Science Grand Challenges

A Synthesis from EarthCube End-User Workshops

Last updated 24 November 2014

Contents

Introduction

In order to reach out to potential end-users of the National Science Foundation's (NSF) EarthCube initiative, NSF funded a series of two dozen EarthCube domain end-user workshops throughout mid-2012 to late 2013, targeting a broad spectrum of Earth, atmosphere, ocean, and related scientists, including senior and early career scientists. The purpose of these workshops was to allow geoscience communities to articulate and document their cyberinfrastructure needs and what they would like to do in the future, in terms of accessing data and information within and outside their disciplines.

A specific goal of these workshops was to gather requirements on EarthCube science-drivers, data utilities, user-interfaces, modeling software, tools, and other needs so that EarthCube can be designed to help geoscientists more easily do the science they want and need to do. More specifically, science that helps address the NSF's GEO Vision, 2009: fostering a sustainable future through a better understanding of our complex and changing planet.

This document is a collection of all the science grand challenges identified during these end-user workshops. Workshop participants were asked to identify and prioritize a set of unifying science drivers and challenges they face in required to pursue key science questions.

Articulating Cyberinfrastructure Needs of the Ocean Ecosystem Dynamics Community

1. How will ocean acidification, warming, and hypoxia affect marine ecosystems? How do these phenomena affect the organisms? What species will be most impacted? What will be the impacts on ecosystem structure and dynamics?

2. As ocean circulation changes in response to warming and changing salinity, how will marine organisms respond to resultant changes in their environment? What will be the impacts on productivity, species distributions, and carbon flux in the ocean? How will such ecosystem changes affect human populations?

3. How will species distributions, life histories, and species interactions in polar oceans be impacted by the changes in ice cover? How will these polar changes impact other regions including changes in weather patterns? What will be the bottom up effects on marine coastal and open ocean ecosystems?

4. How do different physical, chemical, and biological processes come together to create more complex emergent properties in ocean ecosystems? What are the feedback loops among these processes and how do they give rise to biological, chemical, and genetic diversity in the oceans?

5. How are anthropogenic impacts other than climate change, such as eutrophication, overfishing, and coastal development influencing marine ecosystems? How does this influence our approach to ecosystem based management and conservation?

Bringing Geochronology into the EarthCube Framework

1. The primary scientific driver identified during the workshop if EarthCube were successful would be to understand and test hypotheses about the underlying controls on, and the relationships between, major earth systems. Achievement of this goal will entail establishing:
    - a robust, unified chronological **framework** for all earth history;
    - **correlation** of earth system records across a range of nested spatio-temporal scales;
    - **causality** between forcing, responses and feedbacks, including leads and lags;
    - **rates** of change of fundamental earth system processes.

2. The above provides a general framework for the goals of the geochronologic community within EarthCube. Specific examples of scientific opportunities and challenges facing the geochronologic community over the next 15 years that will lead to resolution of the dynamic interactions among. Earth systems include, but are not limited to:
    - the construction of a digital absolute geologic time scale used to resolve the times and drivers for biologic extinction as well as the rates of biologic recovery and evolution;
    - the pace, magnitude and drivers of climate change through earth history (e.g., the carbon cycle, oxygen, sea level, ocean chemistry);
    - addition of a 4th dimension to the construction and evolution of the North American continent, providing knowledge products to be directly integrated with EarthScope data;

- resolving interactions between rates, patterns, and magnitudes of erosion, landscape evolution, and sediment deposition, with climate change and tectonics in deep and more recent time.

Community Based Cyberinfrastructure for Petrology, Geochemistry, and Volcanology

1. Understand the co-evolution of the geo- and biosphere
2. Create a four-dimensional (space-time) description of the chemical and physical state of the Earth including the composition of and extent of fluxes between its major reservoirs -- core, mantle, crust, biosphere and hydrosphere.
3. Understand the role of disequilibrium processes in the formation and evolution of planetary bodies.
4. Integrate observations at volcanoes (e.g. seismic activity, ground deformation, emissions, magma chemistry and petrology, magma physical properties, plate tectonic parameters) in order to (1) forecast and mitigate natural hazards (2) understand and communicate volcanic impacts on society, and (3) to search for natural resources.
5. Map the feedbacks between planetary evolution, plate tectonics, volcanic activity and climate on short and long timescales.
6. Communicate the grand challenges of science to society

Community Modeling

1. How do we integrate and understand multiphysics between highly and weakly coupled systems? e.g. coupled dynamics of fluids, magma, and the solid earth at plate boundaries; co-evolution of hydrologic, geomorphic, critical zone and the deeper subsurface in the face of climate and tectonic drivers.

2. How do we integrate and understand the impacts of anthropogenic activity? e.g. feedback between components of the hydrologic cycle, atmosphere, and biosphere and land use and climate change and the role of human activities in these changes and implications for the quality and availability of water for drinking and other uses under increasing demands and scarcity.

3. How do we integrate the large degree of spatial and temporal variability in our models? Problems in the geosciences span time scales of $<10^{-6}$ to $>10^{15}$ secs to length scales of $<<10^{-6}$ to $>10^{6}$ m challenging the limits of both methodology and technology. This is unattainable purely by increasing resolution and necessitates the development of multiscaling modeling methods. Methods must account for the translations of variables in time and space, coupling between models, model (non)smoothness and uncertainties (whether numerical or data driven).

