

University of Montana

ScholarWorks at University of Montana

Biological Sciences Faculty Publications

Biological Sciences

3-15-2011

Automated Upscaling of River Networks for Macroscale Hydrological Modeling

Huan Wu

John S. Kimball

Nate Mantua

Jack Arthur Stanford

The University of Montana, jack.stanford@umontana.edu

Follow this and additional works at: https://scholarworks.umt.edu/biosci_pubs



Part of the [Biology Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Wu, Huan; Kimball, John S.; Mantua, Nate; and Stanford, Jack Arthur, "Automated Upscaling of River Networks for Macroscale Hydrological Modeling" (2011). *Biological Sciences Faculty Publications*. 96. https://scholarworks.umt.edu/biosci_pubs/96

This Article is brought to you for free and open access by the Biological Sciences at ScholarWorks at University of Montana. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

Automated upscaling of river networks for macroscale hydrological modeling

Huan Wu,^{1,2} John S. Kimball,^{1,2} Nate Mantua,³ and Jack Stanford²

Received 6 November 2009; revised 5 January 2011; accepted 13 January 2011; published 15 March 2011.

[1] We developed a hierarchical dominant river tracing (DRT) algorithm for automated extraction and spatial upscaling of basin flow directions and river networks using fine-scale hydrography inputs (e.g., flow direction, river networks, and flow accumulation). In contrast with previous upscaling methods, the DRT algorithm utilizes information on global and local drainage patterns from baseline fine-scale hydrography to determine upscaled flow directions and other critical variables including upscaled basin area, basin shape, and river lengths. The DRT algorithm preserves the original baseline hierarchical drainage structure by tracing each entire flow path from headwater to river mouth at fine scale while prioritizing successively higher order basins and rivers for tracing. We applied the algorithm to produce a series of global hydrography data sets from $1/16^\circ$ to 2° spatial scales in two geographic projections (WGS84 and Lambert azimuthal equal area). The DRT results were evaluated against other alternative upscaling methods and hydrography data sets for continental U.S. and global domains. These results show favorable DRT upscaling performance in preserving baseline fine-scale river network information including: (1) improved, automated extraction of flow directions and river networks at any spatial scale without the need for manual correction; (2) consistency of river network, basin shape, basin area, river length, and basin internal drainage structure between upscaled and baseline fine-scale hydrography; and (3) performance largely independent of spatial scale, geographic region, and projection. The results of this study include an initial set of DRT upscaled global hydrography maps derived from HYDRO1K baseline fine-scale hydrography inputs; these digital data are available online for public access at <ftp://ftp.ntsg.umt.edu/pub/data/DRT/>.

Citation:

Wu, H., J. S. Kimball, N. Mantua, and J. Stanford (2011), Automated upscaling of river networks for macroscale hydrological modeling, *Water Resour. Res.*, 47, W03517, doi:10.1029/2009WR008871.

1. Introduction

[2] Much effort has been devoted to improving the characterization and simulation of precipitation runoff and routing in global land surface models (LSMs), which define lower boundary conditions for atmospheric general circulation models (GCMs). These LSMs include macroscale hydrological models (MHMs), such as the variable infiltration capacity (VIC) model that provide relatively complete representation of the terrestrial water budget in which the lateral flow of water is a necessary component [Liston *et al.*, 1994; Miller *et al.*, 1994; Nijssen *et al.*, 1997; Olivera *et al.*, 2000; Arora *et al.*, 2001]. Comparisons between simulated streamflow and observed hydrograph data are also widely used for calibration and validation of LSMs. Regional upscaling of river networks and flow directions to

relatively coarse (e.g., $1/16^\circ$ to 2°) spatial scales commensurate with GCM, LSM, and MHM simulations is necessary for representing the lateral movement of water, sediment, and nutrients in continental and global scale modeling studies by providing the necessary flow paths and river networks for runoff routing. Coarse-scale hydrography variables such as river flow direction (FDR), river network structure, flow accumulation area (FAC), and flow distance represent a gross simplification of complex landscapes and cannot be obtained by direct measurements. Coarse-scale hydrography is generally determined by abstracting (upsampling) from relatively fine-scale hydrography that is more representative of real world conditions. The earliest coarse-scale hydrography products have been derived through manual interpretation and digitization of published maps [e.g., Vörösmarty *et al.*, 1989; Miller *et al.*, 1994; Liston *et al.*, 1994; Marengo *et al.*, 1994; Costa and Foley, 1997]. However, manual corrections can be subjective and labor intensive, particularly for larger regions, and must also be repeated when deriving similar products at different spatial scales [O'Donnell *et al.*, 1999]. Regional- or continental-scale applications can involve millions of grid cells, making manual processing largely impractical and requiring the use of more automated methods.

