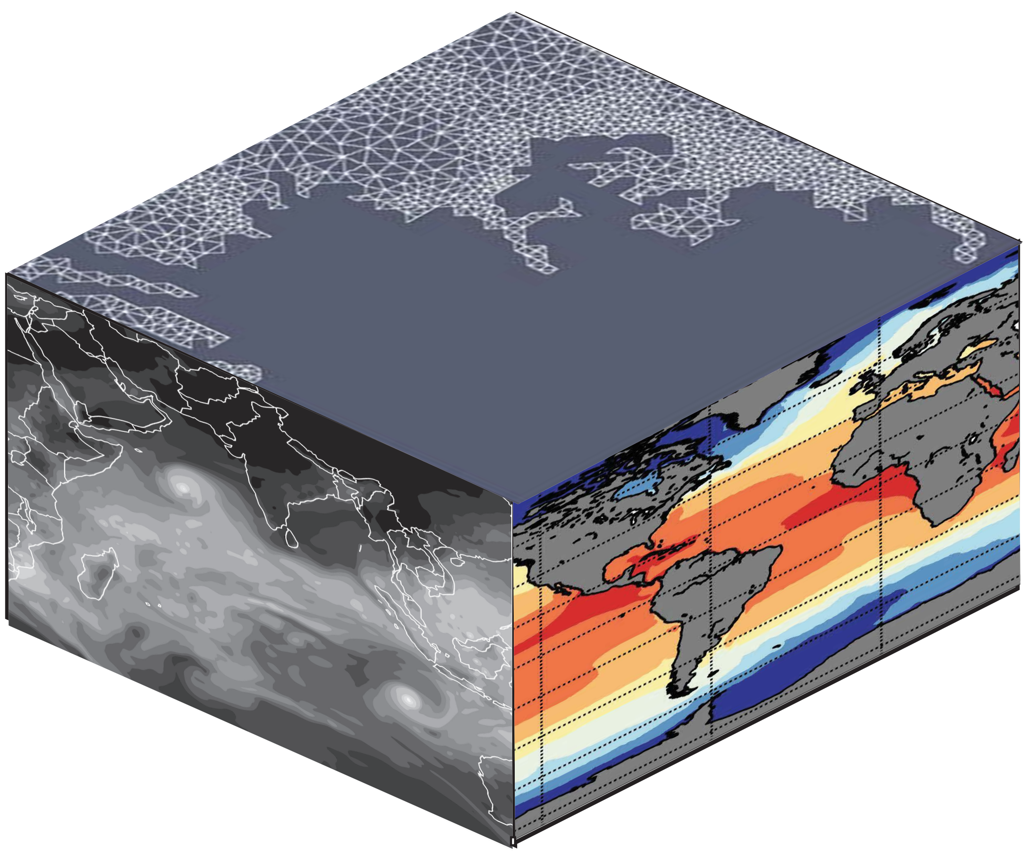
Advancing Cross-Cutting Ideas for Computational Climate Science

Workshop Report

Sponsored by
U.S. Department of Energy
Office of Biological and Environmental Research
and
Office of Advanced Scientific Computing Research

September 12–13, 2016
Rockville, MD



Cover

Clockwise, beginning with the top of the cube:

Description: Mesh representation of the coastlines of Greenland within the CISM-Albany model using a new algebraic multigrid solver, an ice sheet component option within ACME.

Courtesy of Ray Tuminaro, Mauro Perego, Irina Tezaur, Andrew Salinger, Sandia National Laboratories, and Stephen Price, Los Alamos National Laboratory.

Description: Monthly averaged Sea Surface Temperature from years 71-100 of a test simulation of the beta v1 ACME model with active ocean, sea ice, atmosphere, and land surface components.

Courtesy of Milena Veneziani, Los Alamos National Laboratory, and the DOE ACME ocean model development team.

Description: Snapshot of instantaneous integrated water vapor from a development version of the ACME model atmosphere component.

Courtesy of Kate Evans, Oak Ridge National Laboratory, and the DOE ACME atmosphere model development team.

Availability of This Report

This report is available, at no cost, at http://www.osti.gov/bridge. It is also available on paper to the U.S. Department of Energy and its contractors, for a processing fee, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone (865) 576-8401
fax (865) 576-5728
reports@adonis.osti.gov

Please cite this report as follows: E. Ng, K. J. Evans, et al. (2017) Advancing Cross-Cutting Ideas for Computational Climate Science Workshop Report, Rockville, MD. Sept. 11–12, 2016. DOI:10.2172/1341564

Advancing Cross-Cutting Ideas for Computational Climate Science

Workshop Report

Sponsored by
U.S. Department of Energy
Office of Biological and Environmental Research
and
Office of Advanced Scientific Computing Research

September 12–13, 2016
Rockville, MD

Workshop Chairs
Esmond G. Ng (Lawrence Berkeley National Laboratory)
Katherine Evans (Oak Ridge National Laboratory)
Program Committee
Peter Caldwell (Lawrence Livermore National Laboratory)
Forrest M. Hoffman (Oak Ridge National Laboratory)
Charles Jackson (University of Texas at Austin)
Kerstin Kleese Van Dam (Brookhaven National Laboratory)
Lai-Yung (Ruby) Leung (Pacific Northwest National Laboratory)
Dan Martin (Lawrence Berkeley National Laboratory)
George Ostrouchov (Oak Ridge National Laboratory)
Raymond Tuminaro (Sandia National Laboratories)
Paul Ullrich (University of California, Davis)
Stefan M. Wild (Argonne National Laboratory)
Samuel Williams (Lawrence Berkeley National Laboratory)

DOE/ASCR Point of Contact
Randall Laviolette
DOE/BER Point of Contact
Dorothy Koch

Contents

Table of Contents . . . i
Executive Summary . . . iii

1 Introduction **1**
1.1 Purpose . . . 1
1.2 Workshop overview . . . 1
1.3 Goals and structure of the report . . . 2

2 Big Ideas **2**
2.1 Process-resolving models . . . 2
2.1.1 Grand challenge . . . 2
2.1.2 Opportunities and potential solutions . . . 2
2.1.3 Example: Global cloud feedbacks . . . 3
2.1.4 Future directions . . . 4
2.2 Integrated model credibility . . . 4
2.2.1 Grand challenge . . . 4
2.2.2 Opportunities and potential solutions . . . 5
2.2.3 Example: Uncertainty in water vapor feedbacks . . . 7
2.2.4 Future directions . . . 7
2.3 Understanding and managing climate system complexity . . . 8
2.3.1 Grand challenge . . . 8
2.3.2 Opportunities and potential solutions . . . 9
2.3.3 Example: Ocean biogeochemical modeling . . . 9
2.3.4 Future directions . . . 10
2.4 Continuum model framework . . . 10
2.4.1 Grand challenge . . . 10
2.4.2 Opportunities and potential solutions . . . 11
2.4.3 Example: Pore-to-cloud continuum . . . 11
2.4.4 Future directions . . . 11

3 Cross-Cutting Issues **13**
3.1 Addressing model (software) complexity . . . 13
3.1.1 Grand challenge . . . 13
3.1.2 Opportunities and potential solutions . . . 13
3.1.3 Future directions . . . 14
3.2 New hardware . . . 15
3.2.1 Grand challenge . . . 15
3.2.2 Opportunities and potential solutions . . . 16
3.2.3 Future directions . . . 16
3.3 Model performance and time to solution . . . 18
3.3.1 Grand challenge . . . 18
3.3.2 Opportunities and potential solutions . . . 18
3.3.3 Future directions . . . 20
3.4 Data management and analysis . . . 20
3.4.1 Grand challenge . . . 20
3.4.2 Opportunities and potential solutions . . . 21

3.4.3 Future directions . . . 22

4 Structure of Collaboration **22**
4.1 Existing tools and collaborations . . . 22
4.2 Facilitation of communication across disciplines . . . 23
4.3 Outreach and learning as potential solutions . . . 24
4.4 Possible follow-on activities . . . 25

5 Conclusions **25**

Acknowledgments **27**

References **28**

Appendices **32**
List of Acronyms . . . 32
Workshop Participants . . . 33
Workshop Agenda . . . 35
Contributed Ideas Papers . . . 39

Executive Summary

This report presents results from the DOE-sponsored workshop entitled "Advancing X-Cutting Ideas for Computational Climate Science Workshop," known as AXICCS, held on September 12–13, 2016 in Rockville, MD. The workshop brought together experts in climate science, computational climate science, computer science, and mathematics to discuss interesting but unsolved science questions regarding climate modeling and simulation, promote collaboration among the diverse scientists in attendance, and brainstorm about possible tools and capabilities that could be developed to help address identified computational climate challenges.

Several research opportunities emerged from discussions at the workshop that participants believed could significantly advance climate science. These include (1) process-resolving models to provide insight into important processes and features of interest and inform the development of more advanced physical parameterizations, (2) a community effort to develop and provide integrated model credibility, (3) incorporating, organizing, and managing increasingly connected model components that improve model fidelity and potentially complexity, and (4) treating Earth system models as one interconnected organism without numerical- or data-based boundaries that limit interactions.

Participants also identified several cross-cutting advances in mathematics, computer science, and computational science that would be needed to enable one or more of these big ideas. It is critical to address the need for organized, verified, and optimized software, which enables the models to grow and continue to provide solutions in which the community can have confidence. Effectively utilizing the newest computer hardware enables simulation efficiency and the ability to handle output from increasingly complex and detailed models. This will be accomplished through hierarchical multiscale algorithms in tandem with new strategies for data handling, analysis, and storage.

These big ideas and cross-cutting technologies for enabling breakthrough climate simulation advancements also need the "glue" of outreach and learning across relevant scientific domains to be successful. The workshop identified several strategies to allow productive, continuous engagement with those who have a broad knowledge of the various parts of the problem. Specific ideas to foster education and tools to make material progress were discussed. Examples include follow-on cross-cutting meetings that enable unstructured discussions of the types this workshop fostered. A concerted effort to recruit undergraduate and graduate students from all relevant domains and provide them experience, training, and networking across their immediate expertise is needed. This will broaden and expand their exposure to the future needs and solutions and provide a pipeline of scientists with a diversity of knowledge and know-how. Providing real-world experience with subject matter experts from multiple angles may also motivate the students to attack these problems and even come up with the missing solutions.

1 Introduction

1.1 Purpose

Beyond the certainty of globally increasing atmospheric temperatures, there are many remaining unknowns regarding the specifics of how our Earth's climate will be affected over the next several years and into the next century. Examples include lengthening droughts, stronger and more quickly developing hurricanes, more frequent severe winter storms in the northeastern United States, and rising sea levels due to collapsing ice sheets in Antarctica and Greenland. These more localized changes have important scientific, economic, and societal impacts. As the scientific community explores these critical aspects of climate change, with the aim of informing stakeholders over the next 10+ years, there is a growing recognition of the expanding requirements for multiscale, global, coupled Earth System Models (ESMs) as a tool for climate scientists. Such ESMs are expected to provide much more detail and fidelity through the addition of new physics, chemistry, and biology and the use of high spatial and temporal resolutions. They are also expected to provide improved accuracy and automated diagnostics highlighting the model credibility and degree of uncertainties in their predictions. As a result, there are significant challenges to overcome in terms of their construction, algorithmic developments, and computational requirements. Also, as computation capability transitions from petascale to exascale over the next 5+ years, high-performance computer systems are expected to become larger and more complex. This poses additional challenges in terms of the ability to execute and process information from ESMs robustly and efficiently on current and future high-performance computer systems.

Advances in applied mathematics and computer science are crucial to overcoming many of the algorithmic and computational challenges that climate scientists will be facing. A multidisciplinary approach, involving climate scientists, applied mathematicians, and computer scientists, is required to tackle the issues properly and make the necessary breakthroughs in climate science. While many successes have been demonstrated through existing programmatic investments, we propose that longer term and broader efforts are also necessary to realize the promise of new scientific discoveries.

1.2 Workshop overview

Recognizing the need for much closer collaborations across multiple domains to meet the challenges outlined above, a grassroots effort to generate fresh thinking was initiated by a group of climate scientists, applied mathematicians, and computer scientists. The goal was to discover and chart the optimal directions for climate modeling and motivate the latest and as yet to be uncovered developments in mathematics and computer science that will be needed to address new scientific and computational requirements. The first step forward was a workshop, which was convened September 12–13, 2016, in Rockville, MD, with the objective of discussing bold new computational ideas to address longer term science needs for climate modeling. An open invitation was extended to researchers involved with any combination of climate science, applied mathematics, and computer science to submit "ideas" papers on any topic that addressed the future needs of climate science and/or how mathematics and computer science advances could be leveraged to help. Lead authors or their designees were invited to attend the workshop. All authors were encouraged to read and consider the papers before the event to begin conversations and generate ideas and solutions. A total of 59 ideas papers were accepted, and each of them is included in the Appendix of this report.

The workshop was structured to provide ample opportunities for computational climate scientists to present the state of the science and critical bottlenecks to research progress and for applied mathematicians and computer scientists to offer potential solutions. Several rounds of follow-up discussions, in which both groups discussed issues and solutions, provided a venue for the further maturation of ideas. To generate vigorous thought and discussion, several plenary talks and a poster session were held to exchange ideas from the papers. This report summarizes the challenges, strategies to address them, and possible avenues for execution that were identified by the participants at the workshop.

