

Chapman Conference on Scientific Challenges Pertaining to Space Weather Forecasting Including Extremes Discussion Notes

Friday, February 15, 2019

Scribes: Anthony Mannucci, Xing Meng, Anna Kelbert. Assistance from Delores Knipp

Note: To encourage free communication during and after the meeting, comments are not attributed.

This document reports on two Discussion Topics on Day 4 of the Chapman Conference
Scientific Challenges Pertaining to Space Weather Forecasting Including Extremes

- 1) **Prediction and Models**
- 2) **Possible Meeting Outcomes**

1. *Brief morning discussion: Prediction and Models*

“Convergence”: missing from the meeting were dynamicists who think a great deal about predictability. These dynamicists can be found in the Non-Linear Geophysics AGU section, who attend a meeting in La Jolla run by Bruce Tsurutani every few years.

Regarding the scientific role of numerical models (versus operational): in reconnection physics, my understanding is that simulations are the fundamental way that theoretical knowledge is communicated. An expert should review this idea.

Is the GIC (ground induced currents) problem an example of temporal/spatial scales that make the prediction problem intractable? For example, three orders of magnitude variations in conductivity near Earth’s surface, over few-km spatial scales. This is relevant to the GIC application. Is finding a creative solution to this problem a science issue, an operational issue, or what?

The standard “understanding to prediction” paradigm: a model development task accompanies a scientific investigation. The model works well for that investigation, but actually only partially and for selected events. The use of this model for prediction is now deemed an operational task. There’s lots that’s wrong with this paradigm, and there is a National Academies study that backs this up. “Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification”, National Research Council, DOI 10.17226/13395.

The role of “reduced” models.

In what areas are observations the main driver of increases in predictive accuracy, and in what areas are modeling improvements the answer? There is analogy to weather, where the

community believes that improvement over time has been largely due to improved observations. Do they really know?

In any given scenario, when can one be sure that understanding is the limit to prediction? We cannot just state “we need improved understanding so that we need to predict this better”.

2. Afternoon Discussion: Toward Meeting Outcomes

General Discussion

Desired: a “grand challenge” paper relating to the science of space weather prediction.

Geospace Environment Modeling (GEM, NSF program) meetings: overarching goal is to understand. Compare models and gain insight. Forecast busts. Start with these. Events from the recent era. Compare to higher resolution modeling runs. Role of preconditioning.

Why is this a grand challenge (GC)? GC is where we want to be many years from now. GEM challenges are not “grand”, they are focused steps.

GC would involve a broader effort.

What are observations telling us that predictive models totally miss? Don’t look at success, but focus on failures. Failures are opportunities to learn. Why can’t we publish a paper where the model output is unrealistic?

Need to talk with forecasters on why there is a “bust”. Forecasters issue forecasts, not models per se.

GC needs to have a funding imperative. Can we understand the failure modes of our models? Uncertainties of the models. This issue cuts across multiple disciplines.

Journal editors want “positive results”. One can only publish positive results. We need to evolve past this.

Anna Kelbert suggested a detailed domain knowledge-led discussion on temporal and spatial scales. Another suggestion was to have a facilitated discussion on what can realistically be achieved in forecasting space weather. These are quite related. Earlier, we had a suggestion of bringing groups together that are not generally talking, but that was probably meant to suggest we have already not done that. This may be the user issue.

Start drafting an outline with sections. Work in serial or parallel?

Possible section: what does the “understanding to prediction” paradigm mean?

Are there “transformative results” related to extremes?

Possible Meeting Outcomes (all outcomes citable with doi:)

Title 1: Space weather forecasting: scientific understanding, predictability, uncertainty and observations (refereed) -- too broad?

State of the art, tools. What can we not do? Where we want to be, how we get there. We need to be more specific.

Can we understand failures modes?

Cultural issues within the community? Use real examples. What are some new paths? What don't we understand about SWx forecasting? Forward looking approaches.

Special collection in a refereed journal from this meeting: target a ratio of 75%/25% science vs commentary.

Need to review state of the art to pull this off. See Vourlidis' talk -- way forward with tools and observations.

Leverage terrestrial weather to identify how to move forward. How do we get to the next step and learn from terrestrial weather. Siscoe publications.

Here I think we are missing another way of doing “space weather forecasting”: Empirical models. Machine learning (ML) may be useful. And where one doesn't understand the physics, this will aid scientific understanding. What does it mean when an empirical model works better than a first principles (FP) model?

Empirical versus first principle model is also a false dichotomy. FP models contain empirical inputs.

We need a self-assessment. We start with empirical models. MSIS for example (climatological model of the thermosphere composition). ML reveals insights into what quantities are related. Let's use the best of what's out there. The more accurate our models are, no matter how, will teach us and is societally relevant.

ML models *are* empirical models.

Cultural issues: we are engaging in gross over-generalizations. Can be uncomfortable. We must be careful. Let's distinguish different cases.

ML needs training data. Empirical describes well the nonlinear relationships. Data assimilation (DA) is about informing FP dynamics. Both methods should co-exist.

