* during 18–19 March 1978. The insets illustrate typical polar ionospheric equivalent current systems for three phases of a substorm, deduced from ground-based magnetic records by *Kamide et al.* [1982].

[5] A quantitative separation of the components DP1 and DP2 in the current systems have been attempted by *Sun et al.* [1998, 2000]. They applied the NOC method (method of Natural Orthogonal Components) to the data supplied by *Kamide et al.* [1982] and obtained a series of polar current eigenmodes. Comparison of the eigenmode patterns with the current systems of the driven and unloading processes suggested that the first and second modes correspond to DP2 and DP1, respectively.

[6] During the past couple of decades, several sophisticated models of the polar currents have been established, which are certainly needed for understanding magnetospheric-ionospheric coupling and space weather predictions. In addition, concise and convenient models are also useful for predicting or nowcasting most essential characteristics of the current system, such as current intensity, pattern, and position of the current center. These models require relatively few inputs and much easier available data and are able to promptly give the principal features of the current system with detailed structures omitted. The purpose of this paper is to propose a convenient and simplified model for nowcasting basic characteristics of the polar currents.

2. Data and Analysis

[7] The data used in this paper are 576 ionospheric equivalent current systems (IECS) computed every 5 min over the interval 18–19 March 1978 [*Kamide et al.*, 1981]. The IECS cover a latitudinal range of 50°–90° at 1° intervals and a longitudinal range of 0°–360° (or 0000–2400 LT) at 15° intervals (or 1 h). The IECS is assumed to flow in the ionosphere at a distance of $a + h$ from the center of the Earth, where a is the Earth's radius and h is the height

of the ionosphere above the Earth's surface. It is represented by current function (in amperes) as follows:

$$J(\theta, \lambda) = \sum_{n=0}^{\infty} J_n(\theta, \lambda) \tag{1}$$

$$J_n(\theta, \lambda) = -\frac{1}{\mu_0} \frac{2n+1}{n+1} \left(\frac{a+h}{r} \right)^n U_n^e(\theta, \lambda) \tag{2}$$

where $\sum U_n^e(\theta, \lambda)$ is external magnetic potential deduced from 71 stations along six meridian chains in the Northern Hemisphere. The southward and eastward components of the current density (in amperes/m) can be calculated from J by

$$I_\theta(\theta, \lambda) = \frac{1}{r \sin \theta} \frac{\partial J(\theta, \lambda)}{\partial \lambda} \tag{3}$$

$$I_\lambda(\theta, \lambda) = -\frac{1}{r} \frac{\partial J(\theta, \lambda)}{\partial \theta} \tag{4}$$

[8] The IECS pattern is represented by contours of the current function, as shown in Figure 1. In general, the IECS structure is rather complicated. Omitting small-scale details and random distortions, one can recognize the two-cell currents of DP2, which converge at the Harang discontinuity just before midnight, then flowing sunward across the polar cap, eventually diverging near the forenoon cusp region, returning to their respective electrojets. At the convergence (or divergence) point there should be major
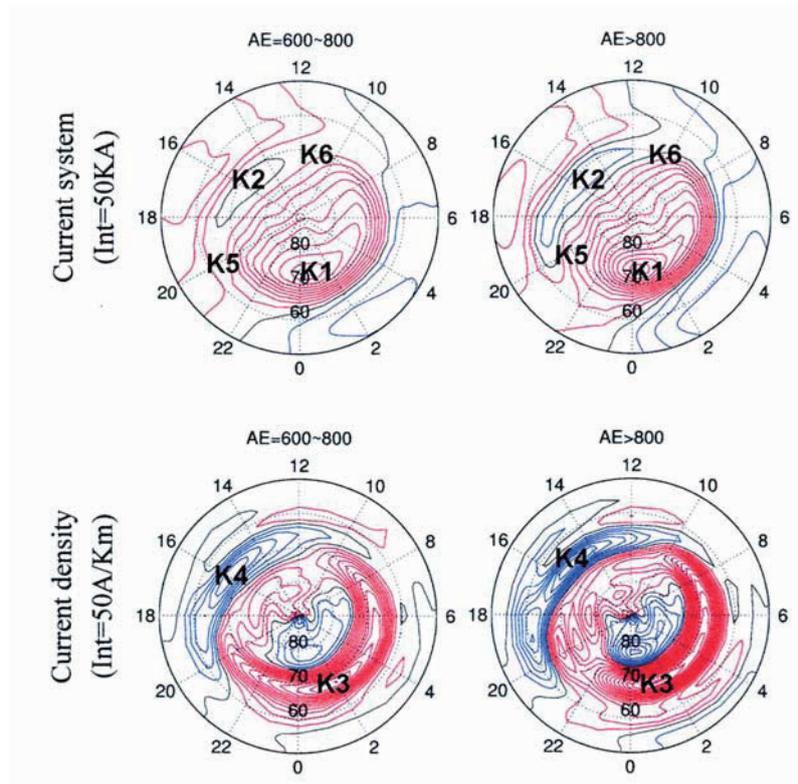
Figure 2. Average ionospheric equivalent (top) current systems and (bottom) current densities for $AE =$ 600–800 nT (left) and $AE > 800$ nT (right). The "key points" in the model are indicated by K1, K2, K3, K4, K5, and K6 (see text).

northward (or southward) currents. It is noted that the real current system in the polar region is three dimensional, including field-aligned currents [*Baumjohann*, 1983].

[9] Figure 2 shows two average equivalent current patterns for $AE = 600$–800 nT and $AE > 800$ nT, which are obtained by superposing 88 and 27 individual patterns, respectively, in order to eliminate complicated irregular fluctuations and small-scale structures. It is noted in Figure 2 that the principal features of the polar current system are described by six "key points": (1) K1, the center of the clockwise cell of the DP2 currents in morning sector, (2) K2, the center of the counterclockwise cell of the DP2 currents in evening sector, (3) K3, the point of the maximum westward current density in morning sector, (4) K4, the point of the maximum eastward current density in evening sector, (5) K5, the point of the maximum northward current density near Harang discontinuity, (6) K6, the point of the maximum southward current density near the cusp region.

[10] Each key point is described by three parameters: current intensity (total current J for K1 and K2, and current density I for K3, K4, K5, and K6), magnetic local time T, and geomagnetic latitude Φ. All of the 18 parameters determine the geometry and strength of the polar region currents.

[11] In order to eliminate irregular fluctuations and random distortions in the current patterns, the 576 samples are divided into 10 groups according to AE ranges: the first group for $AE = 0 \sim 100$ nT, the second for $AE = 100 \sim$ 200 nT.......the 10th for $AE = 900 \sim 1000$ nT. Then the average current pattern for each AE group is calculated by

superposing all current functions included in this group. We confine our analysis to the range $AE = 0 \sim 1000$ nT because only three samples are beyond this range. The average current systems so obtained are illustrated in Figure 3, where the AE range, sample number, and contour interval for each AE group are indicated.

[12] It is noted in Figure 3 that the principal feature of the current system is a two-cell structure, including a major clockwise cell (assigned as positive) in the early morning and a minor counterclockwise cell around evening (assigned as negative). The total currents of the two cells increase with increasing AE.

[13] Figure 4 illustrates the distributions of E-W current density for the 10 AE groups, where westward current is assigned as positive (red color). The most outstanding features are westward electrojet in midnight-morning sector and eastward electrojet (blue color) in afternoon-evening sector.

[14] It is interesting to note a "double peak" structure of the westward electrojet in Figure 4 for $AE \sim 200$–300 and 900–1000, one near midnight and another near dawn. This conforms to the view that westward electrojet has two contributors: the directly driven westward electrojet in DP2 with its peak near dawn and the substorm expansion phase westward electrojet in DP1 with its peak near midnight. The other ranges show that when AE increases, the midnight peak increases much faster than the dawn peak and two peaks combine together, forming an enhanced electrojet belt. In other words, during disturbance periods (typically, substorms), enhancement of DP1 due to the
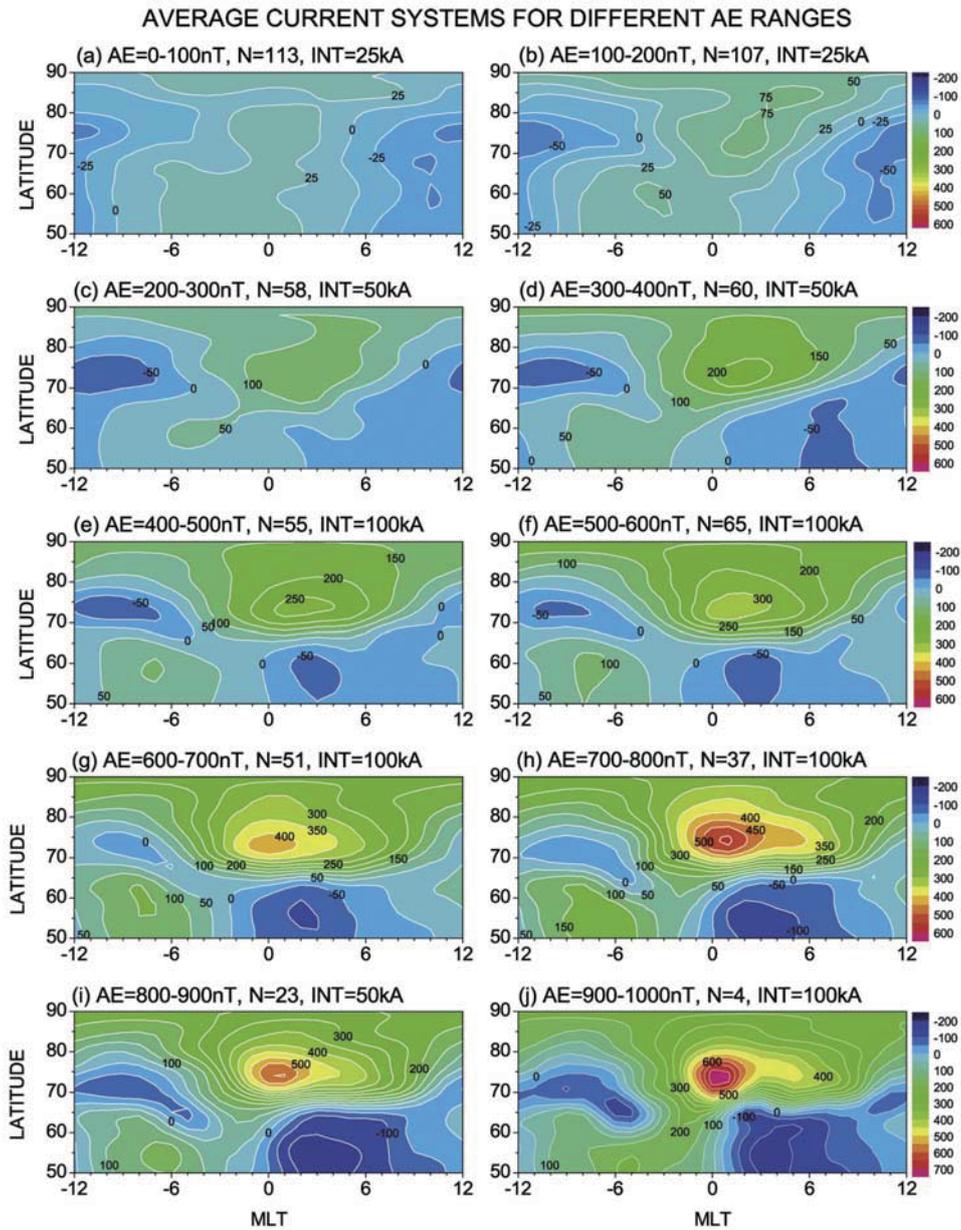
Figure 3. Average equivalent current systems for 10 *AE* groups. The numbers at the top of each figure are *AE* range, sample number, and contour interval for the group.