1.3 Goals and structure of the report

The purpose of this report is to summarize the findings at the workshop and identify potential opportunities to advance climate research through collaborations among climate scientists, applied mathematicians, and computer scientists. In Section 2, several grand challenges in climate science are presented, together with examples and descriptions of future directions. In Section 3, a number of areas in applied mathematics and computer science that were identified at the workshop as relevant to computational climate modeling are presented. Issues related to the structure of collaborations and outreach were also discussed at the workshop and are summarized in Section 4. The report ends with some concluding remarks in Section 5.

2 Big Ideas

2.1 Process-resolving models

2.1.1 Grand challenge

Key processes, particularly convection and cloud processes [Papers 19, 23], cryosphere feedbacks [Papers 3, 7, 22, 29, 47], biogeochemistry [Paper 58], and human-environment interactions [Paper 13]—along with their emergent features such as organized convection and surface heterogeneity—are poorly represented in modern global ESMs [28]. This is largely because the native scales of those processes are too small to be explicitly simulated, so they are mostly parameterized in these models. With intrinsic assumptions about the subgrid processes and their interactions with the explicitly resolved larger-scale environment, subgrid parameterizations have large structural and parametric uncertainties that are not well constrained by observations. In addition, many parameterizations used in current ESMs do not demonstrate proper convergence with increasing temporal or spatial resolution [Papers 44, 53]. In order to improve the representation of fine-scale processes and their upscaled effects in global modeling systems, efforts have emerged to develop specialized models that are built with a minimal set of assumptions. These *process-resolving models* provide insight into the emergent properties that occur in association with the processes and features of interest and can inform the development of advanced physical parameterizations for global climate modeling systems.

2.1.2 Opportunities and potential solutions

Several capabilities have been proposed to advance the development of cutting-edge process-resolving models:

- Advances in the development of lightweight and massively scalable numerical methods that can handle multiple scales of behavior are needed to ensure rapid throughput at extremely high model resolutions. This research aims to construct global process-resolving models with the capability to better utilize computational resources to maximize years of simulation per wall-clock day.

- Improvements in the capability to place computation "where it is needed." For instance, boundary layer clouds are currently poorly represented in Global Climate Models (GCMs) due to models' relatively coarse vertical resolution, leading to significant uncertainty in computations of energy balance. In the horizontal direction, the most obvious improvements in model performance occur due to better representation of topography; using the variable-resolution capabilities, where available. Grid resolution can be enhanced over rough topography but kept relatively coarse elsewhere. In this sense, optimal configurations of the model can be found that produce the best representation of the climate at minimal computational cost.

- If the former point identifies moving regions of needed attention, research could be focused on the ability to adapt computational meshes during run-time to reduce model error and improve the representation of fine-scale features [Papers 39, 42, 49].

- Load-balancing strategies capable of distributing a heterogeneous workload among processors could potentially do so over heterogeneous architectures as well [Paper 31].

- Development of methods for rapid calibration of free parameters with climate observations. Traditional approaches that rely on long-term climate simulations to tune the top-of-the-atmosphere radiative balance will no longer be feasible, so alternative techniques are needed [Paper 25].

- Development of techniques for handling of output data associated with these simulations. Storing all prognostic variables on disk is likely infeasible given the high temporal and spatial resolutions required by these models.

- A co-design effort to define domain-specific languages relevant for translating models of physical processes to heterogeneous architectures [Paper 16].

Many of these capabilities can only be attained by bringing together Earth system researchers and scientists, applied mathematicians, and computer scientists. A guiding objective of this work would be to optimize the scientific value produced for each unit of computational effort, including a focus on computational methods that are capable of reaching peak performance.

2.1.3 Example: Global cloud feedbacks

Among the processes described above, global cloud feedbacks are perhaps the largest contributor to climate model errors, and one of the largest uncertainties in future projections of global climate. A precise and accurate representation of underlying feedback processes requires extremely high model resolution that is unattainable with present general-purpose modeling frameworks such as the U.S. Department of Energy (DOE)'s Accelerated Climate Model for Energy (ACME) or the National Science Foundation's Community Earth System Model (CESM). It is difficult to adapt present frameworks to modeling controlling mechanisms due to the traditional separation between physics and dynamics that is assumed in coarse-resolution atmosphere models. Consequently, efforts have focused on developing extremely high-resolution Large Eddy Simulation (LES) and cloud-resolving models (CRMs) at local and regional scales [22, 33]. Further, the recent development of Superparameterizations and Ultraparameterizations [21, 29] has highlighted a potential avenue by which cloud-resolving models embedded within global climate models as a parameterization can directly improve simulation skill. The success of this methodology for improving emergent atmospheric features such as the Madden-Julian Oscillation (MJO) has suggested the benefits of process-resolving parameterizations [41].

The aforementioned modeling approaches have charted a path toward global cloud-resolution model development efforts now under way at several modeling centers [3, 34]. Global CRMs are potentially desirable future replacements for the atmospheric component of ESMs, as their high spatial resolution means that convective and macrophysical parameterizations are no longer needed, in turn leading to reduced model uncertainties. Transformative advances in computing have already availed substantial supercomputing power, with further rapid advancements in leadership-class facilities (LCFs) expected in the near future. Nonetheless, present GCMs leverage only a fraction of available computing power (see Section 3.3).

2.1.4 Future directions

Many outstanding questions remain on the feasibility of scaling process-resolving models to the global domain. In particular, it is unclear at present what the needed basic assumptions are to constrain the "search space" of possible models. For example, for CRMs: Do we require low-order or high-order methods? Finite volume or finite element dynamics? Column-wise or 3D radiation? What choice of microphysics parameterization? Further, future models will likely need alternative representations of topography; the improved representation of rapidly varying orography at high resolution may make terrain-following coordinates untenable. In its place, should we employ immersed-boundary or cut-cell technologies [Paper 59]?

Additional investigation will be needed in order to determine how to best tackle these questions, particularly in light of future hardware and software infrastructure. A larger challenge is to resolve processes that connect multiple components (atmosphere, ocean, land, cryosphere) in the coupled Earth system. For example, resolving continental scale land surface hydrology processes requires model developments in multiple model components. One example of this at continental scales has provided important insights on how groundwater flow influences evapotranspiration [24]. However, coupled, global ESMs that include process-resolving models for all components will exceed the capacity of exascale computing, and pose additional challenges for model coupling that will require advances in applied mathematics, computational science, and software engineering.

2.2 Integrated model credibility

2.2.1 Grand challenge

Providing model credibility for ESMs has always been a priority; it is necessary for building confidence in model output and analysis. However, the complexity of fully coupled global ESM has prevented a comprehensive and integrated approach to all aspects of credibility. A good fraction of ideas papers submitted to AXICCS concerned the many facets of creating a more integrated approach to model credibility including, for example, strategies emphasizing verification, validation, and uncertainty quantification [Papers 14, 20, 25, 26, 34, 38, 51], new ideas for exploiting scientific or computational insights to allow greater efficiencies in sampling sources of uncertainty [Papers 4, 8, 9, 36, 37, 38, 46, 55], and ideas for leveraging code design to more easily synthesize models and observations [Papers 2, 3, 15, 35]. The topic that most undermines efforts to assess model credibility is the challenge to quantify the effects of model biases on model predictions, particularly when extrapolating into regimes for which we do not have observational data [27]. Here we emphasize structural uncertainty to call attention to the important roles that scientists have within a domain long dominated by applied mathematics and statistics. However, the ideas that were presented could apply to many aspects of the challenges that are faced to assess model credibility given the often ad hoc nature of building climate models.

The current paradigm for addressing structural uncertainties at the global scale is to compare climate predictions from simulations across the Coupled Model Intercomparison Project (CMIP) ensemble. Predictions that are shared among multiple independently developed models are a powerful indicator of robustness insofar as many models differ numerically and exploit different notions for how sub-gridscale processes are represented (known unknowns). However, these models are not entirely independent and do not identify processes for which we are not aware (unknown unknowns). The most important additional ingredients for assessing model credibility are our scientific understanding of the processes that are responsible for reliable predictions, robust comparisons to observations within a historical context, and our assessment of the skill of individual models to capture these processes. Some growth in understanding may be elicited from the analysis of multi-model ensembles. However, pinpointing cause and effect within a multi-model ensemble can be very difficult when one does not have the ability to perform more controlled experiments. By embracing the challenge posed by structural uncertainty to include information about data, models, and sensitivities within a single model framework, one can better align the scientific, math, and computational solutions that are outlined below.

At present, it takes the concerted effort and computational resources of an entire community to develop and test an advanced climate system model. This expense is a major factor limiting efforts to explore alternate, credible solutions to simulating climate phenomena. The sampling process is also stymied by system complexity and the dimensionality of sources of uncertainty. Strategies for meeting limitations must be addressed if we are to make any practical progress on even small steps in applying advances in the applied math and statistics communities. Moreover, there is not a lot of shared experience between communities working on uncertainty quantification [Papers 17, 18, 24] and climate system models and their components. Any practical solutions will need to exploit our understanding of the system and provide more theoretical expertise across disciplines so that solutions make the most of limited resources.

2.2.2 Opportunities and potential solutions

A set of directions that would put us in a much stronger position to assess model credibility is listed below in no particular order. The challenge of accounting for structural uncertainty may be viewed as something that we can have in mind, even as we make solid progress in developing more tools and gaining experience in how to apply more formal strategies to climate model development and testing.

- Embedded error modeling addresses the question of how model errors affect predictions [Paper 17]. The approach treats model error by "embedding" it in some of the key model parameters, which means those parameters are treated as random variables. The distribution of these random variables is calibrated against data on the model outputs. Consequently, a model ensemble for sampled values of these embedded model parameters represents the effect of model error or structural uncertainty on the model outcomes so it can be studied. This approach is sometimes contrasted with Gaussian Process modeling [19], which uses a discrepancy term added to model output within assessments of model likelihood as a kind of discount when testing models against observations. The embedded error model is intended to represent the effects of errors closer to their origin, before those errors propagate, and as such, can capture the effect of model error on all model outcomes, even the ones for which no calibration data are available. This approach is well suited to complex systems where it may not be clear how errors develop. Although embedded error modeling has not been applied to phenomena as multifaceted as the climate system, it is perhaps the most direct approach to quantifying structural errors on predictions. It also depends on the close collaboration between applied mathematicians and climate scientists for its success.

- Data assimilation (DA) encompasses a broad range of mathematical methods to combine data and models. These methods may differ widely in their degree of sophistication and application purposes. Common goals of these model–observation syntheses are to allow models to reside in a state close to what is being observed and to do so by taking advantage of known sources of uncertainty in observations and models. This formal approach has several scientific and computational advantages [Papers 22, 29, 35, 38]. First, it allows a model developer to test predictions in hindcast-mode of newly introduced processes. This is what is being done already through the strategy of developing climate system component models independently using observations (often in the form of climatologies) as boundary conditions. This approach is simplistic because the approach neither offers a formal way to calibrate uncertain parameters, nor does it provide a formal mechanism to account for observational or boundary condition uncertainties when these data are prescribed. Second, from a computational point of view, DA allows us to initialize models for prediction without an expensive spin-up. To the extent that the time evolution of the model to the inferred initial state carries information that is important for predictability, we will need DA that can follow nonsteady behavior and time-resolved observation that constrains transients. We note that while dynamic equilibrium provides a useful simplifying concept for inferring time-mean behavior, it may be of limited value for predicting changes in a real-world context that is governed by nonsteady dynamics and implied time scales. Third, the use of formal DA methods allows for a more quantitative and targeted approach to selecting phenomena and identifying observations that would be the most informative for testing hypotheses and estimating uncertainties in model development. A powerful, computationally efficient DA tool being explored is the adjoint or Lagrange multiplier method to solve gradient-based optimization problems. In contrast to this intrusive method, nonintrusive approaches that operate on large complex systems are also being developed. They may have advantages in circumstances of high nonlinear dynamics. In many cases, it helps to have the application and mathematics communities working in tandem to tailor the approach to take advantage of application-specific knowledge of the physics and observations.

- Multi-fidelity methods: The computational expense of initialization and running experiments is a major factor limiting our ability to explore sources of uncertainty in model predictions. Several papers centered around ideas for improving computational efficiencies beyond numerics. One of the strategies that was mentioned repeatedly in the papers and discussions is the use of multi-fidelity methods [Papers 3, 8, 17, 22, 25, 26]. The idea is to take advantage of what can be learned from cheaper, possibly simplified versions of high-fidelity models to help anticipate what would be important to learn from high-fidelity experiments. Climate scientists often like making use of simplified versions of models for gaining scientific insights into phenomena. The potential exists for both developing a modeling hierarchy and exploiting them within a uncertainty quantification framework using multi-fidelity methods. The challenging aspect of this effort stems from the sometimes tenuous relationship that exists between phenomena at different levels of the hierarchy. Model versions with increasing resolution are perhaps the most straightforward examples of this as is done in Multilevel Monte Carlo (MLMC) [Paper 22]; however, one could envision different kinds of hierarchies based on levels of interactivity/feedbacks permitted. More experience is needed with both the science and mathematics of models with varying degrees of complexity to make this approach feasible.

