Topological relationships-based flow direction modeling: mesh-independent river networks representation

Chang Liao[1], Tian Zhou[1], Donghui Xu[1], Matt Cooper[1], Darren Engwirda[2], Hong-Yi Li[3], L. Ruby Leung[1]

[1]Atmospheric Sciences and Global Change, Pacific Northwest National Laboratory, Richland, WA, USA
[2]T-3 Fluid Dynamics and Solid Mechanics Group, Los Alamos National Laboratory, Los Alamos, NM, USA
[3]University of Houston, Houston, TX, USA

Key Points:

- We use graph theory algorithms to simplify real-world river networks.
- Topological relationships are reconstructed from simplified river networks and mesh intersections.
- Topological relationships can be used to model flow direction field and flow routing parameters.

Corresponding author: Chang Liao, chang.liao@pnnl.gov

Abstract

River networks are important features in surface hydrology. However, accurately representing river networks in spatially distributed hydrologic and Earth system models is often sensitive to the model's spatial resolution. Specifically, river networks are often misrepresented because of the mismatch between the model's spatial resolution and river network details, resulting in significant uncertainty in the projected flow direction. In this study, we developed a topological relationships-based river network representation method for spatially distributed hydrologic models. This novel method uses (1) graph theory algorithms to simplify real-world vector-based river networks and assist in mesh generation; and (2) a topological relationship-based method to reconstruct conceptual river networks. The main advantages of our method are that (1) it combines the strengths of vector-based and DEM raster-based river network extraction methods; and (2) it is mesh-independent and can be applied to both structured and unstructured meshes. This method paves a path for advanced terrain analysis and hydrologic modeling across different scales.

Plain Language Summary

Representing rivers in hydrologic models is difficult because river networks are often very complex. Existing methods generally rely on elevation differences between land and rivers or image processing to define river networks in computer models. In this study, we combine the strengths of two existing methods and develop a topology-based method. This follows river channels and defines river networks in a way that works for any grid system and spatial resolution. The products of this method can be used to improve hydrologic models.

1 Introduction

River networks are important features in hydrologic and Earth system modeling(Jolley & Wheater, 1997; Wu et al., 2012; Liao et al., 2019). Real-world river networks are complex, depending on landscape features such as elevation, aspect, and lithology. Moreover, the fractal nature of river networks means that they are approximately scale-free(Tarboton et al., 1988) without a well-defined spatial resolution at which they should be represented(Davies & Bell, 2009; Yamazaki et al., 2009). As a result, hydrologic models often use a conceptual representation of river networks. To date, there are mainly three methods for representing river networks, each with its own advantages and disadvantages. However, limitations remain when representing river networks due to the resolution mismatch and their interactions with other hydrologic features (e.g., ocean).

There are two methods useful at the watershed/regional scale, the first of which is the vector-based river networks analysis method. This method uses vector datasets to represent the river networks and their topological relationships(Mizukami et al., 2016; Lin et al., 2018). Vector datasets are often provided by public agencies, e.g., the United States Geological Survey (USGS), or are digitized from satellite image processing, e.g., vectorization-based river channel extraction. Various graph theory algorithms are then used to perform quality control and network analysis based on the vector river networks(Lindsay et al., 2019). Lindsay, et al. reviewed several potential issues in existing vector river networks datasets: (1) the vertex coordinates of river flowlines may not exactly overlap with the actual locations due to digitization error and floating-point rounding error. (2) the starting and ending-vertices of a river flowline may be reversed during spatial analysis, resulting in an opposite flow direction. (3) vector datasets obtained from different sources generally use different spatial references and cannot be used or combined directly. Even without these issues, vector datasets still require other pre-processing before use. For example, braided rivers are not universally supported in hydrologic models as multiple flow

directions are not always supported. Although the vector-based river networks representation method is computationally efficient and scale-independent, it has limitations. First, the resulting vector networks are not explicitly linked with the rectangle mesh system commonly used in hydrologic models. As a result, the vector-based river networks method is limited to flow routing and cannot be easily coupled with other hydrologic processes. Second, this method can only be applied to areas with vector datasets unless it's combined with advanced terrain analysis.

The second method is the raster digital elevation model (DEM)-based river networks extraction method(Tarboton, 2003; Esri Water Resources Team, 2011; Yamazaki et al., 2019). This method generally involves several steps: (1) calculation of cell-to-cell flow direction for each raster cell based on elevation differences (e.g., D4/D8 algorithm); (2) calculation of flow accumulation based on flow direction; and (3) definition of river cells using an accumulation threshold. Because of local depressions in the DEM, a depression-filling operation is often required before step 1 to guarantee that water can flow out of each cell. The raster DEM-based river networks extraction method is widely used with a few limitations. First, this method is very sensitive to the spatial resolution and accuracy of the DEM. In general, it only performs well for high-resolution ($< 1\,\text{km}$) DEMs(Goulden et al., 2014; Sood & Smakhtin, 2015). Second, because the derived flow direction in step 1 relies on elevation differences, it is less accurate in flat areas with fewer topographic variations. As a result, obtaining unrealistic river networks is common. To address these limitations, the "stream burning" technique (also called "DEM reconditioning") is often used to lower elevations within and near river channels so that water always flows into the river cells(Hellweger & Maidment, 1997; Lindsay, 2016a). This technique requires an additional river network dataset, typically in vector format. The user-provided river networks dataset is often converted into a binary mask with the same spatial resolution as the DEM(Lindsay, 2016b). Because the binary mask does not describe the upstream-downstream topological relationships between mesh cells, it cannot accurately capture the meander, confluence, and parallel river. In this study, we define parallel rivers as rivers running in adjacent mesh cells side by side. As a result, it often produces incorrect flow directions in these locations (Figures S1-S6). The extensive modifications to the elevations of river networks and riparian-zone can result in large biases in slope calculations (river channel slope and riparian zone slope), which will significantly impact the flow routing and flooding processes in hydrologic models(Shelef & Hilley, 2013). To remediate this, slopes are sometimes calculated from the original DEM(Lindsay, 2016b), but this requires additional adjustments to consider local depressions. Similar issues also arise in the terrain analysis depression filling process. Several studies used additional topological information in the hybrid breaching-filling algorithm to minimize modifications to elevation(Lindsay, 2016a). However, this hybrid approach is not readily available and is not used in most Geographic Information System (GIS) or hydrologic models. The raster DEM-based method is often not directly used at large spatial domains due to the computational cost. Instead, the study domain is often separated into small tiles, which are merged together after applying the method.

The third method is used on a continental to global scale. Because the target spatial resolution (around 10 km to 200 km) is much coarser compared to regional-scale applications(Sood & Smakhtin, 2015), this method often assumes that there is always one major river channel and many tributaries within each mesh cell(Wu et al., 2011). Due to the resolution mismatch between fine-scale river networks and mesh cell size, their topological relationships are even more complex(Shaw et al., 2005). For example, multiple major rivers may run in the same cell or in parallel. To maintain the large-scale flow pattern, many models use high spatial resolution datasets (e.g., results from the raster DEM-based method) as guidance to define the coarse resolution cell-to-cell flow direction(Fekete et al., 2001; Davies & Bell, 2009; Wu et al., 2011). This type of method is also often referred to as the upscaling method. For example, the Cell Outlet Tracing with an Area Threshold (COTAT)(Reed, 2003) and Network Tracing Method (NTM)(Wu et al., 2012) use either high-resolution

raster or vector-based river networks to guide their coarse resolution flow directions. Without increasing the mesh resolution, the upscale method often needs to modify the major river locations to avoid aforementioned issues(Wu et al., 2012; Eilander et al., 2021; Munier & Decharme, 2022). Although this method captures the flowlines, the shift of locations can result in unrealistic spatial distribution of model outputs, for example, when floodplain inundation is of particular interest(Yamazaki et al., 2011; Decharme et al., 2012; Luo et al., 2017; Mao et al., 2019; Zhou et al., 2020). One possible solution is to use an unstructured mesh with regional refinement for the regions with complex river networks; however, all existing model implementations only support the structured rectangle meshes. Taken together, the existing three methods have different advantages and disadvantages, and they are often used in different scales, applications, and hydrologic communities (Table S1).

To date, all the spatially-distributed flow routing models (except vector-based) are limited to the structured rectangle meshes, and cannot be seamlessly coupled with other unstructured mesh-based numerical models. Moreover, existing methods mainly focus on projecting existing river networks onto prescribed structured rectangle meshes. Less attention has been paid to unstructured meshes, which allow certain hydrologic features, such as river networks, to be burnt into the meshes. Many studies have attempted to "burn" river networks into meshes, with most using the Triangular Irregular Network (TIN) approach(Kreveld & Silveira, 2011; Coon et al., 2019). However, existing TIN-based methods do not generally incorporate stream burning or depression filling methods(Ivanov et al., 2004; Coon et al., 2019). As a result, there is still uncertainty in flow direction and slope calculations.

In recent years, model development based on unstructured meshes is an emerging area of interest in hydrologic and Earth system models because it provides several advantages(Engwirda & Liao, 2021): (1) unstructured mesh refinement can be used to define specific regions of interest (ROIs). Because hydrologic features such as river networks, dams, and coastal lines do not align well with the rectangle meshes used in numerical models, spatial interpolation and approximation are often needed. In contrast, unstructured meshes provide the flexibility to represent these features reasonably well through variable size and rotation, resulting in reduced model uncertainty. (2) spatial interpolation will be significantly reduced or removed in a unified unstructured mesh framework that includes all the model components, e.g., ocean, land, and river. These components can exchange fluxes seamlessly at their interfaces(Liao et al., 2022). (3) unstructured meshes can be used to balance spatial resolution and computational cost through variable resolution, critical for large-scale hydrologic and Earth system models.

To the authors' knowledge, there is no river networks representation method designed for the unstructured mesh system(Paz & Collischonn, 2007; Sood & Smakhtin, 2015; de Azeredo Freitas et al., 2016; Hyvaluoma, 2017; Hsu, 2020). For hydrologic models and Earth system models, this requires a new river networks representation and flow direction method that supports unstructured meshes. In this two-part study, we introduce a novel method that combines the strengths of existing methods to produce river networks and flow direction for any mesh system. In part 1, we mainly focus on the topological relationship-based river network representation model (PyFlowline)(Liao & Cooper, 2022). In part 2, we will demonstrate how to use the topological relationship in depression filling and flow direction modeling in the HexWatershed model(Liao et al., 2020, 2022). Part 1 of the study is organized as follows. We first introduce the model algorithms. Then we apply this model to a coastal watershed, the Susquehanna river basin (SRB), using different model configurations, and evaluate the model performance against real-world river networks using several metrics including river length and area of differences. Last, we discuss the limitations and future applications in hydrologic and Earth system models.