unloading process is much more significant than the increase of DP2. Similar double peak structure can be seen in eastward electrojet, especially for intensely disturbed periods.

[15] As a first step in the analysis, the 18 parameters of the six key points for each of the 576 samples are found from its current pattern. Then the "average parameters" are calculated for each of the 10 *AE* groups. Finally, the *AE* dependence of each parameter is deduced by linear fitting.

[16] Figure 5 shows the obtained *AE* dependences of the 18 parameters for six key points. The regression coefficients are listed in Table 1.

[17] It is noted in Figure 5 and Table 1 that when AE increases, the total current J_{CCW} of the evening cell K2 is enhanced at a rate of 0.100 kA/nT, while a much greater

enhancement rate of 0.806 kA/nT is seen in the current J_{CW} of the morning cell K1, implying the important role of the unloading process in the formation of the equivalent polar current system during disturbances. A similar contrast is noted in the current densities I_E (K4, enhancement rate 0.431 A/km/nT) and I_W (K3, enhancement rate 1.382 A/km/nT). In addition, the N-S current densities I_N and I_S for K5 and K6 are simultaneously enhanced linearly.

[18] When *AE* increases, all key points systematically move in both local time (T), and geomagnetic latitude (Φ), as shown by arrows in Figure 6. It is noted in Figure 6 that all key points, except for K3 (the maximum westward current density I_W in morning sector), concentrate towards midnight with increasing *AE*. At the same time, the key
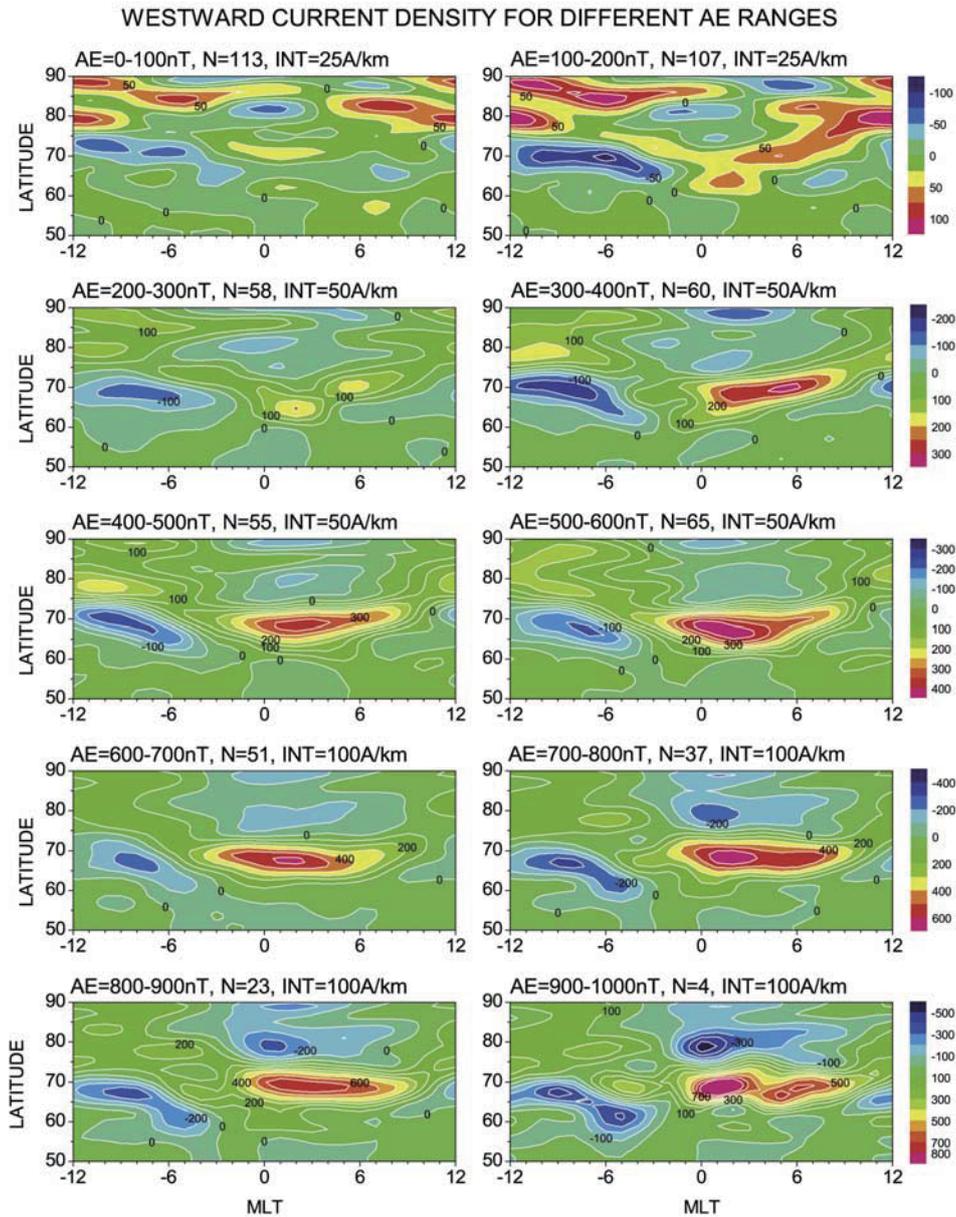
Figure 4. Same as Figure 3, but for the average E-W current density. Different color scales are for different figures.

points K2, K4, and K6 shift equatorward, while the key points K1, K3, and K5 shift poleward.

[19] The most remarkable movements occur for K2 (the evening current cell) and K4 (the eastward electrojet), which are associated with equatorward expansion of the auroral oval during substorms [*Akasofu*, 1977]. When *AE* increases from 0 nT to 1000 nT, K2 moves toward midnight by 2.9 h and equatorward by 10°, while K4 moves toward midnight by about 1 h and equatorward by 7°.

[20] The poleward shift of the key points K1 and K3 is a peculiar feature in present model. A possible explanation is as follows. As mentioned above, the westward electrojet comes from two contributors: the directly driven westward electrojet and the substorm expansion phase westward electrojet. They have different latitudinal motions when *AE* increases. Since substorm expansion phase electrojet

contributes significantly to *AE* during strong activity, the high-latitude edge of the expanded oval in midnight sector shifts poleward, although directly driven electrojet in the morning sector shifts equatorward. New research by *Ahn et al.* [2005] has found that the region of brightest auroras does not coincide with the area of strongest magnetic perturbations. The auroral electrojet is found to remain located adjacent to and poleward of the region of bright auroral luminosity.

3. Discussion
3.1. Advantage of the KP Model

[21] Compared with some sophisticated models of magnetospheric activity with many inputs and outputs, the KP model developed in this paper is a concise, convenient, and
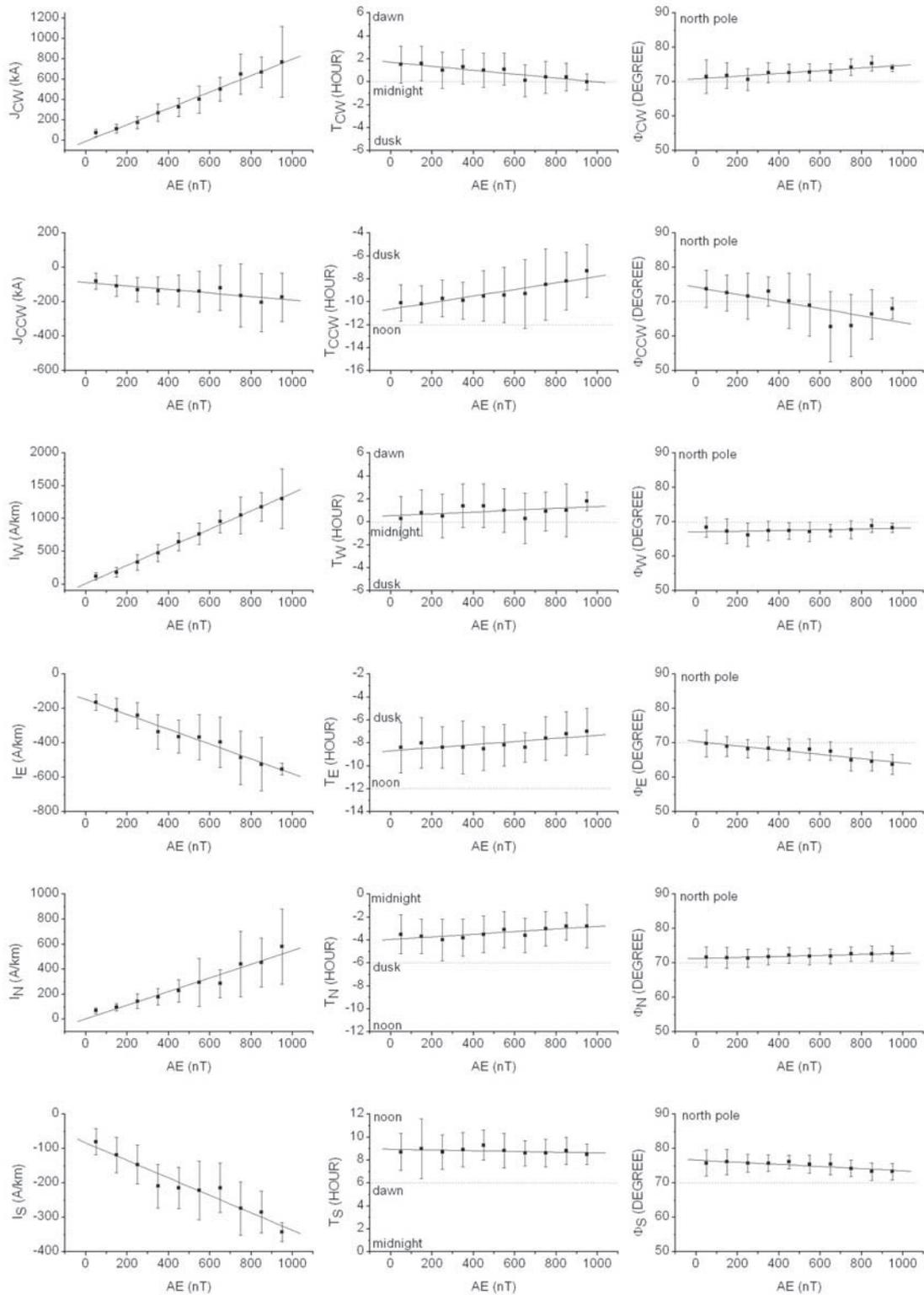
Figure 5. Linear fitting of *AE* dependences of 18 parameters for six key points. The left column shows current intensities, the middle column shows local times of the key points, and the right column shows their geomagnetic latitudes. For clarity, the horizontal dotted lines indicate the times of noon, midnight, dawn, and dusk in the middle and the latitude 700° on the right.

