A Vector-Based River Routing Model for Earth System Models: Parallelization and Global Applications

Naoki Mizukami1, Martyn P. Clark2,3, Shervan Gharari2, Erik Kluzek2, Ming Pan4, Peirong Lin5,6, Hylke E. Beck2, and Dai Yamazaki8

1National Center for Atmospheric Research, Boulder, CO, USA, 2Centre for Hydrology, University of Saskatchewan, Canmore, AB, Canada, 3Department of Geography and Planning, University of Saskatchewan, Saskatoon, SK, Canada, 4Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA, 5Princeton University, Princeton, NJ, USA, 6Institution of Remote Sensing and GIS, Peking University, Beijing, China, 7GloH2O, Almere, Netherlands, 8Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

Abstract A vector-river network explicitly uses realistic geometries of river reaches and catchments for spatial discretization in a river model. This enables improving the accuracy of the physical properties of the modeled river system, compared to a gridded river network that has been used in Earth System Models. With a finer-scale river network, resolving smaller-scale river reaches, there is a need for efficient methods to route streamflow and its constituents throughout the river network. The purpose of this study is twofold: (1) develop a new method to decompose river networks into hydrologically independent tributary domains, where routing computations can be performed in parallel; and (2) perform global river routing simulations with two global river networks, with different scales, to examine the computational efficiency and the differences in discharge simulations at various temporal scales. The new parallelization method uses a hierarchical decomposition strategy, where each decomposed tributary is further decomposed into many sub-tributary domains, enabling hybrid parallel computing. This parallelization scheme has excellent computational scaling for the global domain where it is straightforward to distribute computations across many independent river basins. However, parallel computing for a single large basin remains challenging. The global routing experiments show that the scale of the vector-river network has less impact on the discharge simulations than the runoff input that is generated by the combination of land surface model and meteorological forcing. The scale of vector-river networks needs to consider the scale of local hydrologic features such as lakes that are to be resolved in the network.

Plain Language Summary This study introduces a parallel computing method for river models that simulate discharges at many nested river reaches. The new parallelization method enables us to reduce the computation time over the global domain including numerous independent river basins, but we face difficulty in reducing the computation time for simulations over single river basins even using increased computing resources. Nevertheless, the improved computational efficiency enables us to perform a series of daily global river discharge simulations, one of which simulates at ∼3 million reaches. We found that the simulation using coarser scale river data including 300,000 reaches were more similar to the results with the fine-scale data for longer time scale analysis such as annual scale than the daily scale analysis.

1. Introduction

Terrestrial river systems play an important role in the global water cycle. On average, rivers transport ∼44,000 km³ of annual freshwater into the ocean globally (Clark et al., 2015; Dai & Trenberth, 2002), which is equivalent to ∼35% of terrestrial annual precipitation (Dai & Trenberth, 2002). Interannual variability and long-term trends of river discharge are affected by both spatio-temporal variability of climate forcing (Dai et al., 2009) and human interventions on river systems (Shin et al., 2019, 2020; Zajac et al., 2017). Along with terrestrial runoff, rivers transport constituents such as heat and nutrients into the ocean, affecting global biogeochemical cycles. Besides the use of the river models in Earth System models, river models are extensively used for various applications such as flood forecasting and water security assessments.
Large-domain river routing models have historically been grid-based (e.g., Koirala et al., 2014; Li et al., 2015; Lohmann et al., 1998; Thober et al., 2019), where water moves to one of its neighboring grid boxes. In this approach, the river network data inherits the gridded digital elevation model (DEM) from which the river network is derived (Yamazaki et al., 2013). Many global river modeling application still uses a resolution equal to land model grids (e.g., Döll et al., 2003; Sutanudjaja et al., 2018; van Beek et al., 2011), which may not provide a sufficiently accurate representation of the river system. The potential issues caused by degradation of river model resolutions are (1) misrepresentation of drainage basins (Nguyen-Quang et al., 2018), causing misallocation of terrestrial runoff and thus water balance accounting errors in a particular basin; and (2) difficulty in representing locally relevant features in river systems such as lakes and human water uses (Wada et al., 2016), as well as difficulties in locating gauges in the network.

With global high-resolution (<1 km) hydrography data sets available (Lehner et al., 2008; Yamazaki et al., 2019), it is possible to increase the spatial resolution of the grid-based river network to overcome the issues regarding such river network representation. However, the number of grid boxes or computation costs increases rapidly with increasing resolution. To improve the quality of the coarse-resolution grid-river network, there are several efforts into developing river network upscaling methods, where fine-scale hydrographic characteristics, derived from high-resolution elevation data, are quantified with sub-grid topographic features within coarser grid box (Li et al., 2013; Wu et al., 2012; Yamazaki et al., 2009). Nevertheless, building topology (connectivity) between grid box and additional hydrologic features (e.g., lakes) is not straightforward because a lake, for example, is represented as the fractional lake area within a grid box, or may exist across multiple grid boxes.

Another pathway to construct river networks is to delineate realistic geometries of catchments and associated river reaches by tracing river flowlines in high-resolution DEMs. In this approach, DEM grid boxes are grouped together to form a catchment polygon and a single flowline is assigned to each catchment. This type of river network, called “vector-river network” began to be used for large domain river routing models (Gochis et al., 2018; Lin et al., 2019; Mizukami et al., 2016; Paiva et al., 2013; Siqueira et al., 2018). The spatial “scale” of the vector-river network is defined as the size of delineated catchment and the corresponding length of the river reaches. This scale is determined by the channelization threshold used to construct the vector-river network (the channelization threshold is the minimum upstream area or flow accumulation where a river reach begins to be defined). Lateral movement of runoff over the area below this threshold is simulated by overland and subsurface flow process while the channel routing is explicitly calculated along the defined reaches.

One of the advantages of using vector-river networks is the efficiency in representing the river system compared to high resolution gridded river networks (Lehner & Grill, 2013). Because spatial discretization is physical based, a vector-based network provides much higher fidelity of real river system than a gridded-based river network if the same number of spatial elements (grid boxes or catchments) is used. Also, other hydrologic features such as lakes and irrigation lines can be provided as vector features (polygon, point, and line) and directly linked to the catchments and river reaches.