2 Methods

2.1 Overview

Conceptually, any river channel can be represented using three basic graph elements: vertex, edge, and flowline (Figure S7). River networks can be viewed as collections of these three elements. After converting existing river network datasets into these basic elements, we can then use graph algorithms to extract topological relationships including connectivity and direction. For notation, a single letter/number will be used for the vertex (e.g., vertex A) and a sequence of letters/numbers will be used for the directed edge and flowline (e.g., edge A->B). The essence of our method is that in the simplest scenario, in any type of mesh system, a river channel always intersects (enters and exits) a mesh cell on two different edges (or, less likely, its vertices) (Figure 1), unless it is either a headwater or river mouth in which either the starting or ending vertex lies within the mesh cell.

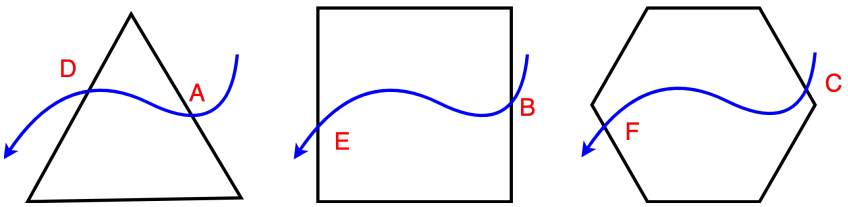

Figure 1. Illustration of a river channel entering and exiting a mesh cell from its edges. Blue curves are river channels. Black polygons are mesh cells (triangle, square, and hexagon). Regardless of the cell type, the intersection between a river channel and a mesh cell always results in two vertices on mesh cell edges, such as A/D, B/E, and C/F pairs, and a directed flowline within, such as A->D, B->E, and C->F.

Based on the intersections between the river channels and mesh system, real-world river networks can be represented digitally as collections of the individual "reach" within each mesh cell (e.g., flowline A->D and B->E in Figure 1). Based on the intersections, topological relationships (e.g., which cells are upstream of the current cell) between mesh cells can be built. As a result, river networks can be consistently preserved in any mesh system. To achieve this, our method consists of several major steps, illustrated in Figure 2 and described in detail in the following sections.

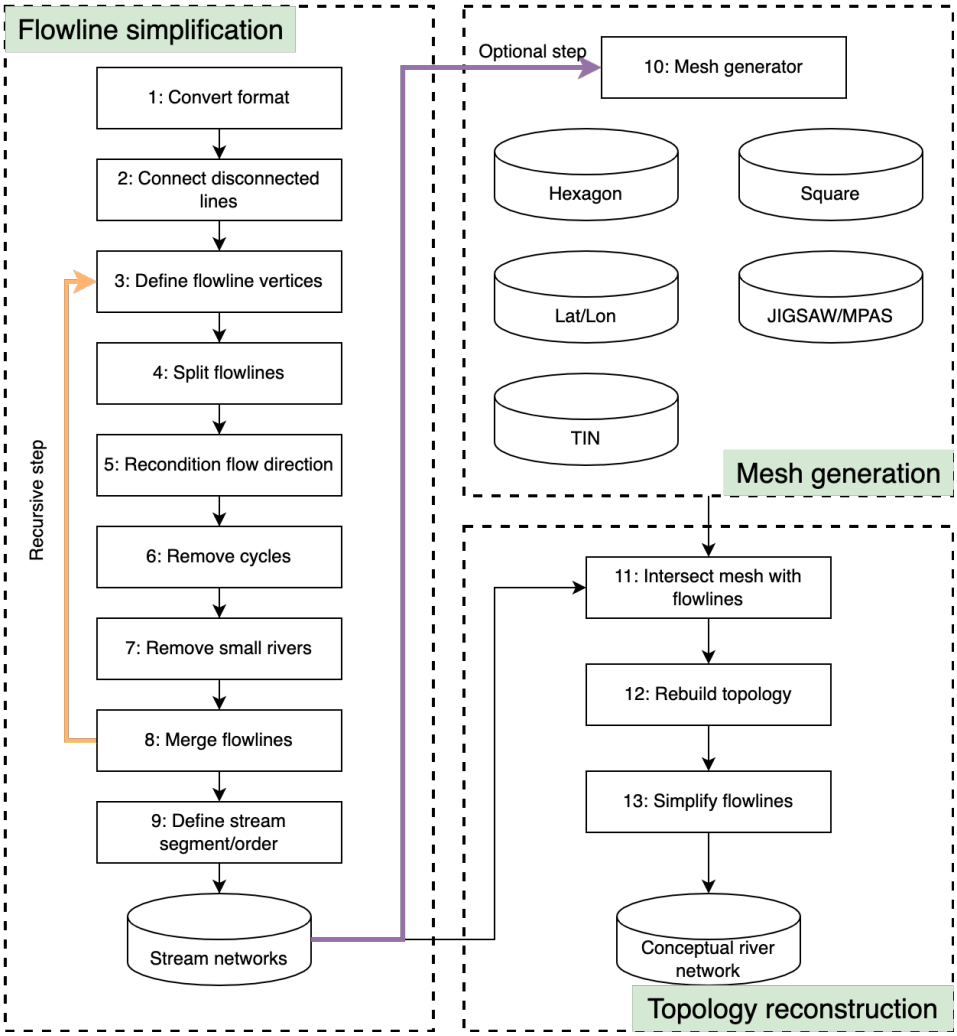

Figure 2. The workflow of the topological relationship-based river networks representation method. The workflow includes three major components: flowline simplification, mesh generation, and intersection-based topology reconstruction. Each component contains one or more steps with indices. In flowline simplification, several steps (highlighted by the yellow arrow) may be run recursively. The output from flowline simplification can be optionally used for mesh generation (purple arrow). Both structured (e.g., rectangle lat-lon and hexagon) and unstructured meshes (e.g., MPAS mesh) are supported. Step 13 is a combination of several steps from Step 3 to 9.

Besides, because our method is completely built upon the geodesic coordinate system (Text S1), it can be applied to both regional and global scales.

2.2 Flowline simplification

To address potential issues in existing vector datasets, we developed a list of algorithms (Steps 1 to 9 in Figure 2) to pre-process the vector flowlines. In practice, these algorithms may be run in different combinations and orders depending on the datasets. From now on flowline pre-processing, as either individual steps or a collection of several steps, is referred to as flowline simplification. And the outputs from flowline simplification are referred to as simplified flowlines. Details of the algorithms in Steps 1 to 9 are

provided in the supplementary materials (Text S2). Additionally, our method provides the option to burn dams and associated flowlines in the workflow. For flowline simplification, this means that other than the prescribed high-order flowlines, the model also includes all the downstream flowlines of the user-provided dams.

2.3 Mesh generation

Mesh generation is the process that defines the spatial domain discretization(Liao et al., 2020). Structured meshes such as the geographic coordinate system (GCS) rectangle mesh (latitude-longitude), projected coordination system (PCS) rectangle mesh (square), and hexagonal mesh, can be generated using various GIS computer programs (e.g., QGIS MMQGIS)(Minn, 2015) or with Python scripting(Liao et al., 2022). This usually involves the following steps: (1) obtaining the spatial extent of the study domain; (2) setting the lower-left or upper-left corner as the origin; (3) calculating the number of rows and columns based on the desired resolution and spatial extent; (4) calculating the vertex coordinates of each mesh cell; and (5) exporting all the mesh cells to a GIS file format.

For unstructured meshes such as the TIN or the Model for Prediction Across Scales (MPAS) mesh (Text S3)(Ringler et al., 2013), advanced mesh generators such as JIGSAW are available(Engwirda, 2014; Sahr, 2015). These mesh generators can generally apply mesh refinement near some ROIs, such as river networks and/or coastal areas. Some generators also allow mesh grid centers to align with predefined polylines (e.g., river flowlines or coastal lines) and points (e.g., dams) to satisfy particular modeling needs. In our study, we mainly used the JIGSAW mesh generator to produce an MPAS mesh. Specifically, we used simplified flowlines, dam locations, and coastal lines during the mesh refinement process so that the MPAS mesh cells align with these hydrologic features(Engwirda, 2017).

To maintain consistency, we use "equivalent" resolution which is the square root of the mesh cell spherical sector area, to define mesh resolution (Equation 1, and Text S1)(Liao et al., 2022). For a structured GCS rectangle mesh, the equivalent resolution changes with latitude. For unstructured variable resolution meshes such as TIN and MPAS meshes, the equivalent resolution only applies to a certain mesh cell.

$$\mathrm{Res}_c = \sqrt{\mathrm{Area}_c} \tag{1}$$

where Res_c is the "equivalent" resolution (m); and Area_c is the mesh cell spherical sector area (m^2).

Mesh generation is not the focus of this study and has been extensively explored in relevant communities(Minn, 2015; Sahr, 2015; Engwirda, 2017). In this study, we use relatively simple meshes and existing mesh generators as test cases to demonstrate our new method.

2.4 Topological relationships reconstruction

Intersecting the simplified flowlines and meshes breaks stream segments into stream reaches. For simplicity, we redefine the simplified flowlines using stream segments and stream reaches (Figure S12). Because both the stream segment index and stream order information are defined during the flowline simplification step (Step 9 in Figure 2), they are directly assigned to each stream reach. A mesh cell may contain one or more internal reaches if it intersects with the simplified flowlines.

Reconstruction of the mesh cell topological relationships starts from the user-provided approximate outlet location (latitude and longitude coordinates) and searches from the outlet to upstream headwater in reverse. Figure 3 illustrates this process in the simplest scenario with only three stream segments in a rectangle mesh. The model also allows a

segment to take a "shortcut" (Red dashed edge F->I in Figure 3) when it passes a mesh cell in a short distance (Reach 3->2 within cell E), which produces the classical D8 diagonal travel path (only in the rectangle mesh). This short distance is defined as a "shortcut" threshold parameter (Table S2). The resulting topological relationships are stored and expressed as conceptual flowlines, which connect one cell center to another (blue and red dashed arrows/edges).

