The challenges and rewards of running a Geospace Environment Modeling Challenge

H. Hietala[1,2,3], A. P. Dimmock[4,5], Y. Zou[6], K. Garcia-Sage[7]

[1]The Blackett Laboratory, Imperial College London, United Kingdom
[2]Space Research Laboratory, Dept. of Physics and Astronomy, University of Turku, Turku, Finland
[3]Dept. of Earth, Planetary, and Space Sciences, University of California, Los Angeles, USA
[4]Swedish Institute of Space Physics, Uppsala, Sweden
[5]School of Electrical Engineering, Aalto University, Espoo, Finland
[6]Department of Space Science, University of Alabama in Huntsville, USA
[7]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Key Points:

- GEM Challenges bring people together to compare models and observations to advance our understanding of solar wind-magnetosphere interaction
- We recount our experiences as the organizers of the GEM Dayside Kinetics Challenge to aid and inspire future participants of such activities
- We give advice on management and funding, and demonstrate the importance of openly documenting the activities

Corresponding author: Heli Hietala, `h.hietala@imperial.ac.uk`

Abstract

Geospace Environment Modeling (GEM) is a community-driven, National Science Foundation-sponsored research program investigating the physics of the Earth's magnetosphere and its coupling to the solar wind and the atmosphere. This commentary provides an introduction to a Special Issue collating recent studies related to a GEM Challenge on kinetic plasma processes in the dayside magnetosphere during southward interplanetary magnetic field conditions. We also recount our experiences of organising such a collaborative activity, where modelers and observers compare their results, i.e., of the human side of bringing researchers together. We give suggestions on planning, managing, funding, and documenting these activities, which provide valuable opportunities to advance the field.

Plain Language Summary

Geospace Environment Modeling (GEM) is a community-driven, National Science Foundation-sponsored research program investigating the physics of the Earth's magnetosphere and its coupling to the solar wind and the atmosphere. An integral part of the program are so-called "Challenges", which bring people together to compare models and observations in order to advance our understanding of the near-Earth space environment. This commentary provides an introduction to a Special Issue collating recent studies related to one such collaborative effort. We also share our experiences as early-career scientists organising such an activity, to aid those who might take part in such endeavours in the future. We give suggestions on planning, managing, funding, and documenting the activities.

1 Introduction

This Special Issue brings together recent studies related to the NSF Geospace Environment Modeling (GEM) Challenge on dayside kinetic processes during southward interplanetary magnetic field conditions, advancing our understanding of the solar wind-magnetosphere coupling. From the start of GEM in 1989 up until now, several challenges and campaigns have addressed various questions throughout geospace (e.g., Lyons, 1998; Birn et al., 2001; Yu et al., 2019). Over the past 30 years, GEM has grown drastically, new scientists have entered the field, and the field of geospace science as a whole has shifted. Concurrently, technology has advanced and the number of models and their sophistication has increased. Given these changes, it is useful to demonstrate to the space physics community what a GEM Challenge looks like now. We would like the Dayside Kinetics Challenge not only to drive progress in the dayside kinetic processes that were our scientific focus, but also to inspire current and future GEM Challenge efforts. To this aim, we wish to use this opportunity to share our experiences as early-career scientists organising such an activity. We hope that our account is helpful to those who might lead or take part in such endeavours in the future.

What is a "GEM Challenge"? It is an activity where researchers come together to compare different models and observations to gain insight into the workings of both the numerical codes and the magnetospheric phenomena. Typically, one or several time intervals are chosen, e.g., a geomagnetic storm, and the challenge is then for models to match particular observed metrics. Ideally, the participants collect observations of magnetospheric phenomena and their drivers, to be used for validation. They try to quantify agreement/disagreement between data sets and models, and determine reasons for data/model, model/model, and data/data differences. This then leads to further development of both the models and the observatories. Ultimately, the Challenges advance our understanding of various multi-scale plasma processes and their role in solar wind-magnetosphere interaction.

