Develop a weather-aware climate model to understand and predict extremes and associated power outages and renewable energy shortages with uncertainty-aware and physics-informed machine learning

Authors

Jiwen Fan[1], Zhangshuan (Jason) Hou[1], Joseph Hardin[1], Paul O'Gorman[2], Jessika Trancik[2], John Allen[3], Peeyush Kumar[4], Ranveer Chandra[4], Jingyu Wang[1], Lai-Yung (Ruby) Leung[1]
[1] Pacific Northwest National Laboratory; [2] Massachusetts Institute of Technology; [3] Central Michigan University; [4] Microsoft Corporation

Focal Area(s)

The focus area is predictive modeling through the use of AI techniques and AI-derived model components with a particular emphasis on extreme weather in Atmospheric Science and power outages and shortages in Energy Science.

Science Challenge

Predicting weather extremes (e.g., heavy precipitation, strong wind, and large hailstones), and weather-related power system outages and shortages can mitigate economic losses, save lives, support renewables integration, and improve power system resiliency. However, currently, the poor reliability and large uncertainty associated with the weather extreme prediction in the current climate models make the problem intractable. The key challenges are: (1) physical factors like green-house gases (GHGs), aerosols, and land use and land cover (LULC) can significantly impact extreme storms, but the understanding of these impacts is limited, particularly globally; (2) the convective permitting resolutions needed to model severe convective storms and their impacts are computationally prohibitive with global climate models (GCMs); (3) interactions between weather extremes and power system outages are complex and subject to great uncertainty. Current outage prediction models are short lead (~ 3 days), which do not allow for long-time planning of energy production and distribution. Moreover, we have limited capacity to predict weather events leading to sustained shortages in a renewable-energy-dominated power system. These challenges drive motivation for mechanistic understanding and reliable and efficient predictive modeling of extremes and their impacts from the sub-seasonal to long term projections.

Rationale

The major barriers to progress are:

- Drivers impacting severe convective storms are poorly understood due to (a) limited knowledge of sub-grid processes (e.g. convection, microphysics, turbulence, and land surface) and interactive processes [e.g. cloud microphysics feedback to dynamics, aerosol-cloud interactions (ACI), and land-atmosphere interactions (LAI)], that shape storm properties, (b) various impacts of physical factors like GHGs, aerosols, and LULC, and (c) small samples of extreme events make it difficult to gain a robust understanding of a hierarchical system of weather extremes.
- Traditional approaches to address model biases are ineffectual due to system complexity and process nonlinearity. Improving physics is particularly challenging since microphysics schemes suitable for ACI and land schemes with resolved urban physics for LAI are not practical in GCMs. Preliminary studies have employed artificial intelligence (AI) and machine learning (ML) for subgrid processes in both the cloud-resolving model (CRM) or convection-permitting model (CPM) simulations for GCMs. However, the employed CRM/CPM simulations were typically highly idealized (e.g. aquaplanets) and did not include anthropogenic signals, ACI, and LAI interactions. These processes are critical given large increases in anthropogenic aerosols and predicted ~ 6 times urban land expansion by end of the century. Also, those ML approaches have considerable potential at finer grid spacings (25 km and below) since they can be trained specifically for a given grid spacing.

DOI: https://doi.org/10.2172/1769695
DOE BER, AI4ESP 2021 #AI4ESP1041 (https://www.ai4esp.org)

Develop a weather-aware climate model to understand and predict extremes and associated power outages and renewable energy shortages with uncertainty-aware and physics-informed machine learning

Advances in AI/ML can help make these problems tractable, by developing and integrating multiple nonlinear ML algorithms (e.g., classification, explainable AI, physics-informed ML, generative networks) for system simplification, parameter ranking, calibration, and data augmentation.

We identify the following research needs to surmount the aforementioned challenges with AI/ML: **(1)** Using ML for observational data reconstruction, extreme event data labeling, and augmentation, to obtain enough sample sizes for robust ML; **(2)** Building an ML-integrated GCM to enable simulation of extreme storms and associated weather hazards from the sub-seasonal to long-term projection scales, by dynamically adjusting model dynamics, thermodynamics, and physics using ML parameterizations. This would efficiently account for observations and the effects of sub-grid processes (convection, microphysics, and turbulence) and interaction processes (i.e., ACI and LAI) through high-resolution simulations with explicit physics; **(3)** Improving understanding drivers of storms at various scales using the classification of large-scale and synoptical-scale environments by integrating high-resolution model simulations (CRM/CPM scales), observations, and reanalysis with physics-informed interpretable ML. **(4)** Building an ML outage and renewables shortage prediction model from observations and reanalysis capable of application to GCMs, allowing (a) improved understanding of interactions between weather extremes and outage/shortage prediction, and (b) predicted likelihood of power outages and renewables shortage events (e.g., spatiotemporal frequency, magnitude, duration) under a changing climate for both individual (e.g., wildfires, damaging winds) and joint likelihoods of these hazards.

