Nathan Korinek

nkorinek@proton.me 303-214-8053 github.com/nkorinek linkedin.com/in/nathan-korinek-474148300 ORCID: orcid.org/0000-0003-0859-7246 nkorinek.com

Work Experience

ORISE Climate Data Fellow

Climate Adaptation Science Center (CASC), U.S. Geological Survey (USGS), Nov 2024 – Present

- Process multi-terabyte climate datasets using Python on USGS's HPC systems, applying parallel computing and data integrity checks to prepare Zarr data stores optimized for the cloud.
- Convert and validate datasets for inclusion in public STAC (SpatioTemporal Asset Catalog) repositories; five datasets processed, two in active upload pipeline.
- Ensure CF-compliant metadata and perform sample workflows to confirm data quality and usability.
- Access cloud data and transfer large COG, NetCDF, and Zarr datasets between HPC and cloud storage.
- Lead onboarding and informal training sessions for team members on cloud workflows, HPC usage, and geospatial Python tools.

Graduate Research Assistant

Earth Lab, University of Colorado Boulder, Aug 2022 - Sep 2024

- Performed geospatial and time series analysis on terabytes of ecological and climactic datasets with Python and R for four publications at the lab, working collaboratively and focusing on efficiency and cloud computing.
- Researched wildfire intensity and severity over time. Synthesized data including Fire Radiative Power, wildfire boundaries, burn indices,
- Published and presented work using linear and random forest models to quantify and model the effect of previous burns on the severity of fires in reburned areas.
- Responsible for visualizing geospatial data in Python to convey findings, including creating figures that are published in scientific papers and conference posters.
- Assisted with planning and piloting UAS flights to collect multispectral imagery and structure from motion over forested areas in Colorado, Wyoming, and Montana.

Data Scientist Software Developer

Earth Lab, University of Colorado Boulder, Jan 2019 - Aug 2022

- Developed and maintained production quality open source Python packages for earth analytics and Jupyter notebook grading. The packages have 72,000 and 16,000 conda-forge downloads, respectively.
- Added content and maintained earthdatascience.org, an open education website for earth data science and analysis with over 100,000 monthly visitors. Over 20 lessons have been used for five years in two classes CU Boulder offers.
- Wrote documentation and created tests for existing packages, as well as for developing and testing additions to packages.
- Worked with tribal and community colleges to develop Earth Data Science Corps, a program that hundreds of people in under served communities have gone through that has run since 2020.
- Co-developed and contributed the .clip() function to GeoPandas.

Education

Master of Arts in Geography, Aug 2022 – May 2024

University of Colorado Boulder, Boulder, CO GPA: 3.96

Relevant Coursework: Machine Learning, Neural Networks, Uncrewed Aerial Systems, Forest Ecology

Bachelor of Arts in Geography, Aug 2017 – May 2019

University of Colorado Boulder, Boulder, CO

Certificate in Geographic Information Systems and Computational Science

GPA: 3.76

Relevant Coursework: Earth Data Science, Earth and Climate Systems, Computer Science, Geospatial Statistics, Geographic Information Systems

Skills

Programming Languages: Python, R

Machine Learning Familiarity: Artificial Neural Networks, Convolutional Neural Networks, Long Short-Term Memory Networks, Random Forest, K Nearest Neighbor, Tensorflow, Keras, Scikit-learn
Software: Git, GitHub, Linux, Cloud Computing, Metashape, ArcGIS Suite, ENVI, Google Earth Engine
Geospatial Data Processing: GeoPandas, Rioxarray, Xarray, Dask, Rasterio, Terra, ArcPy, ArcGIS, QGIS, Cloud Computing, Parallelization, Data Visualization, Image Processing, Data Science, Computer Vision, Data Synthesis, APIs, UAS Data Collection, UAS Piloting, Open Source, Open Science

Projects

Machine Learning for glacier identification, Dec 2023

• Developed a semantic segmentation neural network that used pre-trained MobileNetV2 and pix2pix to identify glaciers from Sentinel-2 imagery within Chile with 93% test accuracy.

Watershed Analysis for lake carbon measurements, Feb 2023

- Created geospatial Python workflows based on a description of the desired output from a collaborator.
- Developed Python scripts to delineate lake upstream catchments using flow accumulation data, and generated statistical summaries of elevation, tree cover, and NDVI across those areas.

Machine Learning to predict vegetation recovery post-fire, Dec 2022

- Performed data curation and synthesis to create a dataset of fires' vegetation recovery in Colorado.
- Helped develop a Multi-Task LSTM to predict vegetation recovery trajectories.

Publications

Cyberinfrastructure deployments on public research clouds enable accessible Environmental Data Science education, Jul 2023

 ${\rm doi.org}/10.1145/3569951.3597606$

abc-classroom: Tools to automate GitHub classroom and autograding workflows, Feb 2022 doi.org/10.5281/zenodo.6026436

Intermediate Earth Data Science Textbook: Open Education Textbook, Oct 2021 doi.org/10.5281/zenodo.10864778

EarthPy: A Python package that makes it easier to explore and plot raster and vector data using open source Python tools, Nov 2019 doi.org/10.21105/joss.01886

Presentations

ESIIL Innovation Summit 2024, May 2024

• Facilitated a diverse and inclusive team of scientists from varying backgrounds to develop and start work on a paper using Machine Learning to investigate down stream water quality from fires in the western US.

2023 American Geophysical Union Poster Presentation, Dec 2023

• Intersecting fire, insect, and drought disturbance and the fate of western US forests in a changing climate.

2023 Forest Resiliency Data Synthesis Working Group, Feb 2023

• Created lessons for and held sessions related to GEDI data, ECOSTRESS data, NEON Hyperspectral Data, and K-means clustering.

Earth Data Science Corps (EDSC), Jul 2020, 2021, 2022

- Collaborated with tribal and community colleges to create curriculum and teach lessons on earth data science in Python for a 1 month long summer program (EDSC).
- Discussed potential projects with teams and individuals, talked about feasibility and how to execute ideas. Aided in producing 13 geospatial projects over three years.