Machine learning climate variability

Ji Hwan Park
Computational Science Initiative
Brookhaven National Laboratory
Upton, NY 11973, USA
parkj@bnl.gov

Shinjae Yoo
Computational Science Initiative
Brookhaven National Laboratory
Upton, NY 11973, USA
sjyoo@bnl.gov

Balu Nadiga
Los Alamos National Laboratory
Los Alamos, NM 87545, USA
balu@lanl.gov

Abstract

While comprehensive Earth System Models (ESMs) are the best tools available to understand climate system details, including its variability and ability to predict changes in the system, the immense computational burden associated with them opens up the possibility that reduced-order dynamical representations of such ESMs will enable and facilitate their use in a broad range of applications. By building on recent developments in deep neural networks, this work employs a hierarchy of such networks, ranging from multilayer perceptron to convolutional long short-term memory networks, to develop reduced-order dynamical descriptions of the spatio-temporal variability of temperature in a particular setting of a popular ESM. Upon evaluating such reduced descriptions from the point of view of predictive skill and predictability of climate more generally, it is evident that the more sophisticated of these network architectures is able to perform skillfully at lead times of up to about a year. Furthermore, this approach succeeds in capturing features of climate that have a basis in climate dynamics and can control predictability, further validating the reduced-order dynamical description. Related issues and further perspectives are also considered.

1   Introduction

Earth System Models (ESMs) that comprise atmosphere-ocean general circulation models coupled to other earth system components, such as ice sheets, land surface, terrestrial biosphere, and glaciers, are central to developing our understanding of climate [e.g., Chapter 9 in 1]. However, the immense computational infrastructure required and cost incurred in running such ESMs precludes their direct use in various applications to further develop our understanding of climate. Therefore, if validated reduced-order dynamical descriptions of ESMs can be developed, they would be useful, as stand-ins for ESMs in such applications [e.g., see 2, 3, 4]. However, in developing these types of reduced-order dynamical descriptions, one must deal with the issue of predictability of a complex dynamical system, in this case, the climate system.

Predictability of the climate system arises (a) from natural climate variability, i.e., variability internal to the climate system under conditions of constant external forcing, and (b) from the climate system's response to varying external forcing. Following [5], they are termed as predictabilities of the first and second kind, respectively.

A prediction of the first kind involves being able to accurately track the future evolution of the climate system after estimating its current state. Therefore, the skill of such a prediction is limited, on one hand, by errors and uncertainties in the model used to approximate the climate system's evolution and, on the other hand, by how errors and uncertainties in the initial condition evolve. Likewise, the skill in a prediction of the second kind is affected not only by model error but also by errors and uncertainties in specifying external forcing.

In the framework of comprehensive ESM, there are many reasons why making predictions of the first kind can be more difficult than modeling the climate system's response to secular changes in external forcing, such as due to greenhouse gases (e.g., refer to [6, 7]). These can include the scientific challenges involved in estimating the state of the climate system with sufficient accuracy, as well as the complex, multiscale, and chaotic dynamical nature of the climate system that complicates the process of accounting for uncertainty in the future evolution of errors in the initial state estimate. As such, our ability to produce longer-term projections that are controlled by external-forcing-related predictability is better developed than our ability to produce near-term (interannual) predictions controlled by natural-variability-related predictability. However, it also should be noted that the climate system's response to external forcing can be modulated by natural variability, leading to the response to external forcing being amplified or mitigated on certain timescales of natural variability.

Remaining in the ESM framework, while initialized predictions of climate seek to augment the external-forcing-related predictability realized in uninitialized long-term projections via predictability related to natural variability, there are a number of issues that remain to be resolved before such initialized predictions are skillful. For example, in many ESMs, observation-based initialization in the presence of model bias leads to a rather rapid departure of the initialized prediction trajectory from observations, necessitating post-processing of the predictions before they can show any skill at all. For these reasons, we concern ourselves with a deep learning approach to the more difficult problem of predicting the natural variability of climate as represented in an ESM in the present article.

2 Data and Method

Because the climate system is driven by solar radiation and one of the main fields of interest is the resulting temperature distribution on the Earth's surface, we seek to model the spatio-temporal distribution of surface air temperature. Further, given the shortness of the observed climate record, we focus on modeling the surface air temperature evolution in an ESM experiment—the pre-industrial control experiment (external forcing is held fixed) of a popular ESM, the Community Earth System Model (CESM2). This data are publicly available from the Coupled Model Intercomparison Project (CMIP) archive (https://esgfnode.llnl.gov/projects/cmip6), it mirrors and spans a period of 1200 years, and monthly averaged fields are used. As the seasonal cycle is highly predictable, we remove the climatological seasonal cycle and seek to model deviations from the seasonal cycle—variability caused by various nonlinear processes and feedbacks internal to the climate system because the external forcing is held fixed
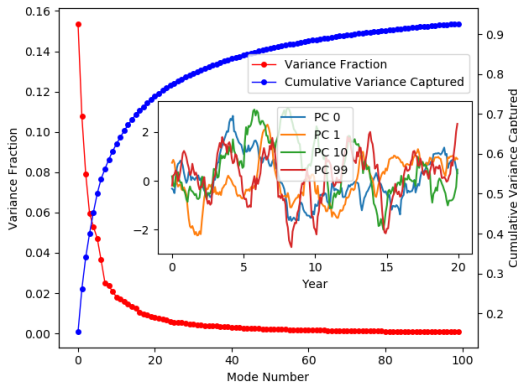

