ELSEVIER

Journal of Atmospheric and Solar-Terrestrial Physics 69 (2007) 3–17

Journal of ATMOSPHERIC AND SOLAR-TERRESTRIAL PHYSICS

www.elsevier.com/locate/jastp

Present status and future challenges of modeling the Sun–Earth end-to-end system

D.N. Baker[a,*], M.J. Wiltberger[b], R.S. Weigel[a], S.R. Elkington[a]

[a]*Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO 80309-0590, USA*
[b]*High Altitude Observatory, National Center for Atmospheric Research, Boulder, CO 80301, USA*

Received 25 April 2005; accepted 5 July 2006
Available online 2 January 2007

Abstract

Several research groups are presently modeling key portions of the solar–terrestrial environment. Large-scale models of the solar corona, of the interplanetary medium, and of the coupled magnetosphere–ionosphere system are providing important insights into the dynamics and temporal evolution of these regions during normal, as well as disturbed, conditions. Some collaborative groups are now working on even more grand syntheses of models and are striving for a true end-to-end modeling capability. The challenges of such integrated modeling are immense: vastly different spatial and temporal scales must be addressed and codes must be synchronized to a high degree. Because much of the underlying physics at micro- and meso-scales remains unknown (or only poorly known), many key interface regions must be treated using empirical or semi-empirical methods. Validating the results from such global-scale models is, in itself, a major undertaking. And last, but by no means least, the question of how available (sparse) data are assimilated into end-to-end models remains a major challenge to the community. This paper provides an overview of the present status and future challenges in end-to-end modeling of the Sun–Earth system.

Keywords: Sun–Earth coupling; Global modeling; Geomagnetic storms; Magnetospheric dynamics; Data assimilation

1. Introduction

Recent years have witnessed remarkable progress in the development of empirical and physics-based models of various parts of the solar–terrestrial system (see JASTP special issue, October–November, 2004). Perhaps even more exciting is the progress now being made in coupling together empirical models from the Sun to the Earth (e.g., Baker et al., 2004a), coupling physics-based models over this same domain (e.g., Luhmann et al., 2004), and dramatic development of true predictive models (Siscoe et al., 2004) that can serve the entire space weather user community. The prospects look good for rapid advancement of these efforts toward an accurate, real-time forecast scheme built on acquisition of solar, heliospheric, and near-Earth data (e.g., Siscoe et al., 2004).

However, despite the impressive efforts presently underway, we should not underestimate the significant challenges that lay before the Sun–Earth connections community. There remains a daunting set of tasks to improve the basic building blocks of

*Corresponding author. Tel.: +303 492 4509; fax: +303 735 4843.
E-mail address: Daniel.baker@lasp.colorado.edu (D.N. Baker).

1364-6826/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jastp.2006.07.017

models, to develop suitable code-coupling strategies, to ingest, and assimilate appropriate data sets, and ultimately to provide accessible specification and forecast products that meet users' needs.

This paper describes some of the current elements of empirical and physics-based models for different parts of the Sun–Earth system. It also describes some of the types of data that can be used to "drive" models and/or assess their performance. In each instance, we attempt to describe and evaluate key challenges that confront the community. The paper concludes with a description of planned convergences of empirical and physics models to produce a continuously improving forecast model.

2. End-to-end modeling: the empirical chain

It is envisioned that the ultimate product of the Center for Integrated Space Weather Modeling (CISM) will be a single, physics-based (i.e., "forward") model which will describe with good accuracy the evolution of solar wind elements, the propagation of these elements from the Sun out to Earth's orbit, and the subsequent interaction of the solar wind with the coupled magnetosphere–ionosphere–atmosphere system (Hughes and Hudson, 2004). It is expected that the backbone of this coupled model will be based on magnetohydrodynamic (MHD) numerical codes. Considerable work (as will be described here) must yet be done to assure that such forward models can provide a complete, accurate, and robust description of the Sun–Earth system under realistic conditions.

The challenges of producing an effective end-to-end physics-based model at the present time are considerable. The CISM team has chosen to utilize empirical, semi-empirical, and inverse models to provide a present-day, state-of-the-art forecast model (FM) of the Sun–Earth system (Baker et al., 2004a). As shown in the lower part of the flow diagram in Fig. 1, it is possible to use previously developed empirical methods to observe the Sun, specify the solar boundary conditions, follow the subsequent solar wind propagation to 1 astronomical unit (AU), and thereby forecast solar wind parameters such as mass density (ρ), speed (V), and magnetic field strength (B). With such parameters forecasted in the vicinity of Earth's dayside magnetopause, it is further possible to predict quantities such as geomagnetic indices (e.g., Ap, Kp, or Dst), magnetic field fluctuations at mid- and high-latitudes, global magnetospheric field configurations, and the relativistic electron fluxes in Earth's radiation belts.

As shown in Fig. 1 (top), it has been common now and in the recent past to use measurements from a solar wind monitor at the upstream Lagrangian point (L1) to drive empirical forecast models. However, using L1 data typically allows for only a 30–60 min lead-time forecast. Thus, we have chosen to explore the possibility of pushing the methods back as far as possible in time in order to give 3- to 4-day lead times. The Wang-Sheeley-Arge (WSA) method is used to predict the solar wind speed (and certain other parameters) near Earth's orbit based upon solar surface measurements (Arge et al., 2004; Siscoe et al., 2004). This approach gives a predicted solar wind time series with a lead time of 3–4 days. Such forecasted solar wind conditions can then be convolved with linear and nonlinear filters in order to provide a predicted set of geomagnetic indices or various particle flux estimates.

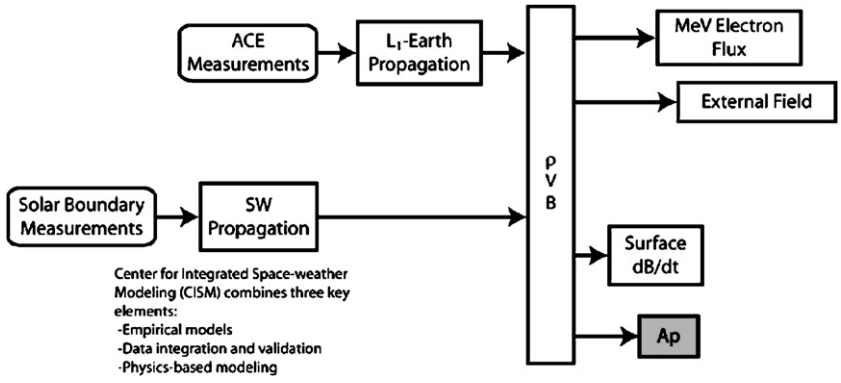

Fig. 1. Linked empirical models being developed within the CISM framework in order to provide forecasts of key parameters and indices as shown to the right (from Baker et al., 2004a).

Baker et al. (2004a) presented a demonstration of an end-to-end empirical forecast of relativistic electrons in the outer Van Allen radiation belt. Prior work showed that radiation belt electron fluxes are highly dependent on the speed of the solar wind striking the magnetosphere. Filters were developed that predict electron fluxes using the WSA estimates of solar-wind speed at L1. This allows for several days lead times. The prediction efficiency (PE) provided by these filters was compared with filters developed to use 3–4 day old values of the solar wind velocity measured at L1 and 3–4 day old values of the measured electron fluxes themselves. It was found that the WSA method provides PEs of the electron flux that are slightly lower than that provided by using old L1 or the autocorrelated electron flux data. However, continuous improvements in the methods have occurred within the past year or so.

An important point concerning Fig. 1 is that the empirical end-to-end models being developed and used today have enduring utility. The current performance of such models constitutes the baseline against which the physics-based models are compared now (and in the future). It is found that the numerical models, at least initially, do not perform as well as the highly tuned, specialized empirical models, or the "inverse" models that were developed in locations that have a long record of historical data available (Weigel et al., 2003; Vassiliadis et al., 2002). However, in time we expect that the forward models will improve and will in many cases outperform the existing models. A key aspect of the National Space Weather Program (NSWP, 2000) is to provide metrics to judge model performance and improvement. The empirical baseline will provide a comparator against which forward models and the composite forecast models can be judged now and over the course of time (see Spence et al., 2004).

