Separating Climate Signals with Machine Learning

Authors

Ankur Mahesh (UC Berkeley, LBNL), Mark Risser (LBNL), Chris Paciorek (LBNL), Huanping Huang (LBNL), William Drew Collins (UC Berkeley, LBNL)

Focal Area

We use machine learning (ML) for the attribution of seasonal and extreme precipitation to different signals in the climate system. We compare ML-based attribution methods to traditional detection and attribution (D&A) techniques. Our project aligns with Focal Area 3 of the AI for Earth System Predictability Call for White Papers: "insight gleaned from complex data (both observed and simulated)."

Science Challenge

To better prepare for future changes in the water cycle, it is vital to understand the regional effects of anthropogenic climate change and internal climate variability on precipitation. With existing methods, there is uncertainty regarding the regional patterns of these climate processes. We propose using machine learning to decompose seasonal total and extreme precipitation (calculated by Risser et. al., 2019) into components caused by each of the following signals: anthropogenic forcing, internal variability, and modes of natural climate variability (such as El Niño/ Southern Oscillation, Atlantic Multidecadal Oscillation, and the Pacific North American teleconnection).

Rationale

Currently, the anthropogenic forced component of precipitation is calculated by averaging a large ensemble with perturbed initial conditions (Lehner et. al., 2020). This does not readily provide the forced component of observations or a single model realization (Barnes et. al., 2019), and it requires computationally expensive large ensembles for proper separation of the anthropogenic and internal variability components. Another method involves regressing the precipitation time series onto the time series of external forcing and climate modes of variability (Thompson et. al., 2015, McKinnon et. al., 2018). This method does not accommodate possible nonlinear relationships between the forcing and the precipitation time series.

In both observations and models, machine learning has the capability to identify spatial, nonlinear relationships between spatial climate fields and time series of the above climate signals (Wills et. al., 2020). Barnes et. al. (2020) and Sippel et. al. (2020) train ML models (they use nonlinear and linear ML models, respectively) to detect influences of anthropogenic climate change in precipitation and humidity data. Using interpretability methods to visualize their trained ML models, they identify the indicator patterns and spatial fingerprint of anthropogenic climate change. We propose extending the trained ML models of Sippel et. al. to account for nonlinear relationships, and we propose extending those of Barnes et. al. to also quantify the *anthropogenic and natural variability contributions* of extreme precipitation in CONUS. Initially, we plan to use ML for signal separation of monthly precipitation data in CONUS; however, over the next decade, we envision that similar ML methods can be used to quantify the

DOI: https://doi.org/10.2172/1769778
DOE BER, AI4ESP 2021 #AI4ESP1088 (https://www.ai4esp.org)

anthropogenic and natural variability contributions to a range of spatial and temporal time scales. In particular, we also envision that they can be used to conduct attribution studies on specific classes of extreme weather events, such as tropical cyclones, which currently use high-resolution dynamical climate models (Patricola and Wehner, 2018).

Narrative

We present a generative adversarial network (GAN) problem formulation, utilizing two neural networks. First, given a 2D precipitation field at a single timestep, the "discriminator" neural network is trained to identify the index of a climate signal. (The index of the climate signal could be the ONI index for El Nino or the greenhouse-gas forcing (W/m^2) for anthropogenic climate change.) Second, the "generator" neural network is trained to transform the 2D field to completely remove the influence of the climate signal; the generator is trained to fool the discriminator. In this way, the generator learns the component of each climate signal.

Using data from reanalysis and the Detection and Attribution Model Intercomparison Project (Gillett et. al., 2021), we compare the forced component calculated by the GAN to those calculated by regression-based methods (Risser et. al., 2021) and large ensembles of dynamical simulations (Deser et. al., 2020). Over the next decade, we propose to address the following questions in the use of machine learning for D&A.

What criteria should be used for comparing dynamical, statistical, and ML-based D&A methods?

We propose a comparison between dynamical D&A methods, traditional statistics-based D&A methods, and the proposed GAN. First, we will compare each method based on a residual consistency test (Allen et. al., 1999, Kirchmeier-Young et. al., 2020) to determine if the residual between total precipitation and the forced component is consistent with expected natural variability. We also will evaluate the methods by testing if the discriminator can identify a climate signal after its corresponding component has been removed.

Second, we plan to evaluate each method based on its interpretability: while dynamical and regression models are traditionally thought to be more interpretable, we aim to use interpretability methods for neural networks (Toms et. al., 2020) to determine if it is possible to obtain physical insight from the output. In particular, Layerwise Relevance Propagation (Barnes et. al., 2020) and saliency maps are two interpretability techniques that could aid this process; they highlight the grid cells that the trained neural network most relies on when making its prediction. We will study the climate dynamics of these grid cells to understand more deeply the rationale behind the neural network's predictions.

Third, we will compare the associated uncertainty of each method. Dynamical models leverage a multimodel ensemble, traditional statistical methods use block bootstrapping to estimate uncertainty in the calculated regression coefficients (Risser et. al., 2021), and neural networks utilize Bayesian methods, such as dropout (Gal and Ghahrahmani 2016).

Finally, we will compare methods based on their computational expense and their ability to support D&A studies on datasets with large spatial and temporal resolution.

How can ML-based D&A models learn from heterogeneous climate datasets?

Because ML models require large amounts of training data, we plan to train the GAN using both observations and climate simulations. Previous studies have successfully trained an ML model on a suite of climate simulations and applied that model to observations (Ham et. al., 2019). We propose exploring "domain adaptation" strategies (Wang et. al., 2018) to train a neural network on one set of domains, such as climate simulations and applying it to a new domain, such as observations. As the climate itself changes over the course of the next decade, training on climate simulations will be crucial to ensure that the GANs' performance remains constant on the changing distribution of observations.

