Descriptive Analysis of the Global Climate System and Predictive Modeling for Uncertainty Reduction in Climate Projections using Complex Networks

Karsten Steinhaeuser (ksteinha@nd.edu)
Department of Computer Science & Engineering
Interdisciplinary Center for Network Science & Applications University of Notre Dame, Notre Dame, IN
Geographic Information Science & Technology Group, Oak Ridge National Laboratory, Oak Ridge, TN
Advisor: Nitesh V. Chawla (Notre Dame) *Co-Advisor:* Auroop R. Ganguly (ORNL)

Introduction & Motivation

As evidence in support of anthropogenic climate change continues to mount [1], the study of climate has become a focus of scientific research, political attention, and socioeconomic concern. There are many different aspects to studying climate, from the collection and analysis of historical/observed data to the understanding of current climate and its underlying physical processes and the development of models for the projection of future trends. And while considerable progress has been made in each of the areas, significant challenges still remain; for example, the Biological and Environmental Research Advisory Committee has identified the following Outstanding Grand Challenges in Climate Change Research to guide the U.S. Department of Energy strategic planning [2]:

1. Characterize the Earth's current climate, and its evolution over the last century to its present state.
2. Predict regional climate change for the next several decades.
3. Simulate Earth System changes and their consequences over centuries.

Our work focuses mainly on the second of these objectives, namely, projections of changes in regional climate and their impacts on natural and man-made systems. Traditionally, projections of future climate are based primarily on physically-based computational models. However, an incomplete understanding of the physical processes involved (as well as a number of other factors such as constraints on spatial and temporal resolution) limits their usefulness for projections at sub-continental scale (i.e., regional or local). Yet it is precisely at those scales that these data are desperately needed. From impacts assessments to policy decisions regarding mitigation or adaptation strategies, climate data must be relevant and credible at the level of states, counties, watersheds, etc. To this end, we are developing data-guided methods that can supplement computational climate models and reduce uncertainty in projections of regional climate.

Methodology

The intuition behind our approach is to extract patterns and predictive models from observed climate data [3], validate their predictive skills, and then integrate them with projections of future climate (Figure 1 shows a schematic diagram of this process). One of the assumptions underlying our particular methodology is that the climate system can be represented as a complex network [4], wherein spatial locations correspond to nodes in the network and weighted edges are created based on climatic relationships between them. This data representation, called a *climate network*, has been shown to exhibit interesting properties in the domain context [4][5].

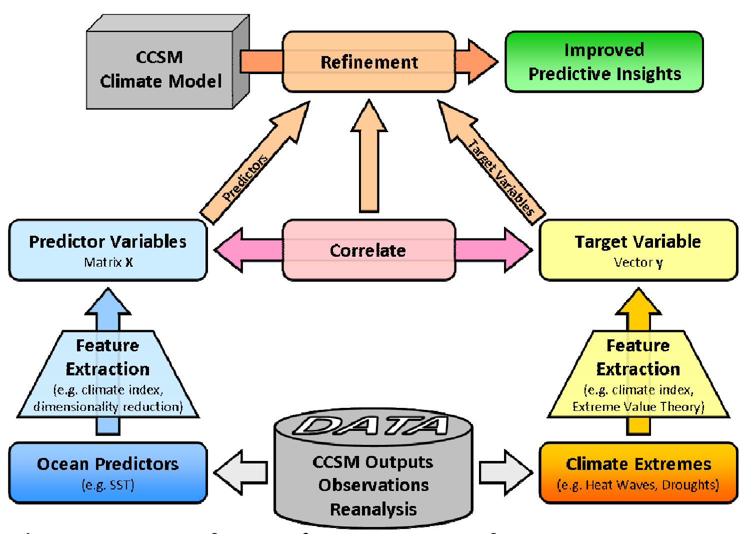

Figure 1. Process diagram for uncertainty reduction in projections of regional climate using data-guided models and insights.

Prior work involving climate networks has largely focused on what we call *descriptive analysis*, that is, studying the global network properties and interpreting them in the domain context (e.g., [4]). However, we propose that the use of complex networks can be extended to include *predictive modeling* in climate within what we have termed a 'unified framework', wherein we utilize networks to discover relationships between ocean climate indicators and land climate from observed data. Specifically, using community detection on climate networks we identify ocean clusters from the data, i.e., regions of homogeneous climatologic behavior. Cluster averages are then treated as potential climate indicators by using them as inputs to a predictive (regression) model for land climate. Thus, we are effectively using data mining techniques to "learn" the unknown physical relationships in the global climate system. We can test the predictive skill of these models, i.e., the validity of these relationships, using separate training and testing sets with held-out data. Our experimental results show that the data-mined models indeed have additional predictive power beyond simple autoregression, used as a baseline method for comparison. Presently, we are investigating several extensions of this approach, such as using nonlinear correlation measures for network construction and evaluating a range of different regression functions including linear regression, regression trees, support vector regression, and neural networks. Once we have determined the optimal representation and modeling technique, we will incorporate climate model data into this computational framework to obtain refined projections of future climate.

Discussion & Conclusions

The two primary motivating factors for this approach are (1) the existence of such relationships is well-documented in the domain literature, and (2) climate models have been shown to perform better over the oceans than over land. In addition, cognizant of the requirements for impacts studies and policy decisions, we are of course most interested in land climatology. This work has already led to some novel insights and promoted the use of complex networks in modeling the global climate system. We believe that this interdisciplinary research on the intersection of climate science and data mining ultimately represents an opportunity of mutual benefit: climate and environmental sciences pose challenging research questions that motivate the development of novel data mining techniques; in turn, data mining offers innovative solutions beyond the traditional scope of the domain sciences, which address present needs for impacts assessment and policy-making.

References

[1] Intergovernmental Panel on Climate Change (IPCC). "Fourth Assessment Report: Climate Change 2007." Cambridge Univ. Press, Cambridge, UK, 2007.

[2] Biological and Environmental Research Advisory Committee (BERAC). "Identifying Outstanding Grand Challenges in Climate Change Research: Guiding DOE's Strategic Planning." Office of Biological and Environmental Research, Washington, DC, 2008.

[3] E. Kalnay et al. "The NCEP/NCAR 40-Year Reanalysis Project". BAMS, 77(3): 437-470, 1996.

[4] A.A. Tsonis and P.J. Roebber. "The Architecture of the Climate Network." *Physica A: Statisctical and Theoretical Physics*, 333, 497-504, 2004.

[5] K. Steinhaeuser, N.V. Chawla, A.R. Ganguly. "An Exploration of Climate Data Using Complex Networks." *ACM SIGKDD Explorations*, 12(1), 2010 (to appear).

[6] K. Steinhaeuser, N.V. Chawla, A.R. Ganguly. "Complex Networks as Unified Framework for Descriptive Analysis and Predictive Modeling in Climate." (Article in review).

Acknowledgments – This work was supported by the National Science Foundation under Grant 0826958. The research was performed as part of a project titled "Uncertainty Assessment and Reduction for Climate Extremes and Climate Change Impacts", which in turn was funded in FY2009 by the initiative called "Understanding Climate Change Impact: Energy, Carbon, and Water Initiative", within the LDRD Program of the Oak Ridge National Laboratory, managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract DE-AC05-00OR22725.