CISM Data Assimilation Plan

September, 2005

Executive Summary

CISM is developing and exploiting a variety of techniques for ingesting observational data in its numerical models. These include data assimilation as well as other techniques that are needed for the modeling regimes and measurement availability in the space weather system. Our goal is to *identify and apply those techniques for using measured data that most effectively advance the models' characterizations of the system*. These are categorized as: (1) areas where model performance will most directly benefit by incorporation of measurement data; (2) areas where the impact of incorporating measurement data is less compelling, but which are worth exploring; (3) areas where measurements are not currently available, but possible future observations have significant potential.

Category 1:

- Synoptic analysis of solar magnetograms
- Velocity and orientation of coronal mass ejections
- Radiation belt specification and prediction
- Solar wind measurements driving coupled geospace models
- Ionospheric specification model for initialization of coupled geospace models
- Atmospheric forcing of thermosphere/ionosphere system
- Solar irradiance measurements driving coupled geospace models

Category 2:

- Adjustment of solar wind model using upstream measurements
- Use of multiple solar wind measurements
- Active region techniques, e.g. LCT, to establish horizontal velocity fields (CME init.)

Category 3:

- STEREO observations of coronal mass ejections
- Incorporation of vector and chromospheric magnetograms to improve CME initiation
- Use of solar wind measurements well off of the sun-earth line (STEREO)

Introduction

Space weather models traverse a range of physical regimes in which plasma and neutral components with very diverse properties interact through many different mechanisms. The causes, scale size, and persistence of spatial-temporal variations from average values, i.e., the weather, vary dramatically across this system, as does the availability and type of measurement data for the physically important parameters.

The *objectives* of space weather modeling are similarly diverse. They include detailed retrospective modeling of events to advance basic scientific understanding, forecasting capabilities of operationally important quantities with lead times of hours or days, "nowcasting" specifications of the environment, and development of a reliable global context for interpreting and understanding detailed, but sparsely sampled in-situ measurements or remote sensing measurements.

Model fidelity is determined not only by the models' inherent capabilities, but by the data that drive and guide them. The constraining factors for measurement data include sample density and frequency, the limited set of available measurement parameters, reliability and timeliness of data availability, and measurement uncertainty. Data may also be directly relatable to model results, e.g. an in-situ density measurement, or it may only be indirectly relatable, e.g. by reconstruction of an image of emissions based on density through further processing of model results. CISM's goal is to identify and apply those techniques for using measured data that most effectively advance the models' characterizations of the system.

CISM is assessing, testing, and applying an array of tools and methods for using measurement data in support of our modeling objectives. These include the application of Kalman filter assimilation in regimes where this technique is effective; developing methods for automatically preprocessing and updating critical driving data in cases where uncertainties and restrictions in those data are a limiting factor in the models' performance; and novel approaches for incorporating observational data in ways that can have a significant impact on model results. It is important to note that CISM's approach to data assimilation is necessarily broader than the conventional meteorological techniques because of the very different nature of critical parts of the CISM modeling problem. For example, although the solar and coronal information that is the driver of the entire chain of the coupled space weather simulation is only obtainable by remote sensing, we nevertheless can use these data to actively adjust the simulation results upstream of the Earth.

As with all aspects of CISM's development, it is important to target our resources where they will have the greatest impact. We outline below CISM's ongoing development for: (1) areas in which model performance will most directly and significantly benefit by additions to, or improvements in, incorporating measurement data; (2) areas in which the impact of incorporating measurement data are less compelling or less clear, but which are worth exploring because of the potential for a significant improvement; (3) areas in which measurements are not currently available, but for which the possible future availability of a particular measurement set has the potential for significantly improving model performance.

Solar Photosphere and Corona

Because solar observations provide the most distant “upstream view” of future conditions, optimal use of these data has a high impact for those objectives in which a predictive lead-time of days to weeks, instead of minutes or hours, is required. Line-of-sight magnetograms from ground- and space-based observatories currently provide the only routinely available measurement to define the inner boundary condition for driving realistic coronal, heliospheric, and sun-to-earth models. As such they are uniquely important to the entire model chain. CISM is applying a number of advanced techniques to improve how these magnetograms are incorporated in the models and to assimilate newly acquired observations into the “synoptic” specification. We are also performing research to establish the viability of using other types of observations to more reliably define the state of this key region.

