

Kyle J. C. Hall

(908) 358-3176 | hallkjc01@gmail.com | <https://kjhall01.github.io/>

Software Engineer and Data Scientist building the next generation of climate forecasting tools in order to increase society's resilience against climate variability and change.

Relevant Experience

International Research Institute for Climate and Society (IRI), Columbia University Palisades, NY
Staff Associate I February 2021 – Present

- Primary software engineer and lead designer for the development of IRI's "Python interface to the Climate Predictability Tool (CPT)", PyCPT.
- Leading the release of the second major version of PyCPT while serving as the main point of contact for users of the current operational version.
- Supporting international partners in the development and implementation of operational climate forecasts using PyCPT by traveling to partner countries and leading capacity building sessions.
- Data Science subject matter expert for the development of grant applications and research proposals. Advising on machine learning theory and implementations with emphasis on state-of-the-art deep learning frameworks like Keras, Tensorflow, PyTorch, and high-performance computing tools like Dask.
- Data Engineer for the IRI Data Library's North American Multimodel Ensemble (NMME), S2S, Copernicus Climate Change Service (C3S), and SubX data pipelines. Incorporating modern task management tools like Airflow into IRI workflows. Maintaining a broad knowledge of the climate models under management by performing literature review and assisting users with the NMME, C3S, SubX and S2S datasets.
- Developed XCast (<https://kjhall01.github.io/xcast>), a high-performance, parallelized Python climate data science library built on Dask and Xarray. Iteratively co-designed XCast by consulting with climate scientists and community stakeholders to identify needs and pain points present in the Python climate data science ecosystem.
- Documented and profiled XCast's performance and implemented climate forecasting case studies. Presented results at conferences and published peer-reviewed articles.

Casual Hire September 2020 – February 2021

- Built new features for PyCPT and supported remote capacity-building sessions on PyCPT climate forecasting with international meteorological groups. Served as a point of contact for PyCPT users.

Graduate Research Assistant - ACToday Vietnam October 2019 – May 2020

- Assisted with debugging the PyCPT code base, and implemented new types of climate data visualizations.

Segal Consulting, National Health Technical Services (NHTS) New York, NY
Health Benefits Data Analyst June 2018 – May 2019

- Assisted clients in developing data-driven health benefits strategies by producing tailored reports and projections.
- Developed a strategy for predicting diabetes diagnoses in pre-diabetic plan members using machine learning.

National Health Technical Services Intern Summer 2016

- Reported on client healthcare cost trends by querying Segal's Health Data Warehouse using Python and SQL.

NASA Goddard Space Flight Center Greenbelt, MD
Earth Sciences Directorate Intern Summer 2017

- Explored ways to use machine learning to increase the speed of atmospheric radiative transfer simulations.
- Trained in High-Performance Computing in Python and Fortran during a University of Virginia summer program.

Education

Columbia University in the City of New York New York, NY
Master of Arts in Climate & Society February 2021

National Security Education Program (Remote) Madison, WI & Malang, Indonesia
David L. Boren Fellowship Recipient - Indonesian Flagship Language Initiative Fall 2020

The College of William & Mary Williamsburg, VA
Bachelor of Science in Computer Science, Economics Minor May 2018

Skills

Climate Data Science

- Manipulating and analyzing gridded raster data and satellite imagery using Python tools like Xarray, Dask, Scikit-Learn, and state-of-the-art deep learning frameworks like PyTorch, Tensorflow, and Keras.

Software Development & High Performance Computing

- Supporting team members through code review and partner programming.
- Driving innovation with design thinking and rapid prototyping.
- Designing, Implementing, Documenting and Distributing software using Git Version Control and Anaconda.

Publications

K. Hall & N. Acharya. (2022) XCast: A Python Climate Data Science Toolkit. *Frontiers in Climate*. {DOI pending publication} Pending Publication (2022)

K. J. C. Hall & N. Acharya. 2022. *Introducing XCast: A High-Performance Data Science Toolkit for Climate Forecasting. In Proceedings of NCAR UCAR Software Engineering Assembly Improving Scientific Software Conference (ISS'22)*. ACM, New York, NY, USA Pending Publication (2022)

Acharya et. al., *On the next generation (NextGen) seasonal prediction system to enhance climate services over Ethiopia, Climate Services*, <https://doi.org/10.1016/j.cliser.2021.100272>. December 2021

ECMWF Challenge to improve Sub-seasonal to Seasonal Predictions using Artificial Intelligence - POELM <https://renkulab.io/gitlab/kjhall01/s2s-ai-challenge-kjhall01> October 2021

Acharya, N., & Hall, K. (2021). *PyELM-MME: A python platform for extreme learning machine based multi-model ensemble. Proceedings of the 2021 Improving Scientific Software Conference p.1-4*. <https://doi.org/10.26024/p6mv-en77> July 2021

Conference Presentations

NCICS Climate Informatics 2022 - *Learning to construct Probabilistic Multi-Model Ensemble for Seasonal Predictions* May 2022

NCAR UCAR SEA Improving Scientific Software Speaker - *XCast: A High-Performance Python Data Science Toolkit for Climate Forecasting* - <https://github.com/kjhall01/xcast> April 2022

ECMWF Machine Learning Workshop- *Sub-Seasonal Probabilistic Precipitation Forecasting using Extreme Learning Machine* March 2022

Seventh WMO International Workshop on Monsoons (IWM-7) - *A Machine Learning Approach for Probabilistic Multi-Model Ensemble Predictions of Indian Summer Monsoon Rainfall* March 2022

IRI Seminar: *Redesigning the IRI's Python Interface to the Climate Predictability Tool (PyCPTv2)* February 2022

IRI Seminar: *NextGen Ensemble Forecasting Training-of-Trainers* October 2021

NCAR UCAR SEA Improving Scientific Software Speaker - *PyELM-MME: A Python Platform For Extreme Learning Machine (ELM) based Multi-Model Ensemble* March 2021

Research Interests

Modeling the Earth System and Climate with Machine Learning

- Improving statistical post processing of climate model forecasts using both deep learning and lightweight, explainable AI methods.
- Learning the physical processes of the climate with statistical and machine learning tools.
- Producing actionable insights from satellite images and observations with traditional and deep learning-based Computer Vision techniques.

Climate Services & Capacity Building

- Implementing operational forecasts using state-of-the-art technology and science.
- Empowering stakeholders to access, analyze, interpret and communicate climate information themselves.