Table 1. Regression Coefficients of the Key Point (Ki, i = 1, 2 . . . 6) Parameters

Parameters	Fitting $y = a_0 + a_1 x$ $x = AE$ index, nT	
	a_0	a_1
K1 (CW) Clockwise Cell		
J_{CW}, kA	−7.9	0.806
T_{CW}, h	1.7	−0.00172
Φ_{CW}, deg	70.9	0.00396
K2 (CCW) Counterclockwise Cell		
J_{CCW}, kA	−88.6	−0.100
T_{CCW}, h	−10.6	0.00285
Φ_{CCW}, deg	74.2	0.0100
K3 (W) Maximum Westward Density		
I_W, A/km	10.0	1.382
T_W, h	0.54	0.0008
Φ_W, deg	67.1	0.0011
K4 (E) Maximum Eastward Density		
I_E, A/km	−148.1	−0.431
T_E, h	−8.7	0.00136
Φ_E, deg	70.4	−0.00626
K5 (N) Maximum Northward Density		
I_N, A/km	4.7	0.5439
T_N, h	−3.9	0.00113
Φ_N, deg	71.3	0.00147
K6 (S) Maximum Southward Density		
I_S, A/km	−83.7	−0.2539
T_S, h	8.9	−0.000309
Φ_S, deg	76.7	−0.00312

simplified model. After the equivalent current systems are obtained from the magnetic data at 71 stations, the input to the model is AE index, which is easily and timely available for users. The output of the model is an equivalent current

system in the polar region, which is essential to description of the space environment. This model can be used to nowcast the most important characteristics of the polar current system, such as the overall current pattern, current intensity, and the key point positions, although the detailed structure, irregular fluctuations, and random distortions in the current system are omitted.

3.2. Limitation of the KP Model

[22] The KP model is a concise, convenient, and simplified model, involving computation of the equivalent ionospheric current, instead of the real three-dimensional current system. In fact, field-aligned currents play a key role in polar region currents and associated magnetic disturbances. Using the KRM method or MIT technique, one can deduce the three dimensional current system, which requires an assumed or experimentally determined ionospheric conductivity model and involves complicated calculations.

[23] It should be pointed out that the magnetic measurements from which the current systems are inferred come from a widely scattered distribution of magnetometer stations that are irregularly distributed globally, and in some regions there is very poor coverage. Therefore one should not expect to believe in small-scale details of the equivalent current structure where there are no data to constrain the results. Fortunately, the KP model only concerns with large-scale structures, averaging out the small-scale details when calculating average current patterns for different AE ranges.

3.3. Input and Output

[24] In this paper, the AE index is chosen as the input to the KP model after comparison with AL and AU. By the definition of the auroral electrojet indices, AU is dependent on the maximum eastward electrojet, represented by I_E of key point K4. However, the present analysis shows that the
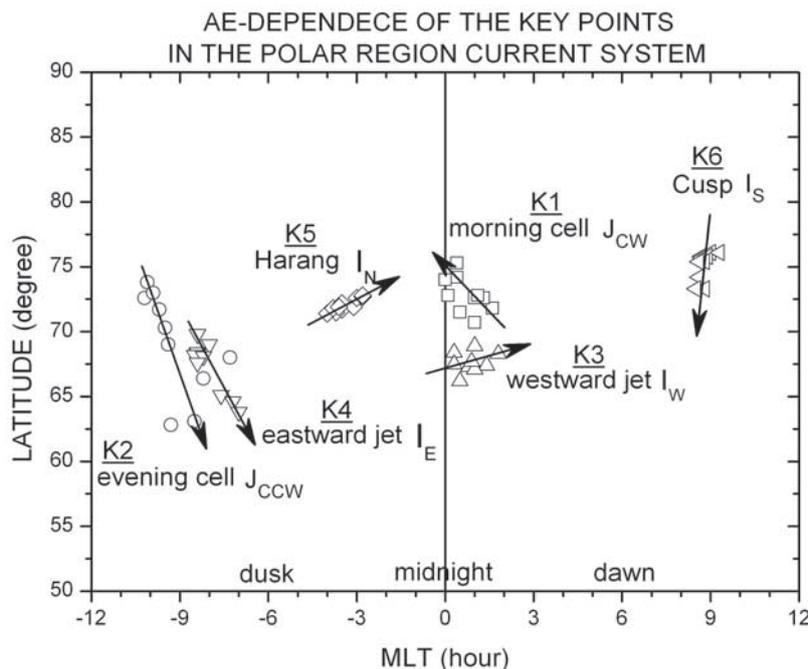
Figure 6. AE dependence of the key point positions. The arrows indicate movement tendencies of the key points with increasing AE.

correlation coefficient of I_E with AU is less than those with AL and AE. One of possible reasons is the large movement of the key point K4 (see Figure 6). During strong substorms, K4 moves equatorward beyond the observation of AE stations, which are confined in a relatively narrow geomagnetic latitudinal belt from $60.7°$ to $70.8°$. In such cases, the magnitude of AU reaches "saturation," although auroral currents may continue increasing [*Akasofu*, 1977; *Kamide and Akasofu*, 1983]. The saturation phenomenon in AE, AU, and AL indices lowers reliability of large values of AE, and also limits extension of the KP model to very intense substorms. Since the data used in this paper are limited, the present KP model is valid for $AE < 1000$ nT. A further extension of the KP model to much more intense substorms requires more data covering $AE > 1000$ nT. However, as indicated above, this extension would be limited by low reliability of large AE and its saturation during strong substorms.

[25] Another possible candidate input parameter is the PC index, that represents the equivalent current J_{PC} flowing across the polar cap and which can be estimated with data from only one station [*Vennerstrom and Friis-Christensen*, 1987].

4. Summary

[26] 1. The equivalent ionospheric current system in the polar region always shows dominant DP1 and DP2 components. This is the basis of the key points model for the polar region current system.

[27] 2. Six key points in the KP model are the centers of the two current cells in DP2, the maximum eastward and westward electrojets, and the maximum northward and southward current densities. These points determine the basic features of polar region current system.

[28] 3. As AE increases, the total currents for K1 and K2 and the current intensities for K3, K4, K5, and K6 are simultaneously enhanced linearly.

[29] 4. As AE increases, most of the key points concentrate toward midnight and, at the same time, shift equatorward. The most remarkable movement occurs at the evening current cell (K2) and the eastward electrojet (K4). When AE increases from 0 nT to 1000 nT, the key point K2 moves toward midnight by 2.9 h and equatorward by $10°$, while K4 moves toward midnight by about 1 h and equatorward by $7°$.

[30] 5. A peculiar feature in present model is poleward shift of the key points K1 and K3. A possible explanation is the poleward expansion of substorm expansion phase electrojet during strong activity.

References

Ahn, B.-H., G. X. Chen, W. Sun, J. W. Gjerloev, Y. Kamide, J. B. Sigwarth, and L. A. Frank (2005), Equatorward expansion of the westward electrojet during magnetically disturbed periods, *J. Geophys. Res.*, *110*, A01305, doi:10.1029/2004JA010553.

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

Akasofu, S.-I. (1977), *Physics of Magnetospheric Substorms*, pp. 381–472, D. Reidel, Boston.

Akasofu, S.-I. (1979), *Dynamics of the Magnetosphere*, 447 pp., D. Reidel, Boston.

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

Chen, H.-F., G.-X. Chen, F.-L. Peng, and W.-Y. Xu (2000), Analysis of S_q^p current system by using corrected geomagnetic coordinates, *Chin. J. Polar Sci.*, *11*, 59.

Clauer, C. R., and Y. Kamide (1985), DP1 and DP2 current systems for the March 22, 1979 substorm, *J. Geophys. Res.*, *90*, 1343.

Kamide, Y., and S.-I. Akasofu (1983), Notes on the auroral electrojet indices, *Rev. Geophys.*, *21*, 1647.

Kamide, Y., and S. Kokubun (1996), Two-component auroral electrojet: Importance for substorm studies, *J. Geophys. Res.*, *101*, 13,027.

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

Kamide, Y., et al. (1982), Global distribution of ionospheric and field-aligned currents during substorms as determined from six IMS meridian chains of magnetometers: Initial results, *J. Geophys.Res.*, *87*, 8228.

Rostoker, G. (1969), Classification of polar magnetic disturbances, *J. Geophys. Res.*, *74*, 5161.

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

Rostoker, G., S.-I. Akasofu, W. Baumjohann, Y. Kamide, and R. L. McPherron (1987), The roles of directly input of energy from the solar wind and unloading of stored magnetotail energy in driving magnetospheric substorms, *Space Sci. Rev.*, *46*, 93.