- synoptic map data interpolation in the poorly observed polar regions of the photosphere
- accounting for differential rotation and meridional circulation by "evolving" synoptic maps
- updating synoptic maps to accommodate the most recent information (synoptic maps by necessity include "old" information from the unobserved solar hemisphere, but can be improved by embedding the latest or high-cadence magnetograms in a Carrington (27-day) map
- data calibration for multiple observatories, providing greater accuracy and more continuous magnetograph coverage
- observations to define Cone Model initialization, using coronagraph data to characterize the injected CME transient in terms of a set of parameters
- observations of active region evolution to launch more realistic CMEs, including local correlation tracking to establish horizontal velocity fields in the active region
- use of vector magnetograms and chromospheric magnetograms to improve CME initiation procedures
- STEREO observations for multi-perspective imaging to improve the cone model parameters and image-based initiation procedures

Interplanetary Solar Wind

At a given location, for example the L1 libration point where the solar wind parameters that drive geospace models are obtained, the state of the supersonic interplanetary solar wind has a very low persistence. That is, unlike the ocean or atmosphere, its future state does not depend on the current state, rather it is depends on what flows from upstream. A perturbation introduced at the L1 point does not affect conditions there minutes later. Similarly, measurements at L1 have little value in predicting important changes that may occur minutes or hours later. Therefore the effective use of data at L1 to improve model performance requires different approaches than are used in regions having more significant persistence.

The CISM Geospace Models, coupled magnetosphere-ionosphere and standalone, can be driven by either upstream models or by L1 measurements. The direct use of L1 data provides the most accurate input to the geospace models with an hour lead-time, as compared to the greater lead-time and greater uncertainty of the upstream models.

In the simulated sun-to-atmosphere coupled system, the observations that drive the entire chain are the solar magnetograms, which provide the innermost boundary conditions at the Sun. It is therefore necessary to invoke a different kind of data assimilation than is used in atmospheric modeling where densely sampled multipoint in-situ and other measurements are available. For example, in this case it may be effective to use the comparison of upstream simulation results with the in-situ plasma and field measurements obtained by one or more in-situ spacecraft to actively adjust the solar boundary conditions to bring the measurements and model into closer agreement. This correction could then feedback into the simulation for the subsequent time steps.

CISM is doing research to assess several techniques that hold promise for applying data assimilation to improve the CISM interplanetary modeling of the upstream parameters that drive the geospace models. These include:

- Using L1 data to adjust solar boundary condition.
  CISM is exploring the possibility of adjusting solar boundary conditions based on L1 data comparisons by 1) investigating the sensitivity of the model results to the different boundary condition attributes (e.g. polar corrections in the synoptic maps) to understand the effects of each adjustable model parameter, and 2) investigating approaches to varying these adjustable parameters based on the data comparisons. These investigations are most reasonably done initially using the empirical models initially. When effective techniques are identified, they can then be applied to the MHD models where appropriate.
- Assimilation of multiple solar wind measurements.
  Incorporating the useful results of a collaboration between Dusan Odstricl and GSFC to assimilate solar wind measurement data from multiple sources along separate streamlines (rather than discrete points) using an optimal interpolation assimilation method. The impact on space weather modeling of the more accurate description of transverse upstream solar wind profiles that this method will provide has not been established, although Lyon et al. have experimented with shock fronts incident on the magnetosphere at an oblique (rather than head-on) angle.
- Measurements separated from earth-sun line.
  Using STEREO measurements, in concert with ACE data and the CISM Solar Wind models, to determine whether using the Earth-trailing spacecraft, and the assumption of corotating structure, improves advance predictions of upstream parameters

Magnetosphere

The magnetosphere includes some components with a relatively high degree of persistence (such as the inner radiation belt), some that are strongly driven, such as the position of the magnetopause, and some that are a mix, such as the extent of the plasmasphere and the ring current intensity. Although magnetospheric measurement data are relatively abundant compared to interplanetary space, the spatial and temporal sampling is sparse compared with many important phenomena for space weather. The application of data assimilation in the magnetosphere must be selected thoughtfully to have a valuable impact. While we discuss in this section the use of magnetospheric data, it's important to note that because of the tightly coupled nature of the magnetosphere and ionosphere, model performance in both regions can be affected by assimilation of data in the other.