Undoubtedly, the most famous venture was the GEM Reconnection Challenge. Different models of magnetic reconnection were run using the same 2D configuration and specified initial and boundary conditions to find out which physics is required to produce fast reconnection. The paper summarizing the results, Birn et al. (2001), has been cited over 800 times to date, and along with the seven other Reconnection Challenge papers in that issue, the conclusions of the GEM Reconnection Challenge are staple materials of space physics courses.

There have been many Challenges over the years, e.g., Lyons (1998), Birn et al. (2005), Liemohn (2006), Pulkkinen et al. (2011), Rastätter et al. (2011), Rastätter et al. (2013), Tu et al. (2019), and Yu et al. (2019). Their topics have ranged from ionospheric flows to geosynchronous magnetic field to spacecraft surface charging. The Challenges clearly remain an integral part of the GEM activities. For instance, the GEM Focus Groups, five year umbrellas for activities on a given topic (led by three to five co-chairs), are selected through a competitive process. In this procedure, the team proposing a Focus Group is generally asked what kind of Challenges they will be running. Inevitably, the success of the Reconnection Challenge casts a long shadow.

2 The Dayside Kinetics Southward IMF Challenge: a chronology

When we proposed the Focus Group on *Dayside Kinetic Processes in Global Solar Wind–Magnetosphere Interaction* in December 2015, we had the following ambitious plan for our five year term: We would have a series of Challenges. We would start with one event with a simple interplanetary magnetic field (IMF) configuration from the first dayside season of the brand new Magnetospheric Multi-Scale (MMS) mission, launched in spring 2015. We would then diverge into *statistics* and *events* based branches. Four years on and here we are determined to wrap up the first phase of our grand plan before the end of our term.

How did this particular Challenge unfold? In the 2016 Summer Workshop, the first for the newly accepted Focus Group, we held a kick-off session to discuss what the community would want from the first Challenge. Before the workshop, we had solicited suggestions for possible challenge events in our Focus Group's very first announcement on the GEM Messenger newsletter. We had reached out to several simulation groups, who were enthusiastic about the prospect of a Challenge on dayside kinetic processes. We had a vibrant discussion among the 30–40 session participants on the science priorities, specifics, and metrics. Because the state-of-the-art global kinetic models were not yet all able to run in 3D, we discussed the possible merits of three, 30–45 minute runs with different 2D geometries: a southward IMF polar plane run, a northward IMF polar plane run, and an equatorial plane run. We had presentations on both the available models and observations. These included an event from November 18, 2015, 01:50–03:00 UT with southward IMF, MMS-Geotail magnetopause conjunction and SuperDARN radar measurements, introduced by Kitamura et al. (2016). In the fall, we organisers searched for further suitable events (southward IMF and multi-spacecraft conjunction, reducible into a 2D plane if needed), focusing on the ones that were published as part of the MMS first results. We did not find other events that would merit to be put up for a vote. In the 2016 mini-GEM, held the day before the American Geophysical Union Fall Meeting, we put that event forward as the primary challenge event. One of the challenging aspects of the event was the requirement to use a tilted dipole, as it was not routinely included in most global kinetic models at the time. We also introduced a few secondary events with different IMF orientations (as we were still thinking about a set of Challenges based on multiple events).

In spring 2017, we formally announced the Challenge, hosted on the website of NASA's Community Coordinated Modeling Center (CCMC). About a dozen observers had by then agreed to participate, and three modeling groups signed up. In a 2017 Summer Workshop session of about 40 participants, we had seven presentations of the observations made throughout the magnetosphere during the event, as well as two on some very preliminary simulations. Naturally, the different observations seemed rather unconnected at that time, as this was the first time they were brought together. In the 2017 mini-GEM, we had presentations of the first simulation runs made of the event. These seemed, again, rather unconnected with the observations.

