NSF Workshop on High-Performance Distributed Computing and Polar Sciences

Rutgers University New Brunswick, New Jersey, 4–5 December 2014

Iceberg by Ilulissat, Greenland (courtesy: Asa Rennermalm)

Report Editors:

Shantenu Jha (Rutgers), Lynn Yarmey (NSIDC)

Report Contributors:

Andrew Carleton (PSU), Anna Liljedahl (UAF), Jarek Nabrzyski (Notre Dame), Toni Rosati (NSIDC), Asa Rennermalm (Rutgers), Jane Wyngaard (NASA)

Graphics: Kristina Davis (Notre Dame)

Supraglacial stream on the Greenland ice sheet (courtesy: Asa Rennermalm)

Executive Summary

Climate change in the 20th and 21st century is dramatically changing the polar regions. This is documented by numerous studies, for example as thawing permafrost, retreating Arctic sea ice and accelerating mass loss from glaciers and ice sheets. These changes may have widespread consequences for many aspects of the earth systems, e.g. carbon budget, food and water security, sea levels, and freshwater input to oceans. To understand the changing polar regions and their global impacts, scientists are increasingly using very large datasets derived from high-resolution satellite imagery, airborne missions, and computer modeling. However, advanced cyberinfrastructure, and in particular, high-performance distributed computing (HPDC) remains an underutilized resource within the polar science community.

To explore the opportunities for addressing this gap and increasing the collaboration between the polar science and HPDC communities, the workshop "High-Performance & Distributed Computing for Polar Sciences: Applications, Cyberinfrastructure and Opportunities" brought together polar scientists, HPDC experts, and data practitioners at Rutgers University in New Brunswick, New Jersey on December 4 and 5, 2014. Approximately thirty U.S.-based researchers gathered for two days of presentations and discussions centered on two questions: 1) How can polar science benefit from HPDC? and 2) What are the challenges in bringing HPDC and polar sciences together?

Through workshop discussions, participants agreed that processing the ever-expanding catalog of high-resolution digital satellite imagery, and running model simulations of polar region dynamics, provide key opportunities for polar science and HPDC collaboration to advance both fields. Despite the potential of these opportunities, a number of challenges currently exist preventing progress. Some example obstacles to collaboration are the knowledge gap, simple access mechanisms to HPDC resources and lower barriers to access HPDC.

Workshop participants discussed many ways to close this gap, inter alia including how to increase data discovery and make connections between data repositories with computing and data processing facilities. Articulating and addressing the heterogeneity of HPDC solutions, whilst improving the simplicity of HPDC resource use (and understanding) were recurring themes. Greater adoption of HPDC might be facilitated by making software products commonly used among polar scientists available on HPDC platforms. Additionally, there are socio-technical and cultural barriers that need addressing.

Participants found the workshop, with adequate time for discussions, very educational and helpful, and there was unanimous consensus that such efforts needed to be sustained in order to understand how to convert aspirations into a plan and subsequent action. Recommendations from the Workshop include the following:

- Continue cross-community engagement to build common directions
- Promote awareness of HPDC training resources for polar scientists
- Work towards a roadmap for HPDC uptake in the polar sciences

1 Introduction and Context

Climate change in the 20th and 21st century is dramatically affecting the polar regions and carries additional risk for global physical, social and economic systems. Recent studies indicate accelerating permafrost degradation, Arctic sea ice retreat, and mass loss from ice sheets (Greenland and Antarctic) and mountain glaciers, among other changes. Understanding the changing polar regions and connections to global climate requires a combination of field data, high-resolution observations from satellites, airborne imagery, and computer model outputs. Computational approaches have the potential to support faster and more fine-grained integration and analysis of these and other data types, thus increasing the efficiency of analysing and understanding the complex processes that are rapidly changing our climate.

Despite these data- and compute-intensive scientific needs, **polar science is inadequately represented, both in practice and in perception, in its use of high-performance and distributed cyberinfrastructure (HPDC).** To address this issue, an NSF funded workshop was organized to bring together the polar sciences community with High-Performance & Distributed Computing (HPDC) community to collectively explore, examine and identify the opportunities and barriers in the use of HPDC tools, techniques and resources to solve polar science research questions.

To establish that there are multiple polar science problems that could benefit from HPDC but are currently not to the fullest extent or potential, we highlight three important problems that were discussed at the different venues in the workshop.

- **Quantifying ice sheet mass loss.** Mass loss from Greenland and Antarctic ice sheets contributes large amounts of freshwater to global oceans, which may affect large scale ocean circulation patterns and global sea level rise. Both of which are of huge global socio-economic concern. While it is clear that ice sheets is one of the major drivers for current sea level rise, future impacts cannot be fully resolved until dynamic mass loss at ice sheet marine terminating outlet glaciers are better understood. Dynamic mass losses are estimated by models capable of simulating ice flow and ice thickness over time. These model come in varying degrees of complexity, where the most complex models have very large computational demands. To unlock the full potential of these models advanced and scalable cyberinfrastructure is needed. This includes integrated data and computational capabilities, which would allow huge increase in the number of ensemble runs and sensitivity studies that can be done.
- **Analysis of high-resolution satellite remote sensing imagery**. Unprecedented volume of high-resolution commercial satellite remote sensing imagery is being collected over the polar regions. This type of data may revolutionize many fields as surface properties over land, ice, ocean can be resolved with greater detail including: geomorphology, surface expression of permafrost thaw and erosion, vegetation change, population change among animals (e.g. penguins), and development of meltwater ponds

and streams on ice sheets, glaciers and sea ice. Before this imagery can be used for scientific application, several computationally intensive pre-processing steps are required (e.g. ortorecification, image mosaicing, creating elevation models from stereo pair imagery). Once imagery is available for scientific use, all applications on larger area and longer time periods will require HPDC.