Vector-based river networks can still pose substantial computational costs for routing as the size of delineated catchments is smaller and hence the number of reaches in the river network increases. Such computational costs can become a bottleneck as we implement transport processes beyond water (e.g., heat, nitrogen, and phosphorous). Therefore, effective methods to parallelize network routing models are beneficial. The spatial decomposition of river networks is complicated because river networks have complex topological relationships between upstream and downstream reaches, that is, the hierarchical tree-like structure. Currently, there are few documented river network decomposition methods that can be applied to the river transport model, except for enhancing the efficiency of the Muskingum routing model in the RAPID model (David, Maidment, et al., 2011; David, Habets, et al., 2011).

The main goal of this modeling work is to develop a river network domain decomposition method that is applicable to parallel computations for any river transport processes, and implement it into a large domain, fine-scale vector-based river routing model. For this objective, we evaluate methods to spatially decompose a fine-scale river network into hydrologically independent subbasins where river transport simulation is performed independently per subbasin. The developed parallelization method enables performing a set of
off-line, global river routing simulations. We produce global simulations using two vector-based river networks with different spatial scales, and assess how the spatial scale of modeled river networks affects global streamflow simulations.

The study is structured as follows. Section 2 provides a brief description of the river routing model used in this study (mizuRoute: Mizukami et al., 2016), with the focus on river network topology, and explains the domain decomposition method. Section 3 describes the data sets including the vector-river network and runoff data sets, which is used for a series of global river discharge simulations to evaluate model performance (i.e., computation efficiency) and effects of model configurations (network data, runoff, and river routing schemes) on the simulations. Section 4 discusses the results. In Section 5, we also discuss future research opportunities. Finally, conclusions are provided in Section 6.

2. The mizuRoute Vector-Based Routing Model

In this study, we use the mizuRoute model (Mizukami et al., 2016). mizuRoute includes two distinct routing schemes: (1) the impulse response function (IRF) and (2) the kinematic wave tracking (KWT) algorithm. IRF routing is also known as the unit-hydrograph model and incurs small computational cost. KWT uses Lagrangian methods to track the propagation of particles with wave celerity computed using Manning’s equation through the river network. KWT is more computationally demanding than IRF. The details of both routing methods are provided in Mizukami et al. (2016). For off-line modeling, mizuRoute uses an areally weighted average to remap runoff from the spatial elements of land models (e.g., grid boxes) to catchments; mizuRoute routes remapped runoff through individual reaches from upstream to the outlet with either the IRF or KWT routing schemes.

2.1. River Network Topology

A vector-river network includes two data sets—polygons defining individual catchments and lines defining each river reach (Figure 1 illustrates these two spatial features). The data sets also include various attributes that provide physical information such as catchment area, and the length and slope of each river reach. Among the attributes, topological information (geometry connectivity), including catchment-to-reach
connectivity (e.g., c1 is linked to s1) and neighboring reach-to-reach connectivity (e.g., s1 is linked to s3), is key to vector-based river network routing.

mizuRoute first performs river network topology augmentation based on the reach-to-reach connectivity: reach ID and its downstream reach ID. River network augmentation produces more detailed topological information, including identifying immediate upstream reaches (possibly multiple upstream reaches, e.g., s3 has s1 and s2 in Figure 1), as well as identifying all the upstream reaches, all the contributory catchments for each river reach, and the reach processing order—the sequential order of reaches in which channel routing is performed. Channel routing is performed per reach in the reach processing order regardless of routing methods used.

The number of total upstream reaches is used for the spatial decomposition of the river network as explained in Section 2.2. The efficiency of the network augmentation process is critical when applying mizuRoute to large domain, fine-scale river data sets.

2.2. Proposed River Network Domain Decomposition

With the use of a single processor, routing is performed one by one, as determined by the reach processing order. With an increasing number of reach elements (e.g., a few million reaches in the fine-scale river network), processing each reach sequentially with one processor becomes impractical. Therefore, parallel computing is necessary. As opposed to the atmospheric and ocean models where fluxes interact with one of the immediate neighboring elements, making the domain decomposition relatively straightforward, river routing is performed for a tree-like river network structure, and neighboring reaches and catchments are not necessarily connected in terms of flux movement. Because of this, it is challenging to define domain decompositions that produce hydrologically connected and well-balanced workload (i.e., number of reaches) per domain.

One approach to the domain decomposition of river networks is by David et al. (2013, 2015) for the RAPID routing model. The RAPID model uses a linear Muskingum routing algorithm (i.e., a matrix-formed Muskingum equation) enabling the solution of all the locations at the same time at each time step. However, with a larger river network, the large matrix size becomes computationally expensive. Given this, they developed domain decomposition methods for solving this matrix-based Muskingum equation in parallel. Their decomposition is based on “radius of downstream influence,” which is based on the idea that upstream conditions far upstream of a reach of interest do not influence the reach of interest.

Our domain decomposition strategy for mizuRoute adopts a two-level hierarchical structure. The first level of the decomposition utilizes distributed memory (DM) parallelism where we use explicit message passing interface (MPI) programs. The second level of the decomposition is applied for each subdomain decomposed at the first level decomposition and utilizes shared memory (SM) parallelism within the DM processor via OpenMP (OMP).

2.2.1. First Level Domain Decomposition for Distributed Memory

We developed a simple and objective topological-based domain decomposition strategy that splits the whole domain into (1) tributary domains that are independent of other tributary domains, and (2) mainstem domains that are required to be processed after the tributary domains. Figure 2 illustrates the overall computational flows along with communications between DM processors.

Figure 3 illustrates the tributary domains and the mainstem domain resulting from the domain decompositions for a single-outlet river basin, the Mississippi basin. The method uses the following topological information: (1) the total number of upstream reaches for each reach \( N_{\text{up}} \), as shown in Figure 3a; and (2) the immediate upstream reach identifier for each reach (upSegID). Although we illustrate the domain decomposition method using the vector-based network, the method can be applicable to the gridded-based network where downstream and upstream grid boxes are easily identified using flow direction information. The DM domain decomposition procedure for a single river basin is computed as follows:
Determine the maximum upstream threshold for the tributary domains \( N_{\text{trib,max}} \). This is based on the total number of reaches in the entire river network \( N_{\text{rch}} \) and the number of DM processors \( N_{\text{proc}} \), with \( N_{\text{trib,max}} = N_{\text{rch}} / N_{\text{proc}} \).