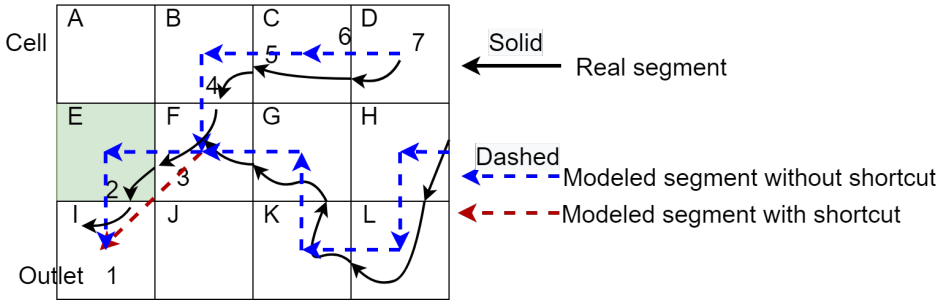

Figure 3. Illustration of the topological relationships reconstruction in a 3-row by 4-column rectangle mesh. The letters in the top-left corner of each cell, from A to L, represent mesh cell IDs. The numbers from 1 to 7 represent the start, end, and intersected vertices. Colored solid arrows are intersected real stream reaches within each mesh cell. Starting from outlet vertex 1, the algorithm searches reversely and reconstructs the cell topological relationships as D->C->B->F, H->L->K->G->F, and F->E->I in blue dashed arrows. Optionally, if the D8 diagonal path is turned on because the reach (3->2) length within grid E is less than the user-provided threshold, the algorithm omits the grid E and takes a shortcut (red dashed arrow). Because the algorithm strictly follows topology relationships, it precisely captures confluence (Grid F), parallel river (C->B and G->F), and meander (H->L->K->G).

With these capabilities, the model can represent the river meander, confluence, and parallel river in any mesh system (Figures S4-S6). In some cases, a stream segment may enter and exit a mesh cell multiple times on the same cell edge, which results in a "cycling" effect (e.g., F->E->B->E->D in Figure S10). To address this issue, the cycling removal algorithm (Step 6 in Figure 2) is applied again to remove the loop (e.g., the final topological relationships are F->E->D). It is also possible that multiple conceptual flowlines enter and exit the same cell. Therefore, some flowline simplification algorithms are reused (Step 13 in Figure 2).

3 Model application

3.1 Study area

The Susquehanna river basin (SRB) is a major river basin located in the Mid-Atlantic region of the United States. The total drainage area of the SRB is about 7.1×10^4 km^2. Its surface elevation ranges from 0 m to more than 900 m. It contains both relatively mild and steep surface slopes, ranging to 30 degrees in some areas (Figure 4). Spatial datasets and maps were produced using the Python packages including Matplotlib and GDAL(Hunter, 2007; Gillies & others, 2007; GDAL/OGR contributors, 2019; Liao, 2022).

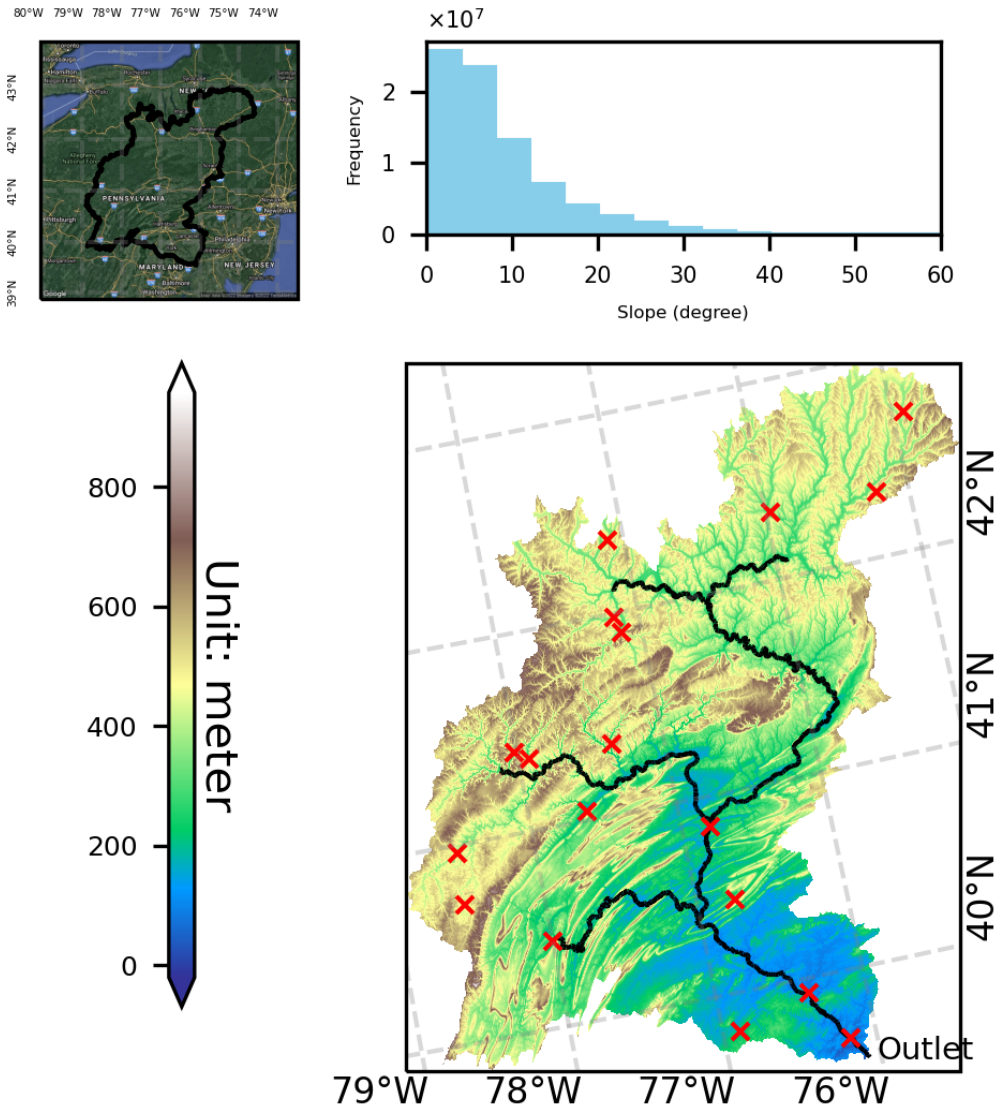

Figure 4. The spatial location, surface elevation, and surface slope distribution (based on DEM) of the Susquehanna river basin. The upper left red polygon is the Watershed Boundary Dataset watershed boundary on Google Maps; the upper right is the histogram of surface slope (degree), and the bottom is the topographic map (m). In the topographic map, the black lines are major river channels. The red crosses are major dams. The outlet is in the lower right corner.

Because the river outlet of the SRB is into Delaware Bay, the study area is prone to sea-level rise, storm surge, and other extreme event-induced flooding. Due to the resolution mismatch and mesh differences between the land, river, and ocean in Earth system models, the study area is generally poorly represented, especially near the coastal lines(Feng et al., 2022).

3.2 Data

We collected river networks, watershed boundary, and dam locations from the United States National Hydrography Dataset Plus High Resolution (NHD Plus HR) and Watershed Boundary Dataset(USGS, 2013). Because NHDPlus HR contains more than 10,000 flowlines in the SRB (Figure 5), we used stream order (higher and equal to 6 or 7) to reduce the number to 120 (Figure 5)(Tarboton et al., 1991).

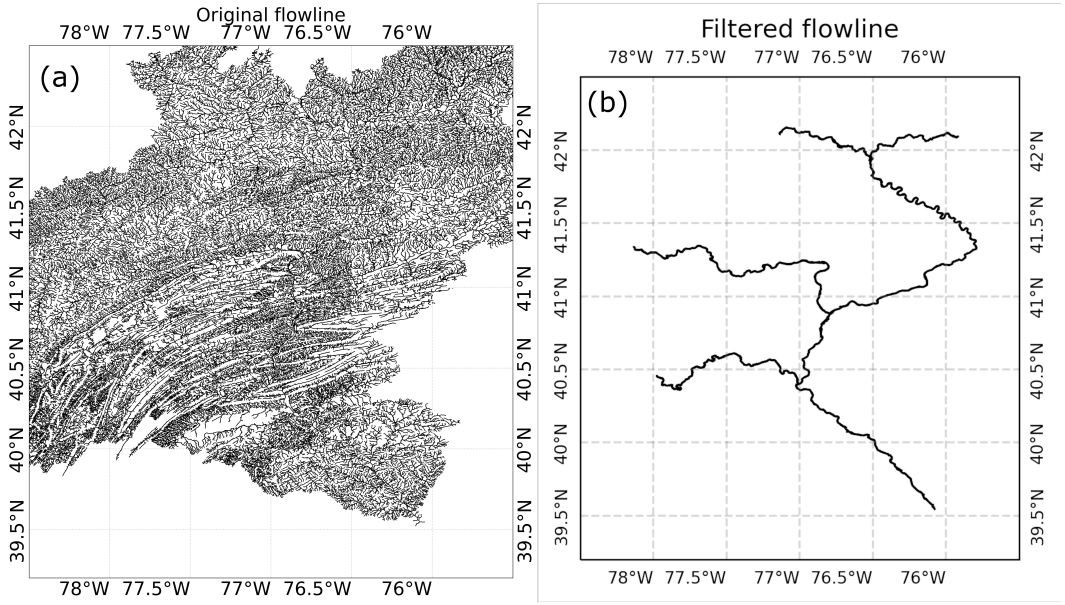

Figure 5. The spatial distribution of river flowlines from NHDPlus HR for the Susquehanna river basin. (a) is the raw data with all flowlines (more than 10k flowlines) and (b) shows filtered flowlines with stream orders higher or equal to 7. A total of 120 river flowlines are retained.

3.3 Model setup

Major model configuration parameters are listed in Table S2. To evaluate the sensitivity of model performance to spatial resolution and meshes, we ran the model with different configurations with case indices used for illustrations (Table 1).

Table 1. Simulation configurations with case indices. The illustrations and analyses all use the same indices.

Mesh/Resolution	50 km	10 km	5 km
Lat-lon	1	2	3
Square	4	5	6
Hexagon	7	8	9
MPAS	10/11/12		

For the structured meshes, i.e., GCS rectangle mesh (lat-lon), PCS rectangle mesh (square), hexagon, we ran 3 different spatial resolutions (50 km, 10 km, and 5 km) with the same stream order threshold 7. The square mesh Case 9 results are used for comparison with the traditional raster DEM-based method (Text S4, Figure S13).

For the unstructured variable resolution ($3 \sim 10$ km) MPAS mesh, we ran 3 cases: (1) Case 10 with stream order threshold 7 (the same with structured-mesh cases); (2) Case 11 with stream order threshold 6. This case is used to evaluate the performance at different levels of flowline details; (3) Case 12 with stream order threshold 7 and additional dams burnt in the flowlines and mesh. This case is used to illustrate the dam burning capability.

