Summer Internship Opportunity: Work with NOAA Scientists and Learn to Integrate Mathematics and Fisheries Science

The Northwest Fisheries Science Center (NWFSC) and the University of Washington request applications for students in the Mathematical Sciences for a summer internship at the NWFSC. Interns will spend summer (~16 June – 15 September) working on a research project that integrates mathematics with the science that informs fishery managers. A stipend of $6,300 will be provided from the Usha and S. Rao Varanasi SAFS Faculty Endowment for Student Support, the NWFSC, and the School of Aquatic and Fishery Sciences (SAFS). The successful applicant will also be provided with office space at the NWFSC or SAFS and a NWFSC mentor.

Although any projects related to sustainable management of west coast fish resources would be considered, the following projects are already available and mentors identified, which each of these projects looking for two or more interns this year:

1. **Combining survey data to quantify spatiotemporal variation in fish populations across the Northeast Pacific Ocean.**
   Mentors: Eric Ward (NWFSC), Kelli Johnson (NOAA Fisheries, Office of Science and Technology), Kiva Oken (NWFSC), Chantel Wetzel (NWFSC), Sean Anderson (Fisheries and Oceans Canada), Lewis Barnett (Alaska Fisheries Science Center).

2. **Development of machine- and deep-learning models for processing remote-sensing data.**
   Mentor: Eli Holmes (NWFSC).

ELIGIBILITY
Must be a currently enrolled UW (Seattle or Tacoma campus) undergraduate student graduating in Spring 2025 or after

HOW TO APPLY
To apply for this internship, submit your application to this form https://forms.gle/fFAvrvxT3MGEF8nt9 by March 15, 2025.

- Application Materials (in one pdf). Save as "LastnameFirstname_MML2025.pdf" (where Lastname and Firstname are your name)
  - o Recent Resumé
  - o Unofficial UW Transcript
  - o Letter of Interest (maximum of four pages) – include the name of the project that most interests you and why; tell us about yourself and your research interests; explain how the internship will further your studies and career; include other information the selection committee should be aware of, such as what it means to you to have a commitment to diversity, equity, and inclusion.

DEADLINE FOR SUBMISSION
March 15, 2025

DECISIONS
Award notifications will be made by April 25, 2025

The University of Washington is an affirmative action and equal opportunity employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, sexual orientation, gender identity, gender expression, national origin, age, protected veteran or disabled status, or genetic information.

Project 1: Combining survey data to quantify spatiotemporal variation in fish populations across the Northeast Pacific Ocean

Mentors: Drs. Eric Ward, Kelli Johnson (S&T), Kiva Oken, Chantel Wetzel, Sean Anderson (DFO), Lewis Barnett (AFSC)

Along with collaborators from the Alaska Fisheries Science Center (AFSC) and Fisheries and Oceans Canada (DFO), we have assembled a coastwide dataset, combining catches of 50+ groundfish species in scientific bottom trawl surveys, 2003–present. Spatially, this joint dataset represents scientific surveys spanning the contiguous range from the U.S./Mexico West Coast border northward through British Columbia to the Gulf of Alaska and Bering Sea (Alaska). Several previous efforts from our group have used this dataset to answer questions related to large-scale changes in species distributions, quantifying changes in population biomass through time, and better understanding shifts in species' thermal niches. This previous work represents the tip of the iceberg of the potential of these data, however, and we are interested in using these data to answer pressing questions affecting the biology and management of these species in a rapidly changing environment. The population status for many of these species is not evaluated often, thus quantifying population change in space and time across a large group of species has the potential to rapidly inform decision making (for both sustainable fisheries and conservation) and bring new understanding to overlooked species. Some of the research topics we are interested in having students collaborate with us on include the following:

(1) In addition to fish abundance and biomass (numbers and weights of species catches) and biological (length, weight, age, etc.) information, the survey data we have collected measure environmental variables, such as temperature and oxygen. Recent applications of species distribution models (SDMs) have tried to quantify changes in species' thermal preferences (Ward et al. 2024), and we are interested in expanding this to include other environmental variables and depth.

