

# Yu Cheng

fischcheng@gmail.com • +1 (786) 503-3366 • <https://fischcheng.github.io>  
Seattle, WA, USA

## SUMMARY

Ph.D. climate modeler with expertise in processing, analyzing and visualizing largescale datasets using open-source softwares.

## SKILLS

- Data Analysis/Machine Learning - Python: pandas, statsmodels, numpy, scipy, xarray, jupyter-notebook, scikit-learn, BeautifulSoup, PySpark; Matlab, SQL, PostgreSQL, R, CDO (climate data operators), NCL (NCAR command language)
- Data Visualization - Python: matplotlib, cartopy, bokeh; NCL, Matlab
- General programming - Python, Fortran, UNIX shell, Git, Docker, Node.js

## PROFESSIONAL EXPERIENCE

**RSMAS, University of Miami**, Miami, Florida, USA

Research Assistant

Aug 2012 – May 2018

- Studied Agulhas leakage, its variability and regional climate impacts, using an advanced high-resolution global coupled climate model with ocean eddy-resolving capability.
- Developed a strategy to quantify Agulhas leakage in the model by releasing and tracking over  $10^6$  virtual particles, resulting in 3 peer-reviewed articles. Used the libraries including matplotlib, cartopy and seaborn to prepare publication-quality figures.
- Used libraries: xarray, numpy and scipy to process and analyze model outputs ( $10^2$  Tb) using techniques such as filtering, cross-spectral analysis, cross-correlation and empirical orthogonal functions.
- Analyzed and visualized Agulhas leakage time-series using Python libraries including pandas and statsmodels, developed a bootstrapping significance test for linear regression between different climate indices.

**Consortium for Climate Change Study**, Academia Sinica, Taipei, Taiwan

Research Assistant

Sep 2011 – Jul 2012

- Collaborated with staffs to develop shell scripts to systematically archive and manage the Coupled Model Inter-comparison Project Phase 5 (CMIP5) model outputs to local servers.
- Processed climate model outputs, such as removing seasonal cycle, calculating anomalies, and climatology mean, using Fortran and CDO.

## PROJECTS

**Automated pipeline to visualize satellite observed ocean surface conditions**

Fall 2013

- Coordinated with two research cruises to create an automated pipeline to visualize near-realtime ocean surface conditions using both MODIS satellite observations and RTOFS ocean forecasts.
- The python-written pipeline downloads data in the *netcdf* format from government websites, plot the data using NCL, and update the plots to a shared dropbox folder, accessible by the chief scientists onboard.

**Visualizing leakage particle pathways**

Fall 2014

- Created a movie to show when and where millions of virtual particles, tagged with volume transports, cross the GoodHope line, using tools including imagemagick and python-matplotlib.

**Recommender system for tennis racquets and strings**

Ongoing

- Building a recommender system based on user and retailer reviews of tennis racquet and strings, using dataset scrapped from multiple websites including Stringforum and Tennis-warehouse.
- Objective: provide a best combination of tennis racquet and string when one provide his/her last three or five setups.

## EDUCATION

**RSMAS, University of Miami**, Miami, Florida, USA

Ph.D. in Meteorology and Physical Oceanography

Aug 2012 – May 2018

- Thesis: Agulhas leakage variability and its climate implications in a coupled system

**National Taiwan University**, Taipei, Taiwan

B.S. in Atmospheric Sciences

Sep 2006 – Jun 2010

- Relevant courses: Applied Mathematics I (Linear Algebra) and II (Ordinary Differential Equations), Statistic for Atmospheric Sciences, Applied Statistics, Scientific Computing (Fortran and Matlab)

**MOOC courses**

Sep 2017 – Present

- Machine Learning (Coursera), MySQL bootcamp (Udemy), Machine learning with Python skill track (four courses including supervised, unsupervised, deep learning and a case study), Parallel Computing with Dask, Time-Series analysis, Intro to PySpark, Statistical thinking (Datacamp), Completed 35% of Dataquest.io missions (including 7 guided projects)