Advancing the Predictability of Water Cycle Phenomena via the Application of AI to Model Ensemble Simulations and Observations

Stephen A. Klein[1], Elizabeth A. Barnes[2], Céline Bonfils[1], Paul J. Durack[1], André Goncalves[3], Alex D. Hall[4], Jiwoo Lee[1], Hsi-Yen Ma[1], Gavin D. Madakumbura[4], Ana Ordoñez[1], Giuliana Pallotta[1], Stephen Po-Chedley[1], Mark D. Zelinka[1], Chengzhu (Jill) Zhang[5]

[1]PCMDI/LLNL, [2]Colorado State University,[3]Computation/LLNL, [4]UCLA, [5]E3SM Project/LLNL

Focal Area. Insight gleaned from complex data using AI, eXplainable AI ("XAI")

Abstract. AI has the power to identify new pathways to extended predictability of the water cycle via its application to model ensembles together with accompanying observations. We discuss both scientific and technological aspects of this challenge, and address the parts of the MODEX approach involving "Model simulations, evaluation, analysis, and benchmarking" and "Identification of key knowledge gaps". The widespread incorporation of AI into model analysis will significantly advance Earth system predictability and our predictive capabilities.

Introduction and Goal. It is well accepted that multi-model predictions are superior to those based upon a single model, and that the study of output from an ensemble of models is crucial to assessing uncertainty and identifying pathways to improved predictions. Superiority is achieved because bringing models together encompasses a greater collective wisdom while cancelling out random errors. Model skill in simulating past events is often correlated to the skill in simulating future events, allowing for the profound possibility that by studying multi-model datasets one can identify the processes most relevant to projections.

Given this immense utility, substantial resources have been expended to create multi-model datasets. For example, DOE led the creation over several decades of a multi-model dataset[1] whose analysis underpins the Intergovernmental Panel on Climate Change reports. Although thousands of papers have been written using this data, the data (totaling tens of petabytes) are inadequately explored relative to its underlying complexity. Adequate exploration of large datasets is a growing problem as future models will have much increased spatial resolution (global storm resolving models[2]) and increased high-frequency output to study finer scale features. Observational data are also growing in volume and their uncertainties have been inadequately explored.

The revolution in AI can address these issues in 2 ways. First, new methods (convolutional neural networks, transfer learning, etc.) permit "deep learning" and the discovery of complex structures within large datasets. Second, new XAI methods (layerwise relevance propagation, gradient methods, etc.) allow one to open the black box and peer inside. This is a game changer for scientific research as it means that XAI can be used to discover new science (e.g.[3]), allowing researchers to ask "why?" as they normally would, but now with the power of AI.

DOI: https://doi.org/10.2172/1769656
DOE BER, AI4ESP 2021 #AI4ESP1069 (https://www.ai4esp.org)

We assert that major advancements in the predictability of water cycle phenomena on time-scales from sub-seasonal to multi-decadal will result from the widespread application of AI to simulation output from coordinated multi-model ensembles and accompanying observations. **AI will identify new sources of predictability and understand their implications for model improvement.** Advancements will come from four primary **Analysis Strategies:**

1. AI-based prediction models: Trained on model ensemble simulations and convolved with observations, neural networks have already produced improved predictive skill for the El Niño-Southern Oscillation, a major mode of inter-annual variability that affects hydroclimate worldwide[4]. These prediction methods can be extended for a wide range of water cycle quantities such as the Madden-Julian Oscillation (MJO) or precipitation over the United States[5]. Importantly, AI derives skill from both multi-model simulations and observations through transfer learning (e.g.[6]). Hybrid AI-physics models are also possible, as is invoking physics throughout the training process itself. Improved uncertainty estimates are also possible through Bayesian approaches (e.g. dropout[7]). Physical credibility of these models must be probed with XAI that identifies the processes most constraining predictions.

2. Identification of processes relevant to predictability: Inter-model prediction differences are often correlated to inter-model differences in simulations of past quantities. When combined with observed values of the past quantities, model predictions can be constrained. These 'emergent constraints'[8] are difficult to locate and may be spurious, requiring confirmation of their physical basis. However, they are extremely valuable because they inform model developers and observationalists as to where their improvement efforts would have the greatest impact in reducing projection uncertainty. AI methods permit more extensive and deep searches for multi-variable and non-linear constraints. XAI is also better suited in identifying the processes underlying the relationships and avoiding identification of spurious constraints[9].

3. Signal Separation: On both long (decadal) and short-time scales (individual events), it is desirable to distinguish the multiple influences (greenhouse gases vs. aerosols, or natural vs. human) leading to water cycle fluctuations and to separate these forced signals from internal climate variability. Such separation is an essential step in assessing models and knowing how much confidence to place in their predictions. XAI methods, including regularized neural networks, can be used to separate signals in both time, space and across variables by leveraging information across all of these domains and identifying the patterns that distinguish the different drivers of the signals. Indeed, XAI's advantage in identifying the nonlinear evolution of a signal has allowed for the detection of the human impact on extreme precipitation[10].

4. AI-based metrics: Metrics are scalar measures of model skill that are extensively used in model development to rank models. Water cycle extremes, such as atmospheric rivers, tropical cyclones, or blocking anticyclones (that trigger droughts over land), often involve complex interactions between water-cycle processes and the environment. This complexity requires

advanced methods to fully characterize the relevant phenomena. AI methods (e.g. clustering techniques) will lead to better metrics of model performance and identify pathways for model improvement[11]. AI methods could discover new relationships between performance metrics and predictions, as well as determine the extent to which models are unique and independent.

