

# pyflowline: a mesh-independent river network generator for hydrologic models

Chang Liao <sup>1</sup> and Matt G Cooper <sup>1</sup>

<sup>1</sup> Atmospheric, Climate, and Earth Sciences, Pacific Northwest National Laboratory, Richland, WA, USA

DOI: [10.21105/joss.05446](https://doi.org/10.21105/joss.05446)

## Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: Hauke Schulz 

## Reviewers:

- [@smchartrand](#)
- [@andres-patrignani](#)

Submitted: 27 March 2023

Published: 06 November 2023

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

## Summary

River networks are crucial in hydrologic and Earth system models. Accurately representing river networks in hydrologic models requires considering the model's spatial resolution and computational mesh. However, current river network representation methods often have several limitations: (1) vector-based; (2) they perform poorly at coarse resolution (3) they do not support unstructured meshes. To overcome these limitations, we developed PyFlowline, a Python package that generates mesh-independent river networks. With PyFlowline, hydrologic modelers can generate conceptual river networks at various spatial resolutions for both structured and unstructured computational meshes. The generated river network datasets can be used by hydrologic models across scales.

## Statement of need

For hydrologic modelers, river networks are a key input for hydrologic models. While some hydrologic models accept vector-based river networks ([Mizukami et al., 2016](#); [Schwenk & Hariharan, 2021](#)), others only accept mesh cell-based representations, or require a generation method from a vector-based river network. Currently, generating a mesh cell-based river network from a given vector-based river network and arbitrary computational mesh is a major challenge. Existing methods are typically limited to structured rectangular meshes, such as 30m x 30m cartesian grids for high-resolution watershed-scale modeling or 0.5 degree x 0.5 degree geographic grids for global climate modeling. In PyFlowline, we define structured meshes (e.g., lat-lon, raster files with projections, and hexagon) as those with fixed cell sizes and shapes and unstructured meshes as those with variable cell sizes and shapes.

Structured mesh-based methods use fixed cartesian or geographic cell sizes, which have several limitations: (1) they cannot resolve fine-scale river networks at coarse cell resolutions (>1km), and (2) they cannot be seamlessly coupled with other unstructured mesh-based hydrologic models such as oceanic models ([Engwirda & Liao, 2021](#)). In contrast, unstructured meshes offer a flexible structure with variable grid-cell sizes and shapes. This flexibility makes them ideal for adapting to complex geometry such as river channels and coastlines. Besides, unstructured meshes provide the flexibility to couple different hydrologic models under a unified framework. Thus, unstructured meshes are increasingly being adopted in hydrologic modeling.

Although unstructured meshes offer these flexibilities, additional efforts are required to generate conceptual river networks that capture real-world river networks across different spatial scales.

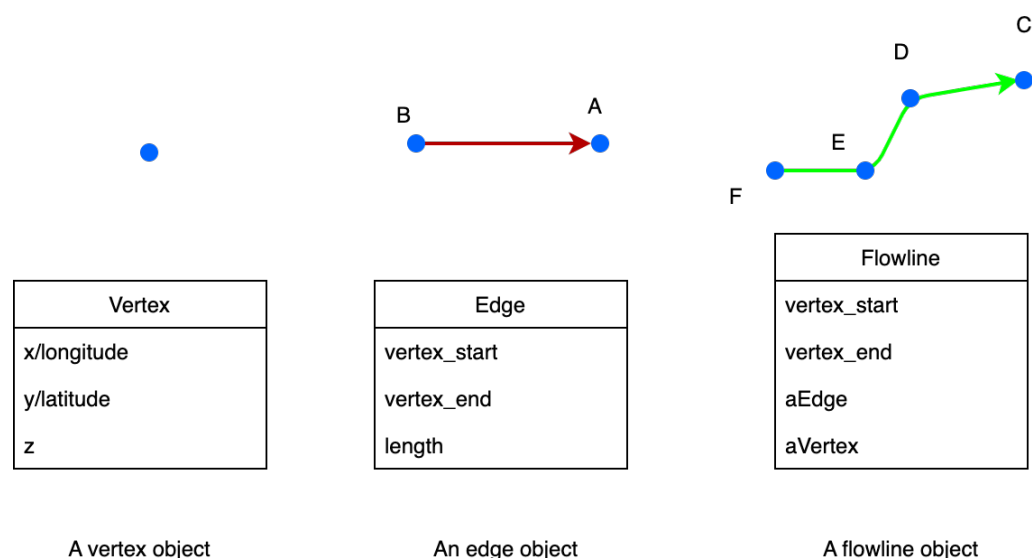
A mesh-independent river network representation method that preserves topological relationships across scales could address this limitation. PyFlowline is a Python package that provides a framework for generating river networks for hydrologic models, meeting the identified need. Using an object-oriented programming approach, PyFlowline represents river network elements and mesh cell relationships. It relies on open-source Python libraries like GDAL and Cython for

data input/output and spatial data operations.

The computational geometry algorithms used in PyFlowline are designed and implemented using a unified spherical framework, making it suitable for regional and global-scale simulations. PyFlowline uses topological relationships to capture the river networks so they are preserved even at coarse spatial resolutions. Moreover, PyFlowline is mesh-independent, supporting both structured and unstructured meshes. It can quickly adopt other mesh types, such as triangulated irregular networks (TIN) or discrete global grid systems (DGGs) (Sahr, 2011). PyFlowline is a core component of the HexWatershed model, a mesh-independent flow direction model. Several scientific studies focusing on coupled Earth system models (Feng et al., 2022; Liao, Zhou, Xu, Cooper, et al., 2023; Liao, Zhou, Xu, Tan, et al., 2023) have utilized PyFlowline. A workshop tutorial has also been provided online and in person to support its implementation.

