Kyle J. C. Hall

(908) 358-3176 | hallkjc01@gmail.com

A climate data scientist seeking to increase society's resilience against climate change and climate variability by contributing to the improvement of quantitative climate forecasting and climate modeling

Relevant Experience

International Research Institute for Climate and Society (IRI), Columbia University **Palisades, NY**

Staff Associate I *February 2021 – Present*

- Currently serving as principal architect of the second major version of the IRI's Python interface to the Climate Predictability Tool (CPT), PyCPT Version 2. Implementing all component Python modules.
- Active manager of the IRI Data Library's North American Multimodel Ensemble (NMME), S2S, Copernicus Climate Change Service (C3S), and SubX data pipelines. Maintained a broad knowledge of the models under management by attending monthly NMME telecons and performing literature review.
- Served as the primary point of contact for users requesting assistance with NMME, C3S, SubX and S2S datasets.
- Original designer and active maintainer of CPT-Tools, an Intake-based Python API for cataloging and downloading General Circulation Model (GCM) forecasts and observation datasets from the IRI Data Library.
- Designer and developer of CPT-Core, a Python library for concurrently executing fortran-based statistical GCM post-processing in CPT. Contributed to the development of CPT.
- Machine Learning subject matter expert during proposal writing, and Python programming expert during climate forecasting capacity building efforts with international meteorological groups in East Africa and South Asia.
- Packaging and distributing PyCPT V2 code with Anaconda, and managing IRI Climate Group software development on Github using a Pull Request workflow.

Casual Hire *September 2020 – February 2021*

- Built new features for the original version of PyCPT, including cross-platform support and testing automation.
- Supported remote capacity-building sessions on PyCPT climate forecasting with international meteorological groups by assisting the primary trainer and providing ad hoc PyCPT support.

Graduate Research Assistant - ACToday Vietnam *October 2019 – May 2020*

- Built new features for PyCPT, including probabilistic and deterministic climate forecast visualizations.

Segal Consulting, National Health Technical Services (NHTS) **New York, NY**

Health Benefits Data Analyst *June 2018 – May 2019*

- Collaborated with the NHTS team to help clients to develop data-driven health benefits strategies.
- Developed a strategy for predicting diabetes diagnoses in pre-diabetic plan members using machine learning.

National Health Technical Services Intern *Summer 2016*

- Performed analysis and generated reports on client data in Segal's Health Data Warehouse using Python and SQL.

NASA Goddard Space Flight Center **Greenbelt, MD**

Earth Sciences Directorate Intern *Summer 2017*

- Applied machine learning to increase the speed of atmospheric radiative transfer simulation with high accuracy.
- Performed Atmospheric Radiative Transfer Simulator (ARTS) experiments under guidance from staff scientists.
- Completed the University of Virginia's Advanced Computing for Earth Sciences (ACES) summer program.
- Trained in High-Performance Computing in Python and Fortran using the UVA supercomputer cluster.

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*

Publications

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*

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*

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

Conference Presentations

NCICS Climate Informatics 2022 - Learning to construct Probabilistic Multi-Model Ensemble for Seasonal Predictions *(Upcoming) 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 *(Upcoming) April 2022*

ECMWF Machine Learning Workshop- Sub-Seasonal Probabilistic Precipitation Forecasting using Extreme Learning Machine *(Upcoming) 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 *(Upcoming) 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 machine learning-based non-parametric probabilistic multi-model ensembles and bias correction.
- Using supervised machine learning models to approximate the physical processes of the climate, and determining how the trained models can be interpreted physically.

Enhancing Scientific Computing

- Optimizing computationally expensive climate models and physical simulations by implementing parallelism and out-of-core computing.
- Emulating climate models and physical simulations using deep learning to improve scalability and portability.
- Eliminating logistical computing problems for researchers by distributing research code on Github and industry-standard package managers like Anaconda.

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.

Skills

Climate Data Science

- Manipulating and analyzing gridded data using Python tools like Xarray, Dask, and Scikit-Learn.
- Post-Processing gridded climate data to produce climatologies, skill maps, spatial loadings and other metrics.
- Visualizing gridded climate data and building maprooms using Python tools like Cartopy, Dash and Plotly.

Software Development & High Performance Computing

- Designing, Implementing, Documenting and Distributing Python code using Git Version Control and Anaconda.
- Working with compiled code on institutional cluster computing resources using tools like SLURM and Dask
- Leading and administrating software development workflows.