- Emergent constraints: One approach for reducing uncertainties using contemporary observations is to identify relationships between contemporary variability and future trends from a suite of model results. When such a relationship is found and contemporary variability can be bounded with observations, future trends are thereby constrained. This strategy was employed by Hall and Qu [14], who evaluated the strength of the springtime snow albedo feedback ($\Delta\alpha_S/\Delta T_S$) from 17 models used for the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report and compared them with the observed springtime snow albedo feedback from the International Satellite Cloud Climatology Project (ISCCP) and ERA-40 reanalysis data. This approach was recently applied to the carbon cycle by Cox et al. [7] and to constrain carbon cycle feedbacks on future atmospheric CO_2 levels by [15]. While this method provides useful insights into the influence of model biases on future projections, it may not directly identify specific improvements to model structures that would result in reduced projection uncertainties. Methods for representing and sampling structural uncertainty within a single modeling framework could be used to determine and test metrics that may be more robust to structural uncertainty.

2.2.3 Example: Uncertainty in water vapor feedbacks

Water vapor feedbacks have twice the impact of greenhouse gases but are not a major source of uncertainty. The process that scientists took to gain confidence in this result provides an interesting case study for our discussion on model credibility. We know of no law that regulates size of the water vapor feedback. However, virtually none of the spread in multi-model simulations of 20$^{\text{th}}$ century climate originated from differences in the water vapor feedback [4], maintaining a virtual lock on relative humidity even as global climates evolve. For global warming, the region that is most critical to water vapor feedbacks is where the concentrations are maintained in the upper troposphere. Soden et al. [37] looked at interannual variations in upper troposphere moisture within satellite observations and compared them to climate model simulations with and without the water vapor feedback. The model was only able to reproduce the observed variations with water vapor feedbacks enabled. This gave confidence that the models capture these feedbacks in the right amplitude and location. The argument made use of multi-models to determine the relationship between an observable (water vapor) and the amplitude of global warming. Theory and experiments identified the upper troposphere as being critical to predictions, and an experiment was set up with a modified version of the model to test whether models were capturing this feedback for the right reasons. This example underscores the importance of developing methodologies for identifying sources of uncertainty and identifying the observables that can be used to test model skill.

2.2.4 Future directions

While there is general enthusiasm for these topics, the scientific, mathematical, and computer science communities need more experience in understanding the science questions, models, and methods for model–data synthesis. For this reason, it may be prudent to start with easier targets such as inverse and forward modeling of sets of parameters deemed to be important to a tractable science question (e.g., cloud feedbacks affecting climate sensitivity, uncertainties in ice flow initialization on rates of sea level rise). For instance, one could focus on the processes by which observations are used to configure the components of a climate system model to generate realistic modes of variability such as the El Niño Southern Oscillation (ENSO). Once these systems are in place, one could begin to explore how the formalized process could be improved through modeling hierarchies and embedded error modeling. The development of component model adjoints and the communities of individuals who are able to make use of such capabilities should be fostered. The availability of model adjoints will be especially important for systems that have long-term memories such as the ocean circulation, ice sheet states, and terrestrial and marine biogeochemistry, where the costs associated with spin-up are prohibitively expensive and time-consuming. Clearly, any improvements that result in greater computational efficiencies or throughput of simulating coupled Earth system processes also result in a greater capacity to assess model credibility.

2.3 Understanding and managing climate system complexity

2.3.1 Grand challenge

Earth's climate system consists of a balance between many complex and interrelated processes. Omitting or misrepresenting any of these processes or their interactions in Earth system models will lead to unrealistic predictions. Over the years, new processes have been added to climate models in an attempt to reduce model bias. Several papers from this workshop describe additional processes that should be included in comprehensive climate models. Developing a comprehensive list of processes critical to climate prediction is an ongoing and important task.

It is necessary to include all relevant processes to improve climate prediction, but getting them to work well together is challenging. Teasing out process deficiencies from observations (in which all processes are always operating and interacting) is tricky but critically important. Further work on techniques for decomposing observations into useful process-level information for model development and validation would be very useful. Sufficient testing is a constant deficiency in climate models because there are so many features and interactions to track.

Another difficulty in developing comprehensive climate models is the need for tuning to create the best possible model out of a collection of processes. In particular, imperfections in one part of the model will induce errors in other related pieces of the model. For example, even perfect models for interactive vegetation or ice sheets will give poor results if provided with inaccurate rainfall information. To avoid this, calibration is often performed on combinations of parameterized processes (rather than focusing on one parameterization at a time). In this context, improving one parameterization can disturb the balance of processes achieved through calibration that contains cancelling errors. As a result, improved representation of the Earth system often results in decreased model skill [16]. More effort to develop modularity, with tests that exercise various processes and also highlight tuning knobs that may not be linked to Earth system behavior, is needed.

Another issue is that the processes included in climate models are often not well understood. Interactions between multiple processes are particularly critical for capturing climate feedbacks. One approach to understanding model uncertainty is to perform a sensitivity analysis, whereby ensembles of simulations are performed with different values for uncertain parameters. Even if ensemble sensitivity analysis was exhaustive in its span of parameter space (which is itself untenable computationally), this approach is not sufficient for quantifying uncertainty because the structure of the parameterizations (as embodied in their equations and closure assumptions) may also be wrong [Papers 17, 36]. In addition, parametric uncertainty cannot capture the effect of processes that are not yet included in models. Parameter sensitivity studies may not capture data uncertainties; the initial and forcing conditions that drive the model are derived from information that is imperfect to an undefined degree [Paper 35] and requires methods such as data assimilation [Papers 18, 22]. Developing a methodology that addresses all these uncertainties is an important challenge for the climate and mathematics communities.

Capturing the important timescales in the climate system (including extreme weather events, ENSO, biogeochemical cycling, and deep ocean circulation) requires ensembles of long runs with high spatial resolution and full system complexity. These requirements are impossible to meet with current computational capabilities. Thus improving model efficiency and scalability are required for solving the climate complexity challenge.

2.3.2 Opportunities and potential solutions

This workshop identified several important processes currently missing from climate models, including, in particular, better boundary treatments [Paper 59], biogeochemical cycling [Papers 37, 43], subglacial hydrology [Paper 29], turbulence [Papers 8, 57], and human and urban systems [Papers 13, 57]. Better integration of multiple processes was also suggested, such as combining models of the hydrological and ecosystems cycles [Paper 58] and coupling iceberg calving with global ice sheet flow [Paper 7]. There are other highlighted areas where better processes are needed, but they have not yet been developed to sufficient maturity [Paper 48]. In some cases, processes need to be treated with better numerics, so that they are more accurate representations of the climate system [Papers 11, 44].

The computational challenges surrounding model complexity, while daunting, can be tackled. Section 3.3 describes many opportunities to improve algorithmic efficiency. It is also worth noting that additional processes or more sophisticated process representations can be added to existing models *without increasing wall-clock time* by running processes in parallel [Papers 4, 28, 33]. Implementing parallel physics is, however, a challenging research question because processes running in parallel don't know what each other is doing, which can lead to potential conservation errors [Paper 54]. Instead of executing individual processes in parallel, performing many simulations in parallel provides a quick way to create large ensembles. This approach to parametric sensitivity analysis is particularly useful when the timescale of interest is relatively short [Papers 34, 46]. Automation is useful for efficiently finding parameter settings that optimize model skill in these ensembles [Paper 24]. The need for process complexity can be balanced with computational constraints by using detailed offline calculations to inform online models with minimal expense [Paper 37]. On a similar note, models with very fine resolution and at high numerical accuracy can be used as benchmarks for global models to ensure that the impact of smaller scales are captured in large-scale models [Paper 23].

There are also several steps we could take to improve our ability to make sense of output from complex simulations [Paper 37]. When simulations have some processes replaced by prescribed observed values, one can isolate the process interactions responsible for comprehensive model behavior of interest. The plenary talk given by Chris Bretherton emphasizes the use of idealizations to distill the essence of complex systems into understandable pieces. Breaking climate models into individual components is also useful for avoiding compensating errors while tuning. An easy-to-use hierarchy of increasingly idealized climate-model configurations would make isolating the source of model behavior and avoiding compensating errors easier. Comprehensive model evaluation packages will also be important for ensuring that comprehensive models capture all aspects of the climate for which they are responsible [Papers 26, 51].

2.3.3 Example: Ocean biogeochemical modeling

Ocean biogeochemical models [Papers 37, 43] provide an example of the challenges and opportunities that come with complexity in Earth system modeling. The goal of such models is to quantify CO_2 uptake and to predict responses of marine productivity to climate change. Such predictions potentially involve 100s of species that respond to varying levels of different kinds of nutrients and sunlight. As an alternative to this traditional approach, one paper [43] suggested taking a new genome-based approach to predicting the biologic response to changing environments by using bioinformatics to predict the biochemical pathways and estimate fluxes through ecosystem-scale biochemical networks.

An additional challenge is that biogeochemical models require decades to centuries to equilibrate, making the development and testing of new ideas expensive and time-consuming. Thus all the strategies that were mentioned concerning model credibility (Section 2.2) also apply here. A range of options exists for making the scientific and computational problem more tractable. First is that the local problem is not as complex as it seems since there are usually only a few factors that dominate the evolution of the local system. Which factors are important can vary in time and space and may not always be obvious. Some ideas to manage this complexity include (1) building statistical models to predict the behavior of unrepresented species, (2) adaptively learning the minimum sufficient representation needed to resolve a behavior of interest, (3) making use of process-resolving column models to inform parameterizations (in an analogy to the cloud-resolving paradigm mentioned above), and/or (4) applying new advances in stochastic parameterization to better represent scale-dependent parameters within a coarser grid cell. Since observations of the primary variables (species and nutrients) are limited, computational approaches [Paper 38] will need to be developed that optimally make use of available observations while identifying what new observations would help constrain uncertainties in model predictions.

2.3.4 Future directions

Climate model complexity has reached a critical state that demands we take a new approach. More formal approaches to model reduction, initialization, calibration, and testing will be needed to keep track of future model versions. Maintaining a process-level understanding of model behavior will be challenging in the face of this formalization. Many of the ideas proposed during this meeting involve exploiting a process-level understanding to simplify the bigger challenge of understanding a complex model. Much work needs to be done to bring together these strategies into a broader work flow that synthesizes the learning that takes place within simpler contexts so that the level of complexity that remains for the fuller system is reduced. At present, we use observational data to accept or reject a parameterization based on how it performs within a coupled climate system model. It would be preferable if we had developed a set of observation-based challenges that test and possibly tune/calibrate individual model processes. Metrics and parameters could be discovered that associate the level of accuracy required to achieve a desired level of skill. The problems listed above are examples of problems that can help the community develop the capacity, knowledge, and tools for driving new advances in climate system modeling.

2.4 Continuum model framework

2.4.1 Grand challenge

As Earth system models become increasingly complex, coupling states and fluxes across component models with strongly interacting processes (e.g., land to atmosphere), finer discretizations, and growing degrees of freedom pose a significant opportunity to improve the methodology by which climate models connect processes across existing boundaries. However, stronger and more integrated couplings of separate Earth system components also create a larger and more complex computational challenge and require a mathematically rigorous coupling strategy to handle the data transfers [Paper 1]. Reduced-order models employing a continuum methodology could have significantly fewer degrees of freedom, provide tight coupling for critical processes spanning traditional component models, eliminate errors associated with coupling (e.g., temporal averaging, remapping, converting units, and computing derived quantities), and reduce communications, computation, and latency incurred from maintaining component model structures. In addition, correlated information can be treated consistently and numerical instabilities minimized since the number of different solvers can be reduced to a few or even one for some process representations. In a continuum model framework, the role of the coupler component to provide mediated data exchanges can be altered to drive other features such as solver execution or performance tracking, which could improve scalability and overall computational efficiency.

2.4.2 Opportunities and potential solutions

By taking a holistic top-down view of the entire Earth system, one can envision the design of integrated multiscale, multiphysics models of critical energy, water, and carbon and nutrient processes that span ocean, land, sea and land ice, and atmosphere realms and instead separate processes by physical coupling rather than "political" or traditional domain science boundaries. Focused on complete algorithmic descriptions of all processes, couplings would be synchronized based on their degree of strong-versus-weak and tight-versus-loose connections as defined by the computational science community [20].

Finer discretizations over continuous domains could be solved with functional approaches, which represent an infinite discretization and provide a flexible and rigorous platform for component model coupling. This encompasses data assimilation and UQ [Paper 22], calibration [Paper 17], and data analysis [Paper 55]. Functional approaches replace computational complexity due to finer discretization with mathematical complexity of handling flexible functions instead of data points. These approaches already generate very substantial activity in mathematics and statistics research and are only beginning to be noticed in climate research [Papers 22 and 55].

It is unnecessary to build a single model spanning all realms and process timescales, so a key strategy will focus on the determination of how each Earth system process is related to others and performing a prioritization based on available computing resources and algorithmic techniques. This will require additional scientific research and development into process understanding, algorithmic tools, and computational impacts of tight coupling of processes when developing parallelization techniques. Evaluating and constructing process parameterizations that should be solved simultaneously to a converged state represents a significant climate and computational science research challenge.