3. Status and key challenges of data acquisition and physical modeling

Unless one wishes simply to model an idealized solar–terrestrial interaction, there is a need for strong interplay between observation and analyzed data, on the one hand, and physical models on the other hand. Observations can set initial conditions and can introduce realistic dynamical evolution into numerical simulations of the Sun–Earth system. Specifically, space-based and ground-based measurements of the Sun reveal the magnetic structure of the solar corona and tell where and when active regions may be operative. Solar wind measurements at key points in interplanetary space can tell about large-scale stream structures and can help inform physical models about transient events such as coronal mass ejections (CMEs) and shock waves. Near-Earth space measurements and ground-based observations can provide key information about the state of the magnetosphere–ionosphere system and can help guide (or drive) physical models.

Present global modeling depends on key types of data and it is anticipated that this will continue to be the case into the future. It is likely that improvements in data assimilation methods (e.g., Schunk et al., 2004) will help substantially to improve the accuracy and timeliness of forecast models. A key to a successful predictive modeling strategy is to have continuous, reliable, and well-calibrated data to drive physics-based models. This requires that both the operational agencies (NOAA, DoD, DOE) and scientific agencies (NASA and NSF) provide high-quality, near real-time information. During the present development phase for models, of course, it is possible to use archival data and the 'real-time' requirement can be relaxed.

This section will examine in sequence the various regional domains of the solar–terrestrial system. We will describe observational data that inform the modeling in that region and describe the types of physical models that are under development. We then discuss the challenges and issues that loom as the community moves toward creating an end-to-end modeling framework.

3.1. Solar physics issues and challenges

The data that presently best inform either empirical or physical models of what is happening on, and near, the Sun comes from photospheric field observations. Synoptic maps of the line-of-sight (LOS) magnetograms from ground-based (e.g., Mount Wilson) or space-based (e.g., SOHO) observatories can be used (e.g., Arge et al., 2004). Frequently updated synoptic maps from the Mount Wilson Solar Observatory have proven to be particularly useful and relevant for WSA empirical modeling.

Fig. 2(a) shows a synoptic map of the radial photospheric fields for Carrington Rotation 1961 (courtesy of S. MacGregor). Red colored regions are radially outward-directed fields and dark gray regions are radially inward-directed field lines. Such

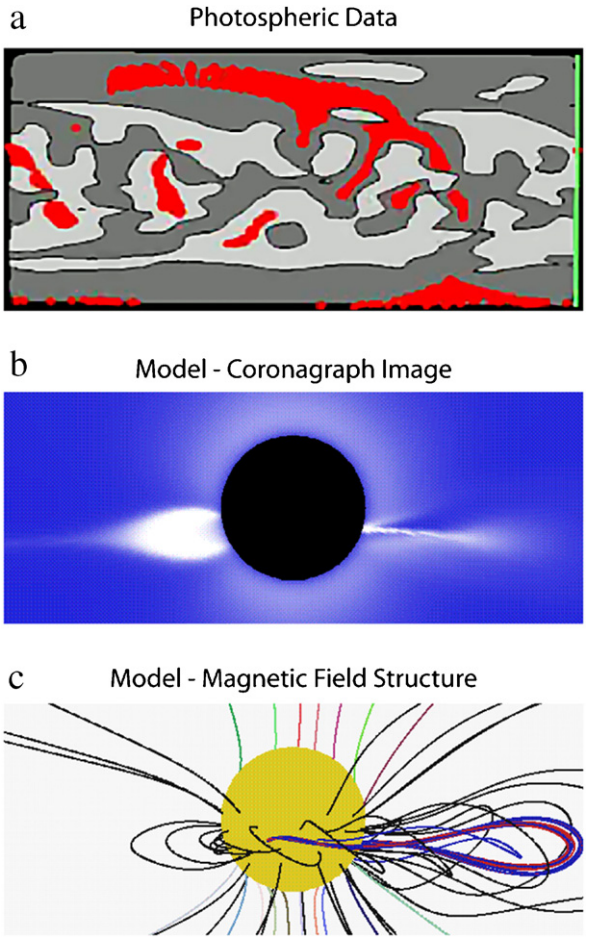

Fig. 2. (a) Radial magnetic field data in the photosphere derived from ground-based measurements for Carrington Rotation 1961 (courtesy of S. MacGregor), (b) 3D MHD simulation of a CME triggered by emerging flux. Scattered white light coronagraph images constructed from the simulation as if the eruption was viewed on the limb arc also shown. Emergence of opposite polarity flux leads to formation of a stable flux rope, and (c) further emergence causing eruption is illustrated by a snapshot of coronal magnetic field lines (courtesy of J. Linker).

synoptic maps permit coronal modeling of fields and provide a basis for estimating solar wind outflow characteristics over the entire solar surface (Arge et al., 2004, and references therein).

Fig. 2(b) shows a simulated coronagraph image (courtesy of J. Linker) portraying a coronal mass ejection. This was created using the magnetohydrodynamics around a sphere (MAS) MHD model. MAS is a 3D representation of the solar corona used to approximate coronal thermodynamical properties and consequences. It uses finite resistivity and viscosity, consistent with the expected effects of coronal processes. The computational grid extends from the photosphere to an outer boundary at 30 solar radii (R_S). The cell spacing is graduated in both the radial and latitudinal dimensions, allowing for higher spatial resolution along a latitude plane of interest, such as the equator. Simulations use a time-dependent potential (current-free) description of the coronal magnetic field as a boundary condition that is based on a global photospheric field map either created or derived from magnetograph observations (see Fig. 2(a)). A uniform boundary density and Parker solar wind outflow completes the specification of initial and boundary conditions.

Time-dependent photospheric inner-boundary conditions are implemented in the MAS code as a finite tangential electric field that is consistent with the time derivative of the boundary magnetic field. The flows implied by this electric field can result in important changes such as flux cancellation at a magnetic neutral line. This is thought to be a key element of coronal eruptions. For evolving magnetic field boundary conditions, the modeled coronal fields open and close with time, producing transient structures with a variety of properties and geometries.

MAS coronal simulations can reproduce the coronal streamer geometrics observed in eclipse and coronagraph pictures when observed photospheric field synoptic maps are used to describe inner boundary conditions (see Luhmann et al., 2004 and references therein). The solar wind produced in this coronal model has features in common with observations. It has a speed gradient at the boundary of open and closed field lines (although the speeds are too low and the flow gradients are weaker than inferred from solar wind observations; the outflow speed, as well as the contrast between the low and high speeds, must be increased by use of a more accurate treatment of the energy equation). The MAS model outflow is supercritical, or supermagnetosonic, at the outer boundary of the MAS grid at $30R_S$. It produces interesting CME-like magnetic field structures as shown in Fig. 2(c) (which corresponds to the 'coronagraph' image in Fig. 2(b)).

From present work in the CISM program and in other programs presently underway, it is clear that substantial progress is being made in the physical modeling of the solar corona and the description of both slow and fast solar wind outflows. However, substantial work remains to be done before one can

use direct solar magnetic field observations to predict where, when, and how a major eruptive event (i.e., CME) will occur. At present, these eruptive features must still be artificially initiated.

There are many key questions and challenges in solar physics to be addressed in the next several years. Among these are:

- What leads to the remarkable 11-year solar activity cycle?
- Can we explain what leads to the Hale polarity law of solar active regions or the hemispheric dependence of magnetic helicity?
- How is the large-scale magnetic field of the Sun maintained?
- How is the large-scale field related to small-scale dynamics (that operates on local convective scales)?
- Quite importantly for space weather modeling, what is the fundamental mechanism that causes a coronal helmet streamer or other such structure to open as a CME?

Addressing such key questions through concerted modeling and continued data analysis will be crucial for future progress in space weather forecasting.