2. Identify the mainstem domain. This is accomplished simply by identifying the reaches where the total number of upstream reaches, \( N_{\text{ups}} \), is greater than \( N_{\text{trib,max}} \). The mainstem is denoted in thick black line in Figures 3b and 3c.

3. Identify the tributary domains. Since all the reaches other than mainstem domain belong to the tributary domains, the outlets of each tributary domain that flow into the mainstem are quickly identified using the immediate upstream reach identifier (upSegID) of the mainstem reach. Such “subbasins” defined at the outlets are hydrologically independent.

4. Allocate the tributary and mainstem domains to all the available DM processors (a primary processor and secondary DM processors). The mainstem domain is assigned to the primary processor. The subbasins in the tributary domain are allocated such that all the DM processors have an almost equal number of reaches to be processed. In Figure 3b, for example, the subbasins with the same color are allocated to the same processors (e.g., two cyan subbasins are in the same processors in domain 4 cases). The subbasins are sorted in size (i.e., number of reaches) from the largest to the smallest, and then the sorted subbasins are assigned in that order to the DM processor that has the least reach number assigned so far. Reach flow information at these tributary outlets need to be transferred from the secondary processors to the primary processor where routing computation for mainstem reaches are performed. To reduce the communication cost between primary DM processor and secondary DM processors, numerous small subbasins (see light gray reaches in Figure 3c) are allocated in the primary processor. Such small subbasin reaches are processed concurrently with the other “larger” subbasins computations that occur at the secondary processors, because the primary processor is available for tributary routing while mainstem routing is waiting for the tributary routing completion at the secondary processors (See Figure 2).

5. A halo reach needs to be identified at each junction between tributary and mainstem, that is, one upstream tributary reach for the mainstem domain and one downstream mainstem reach in the tributary domain. This halo reach holds the flow information in downstream and upstream reaches communicated across the two domains. This is necessary for routing at the mainstem reach just downstream of the tributaries as well as at the most downstream of the tributary domains when the model accounts for the backwater effect from the downstream.

For multioutlet river systems, for example, continental, or global domains, or drainage areas that include endorheic basins, the DM domain decomposition process requires one additional step:
Evaluate \( N_{\text{ups}} \) at the outlet reach of all the river basins and assign the river basins to one tributary domain if \( N_{\text{ups}} < N_{\text{trib\_max}} \).

For “large” river basin with \( N_{\text{ups}} \) at the outlet reach greater than \( N_{\text{trib\_max}} \), follow step 1 through step 3 of a single river basin decomposition described previously.

Domain allocation to DM processors also follows step 4 described for a single river basin decomposition.

### 2.2.2. Second Level Decomposition for Shared Memory

With increasing DM processors or number of domains, \( N_{\text{trib\_max}} \) gets smaller, resulting in increasing mainstem size (shown in Figure 3b). To suppress the increasing computational cost of mainstem routing, the mainstem domain is further decomposed into subdomains, here using Strahler Stream order as illustrated in Figure 4. The subdomain called “branch” within each stream order of mainstem (e.g., five independent, first order branches in the top panel in Figure 4) can be parallelized with shared memory, OpenMP parallel scheme. Individual branches in the stream order have different number of reaches, and the branch with the largest number of reaches can be bottleneck.

Likewise, each tributary domain, resulting from DM domain decomposition, can be further divided into smaller domains (sub-tributary domains) in the same approach as DM domain decomposition. Visualization of this hierarchical decomposition is given in Figure 5. These sub-tributaries composed of tributaries
This domain decomposition process is performed in the initialization phase during the model run. Note that reach routing computation within each tributary domain (tributary route) and mainstem domain (mainstem route) are performed in the reach processing order within the domain.

3. Data Sets and Model Simulations

3.1. River Network Data Sets

We consider two vector-based river network data sets that are available for the global domain and potentially applicable for Earth System modeling (see Table 1). The first data set, MERIT-Hydro, is the river network vectorized from the 3-arc sec (∼90 m) resolution raster hydrography data set called Multi-Error-Removed-Improved-Terrain (MERIT) Hydro (Yamazaki et al., 2019), consisting of nearly 3 million river reaches and catchments over the globe with a channelization threshold of 25 km² (Lin et al., 2019). Each raster cell is defined as a river reach if its upstream contributing area exceeds this threshold. The second data set is the Hydrologic Derivatives for Modeling and Applications (HDMAs: Verdin, 2017), derived from combination of three DEMs—HydroSHEDS, GMTED2010, and SRTM. There is 4% difference in the endorheic

Figure 4. Illustration of mainstem domain decompositions for 16 DM domains (top panel) and 48 distributed memory (DM) domains (bottom panels) for Mississippi basin.
basins between the two networks. This originates from different fine-resolution DEMs used for the catchment delineation. HDMA consists of 295,335 river reaches and catchments over the globe with a channelization threshold of 25 km$^2$. The topological information in HDMA is encoded using Pfafstetter codes (Verdin & Verdin, 1999), simplifying the organization and subsetting of the data set. The spatial scale of catchment of the MERIT-Hydro network is $\sim$10 times smaller than that of the HDMA network, which is equivalent to the difference in the number of catchments between the two networks, though the reach length of MERIT hydro is overall 2–3 times shorter than the HDMA network.