Sun, W., W.-Y. Xu, and S.-I. Akasofu (1998), Mathematical separation of directly driven and unloading components in the ionospheric equivalent currents during substorms, *J. Geophys. Res.*, *103*, 11,695.

Sun, W., W.-Y. Xu, and S.-I. Akasofu (2000), An improved method to deduce the unloading component for magnetospheric substorms, *J. Geophys. Res.*, *105*, 13,131.

Vennerstrom, S., and E. Friis-Christensen (1987), On the role of IMF B-y in generating the electric field responsible for the flow across the polar cap, *J. Geophys. Res.*, *92*, 195.

―――――――――――――
B. Chen, G.-X. Chen, A.-M. Du, X.-C. Liu, Y.-Y. Wu, and W.-Y. Xu, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China. (wyxu@mail.igcas.ac.cn)

Some ubiquitous features of the mesospheric Fe and Na layer borders from simultaneous and common-volume Fe and Na lidar observations

Fan Yi,[1,2,3] Shaodong Zhang,[1,2,3] Xianchang Yue,[1,2,3] Yujin He,[1,2,3] Changming Yu,[1,2,3] Chunming Huang,[1,2,3] and Wei Li[1,2,3]

[1] High-accuracy atom density profiles, obtained by the simultaneous and common-volume Fe and Na lidar measurements at Wuhan, China (30.5°N, 114.4°E), reveal some ubiquitous features of the Fe and Na layers on their borders. The Fe and Na lower boundaries show consistently a delicate stratification in which the lower boundary of the Fe layer is in general slightly higher than or coincident with that of the Na layer, with an overall mean altitude difference being about 0.2 km. Despite the existence of considerable vertical displacements, the two lower boundaries vary always following almost the same track. The overall correlation coefficient between them is as high as 0.96. This ubiquitous delicate stratification of the measured lower boundaries (nearly coincident density cutoff) suggests strongly that the undersides of Fe and Na layers are controlled by the same or very similar processes. The upper boundary of the Na layer is always several kilometers higher than that of the Fe layer. A relatively weak positive correlation is also persistently observed between the two upper boundaries. Weak sporadic layering events frequently occur on the upper extent of the metal layers. They may impair the correct determination of the upper boundaries of the normal metal layers and consequently weaken the correlation. Both the Fe and Na layers often show an evidently steeper density gradient on the underside than on the upper extent, and the borders of the Fe layer are clearly steeper than those of the Na layer. The explanation to these ubiquitous features needs further experimental and modeling efforts.

1. Introduction

[2] Many species of free metal atoms gather in the mesopause region from 80 to 110 km to form layers (Fe, Na, K, Ca, and Li) detectable by ground-based lidars. These metal atoms are believed to result from meteoric ablation. On the basis of independent lidar measurements from many locations, the average properties and seasonal variations of these layers have been respectively investigated. Fe is the most abundant of the mesospheric metallic atoms, with an annual mean column abundance roughly twice that of Na [*Kane and Gardner*, 1993]. The mean column abundances of K and Ca are about two orders smaller [*Eska et al.*, 1999; *Gerding et al.*, 2000]. The mean peak heights (or mean centroid heights) of these metal atom layers exhibit some difference. The Fe layer has the lowest peak height (~88 km) among these metal atom layers. In comparison with the peak of the Fe layer, the Na peak is about 3 km

higher, while both the Ca and K peaks are roughly 2–3 km higher [*Kane and Gardner*, 1993; *Eska et al.*, 1998; *Plane et al.*, 1999; *Gerding et al.*, 2000; *Friedman et al.*, 2002]. At midlatitudes the Na layer is the widest of these layers, with a full width at half-maximum (FWHM) over 10 km. The Fe and K layer FWHMs are 1–2 km narrower than that of the Na layer. The Ca layer is the narrowest among these layers. The seasonal variations of these layers show some similarities as well as differences. At midlatitudes, the Fe and Na layers have early winter maxima and midsummer minima in column abundance, while the K and Ca layers display a semiannual variation with a prominent second maximum in midsummer [*Eska et al.*, 1998; *States and Gardner*, 1999; *Gerding et al.*, 2000; *Plane*, 2003]. The centroid heights of these layers are characterized by a strong semiannual variation. The K layer has the lowest altitude in winter [*Kane and Gardner*, 1993; *Eska et al.*, 1998], whereas the Na, Fe, and Ca layers have the lowest altitude in summer [*Kane and Gardner*, 1993; *Gerding et al.*, 2000]. The layer widths of all layers show maxima in winter. At low latitudes, the seasonal behaviors of the K and Fe layers are clearly different from the corresponding results observed at midlatitude [*Friedman et al.*, 2002; *Raizada and Tepley*, 2003]. In particular, the centroid height and RMS width of

[1]School of Electronic Information, Wuhan University, Wuhan, China.
[2]Key Laboratory of Geospace Environment and Geodesy, Ministry of Education, Wuhan, China.
[3]State Observatory for Atmospheric Remote Sensing, Wuhan, China.

the Fe layer are nearly out of phase with respect to its midlatitude counterparts [*Raizada and Tepley*, 2003], while the Fe column abundance lags that at midlatitude by about 3 months. The latitudinal differences in Fe layer properties have been explained by the latitudinal differences in the annual variations of [O], [O_3], and [O_2], which result in the different season dependence of the Fe chemistry at two latitudes.

[3] Several one-dimensional gas-phase chemical models corresponding to different metal species such as Na, Fe, K, and Ca have been respectively developed to explain these observational results [*Helmer et al.*, 1998; *Plane et al.*, 1999; *Eska et al.*, 1999; *Gerding et al.*, 2000; *Delgado et al.*, 2006]. The Fe and Na models among others successfully reproduced some primary features of the normal Fe and Na layers observed at midlatitudes [*Helmer et al.*, 1998; *Plane et al.*, 1999], for instance, the layer shapes and the seasonal variations in the layer parameters. However, the K model could satisfactorily reproduce the seasonal behavior of the K layer observed at a midlatitude site only after the wintertime K influx from meteoric ablation had been reduced by 30% compared to the summer [*Eska et al.*, 1999]. Furthermore, the Fe and Na layers observed during the austral winter at the South Pole could be well modeled if the Fe and Na influxes from meteoric ablation were increased by a factor of 1.7 for these months [*Gardner et al.*, 2005]. These results appear to reflect two facts: the first is that chemistry alone cannot explain all the seasonal differences [*Raizada et al.*, 2004], and second, we currently lack a quantitative description on the deposition flux of the metals (the influx of extraterrestrial material and ablation efficiency of various metal species).

[4] In fact, the idea that meteoric deposition is the major source of mesospheric metal atoms is currently appealing in qualitative way because hitherto there is no direct and indirect measurement for the total vertical flux of any mesospheric metal species produced by meteoric ablation. In the existing chemical models, the total vertical flux for each metal species was specified to a value which resulted in good agreement between the observed and simulated metal atom profiles [*Helmer et al.*, 1998; *Gardner et al.*, 2005]. In order to understand the formation, structure, and variability of the metal atom layers, one needs to have quantitative knowledge about the source, such as the micrometeor flux as a function of time and geographic grid, the distributions of micrometeor particle masses, mass densities and velocities, the chemical composition of micrometeor particles, the evaporation processes including their dependences on particle mass, density, velocity, direction, and altitude [*Gerding et al.*, 1999]. *Janches et al.* [2006] have made a modeling effort in constructing a global micrometeor input function based on the radar observations of ionized meteor input rate. The model assumes that the detected particles come from three radiant distributions. The apex source flux makes up 70% of the total number of micrometeoroid particles, while the other 30% is provided by the combined distribution of two other sources which are concentrated 80° in ecliptic longitude to each side of the apex. The modeling results reproduce successfully some diurnal variation features of the ionized meteor input rates observed by high power and large-aperture radars at different locations in different seasons. For instance, the ionized

meteor input rate observed at each location almost always shows a consistent increase throughout each night with a maximum near dawn. The magnitude of the rate variation during each night is beyond one order. Different from the nocturnal variation feature of the observed ionized meteor input rate, the observed Fe and Na abundances do not show a consistent increase during each night [*Kane and Gardner*, 1993]. Moreover, in contrast to more than one order of the variations in the ionized meteor rate during each night, the Fe and Na abundances show often a relatively small nocturnal variation. In order to clarify whether there was a connection between the ionized meteor input and the Fe and Na layer abundances, we compared the temporal variations of the ionized meteor input rate and the total Fe and Na abundances observed simultaneously by closely colocated meteor radar and two resonance fluorescence lidars at Wuhan [*Yi et al.*, 2007]. It was found that the ionized meteor input rate recorded by radar had no clear-cut relation with the abundance enhancements of the mesospheric metal atom layer. This is accordant with the observations by *Gerding et al.* [1999], who discerned the inconsistency between the occurrences of the lidar-observed atom meteor trail enhancements and the increased ionized meteor activity recorded by forward meteor scatter radio stations.

[5] As pointed out by *Janches et al.* [2006], in the model of the mesospheric Na layer, *Plane* [2004] used a micrometeor input function including seasonal and diurnal variations based on the radio meteor survey (ionized meteor input rate) by *Yrjölä and Jenniskens* [1998]. However, the observations mentioned above indicate that the ionized meteor input rate has no clear-cut relation with the metal atom layer abundance. Therefore it is necessary to give an explanation for this phenomenon. A meteor radar only samples a subset of the mass/velocity distribution of entering meteoroids. If the idea that meteoric ablation is the major source of the mesospheric metal atoms is still maintained, this phenomenon appears to imply that the features (mass/velocity distribution or flux, etc.) of the subset associated with the ionized meteoric trails differ in general from those of that makes primary contribution to the metal atom layers. Hence we conjecture that the micrometeoroids, that are too small to yield the ionized meteoric trails observed by meteor radars but have a large population, may be the source of the mesospheric metal atoms.