3.2. *Interplanetary physics issues and challenges*

The WSA model is a combined empirical and physics-based representation of the global solar wind flow. It can be used to predict the ambient solar wind speed and interplanetary magnetic field (IMF) polarity at Earth (as well as any other point in the heliosphere). It is therefore useful for forecasting recurrent interplanetary disturbances. It is an improved version of the original Wang and Sheeley (WS) model (Wang and Sheeley, 1992). The model uses ground-based line-of-sight observations of the surface of the Sun's magnetic field as input to a magnetostatic potential field source surface coronal expansion (see Fig. 2(a)). The effects of outward flows in the corona, which are not explicitly contained in the model are included by the imposition of radial boundary conditions at a Sun-centered sphere located at $2.5R_S$.

In the original model, the interplanetary solar wind velocity was determined at each point on the source surface by the relative expansion of the magnetic field from the photospheric base up to that point. This relation between the expansion factor and wind speed is empirical and follows from correlation studies utilizing source surface expansion maps and near-Earth spacecraft observations of solar wind speed (Arge et al., 2004). In addition, the original WS model assumed that the solar wind flow propagates kinematically out from the source surface to Earth, with all dynamics neglected. Several improvements were made to the WS model to bring it to its present state of development (Arge et al., 2004).

Fig. 3(a) is a snapshot from an animation of the WSA output. This shows a color-coded portrayal of solar wind speeds near the ecliptic plane for Carrington Rotation 1896. The particular period shown is 30 May 1995 at 1715 UT. Broad regions of

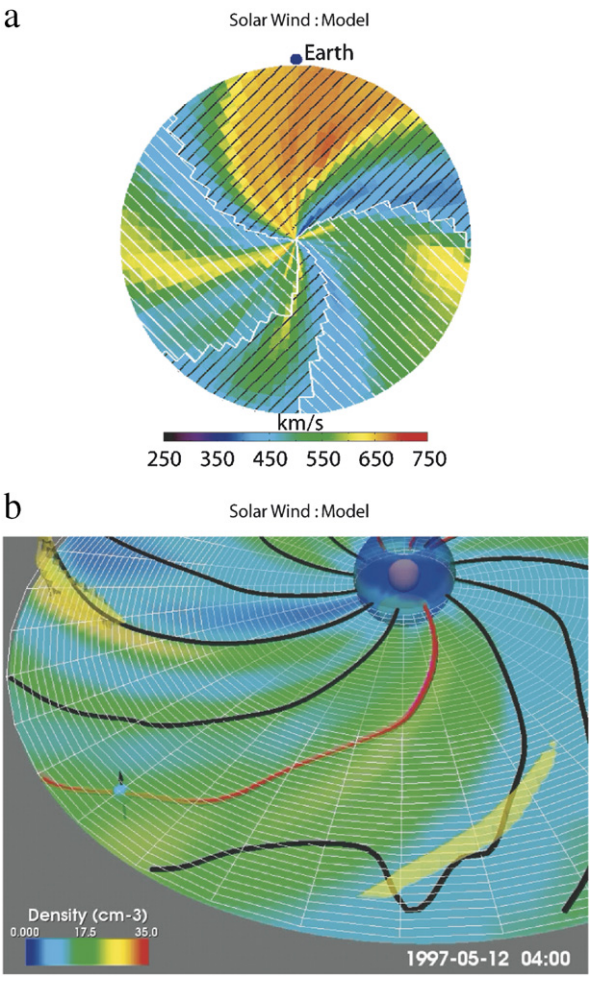

Fig. 3. (a) View of the Wang–Sheeley–Arge model output for a period in May 1995 as described in the text, and (b) the shape and progression of evolving solar wind and coronal transient for a period in May 1997 as simulated by the ENLIL model (courtesy of D. Odstrcil).

high-speed solar wind reaching the vicinity of the Earth are shown for this case, while elsewhere in the inner heliosphere there are broad swaths of low-speed solar wind flow. The entire pattern evolves with time and has a Parker spiral geometry. The WSA model also gives toward-away IMF information. Such data-based model outputs can be very valuable for giving several day lead time forecasts of major space weather events (e.g., Siscoe et al., 2004).

A more physics-based alternative model is shown in Fig. 3(b). In this approach (Odstrcil et al., 2004), the solar wind and coronal transients in the outer grid layers of the MAS model are used to drive the inner boundary of the Enlil solar wind simulation. Enlil is a 3D ideal MHD model (Odstrcil et al., 2002) designed to treat supercritical outflows in the limit where resistivity and viscosity are minimal. Like the MAS code, it is based on a polytropic equation of state. The Enlil model domain extends from the solar equator to within $\pm$ 60° thereby concentrating the uniform latitude–longitude spherical grid on the region of heliospace influencing Earth. This grid has sufficient latitudinal range to minimize the effects of neglecting high latitude behavior on low latitudes. The relative properties of the MAS and Enlil grids are described in detail by Odstrcil et al. (2004). The period shown in Fig. 3(b) is from a CME-dominated event in May 1997 as a solar transient propagates outward toward the Earth. The fidelity of such simulation results can be checked by comparing with ACE or other data from the L1 or other upstream monitoring locations.

Substantial progress is being made in coupling the MAS and ENLIL physics models (e.g., Luhmann et al., 2004) and also in tying together the empirical, semi-empirical, and physics-based (i.e., WSA–ENLIL) models (see Arge et al., 2004). Thus, the prospects are quite good that the next year or two will see dramatic improvements in the modeling of inner heliospheric and near-Earth solar wind conditions. Nonetheless, several significant challenges remain:

- How will we be able to model the speed, trajectory, and temporal evolution of CMEs as they propagate outward toward Earth?
- Can we find methods that can give early, accurate information about the IMF strength and orientation within the outflowing solar wind?
- Can we produce high-accuracy forecasts of solar wind plasma and IMF conditions (ambient and transients) to drive space weather models?

3.3. Magnetosheath and magnetopause issues and challenges

At the L1 point and closer to the dayside magnetopause of the Earth, we enter a realm where much empirical work has been done comparing solar wind dynamic pressure, IMF properties, etc. with the magnetospheric size and configuration (e.g., Shue et al., 1998). As shown in Fig. 4(a), for example, the magnetopause position can be dynamically predicted using upstream solar-wind conditions. The particular case under study occurred on 31 March 2001 when a very high-pressure solar wind compressed the magnetosphere to such an extent that the magnetopause—and even the bow shock—was estimated to have been pushed inside of geostationary Earth orbit (GEO). The figure shows magnetopause (Shue et al., 1998) and bow shock (Farris and Russell, 1994) boundary position estimates for one point in time.

A promising avenue of research now under way is to use actual solar wind measurements to drive global MHD models of the magnetosphere. An example is shown in Fig. 4(b) where the Lyon–Fedder–Mobarry (LFM) 3D MHD model is run for a period on 6 April 2000 using solar wind measurements at L1 (results courtesy of C. Huang). The figure plots the log of plasma density and shows that around 21:47 UT on this day the magnetopause was so compressed that the GOES 12 spacecraft (shown as the white ball with the arrow) was well outside the magnetosphere and in the dayside magnetosheath. The arrow shows the instantaneous magnetic field direction measured by GOES. The field was nearly due south, which is greatly different than the usual stable, northward magnetospheric field that GOES operators use to orient the spacecraft. Clearly, global modeling and visualization could help spacecraft operators understand when and why their satellites are not operating as expected.

As this example shows, given upstream solar wind and IMF information (either from measurements or from models), MHD models of the magnetosphere could provide important space weather products and information. There are, however, issues and questions remaining:

- Can we successfully use solar and/or interplanetary modeling (as described in the last section) to forecast far in advance the magnetopause standoff locations?