Figure 1: Details of PCA used in PCA-LSTM. Variance fraction captured by individual modes and the cumulative variance captured are shown in the main panel. The top 20 modes are retained in this study and explain in excess of 70% of the variance. The inset shows time variations of the normalized principal components over a 20-year period.

from year to year. The data dimension is $14389 \times 192 \times 288$ and about 4.5 GB.

To model this time-evolving data, we consider a range of deep neural networks, ranging from the multilayer perceptrons (MLP) to a variety of long short-term memory (LSTM) networks [8]. As for LSTMs, we consider ones that do not account for spatial correlations, as well as others that account for them using principal components (PCA+LSTM) and through convolutions (convLSTM) [9].
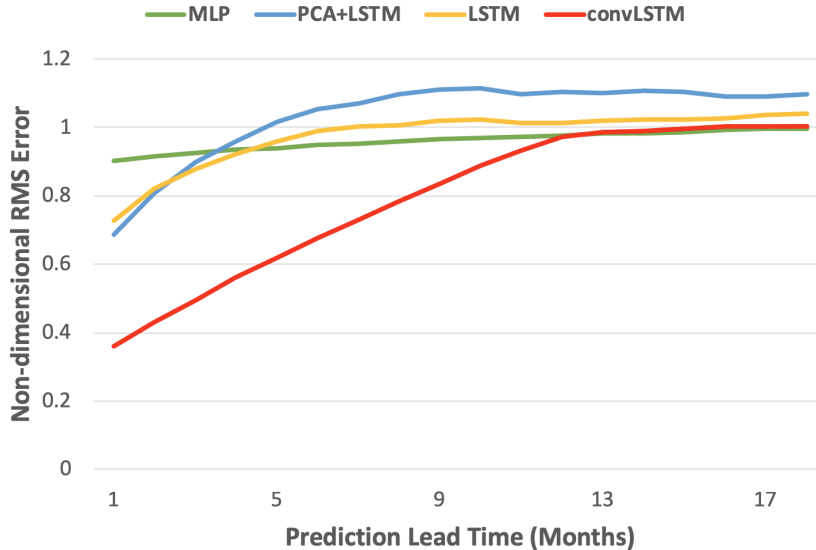
2

Figure 2: Non-dimensional root mean square (RMS) error comparison. We compare convolutional LSTM (convLSTM) to baseline algorithms: Multilayer perceptron (MLP), PCA+LSTM, and LSTM.

3 Results and Discussion

Setup. In a preliminary set of experiments, we considered the MLP, the PCA+LSTM, and the LSTM networks without explicit modeling spatial structure. Then, we compared them with an explicit spatial model, convLSTM. The MLP used in this experiment had two layers. The first layer featured 512 hidden states, and the second layer had 256 hidden states. The LSTM had three 3000-node layers. In PCA+LSTM, we first performed principal component analysis and applied the LSTM with two 256-node layers on the first 20 principal components (Fig. 1). Finally, the convLSTM network used in this experiment had two layers with a kernel size 5×5 and circular padding. Each layer included eight hidden states. In this set of experiments, the dataset was split three ways as train, validation, and test in a 60:20:20 ratio.

Model complexity vs. prediction skill. From MLP to convLSTM, the increased model complexity led to both increasing training data requirements and longer training times. However, the network's increasing complexity proved fruitful in leading to improved predictive skill (shown in Fig. 2).

We recall that the climatological season cycle has been removed given the fact that such cycles are easily predicted. Because PCA+LSTM and LSTM capture complex temporal behavior better, they perform slightly better than MLP. The slightly worse performance of PCA-LSTM when compared to LSTM is likely attributable to the fact that the top 20 components considered explain only about 70% of the variability. Nonlinear encoding of spatial relationships in conv-LSTM probably explains its better performance. In this case, predictions are skillful at lead times up to about a year. This result highlights the importance of simultaneous modeling of spatial and temporal features in achieving improved predictive skill. Needless to mention, the improved skill of convLSTM comes at the cost of a significantly higher level of model complexity: it required the full 32 GB of GPU memory of an NVIDIA V100 GPU card. For this experiment, we used the NVIDIA DGX-2.

Prediction error map analysis. Fig. 3 depicts a prediction error pattern map. Errors are shown at a lead time of one month, but longer lead times show similar qualitative behavior. A couple of standout features include: (a) errors tend to be meridionally structured with greater errors at mid and high latitudes, and lower errors at equatorial and tropical latitudesand, (b) in a given latitude band, errors tend to be higher over land than over oceans, and errors over land tend to increase with additional distance from oceans. This is true of the Arctic and Antarctic as well.