[6] Since the formation of the metal atom layers is related to the input of gas-phase metal materials from meteoroid ablation, dynamic transport and chemical reactions, exploring the relationship between various layers in the mesopause region may help clarify which process is playing a predominant role in the layer formation, particularly given the lack of quantitative information on the metal deposition flux. The simultaneous and common-volume observations of the normal Fe and Na layers at a midlatitude site showed that even though the individual density profiles appeared quite different, the temporal variations of Fe and Na layer parameters (abundances, centroid heights, and RMS widths) were evidently similar in most cases [*Kane and Gardner*, 1993]. This phenomenon was ascribed to dynamic effects [*Kane and Gardner*, 1993]. From concurrent measurements of the charged dust and atomic metals (Fe and Na) in the polar winter mesosphere by sounding rocket and lidars, some interesting relations between the charged dust and the

Fe layer have been found [*Lynch et al.*, 2005; *Gelinas et al.*, 2005]. There exists a strong correlation between the dust density and the Fe density for an altitude range from 82 to 95 km over all four rocket flights [*Lynch et al.*, 2005]. Although the dust layer lay a few kilometers lower than the Fe layer, its shape was similar to the shape of the Fe layer [*Gelinas et al.*, 2005]. The two species appeared to move in unison during the three rocket launches on one night, with the layers compressing and tracking downward together. These relations are surprising, as the density distribution and variation of the dust layer depend on local mesospheric dynamics (eddy diffusion, neutral wind, etc.) and characteristics of the incident meteoroids (velocity, composition, etc.) [*Hunten et al.*, 1980; *Kalashnikova et al.*, 2000] rather than the gas-phase chemistry. Obviously, clarifying the relations needs further experimental and modeling efforts.

[7] Upper extent and underside are critical regions of a normal metal atom layer, where metal atoms manifest a transition from absence to presence (or vice versa). Such transition regions may facilitate identifying the practical physics and chemistry behind the observed layer structure because therein some production mechanisms of the metal atoms likely begin or cease to operate or some removal and/or conversion actions are likely strong enough to cause a cutoff of metal atom density. Therefore exploring the behaviors at the metal atom layer boundaries may help to recognize the processes leading to the formation of the normal metal layer. Recently, it has been revealed by comparing simultaneous and common-volume observations of different resonance fluorescence lidars that the topside layers above 110 km for different metal atoms have remarkably similar seasonal characteristics [*Höffner and Friedman*, 2004, 2005]. In addition, a qualitative agreement is also found for K at two distinct latitudes (54°N and 18°N) [*Höffner and Friedman*, 2004]. The seasonal characteristics differ clearly from those of the main layer below. The calculated metal abundance ratios are fairly constant over an altitude range within the topside layers throughout all seasons, which are in reasonably good agreement with metal abundance ratios determined from simultaneous measurements inside single meteor trails within the main layer. A comparison of metal densities at 113 km altitude with known meteor showers indicates a strong influence of shower meteoroids on the topside layers. Moreover, the metal density enhancement in summer correlates well with the seasonal variation of sporadic micrometeoroid input independent of meteor showers. Taking these evidences into consideration, *Höffner and Friedman* [2004] argued a direct link between ablating meteoroids and the topside metal layers. The findings about the metal layer topside [*Höffner and Friedman*, 2004, 2005] are calling for more attention to the metal layer upper extent and underside.

[8] Lidar measurements also reveal another layer form called "sporadic metal layers." The first finding about them was from São José Dos Campos (23°S, 46°W) by *Clemesha et al.* [1978]. Sporadic metal layers are characterized by large density enhancements in a narrow altitude range [*Hansen and von Zahn*, 1990; *Clemsha*, 1995; *Clemsha et al.*, 1999; *Yi et al.*, 2002]. They occur usually in the upper part of the normal metal layers and have no clear-cut relation with the normal layers. The high correlation in time

and space between Na_s and E_s layers led many scientists to explore a possibility of Na_s layer formation via the neutralization of a reservoir of Na^+. Taking the idea proposed by *Cox and Plane* [1998], *Collins et al.* [2002] successfully simulated the sporadic layering events observed simultaneously by the Na lidar and incoherent scatter radar at Arecibo. Their modeling work showed an excellent agreement with observations in five out of a total of six case studies. However, some additional observational facts are worth noting. It has been observed that the amount of Na^+ ions in E_s layers is not sufficient to form large Na_s layers [*Kane et al.*, 1993]. An observation at Arecibo Observatory showed that the electron density peak of a E_s layer lagged behind that of the correlated Na_s layer [*Friedman et al.*, 2000]. This phenomenon makes the E_s an unlikely source for the formation of sporadic atomic layers because it appeared to contradict the conventional cause/effect relation in the ion reservoir neutralization mechanism. The simultaneous and common-volume Fe and Na lidar observations indicate that the Fe_s and Na_s layers occur in overlapping altitude ranges and moved following almost the same track [*Yi et al.*, 2007]. On occasion, the Fe_s and Na_s layers exactly simultaneously reach their maximum peak densities at nearly the same altitude. These observational results suggest that Fe_s and Na_s layers come from the same source substance and may be formed via the same or similar atomization process. For the majority of the sporadic Na layers observed by a Na wind/temperature lidar at low latitudes, their temperatures were significantly warmer than the mean [*Qian et al.*, 1998]. The narrow-band Na lidar measures the Na atom temperature. There exists a possibility that the Na atoms in the sporadic Na layers are freshly generated by more violent meteoric ablation and then have evidently higher temperature than the mean. A systematic study on the variations of occurrence of sporadic Ca and Ca^+ layers indicates that the neutral and ionized Ca layers are not closely correlated [*Gerding et al.*, 2001]. This result is inconsistent with the recombination mechanism that sporadic layers are produced by the neutralization of metal ions. On the other hand, it seems to support that the neutral and ionized metal layers may come from different meteoric source if the metal atoms and ions are believed to come from the direct meteoric ablation. Obviously, clarifying the formation mechanism of sporadic layering events needs further observational and modeling efforts.

[9] In order to study the mesospheric Fe and Na layers at 30°N, we established two independent resonance fluorescence lidars [*Yi et al.*, 2007]. The two lidar systems have the same altitude and time resolutions (96 m and 5 min) as well as the identical starting and ending times for each sampling. This allows us to make a detailed comparison between Fe and Na density profiles. From one year of simultaneous and nearly common-volume Fe and Na lidar measurements at Wuhan (30.5°N, 114.4°E), we have found that the lower boundaries of the normal Fe and Na layers consistently show a quasi-overlap [*Yi et al.*, 2007]. In addition, the normal Na layer always showed a few kilometers more extension on its top side than the normal Fe layer. These features can be clearly seen also from a comparison of simultaneously observed Fe and Na layers at high, mid and low latitudes respectively [*Kane and Gardner*, 1993; *Lynch et al.*, 2005; *Shibata et al.*, 2006]. Thus these features may
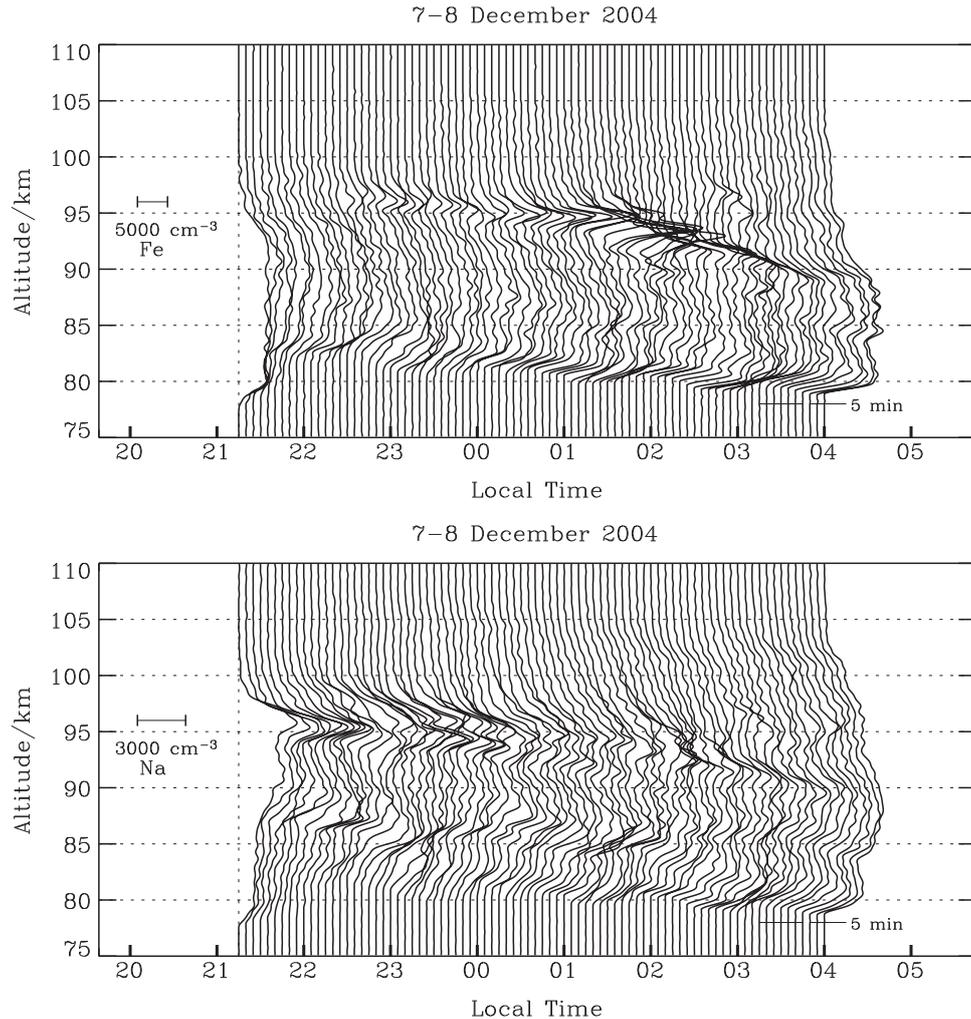
Figure 1. Sequences of density profiles measured during the night of 7–8 December 2004 for (top) Fe and (bottom) Na. Note that the undersides of the normal Fe and Na layers followed nearly the exact movements and occurred at very similar altitudes.

be universal, in which case they may provide a connection between the two different metallic species and reflect the physics and/or chemistry in their formation. Therefore, here we examine further the relationship between the normal Fe and Na layers on their underside and upper extent.