What hardware/ software requirements are necessary to run ML-based D&A at scale?

In ML workflows, training a model (i.e. optimizing the weights of a neural network) is far more computationally expensive than inference (i.e. applying a trained model to a given input). As hybrid supercomputers with both GPU and CPU resources come online, we will explore the optimal allocation of CPUs and GPUs. We will use GPUs for training, as they provide a crucial speed-up in running the optimization algorithm. However, for inference, we plan to explore the feasibility of CPUs. While GPUs are still faster, existing CPU resources and their capability for massive parallelism, may offer satisfactory performance for inference, allowing GPUs to be preserved for training tasks. We plan to use the Toolkit for Extreme Climate Analysis, a software package that allows for parallelizing climate science workflows on high-performance computing resources, to conduct this comparison.

Additionally, GPU memory is a major bottleneck when training neural networks on high-resolution climate simulations. As we conduct D&A studies on varied datasets, we propose exploring the tradeoff between deeper neural networks (which may offer improved performance but require increased GPU memory) and shallower neural networks (which are more tractable for high-resolution climate datasets). To tackle this problem, we aim to explore methods to reduce the size of neural networks, such as pruning (Luo et. al., 2017), distilling the knowledge of an ensemble of neural networks (Hinton et. al., 2015), and optimizing hyperparameters such as the number of layers.

FAIR

We align our research with FAIR (Findable, Accessible, Interoperable, and Reusable) Principles, and we intend to open-source the three major components of our work: (1) our code, (2) our trained neural network models, and (3) a time series of the components of precipitation caused by each of the climate signals. These three components allow for replication of our results, extension of our trained neural network to new datasets, and the use of our calculated components as a benchmark, respectively. Through open-sourcing these components, we hope other scientists can readily build on our results without needing to repeat our analysis and processing. We will store our code on Github, and we will host our trained models and dataset on a public portal at NERSC.

References (Optional)

Allen, Myles R., and Simon FB Tett. "Checking for model consistency in optimal fingerprinting." Climate Dynamics 15.6 (1999): 419-434.

Barnes, Elizabeth A., et al. "Viewing forced climate patterns through an AI lens." Geophysical Research Letters 46.22 (2019): 13389-13398.

Barnes, Elizabeth A., et al. "Indicator patterns of forced change learned by an artificial neural network." Journal of Advances in Modeling Earth Systems 12.9 (2020): e2020MS002195.

Deser, C., et al. "Insights from Earth system model initial-condition large ensembles and future prospects." *Nature Climate Change* (2020): 1-10.

Gal, Yarin, and Zoubin Ghahramani. "Dropout as a bayesian approximation: Representing model uncertainty in deep learning." international conference on machine learning. PMLR, 2016.

Gillett, N. P., et al. "The Detection and Attribution Model Intercomparison Project (DAMIP v1. 0) contribution to CMIP6, Geosci. Model Dev., 9, 3685–3697." (2016).

Ham, Yoo-Geun, Jeong-Hwan Kim, and Jing-Jia Luo. "Deep learning for multi-year ENSO forecasts." Nature 573.7775 (2019): 568-572.

Hinton, Geoffrey, Oriol Vinyals, and Jeff Dean. "Distilling the knowledge in a neural network." arXiv preprint arXiv:1503.02531 (2015).

Kirchmeier-Young, Megan C., and Xuebin Zhang. "Human influence has intensified extreme precipitation in North America." Proceedings of the National Academy of Sciences 117.24 (2020): 13308-13313.

Lehner, Flavio, et al. "Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6." *Earth System Dynamics* 11.2 (2020): 491-508.

Luo, Jian-Hao, Jianxin Wu, and Weiyao Lin. "Thinet: A filter level pruning method for deep neural network compression." *Proceedings of the IEEE international conference on computer vision*. 2017.

McKinnon, Karen A., and Clara Deser. "Internal variability and regional climate trends in an observational large ensemble." Journal of Climate 31.17 (2018): 6783-6802.

Patricola, Christina M., and Michael F. Wehner. "Anthropogenic influences on major tropical cyclone events." Nature 563.7731 (2018): 339-346.

Risser, Mark D., et al. "A probabilistic gridded product for daily precipitation extremes over the United States." Climate Dynamics 53.5 (2019): 2517-2538.

Risser, Mark D., et al. "Quantifying the influence of natural climate variability on in situ measurements of seasonal total and extreme daily precipitation." Climate Dynamics (2021): 1-26.

Sippel, Sebastian, et al. "Climate change now detectable from any single day of weather at global scale." Nature climate change 10.1 (2020): 35-41.

Toms, Benjamin A., Elizabeth A. Barnes, and Imme Ebert‐Uphoff. "Physically interpretable neural networks for the geosciences: Applications to earth system variability." *Journal of Advances in Modeling Earth Systems* 12.9 (2020): e2019MS002002.

Thompson, David WJ, et al. "Quantifying the role of internal climate variability in future climate trends." *Journal of Climate* 28.16 (2015): 6443-6456.

Wang, Mei, and Weihong Deng. "Deep visual domain adaptation: A survey." *Neurocomputing* 312 (2018): 135-153.

Wills, Robert CJ, Sebastian Sippel, and Elizabeth A. Barnes. "Separating forced and unforced components of climate change: The utility of pattern recognition methods in Large Ensembles and observations." Variations 18.2 (2020): 1-10.