- The last problem is a classic "Grand Challenge" for polar and computer scientists, viz., developing timely weather forecasts at sufficiently high spatial and temporal resolutions for the Arctic region. Improved weather forecast models would help both characterize and predict extreme events. Many Arctic-region weather and climate hazards occur on meso-scales (between approximately 50-800 km horizontal lengths, 1-12 h duration), and include an array of locally strong wind phenomena (tip, gap, barrier, katabatic, föhn, low-level jets) and also intense cyclonic storms called "polar lows". The latter, in particular, are most likely to increase in frequency and intensity over the Arctic Ocean and adjacent marginal seas as a result of the sea ice loss. However, meso-scale weather phenomena are notoriously difficult to forecast using numerical weather prediction models: they are poorly—if at all—represented in the sparse observation networks and current atmospheric reanalysis products, which are often used to drive global-scale and nested regional forecast models. At the same time, global models are needed because of the thermal, moisture, and dynamical links between the upstream lower-latitude atmosphere and the Arctic, and which could become more extreme and/or frequent—at least, in certain longitude sectors—as a result of ongoing climate changes. This challenge of improving Arctic-region weather forecasts is tailor-made for HPDC, because current meso-scale models (e.g., the Polar WRF—Weather Research and Forecasting model) can be configured to cover certain sub-regions (e.g., Greenland, northern North Atlantic) but not yet the whole Arctic Ocean and adjacent areas, owing to computational limitations. An alternative—or complementary—approach is to dynamically downscale the output from larger-scale models to particular sub-regions of the Arctic (e.g., Beaufort Sea, Kara and Barents seas), for generating reliable local weather forecasts of, for example, extreme wave heights, icing events, and the gales and hail/snow squalls accompanying polar lows.

The "Grand Challenge" problem, as well as the representative examples, reiterates the potential impact and need for HPDC in the polar sciences. As we will see in Section 2, the current state-of-the-practice of HPDC is inadequate to meet the scientific objectives that the polar community has set out to do. For example, in lieu of large-scale data analytics and scalable computing techniques, existing computational approaches are characterized by local, non-scalable and non-extensible solutions, as well as inefficient data transfer and management techniques. The three representative problems as well as the specific examples in Section 2, clearly articulate the vision of advanced and balanced cyberinfrastructure. In order to transition that vision into reality, many existing gaps must be overcome, as well as a credible and realistic path to that vision established. The NSF funded workshop at Rutgers was organized with this objective.

Before we discuss the structure of the report, we will speak to the intended audience of this report. The three primary groups of people without any sense of prioritization are: (i) The polar scientists and polar community, for whom we hope this document will provide a representative if not motivational set of exemplars of science that seeks to benefit from HPDC. This report will also hopefully provide a condensed but useful reference to the different types of resources -- hardware, software and people/communities, that they can turn to. This is an important objective of this report, as our experience as well as that of workshop participants advised the authors/editors of this report from (incorrectly) "assuming that polar scientists readers of this report will "just" know about XSEDE and/or OSG". We have taken that advice to heart and will grow Section 3. (ii) The second group for whom this report should be a useful read at is the HPDC community broadly defined: resource providers, those who capture and design infrastructure for community requirements, tools and service providers. By discussing a set of polar science applications, the problems and the needs of polar scientists, the hope is that this report provides a useful starting point for solution providers. (iii) Last, but not least the this report is aimed for NSF and possibly other funding/policy agencies, to help them understand the current state-of-practice of the technology and community efforts. We hope that this report will inform future directions that the Foundation will chose to prioritize.

In Section 2 of this report we provide a narrative of existing and representative problems that the polar community is working on, as was communicated by the participants at the workshop. In Section 3, we provide an overview of existing and future NSF HPDC capabilities and offerings. A recurring and underlying consideration was "data"; appreciating the importance of data, in Section 4 we isolate some aspects of data that came to the fore in the workshop. Section 5 provides a narrative of the workshop -- discussions, observations, and possible steps that could be taken to bring the HPDC and the polar sciences community together. Section 6 is a brief distillation of these discussions, observations and perceptions.

2 Polar Science Applications Session Overview

The polar HPDC Workshop started with a science session to begin communicating the broad arena of polar research to the computer scientists. With a total of four presentations, this session was geared to provide an overview of polar science questions revolving around the depletion of land ice, sea ice, and permafrost. A major focus was to introduce the various sources of large datasets that are used in polar research such as from satellite and airborne datasets, as well as climate model outputs. This session also provided an overview of the challenges faced in storage, handling, processing, and dissemination of the several levels of the raw and processed data.

In his keynote address, Paul Morin of the Polar Geospatial Center showed the increasing availability of high-resolution satellite imagery capable of resolving surface features up to half a meter. One immediate application of this dataset is the improved world coastline retrieval, an important component used in coastal monitoring, shipping, and the fisheries industry. Satellite imagery with improved resolution also helps to plan research fieldwork and to conduct science itself, for instance to quantify how glaciers and ice sheets are retreating with changing climate. This presentation also showed the problems associated with processing these high resolution images from raw data to various stages of data reduction. There is an additional problem of transferring these data to users. FTP is a common dissemination mechanism but is time consuming for very large datasets; instead these datasets are usually delivered on hard disks by post. Storage in hard disks runs the risk of these disks failing and loss of data. In addition, integration of various datasets to analyze complex processes of ice mass loss requires computational power that is increasingly becoming inadequate. This talk highlighted the importance of supercomputing and networks that polar scientists can use for their increasing needs of high resolution data processing, integration and analysis.

In post-workshop analysis and discussion, Paul Morin said, *"Cyberinfrastructure is needed for simulation and extracting terrain from high resolution imagery. The more I thought about it, the more that is what I see".* Interestingly as a consequence of interactions and discussions at the NSF workshop, Paul Morin is now on the path towards to using cyberinfrastructure thanks to nascent partnership with XSEDE.

Dr. Anna Liljedahl talked about the thawing permafrost in Alaska in response to climate change. Permafrost in Alaska harbours a unique ecosystem that supports the fragile tundra vegetation and the Arctic flora and fauna. As the permafrost degrades, the polar landscape in these regions sinks, destroying the fragile ecosystem and the creatures that exist there. In addition there is also the problem of the emission of trapped methane in this region. Given the interactions between these processes and the methods of data capture for permafrost, this session highlighted the multi-dataset challenges of polar science, with their corresponding implications for HPDC resource use by the domain.

Rapidly disappearing Arctic sea ice is probably the most visible indicator of global warming and climate change. Dr. Julienne Stroeve discussed sea ice conditions and trends across the coast of Alaska and Greenland. Multi-year ice coverage is decreasing as seasonal melt thins the existing ice pack. This has tremendous impact on albedo and in turn ocean warming and feedback to the atmosphere. The latter has a direct impact on warming the water and atmosphere near Greenland and accelerates melting of the Greenland ice sheet. This talk integrated remote sensing observations with modeling to quantify the loss of sea ice in recent years.