2. Observational Results

[10] The Fe and Na density data employed in this study come from the 195 h simultaneous and nearly common-volume Fe and Na lidar measurements at Wuhan during 26 nights from 8 March 2004 to 14 March 2005, which have also been used for studying the relationship between sporadic Fe (Fe_s) and Na (Na_s) layers [*Yi et al.*, 2007]. Figure 1 gives a comparison between the time sequences of Fe and Na density profiles measured during the night of 7–8 December 2004. On this winter night, despite the existence of a weak single Fe_s layer near 93.0 km around 0140 LT the absence of any high-altitude sporadic layering event (~100 km or above) makes it possible to determine the upper boundaries of the normal layers. As seen in Figure 1, the main bodies of the simultaneously observed

Fe and Na layers appeared quite different at any given time, but their undersides followed one another nearly exactly. The normal Fe layer is several kilometers narrower than the Na layer nearly at all time, and this difference was primarily reflected in the extent of the upper edge of the layer.

[11] For quantitatively characterizing the border characteristics of the normal Fe and Na layers, we define their upper and lower boundaries in terms of a given small density criterion. In more detail, the upper (or lower) boundary for each density profile is defined as such an altitude where the density value of the normal layer becomes equal to or less than the given density criterion as one moves from the layer inside to its high (or low) border. An idealized choice of the density criterion is the cutoff point of the normal layer (null density point). However, near the cutoff altitude the lidar detected photon counts are close to a background noise level so that the upper (or lower) boundary obtained in terms of the null density criterion may have a large uncertainty. For choosing a suitable density criterion we calculate the detection thresholds of Fe and Na atoms between about 75 and 110 km for
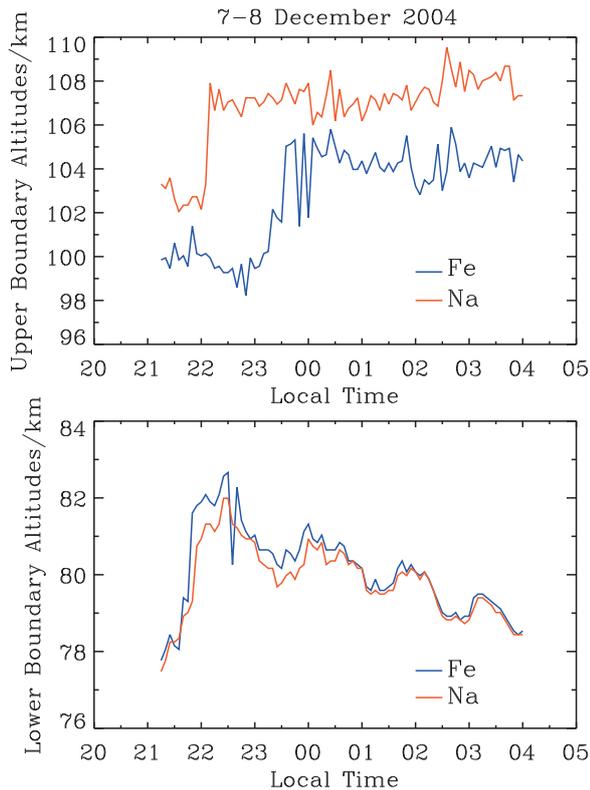
Figure 2. (top) Upper and (bottom) lower boundary altitudes of the normal Fe and Na layers as a function of time on the night of 7–8 December 2004. Note that the two lower boundaries varied following almost the same track and the lower boundary of the Fe layer was in general slightly higher than or coincident with that of the Na layer, with a mean altitude difference being about 0.3 km. The upper boundary of the Fe layer was always lower than that of the Na layer, with a mean altitude difference being about −4 km.

all density profiles shown in Figure 1. It is noticed that their maxima are 23 cm^{-3} (Fe) and 7 cm^{-3} (Na), respectively. Thus, the density criterion for both the Fe and Na layers is set to 30 cm^{-3}. According to the above definition, the upper and lower boundaries of the normal Fe and Na layers are determined respectively and plotted versus time in Figure 2. As seen from the lower panel, the lower boundaries of the Fe and Na layers show consistently a delicate stratification that the lower boundary of the Fe layer is in general slightly higher than or coincident with that of the Na layer, with a mean altitude difference being about 0.3 km (three exceptional time points, i.e., 2130, 2135, and 2235 LT, may result from the low signal-to-noise of our Fe lidar). The two lower boundaries exhibit also considerable vertical displacements on a timescale of a few tens of minutes, particularly with a displacement magnitude of 4.8 km between 2115 and 2230 LT, which modulate the delicate stratification. Despite the existence of the considerable vertical displacements, the two lower boundaries vary following almost the same track with a correlation coefficient being as high as ∼0.93. Here, it is interesting to examine whether the observed DC offset in low boundary altitudes is a geophysical phenomenon or

an effect of zenith pointing errors between the two lidar beams. By using a new method developed by our group [*Liu et al.*, 2005], the optical axes of our Fe and Na lidar telescopes have been adjusted to point to the zenith accurately (with a maximum error of 0.2 mrad in off-zenith angle). For each lidar system in our group there is an autocollimating device above the computer-controlled gimbal-mounted mirror that directs the laser beam toward to the zenith. The autocollimating device guarantees the zenith pointing error of a laser beam being within 0.15 mrad. Hence, the maximum altitude difference induced by the zenith pointing error between the two lidar beams would be roughly 1.8×10^{-3} m that is too small to yield the observed DC offset of in altitudes (hundreds of meters or more). Therefore, the observed DC offset in low boundary altitudes results from geophysical cause. The above observed results suggest strongly that the undersides of the Fe and Na layers are controlled by the same or very similar mechanisms.

[12] As shown in the upper panel of Figure 2, the upper boundary of the Fe layer is obviously lower than that of the Na layer, with a mean altitude difference of about 4 km. The correlation coefficient between them (the two upper boundaries) has a value of ∼0.50, which is obviously lower compared to that of the lower boundaries. However, if the density criterion is set to 200 cm^{-3} the correlation coefficient grows to ∼0.72, while the corresponding value for the lower boundaries increases to ∼0.98. This result appears to indicate that at those altitudes more close to the null density points, the stronger noise fluctuation would reduce the correlation between the upper (or lower) boundaries of the Fe and Na layers. In order to understand why the correlation between the two upper boundaries is obviously weaker than that between the two lower boundaries, we have surveyed all the Fe and Na density profiles shown in Figure 1. It is found that the upper extent of the metal layers is characterized by the frequent occurrence of weak density enhancements in contrast to the underside. They actually represent sporadic layering events with very small absolute peak densities [*Yi et al.*, 2007]. Weak density enhancements occurring frequently on the upper extent may impair the correct determination of the upper boundaries of the normal metal layers. This is confirmed by a fact that the calculated upper boundaries usually show obviously stronger irregular fluctuations on small temporal scales than the lower boundaries (see Figure 2). Therefore, the low correlation between the two upper boundaries is at least partly due to these weak density enhancements prevailing on the upper extent of the metal layers.

[13] It is noted from Figure 2 that the upper boundary appears to be uncorrelated with the lower boundary for each of the metal layers. Combining them and using our boundary condition, we obtain the mean widths of 22.7 km for the Fe layer and 27.0 km for the Na layer. Hence, the Na layer is on average about 4.3 km wider than the Fe layer, and this difference is mostly reflected in the extent of the Na layer top. For further discerning the relationship between the Fe and Na layer structures on the underside and upper extent, a typical example of the individual Fe and Na density profiles is presented in Figure 3. It is observed from Figure 3 that both of the layers show an evidently steeper density gradient on the underside than on the upper extent, and both the Fe underside and upper extent are clearly steeper than those of
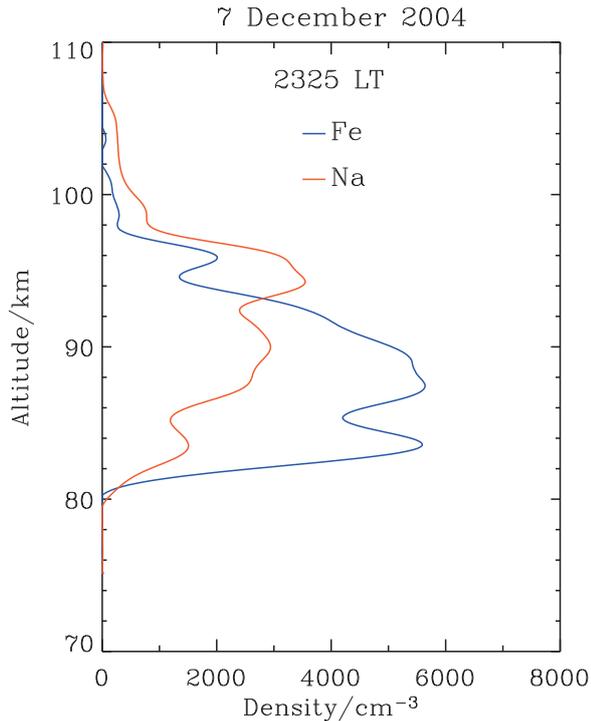
Figure 3. Fe and Na density profiles measured at 2325 LT on 7 December 2004. Note that the upper boundary of the Fe layer is ~5.2 km lower than that of the Na layer, while the lower boundary of the Fe layer is ~0.6 km higher than that of the Na layer. Both of the layers show an evidently steeper density gradient on the underside than on the topside, and both the Fe layer underside and topside are clearly steeper than those of the Na layer.

the Na layer. In order to ascertain whether the density gradient features are prevalent over this night, we have calculated the density scale heights around the upper and lower boundaries for all the Fe and Na density profiles shown in Figure 1. As a result, out of the total 81 pairs of the Fe and Na density profiles, 72 (88%) have the density gradient features. The mean density scale heights around the upper boundary are 0.36 ± 0.15 km for Fe and 1.2 ± 0.75 km for Na, while the corresponding values around the lower boundary are -0.18 ± 0.12 km and -0.27 ± 0.16 km. At the upper boundary of the metal atom layers near 106 km, the diffusion for each metal constituent is believed to be dominated by both molecular and eddy effects. Molecular diffusion tends to drive the density falloff of any given constituent to its natural scale height, while eddy diffusion tends to drive the density falloff to the background atmospheric scale height [*Heinselman*, 1999]. The natural scale heights for Fe and Na atoms have the values of 3.2 and 7.9 km, respectively, at 106 km altitude, which are comparable to the background atmospheric scale height of 6.7 km at this altitude. Noticing the fact that the measured Fe and Na scale heights around their upper boundaries are consistently far less than those values of the natural and background atmospheric scale heights, we believe that the altitude variation in Fe and Na densities on the layer upper extent must represent some features of source and/

or sink for Fe and Na atoms as conjectured earlier by *Tilgner and von Zahn* [1988]. At the lower boundaries of the Fe and Na layers, the consistently negative and small values of the measured Fe and Na scale heights reflect the significant gradients in the source and/or sink for Fe and Na atoms near the lower boundary.