2.4.3 Example: Pore-to-cloud continuum

One example to illustrate the opportunity is the connection of soil moisture reduction through evapotranspiration to cloud condensation, precipitation, and infiltration (the pore-to-cloud continuum). These tightly coupled processes could be treated together and could then be coupled to the rest of the Earth system across interfaces determined to be more loosely connected (see Figure 1).

2.4.4 Future directions

Implementing continuum process representations to construct a new Earth system model is a high-risk endeavor with significant implications for current climate science. A complete reorganization of current Earth system model structures would be required, and implementation of a continuum strategy will necessitate new techniques for balancing computation and communication, with

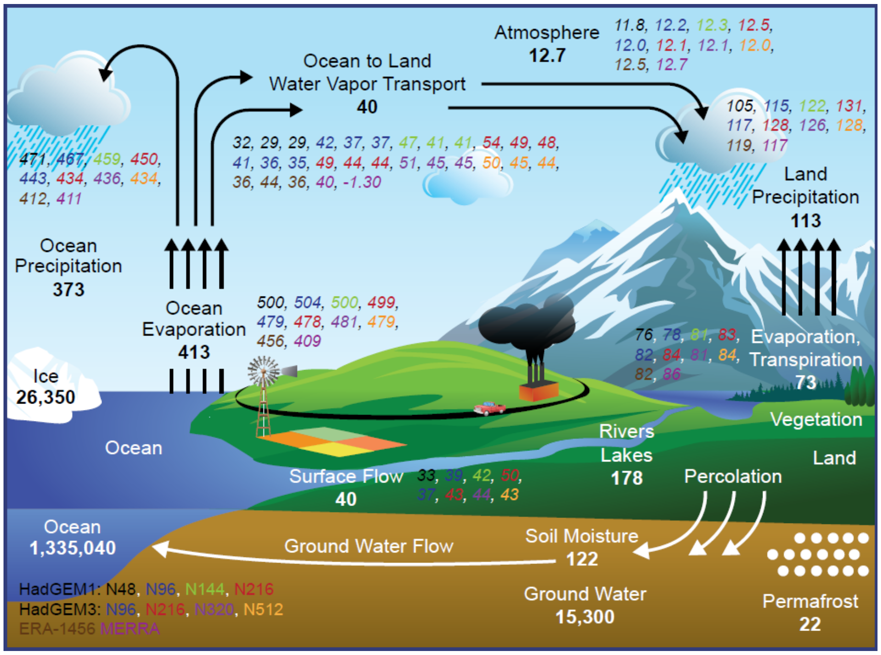


Figure 1: Directly connecting water cycle processes across land, ocean, and atmosphere realms, following a continuum modeling methodology, poses a significant climate and computational science challenge, but this approach could offer a more consistent treatment, eliminate coupler-mediated data exchanges, and improve scalability and computational efficiency. Water storage and flux units are 10^3 km^3 and 10^3 km^3 year^{-1}, respectively.

potentially large payoffs in model performance. Thus, climate scientists must work closely with computational scientists and engineers if such a pioneering effort is to be undertaken.

Component models for land, ocean, sea ice, land ice, and the atmosphere have proved useful for a wide variety of climate science questions, and maintaining the ability to conduct research with component models is important for investigating domain-specific mechanisms. Since continuum modeling is designed to solve equations of state under conditions of mass and energy conservation throughout all parts of the Earth system, this methodology conflicts with the traditional component model structure. However, strategies for modular design of key parameterizations could enable dynamic "plug and play" creation of traditional component models, built from the structures underpinning the continuum process representations. The degree to which code can be reused and the ability to dynamically build numerically stable and accurate models of sub-components are open questions.

One path forward for initiating a research effort to investigate continuum modeling strategies for the design of a new Earth system model would be to select one or two continuum processes to implement. The pore-to-cloud example described above and illustrated in Figure 1 is a prime candidate of relevance to ongoing DOE research. Having strong connections to the mathematics community (e.g., algorithms, solution techniques, and libraries, Section 3.3) and the statistics community (e.g., functional data analysis, Section 3.4.2) may enable faster development and proper implementation of the correct methodologies. Testing methods for coupling the energy cycle and vegetation processes into this continuum model are key initial research tasks, requiring an integrated multidisciplinary team.

3 Cross-Cutting Issues

3.1 Addressing model (software) complexity

3.1.1 Grand challenge

The simulation of a full Earth system (atmosphere, ocean, land ice, sea ice, and land surface components) requires that a multitude of physical processes be accurately represented over a wide range of temporal and spatial scales. As full simulations are built across several models developed by different domain scientists, managing complexity is a nontrivial but essential task. Tackling this is further complicated by the fact that climate models include a large number of parameters associated with approximated quantities, which are often impossible to determine precisely. Ultimately, climate modeling is intended to inform decision makers about possible future climate scenarios, and this in turn relies on verification and validation. Verification centers on assessing that the computer implementation accurately represents a conceptual description of the model and its solution, which includes partial differential equations (PDEs) as well as other types of models and assumptions. Yet, verification is linked to complexity management.

Verification in the presence of complexity, uncertainty, and simulation parameters is discussed in different contexts within many papers including Papers 10, 14, 26. Numerical accuracy and its relationship with parameter calibration [Papers 44, 53] were highlighted in the presentation associated with Paper 53 using an example based on an existing climate capability. In this case, multiple inaccuracies cancel and give the appearance of a validated solution. This can easily occur when models are erroneously calibrated with defective sub-components, as was called out in Section 2.3. That is, the calibration might falsely correct for deficiencies. Not only does this lead to inaccuracies, but it skews future assessments of potentially improved sub-components (e.g., with better accuracy/stability properties) if the assessment employs poorly calibrated parameters (determined with a previously errant sub-component).

3.1.2 Opportunities and potential solutions

Properly managing simulation complexity provides great benefits to domain scientists, to computational mathematicians, and to computer scientists. Many papers touched on complexity and the needed to improve simulation confidence (e.g., Papers 17, 26, 51). Ideas from computer science and computational mathematics can help manage complexity. This includes ideas from verification, software ideas associated with modularity, software engineering ideas and practices, and software ideas associated with programming environments for cross-platform performance on sophisticated computer hardware.

The presentation associated with Paper 53 highlighted opportunities to improve numerical accuracy/stability or, more generally, enhance verification at the most fundamental level. To address current concerns, rigorous benchmarking and testing of important algorithm kernels must be promoted. These tests must include examples that employ representative grids over realistic geometries (using typical elements types and aspect ratios that are characteristic of how a sub-component is utilized in larger simulations). More generally, tests should include increasing levels of complexity and be made available for those exploring potential sub-component improvements. Such tests might include mesh convergence studies perhaps employing the method of manufactured solutions. In addition to improving simulation confidence, benchmarks and published tests provide entry points for computer scientists and mathematicians to explore new concepts, algorithms, and implementations and to make meaningful comparisons with published results. Most likely, the computer science and applied mathematics communities would develop new benchmarks and publish additional results.

In an another direction, uncertainty quantification techniques can also assist in investigating and analyzing the sources of different types of errors. For example, Paper 17 emphasizes the need to understand both structural errors as well as spatial and temporal discretization errors and how certain novel UQ techniques (based on carefully embedding statistical model error terms in model parameters) can possibly account for structural errors.

An important consideration for improving complexity management is a closer association between mathematical concepts and software components. That is, the software framework should be designed to mirror mathematical relations (e.g., software abstractions based on vectors, fields, matrices, vector spaces, function evaluators, etc.). Mathematics is a universal precise language. Software written in this way is generally accessible to a wide community. For example, the discussion associated with Papers 50 and 54 noted that experimenting with time integrators can be cumbersome if it is not easy to associate any piece of code with the discretized form of the time level update to the ordinary differential equation. Papers 2, 3, and 4 touch on a related theme, that is, tools that help increase modularity and the building of sophisticated software by combining already developed agile components. This idea was adopted in the PISCEES project using the Albany and Chombo frameworks to build the FELIX and BISICLES ice sheet models [Papers 39, 45]. By leveraging preexisting development, newly developed dynamical cores can be built rapidly with advanced features (e.g., AMR or UQ). While these examples are mathematical, there are also software/performance incentives. A number of developments are under way associated with languages that facilitate asynchronous many-task programming as well as capabilities that facilitate performance optimization across multiple platforms (e.g., Intel Phi and NVIDIA GPUs). These software projects and their abstractions allow developers to obtain high performance over a set of different complex hardware platforms at a reduced effort [Papers 12, 28, 31]. It should be noted that existing refactorization efforts that leverage next generation platforms provide an opportunity to introduce much greater modularity and improve the management of complexity overall.

3.1.3 Future directions

The future is likely to bring increased complexities to the science of climate modeling. These increases will be due to the inclusion/combination of more physically realistic models, the greater availability of climate data, and the use of more sophisticated hardware platforms.

At this point, it is unclear exactly how these changes will impact climate model design decisions. However, it is clear that greater modularity and an increased emphasis on sub-component flexibility, testing, and performance will be necessary. This will allow climate scientists to manage the growth in complexity while leveraging contributions and foster collaborations from computational scientists in the applied mathematics and computer science fields. As noted above, modularity should consider a closer association between mathematical concepts and software components. Additionally, future modularity efforts should consider the separation of software implementation from the underlying hardware design while at the same time exposing fine-grain parallelism and allow for the possibility of task-based programming models that can address the heterogeneous nature of climate simulations.

While this section has focused on modularity and the importance of verifying individual modules as a means to addressing complexities, model interactions are extremely important and will grow in importance over the next decade. Paper 1 discusses a framework for addressing important aspects of heterogeneous numerical methods in a mathematically rigorous way. An important idea here revolves around using optimization and control ideas to provide a solid mathematical coupling strategy so that dissimilar numerical methods can function within a unified simulation. Paper 51 focuses on understanding the interactions between models that have been developed independently.

Models that do not directly interact with each other (noncontiguous model components) might have indirect effects on each other, so pragmatically understanding this propagation can be challenging. Despite the challenges of component *fractionation* (where each model has its own data, assumptions, and quality measures), the climate community will need to develop approaches for full-system model evaluation. This in turn requires computer science tasks such as building an infrastructure that can efficiently and flexibly transfer data between a large number of models as well as mathematical tasks such as developing appropriate full system metrics that might involve several model ensembles.

3.2 New hardware

3.2.1 Grand challenge

Historically, processor performance has generally followed Moore's law, the observation that the computing power that can be cost-effectively integrated on an integrated circuit doubles every 24 months [26]. Unfortunately, the performance of climate simulations has not kept pace with the improvements in processor performance because all improvements have been more than consumed by increased degrees of freedom and model complexity. In addition, performance is likely affected by several fundamental changes in processor scaling and architecture.

Despite the end of Dennard scaling (constant power density scaling) [9], processor performance has continued to improve according to Moore's law without exponential increases in power. This has been realized primarily through design changes such as energy-efficient circuits, core architectures designed for energy efficiency at lower frequencies, massive increases in parallelism (multicore, manycore, wide vector), and more software-controlled functionality (e.g., scratch pad memories). Lower frequencies tend to increase latencies and overheads (bottlenecks in strong-scaled climate simulations), while the massive increase in hardware parallelism has demanded a commensurate increase in parallelism expressed in the software. Until recently, many compilers for commodity hardware would fail to vectorize even the simplest stencil codes (let alone iterative point-wise equations or complex microphysics routines). Additionally, users' attempts to thread their simulations often ran afoul of bottlenecks in the maximum theoretical parallelization. Programming models and hardware that once virtualized data movement now increasingly demand programmers micromanage data movement through the memory hierarchy, including exploitation of any data locality. Failure to do so can result in substantial performance degradation.

Over the past decade, while peak processor performance continued to scale well, main memory bandwidth scaled much more slowly [30]. As a result, many processors and accelerators available today have machine balances of less than 0.1 Bytes of memory bandwidth per Flop, which is far below what is needed in many algorithms to achieve good parallel efficiency in general. The future trends in semiconductor manufacturing will only exacerbate and compound these challenges. In the strictest sense, Moore's law is dead (exponential performance progresses at a slower pace than doubling every 2 years). By 2016, we were already observing the implications on vendor processor performance and roadmaps. For example, Intel's "tick-tock" cadence, which alternated between process shrinks and novel architectural changes on a regular 2-year basis for the past decade, has been abandoned with the introduction of Kaby Lake between the 14nm SkyLake architectural "tick" and the 10nm Cannon Lake "tock". Similarly, it required 4 years for Intel's Knights Landing (2016) processor to double the performance of the Knights Corner processor (2012).

Whether the climate science community realizes it or not, they are in a race to the bottom with the semiconductor industry. Whereas the semiconductor industry is driving towards atom-scaled devices with exponentially increasing performance (perhaps no longer every 2 years), the climate community is driving towards ever-finer resolutions (e.g., global cloud-resolving models) with exponentially increasing computational demands. With perhaps only four more planar complimentary metal oxide semiconductor (CMOS) lithographic nodes, it may not be possible for the current trajectory of vendor processor offerings alone to enable more than a 2–4× increase in resolution. Any further increase in resolution will require a corresponding increase in power and system cost (e.g., 2× the resolution requires 8–16× the power and system cost). This suggests that it may not be cost-effective to increase the resolution of climate simulations beyond that attained in the 2025–2030 time frame without radical changes to algorithm, implementation, architecture, and semiconductor technology.