Narrative

Understanding, characterizing, and modeling extreme events are critical for weather and climate predictions, infrastructure design, strategic planning, and minimizing losses to the economy and life. Extreme weather phenomena have caused approximately ~80% of large-scale power outages, and increased frequency and intensity of severe weather under climate change will increase the likelihood of infrastructure outages and shortages. Weather-aware climate models are needed to understand weather extremes/hazards and their impacts under a changing climate to plan and prioritize the power system in the future. Modern ML/AI techniques can accelerate the progression to the weather-aware climate models. These techniques can assist in overcoming (a) observational limitations for extremes that preclude a robust understanding of key drivers or distinguishing anthropogenic forcing from natural variability, (b) computational limitations for resolving important physics processes (convection, microphysics, turbulence, and interaction processes of ACI and LAI) that drive large errors in simulating storms and weather extremes.

We describe the four research needs as detailed above. The innovative aspects are:

(1) Historical extreme reconstruction using ML/AI, addressing a major predictability barrier. ML/AI has shown promise in reconstructing the missing data and the historical data before observation record from related observations instrumentation. Applying ML/AI to reconstruct historical weather extremes is unique, and different techniques such as seq2seq/style transfer or low-dimension predictors are needed. This allows (a) expansion of the record, particularly globally where measurements of extremes are sparse. A more complete record is necessary to gain a robust understanding of physical process-weather extreme interactions and to isolate anthropogenic contributions and (b) providing the samples necessary for different storms/extremes to be classified based on large-scale and synoptic-scale environments using ML to obtain a unique physically relevant systematic understanding and model evaluation.

(2) Development of an ML-integrated weather-aware GCM to simulate extreme storms and associated hazards at a range of time scales. Integrating ML/AI with DOE's climate model E3SM to enable accurate model simulations of storms/extremes at the sub-seasonal and longer leads. ML (e.g., physics-informed neural networks) are needed to not only reduce the computation of the detailed physical models and enable fast, global projections with sub-grid processes considered, but more realistically account for the impacts of physical factors on storms to improve model accuracy. Such impacts are not considered in the previous ML-parameterizations for subgrid processes in GCMs, and these processes are key to reduce uncertainties in the prediction of extremes/hazards. Existing approaches have focused on coarse GCMs, and have excluded subgrid variabilities (e.g. peak wind speeds and precipitation rates) critical to evaluating the impacts on power outages and shortages. Innovatively, we believe targeted ML models for identifying sources of biases and correcting the biases based on clustering of storm systems can provide more accurate predictions. A unique model-data integration can be done through interpretable ANNs to identify systematic model biases and highlight possible observational constraints to reduce the model uncertainty. The accuracy of the ML-integrated E3SM can be further improved in future with the a quantum computing method to accurately solve differential equations in physics-based ML solutions. The quantum methods have been observed to provide higher accuracies in solving differential equations in dynamical systems.

(3) Development of uncertainty-aware, physics-informed neural networks to understand drivers and predictability. Traditionally neural networks have struggled with interpretability and capability to accurately represent uncertainty. Developments in scalable uncertainty-aware physics-informed neural networks have provided paths for both limitations. These techniques, e.g., Bayesian Networks, Neural PDEs, Graph-CNN, Bootstrap DQNs, and dropout NNs, have found wide applicability, and offer promises to address cloud microphysics and atmospheric turbulence parameterizations that include ACI and LAI. Developments in multi-modal machine learning combining weather data and storm feature maps also enable extreme event understanding at high spatial resolutions. Physically informed or constrained ML architectures have key benefits in providing out-of-sample resiliency, training efficiency, and responsiveness to rare events, contrasting correlation-based or causative relationships. These works will improve our understanding of the drivers affecting different types of storms at different time scales, which can be employed to evaluate the ML-integrated GCM illustrated in (2) and further improve model physics.

(4) Linking extreme weather to power outage and renewable shortage prediction. ML methods can allow for post-processing of the weather-aware E3SM data to impact relevant predictions for outages and renewable shortages. The development for weather extreme prediction as described in (2) leads to the potential for such predictions using the GCM data. The ML outage and shortage prediction model can be developed based on observation and reanalysis data and then applied to GCMs. This would (i) allow increased understanding of weather-extreme and outage prediction, and exploration of multi-factor demand and outages (e.g., a heatwave increasing demand and wind-induced outages), and (ii) expand DOE modeling capacity to effectively predict the likelihood of outages and renewable shortage in current and future climates to enhance energy resilience.

Overall, the recommended research would deal with huge volumes of high-resolution observations (e.g. radar, satellite) and model data, which necessitates leveraging the latest 5G and Edge computing to manage and train large amounts of data on the cloud for developing effective and accurate AI models. Also, part of 5G signals operate in mmWave, and the signal propagation between 5G base stations can indicate weather (particularly temperature and precipitation), potential datasets for improving model accuracy. Using AI to integrate data from weather and power stations, radar, satellite, 5G signals to help training high-quality AI models is feasible for us with PNNL 5G lab and PNNL's collaboration with Verizon.

Codes will be open access through Github. Findings will be published via the open-access option. Reproducibility of code to enable integration with various open-source weather-related service frameworks such as ESRI will be made through packaging as a cloud-based dockerized service.