3.4 Results and analysis

3.4.1 Flowline simplification

After model simulation, the river flowlines were substantially simplified. First, the total number of flowlines is reduced from 120 to 7, with 4 headwater vertices, 3 confluence vertices, and 1 outlet vertex (Figure 6). Second, all the small rivers with lengths less than the threshold were removed (Figure 7). Third, all braided rivers were removed (Figure 7). Flowline simplification results are the same from Cases 1 to 10.

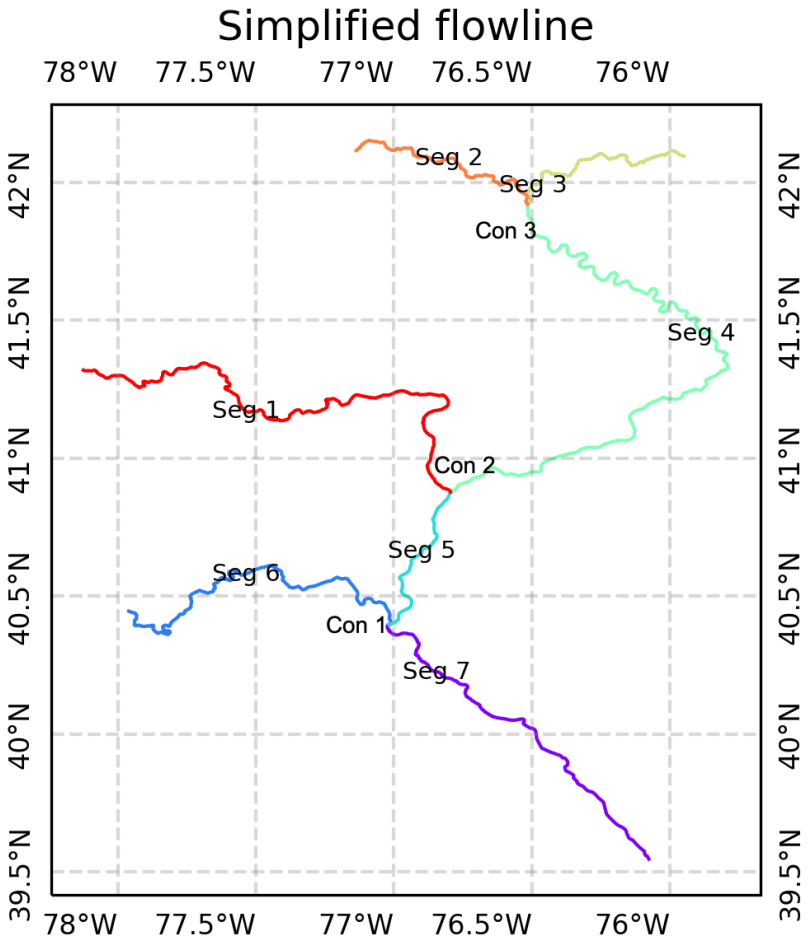

Figure 6. The spatial distribution of flowlines after flowline simplification. (a) Colored line features are 7 flowlines after simplification. Segments 1 to 7 are segment indices. Confluences 1 to 3 are confluence indices.

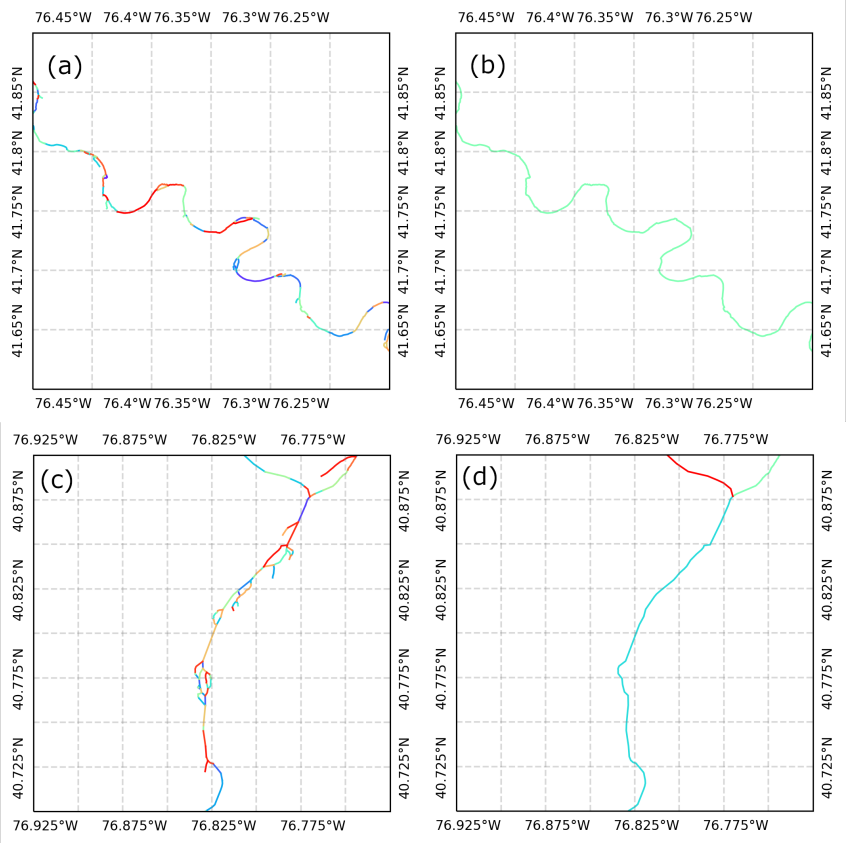

Figure 7. Before and after flowline simplification comparisons. (a) and (b) are zoomed-in views where small rivers are removed near the river meanders; (c) and (d) are zoomed-in views where braided and small rivers are removed.

After removing small and braided rivers, the total length of the flowlines decreased 10.0% (dashed red and blue lines in Figure 8).

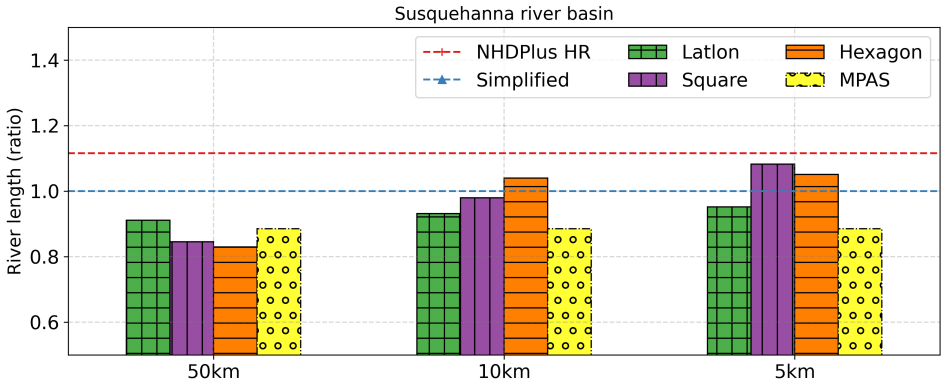

Figure 8. Comparison of the total flowline length from different model configurations. The X-axis is the spatial resolution and the Y-axis is the ratio between modeled and simplified river length. The dashed red and blue lines are the NHDPlus HR and simplified total flowline length. Colored (green, purple, orange, and yellow) bars with textures represent different model configurations.

Depending on user needs, a combination of stream order and dam information can be used in PyFlowline to preserve different levels of details (e.g., tributaries). For example, the simplified flowlines from Cases 11 and 12 are illustrated in Figure S11.

3.4.2 Mesh generation

The generated structured GCS rectangle lat-lon, PCS square, and PCS hexagon meshes are illustrated in Figure 9.

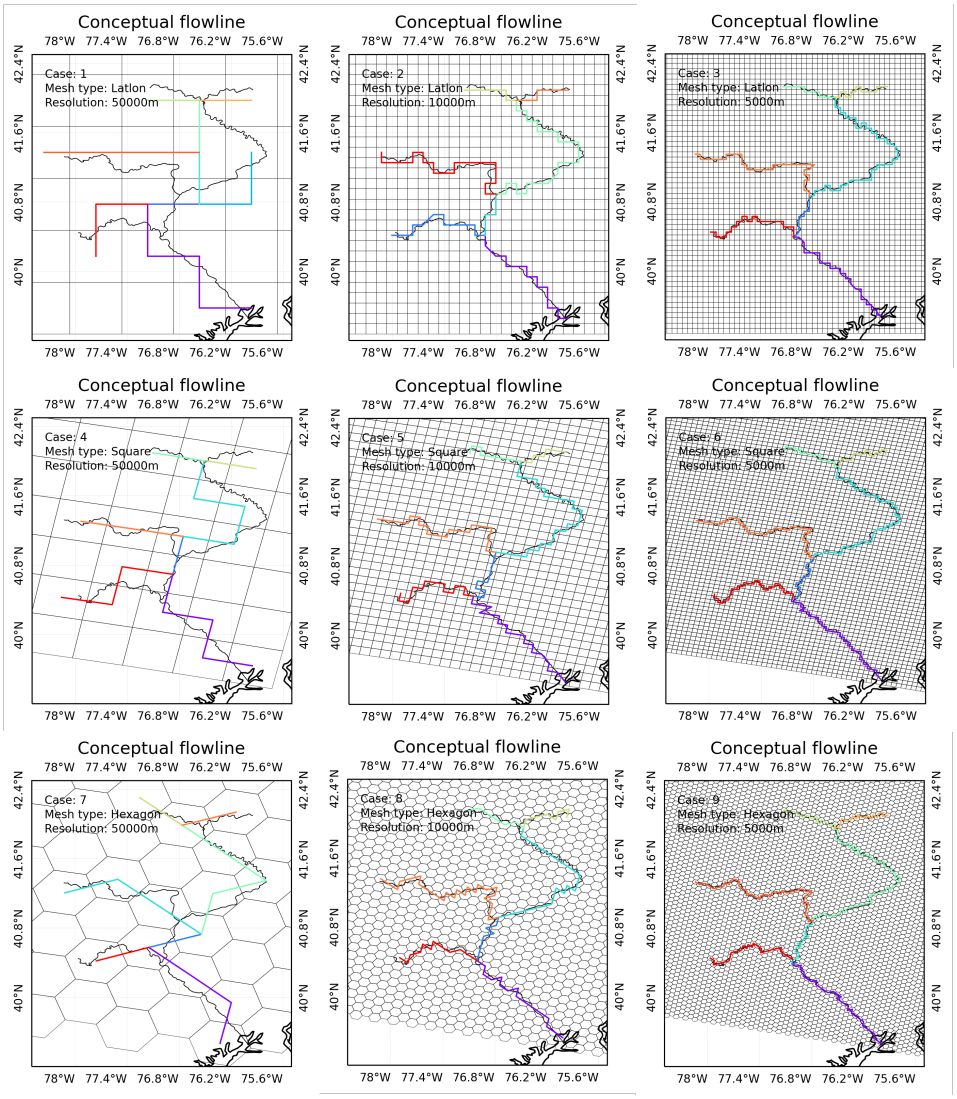

Figure 9. The spatial distributions of modeled conceptual flowlines on the lat-lon, square, and hexagon meshes at multiple spatial resolutions. The layout of the figures and Case indices from 1 to 9 is the same as Table 1. Black line features represent the simplified flowlines. Colored line features represent the conceptual flowlines.