Dr. David Braaten presented the challenges faced with the collection and analysis of airborne geophysical datasets such as the airborne radar. Ice penetrating radar is used to probe into the ice to study deformation, ice dynamics, bedrock topography, and accumulation rates. This is an example of a large dataset that goes through several levels of intermediate processing before it reaches the stage where useful information can be extracted from it. High performance computing will help in processing and analysis of datasets such as these that have multiple levels of intermediate data.

Dr. Anne Nolin presented her work with IceTrendr, a prototype for automated visualization and analysis tool for mapping changes in glacier extents. Glaciers worldwide have been losing mass with impacts on water resources, the hydrological cycle and global sea levels. Remote sensing data from the Landsat satellite project extend back to the 1970's provide an extraordinary resource for mapping glacier extents and other characteristics. IceTrendr is being developed to facilitate the use of these resources among glaciologist, climatologist and educators. Currently, the IceTrendr mines Landsate data back to 1984 for five glaciers and can be used to extract glacier extent, terminus position, time series animations and more

3 High-Performance Distributed Computing Infrastructure

High-performance and distributed computing resources are part of Advanced Cyberinfrastructure for Science (ACI). According to Stewart et al., "cyberinfrastructure consists of computational systems, data and information management, advanced instruments, visualization environments, and people, all linked together by software and advanced networks to improve scholarly productivity and enable knowledge breakthroughs and discoveries not otherwise possible" [5]. In this section we will broadly describe the NSF advanced cyberinfrastructure as well as cloud infrastructure partnerships that are available to the polar community.

There were several talks that provided a detailed overview of existing and upcoming capabilities that NSF infrastructure provides. It was realized, however, that an essential component in the sustainable and effective bridging of polar science problems is, having the right community structures on the one hand, alongside training and education efforts on the other. There is a critical need for community structures and organization. "Other communities" have identified similar needs. As a prototype example of a community till recently new to advanced CI but one that has made progress in the uptake of CI, we turned to the Water Science community.

3.1 Examples of Community Organization:

Hooper presented a perspective on HPDC from the water science community, which spans the disciplines of hydrologic science, civil engineering, ecosystem ecology, and physical geography. This diverse group of scientists have generally not used HPDC and, more broadly, have little experience with community models and community infrastructure. Data analysis and simulation models use a wide variety of languages and statistical packages.

An opportunity exists to engage this community as sensors are more widely deployed; field scientists face data management challenges that are new to them. Scientists who focus on model development also see the need for better code management; new data services being developed by Federal agencies and by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. *(*CUAHSI, [6]) Water Data Center are also recognized as valuable, but not yet integrated into current workflows. Therefore, cyberinfrastructure development must provide a tangible immediate benefit that outweighs the cost of adoption, CI developers need to pay attention to adoption of the tools and provide sufficient training and user support to enable adoption. These challenges have often not been met in the past. It thus reiterates the need for effective community engagement and dissemination programs.

3.2 Advanced Cyberinfrastructure

John Towns, Executive Director for Science and Technology at NCSA and the XSEDE Project Director presented the XSEDE advanced computing infrastructure. Justin Miller from Indiana University's (IU) HPC group presented a new NSF's cloud system JetStream and NSF's data science infrastructure called Wrangler. Anke Kamrath, director for Computing and Operations Services at NCAR presented data-centric infrastructure and operations supporting polar science at NCAR. Such systems as Yellowstone and Erebus were mentioned and future HPC system procurements were discussed.

Justin Miller described Indiana University's current participation in polar research in partnership with the Center for Remote Sensing of Ice Sheets (CReSIS) at the University of Kansas, and presented a technical overview of two National Science Foundation XSEDE high performance computing resources that will available soon to polar researchers. Indiana University has worked closely with CReSIS to develop data collection and processing equipment that is deployed on airborne platforms to Greenland and Antarctica, providing the ability to safely capture and backup data, as well as processing capabilities in the field. IU also provides data management and HPC processing capability for the data once it has returned to the United States. The XSEDE systems Jetstream and Wrangler are HPDC resources that will be available in 2015, providing researchers with access to computing and data analysis platforms. Jetstream in particular will allow for custom virtual machines tailored to polar researchers with software and access to data archives.

It is worth noting that NSF's ACI supports the research, development, acquisition and provision of state-of-the-art CI resources, tools, and services[1]:

- ***Advanced Computing**:* Provide open-science community with state-of-the-art computational systems ranging from loosely coupled clusters to large scale HPC instruments; develop a collaborative and innovative scientific computational environment. Current NSF funded high-end computing systems and facilities include:
  - XSEDE [7], with such resources as e.g. San Diego Supercomputer Center (SDSC) Gordon [8] and SDSC Comet [9, 10], Texas Advanced Computing Center (TACC) Stampede [11] and Wrangler [x], Pittsburgh Supercomputing Center (PSC) Blacklight [12], Indiana University Mason [13], University of Tennessee Darter [14].

[1] Source: Irene M. Qualters, Division Director, Advanced Cyberinfrastructure
National Science foundation: presentation at Oklahoma Supercomputing Symposium 2014.

    - BlueWaters [15]
    - Open Science Grid [16]
    - NCAR/Wyoming Supercomputing Center [17]
    - NSF cloud environments [17]:
        - JetStream (Indiana University, [19])
        - Bridges (PSC, [20])
- ***Data:*** Support scientific communities in the use, sharing and archiving of data by creating building blocks to address community needs in data infrastructure. NSF, through its DIBBS program [47], aims at improving the nation's capacity in data science by investing in the development of infrastructure, building multi-institutional collaborations to increase the number of data scientists and augmenting the ease of using data. Data Observation Network for Earth (DataONE), is a supported by NSF program that aims at supporting "innovative environmental science through a distributed framework and sustainable cyberinfrastructure that meets the needs of science and society for open, persistent, robust, and secure access to well-described and easily discovered Earth observational data" [48].
- ***Networking and Cybersecurity****:* Invest in campus network improvements and re-engineering to support computational and data science. Support transition of cybersecurity research to practice. NSF's Campus Cyberinfrastructure - Network Infrastructure and Engineering program [49] invests in improvements and re-engineering at the campus level to leverage dynamic network services to support a range of scientific data transfers and movement. Dozens of campuses leveraged this program to build ScienceDMZs (research computer networks), and upgrade their networking infrastructure. In 2015 the DIBBs and CC-NIE were merged into one program called CC-DNI [50].
- ***Software:*** Transform innovations in research and education into sustained software resources (shared tools and services) that are an integral part of cyberinfrastructure.
    - Software infrastructure for Sustained Innovation (SI2) is another important program supported by NSF [21]
    - Amazon AWS Scientific Computing Group has helped sponsor and provide computing resources to the NSF DataViz Hackathon for polar CyberInfrastructure [22]
    - National Open Science Grid [16] has also an extensive software program for connecting campus researchers to its distributed infrastructure. OSG Connect tool offers investigators simple and efficient access to distributed high throughput computing resources required by many of today's most challenging problems in science and engineering.

