WHAT CAN WE LEARN ABOUT CLIMATE MODEL RUNS FROM THEIR CAUSAL SIGNATURES?

Dorit Hammerling[1], Allison H. Baker[1], Imme Ebert-Uphoff[2]

Abstract—**This study applies methods from causal discovery theory to the output data of climate models. Causal discovery seeks to identify potential cause-effect relationships from data and is used here to learn so-called *causal signatures* from the data that indicate interactions between the different atmospheric variables. We hope that these causal signatures can act like *finger prints* for the underlying dynamics, and can as such be used in a variety of applications. Sample applications include (1) distinguishing correct model runs from incorrect ones, i.e. providing an additional error check for climate model runs and (2) assessing the impact of data compression on the causal signatures, as a means to determine which type and amount of compression is acceptable. Still being in the early stages of this project, we primarily describe work in progress and future work.**

I. BACKGROUND

The framework of *causal discovery* provides algorithms to identify potential cause-effect relationships from observational data [1], [2]. The output of such algorithms is a graph structure that indicates the potential causal connections of the observed variables. Originally developed for applications in the social sciences and economics, causal discovery is now used successfully in many disciplines including, recently, climate science, e.g. to track interactions between different locations around the globe [3], [4], [5] or to identify interaction patterns between compound climate variables [6], [7].

Here we focus on what we can learn about the relationships between individual atmospheric variables by applying causal discovery to the output of climate models. The data for each run then yields what we call *causal signatures*, i.e. patterns of interaction between the different atmospheric variables, and can be interpreted as *finger prints* of the underlying dynamics in the model. We use the well established framework of *structure learning* for *probabilistic graphical models*, which is described in much detail in [1], [2]. See [7] for the details of the approach used here.

Corresponding author: D. Hammerling, dorith@ucar.edu. [1] National Center for Atmospheric Research (NCAR), Boulder, CO, USA. [2] Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO, USA.

II. DATA

We use publicly available data from the Community Earth System Model (CESM) Large Ensemble (LENS) Community Project [8]. (See also https://www2.cesm.ucar.edu/models/experiments/LENS.) CESM-LENS data currently consists of 38 1-degree CESM simulations from 1920-2100. The ensemble members differ by an initial perturbation to the atmospheric temperature field. For the first stage of this work, we explore daily global spatial average timeseries data from 1920-2005 (86 years), which provides 31,391 data values for each of approximately 50 variables.

III. SAMPLE CAUSAL SIGNATURE PLOT

We first calculated results for the connections between 50 different atmospheric variables, based on daily data and global averages, and using two different data sets and a variety of temporal resolution (e.g. looking for connections between variables that require multiples of $D = 1, 10, 30, 60$ days to travel from cause to effect). The graph results indicated groups of variables that are highly redundant. We subsequently selected a subset of 15 variables that are representative of the whole set and have little redundancy. Fig. 1 shows a sample causal signature plot for those variables for one data set and $D = 1$ day. Repeating this procedure for the second ensemble member showed that most of the connections, and even many of the time scales are identical. However, for larger values of D we get graphs that are more dense and with very different connections. Clearly, for different values of D the algorithm picks up connections of different time scales (fast vs. slow interactions) and so one needs to look at more than one time scale to identify all significant connections.

IV. TARGET APPLICATIONS

A. Climate model software verification

Continually evolving and complex simulation codes such as CESM require frequent software verification and quality assurance to detect errors that may have been introduced into the code or hardware or software environments. To this end, the recently-developed CESM
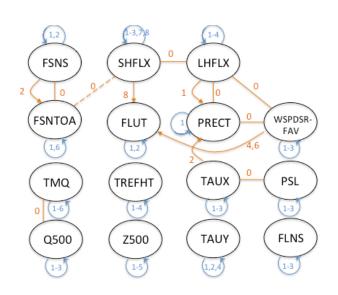
Fig. 1. Sample causal signature plot for $D = 1$. Connections in orange indicate connections between different variables, while blue arrows indicate persistence, i.e. how long the state of a variable strongly affects itself. The numbers next to the connections indicate the number of days from potential cause to potential effect. Dashed lines indicate weak connections.