The generated MPAS mesh is illustrated in Figure 10. Because both the small and braided rivers were removed, JIGSAW successfully aligned cell centers following the simplified flowlines at the desired resolution. JIGSAW also varies cell resolution by considering the distance of the cell center to the simplified flowlines (as well as the coastal lines) through a density function. In this case, a coastal-resolving mesh spacing function is adopted to cluster high model resolution near coastlines. Mesh resolution varies between 3 km near the river outlet/coastal line and 10 km near the domain boundary.

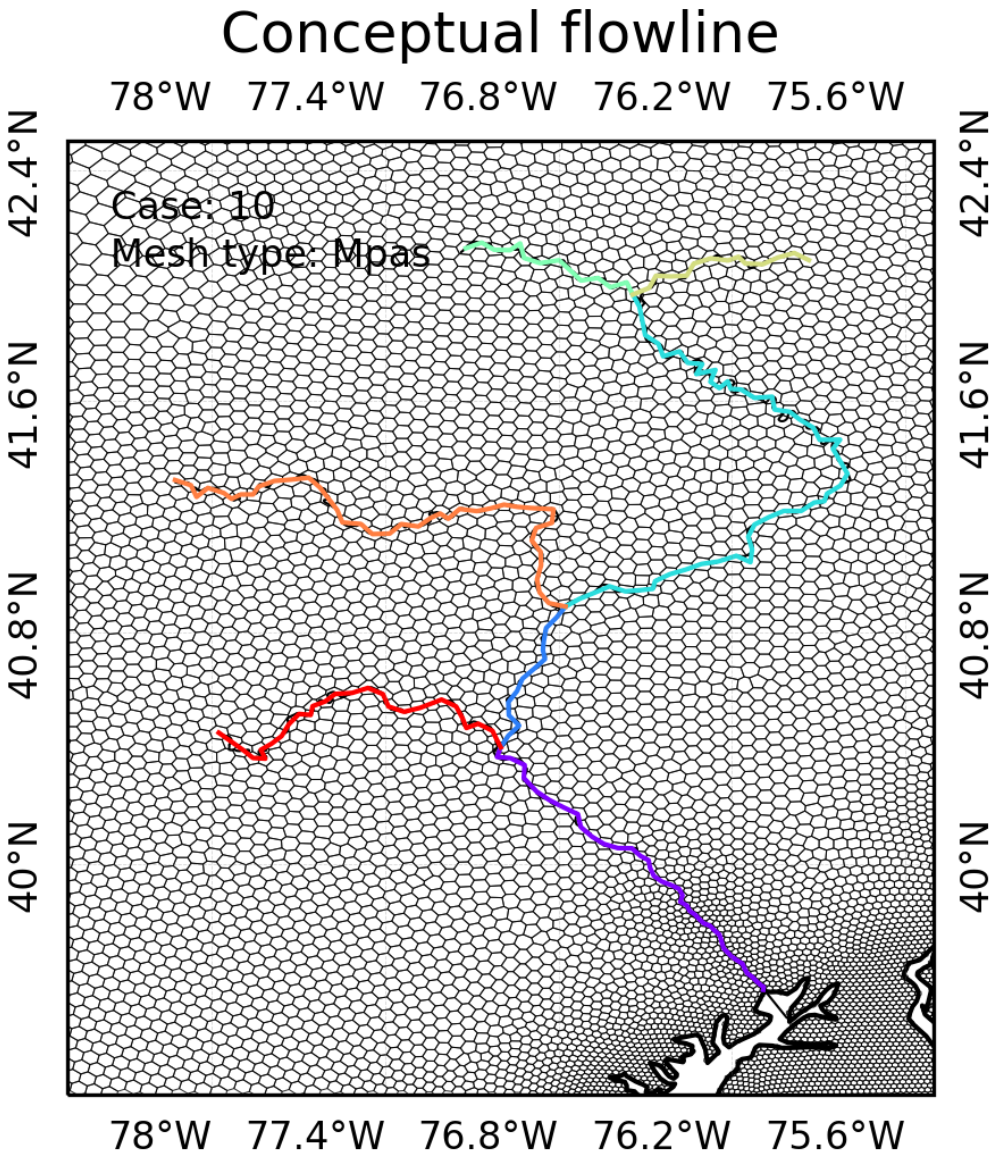

Figure 10. Modeled flowline-guided MPAS mesh from Case 10 (clipped to the study domain) (Table 1). Because the JIGSAW mesh generator considers the simplified flowlines in its density function, the mesh cell center locations align with the flowlines and cell resolutions are higher (~ 3 km) near the flowline than away from it (~ 10 km).

3.4.3 Topological relationship reconstruction

In general, on the structured meshes, the spatial patterns of modeled conceptual flowlines are similar to the simplified flowlines at different spatial resolutions (Figure 9).

The model is able to separate parallel flowlines such as Segments 1 and 6 (Figure 6). It also captures all the river confluences. As the spatial resolution increases, the model can capture more spatial details. At 5 km resolution, the modeled conceptual flowlines

316 can accurately describe the river confluences. The shortcut algorithm also works well,
317 as it produces the classical D8 diagonal paths in many scenarios (e.g., Cases 2 and 5).

318 On the unstructured MPAS mesh, the modeled conceptual flowlines closely follow
319 the provided simplified flowlines (Figures 10 and 11). The model's performance varies
320 with the complexity of the simplified flowline and mesh cell resolution. For example, at
321 moderate resolution (~ 6 km) the modeled conceptual flowline follows the overall pat-
322 tern and captures the sharp U-turn near the river meander (Figure 11). Near Conflu-
323 ence 1 and the outlet (~ 3 km), the modeled conceptual flowlines almost overlap with
324 the provided simplified flowlines (Figure 11). The modeled MPAS mesh-based concep-
325 tual flowlines are closer to the simplified flowlines than the modeled structured meshes-
326 based conceptual flowlines (Section 3.5.2). This improved fidelity is a result of the en-
327 hanced flexibility of a fully unstructured mesh representation. This allows for the close
328 alignment of mesh cells with stream features during the mesh generation phase, which
329 in turn significantly improves the accuracy of the reconstructed river flowlines.

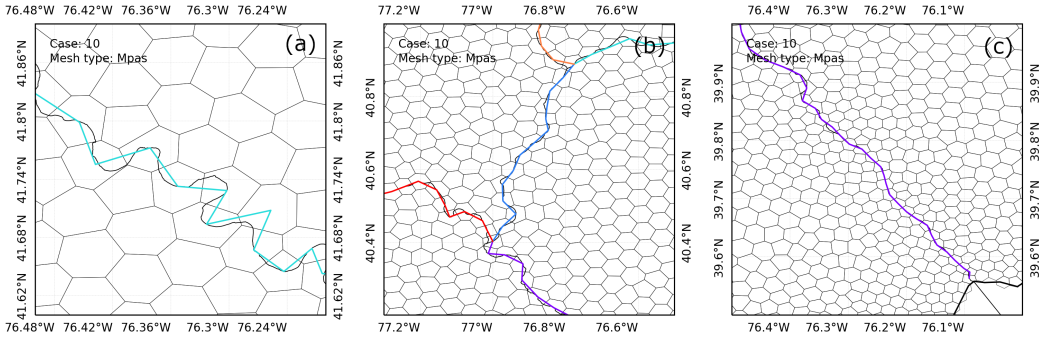

Figure 11. The zoom-in views of the spatial distributions of modeled conceptual flowlines on the MPAS mesh from Case 10 (Table 1). (a). (b), and (c), are the zoomed-in views near the river meander, river confluence, and river outlet, respectively. Black line features are the simplified flowlines. Colored line features are the conceptual flowlines.

330 With increased flowline details, i.e., stream order 6 or dam burnt in, the model is
331 still able to capture the river networks (Figure S11).

332 3.5 Metrics analysis

333 For metrics analysis, we only focus on Cases 1 to 10 as they have the same level
334 of flowline details.

335 *3.5.1 Flowline length*

336 The modeled cell center-based conceptual flowline length, although not always used
337 in hydrologic or Earth system models(Paz & Collischonn, 2007), reflects the closeness
338 between the modeled and real river networks. For structured meshes, the total flowline
339 length increases as the spatial resolution increases (Figure 8). At the same spatial res-
340 olution, the differences between different meshes are around 10.0%. The total length from
341 the MPAS mesh-based flowlines is shorter than the results from equivalent structured
342 meshes (Figure 8). Differences between different segments are mainly influenced by river
343 meander features (Section 3.5.4).

Because the MPAS mesh is designed to follow flowlines whenever possible, its length and area of difference (Section 3.5.2) are the smallest. In contrast, the structured mesh-based flowlines usually generate a "zig-zag" effect, which produces a larger length and area of difference (Figure 12).

