

PROFESSIONAL SUMMARY

Experienced **data engineer** with 5+ years of expertise in **data engineering and analysis**, **data pipelines**, and **machine learning**. Seeking an opportunity at a climate tech company where I can leverage climate data to track, analyze, and forecast environmental challenges, driving innovative solutions for a more sustainable future.

- * **Data Engineering & Analysis:** Leveraged cloud-native tools to design a create a processing and analysis workflow that automates extraction, processing, analysis, and interpretation of cloud-hosted spatial datasets in under 30 seconds at Haven. Contributing author on an ongoing paper calculating extreme precipitation metrics using gridded rainfall data in Hawaii.
- * **Data Pipeline Development:** Built a 13-script address ETL data pipeline using GeoPandas and GDAL/OGR, adding 25+ million new addresses and reducing third-party API calls by over 50% at EarthDefine.
- * **Machine Learning:** Achieved 95% model accuracy in a roads classifier compared to aerial imagery by developing an innovative image augmentation workflow at EarthDefine.

Top skills: Geographic Information Systems (QGIS, ArcGIS) | Python (GeoPandas, NumPy, GDAL/OGR, Rasterio, Shapely, Xarray, DuckDB) | PostgreSQL / PostGIS | Geospatial Data Engineering | Geospatial Data Analysis | Extract, Transform, Load (ETL) | Data Pipeline Development (Airflow, Spark, Sedona) | Machine Learning | Cloud-Native Geospatial (Wherobots, AWS, Iceberg) | Climate Solutions & Sustainability | Systems Thinking | Communication Skills | Cross-Functional Collaboration & Leadership

WORK HISTORY

Climate Data & Geospatial Lead, Haven (volunteer) **March 2025 – Present**

Agentic AI for climate resilience

- * **Data pipeline development, systems thinking:** Lead climate data integration and geospatial strategy. Orchestrated a data extraction pipeline via Airflow to retrieve real-time updates of risk datasets.
- * **Geospatial climate data engineering & analysis:** Leveraged Wherobots to design a create a processing and analysis workflow that automates extraction, processing, analysis, and interpretation of cloud-hosted spatial datasets in under 30 seconds.
- * **Cross-functional collaboration & leadership:** Collaborate with the AI team to integrate agentic AI with spatial datasets.

Senior Geospatial Data Engineer, EarthDefine **July 2024 – Present**

AI powered mapping. Off-the-shelf nationwide geospatial data.

- * **Machine learning, critical thinking:** Achieved 95% model accuracy in a roads classifier compared to aerial imagery, by developing an innovative image augmentation workflow on NAIP imagery.
- * **Data pipeline development:** Led a complete transition from ArcPy to open-source geospatial Python for my address ETL pipeline, cutting processing time per script by at least 50%.

Geospatial Data Engineer, EarthDefine **June 2022 – July 2024**

AI powered mapping. Off-the-shelf nationwide geospatial data.

- * **Data pipeline development:** Built a 13-script address ETL pipeline, adding 25+ million new addresses and reducing third-party API calls by over 50%.
- * **Data pipeline development:** Developed a PostgreSQL/PostGIS database of 215 million addresses with network-wide access, improving query efficiency by over 90% compared to standard GIS methods.

Remote Sensing Researcher, NASA DEVELOP National Program **January – April 2022**

Capacity-building program for early-career professionals. Applying Earth observations to environmental decision-making needs.

- * **Cross-functional collaboration & leadership:** Collaborated with clients and chief scientists to ensure communication of project complied with NASA standards.
- * **Geospatial climate data engineering:** Automated collection of remote sensing NO₂ data using Python APIs, expediting data availability.
- * **Geospatial climate data analysis:** Conducted multidimensional statistical analyses using ArcGIS Pro, correlating satellite and in situ wind measurements.

Independent consultancy working with farmers to modernize their workflow using GIS tools.

- * **Python development:** Scripted a solution to visualize week-to-week vine growth deltas.
- * **Python development:** Launched an agricultural monitoring workflow displaying daily updates of NDVI and water stress maps by automating geospatial data collection and visualization.
- * **Python development:** Achieved 98% automation by creating harvest scheduling and management workflows.

RESEARCH EXPERIENCE

Spatiotemporal Analysis of Precipitation in Hawaii Using High-Resolution Gridded Rainfall Data (ongoing; [link](#))

- * **Geospatial climate data engineering:** Calculated multiple precipitation frequency and intensity metrics across the Hawaiian islands from 1990 – 2024 using R and Python with the goal of understanding long-term climate trends in the state.

Poletti AN, Frierson DMW, Aerenson T, Nikumbh A, Carroll R, et al. (2024) Atmosphere and ocean energy transport in extreme warming scenarios. PLOS Climate 3(2): e0000343. <https://doi.org/10.1371/journal.pclm.0000343>

- * **Raster analysis:** Evaluated outputs from three climate models showing extreme climate sensitivities out to 2300 from the SSP5-8.5 extension scenario of the newly released CMIP6 ensemble.

RELEVANT PROJECTS

Spatiotemporal Analysis of Air Pollutants Collected from Ground and Space Instruments Around the Spring 2022
Guadalupe Mountains and Carlsbad Caverns National Parks, NASA DEVELOP

- * **Geospatial climate data analysis:** Examined average monthly, seasonal, and annual tropospheric column concentrations of NO₂ in Carlsbad Caverns and Guadalupe Mountain National Parks using remote sensing data from OMI and TROPOMI satellites.

Analyzing projected population change in the México-Lerma- Spring 2021
Cutzamala Hydrological Region (MLCHR), 2010-2100, Clark University

- * **Geospatial data analysis (vector):** Conducted space-time cube analysis in ArcGIS Pro to pinpoint population trend hot spots and statistically significant patterns across the five shared socioeconomic pathways (SSP) in the MLCHR region of Mexico. Findings helped identify municipalities at risk of future water shortages.

GIS Suitability of Agrivoltaic Array Installation to Mitigate Climate Stress on Crops, Allegheny College Fall – Spring 2018

- * **Geographic Information Systems (GIS), critical thinking:** Used ArcGIS Desktop geographic information systems (GIS) to create a suitability analysis of optimal locations to install agrivoltaics to mitigate increasing heat stress on crops and provide solar energy to the local community.

EDUCATION

Climatebase Fellow, Cohort 7 March – June 2025
Climatebase Fellowship

Master of Science, GEOGRAPHIC INFORMATION SCIENCE May 2022
Clark University, Worcester, MA

Doctoral Coursework, ATMOSPHERIC SCIENCES August 2019 – January 2021
University of Washington, Seattle, WA

Bachelor of Science, ENVIRONMENTAL SCIENCE May 2018
Allegheny College, Meadville, PA