Ensemble Consistency Testing (CESM-ECT) tool in [9] evaluates new climate simulations against an accepted ensemble distribution composed of yearly global means. The primary motivation for the work presented here is to potentially complement CESM-ECT by detecting subtle changes of relationships between variables on shorter time scales, which might go unnoticed using yearly global means.

We propose to first establish baseline causal signatures from an ensemble of model output that is known to be correct, i.e. to establish the *range* of *correct/acceptable* signatures. (Differences are expected for a variety of reasons, e.g. different initial conditions may trigger certain dynamic mechanisms more strongly than others; the statistical tests used in causal discovery are never perfect; noise in the data.) Secondly, the signature of any new run is calculated on the fly, and compared to the correct signatures to establish whether it is *acceptable* (within range) or not.

B. How much compression is acceptable?

At NCAR we are also studying how compression of the output data of model runs affects the properties of the data [10]. This is a very important question to decide what kind of compression can safely be applied to store the massive amounts of data generated by the models and available to the public. We will test how much the causal signature of the model outputs changes after different compression algorithms have been applied. This may

provide a useful method to determine which type and amount of compression is acceptable, in addition to the means already investigated in [10].

V. FUTURE WORK

Next steps for software verification include (1) calculating and studying the causal signatures for many different data sets (both correct and incorrect sets) and different time scales; (2) developing base lines for *acceptable* results; (3) repeating the procedure using averages over several regions, rather than just global averages. Furthermore, we also expect to find some new and interesting insights into interaction patterns between atmospheric variables from this approach.

REFERENCES

[1] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 2nd ed., Morgan Kaufman Publ., 552 pp, 1988.
[2] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*, Springer Lecture Notes in Statistics. 1st ed. Springer Verlag, 526 pp., 1993.
[3] I. Ebert-Uphoff and Y. Deng, *A New Type of Climate Network based on Probabilistic Graphical Models: Results of Boreal Winter versus Summer*, Geophysical Research Letters, vol. 39, L19701, 7 pages, doi:10.1029/2012GL053269, 2012.
[4] Y. Deng and I. Ebert-Uphoff, *Weakening of Atmospheric Information Flow in a Warming Climate in the Community Climate System Model*, Geophysical Research Letters, 7 pages, Jan 2014.
[5] T. Zerenner, P. Friedrichs, K. Lehnerts and A. Hense, *A Gaussian graphical model approach to climate networks*, Chaos, 23, 023103, 2014.
[6] T. Chu, D. Danks, and C. Glymour, *Data Driven Methods for Nonlinear Granger Causality: Climate Teleconnection Mechanisms*, Tech. Rep. CMU-PHIL-171, Dep. of Philos., Carnegie Mellon Univ., Pittsburgh, Pa., 2005.
[7] I. Ebert-Uphoff, Y. Deng, "Causal Discovery for Climate Research Using Graphical Models", Journal of Climate, Vol. 25, No. 17, Sept 2012, pp. 5648-5665.
[8] J. Kay, C. Deser, A. Phillips, A. Mai, C. Hannay, G. Strand, J. Arblaster, S. Bates, G. Danabasoglu, J. Edwards, J., M. Holland, P. Kushner, J.-F. Lamarque, D. Lawrence, K. Lindsay, A. Middleton, E. Munoz, R. Neale, K. Oleson, L. Polvani, and M. Vertenstein, *The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability*, B. Am. Meteorol. Soc., doi:10.1175/BAMS-D-13-00255.1, in press, 2015.
[9] A. Baker, D. Hammerling, M. Levy, H. Xu, J. Dennis, B. Eaton, J. Edwards, C. Hannay, S. Mickelson, R. Neale, D. Nychka, J. Shollenberger, J. Tribbia, M. Vertenstein, and D. Williamson, *A new ensemble-based consistency test for the Community Earth System Model*, Geoscientific Model Development Discussions, 8, 2015 doi:10.5194/gmdd-8-3823-2015.
[10] A. Baker, H. Xu, J. Dennis, M. Levy, D. Nychka, S. Mickelson, J. Edwards, M. Vertenstein, and A. Wegener, *A methodology for evaluating the impact of data compression on climate simulation data*, in: Proceedings of the 23rd international symposium on High-Performance Parallel and Distributed Computing, HPDC'14, 203–214, 2014.