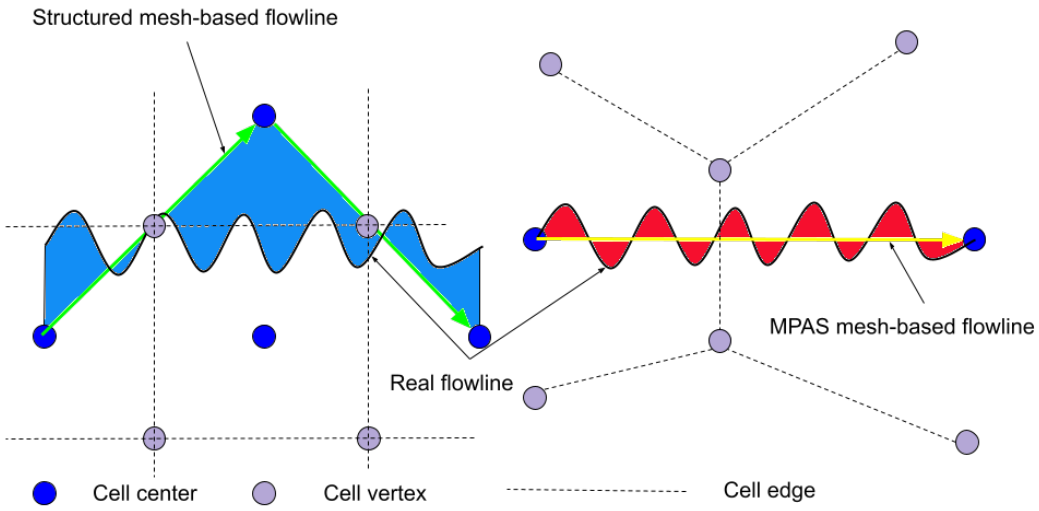

Figure 12. Illustration of the conceptual length and area of difference. The dashed black lines are the cell edges. The blue and purple dots are cell centers and vertices, respectively. The black/green/yellow lines are the real, structured mesh-based, and MPAS mesh-based flowlines. Because the structured mesh-based flowlines (green lines) do not closely follow the real flowlines (black lines), their total length is larger than the MPAS mesh-based flowline (yellow line). The structured mesh-based area of difference (light blue polygons) is larger than the MPAS mesh-based area of difference (red polygons).

3.5.2 Area of difference

Similar to our earlier study(Liao et al., 2022), we compared the area of difference formed by flowline intersections. In this method, we used area to represent line feature discrepancies. If two or more line features intersect, the intersected segments can be used to create enclosed polygons. In general, the smaller the total area, the closer the line features. The area of difference is illustrated in Figure 12. Area of differences can be calculated using Python scripting in the following steps:

1. Convert modeled conceptual flowlines into edge-based flowlines A;
2. Intersect the edge-based flowlines A with the simplified flowlines B to obtain all the vertices list C;
3. Classify the vertices C into different types of vertices (Text S2);
4. Split both A and B using C to obtain a list of flowlines D;
5. Build all the polygons enclosed by connected flowlines in D using a cycling algorithm and calculate their areas;
6. Sum up the areas to obtain the total area of difference.

For structured meshes, the area of difference decreases as the spatial resolution increases. At the same spatial resolution, the hexagon mesh-based area of difference is the

smallest except at the 50 km spatial resolution. For the unstructured mesh, the MPAS mesh-based area of difference is smaller than its equivalent structured mesh results (Figures 13 and 14).

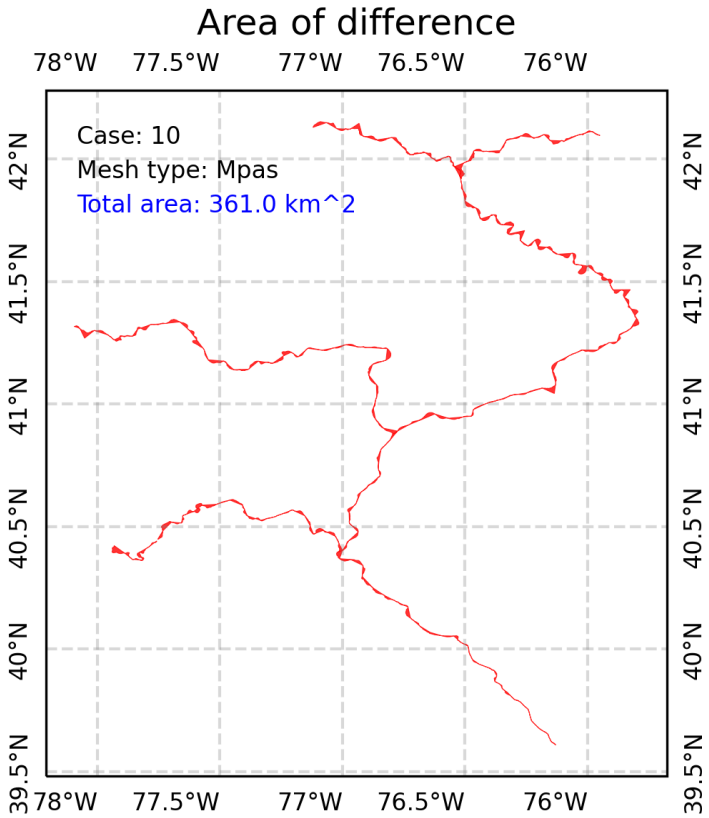

Figure 13. The enclosed area of difference between modeled river networks and simplified NHDPlus flowlines from the MPAS mesh.

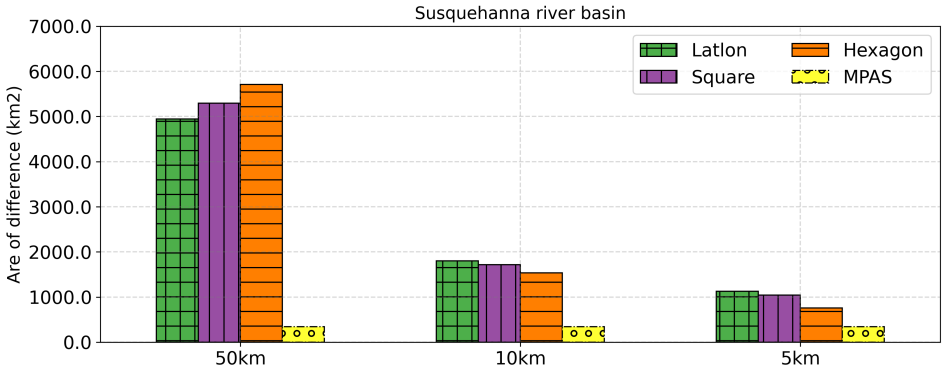

Figure 14. Comparison of the area-of-difference for modeled conceptual flowlines from different model configurations. The X-axis is the spatial resolution and the Y-axis is the area of difference (km^2).

3.5.3 Branching angle

The ramification angle, or branching angle (Text S1), is another important characteristic of river networks(Devauchelle et al., 2012). This metric is calculated using the last incoming flowline edges of a confluence. The results show that conceptual flowlines generally cannot capture the branching angles well, especially at the 50 km spatial resolution. For example, the branching angles at Confluence 2 are all 180 degrees from the structured meshes (Figures 9 and 15). This is because only a limited number of branching angles are supported by the structured meshes. In contrast, the branching angles from the MPAS mesh are more flexible because the mesh provides flexible rotations.

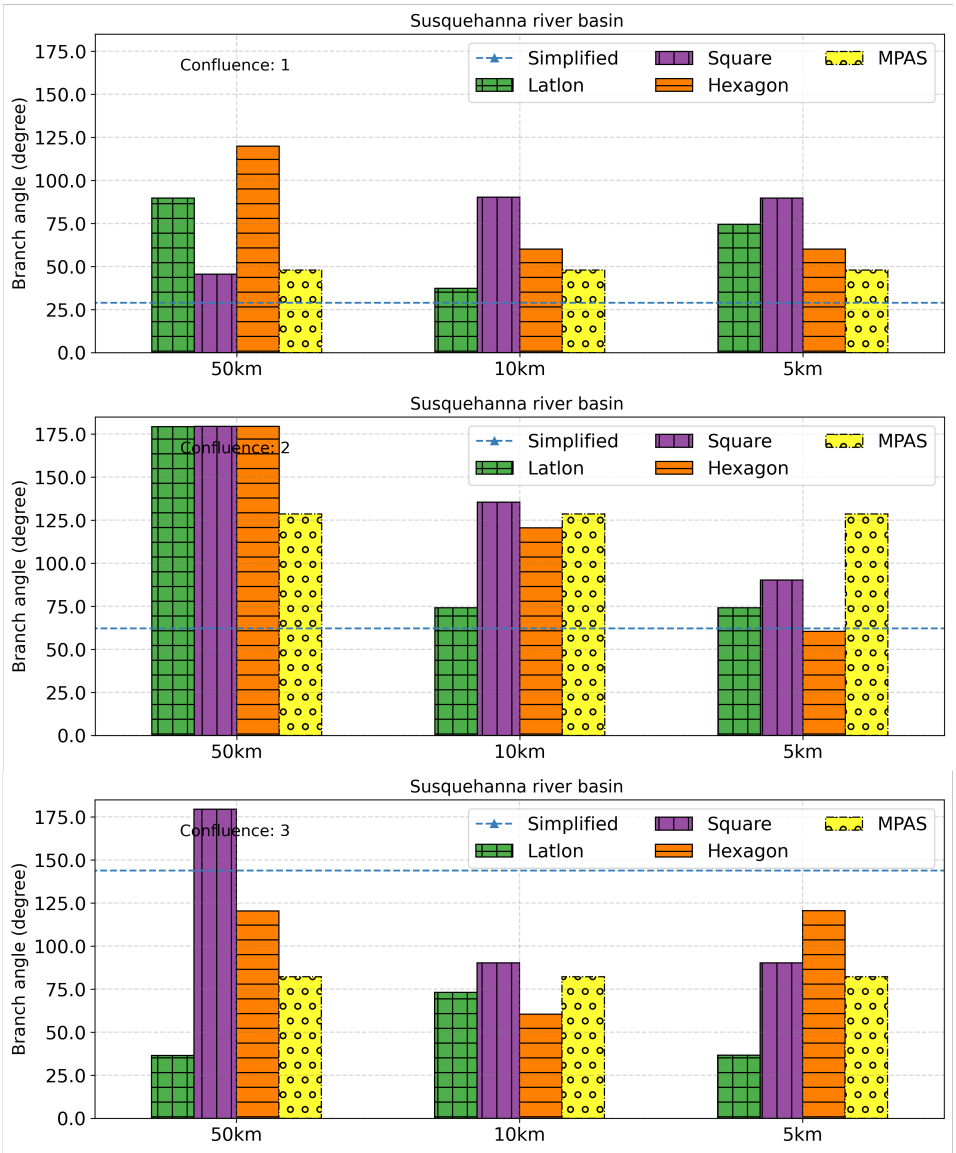

Figure 15. Comparison of the river branch angle from different model configurations for Confluence 1 to 3. The X-axis is the spatial resolution and the Y-axis is the branch angle (degree). The dashed blue line represents results from the simplified flowlines, used as the reference.

3.5.4 River sinuosity

River sinuosity is the ratio between flowline length and valley length (the distance between the starting and ending vertices of a flowline). For example, all the models underestimate river sinuosity for Segment 4, which exhibits many meander features (Figure 16). In contrast, the modeled river sinuosity is much closer to the reference for Segment 7, especially at the 5 km high spatial resolution in part because there are fewer meander features (Figure 16).