3.2.2 Opportunities and potential solutions

Several approaches have been proposed to bridge the potential performance impediments to simulation and analysis.

One obvious strategy to bridge the gap is to improve methods and algorithms. Whereas changes to (or adoption of) a linear solver may reduce computational demands by a constant factor, changes to discretization in space or time (e.g., higher order) can result in an exponential reduction in the computational requirements and mitigate increasing supercomputing costs [Papers 27, 41]. Unfortunately, ensuring a global benefit is an challenging prospect.

Due to the rapid size increase in both observational data and model output, a co-design effort to develop future computational resources to meet the unique needs of large-scale climate data analytics is needed [Paper 56]. Such systems typically require less computational capacity (e.g., fewer processor cores) but would benefit most from large, fast memory systems and high bandwidth input/output (I/O). To meet the growing demands of climate analysis and model benchmarking, a balance must be struck between high computational capacity resources and high-throughput resources at major computing facilities.

Vendors have begun to embrace the concept of benchmarks and proxy applications as a means of specializing their offerings for the computational demands of the various computing centers. Doing so provides them an advantage over their competitors since, with the same process technology and power constraints, they can deliver superior performance. Unfortunately, as vendors are profit motivated to create general solutions that balance the cost-performance trade-off across many computational domains, any solution for climate is likely to be suboptimal. Nevertheless, it has become increasingly cost-effective to design general-purpose, domain-optimized supercomputers specifically for their computational needs.

To mitigate the effects of evermore novel and complex vendor processor offerings, several discussions at the workshop focused around domain-specific languages (DSLs) [Paper 16]. The value of a DSL is its ability to describe the functionality of a method in the context of the domain (e.g., PDEs on structured grids) without prescribing a solution or approach to execution (e.g., Fortran/C with MPI+OpenMP). This separation of concerns is desirable as it amortizes the development cost required to port to each new generation of architecture. Unfortunately, substantial investigation is required to determine what the DSL should look like, and substantial effort is required to develop and maintain a DSL compiler.

3.2.3 Future directions

At the workshop, a number of discussions examined alternatives to the current and emerging computational roadblocks. These alternatives included a number of algorithmic [Papers 5, 6, 30, 40, 41] and model [Papers 13, 33] changes, performance optimization techniques [Papers 4, 28], and advances in semiconductor technology (optical interconnects, carbon nanotubes, etc.) as well as cost-effective semi-custom supercomputers tailored for the needs of climate science [Paper 19].

The past decade has seen the proliferation of IP-based (intellectual property) hardware building blocks for system-on-a-chip processor designs in everything from network routers to cell phones. These building blocks range from memory and network controllers to full, superscalar out-of-order cores (e.g., ARM, Tensilica). More recently, we have witnessed the emergence of open-source hardware, including high-speed I/O circuits and memory controllers as well as full processors (e.g., Berkeley's RISC V core [2]). With a fabrication-less design flow, it is possible not only to tailor the core architecture (out-of-order, cache sizes, vector widths, frequencies, etc.) but also the entire node (including memory) and system (including network) designs to match the computational needs of a scientific problem [11, 23, 36, 40]. Although one may discount this approach, several systems have successfully adopted it, including D.E. Shaw's ANTON and ANTON2 supercomputers designed for extremely fast and efficient molecular dynamics simulations [36]. Sunway's TaihuLight, the world's fastest supercomputer, followed a similar approach and produced a system arguably optimized for solving dense linear systems while still being capable of solving problems from other domains [10, 13].

The cost of such systems should not be underestimated, and several researchers have extensively tabulated the costs for such machines [8], dividing the cost into the nonrecurring engineering costs associated with designing and implementing the system and fabrication and integration costs. As the scale of the procurement increases, the speedup required to reach cost-performance parity with a traditional supercomputer quickly diminishes (e.g., to break even, the custom design needs only be 25% faster—a modest factor for domain-customized hardware—on a $100M procurement). As a result, for large procurements, there is a clear potential cost savings by leveraging the emerging customizable hardware market to tailor the computational capabilities of the supercomputer to the computational requirements of a domain such as climate science.

In the climate community, one of the biggest impediments to exploit emerging manycore and accelerated hardware is the substantial and required optimization and porting to new programming models (CUDA, OpenACC, etc.). By contrast, the basic building blocks of these semi-custom designs are traditional RISC cores, often the same as those found in commodity devices like cell phones. As such, existing C/Fortran code written using standardized programming models is immediately portable. To ensure one may leverage semi-custom instructions without unduly burdening the programmer with intrinsics, an optimizing compiler (cognizant of the new instructions) is auto-generated along with a synthesizable (e.g., Verilog) description of the core.

Substantial preliminary investigation is required to determine the efficacy of a domain-optimized supercomputer for climate. Architecting a domain-optimized supercomputer requires benchmarks to guide optimization. As the design space expands (combinatoric explosion of architectural parameters) and the evaluation infrastructure slows (from proxy hardware down to cycle-accurate simulators), the cost of evaluating configurations becomes expensive. As such, it is incumbent on the climate community to deliver compact, configurable benchmarks to the computer science and applied mathematics communities. These benchmarks must be future looking and designed to proxy the computational methods intended to be used in the 2025–35 time frame. Concurrently, the climate community needs to track the evolution of vendor solutions to quantify the emerging performance gap and to determine whether the architected solution can bridge the gap (constrained by expected procurement funds). Failure to be proactive in this manner will result in hardware and/or software solutions that are poorly matched to the computational challenges with climate science — a continuation of the status quo.

3.3 Model performance and time to solution

3.3.1 Grand challenge

The underlying dynamics driving changes in climate include a wide spectrum of complex interacting physical processes with many different types of nonlinear phenomena that span a range of spatial and temporal scales. These complexities often require high spatial resolution and drive increased model uncertainty, which is addressed with large model ensembles. The push for high-resolution simulations and increased ensemble size has exposed a pressing need for faster time-to-solution. These challenges require a strong collaborative research environment that promotes interaction between computational climate scientists, applied mathematicians, and computer scientists. Domain knowledge about the important physics of the problem is needed to inform appropriate choices for algorithms that best suit the problem. While hardware improvements provide a bedrock for faster completion, progress in the algorithms themselves will be essential in attaining significant reductions in the computational time needed to simulate different features of the climate system. A reduction in solution time allows for a combination of quicker solutions, longer simulation times to consider longer timescales, and/or higher fidelity models that consider more complexity for a given model execution.

3.3.2 Opportunities and potential solutions

To improve solution times, it is expected that climate simulations must leverage and drive advances in time advancement, solvers (nonlinear and linear), adaptivity (spatial and temporal), and where appropriate, the use of ensembles to gather statistics and sensitivities.

Ultimately, understanding the changes in our climate relies on accurately computing the spatial and temporal dynamics of a full Earth system model. This in turn requires improvements in the efficiency and sophistication of the discretization, for example, in finite-element strategies [Paper 6]. These schemes must generally be adapted to specific problem characteristics and must be appropriate for increasingly complex models. As many of the underlying sub-systems involve a wide range of spatial and temporal scales, improvements in model efficiency may come from targeting spatial and temporal resolution at specific processes and regions as needed.

Time-to-solution is also highly dependent on the inter-node communication requirements of spatial discretizations (latency, in particular). New spatial discretizations have the potential to target this cost by increasing the maximum stable time-step size [Paper 5]. Also, many schemes that support arbitrary order-of-accuracy can be tuned to enable improved performance on specific architectures. For instance, L1/L2 cache sizes and vector widths are hardware parameters that can provide dramatically improved performance depending on the order-of-accuracy of the numerical method.

Many advanced numerical techniques used in climate science rely on nonlinear and/or linear solvers. These include implicit time integrators, numerical optimization methods, and some uncertainty quantification/sensitivity analysis approaches. Preconditioner developments are often tied to specific problem classes. This is because an effective preconditioner must roughly approximate the inverse of a linear system, which may include interactions across a range of scales. This underscores the importance of the need for close collaboration between climate scientists and applied mathematicians in developing solvers/preconditioners that are effective for the linear systems arising in climate modeling.

In particular, due to the wide range of temporal scales, efficient time advancement techniques for stiff systems are essential. A number of recent advances in solver technologies have been made to improve solver convergence and robustness. These advances should be further investigated, extended, and adapted to climate situations. This includes implicit [Papers 30, 40], implicit-explicit (IMEX) [Papers 50, 54], operator split, and fully explicit [Paper 5] approaches, although it was widely agreed through ideas papers and discussions that algorithms with large time-step sizes that minimize data transfer across nodes and provide acceptable accuracy are the goal. Future efforts into nonlinear techniques should consider strategies to move beyond time integration advances such as traditional Newton-Krylov (e.g., [17]) to address the increase in model complexity. Two highlighted examples include nonlinear composite combination and nonlinear preconditioning [Paper 30] and Anderson accelerated fixed point methods [Paper 50]. Efforts to provide additional gains are suggested by using Newton-Krylov as a coarse operator for parallel-in-time and spectral-in-time methods [Paper 41]. Additionally, including optimized error constants for specific problems or methods that yield better conditioned linear sub-systems to improve linear solution times is of interest, especially within nonlinear solvers.

For linear problems, there has been significant progress in multigrid, domain decomposition, and physics-based solvers. Linear solver advances include convergence/robustness enhancements for certain problem classes or PDE operators (e.g., incompressible Navier-Stokes operators and/or Maxwell's equations), for problems that contain particular features (e.g., discontinuities or singularities), for certain discretizations (e.g., high-order discontinuous Galerkin methods, mimetic approaches, mixed finite elements), and to address meshes that might be thin with bad aspect ratios or might include boundary layers. There have also been advances in more general solver ideas such as multigrid methods based on K-cycles, compatible relaxation, and bootstrap- or energy-minimization-based multigrid schemes. Additionally, there have been advances in using auxiliary operators to precondition difficult linear systems. The general idea is that the auxiliary operator is more amenable to solver techniques and can still be used to precondition the original operator (e.g., shifted Laplacians for Helmholtz operators). For extremely difficult linear systems, direct solvers may be needed even though they typically have much higher memory and computational requirements. In addition to robustness/convergence enhancements for difficult problems, there continues to be ongoing research to improve solver scalability and performance on advanced/emerging architectures that might include hundreds of thousands of computing units. This includes methods such as communication-avoiding techniques, solvers with increased parallelism (e.g., concurrent processing of levels within a multigrid hierarchy), and solvers that have increased locality (e.g., nonlinear domain decomposition or recursive domain decomposition approaches where algorithm choices take into account the hierarchical nature of specific compute architectures). To leverage recent solver advances within the climate sciences, further research is vital to adapt and enhance solver ideas to the types of realistic complex situations that arise in the modeling of Earth systems. Ultimately, promising solver directions are often driven by a combination of domain-specific knowledge that comes from a combination of climate science and preconditioner expertise.

Many climate processes span a wide range of dynamic scales where fine spatial or temporal resolution is needed in a small subset of the domain in order to correctly model key dynamic processes, while large portions of the domain exhibit much coarser-scale dynamics. Modeling the entire system at the finest spatial and temporal resolutions needed is often both impractical and enormously inefficient. Adaptive mesh refinement (AMR) is an emerging technology that dynamically refines the grid during run-time so as to provide localized improvement in error norms and direct computational cost where it is needed [Papers 39, 42, 49]. Spatial mesh refinement refers to the addition of spatial degrees of freedom to specific areas of the mesh based on a *refinement criterion*. Temporal adaptivity provides the additional capacity to locally modify the time-step size based on the local resolution of the mesh, so as to ensure maximal performance while still satisfying the local Courant-Friedrichs-Lewy (CFL) condition. AMR offers the potential to better represent extreme weather events, such as tropical cyclones or atmospheric rivers, which are similarly associated with fine-scale features that require high model resolution. In ice sheet models, variable-resolution and AMR models have already shown their value due to high-resolution requirements to resolve the dynamics of grounding lines and ice streams, which only occupy very small portions of the continental-scale ice sheets but exert a controlling influence on the dynamics of the entire ice sheets. Research on AMR for global climate models is ongoing, with recent work highlighting its potential [12, 25, 38].

3.3.3 Future directions

Other concerns regarding time to solution consider methods that reclaim computational time due to latency via context switching. One option is to simultaneously compute two ensemble simulations and swap between ensembles when a barrier is reached. Another option is to employ parallel-in-time physics parameterizations and perform physics computations while the barrier is in place. Also, the need for better controls on solution accuracy in coupled systems was highlighted [Paper 53]. In several examples, the output of coupled models is no longer convergent with increasing spatial and temporal resolution. This is an issue that needs to be continually addressed as coupled system complexity increases, since a large part of the increased time to solution is due to the combination of increased model complexity (more sub-models coupled together) and increased model resolution. If the coupled system is not convergent, then it's unclear that the increased cost is actually improving the model results.

Currently much effort is spent tuning physical parameterizations in various models, which is time-consuming and potentially brittle in the context of dynamic changes in climate systems. At some of the breakouts, the idea of replacing physical parameterizations with black-box machine learning or data tools was discussed as a promising alternative that would allow models to tune themselves, using observational data or through deterministic or stochastic [Papers 17, 18, 24] parameterization/superparameterization/ultraparameterization output. This would both improve time to solution for these models and enhance confidence in model output [Papers 25, 52].