4. How do we determine model uncertainty and communicate it to both scientists and lay persons? Uncertainty arises from many sources including data generation and assimilation, model limitations, and poorly understood physical processes or processes represented at an aggregate scale using conceptual or empirical parameters. Models are increasingly being used as tools for “engineering” purposes and hence exert influence on policy, resource management, and exploration.

5. Our workshop did not attempt to develop use cases because of the diversity of problems addressed by models. However, we noted several examples of regions and problems that are

closely connected across space and time, and these provide opportunities for synergy across modeling communities. One example (of many) is modeling science in Cascadia (the Pacific Northwest). This region is a locus of intensive study of geology, geophysics, natural hazards (earthquakes, volcanoes, and landslides), landscape evolution, hydrology, climate, and ecosystems, and provides multiple examples of how models link to data integration, modeling on multiple scales, and the dynamics of coupled Earth systems.

The Cascadia subduction zone hosts a Long Term Ecological Research Network (LTER) site, is a focus area for GeoPrisms, and has extensive observations from EarthScope's US Array and Plate Boundary Observatory. Modeling is being used to understand problems including the role of fluids in the dynamics of subduction, and in the evolution of the landscape. These models integrate remote sensing, geochemical, geophysical, and geological data, with the attendant needs and challenges associated with access to data and interdisciplinary communication, many of which have been discussed at other EarthCube end user domain workshops. There are numerous challenges and opportunities for EarthCube that are directly associated with data acquisition, assimilation, and modeling in such cross-cutting regions or topics of study.

Cyberinfrastructure for Paleogeoscience

Overarching theme: The History and Future of Life and Environment Interactions on Earth

1. Establishment of a 4D framework for life and its environment on Earth. All other community priorities emerge from this primary objective. This framework will integrate across all time scales, regions, taxa, physical/geochemical properties, etc., and enable the ability to extract system state and rate of change at any spatiotemporal moment of interest.
2. Determine climate/ocean/biosphere interactions during times of great change in climate and environment, including extinction events, periods of extreme warmth, and changes over decades to millions of years, including the present geologic transition. Develop detailed characterizations of these past events to inform predictions of future changes.
3. Advance the capability to model the coupled carbon-climate Earth system, deriving the feedbacks, tipping points, and other processes from the paleo record, which is especially critical for deciphering the high magnitude/slow feedback mechanisms (e.g. ice sheet loss, deep ocean circulation) that climate models do not yet fully incorporate.
4. Assimilate paleo observations into process-based Earth System Models to reconstruct Earth history (lat, long, elev, and time), developing a suite of products that facilitate research, inform policy and decision-making (carbon cycle, sustainability, hazards), and deepen the public understanding of environmental vulnerability.
    - The participants recognized the value of recent efforts in characterizing many of the critical science drivers and challenges including:
    - " TRANSITIONS: http://www.sepm.org/CM_Files/ConfSumRpts/TRANSITIONSfinal.pdf
    - " NROES: http://www.nap.edu/openbook.php?record_id=13236
    - " DETELON: http://detelon.org
    - " Conservation Biology Workshop Report: http://www.paleosoc.org/CP_Workshop_Report_Oct_2012.pdf
    - " IGBP PAGES Report 57: http://www.pages-igbp.org

- " NRC 2011 Report: www.nap.edu/catalog/13111.html *[added by steering committee]*

5. Advanced cyberinfrastructure can enable the paleo community to reach these goals by 1) integrating small pieces of information scattered across the long tail (many small science projects), 2) refining sample ages and age uncertainty requisite to meet above challenges, and 3) facilitating a new era of vigorous collaboration across the many subdisciplines within and outside the paleogeosciences (e.g., hazards, paleomagnetics, tectonics, climate impacts, resource management, STEM education).

Deep Seafloor Dynamics & Processes

1. What are the geological/geochemical/physiological/energetic limits of life? What are the boundaries between biological and abiotic control of chemical reactions? How does geochemistry influence microbiology and vice versa? How do we incorporate microbial data into large-scale (global) quantitative geochemical models? How does bioenergetics influence food web dynamics, productivity, energy transfer and nutrient cycles and transform elemental pools between ecosystem compartments? What is the biogeographic, functional and structural distribution of microorganisms and what are the environmental parameters that most influence these distributions? Can these environmental parameters be used as indicators of ecosystem structure and vice versa? How do we define and interpret biomarkers (e.g., paleomicrobiology)? What are the scales of biological responses to disturbance, both natural and anthropogenic and how are these responses reflected in ecosystem connectivity, the relatedness of organisms? Can genetic tools be used to track ecosystem responses to environmental parameters, including adaptation and evolution?

2. What is the architecture of the oceanic lithosphere (including magma processes), and what happens to the plate as it ages from spreading center to subduction zone, as a function of spreading rate, environmental variability, variable crustal architecture? How does plate maturation impact subduction, and what controls the size and cycles of earthquakes in subduction zones? What role does the magma lens play in helping control tectonics/seafloor morphology? Do hot spot/ridge interactions influence the development of oceanic core complexes? What controls the origin, distribution, evolution, and morphology of seafloor features (e.g. seamounts, sulfide mounds), and what is the relationship between these processes/environments on biological communities and mineral resources?