### 3.2. Simulation Design

A series of global routing simulations are performed (Table 2). The model configurations here include the two river networks (HDMA vs. MERIT-Hydro), and two runoff flux data sets produced by the different land models (CLM vs. VIC). Specifically, the routing simulations are performed at a daily time step for the two runoff data sets: (1) runoff from 0.125° Variable Infiltration Capacity model (VIC; Liang et al., 1994) forced with Multi-Source Weighted-Ensemble Precipitation (MSWEP; Beck et al. 2019) data for the time period

#### Table 1

| Properties                        | HDMA            | MERIT-Hydro     |
|-----------------------------------|-----------------|-----------------|
| $^1$Total land area [km$^2$]      | $132.5 \times 10^6$ | $134.7 \times 10^6$ |
| Endorheic area [km$^2$]           | $31.7 \times 10^5$  | $30.5 \times 10^5$  |
| Number of catchments              | 295,335         | 2,996,635       |
| Number of reaches                 | 278,758         | 2,996,635       |
| $^2$Median reach length [km] (range) | 19.2 (8.2–36.8) | 7.0 (3.4–12.8) |
| $^3$Median catchment area [km$^2$] (range) | 362.6 (222.5–588.7) | 36.6 (22.5–59.4) |

$^1$Antarctica and Greenland are excluded. $^2$Parentheses indicate inter-quartile range. Abbreviations: HDMA, Hydrologic Derivatives for Modeling and Applications; MERIT, Multi-Error-Removed-Improved-Terrain.
1980–2013; and (2) runoff from 0.5° Community Land Model version 5.0 (CLM5; Lawrence et al., 2019) forced with the Global Soil Wetness Project forcing data set (GSWP3-v1) for the time period 1989–2013. Two routing schemes, IRF and KWT are used. The goal of this analysis is to evaluate the effects of network data, runoff inputs, and routing schemes, rather than to evaluate the accuracy of the simulations.

4. Results and Discussion

4.1. Computational Scaling

The computational scaling of the parallel routing is evaluated based on 2-year simulations with the MERIT-Hydro network using different hybrid parallelization configurations. We performed the analysis for each routing scheme over global domain as well as Mississippi River basin domain. We ran 16 parallel computing configurations: a permutation of 8 MPI configurations (1, 2, 4, 8, 16, 32, 48, and 64 processors) and 2 OpenMP configuration (No-OpenMP and 16 threads).

Figure 6 shows the scaling of total computational time spent for routing processes, which is the sum of tributary routing, mainstem routing, and communication of reach fluxes between tributary and mainstem reaches, for the global domain. Overall, the total routing computation scales well regardless of routing methods and whether or not OpenMP is used. Use of the OpenMP helps to substantially reduce the routing time. For the global domain that includes numerous independent river basins, all the decomposed domains are tributary domains if a smaller number of DM processors are used. For the global MERIT-Hydro network, the mainstem domain emerges when using 32 DM processors or more. The mainstem routing and communication cost start affecting total routing time with increasing processors due to the increase in the number of mainstem reaches, evident from flattening at 48 processors in KWT scaling for 16-threads OpenMP (Figure 6).

For stand-alone simulations, the model has to input the runoff from the disk and distribute runoff over decomposed domains to each processor. This communication of runoff input (runoff scattering from primary to secondary processors) as well as runoff reading time becomes the bottleneck as more processors are used. This bottleneck will be less pronounced when the routing model is coupled with the Earth System Model.

---

Table 2
Simulation Configurations

| Simulation cases  | Runoff data       | Network data       |
|------------------|-------------------|--------------------|
| VIC-HDMA         | MSWEP VIC-0.125   | HDMA               |
| VIC-MERIT        | MSWEP VIC-0.125   | MERIT-Hydro        |
| CLM-HDMA         | GSWP3-v1 CLM-5.0  | HDMA               |
| CLM-MERIT        | GSWP3-v1 CLM-5.0  | MERIT-Hydro        |

Note. IRF and KWT are performed for each simulation case.

Abbreviations: CLM, Community Land Model; HDMA, Hydrologic Derivatives for Modeling and Applications; KWT, kinematic wave tracking; IRF, impulse response function; MERIT, Multi-Error-Removed-Improved-Terrain.VIC, Variable Infiltration Capacity.

---

Figure 6. Scaling of hybrid parallel routing over the global domain using the Multi-Error-Removed-Improved-Terrain (MERIT)-Hydro network. The x-axis defines the number of distributed memory (DM) processors and the y-axis quantifies the average seconds per time step. Each panel illustrates the scaling for each of the routing schemes. The total-routing time denoted by color lines is the sum of tributary and mainstem routings and communication of reach fluxes between tributaries and mainstems, and the black dash indicates perfect scaling given the elapsed time at no-MPI and no OpenMP.
The computational scaling of the HDMA network is similar to the MERIT-Hydro network (not shown). The computing time for HDMA is 8%–10% of that for the MERIT-Hydro for both routing methods. This reduction of the computation time corresponds to the difference in the number of reaches between the two river networks (Table 1).

Figure 7 shows the scaling for the Mississippi River basin. The parallelization of a single large basin such as Mississippi River basin is quite challenging due to the long main stem of the basin. The total routing time can be impacted by mainstem routing time to greater degree as the size of the mainstem domain increases with increasing number of DM processors, though the computational time for the tributary routing continues to decrease (not shown in Figure 7). As a result, there are no efficiency gains for both routing methods in the Mississippi River when using more than 20–30 DM processors.

4.2. River Discharge Simulations

The computational efficiency achieved by the developed parallelization scheme enables us to perform global river discharge simulations with multiple different settings. We discuss the impact of each model configuration on the discharge simulations at annual, monthly, and daily scales.
4.2.1. Annual Scale

Figure 8 shows the mean annual discharge at the 24 major river outlets for the eight simulations (including two routing schemes) summarized in Table 2. As a reference, the global river discharge data set from Dai (2017) denoted as D17 is also shown. The D17 data set is monthly observed discharge data at the farthest

Figure 8. Effects of runoff inputs, routing methods, and networks on annual discharges based on VIC-MERIT and CLM-MERIT. CLM, Community Land Model; VIC, Variable Infiltration Capacity.
downstream gauges over the globe from 1990 and 2014, with data missing gaps filled by linear regression between CLM3 simulated discharge and the gauge observation (see Dai & Trenberth, 2002; Dai et al., 2009). The D17 discharge data include human impacts on the discharge. The river basins where none of the simulations is close to D17 are river basins including large lakes, such as St. Lawrence.