Water Cycle Quantities Most Amenable to Advances in Predictability. The water cycle's spatio-temporal complexity guarantees that AI will be useful for numerous quantities, from extreme events to large-scale fluctuations (ENSO, MJO, droughts, ...). Model diversity in the simulation of water cycle is very large, resulting in part from the difficulty in representing unresolved water cycle processes. This offers an opportunity for AI to help both through multi-model analyses revealing new pathways to increased predictability and through AI-based metrics identifying which models better represent water-cycle phenomena and which model processes should be improved. AI excels at finding relationships and patterns in large data particularly when our understanding is incomplete and theory lacking (e.g., what determines the number of hurricanes each year?). Other areas ripe for advancement include identifying multi-realm relationships, such as between the ocean, atmosphere, land-surface, and cryosphere (e.g., between precipitation over land and ocean sea-level height[12] and sea-surface salinity[13]). AI can also help identify the multi-realm precursors for complex compound events (such as 'flash droughts').

Technical Challenges Related to the Routine Application of AI to Big Data. The ever-expanding quantity of model and observational data is rapidly becoming too large to be analyzed by the ordinary resources that scientists use to retrieve, store, query, and analyze data. Such challenges may block the major advances in the predictability that are possible from pairing big data with advanced AI. Therefore, it is imperative to provide access to big data computing resources including: (a) cloud computing allocations (*bring the compute nodes to the data*), (b) open-source software to apply AI to big data, and (c) software and hardware to handle memory intensive algorithms and parallel computing (*scalable machine learning*). Developing centralized computing facilities will enhance the speed with which individual scientists can perform sophisticated analyses. Community computing facilities also enable users to leverage one another's insights in tackling existing AI challenges and developing analysis software. This will reduce time-consuming and duplicative technical hurdles.

Furthermore, the importance of effectively deploying AI in the production environments for testing new models is often overlooked. Most machine learning applications evaluating models are developed once for a research paper but are never used again to evaluate future models. A framework is needed to interface the most promising AI metrics with the established software for repeat-use calculations that assist model development. Incorporating AI into DOE-funded tools such as E3SM-diags[14], PCMDI's Metrics Package[15], and the Coordinated Model Evaluation Capabilities[16] among others, will greatly strengthen the evaluation capability of these tools and engage AI in the development of future model versions including those of DOE's E3SM model.

Suggested Expert

Professor Elizabeth A. Barnes is willing to provide a presentation at a workshop.

References

1. Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.
2. Stevens, B., M. Satoh, L. Auger, and co-authors, 2019: DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Prog. Earth Planet Sci.*, **6**, 61, https://doi.org/10.1186/s40645-019-0304-z.
3. Toms, B. A., E. A. Barnes, and I. Ebert-Uphoff, 2020: Physically interpretable neural networks for the geosciences: Applications to Earth system variability. *J. Adv. in Mod. Earth Sys.*, **12**, e2019MS002002, https://doi.org/10.1029/2019MS002002.
4. Ham, Y. G., J. H. Kim, and J. J. Luo, 2019: Deep learning for multi-year ENSO forecasts. *Nature*, **573**, 568–572. https://doi.org/10.1038/s41586-019-1559-7.
5. Pan, B., K. Hsu, A. AghaKouchak, and S. Sorooshian, 2019: Improving precipitation estimation using convolutional neural network. *Water Resources Research*, 55, 2301–2321, https://doi.org/10.1029/2018WR024090.
6. Rasp, S., and N. Thuerey, 2021: Data-driven medium-range weather prediction with a Resnet pretrained on climate simulations: A new model for WeatherBench. *J. Adv. in Mod. Earth Sys.*, https://doi.org/10.1029/2020MS002405.
7. Gal, Y., and Z. Ghahramani, 2015: Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. arXiv:1506.02142v6 [stat.ML].
8. Hall, A., P. Cox, C. Huntingford, and S. Klein, 2019: Progressing emergent constraints on future climate change. *Nature Climate Change*, **9**, 269–278, doi: 10.1038/s41558-019-0436-6.
9. Schlund, M., A. Lauer, P. Gentine, S. C. Sherwood, and V. Eyring, 2020: Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6? *Earth Syst. Dynam.*, 11, 1233-1258, https://doi.org/10.5194/esd-11-1233-2020.
10. Madakumbura, G. D., C. W. Thackeray, J. Norris, N. Goldenson, and A. Hall, 2021: Anthropogenic influence on extreme precipitation over global land areas seen in multiple observational datasets, 10 February 2021, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-227967/v1].

11. Prabhat et al., 2021: ClimateNet: an expert-labeled open dataset and deep learning architecture for enabling high-precision analyses of extreme weather, *Geosci. Model Dev.*, **14**, 107–124, https://doi.org/10.5194/gmd-14-107-2021.

12. Boening, C., J. K. Willis, F. W. Landerer, R. S. Nerem, and J. Fasullo, 2012: The 2011 La Nina: So strong, the oceans fell. *Geophys. Res. Lett.*, **39**, L19602, https://doi.org/10.1029/2012GL053055.

13. Li, L., R. W. Schmitt, C. C. Ummenhofer, and K. B. Karnauskas, 2016: North Atlantic salinity as a predictor of Sahel rainfall. *Sci. Adv.*, **2**, e1501588, https://doi.org/10.1126/sciadv.1501588.

14. https://e3sm.org/resources/tools/diagnostic-tools/e3sm-diagnostics/

15. https://pcmdi.llnl.gov/research/metrics/

16. https://cmec.llnl.gov

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. LLNL-MI-819493.