Radiation belt specification and prediction

The state of earth's radiation belts is an important space weather objective in which data assimilation can have a large impact. CISM is using an extended Kalman Filter with radiation belt measurements and solar wind data to predict the state of the radiation belts as well as the time-varying linear coefficients of the model. Tests using SAMPEX and ACE data have demonstrated that this approach provides a robust method for combining available data with a dynamic model to provide the best possible estimate of the current, and possibly future, state of the electron radiation belt. (Rigler, et al.). This approach represents a very significant improvement over time-stationary linear prediction filter methods. Our next steps are to establish and test the modifications necessary for an operational specification and forecast model, including the assimilation of real-time data and more sophisticated model structures.

SEP measurements with magnetospheric filtering

A numerical model is being developed which is designed to use calculated magnetospheric electric and magnetic fields from LFM, LFM-RCM and LTR to determine SEP fluxes within the magnetosphere from a given energetic particle distribution in the solar wind. Fluxes from the Berkeley SEP model as well as measured solar wind fluxes can be used as input, while measurements at GOES provide a benchmark at geosynchronous orbit that can be incorporated into calculations of absolute flux levels. The evolution of energetic particle distributions is modeled by following Lorentz, or guiding center, test particle trajectories in the time dependent MHD fields using both forward and reverse trajectory methods. Time reversed trajectories, originating from a measured distribution within the magnetosphere, may be followed into the solar wind to obtain an intermediate or provisional solar wind distribution update. The initial magnetospheric distribution is then weighted by assuming a solar wind source population with an appropriate measured or model energy spectrum.

Ionosphere/Thermosphere

The availability of measurement data and the slightly more "persistent" nature of the ionosphere-thermosphere system make it the region in which classic data assimilation can have the greatest impact, at least on short-term predictions and current state specifications. CISM collaborations are taking advantage of the advanced nature of ionosphere-thermosphere data assimilation efforts in the community.

Ionospheric Specification Using Data Assimilation

CISM has established a collaboration with the GAIM group, specifically with Bob Schunk and Jan Sojka at Utah State and Tim Fuller-Rowell at CU/CIRES. Geonhwa Jee, a 2005 PhD from Utah State, is joining HAO/NCAR as a postdoctoral visiting scientist to further facilitate this collaboration. Initial work, using the GAIM ionospheric specification to initialize TING electron densities, successfully demonstrated the technical aspect in ingesting GAIM data into the CISM model. This work also established that the persistence of assimilated density features was still strong after one hour but had essentially vanished after 6 hours, thus helping to quantify the time frame in which such assimilations will have the greatest effect on forecast modeling. The next series of steps in this collaboration will be to investigate and quantify several key areas, including relaxation times at high, middle, and low latitudes in the E-region, F-region, and topside ionosphere at solar maximum, minimum, solstice, and equinox; use of model global winds in the IFM/GAIM model; normalization of the model to TOPEX/JASON mid-latitude TEC data; and comparison of model results with GAIM results for 3-day periods in different conditions to identify model changes needed for better agreement. The next step in use of assimilated ionospheric fields will be designing a system to enable periodic updates of the electron densities in coupled geospace models.

Atmospheric Specification of Lower Boundary of Ionosphere-Thermosphere

The National Center for Environmental Prediction (NCEP) global analysis fields provide stratospheric pressure fields that are used as inputs to the TIME-GCM model to specify the lower boundary condition (at ~30 km). As the coupled geospace models transition to the TIME-GCM for the ionosphere-thermosphere, this method will be employed in a nowcasting sense to provide the atmospheric forcing from below during solar-driven events. Since the stratosphere evolves slowly on the scale of short-term space weather events this approach will suffice for the present. Future work will investigate initialization of coupled full-atmosphere models using traditional meteorological data assimilation.

Measurements of Solar Ultraviolet Radiation

The solar radiation component of upper-atmosphere and ionosphere forcing must be addressed by the CISM project. Initially, this will use empirical models of solar far-ultraviolet, extreme-ultraviolet, and X-ray irradiance based on past and current measurements. For middle-term (~2 day) forecasts, a simple persistence assumption can be employed. The next step in development will be to use daily measurements (e.g., by the TIMED solar EUV experiment) of the solar irradiance in the model, and following that, techniques developed by W.K. Tobiska to perform a several-day extrapolation of solar proxy indices based on the 27-day periodicity will be utilized. Since the thermosphere-ionosphere system responds to (non-flare) solar irradiance variations with a time constant on the order of a day, this approach should be adequate. For flare events, it will be necessary to employ near-real-time observations, e.g., from the GOES X-ray monitors, as available. Solar EUV data from the EVE experiment on the Solar Dynamics Observatory will also be available in near-real-time, and will be utilized when that spacecraft becomes operational.