CSDMS
community surface
dynamics modeling system

University of Colorado — Boulder
Campus Box 450
Boulder, Colorado, 80309-0450
Voice: (303) 492-6985
Fax: (303) 492-6388
Email: gtucker@colorado.edu
http://csdms.colorado.edu/

Further thoughts related to the CORES Study

Over the next decade it will be critical that we, as earth scientists, address questions regarding our understanding of global change and its impacts. In order to address these questions, we need to develop a comprehensive, mechanistic, process-oriented theory of the earth system (core to atmosphere) and its component parts. Research deciphering the history of the earth will contribute to our understanding of ES processes and help us improve forecasting of future change.

There are several challenges to addressing these questions. Earth-surface dynamics encompasses geomorphology, sedimentology, and related domains that deal with the surface of the earth and its sedimentary deposits. One key challenge lies in developing a quantitative, process-oriented theory for the processes that shape landscapes, seascapes, and sedimentary deposits. The ongoing revolution in geodetic data removes a key past impediment to this goal.

A related challenge lies in prediction. Can we develop a predictive capability for geomorphic and sedimentary systems on engineering time scales?

A final related challenge lies in computing. Petabytes of earth-surface data now exist, and meaningful interrogation requires computing. Moreover, given the complexity and nonlinearity of earth-surface processes, computational models are needed to express and explore theory, and to compare theory and data. Can we develop the computational models necessary to express and explore a comprehensive, multi-scale theory of earth's dynamic surface?

Our capacity to address these challenges will depend on infrastructure investments. Software cyberinfrastructure (CI) has become critical to modern geoscience, opening new pathways for discovery, but also requiring investment to create, coordinate, maintain, and share. Examples of geoscience CI include computational model codes, specialized data analysis packages, online data repositories, and various types of middleware. Efficiently addressing key science challenges will require active coordination of community CI, development and promotion of standards, and a sustained effort to maintain, adapt, and develop the existing CI. Achieving this goal will require sustained investment in support facilities, as well as training opportunities for the community. In short: software CI has helped accelerate the pace and scope of discovery, but NSF and other science agencies must recognize that sustaining this acceleration requires serious commitment and investment.

One advantage of research software CI is that there are often aspects that cut across disciplines (for example, standards for software interfaces and data formats are often broadly applicable). NSF could leverage this through closer coordination between programs in GEO and OAC, and by promoting cross-GEO opportunities in CI. Federal agencies often wrestle with similar challenges; for example, both DoE and USGS have made efforts to increase the value and sustainability of investments in computational models by promoting interoperability. Coordinated effort between NSF and agencies like these could be mutually beneficial.

Effectively harnessing the geoscientific computing revolution (a revolution not just in data, but also in the models that help us understand and synthesize those data) will require a new kind of technical expertise for the next generation of geoscientists. Today, the ability to manipulate data using computational tools is indispensable to geoscientific literacy. Few US geoscience departments seem to be adequately addressing this need. NSF GEO could contribute by partnering with programs such as OAC's CyberTraining, and by initiating new opportunities that would help promote a culture of computational thinking and analysis among geoscience undergraduates and graduates.

Although there is a place for centralized (one major facility, guided and shared by community. Example: research vessel) and decentralized (each investigator creates their own infrastructure via small grants. Example: single-PI mass specs) infrastructure, I would argue that a hybrid model results in the most efficient use of limited resources. This hybrid infrastructure consists of a centralized facility that coordinates shared infrastructure among distributed community of users and contributors. Examples include UNAVCO; cyber-oriented facilities e.g. CIG, CSDMS, CUAHSI. For infrastructure that is portable and lends itself to sharing, this third model can be an effective way to maximize coordination with a relatively small investment."

Sincerely,

Greg Tucker, CSDMS Executive Director
Fellow, Cooperative Institute for Research in Environmental Sciences
Professor, Department of Geological Sciences
University of Colorado, Boulder
University of Colorado
Boulder

CSDMS is an NSF sponsored project