Discussion. The climate system, either the natural system or its modeled counterpart, is a complex dynamical system that exhibits variability on a diverse range of spatial and temporal scales. Setting
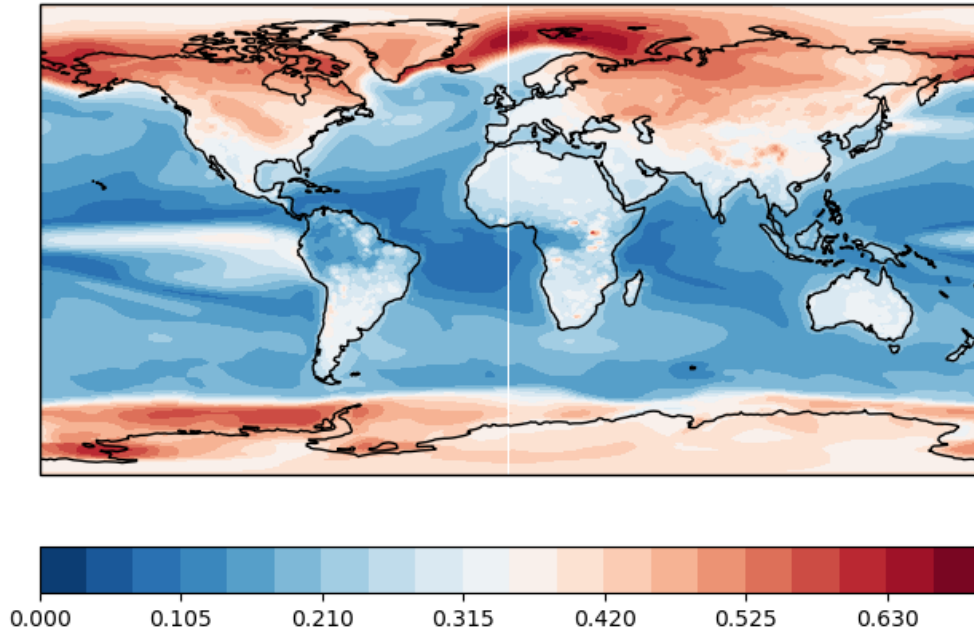
Figure 3: A map of prediction error. The map shows non-dimensional root mean square (RMS) error at a lead time of one month.

aside the seasonal cycle and holding the forcing constant from year to year, most such variability tends to be chaotic. While reduced-order dynamical models of comprehensive ESMs have potential use in varied applications, their development must encompass managing predictability given chaotic variability. We have performed the first experiments on learning climate variability as it occurs in a popular ESM using deep neural networks and find that some of these methods can predict the spatio-temporal variability of surface temperatures up to lead times of about a year. Currently, we are pursuing other architectures[1] and formulations of the problem. For example, we have formulated the prediction problem solely in terms of surface temperature because that climate is mainly a heat transport problem (more correctly, transport of moist static energy in the atmosphere and heat in the ocean and other subsystems). However, it is also the case that the transport is achieved by fluid motions (that are turbulent), and we know *a priori* that surface temperature is affected by numerous other processes. Therefore, it remains to be seen if a formulation that involves other variables (such as velocity and others) can be more skillful. We anticipate that additional investigations along these lines will serve to establish the kinds of methods studied in this work not only as useful reduced-order models of ESMs but also as robust methods for assessing climate predictability.

Acknowledgments

---

[1]For example, LSTNet [10], which also uses convolutional layers but captures both spatial and temporal variations. LSTNet further adopts recurrent and recurrent skip layers that use Gated Recurrent Units for short- and longer-term temporal relationships.

References

[1] Thomas Stocker. *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, 2014.

[2] Balasubramanya Nadiga, Changlin Jiang, and Amir Farimani. Predicting interannual variability of climate using deep learning. *Bulletin of the American Physical Society*, 2019.

[3] Balasubramanya T Nadiga and Nathan M Urban. Improved representation of ocean heat content in energy balance models. *Climatic Change*, 152(3-4):503–516, 2019.

[4] BT Nadiga and NM Urban. Dependence of inferred climate sensitivity on the discrepancy model. *arXiv preprint arXiv:1809.04068*, 2018.

[5] EN Lorenz. Climatic predictability. *The physical basis of climate and climate modelling*, pages 132–136, 1975.

[6] Gerald A Meehl, Lisa Goddard, George Boer, Robert Burgman, Grant Branstator, Christophe Cassou, Susanna Corti, Gokhan Danabasoglu, Francisco Doblas-Reyes, Ed Hawkins, et al. Decadal climate prediction: an update from the trenches. *Bulletin of the American Meteorological Society*, 95(2):243–267, 2014.

[7] Balasubramanya T Nadiga, Tarun Verma, Wilbert Weijer, and Nathan M Urban. Enhancing skill of initialized decadal predictions using a dynamic model of drift. *Geophysical Research Letters*, 2019.

[8] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, November 1997.

[9] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Proceedings of the 28th International Conference on Neural Information Processing Systems*, pages 802–810, 2015.

[10] Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. Modeling long-and short-term temporal patterns with deep neural networks. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 95–104. ACM, 2018.