3. What is the role of the deep ocean and subsurface in obtaining a 4D (spatial and temporal) understanding of global chemical and biological reservoirs, fluxes, and energy transfer? Such a perspective would allow us to address such transformative questions as: How does Earth regulate atmospheric CO2? What are the effects of deep sea biogeochemical processes on modern/ancient global atmospheric chemistry (C,O,S)? What are the relative contributions of biotic and abiotic deep ocean processes to global biogeochemical cycling? How can microbial data be incorporated into large-scale (global) quantitative geochemical models? What are the processes associated with serpentinization, including its diversity, range of environments, and consequences on global elemental cycles? How does the carbonation of peridotites affect global elemental cycles?

4. How do fluids in the subseafloor link thermal, tectonic, seismic, chemical and biological processes in a variety of deep-sea environments? What is the temporal evolution, extent and geometry of fluid flow within oceanic crust? What are the feedbacks between flow and geochemical and geophysical processes? How high within the water column do the fluids go? How does fluid flow effect the transfer of nutrients, energy and heat into habitable zones and

what is the role of fluid flow is establishing geochemical gradients and (micro)niches of habitability within the crust?

Developing a Community Vision of Cyberinfrastructure Needs for Coral Reef Systems Science

Key Science Drivers/ Questions in Coral Reef Science: Participants identified several high-priority science questions that will be the focus of interdisciplinary efforts during the next 5-15 years:

1. What processes are relevant to understanding the biological responses of coral reefs to biotic and abiotic drivers across temporal and spatial scales?
2. What are the mechanisms of coral reef adaptation and acclimatization to climate change?
3. How does symbiosis influence the biology and ecology of coral reef organisms?
4. How does the abundance and diversity of coral reef organisms influence community resilience at local, regional, and global scales?
5. How will invasive species, disease, and parasites disrupt coral reef ecosystem structure and function?

Early Career Strategic Visioning Workshop

1. Participants were all motivated by "grand challenge" geoscience questions concerning global climate change, weather prediction, and other such challenges.
2. This was a highly diverse set of participants, spanning the following geoscience domains (each with distinctive science drivers/challenges):
    - Atmospheric and geospatial sciences (anthropogenic aerosols, climate modeling, earth system science, land use, paleoclimate modeling, space science)
    - Earth Science (biochemistry, carbon cycling, climate change, climate modeling, earth system modeling, earthquakes, geochemistry, geochronology, geodynamics, geoinformatics, geology, geomorphology, geophysics, hydrology, igneous processes, metamorphic petrology, mountain environments, rivers, seismology, tectonics, water cycle)
    - Ocean science (biogeochemical cycling, chemical oceanography, climate change, coastal fluid dynamics, fluid mechanics, geochemistry, magmatic systems, microbiology, ocean acidification, petrology, physical oceanography, remote sensing)
    - Polar science (Antarctic ecology, carbon cycle, climate change, geochemistry, glaciers, ice, ice-ocean interface, meteorology, permafrost, sea ice)
3. Additional participants were from cyber or computer science, social science, and other domains including:
    - Computer science (cognition, machine learning, software)
    - Cyberinfrastructure (algorithms, big data, bioinformatics, climate informatics, cyber data management, data mining, GIS, hydroinformatics, lexical representation, semantics, spatial/temporal data, special databases)
    - Education (disability, disasters, geology, soil and water)

- Engineering (environmental nanotechnology, low temp geochemistry)
- Social science (governance, stakeholder visualization, trust)

4. Current challenges to high-impact, interdisciplinary science:
    - Institutional barriers to interdisciplinary science, particularly the tenure process in universities.
    - Resources and credit for sharing data, tools, models, and software
    - Connecting interdisciplinary research with interdisciplinary education
    - Not being limited to "brute force" accumulation of interdisciplinary data, particularly where the "Z' axis for geochronology is needed.

EarthScope

1. What is the present-day active deformation of the North American continent and how is this deformation related to the seismic activity, the growth and activity of faults, and volcanism?
2. What is the structure of the North American continental crust and underlying lithosphere and how is the structure related to the present day seismic and volcanic activity and over longer geological times to the assembly of the continent and the record of rifting, collision and maintain building over the entire continent?
3. What is the structure of the upper mantle beneath North America and selected regions along the core mantle boundary and how is the structure related to surface geological processes and mantle convection?
4. What is the rupture that unfolds during moderate to large earthquakes and how is that rupture related to the state of stress within the crust, the dynamics of earthquakes, rheology of crustal rocks and the presence of fluids within the crust?
5. How does the movement of aqueous and magmatic fluids influence the pore pressure, temperature, composition, and rheology of the crust and mantle? How does fluid influence lithospheric deformation and mantle flow?
6. Can the EarthScope facilities be used to map water (groundwater, atmospheric water, soil moisture, snowpack, glaciers, and vegetation water content) in time and space in the western United States and Alaska with a resolution that complements other meteorological measurements?

Education

1. Few of the big claims of geoscience (e.g. plate tectonics, global climate change, age of the Earth) can be explored in traditional student laboratory activities in the way that physics students can experiment with forces and accelerations or biology students can experiment with growing plants. Thus, the availability of professional-caliber datasets on the Internet has been transformative, insofar as it has allowed geoscience students to engage, in many cases for the first time, with the data that form the evidentiary basis of the concepts that they are studying. Geoscience education is out ahead of the other sciences in its use of large professionally-collected datasets for undergraduate education, and thus is having to pioneer new pedagogy around teaching and learning with data and models.