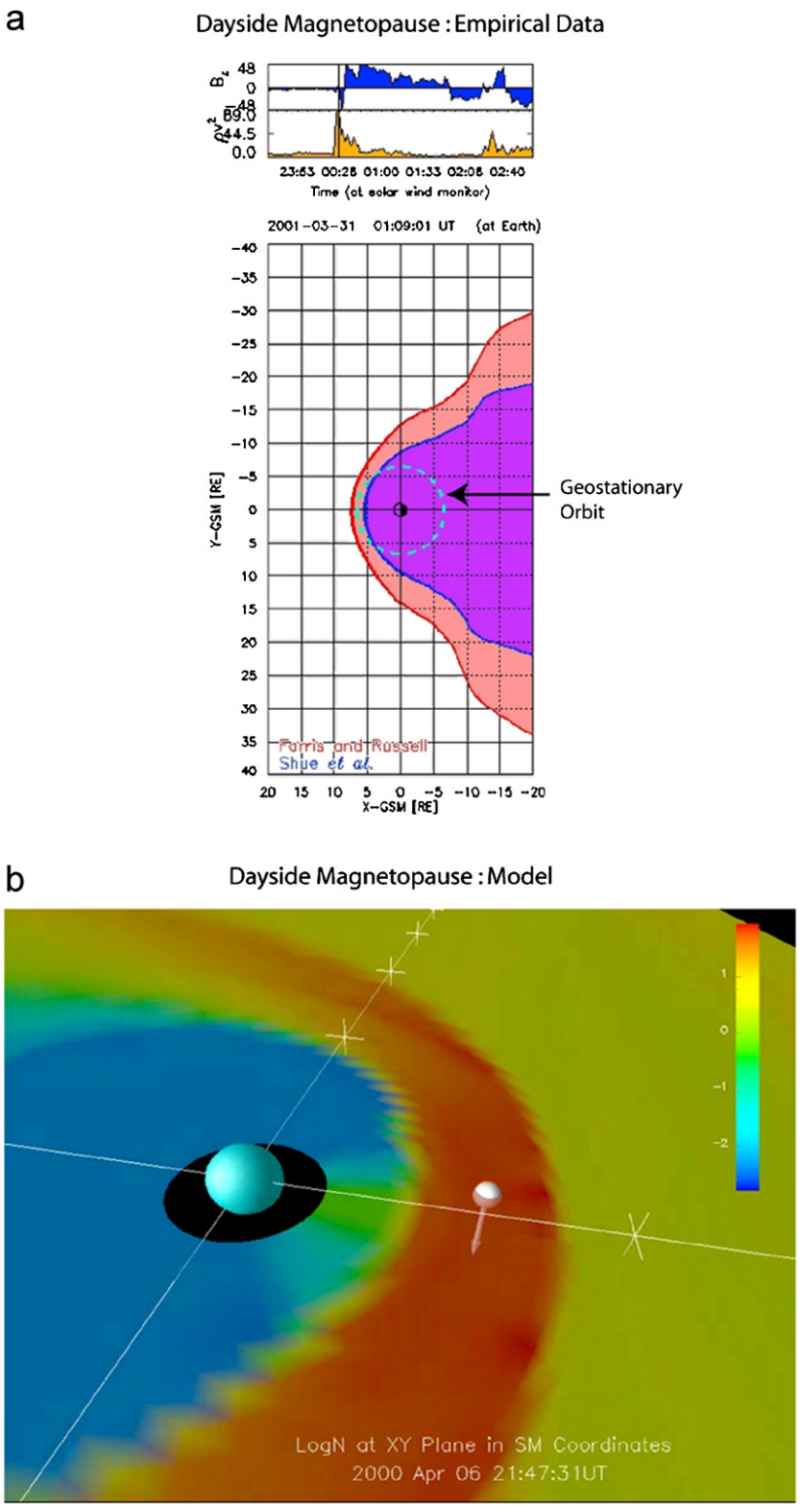

Fig. 4. (a) The estimated bow shock and magnetopause locations for a strong storm period on 31 March 2001 (as described in the text), and (b) a simulation using the LFM code showing a highly compressed magnetopause location during an event on 6 April 2000 (courtesy C. Huang).

- Can we predict accurately the dayside reconnection efficiency and energy coupling strength for typical and/or extreme space weather conditions?
- Can we use present models to forecast solar energetic particle entry into the magnetosphere?
- Can global MHD model provide an adequate description of energy transport from the dayside to the nightside magnetosphere?

3.4. Magnetotail and plasma sheet issues and challenges

A well-known and widely accepted view of magnetospheric dynamics is that the solar wind imparts energy to the magnetosphere by dayside reconnection that is subsequently stored as energy in the magnetotail lobes. Thereafter, the magnetosphere dissipates this stored tail energy via explosive reconnection in the near- or mid-tail plasma sheet that initiates the expansive phase of magnetospheric substorms (e.g., Baker et al., 1996). Substorms produce copious quantities of energetic particles that can be a space weather threat in their own right. Substorms also enhance greatly the large-scale magnetospheric current systems and drive increased magnetosphere–ionosphere coupling. Thus, describing where, when, and how substorms occur is a key requirement for a full understanding of the coupled Sun–Earth system.

Observationally, we have many scientific and operational spacecraft that provide information about substorm processes. For example, Fig. 5(a) shows data from one of the CLUSTER spacecraft for a very well-studied substorm event that occurred on 27 August 2001 (see Baker et al., 2002, 2005a). The upper panel of Fig. 5(a) shows magnetic pressure for the period 00:00–07:00 UT. The middle panel shows plasma bulk flow speeds in the X (or sunward–tailward) sense. The lower panel shows the total magnetic field and B_Z component measured by CLUSTER 1 for the same period (Baker et al., 2005a). We see in these data the classic signature of energy storage (02:00 UT to ~04:00 UT) followed by rapid energy release and dissipation (~04:00–04:30 UT).

As described in the last section, the global configuration of the magnetosphere can be simulated using the LFM model (Lyon et al., 2004 and references therein). The LFM model solves the ideal MHD equations on a nonuniform grid containing the magnetosphere, magnetosheath, and surrounding solar wind, that ranges from 30 Earth radii (RE) upstream to 300 RE downstream, and has lateral and vertical dimensions of 100 RE. The outer boundary conditions applied to the front and sides of the cylindrical domain are the flow conditions of the solar wind plasma. This includes density, temperature and velocity, and the interplanetary magnetic field vector. These are usually taken to be constant across planar fronts flowing from upstream into the model grid, i.e., in the y–z plane in geocentric solar magnetic (GSM) coordinates. In situ measurements such as from the ACE or other upstream spacecraft can provide time-dependent measurements of the driving conditions upstream of the magnetosphere. They are often used as model boundary conditions. Alternatively, solar wind conditions can be obtained from idealized conditions, empirical models, or heliospheric models (see Luhmann et al., 2004). Fig. 5(b) shows a snapshot of electric field strengths (color coding) and plasma flow vectors (arrows) for an LFM simulation of a substorm event on 10 December 1996. Such simulations show the meso-scale features (as alluded to above) concerning substorm onsets and mid-tail reconnection. However, the model also reveals an immense amount of small-scale "channeling" of plasma flow into the inner magnetosphere. This is a topic of current research to compare numerical and observational results (Wiltberger et al., 2000).

Present work with multi-spacecraft observations and global MHD modeling is very exciting. Such work, however, raises questions from a space weather standpoint. Among these questions are:

- Can we forecast with requisite accuracy where, when, and how substorms occur?
- Can we predict the degree of plasma acceleration and the strength of particle transport that will occur for a given solar wind driving condition?
- Will we be able to determine with sufficient lead time where and when important substorm particle injections will occur?
- Can we specify the plasma sheet "source" population that will ultimately constitute the ring current and radiation belts?