Dr. Dan Katz, NSF ACI Program Director in his remarks mentioned the SI2 program and other ACI infrastructure programs.

3.3 Data Foundations for Polar Science and High-Performance Distributed Computing

As integrating components of models, simulations, and visualizations, polar data connect polar science analysis needs and computational resources. polar data are not merely satellite and airborne observations – in situ measurements from the field and network observations from a myriad of local stations are required to get a clear understanding of the complex system of the the Arctic, Antarctic, and the regions they affect.

Polar data have management and use challenges above and beyond typical geodata. Given the complexities of how the data are collected, created, funded, stored, shared and described, integration must not be taken lightly. Services or web access, and high quality data management are required in order for any datasets to be useful in an HPDC environment. The success of HPDC systems and workflows requires that the input data be clear, consistent, and understandable by machines and humans, as well as have the integrity and robustness to measure, articulate and minimize uncertainty.

The NSF 2013 Polar Cyberinfrastructure Workshop [1] was held to, among other objectives, gather the data, cyberinfrastructure, and polar science communities together to build consensus and gather recommendations for a modern cyberinfrastructure that would meet end users' needs, including for High Performance Computing. Challenges articulated during these discussions included the diversity in polar data from data formats, locations, description levels, management and sharing practices, volumes, and more. The 2013 Workshop recommendations included making polar observation accessible through Data-as-a-Service (DAAS) components to better connect especially field data to a computation-heavy science paradigm. A DAAS approach would especially support HPDC work in terms of automated access and processing. Metadata and data standards, including syntax, structure, and semantics, will further enable HPDC access and use, though are currently emerging. Through mechanisms like the Antarctic and Arctic Data Consortium (a2dc) [37], polar data centers are working to coordinate practices to support reuse, including through HPDC resources.

NSF GEO/PLR has also made substantial investments in data management cyberinfrastructure for polar communities. One example is the Advanced Cooperative Data and Information Service (ACADIS), funded by PLR/ARC. ACADIS [23] provides a robust self-publishing interface for researchers where they can easily deposit their data, author metadata, and acquire a citation/DOI. In addition, it also provides publication services and these have been used for remote operational observational systems and model output in an automated mode. This is one example of how distributed systems and workflows might be crafted to integrate HPDC capabilities with data management facilities.

3.4 HPDC Training and Outreach

There are numerous barriers to effective development and use of HPDC in polar sciences. Among these is a lack of awareness of, and connection between, polar scientists and HPDC training resources. In scientific domains with heavy integration of HPDC resources, faculty and students work in teams with computing professionals to take advantage of, and contribute to, emerging capabilities in HPDC, data and software. Two training levels are important. The first is training in using basic HPDC environments and techniques, such as understanding the uses of, and the methods for accessing and "logging in" to national and local supercomputing systems, writing simple job submission and control scripts, managing files, using data analysis tools, and loading appropriate modules. The second level is one of participatory and bi-directional learning, creating an environment for ongoing collaborations through a multidisciplinary computational community. A diverse community where distributed teams can conduct interdisciplinary research is different than the environment often found in a single academic department. This type of collaborative science requires a fundamental change in how we train the future generation of researchers. New programs are needed to provide both the depth of polar science domain knowledge as well as the computer science and applied mathematics to provide a sufficient breadth in applying HPDC. These discussions would align with, and build on, similar education directions discussed at the March 2015 Intelligent Systems for Geosciences (IS-Geo) Workshop held in Washington DC [24, report forthcoming].

Educating the next generation of polar scientists in the opportunities and potential of HPDC resources is key to eventual uptake of HPDC in the polar science. There are many resources available currently to support this goal such as the following: XSEDE training workshops and champions program [25], the Great Lakes Consortium for Petascale Computation virtual summer schools [26], the University of Oklahoma High Performance Computing Workshop Series [27], the Centrality of Advanced Digitally Enabled Science materials [28], the Virtual School of Computational Science and Engineering Summer School [29], the NPO Software and Data Carpentry [30], and the International Summer School on HPC Challenges in Computational Sciences [31].

Representative Example of Education, Outreach and Training,

In addition to HPDC resources, NSF has worked directly with the Amazon AWS Scientific Computing Group to provide computing resources to hackathons and other environments focused on bringing cyberinfrastructure researchers and polar scientists together. Several prototypes and hackathons have been developed using AWS; cloud computing resources provide a long-term potential home for both the code and the data and the applications that leverage them. This includes the NSF DataViz Hackathon [22]; and the AGU hackathon held in December 2014 [32].

There exists great potential for leveraging this model for not only future hackathons in the polar sciences community, but in addition for consideration at the data centers and archive systems that could leverage the code, and applications, along with the data available to provide value to the polar cyberInfrastructure community. When combined with advanced data visualizations, there is a strong potential for designing rapid and responsive interfaces that can better bring scientists to the polar data that is out there. Examples of these visualization frameworks include “mashups” that overlay GIS capabilities on top of polar web services, and that do so using responsive frameworks such as Data-Driven Documents (d3).

HPC and cloud computing can significantly help in these regards as visualizing and interacting with data on the cloud requires elastic capability and resources, especially when considering prototypes that are under rapid development and constraints.

Some examples of end-to-end cloud-based systems under develop for the polar sciences community:

1. GISCube [33] - Visualizations, GIS, and data management for polar and other geosciences data.
2. PolarHub [34]
3. 3d-printing of polar data from large datasets [35]
4. TangeloHub Polar Viz in the Cloud [36]

4 Towards Consilience of Polar Science and HPDC: Workshop narrative

In preparation for workshop session on challenges in bringing polar science to HPDC, and to better target discussion at the workshop, a preliminary community survey was completed at two participating organizations, the University of Alaska Fairbanks and the National Snow and Ice Data Center. The survey included questions on needs for HPDC in project work, experiences with HPDC resources, challenges encountered, and ideas for solutions moving forward. Responses were captured from a diverse set of participants, including polar scientists and graduate students, data/informatics researchers, and software developers. These results were summarized before the polar HPDC workshop and informed session topics and discussion points.