2. We live in a data-infused society. In today's workforce, data isn't only for scientists. In an ever-increasing percentage of professions, from nurse to car mechanic to teacher, adults are expected to be able to make use of data in the daily demands of their work. In our current education system, science and math are the places where students encounter data, and so teaching basic data using skills ("data literacy") in these classes has become a basic workforce

training imperative for all students.

3. Beyond the level of basic data literacy comes a degree of mastery that the workshop participants referred to as "data-savviness." The workshop wove an inspiring vision of the attributes of a data-savvy college graduate, skilled at using data and models to answer difficult questions and solve hard problems, facile with systems thinking and interdisciplinary problem solving, and able to make inferences about process and causality from Earth observations.

4. Our society faces many difficult decisions in the 21st century. The workshop conveners and participants think that better decisions would be made if a larger fraction of the populace used evidence, evidence grounded in data, in making decisions in their personal and professional lives. Undergraduate education is a prime time to establish the habit of mind of using data as an input to answering questions or solving problems. For the suite of decisions that revolve around approaching limits to growth on a finite planet (energy, water and mineral resource limits; environmental degradation; climate change), the relevant data will be of the sort served by EarthCube.

5. As computational models become a much more important part of the geoscientists' toolkit, geoscience education is endeavoring to convey this trend and prepare students to be a part of it. This is proving difficult, in part because students bring forward with them from K-12 the expectation that models are for demonstrating or explaining that which is already known rather than hypotheses to be tested by comparison with data. The epistemology of how scientists actually go about creating new knowledge using external runnable models is not widely understood by teachers or by the public.

6. Pre-college education is on the cusp of change, driven by the advent of the Next Generation Science Standards (NGSS). The NGSS foreground "science and engineering practices," including "analyze and interpret data" and "develop and use models." If and when the NGSS are fully and widely implemented, students will arrive at college with much more knowledge of the Earth, of data and of models—and with an expectation that science education should involve activities in which students construct meaning through active exploration. In the meantime, the prominence of the practices the NGSS is spurring a flurry of education research on the practices, including the data-using and modeling practices.

Engaging the Atmospheric Cloud/Aerosol/Composition Community

1. What are the sources and the removal mechanisms of chemical species in the atmosphere?

2. A lot of work has focused on improving Ozone but other species, for example, Methane and NOx, have been neglected in the process. The entire atmosphere system is sensitive to changes in NOx and needs to be considered. What are the effects of industrial NOx on atmosphere composition?

3. What are the exact roles of the clouds in the cloud systems and in the entire earth system? Several outstanding cloud-related deficiencies in the climate modeling are well documented by the research community and need to be addressed in the next 5-15 years, including the double ITCZ problem, poor MJO, too short and too regular ENSO periodicity, diurnal cycle and frequency of precipitation, inconsistent representation of radiation and clouds.

4. How do clouds affect the cloud feedback on climate sensitivity?

5. What is the role of clouds on biosphere or ecosystems or vice versa?

6. What is the spatial, temporal, size distribution and composition distribution of aerosol particles in the atmosphere and the aerosol particle emissions globally?

7. What are the exact roles of aerosols in the cloud and climate?

8. What is the impact of aerosol on severe marine storms?

9. What are the changes to Cloud Condensation Nuclei (CCN) with changes in aerosol loading?

Engaging the Critical Zone Community to Bridge Long Tail Science with Big Data

The central scientific challenge of the critical zone science community is to **develop a "grand unifying theory" of the critical zone through a theory-model-data fusion approach.** This concept expands on the classical notion of Hans Jenny's state equation for soil formation -- $S = f(cl,o,r,p,t,\ldots)$, where S is for soil, *cl* represents climate, *o* organisms including humans, *r* relief, *p* parent material (lithology), and *t* time -- into a 4D landscape-scale model of coupled physical/chemical /biological processes that frame the critical zone's evolution, function, and response to change. Developing such a grand unifying theory requires answering three broad questions:

1. How do tectonics, lithology, climate and biology co-determine the evolution of critical zone structure and function?
    - "structure" = 3D arrangement of the remnants of physical and chemical weathering from surface to bedrock, and associated spatial patterns in biological communities. It includes properties such as topography, chemical composition, porosity, permeability, and physical structure (cohesion, fracture density, shear strength, and similar properties), as well as biological communities both above and below the surface.
    - "function" = the processes of transforming and transporting energy and materials. CZ function includes all "ecosystem services", including water routing, storage and filtration; biogeochemical transformations such as nutrient, carbon or greenhouse gas uptake/storage/release; sediment flux; and others.

2. What are the drivers of energy and material fluxes (i.e. water, sediment, carbon, nutrients, solutes, etc.) moving through the critical zone?

3. How will critical zone structure, function and evolution respond to human and natural disturbances and over various time and spatial scales?

4. A second, yet equally important, challenge is whether a unified theory of the Critical Zone can create the necessary knowledge base to evaluate the complex issues of supporting sustainable landscapes. Several specific high priority questions were identified to provide detailed examples of the applications of the broader questions above. These were considered high priority in large part because of their immediate relevance to human and ecological sustainability issues.

5. What is the impact of human-induced changes to the nitrogen cycle on the land, air, water, and ecosystem of the critical zone across the scales where science-based management decisions and actions are made (individual land parcels to basin scales)?