[3] Comparisons between upscaled river networks and baseline fine-scale hydrography data can provide reliable

¹Flathead Lake Biological Station, Division of Biological Sciences, University of Montana, Polson, Montana, USA.

²Numerical Terradynamic Simulation Group, College of Forestry and Conservation, University of Montana, Missoula, Montana, USA.

³Climate Impacts Group, Joint Institute for the Study of the Atmosphere and Ocean and School of Aquatic and Fisheries Sciences, University of Washington, Seattle, Washington, USA.

validation of upscaling algorithms, while the accuracy of the derived coarse-scale hydrography should be commensurate with the accuracy of the fine-scale hydrography inputs. Relatively fine-scale and accurate hydrography data are available globally. The hydrologically corrected global HYDRO1K [Gesch *et al.*, 1999; U.S. Geological Survey, 2000] is one of the most widely used global hydrography data sets. As the successor of HYDRO1K, HydroSHEDS is now available in many regions and provides superior scale and quality relative to its predecessor [Lehner *et al.*, 2008]. Given reliable baseline fine-scale hydrography inputs (e.g., HYDRO1K, HydroSHEDS), the essential work of an upscaling algorithm is to preserve the accuracy of the baseline hydrography at coarser spatial scales. Manual corrections can be avoided if an upscaling algorithm can effectively exploit the fine scale information defined from the baseline hydrography inputs. An upscaling algorithm can be considered successful or validated if the accuracy of the derived coarse-scale hydrography is commensurate with the baseline fine-scale hydrography inputs. The purpose of this paper is to put forward a new algorithm for automatic upscaling of river networks that addresses many of the limitations of earlier methods, including coarse-scale distortions of baseline fine-scale hydrography and the need for manual interpretation and correction of associated errors.

2. Background

[4] Increasing efforts have been made over the past decade to develop more efficient algorithms for river network upscaling and generation of coarse-scale hydrography while reducing the need for manual corrections. We recognize two general groups of river network upscaling algorithms, based on the spatial dimension of information used for determining coarse-scale flow directions and river network structure.

[5] The cell-to-cell (CTC) based algorithms use 0-dimension information to define flow direction. The CTC approach employs single representative values for hydrographic variables (e.g., elevation, FAC, aggregated flow direction) on a grid cell-by-cell basis to determine coarse grid-scale flow directions. These representative values are derived from finer-scale hydrography using spatial averaging, resampling, maximum value or vector operation methods. In the CTC approach, subgrid-scale information is only used to derive representative upscale values. No fine-scale information outside of the coarse grid cell under consideration is employed to determine flow direction using the CTC approach. Oki and Sud [1998] averaged elevation information from a 5' digital elevation model (DEM) to obtain a 1° scale global DEM and defined associated flow directions using a maximum downhill (steepest descent) method with pit grid cells and other discontinuous river segments eliminated through manual correction. An approach similar to "stream burning" [Dirmeyer, 1995; Maidment, 1996] was first used for geolocating vector river networks on DEM data and was then used for determining flow directions using the maximum downhill method after pit-removal processing [Graham *et al.*, 1999]. A similar method was used to generate a 0.5° resolution global river network [Renssen and Knoop, 2000]. However, pit removal algorithms [e.g., Jenson and Domingue, 1988; Tarboton

et al., 1991] developed for deriving flow directions and river networks from fine-scale DEM information generally do not work well using coarse-scale DEMs because of inadequate representation of terrain heterogeneity. Therefore, coarse-scale DEM-based upscaling algorithms generally require intensive manual correction. Comanor *et al.* [2000] improved Graham *et al.*'s [1999] upscaled flow directions and river networks by adopting a basin area correction process. A global flow direction data set at 2.81° resolution was derived by defining flow directions of coarser grid cells using corresponding flow directions from a finer-scale (1°) grid with the largest FAC values [Arora and Boer, 1999]. A similar method was used by Döll and Lehner [2002] to generate a global flow direction data set (DDM30) at 1/2° resolution with extensive manual correction. A double-maximum method was developed to track the river beyond the boundary of a selected coarse grid cell using an offset grid cell identified by the largest FAC value [Olivera *et al.*, 2002]. Fekete *et al.* [2001] assigned coarse grid cell flow directions using the inverse of the maximum FAC. A vector addition method was also explored to aggregate subgrid-scale flow directions to define coarse grid flow directions in which only four cardinal directions are derived [Shaw *et al.*, 2005].