Different runoff inputs (VIC vs. CLM) affect the mean annual discharge to a much greater degree than the routing configurations by noting a clear disparity between blue-cyan (CLM) and red-orange (VIC). Figure 9 shows the difference in mean annual runoff depth between CLM and VIC. For direct comparison between two different resolution runoff data (VIC: 0.125° vs. CLM: 0.5°), VIC runoff was aggregated at 0.5° with conservative mapping. Table 3 also summarizes basin area defined with each network, annual runoff volume based on VIC and CLM. Comparing Figures 8 and 9 (and Table 3), larger differences in the annual mean discharge between CLM and VIC can be explained by difference in runoff input over the basin. For example, annual discharge based on VIC runoff is greater than CLM runoff at Congo (2 - river number in Figures 5 and 6), Yenisey (9), Mackenzie (17), and Yukon (22) where corresponding difference in mean runoff over much of drainage basins. The opposite difference is seen in Paraná (8), Uruguay (21), and Mississippi river (6) basins.

Compared to the effect of runoff inputs, the majority of the rivers exhibit little difference among the different river networks (note small blue-cyan difference and red-orange difference in Figure 8). Nevertheless, there are noticeable differences in annual discharge for several rivers, for example, Ob, Xijiang, St. Lawrence, and Danube. For the St Lawrence River, the annual discharge differences between MERIT-Hydro and HDMA are larger than the difference due to runoff inputs. A larger effect of river networks on the annual discharge is seen in the simulations with CLM runoff input (i.e., blue-cyan difference) compared to VIC runoff input (red-orange difference). Table 3 indicate at St. Lawrence, CLM based mean annual runoff volumes are 348 km$^3$/yr for HDMA and 312 km$^3$/yr for MERIT-Hydro and this runoff difference (36 km$^3$/yr) is slightly greater than the difference between CLM and VIC (31 km$^3$/yr for HDMA). One reason for this is that the CLM land mask excludes major water bodies (e.g., Great Lakes), and due to the difference in the catchment size between MERIT-Hydro and HDMA, fractions of the CLM land area covered by the catchments in MERIT-Hydro and HDMA become substantially different along the shoreline of the lakes, which result in substantial difference in the total runoff volumes between the two networks. Meanwhile, VIC produces runoff even over the interior water bodies. Another river basin with large network impacts on discharge is the Ob River. We found that HDMA includes the Endorheic areas seen around the Russia-Kazakhstan border, which MERIT-Hydro correctly exclude. This explains the large difference in the drainage areas.
For an annual scale, routing schemes should have a minimal impact on the simulations because routing schemes affect timing and attenuation of flood wave, but do not alter annual scale water balance at any point of the river reaches.

4.2.2. Monthly Scale

The differences in monthly discharge due to the routing configurations are larger than in the annual simulations (Figure 10). Routing methods do begin to play a role in sensitivity of network choice to the discharge at a monthly scale for certain river basins. First, as shown in top panel in Figure 10, the difference in monthly discharge due to the networks is small but routing schemes have larger impacts (blue and red points do not overlap in many months). Second (Figure 10, middle panel), the difference due to the networks is smaller for one routing method (i.e., IRF), than the other (i.e., KWT). Finally (bottom panel), both routing methods show large difference in monthly discharge between two river networks.

4.2.3. Seasonal Cycle

Figure 11 shows the seasonal cycle on a monthly scale. As expected, runoff inputs have much greater effect on seasonality in the simulations than any other factors (VIC: red-orange vs. CLM: blue-cyan in Figure 11). The difference in the routing schemes (KWT: solid vs. IRF: dash) also causes larger differences in the timing of the annual peak than different river networks for some rivers (e.g., Amazon, Congo, Changjiang, Parana,

| No. | River   | HDMA Basin area (km²) | VIC mean annual runoff (m³/yr) | CLM mean annual runoff (m³/yr) | MERIT-hydro Basin area (km²) | VIC mean annual runoff (m³/yr) | CLM mean annual runoff (m³/yr) |
|-----|---------|-----------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| 1   | Amazon  | 4,671,598             | 5,618                         | 5,049                         | 4,672,174                    | 5,618                         | 5,045                         |
| 2   | Congo   | 3,615,448             | 2,740                         | 1,275                         | 3,604,610                    | 2,737                         | 1,267                         |
| 3   | Orinoco | 821,602               | 1,010                         | 980                           | 820,993                      | 1,009                         | 978                           |
| 4   | Changjiang | 1,678,381          | 919                            | 968                           | 1,676,435                    | 918                           | 965                           |
| 5   | Brahmaputra | 513,949              | 526                            | 485                           | 512,384                      | 524                           | 482                           |
| 6   | Mississippi | 2,884,792            | 675                            | 898                           | 2,902,761                    | 675                           | 895                           |
| 7   | Yenisey | 2,443,926             | 631                            | 492                           | 2,439,896                    | 630                           | 487                           |
| 8   | Parana  | 2,521,089             | 698                            | 943                           | 2,558,636                    | 699                           | 946                           |
| 9   | Lena    | 2,437,270             | 653                            | 371                           | 2,423,299                    | 649                           | 368                           |
| 10  | Mekong  | 548,205               | 528                            | 353                           | 547,283                      | 527                           | 352                           |
| 11  | Tocantins | 757,304               | 425                            | 402                           | 758,458                      | 426                           | 402                           |
| 12  | Ob      | 2,877,919             | 495                            | 515                           | 2,485,279                    | 481                           | 485                           |
| 13  | Ganges  | 928,331               | 450                            | 447                           | 929,118                      | 450                           | 438                           |
| 14  | Irrawaddy | 123,869               | 163                            | 132                           | 124,781                      | 164                           | 132                           |
| 15  | St Lawrence | 773,333              | 317                            | 348                           | 773,359                      | 317                           | 312                           |
| 16  | Amur    | 1876,740              | 290                            | 329                           | 1,742,520                    | 285                           | 321                           |
| 17  | Mackenzie | 1,673,081            | 328                            | 271                           | 1,673,842                    | 328                           | 259                           |
| 18  | Xijiang | 313,546               | 241                            | 234                           | 302,376                      | 232                           | 226                           |
| 19  | Columbia | 601,898               | 200                            | 146                           | 601,360                      | 200                           | 145                           |
| 20  | Magdalena | 258,070              | 345                            | 272                           | 256,838                      | 345                           | 272                           |
| 21  | Uruguay | 244,074               | 150                            | 191                           | 243,648                      | 150                           | 191                           |
| 22  | Yukon   | 819,380               | 199                            | 135                           | 825,514                      | 202                           | 136                           |
| 23  | Danube  | 779,964               | 224                            | 241                           | 785,918                      | 229                           | 245                           |
| 24  | Niger   | 2,062,309             | 282                            | 365                           | 1,990,629                    | 280                           | 362                           |
Both routing schemes have two parameters that affect the simulations (KWT: manning coefficient and channel width factors, and IRF: celerity and diffusivity). These parameters in both methods impact how fast flood waves move and how the flood waves are attenuated. Therefore, this timing difference can be altered via parameter adjustments. Though parameter estimation is beyond the scope in this study, there is a need for more research on the sensitivity of streamflow to routing parameters and how routing parameters can be estimated via calibration.