3.5. Inner-magnetosphere challenges and issues

The inner part of the magnetosphere is in many ways among the most interesting and most challenging from a space weather and modeling standpoint. It encompasses a radius that most operational and commercial spacecraft orbit the Earth and, thus, it

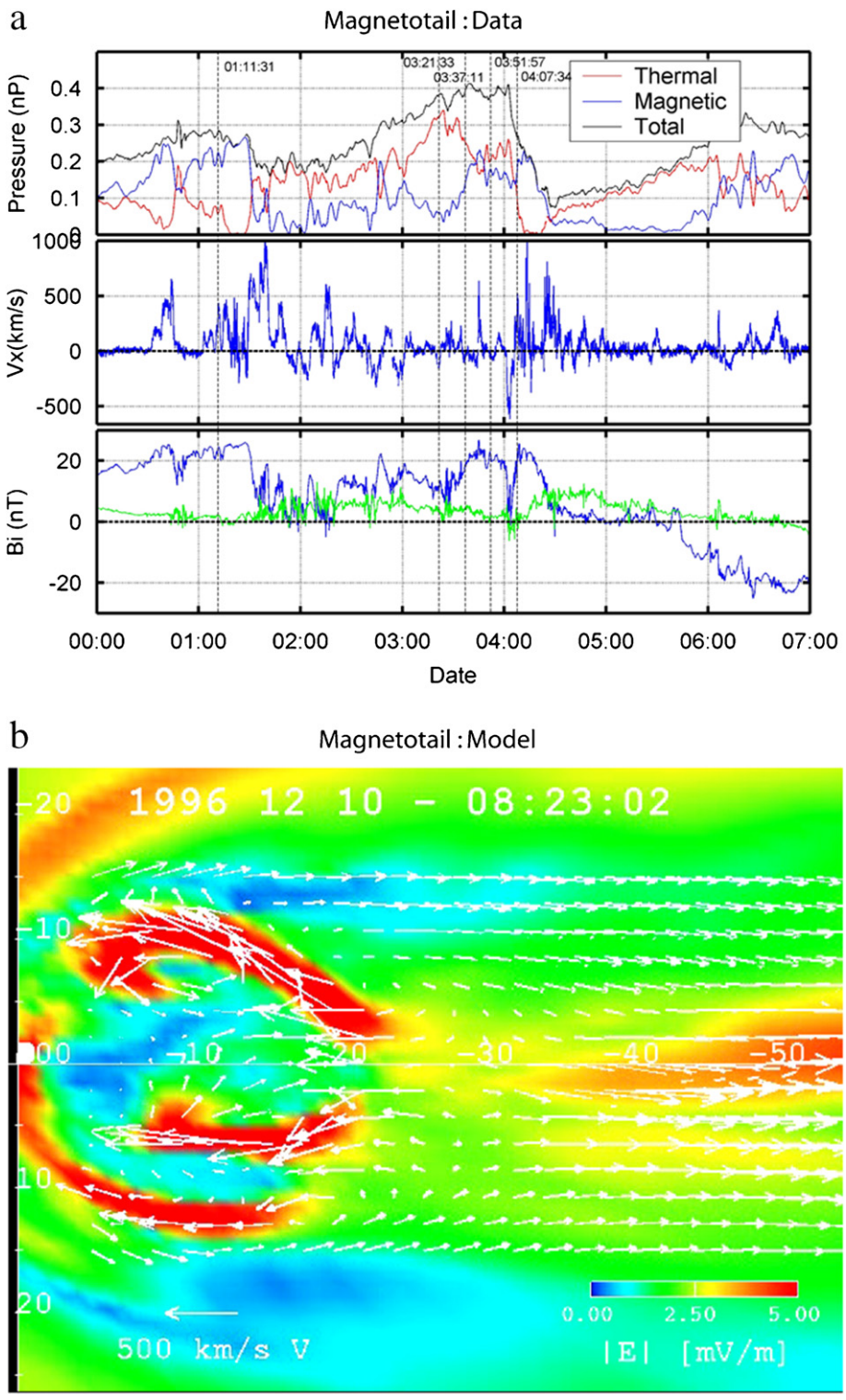

Fig. 5. (a) Data from the Cluster 1 spacecraft showing substorm energy storage and release in the magnetotail on 27 August 2001 (from Baker et al., 2005a), and (b) a model simulation of plasma flows and electric field strengths for a substorm event on 10 December 1996 (from Wiltberger et al., 2000).

is the space weather in this region that affects low-, medium-, and geostationary-altitude satellites (LEO, MEO, and GEO).

As shown schematically by Fig. 6, the inner magnetosphere is the domain where several key plasma constituents abut one another and where often the plasma populations commingle heavily. Fig. 6(a) shows a schematic diagram (courtesy of J. Goldstein) of the cold plasma distribution in the plasmasphere (see, e.g., Sandel et al., 2003). Fig. 6(b) shows a highly schematic portrayal of the medium energy particles (electrons and ions of tens to hundreds of keV) that constitute the terrestrial ring current. Typically, the ring current population lies outward of the main plasmasphere in geocentric distance. Finally, at highest energies are the radiation belt particles that may extend to many MeV in energy and may occupy much of the same spatial domain as the ring current particles. The outer radiation belt electron population is strongly controlled by wave–particle interactions that are closely related to the plasmaspheric plasmas (see Baker et al. 2004b). The relationship of the plasmasphere and inner edge of the electron radiation belt is portrayed schematically in Fig. 6(c).

Scientific spacecraft such as SAMPEX and POLAR, as well as several operational spacecraft such as GOES, GPS, and Los Alamos-instrumented satellites provide a wealth of data about the radiation belt (and ring current) populations. Such data are invaluable for helping to model the development of geomagnetic storms or track the enhancements of the outer radiation belt. A major goal of a program like CISM is to provide an accurate specification and several day forecast of the radiation belt and high-energy ion populations in near-Earth space. We have empirical methods that are working right now (Baker et al. 2004a) to predict radiation belt electrons. The goal is to establish in the future even more successful physics-based approaches (Elkington et al., 2004).

Present physical modeling suggests that low- to moderate-energy electrons can be transported into the inner magnetosphere by convection and then can be acted upon by self-consistent electric fields and ULF waves present there. This has been examined by Elkington et al. (2004). The modeling begins by considering a particularly well-observed magnetic storm interval. There was a large CME-generated magnetic storm that developed very abruptly on 31 March 2001 following the passage of a strong interplanetary shock wave. The ring

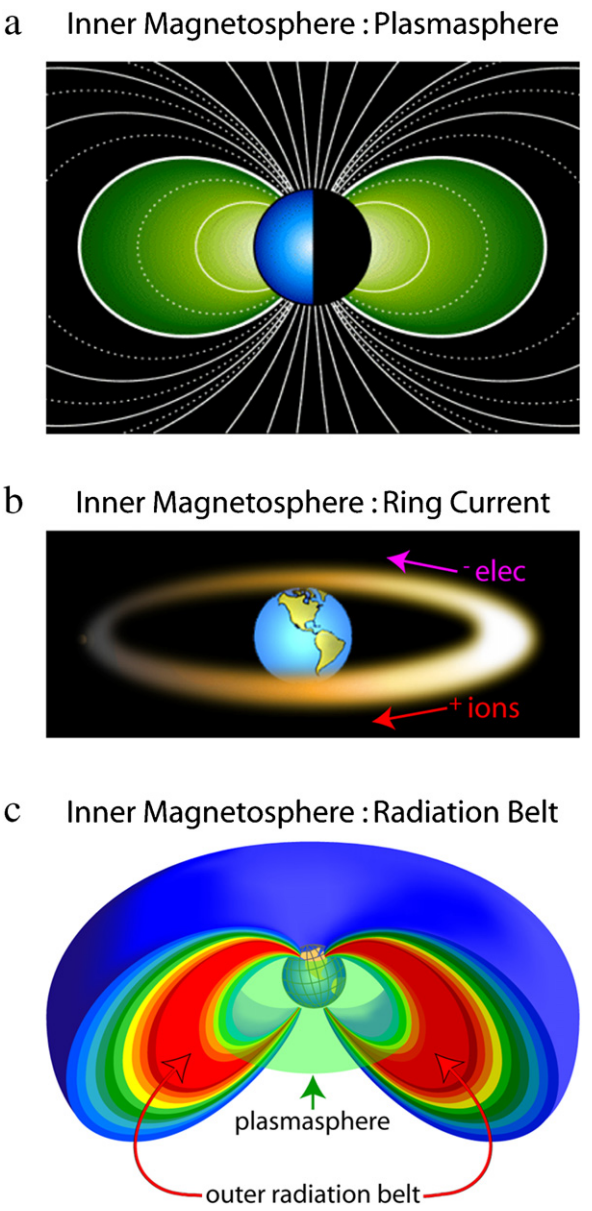

Fig. 6. (a) A schematic of the plasmasphere region in the inner magnetosphere (courtesy of J. Goldstein), (b) an illustration of the terrestrial ring current, and (c) a diagram of the outer Van Allen radiation belt and its relation to the plasmasphere (from Baker et al., 2004b).

current index, D_{st}, dropped strongly to ~ -360 nT, indicating a major geomagnetic storm. It reached a peak at ~09:00 UT on 31 March. A series of powerful magnetospheric substorms followed the shock arrival and these injected large fluxes of medium-energy electrons. Some hours later, the relativistic electron population in the range $L = 3$–6

began to build up greatly. This interval of time has been simulated using the LFM magnetohydrodynamic code. The simulation used the observed upstream solar wind conditions to drive the LFM code. Many of the key features of the observed storm development have been replicated by this numerical simulation (Baker et al., 2005b).