In preparation for the 2018 Summer Workshop, we organisers devised so-called *standard plots*: detailed instructions on how to make (line) plots showing MMS observations in the same panels with simulated data. To proceed with the analysis, two participating modeling teams joined forces with observers. We heard each collaboration present their simulation-observation comparisons on five prescribed topics based on past years' discussions using the new standard plot format: magnetic field and plasma signatures, waves, magnetopause location, and X-line dynamics. The teams also had some time to present any additional findings they found interesting. We also started discussions about a joint Special Collection.

For the 2018 mini-GEM, we requested model-model comparisons and analysis on magnetopause transients, and we announced the upcoming Special Collection. In the 2019 Summer Workshop, we had presentations from some further observation-modelling comparisons, in particular of the magnetopause transients, as well as observation-model-model comparisons. By this point, the number of session attendees had decreased to less than 20, and naturally the number of active Challenge participants is a fraction of the number of people attending the sessions. On the other hand, we also attracted some new contributors at this stage. Finally, the submission deadline for this Special Collection was in fall 2019, and eventually extended to January 2020.

3 Challenges and rewards

For us, the Challenge organisers, the biggest struggle has been maintaining focus, interest, continuity, and communication from year to year. The fact that the same researchers don't continuously attend summer workshops and mini-workshops is only one of the issues. While four years is a relatively short time for science, the people doing the science move institutes, switch topics, their funding stops, different funding (maybe) starts, orbits of their favorite satellite missions change, PhD students graduate, etc. The organizers need to tirelessly reiterate the message of the value of Challenge activities.

Most of the efforts of the organizers go into communication, i.e., emails. We quickly learned that a message on a mailing list or an email starting with "Dear all" will not get any traction. First of all, people need to be individually persuaded to commit their time, essentially for free (as there is no funding dedicated to the Challenges), for something that is most likely not directly in their interest as they are presently funded to do something else. Furthermore, they often need to be convinced to do something that may be relatively trivial for them (i.e., not terribly interesting in their a priori opinion), for which the main significance comes from putting it together with what the other participants have done/are doing/will be doing. Once you do succeed in enlisting their support, you need to make sure that everyone carries out the work as agreed, and without changing their focus along the way.

At times, it was simply frustration all around: The organizers were frustrated because despite sending 5 plus emails over a period of months, everything always seems to get done during the two weeks right before the Summer Workshop. The modelers struggled to understand why the observers are unable to determine "even the simplest things" from their data. The observers were frustrated because the modelers are unable to readily pinpoint why their model produces this or that feature.

There were also moments of astonishment, delight, and accomplishment: When someone volunteers their time, saying: "I would like to look at that." When we saw a modeler and an observer beginning to work together. When we're copied in on an email where, against all our skepticism, two "rival" modellers are sharing their data with each other. When we did amass a 90-minute session worth of presentations on the Challenge in a Summer Workshop, and the year after, and the year after that. When we went through our notes from the past few years and realized that the participants have indeed, over time, addressed all the things that the community suggested in past end-of-session discussions.

We, the Dayside Kinetics chairs, got immeasurable help and encouragement from the chairs of the *Modelling, Methods, and Validation* Focus Group. While we were organising a Challenge for the first time, they had a longer experience of them, and also insight from other Challenges run by other Focus Groups at the same time and before us. At times when we felt like we were not making any progress, trying to live up to our greatly idealised picture of a Challenge, they assured us that we were on the right track: It is already valuable that people are comparing models with observations. The Challenge activities are highlighting the various tools that exist. The people are sharing their data and leveraging data from others. The Challenge has prompted the participants to discuss, e.g., the scaling related to kinetic models, and how to do it. Models are improved based on the comparisons whenever a new run with, e.g., a different inner boundary condition is done. (It does not need to be a whole new version of the complex simulation model.) When you are doing a new Challenge, you need to (constantly try to) come up with new ways to make comparisons. That is normal and to be expected. A single number used as a metric for some previous Challenge may not be good for your purposes, especially if your Challenge is interested in 3D structures evolving in time. The path of development-testing-validation and eventual operation is long; while the magnetohydrodynamic (MHD) models are much further on this track than kinetic models, the important thing is to keep going. Thanks to the Challenge efforts, new, interesting features are emerging for future studies.

