# Network-Based Approaches to Climate Knowledge Discovery

***Climate Knowledge Discovery Workshop; Hamburg, Germany, 30 March to 1 April 2011***

PAGE 425

Do complex networks combined with semantic Web technologies offer the next generation of solutions in climate science? To address this question, a first Climate Knowledge Discovery (CKD) Workshop, hosted by the German Climate Computing Center (Deutsches Klimarechenzentrum (DKRZ)), brought together climate and computer scientists from major American and European laboratories, data centers, and universities, as well as representatives from industry, the broader academic community, and the semantic Web communities.

The participants, representing six countries, were concerned with large-scale Earth system modeling and computational data analysis. The motivation for the meeting was the growing problem that climate scientists generate data faster than it can be interpreted and the need to prepare for further exponential data increases. Current analysis approaches are focused primarily on traditional methods, which are best suited for large-scale phenomena and coarse-resolution data sets. The workshop focused on the open discussion of ideas and technologies to provide the next generation of solutions to cope with the increasing data volumes in climate science.

In the course of the workshop, it was agreed that a suite of new techniques is needed that interpret and link phenomena among different Earth system components and processes at multiple temporal and spatial scales. Tools that use a combination of high-performance analytics with algorithms motivated by network science, nonlinear dynamics, and statistics, as well as data mining and machine learning, could provide unique insights into challenging features of the Earth system, including extreme events and chaotic regimes.

Using complex networks was identified as one very promising solution. By representing the climate system as networks, scientists can improve their understanding of observed climate phenomena and complex relationships in the global climate system, as well as anticipation of the consequences of climate change. Networks constructed from climate data have been shown to detect natural changes in the climate system. There is also the potential to enhance regional climate predictions over land by exploiting atmospheric teleconnections. Such climate networks could be as large as millions or billions of nodes. Investigating data at this massive scale will require advanced parallel or multithreaded computing technologies and software.

The multidisciplinary and interdisciplinary nature of climate research, which requires extensive sharing of data and analysis results, further complicates studies of the Earth system. Workshop participants agreed on the need for a climate-specific ontology for annotating data sets, facilitating analysis workflows, reasoning, and providing a mechanism for data provenance. Ontologies, informally or formally, represent knowledge of a set of concepts and the relationships among those concepts often in specific problem domains. The semantic Web builds on semantic technologies, including ontologies in machine-readable form, to provide a framework that can allow data that may span application, organization, and domain boundaries, to be found, accessed, and used.

A broad initiative is needed both to educate climate researchers about the potential of using knowledge discovery tools and to conduct research into ways and means of applying graph-theoretic techniques to multidisciplinary climate model data. A number of areas for future work were identified at the workshop, including ontologies to describe the relationship between features and events, a CKD test bed based on community climate model data, and further examination of science topics that will allow the climate science and graph-analytic communities to interact on a set of concrete examples. Sample science areas include model intercomparison, climate teleconnections, and scale interaction.

More information on CKD activities is available at https://redmine.dkrz.de/collaboration/projects/ckd-workshop.

The following people contributed to this meeting report: Peter Fox, Rensselaer Polytechnic Institute; Auroop Ganguly, Northeastern University; Jim Kinter, Center for Ocean-Land-Atmosphere Studies; and Karsten Steinhaeuser, University of Minnesota, Twin Cities.

—Reinhard Budich, Max-Planck-Institut für Meteorologie, Hamburg, Germany; Per Nyberg, Cray, Inc., Seattle, Wash.; E-mail: nyberg@cray.com; and Tobias Weigel, DKRZ, Hamburg, Germany