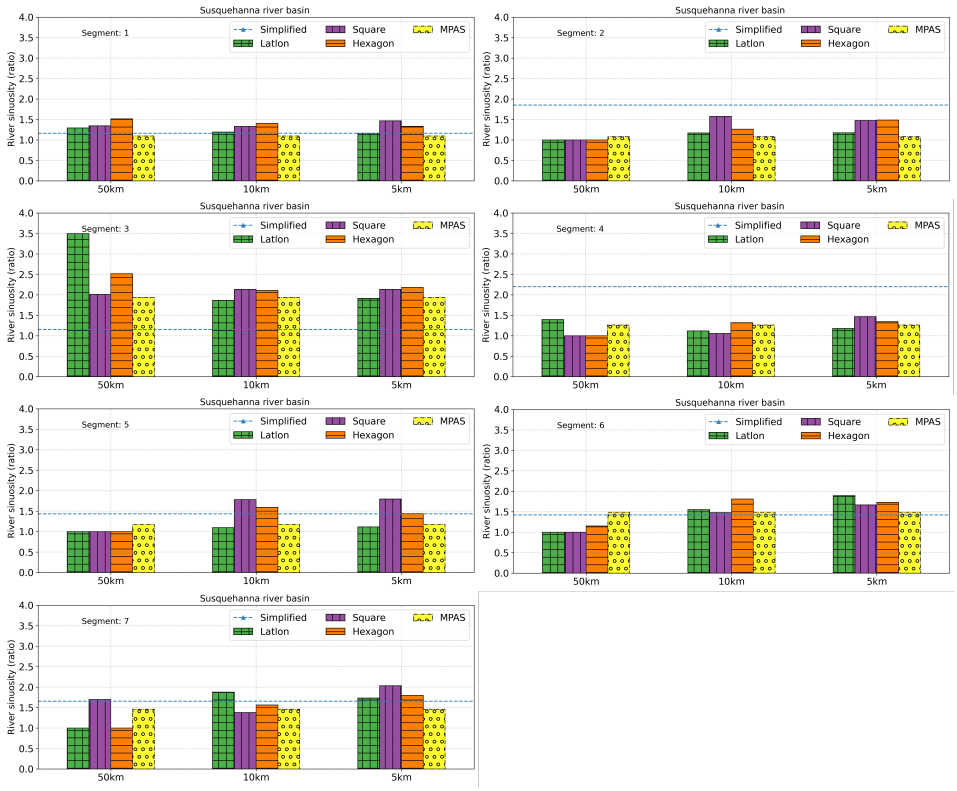

Figure 16. Comparison of river sinuosity for Segment 1 to 7 from different model configurations. The X-axis is the spatial resolution and the Y-axis is the sinuosity (ratio). The dashed blue line is the simplified flowline, used as the reference. Colored (green, purple, orange, and yellow) bars with textures represent different model configurations.

4 Discussion

4.1 Thresholds

4.1.1 Stream order

The stream order used to filter out small rivers from the original NHDPlus HR datasets determines the level of details the model can preserve (Figure 5). This threshold should be tested based on the mesh spatial resolution and hydrologic model application. For example, if low-order rivers are included during the data preparation step, they may be still removed by the small river removal algorithm.

Because stream order is based on the topological relationship between stream segments, it may not reflect the actual drainage area of a river channel. For this reason, the drainage area should also be considered to filter out small rivers.

4.1.2 Small river removal

The "small river" threshold (Table S2) is an important parameter for removing river channels with relatively short lengths. It should be configured considering the mesh spatial resolution. User-provided flowlines may be too fragmented, causing the model to excessively remove flowlines. For this reason, a visual inspection of the raw flowlines is important.

4.1.3 Shortcut path length

Although the "shortcut path" algorithm allows the model to reproduce the classical D8 diagonal travel path in the rectangle meshes, its impact on the topological relationship reconstruction process is not trivial. The structured rectangle mesh particularly exemplifies the effects of the "shortcut path" algorithm because it produces a diagonal travel path that goes through the mesh cell vertex. In contrast, the shortcut remains on the mesh cell edge instead of vertex for the structured hexagon and unstructured MPAS meshes. Thus, the structured hexagon and unstructured MPAS meshes are preferred for coupled surface and subsurface hydrologic simulations because they are more compatible with hydrologic models that assume the flux exchange occur through the cell faces(Liao et al., 2022).

Similar to the small river removal algorithm, this parameter should consider the mesh spatial resolution. Therefore, a fraction between 0.0 and 1.0 of the mesh cell resolution is preferred.

In our model, the flowline length is used to determine whether the conceptual flowline should take a shortcut. Other studies also suggest that the distance from the flowline to the mesh cell center should also be considered(Lindsay, 2016b). A combined approach considering both length and distance may produce even more robust results.

4.2 Quality of modeled conceptual flowlines

In general, the quality of modeled conceptual flowlines increases as the spatial resolution increases. For structured meshes, the model performances are very close and the hexagon mesh is potentially better (Figure 14), which is constant with our earlier study(Liao et al., 2020). The MPAS mesh-based conceptual flowlines are the closest to the real flowlines.

Although the center-to-center flowline length often leads to reduced river length, the length decrease is offset by the "zig-zag" effect (Figure 12). This explains why the total length can be higher than the simplified flowlines from the structured meshes-based results (Figure 8). Because the MPAS mesh cells align with the simplified flowlines, they have a minimized offset effect and the shortest total length (Figure 16).

The area of difference comparison shows that the MPAS mesh produces the closest conceptual flowlines to the simplified flowlines in terms of area followed by the hexagon mesh. Its performance is even better than structured meshes at higher resolution 5 km (Figure 14).

Branching angle analysis shows that the structured meshes generally cannot capture the confluence ramification feature well because they only support a limited number of angles. The MPAS mesh has the potential to resolve this issue, but requires a high-quality mesh (Figure 15).

4.3 Importance of topological relationships

Compared with the vector-based or raster DEM-based river networks extraction methods, our topological relationship-based method overcomes several existing limitations. It precisely follows river channels near river meander and confluence, producing high-quality conceptual river networks.

Given the explicit topological relationships, advanced stream-burning algorithms can be implemented to remove the depressions within river channels. Traditionally, stream burning algorithms blindly lower river channel elevations significantly to force the flow direction. More recently, adaptive stream-burning algorithms can adjust river channel elevations using a hybrid breaching and filling approach(Lindsay, 2016b). However, the

breaching algorithm requires topological relationships, especially near river confluences, to avoid incorrect breaching directions(Liao et al., 2022). Therefore, the topological relationships produced from our method provide an opportunity to use the fully hybrid breaching and filling method to minimize modifications to both land and river elevations.

4.4 Implication for hydrologic and Earth system models

Representing river networks is key to hydrologic models. Our study achieves this through two major considerations. First, the flowline simplification guarantees that the most important river information is preserved while keeping river flowlines in their simplest formats. Second, the topological relationship-based conceptual flowlines further preserve the river network structures while linking them with the mesh system. In this way, the model is able to consistently capture the spatial pattern of river networks.

The support for non-rectangle meshes further improves model performance and demonstrates the advantage of using unstructured meshes in hydrologic and Earth system models. Instead of projecting existing hydrologic features onto the rectangle meshes, we can now directly generate meshes that explicitly define these features. For example, in this study, we burnt flowlines in the MPAS mesh through the JIGSAW density function and in the conceptual flowlines through the topological relationship reconstruction. Our method also burnt dams through the workflow (Figure S11). In the future, we can include other hydrologic features such as lakes and wetlands. This enables us to much more precisely define the geometry of hydrologic features, resulting in improved hydrologic and Earth system models(Liao et al., 2022).

Our method produces high-quality river routing parameters other than the spatial pattern. For example, it keeps track of both actual and conceptual flowline length throughout the processes. However, when the "shortcut" occurs, the corresponding flowline length can be merged into the upstream/downstream, or ignored. More importantly, the topological relationships can be used to minimize the modifications to both land and river elevations in the fully adaptive hybrid breaching and filling method, generating more realistic slopes in both rivers and riparian zones. Because both river length and slope are important river routing parameters, the results from our method can be used to improve flow routing models.

4.5 Limitations

There are several limitations to the current model. First, unlike the small river removal and shortcut path length parameters that can be configured to consider the mesh resolution, the stream order threshold used to filter out small rivers is not directly associated with any physical attribute. Alternatively, drainage area should be a more meaningful metric to filter out small rivers.

Second, the model does not use additional information in the cycling detection algorithm to decide which channel should be preserved if they have the same stream order. River channel width or length may be included to capture the dominant river channel.

Third, our current method only considers rivers in the workflow. Other hydrologic features (e.g., lake and reservoir) are not fully considered. In a complex landscape, these features should also be included so the final flow routing map is consistent for hydrologic and Earth system models.

Last, our model currently only relies on the topological relationship to reconstruct conceptual river networks. It is possible to combine our method with other upscaling methods to consider more complex scenarios. In this case, a more dedicated mesh cell topol-

ogy algorithm is needed because unstructured meshes often do not have the parent-child hierarchical structure.

5 Conclusions

In this study, we developed a mesh-independent topological relationships-based river networks representation model (PyFlowline). We applied the model to the Susquehanna river basin with different configurations. The model evaluation shows that the model performs well and the modeled conceptual river networks are consistent with the real-world river networks. The outputs of our model, especially the topological relationships, should be used for advanced terrain analysis and hydrologic models.

Acknowledgments
This work was supported by the Earth System Model Development program area of the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research as part of the multi-program, collaborative Integrated Coastal Modeling (ICoM) project. The data used for model simulations can be downloaded through the USGS website (`https://www.usgs.gov/national-hydrography`). The PyFlowline model can be installed as a Python package (`https://github.com/changliao1025/pyflowline`). The data and code used in this paper is available from `https://github.com/DOE-ICoM/liao-etal_2022_pyflowline_james`. 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.

Conflict of Interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

Coon, E., Svyatsky, D., Jan, A., Kikinzon, E., Berndt, M., Atchley, A., ... Molins, S. (2019, 9). *Advanced Terrestrial Simulator.* [Computer Software] \url{https://doi.org/10.11578/dc.20190911.1}. Retrieved from `https://doi.org/10.11578/dc.20190911.1` doi: 10.11578/dc.20190911.1

Davies, H. N., & Bell, V. A. (2009). Assessment of methods for extracting low-resolution river networks from high-resolution digital data. *Hydrological Sciences Journal*, *54*(1), 17–28. doi: 10.1623/hysj.54.1.17

de Azeredo Freitas, H. R., da Costa Freitas, C., Rosim, S., & de Freitas Oliveira, J. R. (2016). Drainage networks and watersheds delineation derived from TIN-based digital elevation models. *Computers & Geosciences*, *92*, 21–37.

