# LAUREN ALEXANDRA

## Data Scientist

Sacramento, CA | +916-990-7805 | lauren@laurenalexandra.io | laurenalexandra.io
linkedin.com/in/lauren-alexandra | github.com/lauren-alexandra

## SUMMARY

Lauren has worked on projects supporting efforts to improve understanding of Earth system processes. Her research focus is climate-informed modeling for ecosystem management with a particular emphasis on mitigating the impacts of hydroclimate variability.

## SKILLS

- Machine Learning
- Deep Learning
- Python
- Remote Sensing
- GIS
- Geospatial Analysis

## EXPERIENCE

**Environmental Molecular Sciences Laboratory at Pacific Northwest National Laboratory**
Software Engineer | April 2023 - February 2025

- Applied machine learning models for hydrological forecasting to understand climate risk
- Developed a data and project platform designed for environmental researchers

**Pacific Northwest National Laboratory**
Masters Intern | June 2022 - April 2023

- Implemented machine learning infrastructure for efficient experiment tracking on high-performance computing systems to support reproducibility

**American River Conservancy**
Water Quality Monitor | June 2024 - December 2024

- Collected hydrological data in the Cosumnes River Watershed for UC Davis partner research on wildfire impacts

## EDUCATION

**Present | University of Colorado Boulder**

Earth Data Analytics, Graduate Certificate Program

**2024 | UC Agriculture and Natural Resources**

Climate Steward

**2024 | University of California, Davis**

Geographic Information Systems Specialization

**2022 | Colorado State University**

Artificial Intelligence and Machine Learning, M.S.

## PROJECTS

**Blue Oak Habitat Suitability**

- Late 21st century habitat suitability for Quercus douglasii is evaluated under a medium emissions scenario across national forests in Northern and Southern California.
- Downscaled climate models were compared by site for two periods (2036-2066 and 2066-2096), examining winter and summer mean temperatures (and change relative to historical by °F) as well as winter precipitation (and percent change relative to the historical value).
- Given optimal characteristics for each site variable, a habitat suitability score was identified for each site leveraging a fuzzy Gaussian function to assign scores for each raster cell based on proximity to ideal site values.
- For each site and medium emissions scenario, variable rasters were harmonized with rioxarray across resolution and projection, suitability layers calculated and multiplied, and then combined into a final raster.
- Project: http://laurenalexandra.io/projects/habitat-suitability

**Willow Flycatcher**

- A species distribution for Empidonax traillii is mapped across the Americas.
- The Global Biodiversity Information Facility occurrences data was accessed with the Python client for the GBIF API and queried for the year (2023), species, and coordinates.
- Data for ecoregions were gathered from a RESOLVE shapefile (2017) and read into a GeoDataFrame with geopandas. The data was normalized by space (ecoregion average) and time (monthly average) to account for the sampling effort.
- The visualization produced highlights monthly migration patterns and is interactive due to the sliding widget from the HoloViews panel library.
- Project: https://laurenalexandra.io/projects/species-distribution

**Basin Inflow**

- A deep neural network for forecasting reservoir inflow in the American River Basin.
- The proposal model is composed of stacked LSTM layers, with regularization applied to both (L2, dropout, and recurrent dropout), followed by a densely connected layer.
- Model application and evaluation uses river basin data with temporal coverage from 2008-2022.
- Across the basin and along forks of the river, daily precipitation, temperature, snow water content and depth, and river discharge and stage are employed to predict local reservoir inflow.
- Data was pre-processed with a 30-day exponential moving average to give recent weather events more weight in the forecast, normalized, and made stationary for training. 14 years of data were separated into three sets: training set (2008-2017), validation set (2018–2020), and test set (2020-2022).
- The model achieves a lower MAE score (0.036) in inflow prediction than the baseline model.
- Project: https://laurenalexandra.io/projects/basin-inflow