4.2.4. Daily Scale

In recent years there has been a greater demand for quantitative information on river discharge at finer time scales for the global domain (Bierkens et al., 2015; Lin et al., 2019). This information provides a more comprehensive evaluation of the simulated hydrologic processes at local scales over the global domain. Although we do not focus on evaluating the skill of simulated streamflow compared to observations, we

Figure 10. Scatter plots of monthly discharges based on VIC-MERIT (x-axis) and VIC-HDMA (y-axis) at nine selected rivers. VIC, Variable Infiltration Capacity.

etc.}.
evaluate how the spatial scale of the river network could potentially affect the discharge simulations at daily time scales.

First, we identify how each river network resolves the locations of the gauges reporting daily observed discharge over the globe. Similarly to Lin et al., (2019), we overlayed the total 21,955 gauges compiled by Lin

**Figure 11.** River network effects on 1990–2008 annual cycle of monthly discharge based on VIC-MERIT and VIC-HDMA. VIC, Variable Infiltration Capacity.
et al. (2019) and Beck et al. (2015) over each river network, and we selected gauges based on the proximity between reach and reported gauge location and the drainage area difference between the river network and the gauge data. Here, the gauges with drainage area differences greater than 30% are eliminated. HDMA has 8,395 gauges while MERIT-Hydro has 15,915 gauges. 6,106 gauges are common to both river networks, which is used for subsequent analysis.

At the 6,106 gauges, we computed the mean difference and the absolute mean difference in the fraction between the 33-year daily discharge simulations using the two different river networks as well as using the two different routing methods. Figure 12 shows cumulative distributions of the mean difference (panel a) and absolute mean difference (panel b) between the two routing methods (blue-cyan) and river networks (red-orange). The mean difference between the routing methods (blue-cyan) is very small at the majority of locations if the same river network is used (see Figure 12a). This occurs because the choice of routing methods primarily affects the timing of flow, and differences in timing are averaged out when considering the long-term mean. The network choice does affect the mean difference due to some differences in contributory areas and reaches geometries between the two river networks (see Figure 12a). However, as shown in Figure 12b, the choice of routing methods has larger impact on the magnitude of daily discharge differences than the choice of the river networks. In this case, the mean absolute errors highlight differences in timing and the choice of routing method becomes more important. Figure 13 shows the spatial patterns of the absolute mean difference between the two river networks (panel a) and between the two routing methods (panel b). Both effects (choice of river networks and choice of routing methods) produce similar spatial pattern of absolute mean difference.

5. Limitations

Besides river reaches, lakes (both man-made and natural) are important hydrological features that directly affect river discharge over the land and into the oceans. In the vector representations of river systems, it is straightforward to define topological relationship between river reaches and lakes to generate vector river-lake network for river-lake routing models. This is because the shapes and locations of lake polygons as well as catchment boundaries and river reach lines are accurately represented in the modeled river systems. For example, the hydroLAKES data set (Messager et al., 2016) contains 1.42 million polygon features representing natural lakes and reservoirs with an area greater than 10 ha (0.1 km$^2$) with various physical information on the lakes (surface area, average depth, volume). Once reach-lake topology is built, each element is labeled as reach or lake and the processing order in the river-lake network can be determined in the same manner as reach-only network. In this way, rivers or lakes are identified for each element as cascading through the network. Incorporating lakes in mizuRoute is a topic of ongoing research.

Figure 12. Percent mean difference (left) and percent mean absolute difference (right) in daily flow simulations between routing choices: river network and routing scheme.
While the fine-scale river networks may not improve estimates of the freshwater flux to the ocean based on the results in Section 3.3, such network data does provide explicit connectivity of various sized lakes and reservoirs as well as gauge points to the catchments and river reaches. Moreover, it is straightforward to aggregate fine-scale vector-river networks and then produce simulations at variable spatial scales. Figure 14 shows an example of three different scaled MERIT-Hydro river networks (native scale, reach with second order or higher, and reaches with third order or higher). As shown in Figure 14, the MERIT-Hydro network can be upscaled and mimic the HDMA network. The method for river network aggregation that utilizes Pfafstetter codes together with various criteria (catchment areas, number of upstream reaches, etc.) is under development and can enable the existing network to be upscaled to the desired degree. This will help to understand how local relevant features (e.g., lakes) are connected to the reaches in the network with the different scales and consequently the simulated water cycle is affected at the spatiotemporal scale desired for the specific analysis (climate, flood prediction, etc.). Ultimately, the optimal river network scale could be determined for specific applications.

One of the challenges toward a more complete river model is modeling river distributaries commonly seen in river deltas but occasionally seen inland. Both MERIT-Hydro and HDMA do not represent river distributaries. Moreover, current augmentation of river network topology as well as river network domain decomposition assumes there are no river distributaries present in the river network. Though it seems to be reasonable to assume that river flows are split evenly between the distributary channels, physical mechanisms of water flows in deltas are complex due to interaction between the river flow, tides, and waves from the oceans (Wagner & Mohrig, 2019). Given many of the large deltas host large populations and farmlands, however, incorporating river distributaries into the river model will improve more precise fluvial conditions over the deltas and provide detailed flow information.