Elkington et al. (2004) used the MHD-simulated magnetospheric results for 31 March. Within the self-consistent framework of the magnetic, electric, and low-frequency wave fields in the LFM simulation, energetic test particles were "pushed" to simulate the energetic electron acceleration. For the simulation, the process began with $E = 60\,\mathrm{keV}$ electrons that were started at $20\,R_E$ geocentric distance. Pulses of particles were then launched at 15-s intervals. It was assumed that the first adiabatic invariant of the electrons was conserved. The system then evolves naturally under the MHD E and B fields. Of key importance is the fact that the LFM numerical simulation produces realistic ULF wave fields in the inner magnetosphere. This plays a critical role in the observed acceleration of the electrons.

Fig. 7 (taken from Baker et al., 2005b) shows three snapshots at selected times from the Elkington et al. simulation. This is an X–Y plane cut through the simulation extending from $X = +7R_E$ to $X = -20R_E$. Contours of constant B in the LFM simulation are portrayed. The time associated with each snapshot is shown in the upper corner of each panel. As noted, the electrons were launched from a broad range of local times at $X = -20\ R_E$ with initial energies of 60 keV. The energy of the particles is color-coded throughout the simulation as shown by the color bar in the bottom panel. Electrons that encounter the magnetopause boundary were removed from the simulation.

The simulation results show that a significant fraction of the tail-launched electrons are, in fact, trapped on closed drift paths and are subsequently pumped up in energy to of order 1 MeV. This energized population ends up forming a dense ring of relativistic electrons (Fig. 7(c)) around $L{\sim}4.0$ in the modeled system. This is closely consistent with the observations for this particular event (Baker et al. 2005b and references therein).

Because of computational limitations, MHD models presently do not simulate accurately the innermost part of the magnetosphere. Moreover, multifluid dynamics are needed to treat ring current and other issues in the inner region. At present, there is considerable effort being expended to couple, for example, the "Rice Convection Model" and the LFM MHD model (Lyon et al., 2004).

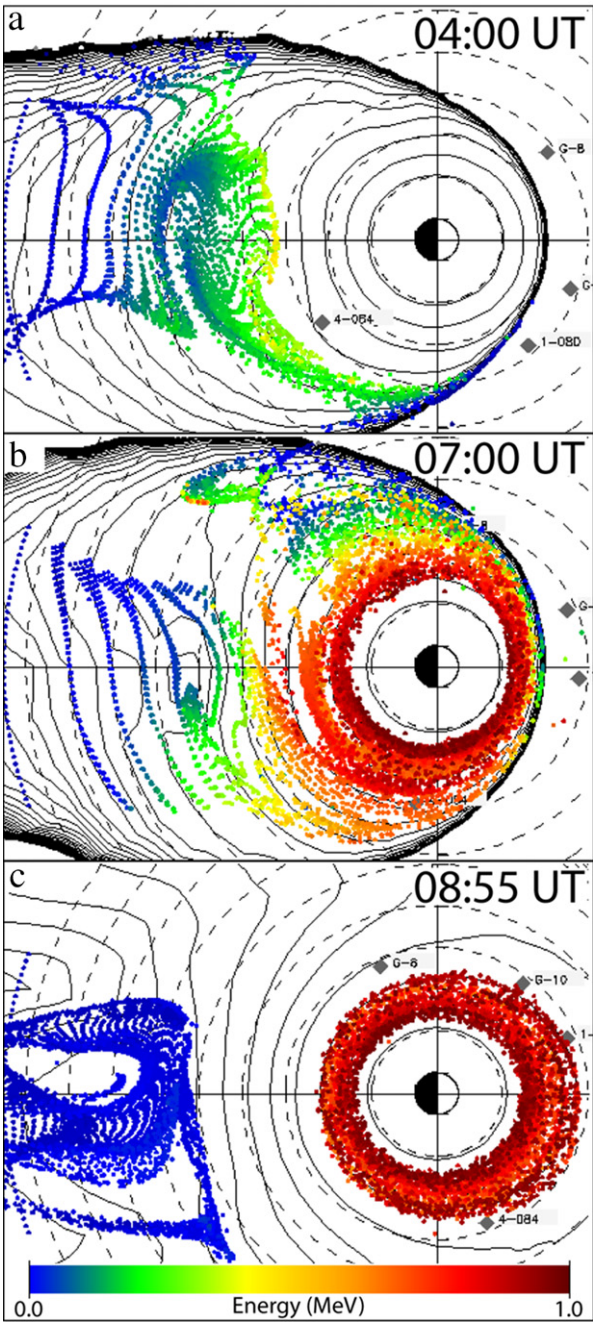

Fig. 7. Snapshots of simulation results for 31 March 2001 showing energetic particle pushing within the Lyon–Fedder–Mobarry MHD code (as described in Elkington et al., 2004). Three times are depicted: (a) 04:00 UT, (b) 07:00 UT, and (c) 08:65 UT (from Baker et al., 2005b).

Given these bodies of work yet to be done, we see several significant issues and questions:

- Can we predict radiation belt configuration and intensity changes from physical first principles?
- Do we have sufficient understanding of wave–particle interactions to describe particle acceleration and loss processes in the inner magnetosphere?
- Can we forecast accurately the occurrence of 'quiet' magnetospheric intervals, or by contrast, major storm intervals?
- Can we successfully couple global magnetospheric models with codes and modules that describe detailed (kinetic) effects in the inner magnetosphere?

3.6. Magnetosphere–ionosphere coupling challenges

Recent spacecraft programs such as the IMAGE mission (Burch et al., 2001) have returned global remote sensing data about the timing, extent, and intensity of auroral luminosity and energy input into the upper atmosphere from the magnetosphere. An example of such information is shown in Fig. 8(a) which is taken from an imaging sequence from the Far Ultraviolet (FUV) experiment onboard IMAGE during the famous Bastille Day event (15 July 2000). During the height of this Great Storm, the auroral region was greatly expanded equatorward and the energy input to the ionosphere was very large.

A challenge for modeling is to couple the global MHD simulations of the magnetosphere to realistic models of the ionosphere and upper atmosphere. The latest in an evolving sequence of ionosphere-atmosphere models is the so-called thermosphere-ionosphere nested grid (TING) model (see Wang et al., 2004). The TING model solves the time-dependent momentum, energy, and continuity equations in a 3D geographic domain. The pressure-level grid extends from an altitude of about 97 km to over 500 km, depending on solar activity, at a resolution of two levels per scale height. A dynamically-coupled ionosphere is calculated considering that transport is significant for O^+. The other major ion species are assumed to be in local chemical equilibrium. The electron density is assumed to be equal to the sum of the ion densities. The inputs driving the model are solar ultraviolet and X-ray radiation at the top of the atmosphere and the electric field and auroral precipitation