## Model features

PyFlowline uses Python's object-oriented programming (OOP) architecture to describe river networks using three essential elements: segments, reaches, and confluences. When applicable, river networks are processed as objects throughout the package.



**Figure 1:** The data model. A vertex class object represents a point on the Earth surface. It has three coordinates. An edge class object represents a directed line between two points. Besides, it has a length attribute. A flowline class object represents a list of connected lines.

PyFlowline provides several key features, including

1. Support for both structured and unstructured meshes, with JSON as the default file I/O format. For geospatial datasets such as vector river networks, GEOJSON is used.
2. Regional and global-scale processing capabilities through the use of fast Cython- and (global-only) AABB tree-based algorithms.
3. Built-in visualization functions (experimental) based on the Python Matplotlib package (Liao, 2022), making it easy to visualize and analyze the PyFlowline model outputs.

## State of the field

Existing river network representation methods often fall into three categories, each with associated strengths and weaknesses:

1. Vector-based. Hydrologic models that use this method can represent fine-scale river networks, but cannot couple river and land without one-to-one mapping between river segments and land model cells (Schwenk & Hariharan, 2021);
2. High-resolution DEM-based. Flow networks derived from structured rectangle-grid DEMs are widely available, but resolving fine-scale river networks requires grids with very high spatial resolution (e.g., 30m x 30m or finer) (Esri Water Resources Team, 2011);
3. Upscaling-based methods address the scale mismatch between coarse-resolution grids and fine-scale river networks, but only support structured geographic grids (e.g., 0.5 degree x 0.5 degree) at coarse resolutions (Wu et al., 2012). This method often cannot provide global coverage, including Greenland and the Antarctic.

PyFlowline is the only modeling software that provides these unique features:

1. It can generate river networks on unstructured meshes;
2. It uses topological relationships to capture river networks precisely;
3. It can be applied at both high and coarse resolutions;
4. It can provide global coverage, including Greenland and the Antarctic.

Model documentation is hosted at <https://pyflowline.readthedocs.io/en/latest/>, including a case study for the Susquehanna River Basin in the Mid-Atlantic region of the United States.

## Acknowledgment

The model described in this repository was supported by the following:

- the Earth System Model Development and Regional and Global Model Analysis program areas of the U.S. Department of Energy, Office of Science, Biological and Environmental Research program as part of the multi-program, collaborative Integrated Coastal Modeling (ICoM) project.
- the Earth System Model Development and Regional and Global Model Analysis program areas of the U.S. Department of Energy, Office of Science, Biological and Environmental Research program as part of the multi-program, collaborative Interdisciplinary Research for Arctic Coastal Environments (InterFACE) project.
- the Next Generation Ecosystem Experiments-Tropics project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research at Pacific Northwest National Laboratory.

A portion of this research was performed using PNNL Research Computing at Pacific Northwest National Laboratory.

PNNL is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830.

## References

- Engwirda, D., & Liao, C. (2021). 'Unified' Laguerre-Power Meshes For Coupled Earth System Modelling. *29th International Meshing Roundtable (IMR), Virtual Conference*. <https://doi.org/10.5281/zenodo.5558988>
- Esri Water Resources Team. (2011). *Arc Hydro Tools - Tutorial [Software]*.
- Feng, D., Tan, Z., Engwirda, D., Liao, C., Xu, D., Bisht, G., Zhou, T., Li, H.-Y., & Leung, L. R. (2022). Investigating coastal backwater effects and flooding in the coastal zone using a global river transport model on an unstructured mesh. *Hydrology and Earth System Sciences*, 26(21), 5473–5491. <https://doi.org/10.5194/hess-26-5473-2022>
- Liao, C. (2022). *PyEarth: A lightweight Python package for Earth science* (Version 0.1.20). Zenodo. <https://doi.org/10.5281/zenodo.6109987>

- Liao, C., Zhou, T., Xu, D., Cooper, M. G., Engwirda, D., Li, H.-Y., & Leung, L. R. (2023). Topological Relationship-Based Flow Direction Modeling: Mesh-Independent River Networks Representation. *Journal of Advances in Modeling Earth Systems*, 15(2), e2022MS003089. <https://doi.org/10.1029/2022MS003089>
- Liao, C., Zhou, T., Xu, D., Tan, Z., Bisht, G., Cooper, M. G., Engwirda, D., Li, H.-Y., & Leung, L. R. (2023). Topological Relationship-Based Flow Direction Modeling: Stream Burning and Depression Filling. *Journal of Advances in Modeling Earth Systems*, 15(11), e2022MS003487. <https://doi.org/10.1029/2022MS003487>
- Mizukami, N., Clark, M. P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., Viger, R. J., Markstrom, S. L., Hay, L. E., Woods, R., Arnold, J. R., & Brekke, L. D. (2016). mizuRoute version 1: A River Network Routing Tool for a Continental Domain Water Resources Applications. *Geoscientific Model Development*, 9(6), 2223–2238. <https://doi.org/10.5194/gmd-9-2223-2016>
- Sahr, K. (2011). User documentation for discrete global grid generation software. *Southern Oregon Univ., Ashland, OR, USA, Tech. Rep. Dggrid Version, 3*.
- Schwenk, J., & Hariharan, J. (2021). RivGraph: Automatic extraction and analysis of river and delta channel network topology. *Journal of Open Source Software*, 6(59), 2952. <https://doi.org/10.21105/joss.02952>
- Wu, H., Kimball, J. S., Li, H., Huang, M., Leung, L. R., & Adler, R. F. (2012). A new global river network database for macroscale hydrologic modeling. *Water Resources Research*, 48(9). <https://doi.org/10.1029/2012WR012313>