[14] As an example from summertime, Figure 4, which is based on the sequences of the Fe and Na density profiles shown by *Yi et al.* [2007] (Figure 1), displays the temporal variations of the upper and lower boundaries for the normal Fe and Na layers observed on the night of 28−29 July 2004. Since the maximum detection thresholds for all the Fe and Na density samplings (5-min) are 13 cm^{-3} and 5 cm^{-3}, respectively, the density criterion here is still taken to be 30 cm^{-3}. Note that the upper boundaries of the normal layers are uncertain between 0015 to 0500 LT because of the presence of the high-altitude Fe$_s$ and Na$_s$ layers [*Yi et al.*, 2007]. It is clear from the upper panel of Figure 4 that the upper boundary of the Fe layer is over 8 km lower than that of the Na layer before 0015 LT. The two upper boundaries exhibit some features similar to those observed in the wintertime example, i.e., strong irregular fluctuations and a low correlation compared with the lower boundaries. Just
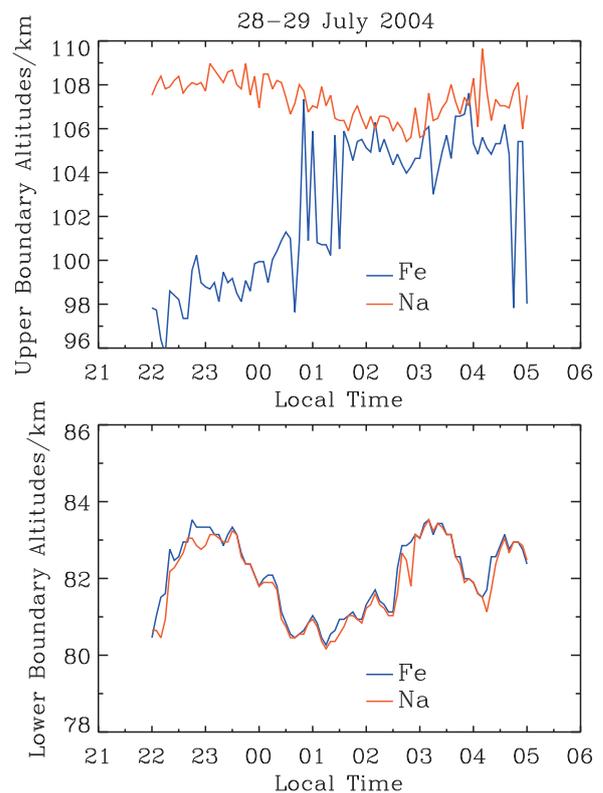
Figure 4. (top) Upper and (bottom) lower boundary altitudes of the normal Fe and Na layers as a function of time on the night of 28−29 July 2004. Note that the two lower boundaries varied following almost the same track and the lower boundary of the Fe layer was in general slightly higher than or coincident with that of the Na layer, with a mean altitude difference being about 0.2 km. The upper boundaries of the normal layers are uncertain between 0015 to 0500 LT because of the presence of the high-altitude Fe$_s$ and Na$_s$ layers.
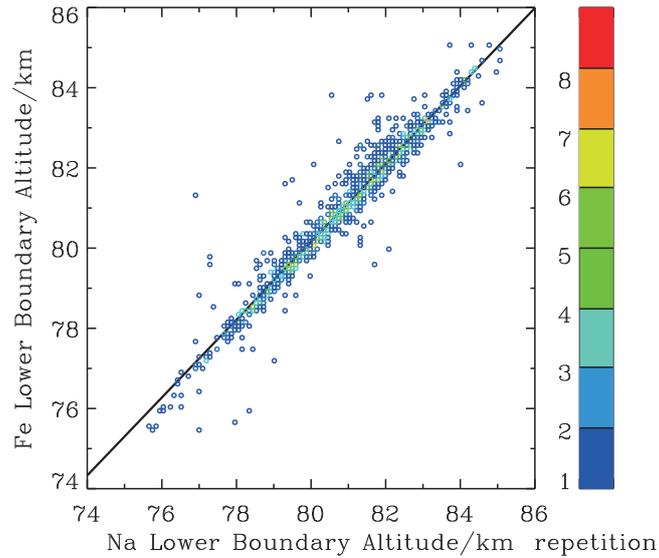
Figure 5. Scatterplot showing the relationship between the lower boundary altitudes of the normal Fe and Na layers based on the 1075 pairs of the high-accuracy Fe and Na density profiles from the 12 nights. The color of each open circle denotes the repetition number on this data point. The straight line is a linear regression fit to the data. The correlation coefficient is 0.96, indicating that it is a ubiquitous phenomenon that the lower boundaries of normal Fe and Na layers follow nearly the same movement. Note that the lower boundary of the Fe layer mostly lies at a slightly higher altitude than or coincides with that of the Na layer.

like the wintertime example above, the lower boundaries of the summertime layers displayed consistently the delicate stratification that the low boundary of the Fe layer was mostly slightly higher than or accordant with that of the Na layer, with a mean altitude difference being about 0.2 km (six exceptional time points may result from the noise fluctuation, see Figure 4, bottom). The two lower boundaries also follow nearly the same movement with a correlation coefficient as high as 0.96. The delicate stratification appears to be modulated by the movement. As shown in the low panels of Figure 2 and 4, the consistently subtle stratification that the lower boundary of the Fe layer is in general slightly higher than or coincident with that of the Na layer makes it unlikely that the correlated variations of the measured lower boundaries are simply caused by gravity waves alone because at the lower boundaries near 80 km altitude the vertical mixing of constituents by turbulence generated by gravity wave breaking tends to spoil the stratification. The gradient features shown in most individual Fe and Na profiles (89%) near the upper and lower boundaries are also consistent with those observed from the wintertime example, i.e., there exists an evidently steeper density gradient on the underside than on the upper extent, and the Fe layer borders are clearly steeper than those of the Na layer. In the summertime example, the mean density scale heights around the upper boundary are 0.40 ± 0.23 km for Fe and 1.7 ± 0.73 km for Na, while the corresponding values around the lower boundary are -0.16 ± 0.03 km and -0.22 ± 0.10 km. The scale height values are comparable with those in the wintertime example.

[15] In order to identify the statistical significance of the above observed results, we have further surveyed all the simultaneous Fe and Na data. Note that out of a total of 26 observation nights, 12 nights (including two examples

shown above) have the maximum detection thresholds less than 30 cm^{-3} due to a variational atmospheric transmission from night to night. These 12 nights with high detection accuracy (high signal-to-noise ratio) allow us to explore the behaviors of the layer boundaries close to the density cutoff altitudes. Figure 5 contains a scatter plot showing the relationship between the lower boundaries of the normal Fe and Na layers based on the 1075 pairs of the high-accuracy Fe and Na density profiles from the 12 nights. The straight line represents a linear regression fit to the data. The overall correlation coefficient is 0.96, indicating that it is a ubiquitous phenomenon that the lower boundaries of the normal Fe and Na layers follow nearly the same movement. Figure 5 reveals also that the lower boundary of the Fe layer mostly lies at a slightly higher altitude than or coincides with that of the Na layer, with a mean altitude difference being about 0.2 km. This is clearly seen in 847 (79%) out of the total 1075 pairs of the Fe and Na density profiles. The remaining 228 pairs (21%) may be ascribed to the noise contamination or tiny structure on underside. The correlation coefficient between the upper boundaries of the Fe and Na layers is low (0.37) because of the impairment of weak density enhancements occurring frequently on the upper extent of the metal layers. However, it is clear that the upper boundary of the Fe layer is always several kilometers lower than that of the Na layer in the absence of high-altitude sporadic layers. Our statistics reveals again that there is no clear-cut correlation between the upper and lower boundaries for each of the metal layers.

[16] Both the normal Fe and Na layers over Wuhan display an evidently steeper density gradient on underside than on upper extent as seen from the two examples above. The feature is consistent with the earlier observations at high and middle latitudes [*Tilgner and von Zahn*, 1988;

Kane and Gardner, 1993]. In order to quantitatively determine the consistency of the feature at our site, we have calculated the density gradients near the upper and lower boundaries for all the high-accuracy Fe and Na density profiles in the absence of sporadic layers around the upper extent of the normal layers. As a result, the feature (steeper density gradient on the underside than on the upper extent) is seen for ∼90% of all the Fe profiles and for 86% of all the Na profiles. In addition, it is also noticed from the Fe and Na density profiles that the Fe layer generally has obviously steeper density gradients on its borders than the Na layer (e.g., Figure 3). This agrees with the available observational fact at other midlatitude sites that the Fe layer has an extremely small bottom scale height compared with the Na layer [*Plane*, 2003].

[17] For the convenience of depicting the observed layer boundaries, we have introduced a digital definition. It should be emphasized that the chosen density criterion (30 cm^{-3}) is an extremely small quantity (comparable to the minimum detectable density) rather than an arbitrarily chosen value. The validity of the definition has been tested further by comparing the calculated boundary altitudes with all the individual pairs of the Fe and Na profiles. It is indicated that the altitude relations between the calculated Fe and Na boundaries are consistent with those seen in nearly all the individual pairs of the Fe and Na profiles. Furthermore, the lower boundary altitude relation, that the lower boundary of Fe layer is in general slightly higher than or coincident with that of Na layer, is also consistent with the observed fact that the Fe underside is generally obviously steeper than that of the Na layer.

[18] In addition, only from a visual observation of all the nights of the sequences of the Fe and Na density profiles observed simultaneously, we can clearly judge that the undersides of the normal Fe and Na layers follow nearly exact movement and occur at nearly the same altitude (e.g., Figure 1). In this case, the correlation coefficient is not sufficient to depict the relation between the Fe and Na lower boundaries. Thus, the undersides of Fe and Na layers are likely to be formed by the same or very similar processes.

