Ying Li

https://yingli9718.github.io/ yingli.atmos@gmail.com
www.linkedin.com/in/yingli-atmos/ (609)-356-6330
US Permanent Resident Vienna, VA

Highlights

- 15+ years experience in machine learning, big data analysis, visualization, scientific and predictive modeling
- Strong verbal communication and writing skills with 50+ conference presentations and 20+ papers in top journals
- Good collaboration skills with many global collaborators on different projects
- AWS Certified Cloud Practitioner, Machine Learning (Coursera), Data Scientist with Python (26 courses on Datacamp)

Experience

- **Colorado State University**
  Research Scientist II *Aug. 2016 - Present*
  - Led a 4-yr project funded by NSF to study the two-way interaction between clouds and large-scale atmospheric circulation and published 8 lead author papers in top academic journals
  - Successfully brought in ∼$1,000,000 in funding from NSF and NASA by leading projects as principal investigator
  - Mentored one PhD student and served as thesis committee member
  - Analyzed and diagnosed periodic patterns in data from observations, idealized conceptual models using techniques including time and space filtering, principal component analysis (PCA), power spectral analysis, regressions. Discovered a robustly periodic behavior that can be used for climate predication
  - Designed and performed numerical experiments by changing the radiative component of clouds in the idealized and coupled general circulation models to investigate the feedbacks from clouds on the climate. Provided directions for reducing the model biases
- *Postdoctoral Scientist* *Jan. 2012 - Jul. 2016*
  - Investigated contributing factors on the formation and variation of cloud vertical structures by studying and analyzing ∼400 millions remotely sensed profile datasets with advanced statistical and data mining approaches. Provided a baseline for evaluating model physical parameterizations in comprehensive climate models
  - Performed large ensemble of high-resolution global climate model simulations on high-performance computing clusters at National Center for Atmospheric Research (NCAR). Improved understanding of the role of the ocean circulation in extratropical climate variability
  - Quantified and visualized inter-model spread in circulation responses to global warming in multi-model climate change simulations. Provided the uncertainty attributions to climate change simulation
- **Princeton University**
  Research Assistant *Sep. 2006 - Nov. 2011*
  - Quantified the relationships between the large-scale atmospheric patterns using techniques including regression, time series analysis, PCA, based on the 2000-yr, high-dimensional, tera-byte scale data from numerical simulations with Earth system models. Provided comprehensive picture of the global teleconnectivity
  - Designed and performed numerical simulations running on high-performance computing clusters at Geophysical Fluid Dynamics Laboratory (GFDL). Provided evidences on the relative roles of surface boundary conditions over the particular sectors in the globe

Side Projects

- Build a COVID-19 dashboard to visualize and track the spread of the virus using Python and Tableau
- Build machine learning models using Python scikit-learn to do sentiment analysis on Amazon review data
- See more at https://yingli9718.github.io/side_project.html

Education

- **Princeton University** *Ph.D. in Atmospheric and Oceanic Sciences*
- **Zhejiang University** *B.S. in Atmospheric Sciences*

Technical Skills

- **Programing tools**: Python (scikit-learn, NumPy, Scipy, Pandas, TensorFlow, Keras, Plotly, BeautifulSoup), R, Fortran, C, SQL, Git, PostgreSQL, MATLAB, Shell scripting, HTML, CSS, LaTex
- **Machine Learning tools**: Linear/Logistic Regression, Classification, Decision Trees, GBM, XGBoost, Random Forest, Neural Networks
- **Statistics**: Regression, Bayesian and Monte Carlo methods, Time Series, PCA, A/B testing
- **Data Visualization tools**: Tableau, Matplotlib, Seaborn, Plotly
- **Big Data & Hadoop**: HDFS, MapReduce, Hive