3.4 Data management and analysis

3.4.1 Grand challenge

Climate is unique in terms of the scale of production, retention, and throughput of simulation output. This is due to the size and diversity of the international community that consumes the output and the breadth of questions spanning broad time and space scales for which the data is tasked to answer. Operationally, the expense of each simulation is such that all data that might possibly be desired or required for post-processing is kept for insurance. As a result, data reduction decisions are particularly difficult. Yet as we approach exascale, hardware constraints will force some novel solutions including relative data reduction via in situ analysis, more extensive utilization of low-power profile co-processors, and checking for potential data corruption. The increasing complexity of models, while a challenge for model throughput, unlocks many opportunities for new avenues for analyzing model output in ways that will provide understanding. However, the expense and complexity of methods to target data from these models present their own challenges of performance.

Uncertainty quantification and model ensembles bring further challenges to analytics that exacerbate the need for scalability and require the development of new techniques. They also bring experimental design issues for optimal exploration of model parameter space so that minimal-size ensembles can deliver information needed for understanding of the parameter space and uncertainty.

3.4.2 Opportunities and potential solutions

Functional data analysis The development of functional data analysis (FDA) in statistical science [Paper 55] presents an opportunity to facilitate a gradual transition to continuum process representations in climate science while addressing data reduction. At a high level, FDA begins with data, converts it to nonparametric functions (splines, wavelets, Fourier, etc.), and remains as functions for further reduction, storage, analysis, and visualization. Addressing the transition to continuum challenge is not unlike the transition from mathematical tables to mathematical functions that took place in the late part of the past century. We trade storage for mathematical complexity. The same can be said about the transition from dense matrix to sparse matrix methods. This has several consequences:

- Reduced storage and data movement from simulation to storage.
- No need to store tables of numbers. Instead, specifications of functions are stored.
- Data analysis is performed in functional space with a smaller memory footprint than traditional methods. This is where the mathematical complexity comes in, but FDA tools already exist for common analytics such as variability attribution, principal components, canonical correlation, clustering, and many other multivariate methods as well as new techniques that are unique to functional data.
- Enables principled model component coupling at any resolution.
- Provides rigorous interpolation schemes for visualization.
- Allows a fallback to traditional methods because data can be reconstructed at any resolution.

Standard decompositions, such as principal components analysis (or empirical orthogonal functions), can be extended to tensor decompositions (also known as higher-order principal components analysis). Algorithms for these are already available [1] even in the functional space, but their application to climate data remains open. Function space sampling algorithms [Paper 22] take the functional approach to reduce computational complexity of Monte Carlo for inference and UQ by sampling functions instead of points [6].

A wealth of methodology for statistical analysis and wrangling of data is encapsulated in the R Environment for Statistical Computing [32] and its thousands of packages. This includes a variety of multivariate methods with many already available to operate in functional space. Recent developments enable R scalability on large systems [5, 35] with various in situ options, including sharing a communicator and data staging. These developments enable novel data analysis algorithm scripting powered by the same scalable mathematical and communication libraries that currently power many petascale simulation codes.

Additional strategies Other ideas were suggested to improve data processing and work flow, which would enhance efficiency outside of the time-to-solution envelope as well as model credibility, through reproducibility.

Clustering and empirical orthogonal functions for variability attribution are commonly used in analysis of climate data. Their ubiquity and their potential data reduction properties make them good candidates for in situ processing. Scalable algorithms for exascale architectures are still lacking, although some have been developed in the past [Paper 56] and more recently in [5, 35], the latter being able to utilize pluggable dense linear algebra co-processor libraries.

In situ analysis as well as fast surrogates [Paper 21] provide an external observer the opportunity to check for potential data corruption. Such data reductions rely on smoothness properties in time or in space to detect anomalies, some of which may be corruptions. As data analysis is based on an explicit or implicit model, it provides an opportunity for building data-based reduced models as smart proxies. This is discussed in [Paper 9] and is also a potential result of the models in all the other approaches discussed in this section. An overarching strategy to automate large portions of the climate model work flow system could shift important but expensive and mundane steps in the model development, analysis, and publication process [Paper 32].

3.4.3 Future directions

Implementing scalable in situ and post-processing analysis algorithms requires common frameworks for data transport and data sharing. Options for in situ analysis include running on the same resources as the simulation, sharing a communicator, or as a service on separate resources through in-memory or burst buffer staging. Establishing middleware standards that enable these choices with a common interface (such as ADIOS [31]) will enable the use of the same analysis software whether in situ or in post-processing.

Analytics challenges brought by ensembles and uncertainty quantification can be addressed by stronger connections to the statistics community, where uncertainty plays a fundamental role. Techniques that include statistical design of experiments can lead to optimal information content in ensembles and simplify analysis. Fundamental mathematics in analytics, such as dense and sparse linear algebra, is available in scalable libraries and continues to be developed by the applied mathematics community. Engaging this as infrastructure in easily reconfigurable ways is needed for scalable analytics. This has already started in the R language [35], but a lot more needs to be done to provide diverse scalable techniques from modern statistical science and machine learning for climate.

4 Structure of Collaboration

4.1 Existing tools and collaborations

Due to the computational demands of modeling the climate system with high fidelity, there is already a rich history of collaboration and communication between applied mathematicians, computer scientists, statisticians, and climate scientists. In fact, climate models have led the way in organizing successful collaborations between these disciplines in the service of improving the state of climate science. One example of such a collaboration was through the DOE-supported ISICLES effort. The IPCC AR5 report [39] called attention to the large uncertainties in predictions of sea level rise resulting from the inadequacy of then-current ice sheet models to the task of modeling the response of the Greenland and Antarctic ice sheets to projected climate forcing. As a result, DOE ASCR, as an expansion of an existing partnership with DOE BER, issued a unique call for ice sheet modeling efforts that would bridge the gap between computational expertise in ASCR and the near-term needs for improved ice sheet modeling efforts within BER. The result was six funded projects that covered a range of ice sheet modeling strategies ranging from Lagrangian particle models to spatial meshes that better capture ice fracture. While some of the projects were more successful than others, the net result was an infusion of applied mathematics and computer science expertise into the relatively young field of ice sheet modeling. This helped to propel the DOE portion of that effort forward to the point where the DOE is now a world leader in ice sheet model sophistication and maturity.

The joint ASCR and BER partnership in the Scientific Discovery through Advanced Computing (SciDAC) Program is a long-term investment from the DOE Office of Science that has propelled climate modeling to petascale and has a new call towards the exascale level of computing. It varies with each call but generally consists of centralized mathematics and computer science “Institutes” that support a number of “Applications Partnerships” in each of the sub-offices in the Office of Science. This model has proved an effective way to organize the deployment of advances into the various application fields. Climate modeling has benefited from this collaboration since the first SciDAC program, and the key driver for success has been the joint participation from applied mathematicians, computer scientists, and a variety of science domain experts with expertise in climate science.

More recently, the ACME Project has stood up to create a global coupled Earth system model that is designed to answer key science questions of interest to the DOE mission. It is achieving its science goals through effective use of DOE leadership-class computing via a significant focus on advanced performance, data management and workflow, and software engineering. Other international climate centers, including the work at the United Kindgom’s European Centre for Medium-Range Weather Forecasts (ECMWF) presented in a plenary talk by George Mozdzynski at the workshop, also recognize the importance of and invest in research to best utilize the largest available computing resources to achieve the best modeling results for weather and climate prediction.

4.2 Facilitation of communication across disciplines

A primary issue is communication across the various disciplines that all feed into a successful climate modeling enterprise. While it is tempting to think of collaboration in terms of tools, as in “which tools facilitate collaboration” or “how can we adapt tools to be of use to climate scientists,” in the breakout session discussions of effective communication, climate scientists and mathematics and computer science specialists invariably dwell on processes rather than concrete tools.

Embedding applied mathematics, statistics, and computer science experts in the domain science side has proven through the successes of SciDAC and in the instantiation of ACME to be a useful strategy for intellectual and technology transfer, particularly from the applied mathematics domain to the climate and computer science domains. For example, applied mathematics researchers in the ISICLES and SciDAC-funded PISCEES ice sheet projects have regularly contributed both to applied mathematics and glaciology-related conferences and journals. Embedding works because it creates a hybrid person who is conversant in both the specific demands of the application space along with the mathematical background to either apply directly or to be able to communicate back to others in the mathematical and computer science sides. The key challenge is for these hybrid scientists to be recognized and motivated in their efforts. Recognizing statistics as distinct from applied mathematics puts more emphasis on the increasing need for modern data analysis and data assimilation required for analysis going forward in addition to current and future simulation science’s use of PDE solution methodologies to advance model execution speed and accuracy.

The importance of meaningful (but tractable) benchmarks came up multiple times during AXICCS discussions. Besides the significant utility in building model confidence, benchmark problems are the way we communicate between fields, making it much easier to transfer information about model attributes and issues rather than simply “kicking something over the fence” back to the applied mathematics and computer science communities. A useful set of benchmark problems also provides a well-defined path to entry into a science domain for the embedded interloper from the mathematics side.

Discussion about colocation of applied mathematicians, computer scientists and climate scientists indicated a likely benefit, although the logistics of such an arrangement is complicated, especially when climate scientists themselves are also not colocated, as in the case of ACME. The biggest challenge is how to ensure continuous engagement among the climate science, mathematics and computer science communities—how to continue to find and encourage people who are effectively trained in climate science, mathematics, and computer science. Usually, job advertisements targeting the full range of postdoc to senior scientist career stages that prioritize personnel with combined mathematics and/or computer science along with Earth science knowledge do not generate many candidates.

4.3 Outreach and learning as potential solutions

Direct marketing to undergraduate and graduate mathematics and computer science students can be useful—the perception that climate science is one of the grand challenge scientific problems of our time is a strong motivator in many instances to get non-climate scientists involved. Workforce training efforts like the DOE-managed Computational and Stewardship Science Graduate Fellowships (CSGF, SSGF) and Office of Science Graduate Student Research (SCGSR) programs provide a useful mechanism for channeling graduate students in the right directions to be able to contribute meaningfully as early-career scientists. For example, the CSGF actively steers the courses of study of its fellows, ensuring solid backgrounds in mathematics and computer science while maintaining a focus on producing true computational scientists who are productive in their respective scientific fields. The newly formed Science Graduate Student Research (SCGSR) program within DOE involves graduate students being hosted at a national laboratory for a 3-month appointment. While this is not much time to develop skills, it does start the process for future development and interaction. Additional venues to train multidisciplinary scientists include topical workshops, webinars, and/or summer schools. A team composed of ACME scientists, named team "HACME," have attended the last two Oak Ridge Leadership Computing "Hackathons." These events connect domain scientists, mathematicians, computer scientists, and compiler vendors to learn how to optimize the performance of their code on the Oak Ridge Leadership Computing Facility (OLCF) machines.

One of the side benefits of the AXICCS workshop was a large number of side conversations between members of the different communities who otherwise would not have met. It is likely that at least some of these conversations will lead to productive lines of work. In the workshop, these informal gatherings became known as the "hot tub" idea—the more chances for people from the different fields to interact, the more chances for meaningful and substantive interaction to take place. While it may be hard to justify in a project-driven and end-result-driven funding ecosystem, such opportunities for cross-cutting contact and discussions stand to enormously benefit the state of climate science as a whole, which will then feed back into the applied mathematics and computer science disciplines. In the past, the annual SciDAC project meetings have provided such opportunities, although by necessity they were limited to those projects already in SciDAC collaborations and covered many subject areas beyond climate as the target area within which to form connections.

Networking opportunities for graduate students and postdoctoral scholars have also been developed through dual-purpose (education and workshop) events such as the Dynamical Core Model Intercomparison Project (DCMIP) summer school and workshop [18]. Every 4 years since 2008 this program has brought together students, postdoctoral researchers, application scientists, model developers, and other experts with the objective of (a) supporting a massive learning initiative directed at dynamical core research and (b) leveraging an enthusiastic cohort of participants to undertake the model intercomparison. The outcome of this intercomparison has been praised by both participants and modeling groups and featured wide support from major federal agencies. The event in 2016 was sponsored by the National Center for Atmospheric Research's (NCAR) Computational Information Systems Laboratory (CISL), NOAA, NASA, DOE, the National Science Foundation (NSF), and the Office of Naval Research (ONR). Similar efforts could likewise be developed for other model intercomparisons that require mathematics and computer science experts to bring together a cohort of experts and students, for instance, one specifically focused on coupling of model components. However, workshops such as these rely on the target model component(s) to have a sufficiently diverse and broad community to make intercomparison worthwhile. Further, such an effort also relies on community agreement on a "standardized" suite of benchmarks to be employed. Nonetheless, projects such as these provide an excellent bridge between laboratories and academia, with positive outcomes for both education and productivity.

4.4 Possible follow-on activities

The discussions at the workshop have generated a collection of big ideas and cross-cutting issues that climate scientists, mathematicians, and computer scientists may want to pursue collaboratively in order to advance climate modeling and simulation. It is important to keep the momentum going so that the big ideas can be realized. In addition, it will be beneficial to engage additional researchers from the three communities beyond this workshop group. This allows the big ideas and cross-cutting issues identified in this report to be fleshed out by bigger groups, and possibly motivating additional ideas and issues. For example, follow-up workshops can be organized to focus on each of the big ideas. Another possibility is to have special journal articles and/or issues calling for research in such areas.