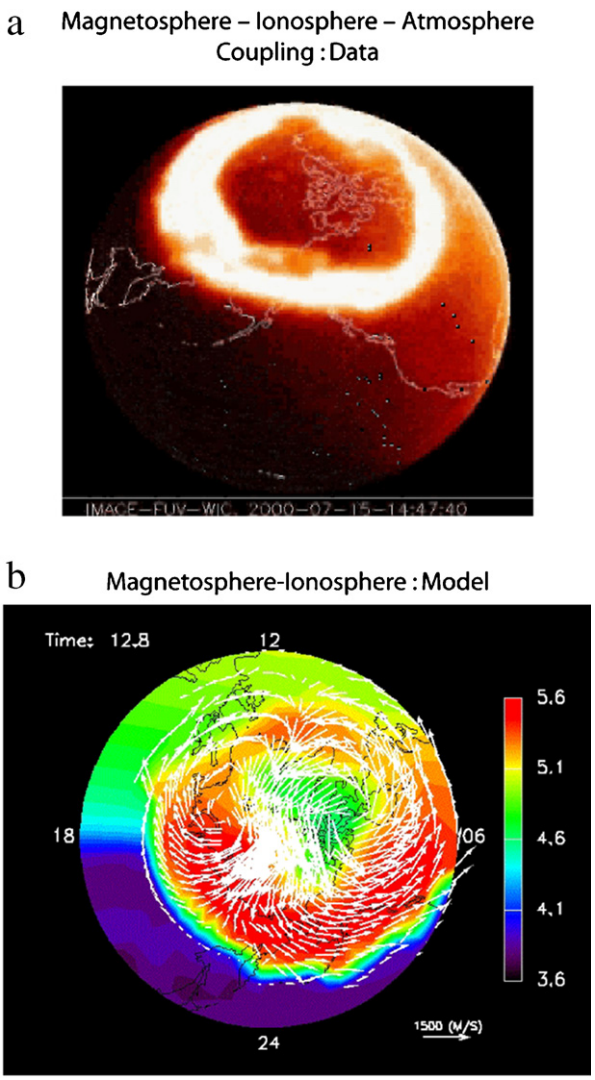

Fig. 8. (a) A snapshot of the auroral development on 15 July 2000 as seen by the FUV experiment on IMAGE, and (b) response of the thermosphere/ionosphere system calculated by the LFM/TING coupled model E-region electron densities ($\log_{10}\mathrm{cm}^{-3}$) at ~120 km with the ion drift pattern superimposed (courtesy of W. Wang).

pattern imposed by the magnetosphere on the upper boundary in the polar regions. Also, the pressure changes due to atmospheric tides are applied to the lower boundary (see Wang et al., 2004).

Two-way coupling between the LFM and TING models is accomplished using a method similar to the procedure employed by the LFM model in coupling to its simplified ionosphere, described by Wiltberger et al. (2004). Solar wind and IMF measurements serve as the LFM external boundary conditions at a 1-min cadence and the MHD

equations are solved using a variable time step. Once every 120 s, auroral electron number flux and characteristic energy are calculated from the field-aligned currents at the inner boundary, using the empirical relationships between currents and particle flux parameters that are employed in the stand alone LFM model. These auroral electron fluxes are interpolated onto a rotating geocentric equatorial grid and are applied to the upper boundary of the TING model. TING then performs a single time step, calculating the conductivity at every grid point and at every pressure level using the ion and neutral densities obtained, and vertically integrating to produce the conductance pattern. The LFM model is then run for another ~480 time steps using that conductance distribution as its inner boundary before another TING calculation. Fig. 8(b) shows a snapshot from an animation of the TING model results during a strong substorm event (courtesy W. Wang). The electron number densities (color coding) and plasma flow patterns (arrows) are shown. Such modeling results comport well with actual substorm observations.

There has been substantial progress in bringing together the magnetosphere and ionosphere physics-based models. However, a number of issues remain:

- How do we better estimate ionospheric "feedback" to the magnetosphere both spatially and temporally?
- Given initial coupled-code successes, can we now gain insight into such key issues as cross polar cap saturation and storm-time potentials?
- Can we integrate present successful data assimilation methods (e.g., Schunk et al., 2004) into magnetosphere–ionosphere physical models?

4. Forecast model development

Siscoe et al. (2004) have noted that a set of linked, physics-based numerical codes capable of predicting environmental conditions at any place and time throughout the solar–terrestrial system needs metrics to tell how well it is performing. To provide metrics, CISM compares the prediction accuracy of the coupled physics-based codes against the corresponding accuracy of data-based models. One such metric that can be generated is skill score: skill score $= (1-\mathrm{MSE}/\mathrm{MSE}_{\mathrm{ref}}) \times 100$, where MSE is the mean-square error in the predictions of the physics-based codes and $\mathrm{MSE}_{\mathrm{ref}}$ is the mean-square error in the predictions of the data-based algorithms. If the MSE of the physics-based codes is bigger than the MSE of the data-based algorithms the skill is negative, otherwise it is positive.

One role of empirical models is to provide baseline metrics for evaluating the performance of the physics-based models of CISM (Spence et al., 2004). The most important property of a baseline model is that it does not change, otherwise it could not serve as a baseline against which to measure the improvement of other models (Siscoe et al., 2004). As discussed by Siscoe et al., a baseline model need not be the best among models available even at the time of its adoption.

To illustrate the strategy, Fig. 9 depicts schematically a line of physics-based, numerical models labeled Forward Models and a line labeled Inverse Models. This simplified, two-line depiction serves to illustrate the point by focusing on the linked models in the chain. In the initial forecast Models, the semi-empirical WSA model can provide the Sun-to-L1 or Sun-to-magnetosphere values of solar wind and IMF parameters that drive a suite of models that forecast operational magnetospheric parameters. At some point in the development of the Forward Models, the MAS-ENLIL code should begin to provide values of solar wind and IMF parameters that surpass the skill of forecast models that use them over that obtained with the baseline WSA. Then the MAS–ENLIL code should replace the baseline WSA model as the provider of solar wind and IMF parameters in the suite of forecast models. Thus, the solar wind–IMF part of the forecast models will have reached the initial stage of the progressive increase in skill relative to a baseline model.

The concept presented here forms some of the baseline testing of models that will be a part of the first-generation CISM forecast model. Fig. 9 demonstrates the relationship between the CISM numerical models (e.g., MHD), the CISM inverse (e.g., filters or data-dived models, empirical, or semi-empirical models), and the CISM forecast model. The performance of the first-generation forecast model will be used as a benchmark. The CISM forward models that are able to predict a quantity of forecasting interest, or are able to exceed the performances of an existing model in the forecast chain, will be integrated into the forecast model. It is anticipated that the forward and inverse models will work in a competitive and complementary way in forming the forecast model. In some cases a model that uses one of these two approaches

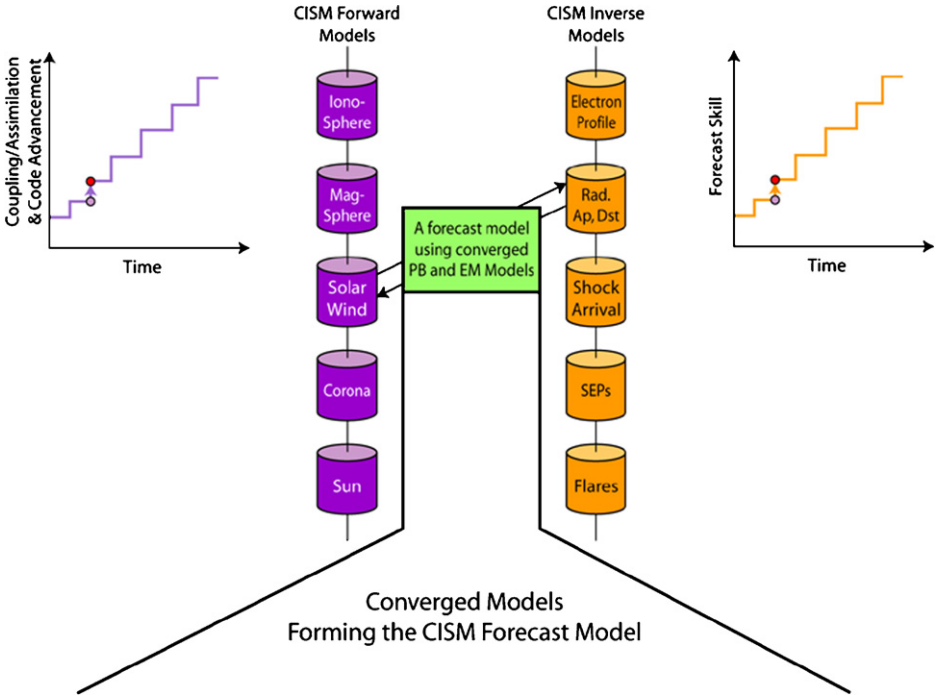

Fig. 9. The coupled forward models (left) and the coupled inverse models (right) being developed within CISM. These two approaches support and inform one another leading to continuous improvement. These models converge in order to form the CISM Forecast Model (FM) which builds upon elements of both the forward and inverse methods (from Baker et al., 2004a).

will replace another. In other cases one of the approaches will fill in a gap where the other is deficient (Baker et al., 2004a).

