

# Kealie Pretzlav, PhD

(510) 213-0267 | kealie@pretzlav.com | <https://www.linkedin.com/in/kealiep/> | <https://github.com/kpretz/>

## SUMMARY

Senior PhD data scientist and hydrologist/geomorphologist with a strong scientific domain knowledge and proven track record of data analysis, modeling, and machine learning applications using Python and SQL for climate, water, and broader environmental concerns. Currently seeking data science, analyst, or modeling positions in the environmental climate field (e.g. water, floods, wildfires, climate, land use, agricultural, soils).

## SKILLS

### Data Science and Modeling

- Python, SQL, APIs, Data Visualization
- Machine Learning with Random Forest, SVM, Logistic Regression; Time series modeling (e.g. LSTM)
- Data expertise: big data, timeseries, geospatial, remote sensing, climate projections, field data, data architecture

### Water Resources

- Hydrology, geomorphology, climate, soils, hydraulics, land use, topography/DEMs
- Software: ArcGIS(Pro), Matlab, Civil 3D, HEC-RAS, HEC-HMS, SRH2D
- Field gaging installation, maintenance, and calibration: e.g. precipitation, streamflow, sediment flux, water quality

### Communication, Work Style, and Leadership

- Able to communicate and build consensus within groups of technical and non-technical people
- Excellent data investigator; finding patterns, QA, answering questions, and relating to natural processes
- Excellent track record of mentorship, developing best practices, and successful project management

## PROJECTS

### Restoration Impact of Water Retention in Mountain Meadow for Climate Resiliency

Quantify restoration effectiveness of water storage in mountain meadow using LSTM and Random Forest machine learning models in Python. Target variable was remote sensing Landsat product NDWI, features were various hydrologic parameters (precipitation, snow depth, streamflow). Results showed restored meadows increased groundwater storage and mid-summer baseflow.

### Pond Inundation and Timing Model (Pond-IT)

Open-source modeling tool to predict pond inundation probabilities over decades of climate projections in Python and Excel VBA as a low-cost estimation climate adaptation tool for land planners. The model uses least squares optimization, fits to available low resolution pond depth field data, and publicly available datasets, combining climate timeseries, sparse field observations, and land use geospatial datasets to produce pond dry-down curves. Final products included technical guidebook, GitHub repo, Excel VBA workbook, and training webinar videos, and two live online training sessions.

### Economic Impact of Floods

Machine learning model in Python to predict to predict US Census median household income by Zip Code Tabulation Area (ZCTA). Data ingested using JSON/web scraping APIs. Geopandas used to derive stream channel length and flood occurrence by ZCTA.

### Smart Rock Motion and Rest Algorithm

Install and recover artificial “smart rocks” in snowmelt-dominated mountain stream to characterize cobble transport during peak flows. Developed Python algorithm to convert high-resolution accelerometer, gyroscope, and magnetometer data to characterize particle motion and rest. 50GB of data processed using pytables.

Verified against direct observations in the laboratory.

More information here: <https://doi.org/10.1029/2020WR028150>

### Real-time Stream Gage Data Analysis Pipeline and Data Architecture

Developed stream gage archival database architecture, best practices, and integration with real-time system.

Created internal database query tool using Docker-deployed jupyter notebooks and SQL for easy data access within company intranet.

## EXPERIENCE

### Balance Hydrologics, Berkeley, CA

Senior Geomorphologist/Hydrologist

May 2015 – Mar 2022

- Project lead for multiple large interdisciplinary projects (biologists, engineers, geologists, ecologists) with multiple stakeholders including parks, conservation, permitting, municipal, and funding agencies
- Led all big data or scientific computing projects, including climate change projections, ML model development, remote sensing data, and open-source modeling tools (Python, VBA) which provided novel services for clients
- Developed internal tools and processes for combining datasets with different temporal and spatial scales, including archival database of field gaging datasets resulting more efficient data workup procedures and data resulting in improved data processing efficiency and capabilities
- Completed successful project grant applications for clients
- Led multi-person geomorphic and hydrologic surveys, including equipment installation, maintenance, and data workup
- Mentored junior staff in company procedures, tools, and scientific observation techniques which bolstering team independence and work efficiency

### Department of Geological Sciences, The University of Texas at Austin

Aug 2011 – Aug 2016

Graduate Teaching and Research Assistant

- Extensive data post-processing of Accelerometer, 3D Velocity, and Terrestrial Lidar data using Python and Matlab
- Designed and programed custom-built accelerometer- and gyroscope-embedded “Smart Rock”
- TA for intro to geology and near-surface geophysical data collection techniques in the field, gave class demonstrations, revised class material

### M-Factor Inc., San Mateo, CA

Jul 2009 – Jun 2011

Operations Data Analyst

- Managed trade and financial datasets for consumer packed goods clients for use in financial analysis software

## EDUCATION

### Springboard Data Scientist Program

Anticipated Summer 2024

6-month intensive certification, machine learning concentration

### Ph.D., Geological Sciences

May 2016

The University of Texas at Austin, Jackson School of Geosciences

Dissertation: *Flash Floods and Unsteady Flows: Sediment Transport, Turbulence, and Bed Surface Armoring*

### Bachelor of Arts, Applied Mathematics, Operations Research Concentration

May 2009

University of California at Berkeley

*References available upon request*