Decharme, B., Alkama, R., Papa, F., Faroux, S., Douville, H., & Prigent, C. (2012). Global off-line evaluation of the ISBA-TRIP flood model. *Climate Dynamics*, *38*(7), 1389–1412.

Devauchelle, O., Petroff, A. P., Seybold, H. F., & Rothman, D. H. (2012). Ramification of stream networks. *Proceedings of the National Academy of Sciences*, *109*(51), 20832–20836.

Eilander, D., Van Verseveld, W., Yamazaki, D., Weerts, A., Winsemius, H. C., & Ward, P. J. (2021). A hydrography upscaling method for scale-invariant

parametrization of distributed hydrological models. *Hydrology and Earth System Sciences*, *25*(9), 5287–5313.

Engwirda, D. (2014). Locally optimal Delaunay-refinement and optimisation-based mesh generation.

Engwirda, D. (2017). JIGSAW-GEO (1.0): locally orthogonal staggered unstructured grid generation for general circulation modelling on the sphere. *Geoscientific Model Development*, *10*(6), 2117–2140.

Engwirda, D., & Liao, C. (2021, 10). 'Unified' Laguerre-Power Meshes for Coupled Earth System Modelling. *Zenodo*. Retrieved from `https://zenodo.org/record/5558988` doi: 10.5281/ZENODO.5558988

Esri Water Resources Team. (2011). *Arc Hydro Tools - Tutorial* (Tech. Rep.).

Fekete, B. M., Vorosmarty, C. J., & Lammers, R. B. (2001). Scaling gridded river networks for macroscale hydrology: Development, analysis, and control of error. *Water Resources Research*, *37*(7), 1955–1967.

Feng, D., Tan, Z., Engwirda, D., Liao, C., Xu, D., Bisht, G., ... Leung, 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 Discussions*, *2022*, 1–31. Retrieved from `https://hess.copernicus.org/preprints/hess-2022-251/` doi: 10.5194/hess-2022-251

GDAL/OGR contributors. (2019). *Geospatial Data Abstraction software Library.* Retrieved from `https://gdal.org`

Gillies, S., & others. (2007). *Shapely: manipulation and analysis of geometric objects.* Retrieved from `https://github.com/Toblerity/Shapely`

Goulden, T., Hopkinson, C., Jamieson, R., & Sterling, S. (2014). Sensitivity of watershed attributes to spatial resolution and interpolation method of LiDAR DEMs in three distinct landscapes. *Water Resources Research*, *50*(3), 1908–1927. doi: 10.1002/2013WR013846

Hellweger, F., & Maidment, D. (1997). AGREE-DEM surface reconditioning system. *Online at http://www. ce. utexas. edu/prof/maidment/gishydro/ferdi/research/agree/agree. html (last accessed March 3, 2005)*.

Hsu, S.-C. (2020). Hex-Urban: Investigating the adoption of hexagonal grids to represent the total urban water cycle within a distributed water balance model.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, *9*(3), 90–95. doi: 10.1109/MCSE.2007.55

Hyvaluoma, J. (2017). Reducing the grid orientation dependence of flow routing on square-grid digital elevation models. *International Journal of Geographical Information Science*, *31*(11), 2272–2285.

Ivanov, V. Y., Vivoni, E. R., Bras, R. L., & Entekhabi, D. (2004). Catchment hydrologic response with a fully distributed triangulated irregular network model. *Water Resources Research*, *40*(11). doi: 10.1029/2004WR003218

Jolley, T. J., & Wheater, H. S. (1997). The introduction of runoff routing into large-scale hydrological models. *Hydrological processes*, *11*(15), 1917–1926.

Kreveld, M. v., & Silveira, R. I. (2011). Embedding rivers in triangulated irregular networks with linear programming. *International Journal of Geographical Information Science*, *25*(4), 615–631.

Liao, C. (2022, 3). *A lightweight Python package for Earth science.* Zenodo. Retrieved from `https://zenodo.org/record/6368652` doi: 10.5281/ZENODO.6368652

Liao, C., & Cooper, M. (2022, 4). *Pyflowline a mesh-independent river network generator for hydrologic models.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.6407299` doi: 10.5281/zenodo.6407299

Liao, C., Tesfa, T., Duan, Z., & Leung, L. R. (2020). Watershed delineation on a hexagonal mesh grid. *Environmental Modelling & Software*, *128*, 104702.

Retrieved from `http://www.sciencedirect.com/science/article/pii/S1364815219308278` doi: 10.1016/j.envsoft.2020.104702

Liao, C., Zhou, T., Xu, D., Barnes, R., Bisht, G., Li, H.-Y., ... Engwirda, D. (2022). Advances in hexagon mesh-based flow direction modeling. *Advances in Water Resources*, 104099. doi: 10.1016/j.advwatres.2021.104099

Liao, C., Zhuang, Q., Leung, L. R., & Guo, L. (2019). Quantifying Dissolved Organic Carbon Dynamics Using a Three-Dimensional Terrestrial Ecosystem Model at High Spatial-Temporal Resolutions. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4489–4512. doi: 10.1029/2019MS001792

Lin, P., Yang, Z.-L., Gochis, D. J., Yu, W., Maidment, D. R., Somos-Valenzuela, M. A., & David, C. H. (2018). Implementation of a vector-based river network routing scheme in the community WRF-Hydro modeling framework for flood discharge simulation. *Environmental Modelling & Software*, *107*, 1–11.

Lindsay, J. B. (2016a). Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models. *Hydrological Processes*, *30*(6), 846–857. doi: 10.1002/hyp.10648

Lindsay, J. B. (2016b). The practice of DEM stream burning revisited. *Earth Surface Processes and Landforms*, *41*(5), 658–668. doi: 10.1002/esp.3888

Lindsay, J. B., Yang, W., & Hornby, D. D. (2019). Drainage network analysis and structuring of topologically noisy vector stream data. *ISPRS International Journal of Geo-Information*, *8*(9), 422.

Luo, X., Hong-Yi, L., Leung, L. R., Tesfa, T. K., Getirana, A., Papa, F., & Hess, L. L. (2017). Modeling surface water dynamics in the Amazon Basin using MOSART-Inundation v1. 0: impacts of geomorphological parameters and river flow representation. *Geoscientific Model Development*, *10*(3), 1233.

Mao, Y., Zhou, T., Leung, L. R., Tesfa, T. K., Li, H., Wang, K., ... Getirana, A. (2019). Flood inundation generation mechanisms and their changes in 1953–2004 in global major river basins. *Journal of Geophysical Research: Atmospheres*, *124*(22), 11672–11692.

Minn, M. (2015). MMQGIS for QGIS Geographic Information System. *Open Source Geospatial Foundation Project.[(accessed on 17 May 2018)]*.

Mizukami, N., Clark, M. P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., ... Woods, R. (2016). mizuRoute version 1: a river network routing tool for a continental domain water resources applications. *Geoscientific Model Development*, *9*(6), 2223–2238.

Munier, S., & Decharme, B. (2022). River network and hydro-geomorphological parameters at 1 / 12 resolution for global hydrological and climate studies. *Earth System Science Data*, *14*(5), 2239–2258.

Paz, A. R., & Collischonn, W. (2007). River reach length and slope estimates for large-scale hydrological models based on a relatively high-resolution digital elevation model. *Journal of Hydrology*, *343*(3-4), 127–139.

Reed, S. M. (2003). Deriving flow directions for coarse-resolution (1–4 km) gridded hydrologic modeling. *Water resources research*, *39*(9).

Ringler, T., Petersen, M., Higdon, R. L., Jacobsen, D., Jones, P. W., & Maltrud, M. (2013). A multi-resolution approach to global ocean modeling. *Ocean Modelling*, *69*, 211–232. doi: 10.1016/j.ocemod.2013.04.010

Sahr, K. (2015). *DGGRID version 6.2 b: User documentation for discrete global grid software.*

Shaw, D. A., Martz, L. W., & Pietroniro, A. (2005). A methodology for preserving channel flow networks and connectivity patterns in large-scale distributed hydrological models. *Hydrological Processes: An International Journal*, *19*(1), 149–168.

Shelef, E., & Hilley, G. E. (2013). Impact of flow routing on catchment area calculations, slope estimates, and numerical simulations of landscape development. *Journal of Geophysical Research: Earth Surface*, *118*(4), 2105–2123.

Sood, A., & Smakhtin, V. (2015). Global hydrological models: a review. *Hydrological Sciences Journal*, *60*(4), 549–565. doi: 10.1080/02626667.2014.950580

Tarboton, D. G. (2003). Terrain analysis using digital elevation models in hydrology. In *23rd esri international users conference, san diego, california* (Vol. 14). Citeseer.

Tarboton, D. G., Bras, R. L., & Rodriguez-Iturbe, I. (1988). The fractal nature of river networks. *Water Resources Research*, *24*(8), 1317–1322.

Tarboton, D. G., Bras, R. L., & Rodriguez-Iturbe, I. (1991). On the extraction of channel networks from digital elevation data. *Hydrological processes*, *5*(1), 81–100. doi: 10.1002/hyp.3360050107

USGS. (2013). *National Hydrography Geodatabase: The National Map viewer available on the World Wide Web, accessed [June 02, 2015], at url [http://viewer.nationalmap.gov/viewer/nhd.html?p=nhd].* Retrieved from `http://viewer.nationalmap.gov/viewer/nhd.html?p=nhd`

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). doi: 10.1029/2012WR012313

Wu, H., Kimball, J. S., Mantua, N., & Stanford, J. (2011). Automated upscaling of river networks for macroscale hydrological modeling. *Water Resources Research*, *47*(3). doi: 10.1029/2009WR008871

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT Hydro: A high-resolution global hydrography map based on latest topography dataset. *Water Resources Research*, *55*(6), 5053–5073. doi: 10.1029/2019WR024873

Yamazaki, D., Kanae, S., Kim, H., & Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, *47*(4).

Yamazaki, D., Oki, T., & Kanae, S. (2009). Deriving a global river network map and its sub-grid topographic characteristics from a fine-resolution flow direction map. *Hydrology and Earth System Sciences*, *13*(11), 2241–2251. doi: 10.5194/hess-13-2241-2009

Zhou, T., Leung, L. R., Leng, G., Voisin, N., Li, H., Craig, A. P., ... Mao, Y. (2020). Global irrigation characteristics and effects simulated by fully coupled land surface, river, and water management models in E3SM. *Journal of Advances in Modeling Earth Systems*, *12*(10), e2020MS002069.