# Paul M. **Hobson, P.E.**

Environmental Analytics · Open Source Development · Data Management and Engineering · Civil Engineering (Licensed in OR)

*Portland, Oregon, USA*

(770) 633-8096 | pmhobson@gmail.com | phobson | pmh2635 | @pmhobson

## Skills

| | |
|---|---|
| **Programming** | Scientific Python (Expert), R (Proficient), C# & .NET (Proficient), C/C++ (Limited), git (Expert), Continuous Integration (Advanced) |
| **Computing** | Spatial Analysis (Expert), Machine Learning (Familiar), Microsoft Azure (Proficient), Google Earth Engine (Familiar) |
| **Web** | Relational Databases (Advanced), REST API Development with C# (Proficient), Python (Proficient), HTML5 (Familiar) |
| **Numeric Models** | SWMM (Familiar), HEC-RAS (Familiar), EFDC (Familiar+), WASP (Limited) |

## Open Source Portfolio

| | |
|---|---|
| **coiled/dask-vertica** | Author – Created a library for asynchronously reading and writing data to Vertica data warehouses via Dask. |
| **matplotlib/matplotlib** | Committer – Contribute to documentation, review pull requests, and maintain statistical plotting functions. |
| **matplotlib/mpl-probscale** | Author – Created library that uses scipy's statistical distributions to implement probability scales for axes. |
| **pygridgen/pygridgen** | Maintainer – Fix bugs and manage releases of a `ctypes` interface to generate curvilinear-orthogonal meshes. |
| **geosyntec/pygridtools** | Author – Created and maintain high-level curvilinear-orthogonal mesh manipulation and visualization tools. |
| **BMP Database/wqio** | Author – Created, maintain statistical methods for assessing the efficacy of stormwater management practices. |
| **phobson/paramnormal** | Author — An API around scipy's `stats.distributions` to redefine distributions based on their classic definitions. |
| **python-metar** | Maintainer – Review pull requests and respond to bug reports for a python library that parses weather reports. |
| **geosyntec/cloudside** | Author – Builds on python-metar to fetch 5-mibute weather data observations from an FTP hosted by the FAA. |
| **Google Summer of Code** | Mentor – Co-mentored a GSOC student whose project goals involved enhancing the plotting capabilities of geopandas. |

## Work History

| | | |
|---|---|---|
| 2022–Present | **Open Source Developer,** Coiled Computing | *Portland, OR* |
| 2021–2022 | **Machine Learning Engineer,** Confluency | *Portland, OR* |
| 2017–2021 | **Project Engineer (VI),** Geosyntec Consultants | *Portland, OR* |
| 2012–2017 | **Water Resources Engineer (III),** Geosyntec Consultants | *Portland, OR* |
| 2008–2012 | **Senior Staff Engineer (II),** Geosyntec Consultants | *Portland, OR* |
| 2006–2008 | **Research and Teaching Assistant,** Georgia Institute of Technology | *Atlanta, GA* |

## Representative Projects & Responsibilities

### Daily Responsibilities | Open Source Developer

*Coiled Computing*

Open Source Maintenance, Customer Engagement, User Research

*2022 - Present*

As an open source developer at Coiled Computing, I took on a wide variety of roles that led the company forward. One critical role was maintaining the Dask ecosystem, a component of the larger scientific python ecosystem that allows scientists and engineers to scale up numeric computing, data analysis, and machine learning workflows onto distributed clusters. Specifically, my role included OSS community management, triaging bug reports, fixing bugs, and enhancing the libraries with new features. To support Coiled's Marketing and Outreach team, I created large, production-scale machine learning pipelines. These included integrating an ML-training & hyperparameter tuning pipeline with the Coiled platform, Dask, XGBoost, and Optuna with Snowflake and Parquet data sources to predict New York City trip times in for-hire vehicles. Additionally, I managed approximately 10 customers and provided short-term computational consulting to them on an as-needed basis.