6. What is the current distribution of soil carbon at global, regional and landscape scales, and given the drivers of these distributions how will soil carbon stocks change in the next 50-100 years?

7. What essential biodiversity and other biological variables are most relevant for characterizing the biological processes that co-determine critical zone structure and function? At what scale are these variables best measured?

Envisioning a Digital Crust for Simulating Continental Scale Subsurface Fluid Flow in Earth System Models

1. A high priority is understanding the evolution and functioning of the earth's critical zone, defined as the thin near-surface layer of the crust that sustains all terrestrial life. Fluid circulation and thus enabled energy, carbon, nutrient and other geochemical fluxes play a critical role in shaping the evolution of terrestrial biosphere and societies. The structure of shallow groundwater flowpaths, and its exchange with surface waters and the vegetation root-zone determine the seasonal water availability to vegetation and aquatic ecosystems, as well as carbon and nitrogen transformation and transport. There is no information on the material properties below the soil survey depth (~1m), preventing interpretation of field observations and modeling efforts across cm to watershed to regional scales.

2. Another high priority for our science is to advance a synthetic understanding of forcing of groundwater flow over many scales. Currently groundwater assessments are done at discrete scales, and information is not typically transferred between scales (upscaling or downscaling). The digital crust effort could provide a means to evaluate forcing of groundwater over a very wide range of scales (local, regional, continental), and to understand linkages between scales (e.g., effects of changing precipitation patterns and sea level on regional-to-continental groundwater levels, cumulative effects of water withdrawals, effects of regional-scale modifications of land use and surface drainage networks), as well as provide the basis for better incorporating groundwater in earth system models in ways that allow us to evaluate two-way feedbacks between groundwater and climate system on much larger and longer timescales then currently possible.

3. Another fundamental science question that bridges several geosciences disciplines and has extreme relevance for society is understanding the role of fluids in seismicity and tectonics. How can we quantify the distribution and magnitude of fluxes from the brittle to the ductile regime; can we better under the interaction between the hydrosphere and lithosphere?

4. Share different interpretations of available data into geologic structures. Data standards and tools are not currently adequate to allow domain scientists to share interpretations or to quantitatively compare and contrast different interpretations of the various kinds of geologic, geophysical, and mineralogical data used to infer geologic structures.

5. Organize the variability, connectivity, averaging, and covariance of disparate physical and chemical properties of the crust within the context of geologic structures. One of the central

challenges identified by having workshop participants discuss their knowledge and challenges within their own disciplines is the fact that the earth appears to have many structures depending on the particular properties used to define the structure, but many applications require synthesizing information on multiple properties (e.g., weathering → temperature, mineralogy, water flow and chemistry, etc.; nitrogen dynamics → temperature, water flow, oxygen, carbon, microbial community structure, active microbial biomass and/or metabolism). At the current time, we don't have a good sense of how these various representations of earth's structure compare, or at what scales different properties average, or how important properties co-vary (or don't co-vary). We also need to directly face the fact that all estimates of structure have a high degree of uncertainty.

6. Advancing our understanding of paleo-reconstructions of depositional environments. A specific example discussed at the workshop was the Gulf Coast. Building a complex model of the 3D geology of depositional environments over several periods of time (from Mesozoic to present day). Mapping these over time gives a much better understanding of the complex stratigraphy of these depositional environments, allows targeted sampling of geologic features to derive source evolution for tectonic investigations, and can aid societal needs such as energy exploration.

7. Another high priority science challenge identified by some participants is to further our understanding what the geologic, geomorphic, and environmental factors that determine the formation of the unique environments – e.g. karst systems. Karst systems exemplify the type of transformative and societal important research the digital crust would enable. Karst covers about 20% of the earth's surface, and are incredible fragile environments - subsidence (natural hazard), water quality, and urban/other planning that needs to understand the impacts of karst geology on water supply, construction, and other issues. Researchers currently studying karst environments often have to create their own datasets from many disparate and regional sources, and these studies often create a wealth of data that are not easily shared, so we do not have a comprehensive picture of global karst research and information.

Experimental Stratigraphy

1. How do we apply technical advances currently underway to experimental methods to create the next major advances in scientific knowledge? This will allow us to answer standing questions as well as ask completely new ones. These methods are likely to include:
    - Tomographic methods for the detailed *in-situ* investigation of strata as they evolve.
    - Long-range particle tracking methods for developing Lagrangian framework theories for sediment transport and deposition.
    - Computational methods for measuring and modeling individual sediment grains in large, complex systems.

2. What framework and model will allow us to gather and distribute large experimental data volumes for broad use beyond the original investigation? This is key to extracting greatest value from experimental data, increasing scientific efficiency at community level, and enhancing collaborations within and beyond the experimentalist community.

3. How can directly coupling laboratory experiments to outcrop-based investigations accelerate

advances in understanding? This approach is an excellent one for addressing major issues including: -Testing field-derived stratigraphic models (i.e. those directly tied to reservoir problems). -Addressing the grand challenge of integrating autostratigraphy and sequence stratigraphy. -Overcoming the community reluctance to incorporate experimentally-derived stratigraphic knowledge into stratigraphic models.