One step this group has taken is to organize a minisymposium at the upcoming SIAM Conference on Computational Science and Engineering, which will be held at the end of February 2017. Several members of the program committee of this workshop will present the findings in this report at the minisymposium. As alluded to earlier, this is not meant to signal the end of the workshop. Rather, it should be considered the beginning of a larger and broader effort.

5 Conclusions

The Advances in Mathematical and Computational Climate Modeling (AXICCS) workshop was different from many workshops that focus on scientific discussions and brainstorming of new research directions. It was unique because the format was designed so that the participants focused on identifying future areas of research that would materially impact climate simulation without restrictions of time, funding, or expertise. A key attribute was to discuss, in tandem, the key mathematical and computer science advances that would be required to pursue new ideas we identified. By allowing discussions across many scientific domains and among colleagues at different stages of their career, the goal was to provide the opportunity for many voices and ideas to be heard and discussed.

Although many new science ideas were considered, we developed several major findings about how to improve Earth system models into more formal strategies within Section 2. These spanned improvements in managing model complexity, strategies to maintain and improve confidence in these complex and interacting models, and the coupling of processes to enable a better understanding of our rich and diverse Earth system and its changes. We also presented a novel path to provide a simulation of the Earth without arbitrary boundaries fostered by community tradition.

We prioritized key tools from the mathematics and computer science communities we thought would be most beneficial to bring these ideas to a more mature state of evaluation and implementation. These included computer science methods such as language and compiler tools to address upcoming and dramatic hardware changes, mathematical tools for model development that will target the new hardware effectively to maximize space and timescale limitations, strategies to treat and minimize software complexity so that climate scientists and non-climate scientists can develop and evaluate code use to produce these large and interacting models, and data analysis and management, which address statistical tools to unlock the new insights within simulation output as well as new methods for data storage and workflow to enable the more complex models to be managed. These have been outlined in Section 3.

Finally, in Section 4, we discussed the possible path forward. Additional opportunities, through more interactions with these and additional climate scientists, would flesh out these ideas and also motivate additional ideas. Next steps would include workshops to generate expansions and specifics of each idea, additional and incentivized interactions with colleagues from the climate, mathematics, and computer science communities beyond the limited attendance of this workshop, and more communication of these groups through traditional and novel methods, for example a special journal article and issue calling for research in these areas.

Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Offices of Biological and Environmental Research (BER) and Advanced Scientific Computing and Research (ASCR). The workshop organizers thank Drs. Dorothy Koch from BER and Randall Laviolette from ASCR for their encouragement and support. Logistics were handled with excellent care by Andrew Flatness and Tracey Vieser. The organizers are grateful to the participants of the workshop, especially the program committee and those who served as plenary speakers, poster and oral presenters, note-takers, and engaged participants in the discussion. The organizers are also grateful to Deborah Stevens, Kathy Jones, and Madeleine Evans for editing the report.

References

[1] G. Allen. Sparse higher-order principal components analysis. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, volume 22, page 27:36, 2012.

[2] K. Asanovi and D. A. Patterson. Instruction sets should be free: The case for RISC-V. Technical Report UCB/EECS-2014-146, EECS Department, University of California, Berkeley, Aug 2014. URL `http://www.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-146.html`.

[3] C. S. Bretherton. Frontiers in multiscale and global simulation of boundary layer clouds, 2016. Presentation at the Workshop on Advancing X-cutting Ideas for Computational Climate Science (AXICCS), Monday, September 12, 2016.

[4] R. Cess, G. L. Potter, J. P. Blanche, G. J. Boer, S. J. Ghan, J. T. Kiehl, H. L. Treut, Z.-X. Li, X.-Z. Liang, J. Mitchell, J.-J. Morcrette, D. A. Randall, M. R. Riches, E. Roeckner, U. Schlese, A. Slingo, K. E. Taylor, W. M. Washington, R. T. Wetherald, and I. Yagai. Interpretation of cloud–climate feedback as produced by 14 atmospheric general circulation models. *Science*, 245(4917):513–516, 1989.

[5] W.-C. Chen, G. Ostrouchov, D. Pugmire, Prabhat, and M. Wehner. A parallel EM algorithm for model-based clustering applied to the exploration of large spatio-temporal data. *Technometrics*, 55(4):513–523, 2013. doi:10.1080/00401706.2013.826146.

[6] S. L. Cotter, G. O. Roberts, A. M. Stuart, D. White, et al. MCMC methods for functions: Modifying old algorithms to make them faster. *Statistical Science*, 28(3):424–446, 2013.

[7] P. M. Cox, D. Pearson, B. B. Booth, P. Friedlingstein, C. Huntingford, C. D. Jones, and C. M. Luke. Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, 494(7437):341–344, Feb. 2013. doi:10.1038/nature11882.

[8] M. Deneroff. Building an SOC: How to do it? What will it cost? In *Workshop on System-on-a-chip design for HPC*, 2014.

[9] R. Dennard, F. Gaensslen, H. Yu, V. Rideout, E. Bassous, and A. Leblanc. Design of ion-implanted MOSFETs with very small physical dimensions. *IEEE Journal of Solid-State Circuits*, 9(5):256–268, 1974.

[10] J. Dongarra. Report on the Sunway TaihuLight system. Technical Report UT-EECS-16-742, University of Tennessee, 2016.

[11] D. Donofrio, L. Oliker, J. Shalf, M. F. Wehner, C. Rowen, J. Krueger, S. Kamil, and M. Mohiyuddin. Energy-efficient computing for extreme-scale science. *IEEE Computer*, 42(11):62–71, 2009. doi:10.1109/MC.2009.353.

[12] J. O. Ferguson, C. Jablonowski, H. Johansen, P. McCorquodale, P. Colella, and P. A. Ullrich. Analyzing the Adaptive Mesh Refinement (AMR) characteristics of a high-order 2D cubed-sphere shallow-water model. *Monthly Weather Review*, 2016. In press.

[13] H. Fu, J. Liao, J. Yang, L. Wang, Z. Song, X. Huang, C. Yang, W. Xue, F. Liu, F. Qiao, W. Zhao, X. Yin, C. Hou, C. Zhang, W. Ge, J. Zhang, Y. Wang, C. Zhou, and G. Yang. The Sunway TaihuLight supercomputer: System and applications. *Science China Information Sciences*, 59(7):072001, 2016. ISSN 1869-1919. doi:10.1007/s11432-016-5588-7. URL `http://dx.doi.org/10.1007/s11432-016-5588-7`.

[14] A. Hall and X. Qu. Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys. Res. Lett.*, 33(3):L03502, Feb. 2006. doi:10.1029/2005GL025127.

[15] F. M. Hoffman, J. T. Randerson, V. K. Arora, Q. Bao, P. Cadule, D. Ji, C. D. Jones, M. Kawamiya, S. Khatiwala, K. Lindsay, A. Obata, E. Shevliakova, K. D. Six, J. F. Tjiputra, E. M. Volodin, and T. Wu. Causes and implications of persistent atmospheric carbon dioxide biases in Earth system models. *J. Geophys. Res. Biogeosci.*, 119(2):141–162, Feb. 2014. doi:10.1002/2013JG002381.

[16] F. Hourdin, T. Mauritsen, A. Gettelman, J.-C. Golaz, V. Balaji, Q. Duan, D. Folini, D. Ji, D. Klocke, Y. Qian, F. Rauser, C. Rio, L. Tomassini, M. Watanabe, and D. Williamson. The art and science of climate model tuning. *Bull. Amer, Meteor. Soc.*, 2017. doi:10.1175/BAMS-D-15-00135.1. URL `http://dx.doi.org/10.1175/BAMS-D-15-00135.1`.

[17] J.-F. Lemieux, S. F. Price, K. J. Evans, D. A. Knoll, A. G. Salinger, and T. Payne. Implementation of the Jacobian-free Newton-Krylov method for solving the first-order ice sheet momentum balance. *J. Comp. Phys.*, 230:6531–6545, 2011. doi:10.1016/j.jcp.2011.04.037.

[18] C. Jablonowski, P. A. Ullrich, J. Kent, K. A. Reed, M. A. Taylor, P. H. Lauritzen, and R. D. Nair. The 2012 dynamical core model intercomparison project (DCMIP). In *AGU Fall Meeting Abstracts, A53C-0160*, 2012. URL `http://fallmeeting.agu.org/2012/files/2012/12/AGU_poster_Jablonowski_Dec_2012.pdf`. (Accessed 3 July 2016).

[19] M. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001. doi:10.1111/1467-9868.00294.

[20] D. E. Keyes, L. C. McInnes, C. Woodward, W. Gropp, E. Myra, M. Pernice, J. Bell, J. Brown, A. Clo, J. Connors, E. Constantinescu, D. Estep, K. Evans, C. Farhat, A. Hakim, G. Hammond, G. Hansen, J. Hill, T. Isaac, X. Jiao, K. Jordan, D. Kaushik, E. Kaxiras, A. Koniges, K. Lee, A. Lott, Q. Lu, J. Magerlein, R. Maxwell, M. McCourt, M. Mehl, R. Pawlowski, A. P. Randles, D. Reynolds, B. Rivire, U. Rde, T. Scheibe, J. Shadid, B. Sheehan, M. Shephard, A. Siegel, B. Smith, X. Tang, C. Wilson, and B. Wohlmuth. Multiphysics simulations: Challenges and opportunities. *Int. J. High Perf. Comp. Apps.*, 27(1):4–83, 2013. doi:10.1177/1094342012468181.

[21] M. Khairoutdinov, D. Randall, and C. DeMott. Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *Journal of the Atmospheric Sciences*, 62(7):2136–2154, 2005.

[22] M. F. Khairoutdinov and D. A. Randall. Cloud resolving modeling of the arm summer 1997 iop: Model formulation, results, uncertainties, and sensitivities. *Journal of the Atmospheric Sciences*, 60(4):607–625, 2003.

[23] J. Krueger, D. Donofrio, J. Shalf, M. Mohiyuddin, S. Williams, L. Oliker, and F.-J. Pfreund. Hardware/software co-design for energy-efficient seismic modeling. In *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '11, pages 73:1–73:12, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0771-0. doi:10.1145/2063384.2063482. URL `http://doi.acm.org/10.1145/2063384.2063482`.

[24] R. M. Maxwell and L. E. Condon. Connections between groundwater flow and transpiration partitioning. *Science*, 353(6297):377–380, 2016.

[25] P. McCorquodale, P. Ullrich, H. Johansen, and P. Colella. An adaptive multiblock high-order finite-volume method for solving the shallow-water equations on the sphere. *Communications in Applied Mathematics and Computational Science*, 10(2):121–162, 2015.

[26] G. E. Moore. Cramming more components onto integrated circuits. *Electronics*, 38(8), April 1965.

[27] National Research Council. *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*. National Academies Press, May 2012. doi:10.17226/13395. URL `https://www.nap.edu/download/13395`.

[28] NRC-Committee. *A National Strategy for Advancing Climate Modeling*. The National Academies Press, 2012. ISBN 9780309259774. URL `http://www.nap.edu/openbook.php?record_id=13430`. Committee on a National Strategy for Advancing Climate Modeling; Board on Atmospheric Studies and Climate; Division on Earth and Life Studies.

[29] H. Parishani. Explicit representation of low clouds in a global atmospheric model: initial results from Ultra-Parameterized CAM5. In *2015 AGU Fall Meeting*. AGU, 2015.

[30] D. A. Patterson. Latency lags bandwith. *Commun. ACM*, 47(10):71–75, Oct. 2004. ISSN 0001-0782. doi:10.1145/1022594.1022596. URL `http://doi.acm.org/10.1145/1022594.1022596`.

[31] N. Podhorszki, Q. Liu, J. Logan, J. Mu, H. Abbasi, J.-Y. Choi, and S. A. Klasky. Adios 1.9 user's manual. Technical report, ORNL/TM-2009/100, 2015.

[32] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL `https://www.R-project.org`.

[33] D. M. Romps. The dry-entropy budget of a moist atmosphere. *Journal of the Atmospheric Sciences*, 65(12):3779–3799, 2008.

[34] M. Satoh, T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S.-I. Iga. Nonhydrostatic icosahedral atmospheric model (nicam) for global cloud resolving simulations. *Journal of Computational Physics*, 227(7):3486–3514, 2008.

[35] D. Schmidt, W.-C. Chen, M. A. Matheson, and G. Ostrouchov. Programming with BIG data in R: Scaling analytics from one to thousands of nodes. *Big Data Research*, 2016. URL `http://dx.doi.org/10.1016/j.bdr.2016.10.002`.

[36] D. E. Shaw, M. M. Deneroff, R. O. Dror, J. S. Kuskin, R. H. Larson, J. K. Salmon, C. Young, B. Batson, K. J. Bowers, J. C. Chao, M. P. Eastwood, J. Gagliardo, J. P. Grossman, C. R. Ho, D. J. Ierardi, I. Kolossváry, J. L. Klepeis, T. Layman, C. McLeavey, M. A. Moraes, R. Mueller, E. C. Priest, Y. Shan, J. Spengler, M. Theobald, B. Towles, and S. C. Wang. Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM*, 51(7): 91–97, July 2008. ISSN 0001-0782. doi:10.1145/1364782.1364802. URL `http://doi.acm.org/10.1145/1364782.1364802`.