### Daily Responsibilities | Machine Learning Engineer

*Confluency*

Technical Leadership, Client Relations, Machine Learning

*2021 - 2022*

As a machine learning engineer at Confluency, I split my time evenly between delivering water resources projects and contributing to a Small Business Innovation Research (SBIR) grant that investigated the uses and applicability of machine learning models to solve urban stormwater and water distribution problems. For my clients, my work focused on developing computational tools in support of efforts to automate numeric model scenario creation, execution, and validation. A large part of that work included spatial and temporal analysis of hydrologic and meteorologic datasets, including synoptic analysis of rainfall records Under the SBIR grant, I developed various machine learning models to predict the occurrence and magnitude of combined sewer overflows (CSOs) based on storm characteristics and antecedent watershed conditions.

### Daily Responsibilities | Project Engineer

*Geosyntec Consultants*

Technical Leadership, Project Management, Staff Mentorship

*2018 - 2021*

My daily duties as project engineer were varied. Annually, I managed approximately $250,000/year of environmental consulting work for industrial, municipal, and federal clients. The technical side of my job broadly related to helping clients restore, protect, and preserve the natural environment through robust data analysis, and statistical and numerical modeling. In doing so, I implemented domain-specific statistical algorithms (e.g., regression-on-order statistics for censored data), collaborate with other consultants and academia by maintaining open source libraries and web APIs built in Python or .NET. For the past five years I have served on the steering committee for the company's global information management and data analysis practice group. I also contributed to proposals and marketing efforts targeting potential clients' requests for proposals and statements of qualifications. Lastly, I happily provided daily, weekly, and *ad hoc* technical and career-focused mentorship to junior staff.

### Iowa Agricultural BMP Assessment

*Iowa Nutrient Research & Education Council*

Data Management, Workflow Automation, Spatial Analysis

*2020 - 2021*

- Incorporated high-resolution DEM and land cover rasters with vector datasets of BMPs within select HUC12 watersheds
- Assessed agricultural phosphorus loss reduction due to implementation of agricultural BMPs in the 1980s, 2006 - 2008, and 2016 - 2018 by HUC12, major land resource area, and state boundaries
- Automated the workflow into a python library with a command line interface

### The International Stormwater BMP Database | https://dot.bmpdatabase.org

*NCHRP & WRF*

Project Management, Web API Development, Statistical Analysis

*2015 - 2020*

- Secured granted funding from the National Cooperative Highway Research Program (NCHRP) and the Water Research Foundation (WRF).
- Lead the migration from Microsoft Access database file to an Azure SQL database with a RESTful API written in C#.
- Authored statistical report summarizing the efficacy of the pollutant removal strategies found in the database. Analysis relied on scientific python and supported further development of `wqio`, which implements regression-on-order-statistics, a technique for imputing left-censored observations.

### Tahoe Regional Stormwater Management Program | Tahoe Data Management System

*Tahoe Resource Conservation District*

Database Design, Web API Development, Statistical Analysis

*2010 - 2021*

- Designed and developed a SQL Server database and desktop frontend to manage stormwater quality, hydrologic, and meteorologic data from the Lake Tahoe Basin in support of their lake clarify initiative.
- Responsible for periodic update, curation, and management of high-frequency environmental monitoring data and database/API performance.
- Developed SQL and C#-based ETL scheme to provide site-specific and basin-wide reports for annual summaries and long-term environmental trends.

### Natural Gas Pipeline Hazard Analysis

*State of Arizona*

Project Management, Geospatial Analysis, Machine Learning

*2016*

- Compiled a comprehensive dataset of natural gas pipeline stream crossings throughout Arizona from disparate data sources including old database and hard-copy maps.
- Automated the collection and assessment of hydrological and geomorphological conditions at pipeline crossings.
- Oversaw supervised machine learning analyses to classify the risk of pipeline exposure based on watershed characteristics and other data.

### Puget Sound Stormwater Heatmap | https://stormwaterheatmap.org

*The Nature Conservancy*

Distributed Data Management, Hydrologic Modeling, Geostatistical Analysis

*2018 - Present*

- Interactive stormwater pollutant heatmap helps city planners in the Puget Sound area prioritize green infrastructure.
- Project required automating >10,000 hydrologic model runs along with results post-processing and uploading to Google Big Query.
- Model results were then combined with high-resolution terrain and landuse raster data and aggregated into a Google Earth Engine layer.
- Full workflow requires handling of tens of terabytes of data.