(2) We are interested in developing workflows for analyzing biological information associated with these samples. Our current dataset (bundled as an R package surveyjoin) includes the total weight of each species in each haul; we are interested in bringing in measurements of individual fish in these same surveys, allowing us to perform coastwide analyses of size (growth), condition (weight at length), and develop life-stage specific biomass indices (Tolimieri et al. 2020).

(3) Summarizing data from surveys to generate high-level understanding of species across the Northeast Pacific Ocean. Following our Canadian colleagues' development of annual reporting (Anderson et al. 2019), we are working on combining data from the USA and Canada into an online dashboard to report changes in biomass, biology, describe how and where samples have been collected, and understand the intensity of fishing pressure.

Project 2: Development of machine- and deep-learning models for processing remote-sensing data

Mentor: **Dr. Eli Holmes**

In this project, you will work on a project using deep learning (e.g., convolutional neural networks) or machine learning (e.g. tree models) for processing remote-sensing data used for ocean ecosystem and fisheries research. There are a variety of different projects that interns can take on depending on their specific interests and goals. Our current main projects for 2025 concern algorithms and applications for filling missing values in ocean satellite data. These gaps are ubiquitous in ocean color data used to estimate chlorophyll and ocean productivity. The gaps occur in two main ways. First, the data are missing due to clouds or glint, and clouds cover around 70% of the ocean at any time. Second, the remote-sensing product may be low-resolution, while we want high-resolution information. The 2024 interns made substantial progress on proof-of-concept models for gap-filling and up-scaling. The 2025 intern(s) will build on this work to develop more flexible regional gap-filling models. Here are some ideas for summer projects but the intern is not limited to these and interns are encouraged to develop their own project based on their interests.

1. Develop a proof of concept for a foundational model for regional gap-filling and up-scaling tasks. Foundational models 'self-learn' a complex data set and then allow one to develop new gap-filling applications.
2. Continue the 2024 work, finish the model validation and help write a paper for submission to a journal
3. Develop of a python package that uses trained deep-learning models to produce gap-filled or up-scaled ocean color data
4. Develop proof of concept approaches for chlorophyll-a estimates when remote-sensing data are fully absent (prior to 1997). This would extend a pilot project by the 2024 interns.
5. Assemble Cloud-ready and AI-Ready datasets for ocean and fisheries data that we will use for tasks 1-3

Past intern projects are listed at this website: https://github.com/SAFS-Varanasi-Internship/

Why is this project important?

Derived from satellite observations, ocean color provides information about the concentration of chlorophyll-a, a proxy for phytoplankton abundance. Phytoplankton form the foundation of the marine food web, supporting zooplankton, which in turn feed many commercially important fish species. Ocean color data also allows us to detect and monitor harmful algal blooms, which produce dangerous toxins and can deplete oxygen levels, impacting fish health and habitats. Clouds produce so many missing values that the data are often hard to use for fisheries research, because the ocean color data is often missing where we have fisheries data. Although there are gap-filled chlorophyll products, these are not available for the full suite of ocean color products, and thus we are not getting the full benefit of all the ocean color data that we have. The gaps in the data directly inhibit our ability to use satellite data in fisheries and ocean ecosystem applications.

What the intern will gain from the project?

This project will give the intern experience working with ocean remote-sensing data and applying image classification and machine learning algorithms to create predictive models. The intern will take part in a research project on novel uses of these approaches in fisheries and have the opportunity to work with others during hackweeks. The intern will get experience with popular Python packages for machine learning (e.g. TensorFlow, PyTorch or Keras). As part of the project, you will also have the opportunity to participate in an earth data hackweek: OceanHackWeek (in Seattle), PACE HackWeek (Baltimore or San Diego) or NASA Earth Data Hackweek (Seattle). In addition, interns have presented their work at conferences. The Varanasi internship gives you an opportunity to develop research skills, and past interns have had high success at acceptance to masters and PhD programs in computer science and data science.

Required background:

The intern will need programming experience in Python to be successful in this project. The intern should be able to read about an algorithm and write code to implement it. The intern should enjoy reading tutorials and then apply the ideas learned to a novel project on a different set of data. Prior experience with machine- and deep-learning (classes or self-study) will be helpful but not required.