Our ultimate modeling goal is to implement the vector-based river model in an Earth System Model. While the main purpose of the river model in an Earth System Model is to feed estimated discharge to the ocean at accurate outlet locations, coupling the river model with the land model is important to simulate the terrestrial component of the hydrological cycle, including water fluxes from the river system to riparian areas.
Water fluxes from the river to the land include (1) water resulting in endorheic basins, which potentially help to simulate dynamics of natural lakes (shrinkage or expansion), (2) over-bank inundated water, which impacts land-atmosphere energy flux through enhanced evaporation (Dadson et al., 2010), and (3) riverbed infiltration causing water loss in the river reach and soil moisture recharge, which depends on river water height relative to water table in soil. Moreover, the water in rivers as well as lakes is lost through evaporation. Such water loss can be computed using the information from the land model.

6. Conclusions

A vector-based river network represents physical features of the river system (river tortuosity, catchment area) more efficiently than high-resolution gridded river network data. This is beneficial for the river routing model, especially for computationally efficiency. However, it is still necessary to parallelize river routing computations to accommodate finer-scale river data sets (e.g., the MERIT-Hydro network) and to implement more complex routing methods (e.g., including water temperature, nitrogen, phosphorous, etc.). To reduce the computation time for large-domain, fine-scale river routing model, we developed a topological-based river network domain decomposition. Our computational scaling results demonstrated improvement in computational efficiencies up to approximately 60 processors for global river network routing. The challenge for parallelization remains for a single large river basin. Our current domain decomposition method (i.e., partitioning into mainstems and tributaries) generates longer mainstems as the number of domains increases. This increased mainstem computation time impacts total routing computational time.

For the simulation of river discharge from larger river basins, a spatial scale of vector river networks (HDMA vs. MERIT-Hydro) has a limited impact on river discharge estimates, particularly at longer time scales. The HDMA and MERIT differences in the catchments and reaches are seen on very fine scales. In other words, MERIT-Hydro network represents small-scale rivers in the headwater catchments many of which are not present in HDMA network data. This is one of the reasons why the choice of the two vector-river network
data sets is shown to have a limited impact on simulations of the river discharge to the oceans. However, fine-scale vector-river network data provides more discharge gauge points which can be used for model evaluations. Furthermore, a finer-scale vector river network resolves smaller lakes and reservoirs that can be included in the river network. The impact of different scales of river-lake network on the discharge simulations at different time scales will need to be evaluated in the future. This will help determine which scale of river network is optimal for the desired model applications. Such optimal river network, together with the parallelization, improves computational solutions of river system dynamics.

Data Availability Statement

All simulations presented in this study are produced by the mizuRoute code available at Zenodo (https://doi.org/10.5281/zenodo.4737837). The mizuRoute outputs used for the analyses is available at Zenodo (https://doi.org/10.5281/zenodo.4733293). The other data sets used in this study are in the public domain and cited within the study. These include (1) MERIT-Hydro river network data (ESRI shapefile) available at https://www.reachhydro.org/home/params/merit-basins; (2) HDMA river network data (ESRI shapefile) available at https://pubs.er.usgs.gov/publication/ds1053; and (3) Dai and Trenberth Global River Flow and Continental Discharge data set at https://rda.ucar.edu/datasets/ds551.0/.

Acknowledgments

This study was funded by the NCAR re-investment fund “Accelerated development of CTSM” and partially by NASA AIST-16 fund (AIST-16-0081) “Climate Risks in the Water Sector: Advancing the Readiness of Emerging Technologies in Climate Downscaling and Hydrologic Modeling”. The Global Water Futures programme also provided support for this work. The simulations and visualization presented here were produced through the Cheyenne and Casper computational resources (https://doi.org/10.5065/D8RX091H) at the NCAR-Wyoming Supercomputing Center supported by the National Science Foundation and operated by NCAR's Computational and Information Systems Laboratory. We would like to thank David Lawrence, Sean Swenson, Keith Oleson, Bill Sacks, Mariana Vertenstein, Jim Edwards, Michael Barlage and Andy Wood for discussions on the parallelization scheme and implementation.

References

Beck, H. E., de Roo, A., & van Dijk, A. I. J. M. (2015). Global maps of streamflow characteristics based on observations from several thousand catchments. *Journal of Hydro meteorology, 16*(4), 1478–1501. https://doi.org/10.1175/jhm-d-14-0155.1

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miraless, D. G., van Dijk, A. I. J. M., et al. (2019). MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bulletin of the American Meteorological Society, 100*, 473–500. https://doi.org/10.1175/BAMS-D-17-0316.1

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., et al. (2015). Hyper-resolution global hydrological modeling: What is next? *Hydrological Processes, 29*, 310–320. https://doi.org/10.1002/hyp.10391

Clark, E. A., Sheffield, J., van Vliet, M. T. H., Nijsen, B., & Lettenmaier, D. P. (2015). Continental runoff into the oceans (1950–2008). *Journal of Hydro meteorology, 16*(4), 1502–1520. https://doi.org/10.1175/JHM-D-14-0183.1

Dadson, S. J., Ashpole, I., Harris, P., Davies, H. N., Clark, D. B., Blyth, E., & Taylor, C. M. (2010). Wetland inundation dynamics in a changing climate: What is next? *Hydrological Processes, 24*, 3142–3145. https://doi.org/10.1002/hyp.8070

Dai, C. H., Habets, F., Maidment, D. R., & Yang, Z.-L. (2011). Rapid applied to the SIM-France model. *Hydrological Processes, 25*, 473–500. https://doi.org/10.1002/hyp.8288

Dai, C. H., Yang, Z.-L., & Famiglietti, J. S. (2013). Quantification of the upstream-to-downstream influence in the Muskingum method and implications for speedup in parallel computations of river flow. *Water Resources Research, 49*, 2783–2800. https://doi.org/10.1002/wrcr.20250

Döll, P., Kaspar, F., & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: Model tuning and validation. *Journal of Hydrology, 270*, 105–134. https://doi.org/10.1016/S0022-1694(02)00283-4

Gochis, D. J., Barlage, M., Dugger, A., FitzGerald, K., Karsten, L., McAllister, M., et al. (2018). The WRF-Hydro modeling system technical description (Version 5.0). NCAR Technical Note, 107.