[37] B. J. Soden, D. L. Jackson, V. Ramaswamy, M. D. Schwarzkopf, and X. Huang. The radiative signature of upper tropospheric moistening. *Science*, 310(5749):841–844, 2005. doi:10.1126/science.1115602.

[38] A. St-Cyr, C. Jablonowski, J. M. Dennis, H. M. Tufo, and S. J. Thomas. A comparison of two shallow-water models with nonconforming adaptive grids. *Mon. Wea. Rev.*, 136(6):1898–1922, 2008.

[39] T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, editors. *Summary for Policymakers*, book section SPM, page 130. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013. ISBN ISBN 978-1-107-66182-0. doi:10.1017/CBO9781107415324.004. URL www.climatechange2013.org.

[40] M. F. Wehner, L. Oliker, J. Shalf, D. Donofrio, L. A. Drummond, R. Heikes, S. Kamil, C. Kono, N. Miller, H. Miura, M. Mohiyuddin, D. Randall, and W.-S. Yang. Hardware/software co-design of global cloud system resolving models. *Journal of Advances in Modeling Earth Systems*, 3(4), 2011. ISSN 1942-2466. doi:10.1029/2011MS000073. URL http://dx.doi.org/10.1029/2011MS000073. M10003.

[41] H. Zhu, H. Hendon, and C. Jakob. Convection in a parameterized and superparameterized model and its role in the representation of the MJO. *Journal of the Atmospheric Sciences*, 66 (9):2796–2811, 2009.

Acronyms used in the report

ACME: Accelerated Climate Model for Energy
AMR: Adaptive Mesh Refinement
ASCR: Advanced Scientific Computing Research
BER: Biological and Environmental Research
CS: Computer Science
CESM: Community Earth System Model
CISL: Computational Information Systems Laboratory
CMIP: Coupled Model Intercomparison Project
CMOS: Complimentary metal oxide semiconductor
CRM: Cloud-resolving Model
CSGF: Computational Science Graduate Fellowship
DA: Data Assimilation
DOE: U.S. Department of Energy
DSL: Domain-Specific Language
ECMWF: European Centre for Medium-Range Weather Forecasts
ENSO: El Niño Southern Oscillation
ESM: Earth System Model
FDA: Functional Data Analysis
GCM: Global Climate Model
IMEX: Implicit-Explicit
IPCC: Intergovernmental Panel on Climate Change
ISCCP: International Satellite Cloud Climatology Project
ISICLES: Ice Sheet Initiative for CLimate ExtremeS
LCF: Leadership-Class Facility
LES: Large Eddy Simulation
MJO: Madden-Julian Oscillation
MLMC: Multilevel Monte Carlo
NCAR: National Center for Atmospheric Research
NASA: National Aeronautics and Space Administration
NOAA: National Oceanographic and Atmosphere Administration
NSF: National Science Foundation
ONR: Office of Naval Research
PDE: Partial Differential Equation
PISCEES: Predicting Ice Sheet and Climate Evolution at Extreme Scales
SciDAC: Scientific Discovery through Advanced Computing
SCGSR: Office of Science Graduate Student Research
SSGF: Stewardship Science Graduate Fellowship
UQ: Uncertainty Quantification

Workshop Organizers

Katherine J. Evans, Oak Ridge National Laboratory
Esmond Ng, Lawrence Berkeley National Laboratory

Program Committee

Peter Caldwell, Lawrence Livermore National Laboratory
Forrest M. Hoffman, Oak Ridge National Laboratory
Charles Jackson, University of Texas, Austin
Kerstin Klees Van Daam, Brookhaven National Laboratory
L. Ruby Leung, Pacific Northwest National Laboratory
Dan Martin, Lawrence Berkeley National Laboratory
George Ostrouchov, Oak Ridge National Laboratory
Ray Tuminaro, Sandia National Laboratories
Paul Ullrich, University of California, Davis
Stefan Wild, Argonne National Laboratory
Samuel Williams, Lawrence Berkeley National Laboratory

Plenary Speakers

Chris Bretherton, University of Washington
William Collins, Lawrence Berkeley National Laboratory
Petros Koumoutsakos, ETH Zurich
George Mozdzynski, European Center for Medium Range Forecasts

Workshop Participants

Melissa Allen, Oak Ridge National Laboratory
Pavel Bochev, Sandia National Laboratories
Jed Brown, University of Colorado, Boulder
Peter Caldwell, Lawrence Livermore National Laboratory
Donna Calhoun, Boise State University
Nathan Collier, Oak Ridge National Laboratory
Emil Constantinescu, Argonne National Laboratory
Bert Debusschere, Sandia National Laboratories
George Delic, HiPERiSM Consulting, LLC
Pedro Diniz, University of Southern California
David Donofrio, Lawrence Berkeley National Laboratory
David Gardner, Lawrence Livermore National Laboratory
David Hall, University of Colorado Boulder
Matthew Hoffman, Los Alamos National Laboratory
Christiane Jablonowski, University of Michigan
Robert Jacob, Argonne National Laboratory
Philip Jones, Los Alamos National Laboratory

Joseph Kennedy, Oak Ridge National Laboratory
Andreas Kreienbuehl, Lawrence Berkeley National Laboratory
Kody Law ,Oak Ridge National Laboratory
Li Li, Penn State University
Yangang Liu , Brookhaven National Laboratory
Fei Lu, UC Berkeley/Lawrence Berkeley National Laboratory
Vijay Mahadevan, Argonne National Laboratory
Salil Mahajan, Oak Ridge National Laboratory
Benjamin Mayer, Oak Ridge National Laboratory
Richard T. Mills, Intel
Sri Hari Krishna Narayanan, Argonne National Laboratory
Matthew Norman, Oak Ridge National Laboratory
George Ostrouchov, Oak Ridge National Laboratory
Maruro Perego, Sandia National Laboratories
Tom Peterka, Argonne National Laboratory
Stan Posey, Nvidia
Stephen Price, Los Alamos National Laboratory
David Randall, Colorado State University
Jaideep Ray, Sandia National Laboratories
Andy Salinger, Sandia National Laboratories
Bill Spotz, Sandia National Laboratories
Georg Stadler, Courant Institute, New York University
Irina Tezaur, Sandia National Laboratories
Raymond Tuminaro, Sandia National Laboratories
Paul Ullrich, University of California, Davis
Nathan Urban, Los Alamos National Laboratory
Haim Waisman, Columbia University
Hui Wan, Pacific Northwest National Laboratory
Dali Wang, Oak Ridge National Laboratory
Shanlin Wang, Los Alamos National Laboratory
Carol Woodward, Lawrence Livermore National Laboratory

Workshop Agenda

Workshop on Advancing X-cutting Ideas for Computational Climate Science (AXICCS)

September 12-14, 2016

Hilton Rockville, 1750 Rockville Pike, Rockville, MD 20852

Agenda

Monday, September 12, 2016

Start Time	***End Time***	***Topic***
9:00AM	9:30AM	Welcome
9:30AM	10:30AM	Plenary: **Bill Collins**, Lawrence Berkeley National Laboratory *Climate Simulation at Impactful Scales: Charge for a New Physics Paradigm*
10:30AM	11:00AM	Coffee Break
11:00AM	12:30PM	Breakout #1

Breakout Topic 1A: Climate Science Problems in Coupling

Moderators: ***Peter Caldwell and Forrest Hoffman***

Speakers: 40 minutes total

Discussion: 50 minutes

L. Li, Y. Shi, C. Duffy. *Building Computational Bridges Across the Water, Ecosystem, and Soil Biogeochemistry Disciplines*

R. Mills and F. Hoffman. *Machine-learning guided, multi-resolution approaches to high-fidelity representation of global hydrology in ESMs*

H. Waisman, J. Bassis, S. Price, R. Tuminaro and I. Tezaur. *A physics based iceberg calving model coupled with a global ice-sheet flow model for accurate assessment of sea level rise*

M. Hoffman, L. Bertagna, M. Gunzburger, M. Perego and Stephen Price. *Realistic Subglacial Hydrology For Improved Ice Sheet-Climate Coupling and Sea Level Prediction*

Breakout Topic 1B: Climate Model Complexity and Scaling

Moderators: ***Ruby Leung and Paul Ullrich***

Speakers: 45 minutes total

Discussion: 45 minutes

M. Allen, M. Branstetter, O. Omitaomu. *Embedded Urban Framework for ACME Regions of Refined Resolution*

P. Bochev, K. Evans, M. Gunzburger and K. Peterson. *Optimization-Based Heterogeneous Numerical Methods: an Abstraction for Mathematically Rigorous Coupling of Earth System Models*

C. Jablonowski, J. Ferguson, H. Johansen and P. Colella. *Transforming Climate Modeling via Scale-Adaptive Computational Techniques*

W. Maslowski, A. Roberts, E. Hunke, F. Giraldo and M. Kopera. *Sea Ice Modeling Across Scales at Exascale and Beyond*

Breakout Topic 1C: Climate Model Ensembles and Uncertainty Quantification

Moderators: ***Charles Jackson and Michael Prather***

Speakers: 40 minutes total

Discussion: 50 minutes

A. Salinger, E. Phipps and J. Fyke. *Embedded Ensembles*
S. Mahajan, K. Evans and M. Norman. *Expanding the Utility of High-Resolution Global Climate Models via Short Ensembles*
S. Price, M. Perego and G. Stadler. *Optimization and Uncertainty Quantification of Ice Sheet Models*
S. Wang, N. Urban, M. Maltrud and Alexandra Jonko. *Automation of parameterization and structure selection of ocean biogeochemical models*

12:30PM	2:00PM	Lunch
2:00PM	2:30PM	Outbriefs from Breakout #1 (all)
2:30PM	3:30PM	Plenary: **Christopher S. Bretherton**, University of Washington *Frontiers in Multiscale and Global Simulation of Boundary Layer Clouds and Their Interactions with Climate*
3:30PM	4:00PM	Coffee Break
4:00PM	5:30PM	Breakout #2: Math and Computer Science Advances

Breakout Topic 2A: Coupling, PDEs, and linear algebra
Moderators: ***Ray Tuminaro and Dan Martin***
Speakers: 45 minutes total
Discussion: 45 minutes
M. Perego, S. Price and A. Salinger. *Next generation implicit solvers and analysis algorithms for ice sheet modeling*
J. Brown. *Higher Standards on the Control of Numerical Accuracy*
M. Norman. *New Temporal and Spatial Algorithms for Atmospheric Climate Models*

Breakout Topic 2B: Optimization and Statistics
Moderators: ***Stefan Wild and George Ostrouchov***
Speakers: 40 minutes total
Discussion: 50 minutes
O. Ghattas and **G. Stadler**. *From Data through Inference to Optimization under Uncertainty: Towards End-to-End Climate Model-Based Decision-Making*
J. Ray, L. Swiler, G. Pau, G. Bisht, F. Hoffman, M. Huang, Z. Hou and X. Chen. *Improving predictive capability of land surface models through robust statistical calibration techniques*
N. Urban. *Climate Model Uncertainty Quantification*

Breakout Topic 2C: Computational Performance and Data Management
Moderators: ***Sam Williams and Kerstin Kleese Van Damm***
Speakers: 40 minutes total
Discussion: 50 minutes
P. Ullrich, G. Jost, B. Lelbach and H. Johansen. *Exascale-Ready Programming Models for Climate*
D. Wang, O. Hernandez, G. Lopez and F. Winkler. *Compiler-based software analysis toolkit for climate model development*

5:30PM	7:00PM	Poster Session

Workshop Agenda

Workshop on Advancing X-cutting Ideas for Computational Climate Science (AXICCS)

September 12-14, 2016

Hilton Rockville, 1750 Rockville Pike, Rockville, MD 20852

Agenda

Tuesday, September 13, 2016

Start Time	*End Time*	*Topic*
9:00AM	9:30AM	Outbriefs from Breakout #2 (all)
9:30AM	10:30AM	Plenary: **Petros Koumoutsakos**, ETH Zurich, Switzerland *The Art of Computational Science: Closing Gaps, Forming Alloys*
10:30AM	11:00AM	Coffee Break
11:00AM	12:30PM	Breakout #3: Climate Response to Math and CS Ideas

Breakout Topic 3A: Same as 1A

Breakout Topic 3B: Same as 1B

Breakout Topic 3C: Same as 1C

12:30PM	2:00PM	Lunch
2:00PM	2:30PM	Outbriefs from Breakout #3 (all)
2:30PM	3:30PM	Plenary: **George Mozdzynski**, European Centre for Medium-Range Weather Forecasts, UK *Addressing Future Scalability and Power Challenges at the European Centre for Medium-Range Weather Forecasts (ECMWF)*
3:30PM	4:00PM	Coffee Break
4:00PM	4:30PM	Wrap-up

Workshop Agenda

Workshop on Advancing X-cutting Ideas for Computational Climate Science (AXICCS)

September 12-14, 2016

Hilton Rockville, 1750 Rockville Pike, Rockville, MD 20852

Agenda

Wednesday, September 14, 2016

Start Time	***End Time***	***Topic***
8:30AM	1:00PM	PC Only Report Writing

A version of this report with the white papers included in the appendix are available online at: https://science.energy.gov/ber/community-resources/