### Sediment and Pollutant Fate & Transport Modeling Projects

*Various Confidential Clients*

Hydraulic Modeling, Software Engineering, Data Management, and Geospatial Analysis

*2012 - Present*

- Utilized three-dimensional hydrodynamic and sediment fate and transport modeling for litigation support projects requiring strict confidentiality.
- My role focused on building tools to support the creation of 3D hydrodynamic models and post-processing model results.
- Projects typically support further development of open source libraries such as `pygridgen` and `pygridtools`.

## Select Presentations and Publications

**NCHRP 25-25 Task 120 Final Report: Use of the State Department of Transportation Portal to the International Stormwater BMP Database | Full text** — *Contributing Author*

Transportation Research Board - National Cooperative Highway Research Program — *March 2020*

- This report details the grant-funded work to create a web portal to the stormwater management data most relevant to state departments of transportation.
- The portal allows engineers and scientists to query the database for stormwater quality data, generate statistical summaries, and access reports.
- An interactive web map provides an alter direct access to data and result summaries.

**`pygridgen` & `pygridtools`: A Successful Collaboration with the Private Sector, Academia, and the Open Source Community | Slides** — *Austin, TX*

Scipy Conference — *July 2017*

- Covered the challenges and successes of making niche, domain-specific tools more accessible to practicing scientists and engineers.
- Discussed my experience as a consulting engineer and maintainer of a python interface to a C-library designed to create curvilinear-orthogonal grids for hydrodynamic and oceanographic modeling.
- Highlighted the contributions of academia, the open-source scientific community, the US federal government, and in most notably the effort and infrastructure provided by conda-forge.

**Quantile and Probability Plots | Slides** — *Portland, OR*

PDX Data Visualisation — *February 2017*

- Presented at a meetup of data visualization professionals to discuss the use a probability plots to visualize the empirical distribution of data.
- Provided an overview of a library/matplotlib extension I authored that allows users to use any distribution in scipy to define a probability scale for an axis.

**Python in Civil and Environmental Engineering | Slides** — *Portland, OR*

PDX Python — *March 2016*

- Presented to a general audience of python enthusiasts to provide general context of the challenges of environmental consulting and the ways in which the rapidly expanding scientific python ecosystem was addressing them.
- I also introduced some of the open source libraries I had authored.

## Education

**Georgia Institute of Technology** — *Aug. 2006 – Dec. 2008*

M.S. Civil Engineering — *Atlanta, GA*

Thesis: *Rheologic and Flume Erosion Characteristics of Georgia Sediments from Bridge Foundations*

**Georgia Institute of Technology** — *Aug. 2001 – May 2006*

B.S. Civil and Environmental Engineering — *Atlanta, GA*

Cooperative education degree, graduated with Highest Honors

## Volunteering

**Officer of the Board of Directors** — *2017 - Present*

Northwest Trail Alliance — *Portland, OR*

Since 2017 I have served as an officer on the all-volunteer board of directors of the Northwest Trail Alliance (NWTA). The mission of NWTA is to advocate for, enhance, and steward trails and access to nature in the area surrounding Portland that extends from the Pacific coast to the west, Mt Hood to the east, and Mt St Helens to the north. I have spent two years as Treasurer of the organization, managing and accounting for roughly $300,000 in assets and funds and maintaining insurance for the organization and its volunteers. I have also spent a year as Secretary of NWTA and two years as Vice President, keeping records, filing necessary paperwork with the State of Oregon to maintain our 501(c)(3) non-profit status and engaging community and land manager stakeholders to advocate for NWTA's mission. In addition to my duties as an officer, I act as a trail crew leader during volunteer events and as a Trail School instructor, teaching aspiring trail stewards the basics of trail design, water management, erosion mitigation, soil and rock types, and best practices for trail sustainability.