3. Summary

[19] The high-accuracy atom density profiles, obtained by the simultaneous and common-volume Fe and Na lidar measurements at Wuhan, China over 1 year, reveal some prevalent relations between the Fe and Na layer borders. The lower boundaries of the Fe and Na layers show consistently a delicate stratification in which the lower boundary of the Fe layer is in general slightly higher than or coincident with that of the Na layer, with an overall mean altitude difference being about 0.2 km. The delicate stratification is often modulated by the highly correlative vertical movements of the two lower boundaries. The overall correlation coefficient between the two lower boundary altitudes is as high as 0.96, indicating that the delicate stratification is a ubiquitous phenomenon. These results suggest strongly that the undersides of Fe and Na layers are controlled by the same or very similar processes. The upper boundary of the Na layer is always several kilometers higher than that of the Fe layer. A relatively weak positive

correlation is also persistently observed between the two upper boundaries. Weak sporadic layering events frequently occur on the upper extent of the metal layers. They may impair the correct determination of the upper boundaries of the normal metal layers and consequently weaken the correlation. Both the Fe and Na layers often show an evidently steeper density gradient on the underside than on the upper extent, and the borders of the Fe layer are clearly steeper than those of the Na layer. The explanation to these observed relations needs further work in both observation and modeling.

References

Clemesha, B. R. (1995), Sporadic neutral metal layers in the mesosphere and lower thermosphere, *J. Atmos. Terr. Phys.*, *57*, 725–736.

Clemesha, B. R., V. W. J. H. Kirchhoff, D. M. Simonich, and H. Takahashi (1978), Evidence of an extraterrestrial source for the mesospheric sodium layer, *Geophys. Res. Lett.*, *5*, 873–876.

Clemesha, B. R., P. P. Batista, and D. M. Simonich (1999), An evaluation of the evidence for ion recombination as a source of sporadic neutral layers in the lower thermosphere, *Adv. Space Res.*, *24*, 547–556.

Collins, S. C., et al. (2002), A study of the role of ion–molecule chemistry in the formation of sporadic sodium layers, *J. Atmos. Sol. Terr. Phys.*, *64*, 845–860.

Cox, R. M., and J. M. C. Plane (1998), An ion-molecule mechanism for the formation of neutral sporadic Na layers, *J. Geophys. Res.*, *103*, 6349–6359.

Delgado, R., B. R. Weiner, and J. S. Friedman (2006), Chemical model for mid-summer lidar observations of mesospheric potassium over the Arecibo Observatory, *Geophys. Res. Lett.*, *33*, L02801, doi:10.1029/2005GL024326.

Eska, V., J. Höffner, and U. Von Zahn (1998), The upper atmosphere potassium layer and its seasonal variability at 54°N, *J. Geophys. Res.*, *103*, 29,207–29,214.

Eska, V., U. von Zahn, and J. M. C. Plane (1999), The terrestrial potassium layer (75–110 km) between 71°S and 54°N: Observations and modeling, *J. Geophys. Res.*, *104*, 17,173–17,186.

Friedman, J. S., S. A. Gonzalez, C. A. Tepley, Q. Zhou, M. P. Sulzer, S. C. Collins, and B. W. Grime (2000), Simultaneous atomic and ion layer enhancements observed in the mesopause region over Arecibo during the Coqui II sounding rocket campaign, *Geophys. Res. Lett.*, *27*, 449–452.

Friedman, J. S., S. C. Collins, R. Delgado, and P. A. Castleberg (2002), Mesospheric potassium layer over the Arecibo Observatory, 18.3°N 66.75°W, *Geophys. Res. Lett.*, *29*(5), 1071, doi:10.1029/2001GL013542.

Gardner, C. S., J. M. C. Plane, W. Pan, T. Vondra, B. J. Murray, and X. Chu (2005), Seasonal variations of the Na and Fe layers at the South Pole and their implications for the chemistry and general circulation of the polar mesosphere, *J. Geophys. Res.*, *110*, D10302, doi:10.1029/2004JD005670.

Gelinas, L. J., K. A. Lynch, M. C. Kelley, R. L. Collins, M. Widholm, E. MacDonald, J. Ulwick, and P. Mace (2005), Mesospheric charged dust layer: Implications for neutral chemistry, *J. Geophys. Res.*, *110*, A01310, doi:10.1029/2004JA010503.

Gerding, M., M. Alpers, J. Höffner, and U. von Zahn (1999), Simultaneous K and Ca lidar observations during a meteor shower on March 6–7, 1997, at Kühlungsborn, Germany, *J. Geophys. Res.*, *104*, 24,689–24,698.

Gerding, M., M. Alpers, U. von Zahn, R. J. Rollason, and J. M. C. Plane (2000), Atmospheric Ca and Ca$^+$ layers: Midlatitude observations and modeling, *J. Geophys. Res.*, *105*, 27,131–27,146.

Gerding, M., M. Alpers, J. Höffner, and U. von Zahn (2001), Sporadic Ca and Ca+ layers at mid-latitudes: Simultaneous observations and implications for their formation, *Ann. Geophys.*, *19*, 47–58.

Hansen, G., and U. von Zahn (1990), Sudden sodium layers in polar latitudes, *J. Atmos. Terr. Phys.*, *52*, 585–608.

Heinselman, C. J. (1999), Auroral effects on meteoric metals in the upper atmosphere, Ph.D. dissertation, Stanford Univ., Stanford, Calif.

Helmer, M., J. M. C. Plane, J. Qian, and C. S. Gardner (1998), A model of meteoric iron in the upper atmosphere, *J. Geophys. Res.*, *103*, 10,913–10,925.

Höffner, J., and J. S. Friedman (2004), The mesospheric metal layer topside: a possible connection to meteoroids, *Atmos. Chem. Phys.*, *4*, 801–808.

Höffner, J., and J. S. Friedman (2005), The mesospheric metal layer topside: Examples of simultaneous metal observations, *J. Atmos. Sol. Terr. Phys.*, *67*, 1226–1237.

Hunten, D. M., R. P. Turco, and O. B. Toon (1980), Smoke and dust particles of meteoric origin in the mesosphere and stratosphere, *J. Atmos. Sci.*, *37*, 1342–1357.

Janches, D., C. J. Heinselman, J. L. Chau, A. Chandran, and R. Woodman (2006), Modeling the global micrometeor input function in the upper atmosphere observed by high power and large aperture radars, *J. Geophys. Res.*, *111*, A07317, doi:10.1029/2006JA011628.

Kalashnikova, O., M. Horanyi, G. E. Thomas, and O. B. Toon (2000), Meteoric smoke production in the atmosphere, *Geophys. Res. Lett.*, *27*, 3293–3296.

Kane, T. J., and C. S. Gardner (1993), Structure and seasonal variability of the nighttime mesospheric Fe layer at midlatitudes, *J. Geophys. Res.*, *98*, 16,875–16,886.

Kane, T. J., C. S. Gardner, Q. Zhou, J. D. Mathews, and C. A. Tepley (1993), Lidar, radar and airglow observations of a prominent sporadic Na/sporadic E layer event at Arecibo during AIDA-89, *J. Atmos. Terr. Phys.*, *55*, 499–511.

Liu, B., F. Yi, and C. M. Yu (2005), Methods for optical adjustment in lidar systems, *Appl. Opt.*, *44*, 1480–1484.

Lynch, K. A., L. J. Gelinas, M. C. Kelley, R. L. Collins, M. Widholm, D. Rau, E. MacDonald, Y. Liu, J. Ulwick, and P. Mace (2005), Multiple sounding rocket observations of charged dust in the polar winter mesosphere, *J. Geophys. Res.*, *110*, A03302, doi:10.1029/2004JA010502.

Plane, J. M. C., C. S. Gardner, J. Yu, C. Y. She, R. R. Garcia, and H. C. Pumphrey (1999), Mesospheric Na layer at 40°N: Modeling and observations, *J. Geophys. Res.*, *104*, 3773–3788.

Plane, J. M. C. (2003), Atmospheric chemistry of meteoric metals, *Chem. Rev.*, *103*, 4963–4984.

Plane, J. M. C. (2004), A new time-resolved model for the mesospheric Na layer: Constraints on the meteor input function, *Atmos. Chem. Phys.*, *4*, 627–638.

Qian, J., Y. Gu, and C. S. Gradner (1998), Characteristics of the sporadic Na layers observed during the airborne lidar and observations of Hawaiian airglow/airborne noctilucent cloud (ALOHA/ANLC-93) campaigns, *J. Geophys. Res.*, *103*, 6333–6347.

Raizada, S., and C. A. Tepley (2003), Seasonal variation of mesospheric iron layers at Arecibo: First results from low-latitudes, *Geophys. Res. Lett.*, *30*(2), 1082, doi:10.1029/2002GL016537.

Raizada, S., C. A. Tepley, D. Janches, J. S. Friedman, Q. Zhou, and J. D. Mathews (2004), Lidar observations of Ca and K metallic layers from Arecibo and comparison with micrometeor sporadic activity, *J. Atmos. Sol. Terr. Phys.*, *66*, 595–606.

Shibata, Y., C. Nagasawa, M. Abo, T. Maruyama, S. Saito, and T. Nakamura (2006), Lidar observations of sporadic Fe and Na layers in the mesopause region over equator, *J. Meteorol. Soc. Jpn.*, *84A*, 317–325.

States, R. J., and C. S. Gardner (1999), Structure of the mesospheric Na layer at 40°N latitude: Seasonal and diurnal variations, *J. Geophys. Res.*, *104*, 11,783–11,798.

Tilgner, C., and U. von Zahn (1988), Average properties of the sodium density distribution as observed at 69°N latitude in winter, *J. Geophys. Res.*, *93*, 8439–8454.

Yi, F., S. D. Zhang, H. J. Zeng, Y. J. He, X. C. Yue, J. B. Liu, H. F. Lv, and D. H. Xiong (2002), Lidar observations of sporadic Na layers over Wuhan (30.5°N, 114.4°N), *Geophys. Res. Lett.*, *29*(9), 1345, doi:10.1029/2001GL014353.

Yi, F., S. Zhang, C. Yu, Y. He, X. Yue, C. Huang, and J. Zhou (2007), Simultaneous observations of sporadic Fe and Na layers by two closely colocated resonance fluorescence lidars at Wuhan (30.5°N, 114.4°E), China, *J. Geophys. Res.*, *112*, D04303, doi:10.1029/2006JD007413.

Yrjölä, I., and P. Jenniskens (1998), Meteor stream activity. VI: A survey of annual meteor activity by means of forward meteor scattering, *Astron. Astrophys.*, *330*, 739–752.

————————————
Y. He, C. Huang, W. Li, F. Yi, C. Yu, X. Yue, and S. Zhang, Department of Electronic Engineering, Wuhan University, Wuhan, 430079, China. (yf@whu.edu.cn)