4 Looking around and ahead

The question, then, naturally arises: How are the other Focus Groups running their Challenges? An example of a successful Challenge activity organized very differently from ours is given by the contemporary *Magnetotail Dipolarization and Its Effects on the Inner Magnetosphere* Focus Group chaired by Christine Gabrielse et al.: In a matter of months in 2018, they run what they called a "Challenge question" on "Can the large-scale dipolarization and/or Substorm Current Wedge be built by an accumulation of many dipolarizing flux bundles?". They began with a panel discussion on the question at the 2018 Summer Workshop. In September, they asked and reached out for contributors for a second stage, leading to a combination of invited participants and volunteer participants. They required the contributors to answer the question above, either by older papers or new analysis, by the end of October. All contributors received a copy of everyone's answers. The participants then had one month to formulate and submit rebuttals, which were forwarded to the appropriate contributors to give them time to consider replies. At the 2018 mini-GEM, the contributors presented their answers and methods. Rebuttals and replies were made. The organizers ordered the talks in such a way that they replied to one another. The result was several papers that are now published or are being worked on. Note, however, that this format required intense management on behalf of the organizers during those months. Were we to start a Challenge activity again, we would definitely begin by considering whether our aims would be better achieved by such a short burst of action or by a more extended program.

Given the amount of support we received from the *Modelling, Methods, and Validation* Focus Group, we are excited that from the beginning of 2020, it has been transformed into a permanent "GEM Resource Group". We hope that this Group will provide continuity, in particular by collating, curating, and distributing the often tacit (silent) knowledge about running Challenge-type activities. By documenting our experiences, we have aspired to contribute to these ends. Based on the information it gathers, the new Resource Group could also develop guidelines on, e.g., the co-authorship of Challenge organizers on papers resulting from Challenge activities.

One of the key problems for Challenge activities, especially for the more extended ones, is the lack of funding. This is understandable considering the numbers: At any given time, there are more than ten GEM Focus Groups in place, presumably most of them running a Challenge of their own, each with several participants. Therefore, it is unthinkable in the present funding situation that multiple teams taking part in a given Challenge would all be successful in obtaining *separate* grants to cover their participation. Yet, by definition, the Challenge activities require multiple participants.

We would recommend the GEM community to consider writing collaborative proposals, one per Challenge. This would naturally result in quite small Full Time Equivalents (FTEs) per person, as there would be many institutes and researchers involved. However, it would also mean that the merits of each Challenge would be evaluated, and if successful, the participants' engagement would be compensated. We believe that such grants would lead to Challenges being more focused, their science questions better formulated, better planning, and more committed participants.

5 The Special Collection

The general purpose of this Special Collection is to document the Challenge event from November 18, 2015, and our understanding of the magnetospheric condition at the time, by gathering the various available data sets and their analyses to a common location. Note, however, that some studies related to the event, namely Kitamura et al. (2016) and Nishimura et al. (2019), have already been published elsewhere. We wish to highlight the current state-of-the-art tools, the capabilities of different methods, as well as their limitations and uncertainties. All manuscripts relevant to the overall topic of dayside kinetic processes during southward IMF, also from those who have not previously participated in the Focus Group activities, have therefore been welcome.

The collection illustrates the effort that goes into analysing the state of the magnetospheric system, even for a short time interval. You are also likely to see that people don't necessarily agree on the interpretation nor converge to a single conclusion. This collection sets the stage for further developments, and aims to enable future benchmarking and validation.

We hope you will enjoy this collection.