Results of the survey highlighted not only the diverse nature of polar science needs and perspectives, but also the differences in perspective between the polar and HPDC communities. The range of needs from respondents included simple computations on very large datasets, condensing massive imagery datasets into usable products, and modeling the evolution of sea ice cover through the 21st century on a regional basis. In addition to the science use cases, emerging polar informatics projects added web crawling for polar data as an example of another computing-intensive processes. The common thread across the diverse replies was that a major challenge to incorporation of HPDC resources was fitting linear data processes into the batch-oriented HPDC computing environments.

Other challenges includes the lack of time, awareness, and expertise needed to leverage HPDC effectively. For instance, respondents pointed to the learning curve needed to access and use HPDC resources, recoding reliable scripts into HPDC-friendly form, and understanding what HPDC resources are available in the first place. Moving data into and out of HPDC environments, unclear HPDC prioritization processes especially for small jobs, and the long turnaround times were also identified as challenges

Based on the survey results, two questions were presented at the workshop to break-out groups:

1. What are important resources, including tools, capabilities, and services, to advance polar research?
    a. Existing resources
    b. Missing resources
    c. Poorly developed resources
2. What are the most important challenges to address?

Break-out groups were structured to include representation from each of three high-level participant categories of Environmental Scientists, Computational Scientists, and self-identified 'Bridging' people. The discussion was wide ranging, a summary of the primary themes to emerge included the following.

1. **Data management**: Enhancing capacity for reanalysis of data was discussed extensively. Data stewardship (storage and cataloging), discoverability, and visualisation, were identified as areas in need of improvement. Non-HPDC participants were unaware of existing tools such as a metadata extractors, format wranglers, databases, search functions, and web portals tailored to scientific data.
2. **Data processing:** While the use of computational models in the polar community is an important use case for HPDC, many investigative avenues which do not use such models are also bottlenecked by data processing. This use case for HPDC is less well known but equally important. Tools and training in using classic HPDC cores, the cloud, and local desktop multi-core processors, to run data analysis and transformation algorithms over many independent data files in parallel, would be highly beneficial. Given the overhead of developing data processing pipelines to exploit such resources, it would also be beneficial if such tools might be easily shared within the community.
3. **Model integration:** Given the diversity of scientific domains within the polar community, tools and methods for facilitating the integration of multiple models (both of the theoretical models and the theoretical expressed as computational models) are needed. Part of enabling such is the development of standardized taxonomies.
4. **Training in HPDC resource use:** While there are already tools and resources available for training scientists on use of HPC tools (emails advertising parallel computing courses and XSEDE notifications were cited), a lack of knowledge around the potential benefit of HPDC resources means scientists are reluctant to allocate their limited time to opt-in. One of the conversations to come out of the workshop was a realisation on the part of

scientists that the HPDC resources might now be sufficiently accessible and beneficial to their science to be worth pursuing. Greater awareness of the value of HPDC training, and possibly more targeted training resources, are needed.

Through each of the discussion themes discussed above, possible areas of progress were identified:

1. **Data management:** While semantic negotiations across multiple domains are still maturing, tools currently exist for transforming and enabling discoverability of scientific data. Apache Tika already recognises HDF and NetCDF, along with standard txt, csv, excel, and various image data formats, enabling auto parsing of such files for metadata extraction and data archiving. Apache Solr, already in use by polar data discovery applications including the ACADIS Gateway[38] and Arctic Data Explorer [39], is a tool that enables search of such data archives. ACADIS and the Arizona State University PolarHub [34] project both offer federated polar data search. Data brokering (ex. EarthCube BCube [40]) is emerging with potential to better discover and transform data.
2. **Data processing:** As part of EarthCube's Geosoft project a scientific tool sharing portal (Turbosoft [41]) and training portal (Softcamp [42]) are being developed. Both could be utilized to enable scientists to make greater use of available compute resources (such as what XSEDE has to offer, but also local multi-core servers and desktops). It is worthwhile to consult with these project teams in regard to what would benefit the polar science community, and to track what they are doing for distribution to the community.
3. **Model integration:** The Community Surface Dynamics Modeling System (CSDMS) at University of Colorado Boulder [43] is making noteworthy progress on enabling multi-model coupling. They have created a methodology for standardising model interfaces that produces more exclusive quantity and attribute names than the CF standard names convention. This allows the immediate evaluation of interoperability of models on an I/O basis. They have added functionality that enables verification of the scientific validity of inter-model interaction given the different assumptions made by each, and the numerical stability of the combination given the internal numerical methods used by each. CSDMS and the polar science community are beginning to partner more closely on incorporating additional models into the CSDMS system.
4. **Training in HPDC use:** The workshop discussion agreed with the results of a more broad 2011 US study [44], which indicated the steep learning curve is one of the primary reason behind the scientific community not using available compute resources. Work can certainly be done towards lowering this barrier, such as by sharing tools made by others, and hosting training workshops (ideally online and targeted at senior researches according to some discussions at this workshop). Some sites already exist offering training in some general purpose tools (such as Software Carpentry [45]), while others more specific to this domain are under development (the aforementioned SoftCamp [42]). However, also developing relevant case studies demonstrating the scientific gains possible, might go further in convincing the community the time investment on their part is worthwhile.