Integrating Real-Time Data into the EarthCube Framework

1. How can we better use real-time data to understand the processes of high impact events or phenomenon and translate that knowledge to better response procedures? Examples of critical cases include, but are not limited to:
    - Improved hurricane track and intensity forecasting; prediction and response to coastal inundation and shoreline breaches
    - Better understanding of tornado and severe convective storm genesis and warning
    - Earthquake and tsunami prediction
    - Better understanding, predicting, and managing of Hydrologic Extremes, e.g. flash floods
    - Early detection of harmful algae blooms
    - Prediction of large solar flare events for assessment of damage satellites.

2. How can we better understand scientifically compelling phenomenon with adaptive real-time, feedback-driven science? The following strategies optimize the scientific values of our measurements and enable new discoveries:
    - Dynamic sampling strategy to collect, analyze, and respond to real-time data
    - Response examples: Moving platforms, changing scan strategies, adjusting flight patterns, automatic adjusted of instrument signal processes, deployment of additional instruments, etc.
    - Using models in conjunction with adaptive strategies to improve sampling
    - Real-time awareness of instrument status to support rapid response to issues and improve data quality
    - Instrument validation (is it responding to its environment and can we adjust the instruments to improve the response?)
    - Tools that enable broad communication and collaboration during real-time mission oriented research
    - Detection and discovery of new, unexpected phenomena that need to be explored further
    - Tracking and sampling of transient phenomena

Integrating the Inland-Waters Geochemistry, Biogeochemistry and Fluvial Sedimentology Communities

The study of dissolved and particulate matter is of relevance to geoscientists and ecologists and encompasses diverse landscape scales and types, element and material cycles, approaches, and data collection contexts. This broad community is highly interdisciplinary; two different breakout groups came to the similar conclusion that this interdisciplinary nature makes it very difficult to label data sets as being "within" vs. "outside" the discipline. The sub-disciplines are complementary and interact with one another. Several unifying themes emerged, containing more detailed questions and challenges:

1. **We are in the era of anthropogenic changes.**
    - Need for understanding current states in relation to historic mechanisms driving the systems. The role of *legacies:* Climate history, soil structure, past disturbance, past land use.
    - What is the magnitude of climate change impacts vs. direct human perturbations such as land use change, aquatic environment modifications, and hydraulic engineering?
    - What are the global trends in carbon export, concentrations, gas evasion fluxes, and burial?
    - How will climate change affect higher latitude changes/creation of wetlands?
    - Trajectories and impacts of wetland degradation and restoration.
    - Advancing understanding of water ecosystem services to address landscape management.

2. **Connectivity:** Lateral linkages via water transport.
    - When and where do hillslope flowpaths connect and disconnect?
    - What is the impact of groundwater connectivity on stream processes?
    - How do we connect flowpaths and systems across scales?

3. **Temporal perspectives:** Predicting time of response to climate change and disturbance across biogeochemical response variables and water body types.
    - Pulsed events, extreme events (hurricanes, landslides).
    - How do seasonality, magnitude and duration of events influence biotic responses?
    - How does temporal variability impact societal needs or benefits?

4. **Spatial perspectives:** Predicting zones of conservation, transformation, propagation within inland water networks across range of response variables.
    - Defining and mapping the time-varying *hydroscape*, including small streams.
    - Upscaling different systems and fluxes to regional, continental, and global scales.

5. **Grand goal: Integrating and translating across spatial scales and forecasting in time needs:**
    - Improving the mechanistic understanding of processes.
    - Increasing spatial and temporal extent and resolution of observations.
    - Dynamics of fluxes, process rates, and system scales.
    - Linking across different types of processes and forcings: physical, biological, chemical, geomorphic, and anthropogenic.
    - Linking understanding of quantity and composition of complex constituent mixtures.
    - Using fine-scale data and understanding to inform global scale understanding.

- Determine how water, sediment and biogeochemical fluxes *throughout* a river basin are connected and affected by event magnitude, duration, sequencing and spatial extent.

6. **Other challenges:**
   - How do relationships between discharge and concentration impact downstream ecosystem function, and how do food webs impact biogeochemical and even hydrologic responses?
   - Estimate global time series of monthly carbon burial fluxes in all aquatic depocenters, and carbon gas exchange fluxes across all water surfaces.
   - How do floodplains function geochemically and geomorphologically?
   - Understanding delta subsidence and retreat due to decreased sediment supplies.

Meetings of Young Researchers in Earth Science (MYRES) V: The Sedimentary Record of Landscape Dynamics

1. What processes are relevant to understanding landscapes and mass flux (i.e. sediment budgets) in the past, present, and future across different temporal and spatial scales?

2. How is sediment generated and changed as it moves through the landscape?

3. How does downstream transmission of Earth-surface materials filter and record the frequency and magnitude of Earth's environmental changes?

4. How does life influence surface processes and transform environmental signals preserved in the sedimentary archive?

5. To what extent do extreme events control landscape evolution and stratigraphy?

6. How do the effects of tectonic and climate conditions propagate through the landscape and depositional system? At what scales?

Ocean 'omics science and technology cyberinfrastructure: current challenges and future requirements

1. How do physical and chemical oceanographic parameters and biological population structure and function co-vary within and between different oceanographic provinces? Do steep physical and chemical gradiients result in steep microbial functional gradients and drive changes in microbial biodiversity? Do feedbacks exist in both directions?