Predictability topic: Lorenz' work for weather, butterfly effect, chaos. Dynamics in space weather could be different. Dynamical system view.

Practical predictability: limits to predictability. Spatial and temporal scales. Some of this is theoretical. Observational studies can characterize these scales. We can understand statistically, at this scale or other. We need not always predict the specifics. The European Center for Medium Range Weather Forecasting (ECMWF, a numerical weather prediction center) does not run ensembles of high-resolution models. They inform statistical models with the larger scale runs.

Empirical versus FP models. Both can make a prediction. FP described what might happen, empirical what has happened. Space weather is about the fluctuations about the climate. Empirical models come from limited locations and apply to the past. FP models give us the connections in the system. Assimilate data to improve the prediction. Increase ability to act and make decisions, decrease marginal cost of operations.

ML versus empirical models: there is a difference as we generally use these terms. ML is interpolating input to output. MSIS was developed with physical understanding. There is a limitation of using ML with no physical understanding.

Where do FP models come from? They come from empirical studies. Substorms. Theoretical idea was reconnection. Various ways of interpreting using data/theory/modeling will work.

Physical understanding is required to develop empirical models. Understanding comes from theory and observations. Understanding comes from observations, but that's not an empirical model.

Physical understanding allows us to know where to target our observations.

Jeffrey Love has shown -- one data point can affect an empirical model. We need to apply theory also.

Space weather is not only forecasting. Climatology and retrospective events also are relevant.

A series of papers within a collection "The scientific challenge of space weather"

Title 2: Meeting report (brief partially refereed?)

Scientific Challenges Pertaining to Space Weather
Forecasting Including Extremes

Pasadena, CA, USA • 11-15 February 2019



Title 3: Chapman Conference on Scientific Aspects ... (non-refereed, comprehensive meeting record, with questions, minutes, survey results, etc.)

Title 4: Space weather and users: transformative approaches (non-refereed)

Why is the user problem not solved?

Progress is being made at Space Weather Week (SWW in the US) and European SWW. These are not workshops. Missing that element. Needs more focus on detailed technical discussion. One cannot discuss science with a military General.

Expectations of scientists versus industry representatives. Scientists may say: "We want your proprietary data." Find mutually beneficial "deals." Manage expectations about direct interactions between scientists and industry.

European Space Agency (ESA) -- lots of user requirements. Requirements don't really meet the user needs. More science interest, but does not resonate with users. Still work to be done in translating user requirements into something usable. It's a slow process.

Translate from engineering requirements to operational requirements. User perspective is an art and a science. Operators use the same data differently, it's an art.

Why aren't users coming up with requirements? Lack of knowledge. They need to learn from us. E.g. developing the appropriate risk indicators. Engineers and scientists need to talk. It's a slow process.

Some fields -- they don't even think about space. E.g. day/night differences not understood. Explain substorms to a roomful of electrical engineers? Have to always restart the discussion. Users need an education.

There is inertia on the user side. Space weather is a minor issue compared to other issues. Comply with regulations. Overwhelming scientific consensus is needed. Resistance to spend resources.

Users want products. Research is secondary.

Title 5: Space weather policy: what the science informs us (non-refereed).

SPA advocacy committee -- a paper came out on policy aspects.

A user-oriented policy paper?

Lack of scientific consensus on the benchmarks.

Concluding: What are the remaining gaps we have not covered?

Listing of how we approach space weather comprehensively.

What have we not discussed: new observations. What are the most critical observations?

World Meteorological Organization: has a list of space weather observations at the OSCAR web site (<https://www.wmo-sat.info/oscar/>).

Thermosphere density data are lacking -- consistency is lacking etc. ESA has standard data sets. Earlier data were processed inconsistently.

Solar EUV -- absolute calibration is not standard. A major modeling problem.

There is low confidence in information -- the existing information is insufficient for making decisions. We need to be forward looking. 557th Weather Wing (Air Force) space weather products do go out to operational users. These data are used now for retrospective studies. The answers needed for operations are not there. We are looking for future products that don't exist yet, new capabilities and new needs. Demand for precision is increasing. Space is being relied on more, atmospheric effects and other effects need to be better specified. Is vulnerability getting higher due to new technology? In other areas is vulnerability reducing? New surge in satellite constellations -- increasing vulnerability.

Science -> engineering -> operational -- does science community talk to the engineers in satcom, gps operators, ground folks -- no dialog. What are system level impacts?

Impacts: (missing?) are we overestimating the impacts, or the opposite?

Redefining what extreme means. We cannot keep crying wolf. Inadequacy of Dst as a useful metric. What is definition of "extreme"? An event may be extreme in one aspect versus a non-event in another aspect. Storms are not necessarily the organizing factor. Where are the extremes really found? What is the quantity of interest? Discoveries can change things. Prompt penetration electric fields example. This explained why certain smart bombs were off-target.

Visualizing results and communication with end users. Science versus end users. Visualizing metrics that are useful. Focus groups with end users.