Position Paper – Aaron Ridley

Computing Innovations

I personally don't feel like there have been many "computing innovations" that have changed the way that science is conducted in geophysics in the last few years. Our field appears to be playing catch-up with many innovations. I would say that they last innovations that have helped are:

1. Access to more cores on supercomputers has allowed massive simulations of the sun, interplanetary space, the magnetosphere, and upper atmosphere. Groups are starting to recognize the power in massive computing and are moving to codes that are more scalable.
2. More groups are starting to work in the area of data assimilation. Some groups have started with 4DVAR, while others have started to work with ensemble Kalman filters. Still others are working on tomographic reconstruction techniques using GPS (for the ionosphere) and imagery of the sun (for the corona).
3. Algorithms have been developed to allow radar data to be processed in a different way than it was 10-15 years ago. This has changed the method for taking radar data and made radars significantly more powerful.
4. The Community Coordinated Modeling Center has gathered a large number of codes that they make available to the community. You can log into their website and request a run of a model. There are two aspects that are interesting about this: (A) non-modelers can run simulations of the geospace environment, and (B) there are thousands of model runs available that can be mined for information.

Science Challenges

Geospace is defined as the near-Earth space environment, which includes the Sun, the space between the Sun and Earth's atmosphere, and other interplanetary locations. Geospace starts around 100 km and extends outwards.

There are three overarching major difficulties in geospace sciences:

1. It is difficult to take measurements of geospace. Most measurements are done with passive optics or active radar/lidar systems. The optical measurements cannot be taken under cloudy conditions. Some optics do not work on the dayside. Radars and lidars are relatively expensive and so can not be installed too many places on the Earth. The also have limitations on the minimum values that they can measure, so sometimes do not work well. There is often a significant trade-off between signal-to-noise and integration time. The question then is how to address scientific questions with very limited data?
2. Geospace has electric fields that drive strong flows over small scales. This makes it incredibly difficult to accurately model the environment for any

significant length of time without making significant assumptions. Both the time step and the grid size of the models need to be small. While the field is moving towards resolving some of the scale sizes, it is quite complicated to conduct ensemble simulations without massive computer resources.

3. The near-Earth space environment is a strongly driven non-linear system. The main driver is the sun's atmosphere, which takes approximately three days to flow from the sun to the Earth. During these three days, there are no measurements of the state, which evolves. So, it is unclear what will actually be driving the near-Earth space environment at any time. Further, because it is strongly driven, the state of the atmosphere can completely change within a few hours if drivers change dramatically. It is truly unclear whether data assimilation is important for the upper atmosphere for prediction, or whether this is only needed for retrospective analysis.

I am sure that there are other issues, such as the data cleanliness that was raised on the telecom. This actually goes into problem 1. For example, how do you tell whether there were clouds over the observation site? If there were possible clouds, can you differentiate between the clouds and geophysical phenomena, such as the aurora? That is an extremely challenging problem.