2. How does 'omic and population plasticity in microbes bolster ecosystem resilience to disturbances? How does global change and environmental disturbance impact genomic repertoires, transcriptional organization, protein and metabolome content, and biogeochemical activity?

3. What are the underlying molecular and boichemical mechanisms that regulate the physiological responses of microbes to environmental change, and their downstream biogeochemical consequences and feedbacks?

4. How do microbial communities in the ocean fluctuate as a function of distance from land, seafloor spreading centers, gyres, and upwelling zones? How do they change as a function of geochemistry, currents, and crustal age? How does this affect the flux of matter and energy in the surface and deep sea?

5. By what microbially-mediated mechanisms does rapid polar climate change affect the budget of greenhouse gases in the context of permafrost thawing and dissolved organic carbon release and transport, in time and space?

6. How can 'omics data be more effectively leveraged into predictive frameworks for understanding ecosystem processes and their future trajectories? How can 'omics data be distilled into tools useful to managers and stakeholders for efficiently monitoring ecosystem change and detecting ecosystem impairment?

Rock Deformation and Mineral Physics Research

1. As planets age and cool, how do their physical & chemical properties and internal structures evolve under the extreme conditions of pressure and temperature? What material transformations occur in complex, multiphase systems within planetary interiors and how do these impact key compositional and rheological boundaries such as the lithosphere asthenosphere boundary and the D'' region at the base of Earth's mantle?

2. What processes determine where earthquakes occur to define the seismogenic zone, and how do they influence the tsunami-generating potential of seismic rupture at subduction megathrusts?

3. What are the factors that dictate the spectrum of fault slip behaviors and the physics of slow earthquakes where self-sustained, quasi-dynamic ruptures propagate at velocities dictated by unknown processes?

4. How do the physical and chemical properties of planetary materials control the dynamics and magnetic behavior of Earth and other planets?

5. How can we best utilize seismological data and models from EarthScope and other sources to determine the composition, temperature, and flow fields that produce tectonic processes on Earth's surface?

Science-Driven Cyberinfrastructure Needs in Solar-Terrestrial Research

The latest NRC Decadal Survey in Solar and Space Physics outlines four overarching key science goals for solar-terrestrial studies in the coming years. Below are more-focused science goals, consistent with the Decadal Survey goals, that we anticipate will benefit most from investments in cyberinfrastructure during the next 5 - 15 years:

- **Understanding the couplings among physically different domains ranging from the solar interior to the Earth's atmosphere**: The advent of "Big Data" (the aggregation of large, complex, heterogeneous data sets) in observations and numerical modeling holds promise for rapid progress in solar-terrestrial research. Space- and ground-based observatories will provide important constraints for models in terms of boundary conditions and synthetic observables. New observational data and computational advances provide new opportunities to develop cutting edge, data-driven models for the evolution of the magnetic flux below and above the solar surface, its influence throughout the heliosphere, and its impact at Earth. New cyberinfrastructure is required to improve our knowledge of the transfer of physical drivers across different physical domains from observational data and numerical simulations.

- **The study of the fundamental processes through which magnetic energy is generated, stored, released, and propagated**: This is critically dependent on an advanced cyberinfrastructure that enhances our ability to assemble, analyze, and visualize multi-instrument, multi-wavelength datasets covering multiple temporal and spatial scales in combination with detailed physical models. The application of computer vision and machine learning techniques to identify features across different physical dimensions and to better mine large, distributed databases will be needed to enable event identification and statistically driven analysis. Of particular interest is understanding the process of magnetic reconnection, the primary mechanism for energy release in solar flares and coronal mass ejections, which controls the occurrence and severity of magnetic storms through transport of mass, energy and momentum both at the sunward side of the magnetosphere and in the magnetotail.
- **Predicting the solar wind and Interplanetary Magnetic Field in the near-Earth environment.** Understanding the origin of magnetic flux structure at the Sun, and how it evolves during magnetic eruption and propagation through the heliosphere to produce the relevant spatial scale of B_z variation near Earth that drives magnetic storms, will depend critically on in situ and remote sensing observations from the *Solar Dynamics Observatory*, *Magnetospheric Multiscale, Solar Probe Plus* and *Solar Orbiter* and other spacecraft, as well as ground-based facilities, combined with modeling techniques capable of simulating CME flux ropes from the Sun to the Earth. The many disparate types of data and the broad range of spatial and temporal scales involved in both observations and models present a substantial cyberinfrastructure challenge.
- **Understanding the acceleration of particles throughout the Sun-Earth system.** Acceleration of electrons and ions, often to extremely high energies, is ubiquitous throughout the solar atmosphere, heliosphere, magnetosphere, and ionosphere, and creates hazards for humans and technological systems (spacecraft, communication and navigation systems, and even aircraft) everywhere within Geospace. In every region, important tasks remain, such as: identifying the acceleration mechanisms that operate in the various regions of the Sun-Earth system; determining which mechanisms are most important at different times and locations; identifying common *vs*. distinct mechanisms in different regions; identifying the more important plasma instabilities that operate in the different regions and the role they play in particle acceleration under varying conditions; and following the propagation of accelerated particles within and across regions of the Sun-Earth system.