Koirala, S., Hibarayashi, Y., Mahendran, R., & Kanase, S. (2014). Global assessment of agreement among streamflow projections using CMIP5 model outputs. *Environmental Research Letters, 9*, 64017. https://doi.org/10.1088/1748-9326/9/6/064017

Lawrence, D. M. (2019). The Community Land Model Version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems, 11*, 4245–4287. https://doi.org/10.1029/2018MS001583

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world’s large river systems. *Hydrological Processes, 27*, 2171–2186. https://doi.org/10.1002/hyp.9740

Lehner, B., Verdin, K., Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions, American Geophysical Union, 89*(10), 93–94. https://doi.org/10.1029/2008EO100001

Li, H.-Y., Leung, L. R., Getirana, A., Huang, M., Wu, H., Xu, Y., et al. (2015). Evaluating global streamflow simulations by a physically based routing model coupled with the Community Land Model. *Journal of Hydro meteorology, 16*, 948–971. https://doi.org/10.1175/jhm-d-14-0079.1

Li, H.-Y., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M., & Leung, L. R. (2013). A physically based runoff routing model for land surface and earth system models. *Journal of Hydro meteorology, 14*, 808–828. https://doi.org/10.1175/jhm-d-12-0151.1
Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, 99, 14415–14428. https://doi.org/10.1029/94jd04083

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., et al. (2019). Global reconstruction of naturalized river flows at 2.94 million reaches. *Water Resources Research*, 55, 6499–6516. https://doi.org/10.1029/2019wr025287

Lohmann, D., Lettenmaier, D. P., Liang, X., Wood, E. F., Boone, A., Chang, S., et al. (1998). The project for intercomparison of land-surface parameterization schemes (PILPS) phase 2c) Red-Arkansas River basin experiment. *Global and Planetary Change*, 19, 161–179. https://doi.org/10.1016/S0921-8181(98)00036-6

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, 7, 13603. https://doi.org/10.1038/ncomms13603

Mizukami, N., Clark, M. P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., et al. (2016). MizuRoute version 1: A river network routing tool for a continental domain water resources applications. *Geoscientific Model Development*, 9, 2223–2238. https://doi.org/10.5194/gmd-9-2223-2016

Nguyen-Quang, T., Polcher, J., Ducharne, A., Arsouze, T., Zhou, X., Schneider, A., & Fita, L. (2018). ORCHIDEE-ROUTING: Revising the river routing scheme using a high-resolution hydrological database. *Geoscientific Model Development*, 11, 4965–4985. https://doi.org/10.5194/gmd-11-4965-2018

Paiva, R. C. D., Buarque, D. C., Collischonn, W., Bonnet, M.-P., Frappart, F., Calmant, S., & Bulhões Mendes, C. A. (2013). Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. *Water Resources Research*, 49, 1226–1243. https://doi.org/10.1002/wrcr.20067

Shin, S., Pokhrel, Y., & Miguez-Macho, G. (2019). High-resolution modeling of reservoir release and storage dynamics at the continental scale. *Water Resources Research*, 55, 787–810. https://doi.org/10.1029/2018WR023025

Shin, S., Pokhrel, Y., Yamazaki, D., Huang, X., Torbick, N., Qi, J., et al. (2020). High-resolution modeling of river-floodplain-reservoir inundation dynamics in the Mekong River Basin. *Water Resources Research*, 56, e2019WR026449. https://doi.org/10.1029/2019WR026449

Siqueira, V. A., Paiva, R. C. D., Fleischmann, A. S., Fan, F. M., Ruhoff, A. L., Pontes, P. R. M., et al. (2018). Toward continental hydrologic-hydrodynamic modeling in South America. *Hydrology and Earth System Sciences*, 22, 4815–4842. https://doi.org/10.5194/hess-22-4815-2018

Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., et al. (2018). PCR-GLOBWB 2: A 5 arcmin global hydrological and water resources model. *Geoscientific Model Development*, 11, 2429–2453. https://doi.org/10.5194/gmd-11-2429-2018

Thober, S., Cuntz, M., Kelblng, M., Kumar, R., Mai, J., & Samaniego, L. (2019). The multiscale routing model mRM v1.0: Simple river routing at resolutions from 1 to 50 km. *Geoscientific Model Development*, 12, 2501–2521. https://doi.org/10.5194/gmd-12-2501-2019

van Beek, L. P. H., Wada, Y., & Bierkens, M. F. P. (2011). Global monthly water stress: 1. Water balance and water availability. *Water Resources Research*, 47. https://doi.org/10.1029/2010wr009791

Verdin, K. L. (2017). Hydrologic derivatives for modeling and analysis—A new global high-resolution database, 24. https://doi.org/10.3133/doi1053

Verdin, K. L., & Verdin, J. P. (1999). A topological system for delineation and codification of the Earth's river basins. *Journal of Hydrology*, 218, 1–12. https://doi.org/10.1016/S0022-1694(99)00011-6

Wada, Y., de Graaf, I. E. M., & van Beek, L. P. H. (2016). High-resolution modeling of human and climate impacts on global water resources. *Journal of Advances in Modeling Earth Systems*, 8, 735–763. https://doi.org/10.1002/2015MS000618

Wagner, W., & Mohrig, D. (2019). Flow and sediment flux asymmetry in a branching channel delta. *Water Resources Research*, 55, 9563–9577. https://doi.org/10.1029/2019WR026050

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, W09701. https://doi.org/10.1029/2012wr012313

Yamazaki, D., de Almeida, G. A. M., & Bates, P. D. (2013). Improving computational efficiency in global river models by implementing the local inertial flow equation and a vector-based river network map. *Water Resources Research*, 49, 7221–7235. https://doi.org/10.1002/wrcr.20552

Yamazaki, D., Ikehama, 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, 5033–5073. https://doi.org/10.1029/2019WR024873

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, 2241–2251. https://doi.org/10.5194/hess-13-2241-2009

Zajac, Z., Revilla-Romero, B., Salamon, P., Burek, P., Hirpa, F. A., & Beck, H. (2017). The impact of lake and reservoir parameterization on global streamflow simulation. *Journal of Hydrology*, 548, 552–568. https://doi.org/10.1016/j.jhydrol.2017.03.022