5. Summary and future directions

This paper has described data sets and models (both empirical and physics-based) that are envisioned to provide the elements for end-to-end solar–terrestrial forecasts. Obviously, modeling directions may change and improve with time. Hence, it is possible (in fact, likely) that the details of the model suite may be modified in future years. Similarly, it is quite likely that new types of data may come on line and data analysis methods will also probably improve and change in ways that only can be guessed at.

In the longer term, we hope and expect that a community modeling effort built on the CISM effort, the Community Coordinated Modeling Center (CCMC), and various Multidisciplinary Research Program of the University Research Initiative (MURI) programs will be established. All the science and operational agencies could and should pool their resources and support such a vision.

In the relatively near term, the direction and content of modeling approaches is relatively clear. A crucially important part of the National Space Weather Program is to identify the greatest needs of the user community and assure that the research community is addressing those needs. By focusing on high priority requirements and working closely with civilian (NOAA) and military (DOD) user entities, it is quite likely that research ideas can be rapidly and efficiently transitioned to an operational environment. It is to be hoped that many scientists doing basic research can participate effectively in this challenging endeavor.

Acknowledgments

The authors thank their many colleagues in the research community for data and graphics used in this presentation. Special thanks are extended to colleagues from the Center for Integrated Space Weather Modeling (CISM). This work was supported by CISM and by other grants from NSF and NASA.

References

Arge, C.N., Luhmann, J.G., Odstrcil, D., Schrijver, C.J., Li, Y., 2004. Stream structure and coronal sources of the solar wind during the May 12th, 1997 CME. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1295–1310.

Baker, D.N., Pulkkinen, T.I., Angelopoulos, V., Baumjohann, W., McPherron, R.L., 1996. The neutral line model of substorms: past results and present view. Journal of Geophysical Research 101, 12,995–13,010.

Baker, D.N., Peterson, W.K., Eriksson, S., Li, X., Blake, J.B., Burch, J.L., Daly, P.W., Dunlop, M.W., Korth, A., Donovan, E., Friedel, R., Fritz, T.A., Frey, H.U., Mende, S.B., Roeder, J., Singer, H.J., 2002. Timing of magnetic reconnection initiation during a global magnetospheric substorm onset. Geophysical Research Letters 29 (24), 2190.

Baker, D.N., Weigel, R.S., Rigler, E.J., McPherron, R.L., Vassiliadis, D., Arge, C.N., Siscoe, G.L., Spence, H.E., 2004a. Sun-to-magnetosphere modeling: CISM forecast model development using linked empirical models. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1491–1497.

Baker, D.N., Kanekal, S.G., Li, X., Monk, S.P., Goldstein, J, Burch, J.L., 2004b. An extreme distortion of the Van Allen belt arising from the 'Hallowe'en' solar storm in 2003. Nature, doi:10.1038/nature03116.

Baker, D.N., McPherron, R.L., Dunlop, M.W., 2005a. Cluster observations of magnetospheric substorm behavior in the near- and mid-tail region, COSPAR 02-A-00234; D3.2-C1.4-0032-02, in press.

Baker, D.N., Elkington, S., Li, X., Wiltberger, M.J., 2005b. Particle acceleration in the inner magnetosphere. in: Physics and Modeling of the Inner Magnetosphere, AGU Monographs, in press.

Burch, J.L., et al., 2001. Views of the Earth's magnetosphere with the IMAGE satellite, Science 291, 619–624.

Elkington, S.R., Wiltberger, M., Chan, A.A., Baker, D.N., 2004. Physical models of the geospace radiation environment. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1371–1388.

Farris, M.H., Russell, C.T., 1994. Determining the standoff distance of the bow shock: Mach number dependence and use of models. Journal of Geophysical Research 99, 17681–17689.

Hughes, W.J., Hudson, M.K., 2004. Editorial. Journal of Atmosphere and Solar–Terrestrial. Physics 66, 1241–1242.

Luhmann, J.G., Solomon, S.C., Linker, J.A., Lyon, J.G., Mikic, Z., Odstrcil, D., Wang, W., Wiltberger, M., 2004. Coupled model simulation of a Sun-to-Earth space weather event. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1243–1256 doi:10.1016j.jastp.2004.04.005.

Lyon, J.G., Fedder, J.A., Mobarry, C.M., 2004. The Lyon–Fedder–Mobarry (LFM) global MHD magnetospheric simulation code. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1333–1350.

National Space Weather Program Council, 2000. National Space Weather Program Implementation Plan, second ed. Office of the Federal Coordinator for Meteorology document FCM-P31-2000.

Odstrcil, D., Linker, J.A., Lionello, R., Mikic, Z., Riley, P., Pizzo, V.J., Luhmann, J.G., 2002. Merging of coronal and heliospheric two-dimensional MHD models. Journal of Geophysical Research 107 doi:2002JA009334.

Odstrcil, D., Pizzo, V.J., Linker, J.A., Riley, P., Lionello, R., Mikic, Z., 2004. Initial coupling of coronal and heliospheric numerical magnetohydrodynamic codes. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1311–1320.

Sandel, B.R., Goldstein, J., Gallagher, D.L., Spasojevic, M., 2003. Extreme ultraviolet imager observations of the structure and dynamics of the plasmasphere. Space Science Review 109, 25–46.

Schunk, R.W., Scherliess, L., Sojka, J.J., Thompson, D.C., 2004. Global assimilation of ionospheric measurements (GAIM). Radio Science 39, RS1S02.

Shue, J.-H., Chao, J.K., Fu, H.C., Russell, C.T., Song, P., Khurana, K.K., Singer, H.J., 1998. A new functional form to study the solar wind control of the magnetopause size and shape. Journal of Geophysical Research 103, 17691.

Siscoe, G.L., Baker, D.N., Weigel, R.W., Hughes, W.J., Spence, H.E., 2004. Roles of empirical modeling within CISM. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1481–1489.

Spence, H., Baker, D.N., Guild, T., Huang, C.-L., Siscoe, G., Weigel, R.S., 2004. CISM metrics plan and initial validation results. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1499–1507.

Vassiliadis, D., Klimas, A.J., Kanekal, S.G., Baker, D.N., Weigel, R.S., 2002. Long-term average, solar-cycle, and seasonal response of magnetospheric energetic electrons to the solar wind speed. Journal of Geophysical Research 107 (A11), 1383.

Wang, Y.-M., Sheeley Jr., N.R., 1992. On potential field models of the solar corona. Astrophysical Journal 392, 310.

Wang, W., Wiltberger, M., Burns, A.G., Solomon, S.C., Killeen, T.L., Maruyama, N., Lyon, J.G., 2004. Initial results from the coupled magnetosphere–ionosphere–thermosphere model: thermosphere–ionosphere responses. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1425–1441.

Weigel, R.S., Klimas, A.J., Vassiliadis, D., 2003. Solar wind coupling to and predictability of ground magnetic fields and their time derivatives. Journal of Geophysical Research 108 (A7).

Wiltberger, M.J., Pulkkinen, T.I., Lyon, J.G., Goodrich, C.C., 2000. MHD simulation of the December 10, 1996 substorm. Journal of Geophysical Research 106, 27649–27663.

Wiltberger, M.J., Wang, W., Burns, A.G., Solomon, S.C., Lyon, J.G., Goodrich, C.C., 2004. Initial results from the coupled magnetosphere ionosphere thermosphere model: magnetospheric and ionospheric responses. Journal of Atmosphere and Solar–Terrestrial Physics 66, 1411–1423.