Understanding and forecasting the effects of forcing on the coupled Ionosphere-Thermosphere-Mesosphere (ITM) system. The ITM system presents a unique challenge in that strong coupling between charged and neutral species dominates physical processes. The system is responsive to external forces, e.g. reconnection, which impose global electric fields and magnetic currents, but also to internal processes, e.g. tropospheric heating and upward transmission of tidal forces, ionospheric instabilities, ion-neutral collisions and frictional drag. The coupled system demands cross-disciplinary study involving data acquired over multiple time and distance scales from ground and space observatories. Our ability to facilitate telecommunication and navigation, prevent catastrophic failure of the power grid during magnetic storms, or protect space assets from collisions demands accurate forecasting of the ITM response to forcing. Unique to this effort, international collaborations often require the participation of poorer countries with desirable locations for observations, but without the means to install instrumentation or distribute data in optimal ways.

Sedimentary Geology

Three overarching societal issues were highlighted as drivers that will condition research within the SGC community over the next 5-15 years. Multiple scientific challenges were identified relative to each driver. The primary theme of SGC is to fully integrate our discipline with Earth, Atmospheric, Oceanic, Biologic, and quantitative sciences in addressing the sedimentary dynamics of Earth and planetary systems from the beginning of time, and the current role of human interactions into a sustainable future.

1. **Driver #1** - Securing the energy and water resources needed for an increasing global population while balancing resources for a sustainable Earth.

   Related research challenges within the SGC community:

   - Predicting lateral spatial heterogeneity in the geometry and physical properties of sedimentary rocks/bodies. This is a necessity for effectively predicting resource distribution, modeling fluid flow, and mitigating contaminant problems.
   - Improved understanding of organic-rich fine-grained sedimentary systems throughout geologic time, particularly their origin and the processes that generate sedimentologic and geochemical heterogeneity.

2. **Driver #2** - Understanding the Earth as a system, the nature of global climate change, and its impact on climate change on life, the environment, and Earth resources.

   Related research challenges within the SGC community:

   - The deep-time sedimentary record must be scrutinized to learn how the Earth's climate system operates in periods of stasis, rapid change, and greenhouse and icehouse conditions. This requires:
     - Continued development of proxies for ancient climate, improvement in existing proxies, and reconciliation between proxies.
     - Analyzing the sedimentary record to identify and understand the components of deep-time climate change (forcing factors, feedbacks, tipping points) and the resultant impact on the deep-time Earth system (changes to hydrologic cycle, weathering, denudation, sediment fluxes, nutrient runoff, ocean circulation, extinctions and originations of life, etc.).
   - Develop a deeper understanding of the interplay between life, the physical and chemical environment in Earth's past, climate, tectonics, environmental change, and sedimentary processes.

3. **Driver #3** – Human activities influence, and in some cases dominate, many Earth surface processes. Understanding those anthropogenic influences will be necessary to minimize risks to society and insure environmental sustainability, particularly in deltaic, coastal zone, reefs, lake, and fluvial settings.

   Related research challenges within the SGC community:

   - Development of morphodynamic models of how sedimentary environments and landscapes responded on daily to millennial scales to climate change, sea-level rise, sediment supply, induced subsidence, engineered structures, etc.

- Determining how to use the sedimentary record to make predictions about future environmental changes, assess critical boundary conditions, quantifying parameters of environmental change, and evaluate rates of change.

Research Challenges Common to Drivers #1-3 above:

- What controls stratigraphic architecture and landscape dynamics? Revisiting the respective roles at varying temporal and spatial scales of autogenic (intrinsic feedback loops) vs. allogenic (climate, tectonics, eustacy) controls.
- Development of geochronological tools that provide more precise and accurate timing of critical events in Earth's history are necessary to meet all other research challenges. Geochronology must address: (i) the timing, duration, and rates of ancient climate change; (ii) how rapidly life responds to environmental change; (iii) the rapidity of geochemical changes in the Earth system; and (iv) the recurrence & magnitude of natural hazard events preserved in the sedimentary record.

Shaping the Development of EarthCube to Enable Advances in Data Assimilation and Ensemble Prediction

1. What are the limits of predictability in the atmosphere? What are the sources of uncertainty/ errors, and how do they feed into predictability?

2. What observations are critically needed to enhance atmospheric predictions, and where? What is the optimal configuration of the observation network?

3. What are the appropriate types, combinations, and configurations of parameterization schemes for high-resolution mesoscale models? How can the errors and biases in these parameterizations be quantified and corrected?

4. What is the optimal ensemble configuration to accurately predict the distribution of possible outcomes? How many ensemble members are needed and how should the ensembles be initialized?

5. What are the advantages and disadvantage of variational versus ensemble-based data assimilation techniques, as well as different types of hybrid approaches?

6. What are the most effective ways to post-process ensemble forecasts to achieve reliable and calibrated probabilistic predictions?

Structural Geology and Tectonics

1. What is the evolution of geological structures in three dimensions and at all spatial scales?

2. How can we use the rock record of deformation to better assess the rheology of the crust and upper mantle in different tectonic settings and over different spatial and temporal scales?

3. What are the timescales of different geological processes (fault motion, magmatism, landscape development, etc.) and how do they interact with each other?

4. How do we integrate between short-term (e.g., earthquakes) and long-term (e.g., mountain building) geological processes?

5. How do landscape development and other processes at the Earth's surface relate to geological structures and processes within the Earth's lithosphere?

6. How do mantle processes influence crustal deformation, and what is the dynamic interplay between magmatism, deformation, and mantle flow?