Acknowledgments
We acknowledge valuable discussions with our fellow Focus Group co-chairs: Xochitl Blanco-Cano and Gabor Toth (Dayside Kinetics) and Michael Liemohn, Robert Redmon, and Lutz Rastaetter (Modeling Methods and Validation). We thank the National Science Foundation (NSF) for supporting the GEM program and, in particular, the GEM Workshops. During the different stages of the Challenge, HH has been supported by National Aeronautics and Space Administration (NASA) grant NNX17AI45G and NASA contract NAS5-02099, the Turku Collegium of Science and Medicine, and a Royal Society University Research Fellowship. KGS has been supported by the Magnetospheric Multiscale Mission (MMS), NASA Heliophysics Internal Scientist Funding Model (HISFM18-0006), and the Community Coordinated Modeling Center (CCMC).

References

Birn, J., Drake, J. F., Shay, M. A., Rogers, B. N., Denton, R. E., Hesse, M., . . . Pritchett, P. L. (2001). Geospace Environmental Modeling (GEM) Magnetic Reconnection Challenge. *Journal of Geophysical Research: Space Physics*, *106*(A3), 3715-3719. doi: 10.1029/1999JA900449

Birn, J., Galsgaard, K., Hesse, M., Hoshino, M., Huba, J., Lapenta, G., . . . Priest, E. R. (2005). Forced magnetic reconnection. *Geophysical Research Letters*, *32*(6). doi: 10.1029/2004GL022058

Kitamura, N., Hasegawa, H., Saito, Y., Shinohara, I., Yokota, S., Nagai, T., . . . Burch, J. L. (2016). Shift of the magnetopause reconnection line to the winter hemisphere under southward IMF conditions: Geotail and MMS observations. *Geophysical Research Letters*, *43*(11), 5581-5588. doi: 10.1002/2016GL069095

Liemohn, M. W. (2006). Introduction to special section on "Results of the National Science Foundation Geospace Environment Modeling Inner Magnetosphere/Storms Assessment Challenge". *Journal of Geophysical Research: Space Physics*, *111*(A11). doi: 10.1029/2006JA011970

Lyons, L. R. (1998). The Geospace Modeling Program Grand Challenge. *Journal of Geophysical Research: Space Physics*, *103*(A7), 14781-14785. doi: 10.1029/98JA00015

Nishimura, Y., Wang, B., Zou, Y., Donovan, E. F., Angelopoulos, V., Moen, J. I., . . . Nagatsuma, T. (2019). Transient solar wind-magnetosphere-ionosphere interaction associated with foreshock and magnetosheath transients and localized magnetopause reconnection. American Geophysical Union (AGU).

Pulkkinen, A., Kuznetsova, M., Ridley, A., Raeder, J., Vapirev, A., Weimer, D., . . . Chulaki, A. (2011). Geospace Environment Modeling 2008–2009 Challenge: Ground magnetic field perturbations. *Space Weather*, *9*(2). doi: 10.1029/2010SW000600

Rastätter, L., Kuznetsova, M. M., Glocer, A., Welling, D., Meng, X., Raeder, J., . . . Gannon, J. (2013). Geospace Environment Modeling 2008–2009 Challenge: Dst index. *Space Weather*, *11*(4), 187-205. doi: 10.1002/swe.20036

Rastätter, L., Kuznetsova, M. M., Vapirev, A., Ridley, A., Wiltberger, M., Pulkkinen, A., . . . Singer, H. J. (2011). Geospace Environment Modeling 2008–2009 Challenge: Geosynchronous magnetic field. *Space Weather*, *9*(4). doi: 10.1029/2010SW000617

Tu, W., Li, W., Albert, J. M., & Morley, S. K. (2019). Quantitative assessment of radiation belt modeling. *Journal of Geophysical Research: Space Physics*, *124*(2), 898-904. doi: 10.1029/2018JA026414

Yu, Y., Rastätter, L., Jordanova, V. K., Zheng, Y., Engel, M., Fok, M.-C., & Kuznetsova, M. M. (2019). Initial results from the GEM Challenge on the spacecraft surface charging environment. *Space Weather*, *17*(2), 299-312. doi: 10.1029/2018SW002031