Here we present, in a semi-structured fashion, some other observations made by the participants or discussion points that arose during the workshop:

1. **Community:**
    a. In order for the sharing of tools to be successful, formal guidelines regarding the appropriate use, installation, and operation of such tools needs to be included. Any tool sharing forum must include such guidelines and find a means of enforcing their posting (perhaps via public user rating). The broader open source software community equally has extensive experience in this domain of sharing software tools, and should therefore perhaps be approached for further input.
    b. Fortunately, there is now much potential for liaising the efforts of the HPDC and polar meteorology communities to focus on improved Arctic-region weather (and climate) predictions; for example, the recently begun World Weather Research Programme (WWRP) Polar Prediction Project (PPP) of 2013-2022, and the planning for the 2016-17 Year of Polar Prediction (YOPP) with intensive observing periods. The Polar HPDC Workshop and subsequent efforts are well-timed given these and other opportunities.
2. **Data**:
    a. We have offered no specific path to addressing the problem of data visualisation.
    b. As a means of extracting file metadata for search.
3. **Polar sciences:**
    a. The CSDMS project mentioned above, while highly beneficial to the computational modelling portion of the polar community, does not directly address the problems of potential ambiguity in inter-domain taxonomy on a data file level. The CSDMS's approach might perhaps, however, be adapted for non-computational models portion of the community too.
4. **Cyberinfrastructure:**
    a. There is a need for balanced cyberinfrastructure, as even within the Polar science community, there are a range of computational requirements.
    b. The majority of polar community science belongs to so called long tail of science, although there are many grand challenges there, too. Many challenges are microcosms of the state of the overall community [46].
    c. Given the nascent state of collaborations among the polar science, HPDC, and data communities, there is a need for communication, consensus, and buy-in from participants across the multiple domains which must be iterative, sustained and multi-dimensional conversations.

5 Recommendations

The following recommendations summarize the workshop discussions and the subsequent report content above.

1. **Continue cross-community engagement** - Polar research has high potential for significant scientific advancement when partnered with HPDC capability. Domain scientists must be engaged by cyber/computer scientists across different levels in order to determine their cyberinfrastructure requirements. Inclusive partnerships are needed to come to a shared understanding of concepts, terms, vocabularies, and overall requirements. HPDC needs for the polar science community must be understood, articulated, integrated, and aligned in the context of national and international cyberinfrastructure.
2. **Promote awareness of HPDC training resources for Polar scientists** - The polar community encompasses a significant portion of users who need tools for exploiting computing power. In addition to strategic technical development and broad consultation with stakeholders, educating a new generation of polar scientists in the skills needed to realize the opportunities and potential of HPDC is key to eventual uptake of HPDC in polar science. Leveraging existing tools and/or assessing educational gaps are first steps in this education process.
3. **Work towards a roadmap for HPDC uptake in the Polar sciences** - With progress on the first two recommendations, an inclusive team with representation from polar science, data, and HPDC communities should design a plan so that new and existing NSF-funded cyberinfrastructure efforts are cognizant of the needs of the polar science community. This plan would contain specific activities to advance HPDC uptake in the polar sciences as well as an ultimate vision for outcomes.

This workshop was a successful next step in a much longer conversation to better connect polar science with HPDC. To continue this progress, it is recognized that caution must be exercised when considering greater data and model interoperability across multiple domains. Domain conventions and assumptions regarding variable units and scales, or collection methodologies that exploit domain appropriate assumptions, or domain taxonomies, need to be made explicitly clear. If not explicit and clear, it is highly likely that inappropriate applications, analyses, and transformations of data would lead to incorrect conclusions or simply waste scientists time. Especially as the Polar-HPDC collaborations emerge, care must be taken to build trust and a common language to support useful progress. Bi-directional sharing of expertise and perspectives will ensure correct, sustainable and extensible design of solutions. With this foundation, HPDC will make it routine to address demanding problems that are simply out of reach currently, as well as make currently difficult computational problems simpler to solve. HPDC will thus help transition polar science into a computation and data driven quantitative science wherein, the sum of HPDC and polar science will be much greater than the sum of its parts.

6 Acknowledgements

This workshop was made possible by NSF Award PLR-1445823 (Program Director: Marco Tedesco). We also thank Ora Titus (ECE) and others who made local arrangements.

7 Additional Report Contributors:

- Indrani Das (Columbia)
- Don Middleton (NCAR)
- Chris Mattman (NASA)
- Matteo Turilli (Rutgers)
- Justin Miller (Indiana)

8 References

[1] Pundsack, J., R. Bell, D. Broderson, G.C. Fox, J. Dozier, J. Helly, W. Li, P. Morin, M. Parsons, A. Roberts, C. Tweedie, and C. Yang. 2013. Report on Workshop on Cyberinfrastructure for Polar Sciences. St. Paul, Minnesota. University of Minnesota Polar Geospatial Center, 17pp.
[2] High-Performance & Distributed Computing for Polar Sciences: Applications, Cyberinfrastructure and Opportunities Workshop URL: https://sites.google.com/site/polarhpdc/
[3] High-Performance & Distributed Computing for Polar Sciences: Applications, Cyberinfrastructure and Opportunities Workshop Agenda: https://docs.google.com/document/d/1y0-T6_cdbN8Ap5RWiv5U2XoKi_7gMtp6Befd4fCxYUk/edit
[4] Polar-HPDC Workshop material Google Drive (open access): http://goo.gl/plb05w
[5] C. A. Stewart, S. Simms, B. Plale, What is Cyberinfrastructure?, SIGUCCS'10, October 24–27, 2010, Norfolk, Virginia, USA.
[6] Consortium of Universities for the Advancement of Hydrologic Science, Inc.
[7] John Towns, Timothy Cockerill, Maytal Dahan, Ian Foster, Kelly Gaither, Andrew Grimshaw, Victor Hazlewood, Scott Lathrop, Dave Lifka, Gregory D. Peterson, Ralph Roskies, J. Ray Scott, Nancy Wilkins-Diehr, "XSEDE: Accelerating Scientific Discovery", Computing in Science & Engineering, vol.16, no. 5, pp. 62-74, Sept.-Oct. 2014, doi:10.1109/MCSE.2014.80
[8] San Diego Supercomputing Center (SDSC) Gordon: http://www.sdsc.edu/us/resources/gordon/
[9] San Diego Supercomputing Center (SDSC) Comet: https://portal.xsede.org/sdsc-comet
[11] Texas Advanced Computing Center (TACC) Stampede: https://www.tacc.utexas.edu/stampede/
[12] Pittsburgh Supercomputing Center Blacklight: https://portal.xsede.org/psc-blacklight
[13] IU Mason: https://kb.iu.edu/d/bbhh
[14] UT Darter: http://www.nics.tennessee.edu/computing-resources/darter
[15] Brett Bode, Michelle Butler, Thom Dunning, William Gropp, Torsten Hoe- fler, Wen-mei Hwu, and William Kramer (alphabetical) (2012). The Blue Waters Super-System for Super-Science. Contemporary HPC Architectures, Jeffery Vetter editor. Sitka Publications, November.Edited by Jeffrey S . Vetter, Chapman and Hall/CRC 2013, Print ISBN:

978-1-4665-6834-1, eBook ISBN: 978-1-4665-6835.
http://www.ncsa.illinois.edu/enabling/bluewaters
[16] R. Pordes, D. Petravick, B. Kramer, D. Olson, M. Livny, A. Roy, P. Avery, K. Blackburn, T. Wenaus, F. Wurthwein, I. Foster, R. Gardner, M. Wilde, A. Blatecky, J. McGee, and R. Quick, "The open science Grid," Journal of Physics: Conference Series, vol. 78, no.1, 2007, pp. 012-057.
http://osg-docdb.opensciencegrid.org/0006/000637/002/OpenScienceGrid-4.pdf
[17] National Center for Atmospheric Research/Wyoming Supercomputing Center: http://nwsc.ucar.edu/
[19] JetStream at Indiana University: http://pti.iu.edu/jetstream/
[20] Bridges at Pittsburgh Supercomputing Center: http://www.psc.edu/index.php/bridges
[21] Software infrastructure for Sustained Innovation (SI2)
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504817
[22] 2014 NSF Polar Cyberinfrastructure DataVis Hackathon, November 3-4 New York City: http://nsf-polar-cyberinfrastructure.github.io/datavis-hackathon/
[23] Moore, J.A., M Serreze, D. Middleton, M.K. Ramamurthy, L Yarmey (2013), The Arctic Cooperative Data and Information System: Data Management Support for the NSF Arctic Research Program, Abstract IN44A-07 presented at 2013 Fall Meeting, AGU, San Francisco, Calif., 9-13 Dec.
[24] 2015 NSF Workshop on Intelligent Systems for Geosciences, March 26-27, Washington D.C.: http://is-geo.org
[25] Extreme Science and Engineering Discovery Environment (XSEDE) Training: https://www.xsede.org/training1
[26] Great Lakes Consortium for Petascale Computation (GLCPC) (2009). Virtual School of Computational Science and Engineering Summer School 2009: Scaling to Petascale. http://www.greatlakesconsortium.org/events/scaling/
[27] OU Supercomputing Center for Education & Research (OSCER) (2015). Supercomputing in Plain English: A High Performance Computing Workshop Series.
http://www.oscer.ou.edu/education.php
[28] Centrality of Advanced Digitally ENabled Science (CADENS). Enabling Discovery materials. http://www.ncsa.illinois.edu/enabling/vis/cadens
[29] Virtual School of Computational Science and Engineering (VSCSE) (2015). Supercomputing for Everyone Series: Performance Tuning Summer School (August 17 - 21, 2015). http://www.vscse.org/summerschool/2014/index.html
[30] Software Carpentry (SOFTCAR) (2015). Workshops
http://software-carpentry.org/workshops/index.html
[31] XSEDE (2015). 2015 International Summer School on HPC Challenges in Computational Sciences. https://www.xsede.org/2015-international-summer-school
[32] AGU hackathon, Dec 2014
[33] GISCube, see: https://giscube.jpl.nasa.gov/tools - Visualizations, GIS, and data management for Polar and other geosciences data.
[34] PolarHub: http://polar.geodacenter.org/polarhub/

[35] 3d-printing of polar data from large datasets,
https://github.com/NSF-Polar-Cyberinfrastructure/datavis-hackathon/issues/50
[36] TangeloHub Polar Viz in the Cloud
https://github.com/NSF-Polar-Cyberinfrastructure/datavis-hackathon/issues/42
[37] Antarctic and Arctic Data Consortium (a2dc): http://www.a2dc.org/
[38] Advanced Cooperative Arctic Data and Information Service (ACADIS) Gateway:
http://aoncadis.org
[39] Arctic Data Explorer (ADE): http://nsidc.org/acadis/search
[40] Bcube http://nsidc.org/research/informatics/bcube
[41] Turbosoft Portal, part of the NSF EarthCube Geosoft project,
http://www.geosoft-earthcube.org/turbosoft-portal
[42] Softcamp, part of the NSF EarthCube Geosoft project, http://earthcube.org/group/geosoft ,
https://github.com/geosoft-earthcube/geosoft-apis
[43] CSDMS http://csdms.colorado.edu/wiki/Main_Page
[44] Prabhu Prakash, Zhang Yun, Ghosh Soumyadee, August David, Huang Jialu, Beard Stephen, Kim Hanjun, Oh Taewook, Jablin Thomas, Johnson Nick, Zoufaly Matthew, Raman Arun, Liu Feng, Walker David, A survey of the practice of computational science, Proceedings State of the Practice Reports, 2011, DOI: 10.1145/2063348.2063374
[45] Software-Capentry, http://software-carpentry.org/lessons.html
[46] http://arxiv.org/pdf/1210.3271.pdf and the follow up SC11 BoF:
https://sites.google.com/site/toaunifiedci/sc11-bof.
[47] NSF DIBBS: http://www.nsf.gov/pubs/2014/nsf14530/nsf14530.htm
[48] DataOne: https://www.dataone.org/
[49] NSF CC-NIE: http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf14521&org=NSF
[50] NSF CC-DNI: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504748

9 Appendix I: Participants and Picture

1. Lynn Yarmey (NSIDC)
2. Paul Morin (Minnesota)
3. Julienne Stroeve (NSIDC)
4. David Braaten (CReSIS)
5. Anne Nolin (Oregon)
6. Anna Liljedahl (UAF)
7. Indrani Das (Columbia)
8. Eric Noble (CCNY)
9. Rick Hooper (CUAHSI)
10. Guido Cervone (PSU)
11. Laura Harding (PSU)
12. Andrew Carleton (PSU)
13. Jane Wyngaard (NASA)
14. Richard Knepper (Indiana)
15. Jarek Nabrzyski (Notre Dame)
16. John Towns (NCSA/UIUC, XSEDE)
17. Don Middleton (NCAR)
18. Anke Kamrath (NCAR)
19. Andre Luckow (HPC Consultant)
20. Larry Murakami (Arctic Region Supercomputing Center)
21. Peter Fox (RPI)
22. Carol Song (Purdue)
23. Asa Rennermalm (Rutgers)
24. Shantenu Jha (Rutgers)
25. Matteo Turilli (Rutgers)
26. Oscar Schofield (Rutgers)

27. Robert Harrison (Stony Brook) remote presentation

NSF
  Marco Tedesco (Polar)
  Daniel S. Katz (ACI/CISE)

Appendix II

Polar Science Problems that motivate and neccessiate HPDC resources.

An important challenge in polar sciences is to quickly and easily integrate observations from the field, satellite, airborne platforms, and computer modeling. Although observations have improved in resolution by leaps and bounds, climate models, both regional and global, have coarser resolution because of limitations in computer processors. High-resolution models (~km scale) are difficult to execute as it is difficult to assimilate high resolution data and the number of input variables. Analysis of global output from these models is also challenging due to the lack of sufficient memory.

Global climate is changing rapidly, with the post-1970s increase in near-surface temperatures varying by latitude zone and region. Temperatures have increased the most over the Arctic Ocean and its adjacent seas and land areas (e.g., Barents Sea, Alaska) and, for the Southern Hemisphere, the Antarctic Peninsula. In both polar regions there is strong feedback between ice, atmosphere and ocean. In the Arctic, the rising temperatures have accompanied reductions in sea-ice extent and also thickness; both in winter and summer, although the downward trend is more marked in the late summer (September) ice minimum. Replacing a more reflective surface (i.e., sea ice, especially if snow-covered) with more melt-ponds and open ocean, results in greater absorption of solar radiation in summer, which raises Arctic Ocean temperatures, providing a positive feedback ("Arctic Amplification") to further reductions in sea ice in subsequent seasons and years. Moreover, a reduced winter-time ice extent increases the heat fluxes from the exposed "warm" ocean to the overlying cold atmosphere, helping initiate or intensify high-latitude cyclonic storms, such as "polar lows". On the western side of the Antarctic Peninsula, wintertime sea ice extent has decreased in recent decades, likely due to the greater strength of the westerly winds which import milder air from lower latitudes of the South Pacific. Over the rest of the Southern Ocean, the wintertime sea-ice extent has increased. Although this seems paradoxical at first, it is consistent with the increased surface wind speeds, which enhance the divergence of the pack ice away from the continent and permit more ice to form in the leads between ice floes. At the same time, there is evidence that the Antarctic sea ice has become thinner as a result of a warmer ocean and the increased flux of fresh water from melting land ice as near-surface air temperatures have risen, especially for West Antarctica.

In turn, there is feedback between the polar and the tropical systems as well. The more rapid increase in Arctic temperatures compared with temperatures in lower latitudes, has decreased the meridional (south-north) temperature gradient, which means an increased tendency to "waviness" (i.e., higher-amplitude troughs and ridges) in the westerly winds over middle and high-middle latitudes of the Northern Hemisphere. In its more extreme form, waviness can resolve into persistent blocking events that produce anomalously warm conditions over the Greenland region but anomalously cool and wet conditions over north-east Canada and western and north-west Europe (e.g., July 2012). Because higher-amplitude waves are also slow moving, and may become fixed over particular longitudes for extended time periods, extreme weather and climate anomalies over middle latitudes—in all seasons—may be seen as resulting from the Arctic sea-ice declines; most notably, the recent warm and dry winters of western North America versus the extreme cold and snowy winters of eastern North America. In the Antarctic, there are strong teleconnections with lower-latitude circulation and climate anomalies on inter-annual time scales, in particular to the El Niño Southern Oscillation (ENSO) of El Niño ("warm") and La Niña ("cold") events that are centered on the South Pacific Ocean. ENSO events affect temperatures, winds, and storm tracks over most of the Southern Hemisphere, but especially over the Ross, Amundsen/Bellingshausen, and Weddell seas of the Antarctic Ocean. Moreover, the primary circulation pattern of the Southern Hemisphere extratropics—the so-called Southern Annular Mode (SAM)—shows a trend to increasingly positive values in recent decades (i.e., stronger westerly winds), possibly a result of the Antarctic ozone hole, which enhances cooling over the continent, a strengthened meridional temperature gradient, and a jet stream (and storm tracks) that displaced closer to Antarctica than in previous decades. In turn, ENSO and SAM may either reinforce or counteract one another, resulting in a range of spatially-varying climate anomalies. Thus, we can truly understand the changing polar climate only through a composite analysis of these ice-ocean-atmosphere systems and how they are influenced by large-scale global changes.

Appendix III

Unedited Remarks from Break-out Session (Day I)

Appendix IV

Representative Examples of HPDC and polar science efforts:

1. One example of a widely used ice sheet model is the Ice Sheet System Model (ISSM) [https://issm.jpl.nasa.gov/]. ISSM an **open source** model funded by the NASA Cryosphere and MAP (Modeling Analysis and Prediction) programs, JPL R&TD (Research, Technology and Development) and the National Science Foundation. This model can currently be run in parallel on clusters or on multi-core computers, for example the NASA Pleiades Super computer [http://www.nas.nasa.gov/hecc/resources/pleiades.html]. It has also be adapted for cloud computing and can be executed from the EC2 Amazon Cloud servers. We hypothesize that making models like ISSM available on XSEDE might further increase model applications.

2. The recent trend of marked reductions in Arctic sea-ice extent and thickness—especially in summer—is opening up the region to environmental stresses (e.g., greater shoreline exposure to erosion from strong winds). The potential for greatly increased human presence and activity with resource extraction and international shipping (military, commercial, leisure cruises) is also increased. Enhanced human exposure to Arctic weather and extreme events is also likely to result from the physical environment changes themselves: reduced sea ice permits greater wave action and also stronger sea-air fluxes of heat and moisture to energize or re-intensify storm systems. Some research suggest increases in storm activity is already taking place. Other extreme events that may become more frequent in the future are anomalous precipitation and wildfires. These events will create challenges to populations living in the Arctic, as well as research and industry.

As the required spatial and also vertical resolutions (e.g., for temperature inversions) increase, computing power must also increase to produce forecasts and warnings that are sufficiently timely to permit evasive or other actions to be taken. Because Arctic weather and climate hazards are often the result of strong sea-air interactions, coupled ocean-atmosphere models offer the best possibility for improving weather predictions, yet these are highly computationally intensive. Moreover, this is not just a question of increasing computer power and strategies for improved ingestion and assimilation of satellite data by the models: processes that can be reasonably parameterized (i.e., approximated) in larger-scale models (e.g., clouds, open-water leads in sea ice, radiative and convective heat fluxes) may need to be physically explicit to accurately forecast meso-scale weather and climate hazards for the Arctic.