[6] The second group of upscaling algorithms utilizes a dominant river segment tracing (DST) approach and quasi 1-dimensional information to determine flow directions. The DST approach determines coarse grid-scale flow directions by employing some baseline subgrid-scale information (e.g., flow directions and river networks). The DST approach generally uses the maximum FAC value to first identify the dominant river segment of a coarse grid cell and then traces the fine-scale dominant river segment along a user-defined distance outside of the coarse grid cell. The river tracing method was first explored by O'Donnell *et al.* [1999] to determine flow direction by tracing the fine-scale river network beyond the boundary of the selected coarse grid cell. A fine-scale flow direction-based algorithm was improved by tracking dominant river segments after they leave coarse grid cells [Arora and Harrison, 2007]. A more advanced river network tracing method was developed to determine the downstream coarse grid cell by tracing fine-scale vector river networks, in which the dominant river segment was defined by the largest upstream flow length [Olivera and Raina, 2003]. The CTC and DST methods have constraints, including (1) only local information, usually defined within a 3 × 3 coarse grid cell matrix centered within the selected coarse grid cell, is employed to determine flow direction of the coarse grid cell, leading to potential misclassification of river segments, especially for complex terrain and associated river networks; and (2) baseline fine-scale hydrography is generally underutilized in the upscaling algorithms. The FAC represents the entire upstream source area, which can be global information reflecting the overall drainage (runoff routing) structure of the given domain. However, in previous approaches the FAC is essentially used as local rather than global information to identify the outlet pixel of individual coarse grid cells and local river segments within the cell; the FAC is used to represent the local river segment but not the entire river. To improve flow directions and reduce large basin aggregation of smaller catchments during spatial upscaling,

Fekete et al. [2001] incorporated sub-basin information to constrain the range of the “uphill” search area within sub-basin cells (the improved method is hereafter referred to as NSABE). However, the NSABE only uses FAC values to determine flow directions within each subbasin, while sub-basin boundary constraints may not provide sufficient global information to identify and preserve river structure during the upscale process, especially when there are multiple rivers (or flow paths) in a subbasin. With these limitations, previous approaches can result in two common error sources. First, these methods cannot guarantee that flow directions derived for adjacent upstream and downstream coarse grid cells are defined from the same dominant river going through them; this is especially problematic when tracing sinuous river networks using a single, representative flow direction, leading to potential errors in upscale flow direction, basin area, and river length calculations. Second, without sufficient global information defined from the baseline fine-scale hydrography, in some situations multiple dominant rivers cannot be correctly identified and conserved when they drain through the same grid cell, leading to potential misclassification of river segments. Therefore, larger basins unreasonably expand at the expense of smaller catchments during the upscaling process. This effect is illustrated in Figure 1a, where both upstream areas of rivers R4 and R2 drain through the same coarse-scale grid cell in row 3 and column 6 of the specified grid (hereafter designated as cell [3,6]); during the upscaling process, the drainage area of R4 is absorbed into the R2 stem river basin, which has a larger FAC area, resulting in excessive growth of the R2 drainage area and a corresponding reduction in the R4 drainage area. The absorbed upstream river

segments are “excluded” from R4 and represent “false” river (R2) segments in the upscaled results. Without global tracing of fine-scale river features during the upscaling process, segments of some globally dominant rivers in complex drainage areas (e.g., coarse grid cells with multiple major rivers such as cell [1,2], cell [2,2], and cell [3,6] in Figure 1a) cannot be correctly identified and preserved by either (CTC or DST) approach, leading to coarse-scale distortions of river shape and length, basin shape, and area. These distortions generally require manual correction and could have significant, negative impacts on hydrological flow-related model calculations, including water routing, sediment transport, and stream temperature simulations. In contrast, the CTC approach tends to produce more false river segments and exclude real river segments, resulting in shorter river lengths and relatively more bias in basin shape and area calculations.

3. Methods

[7] We designed a river upscaling methodology based on the following criteria. In hydrological models using the D8 single flow direction approach, all runoff generated in a grid cell drains into the immediate downstream cell according to an assigned single flow direction (out of eight possible directions). The D8 approach is more appropriate and effective in meeting hydrological model requirements when a greater proportion of runoff from the underlying fine-scale hydrography drains in a consistent flow direction matching the assigned flow direction of the overlying grid cell. An effective upscaling algorithm should assign as many cells as possible with coarse-scale flow directions consistent with

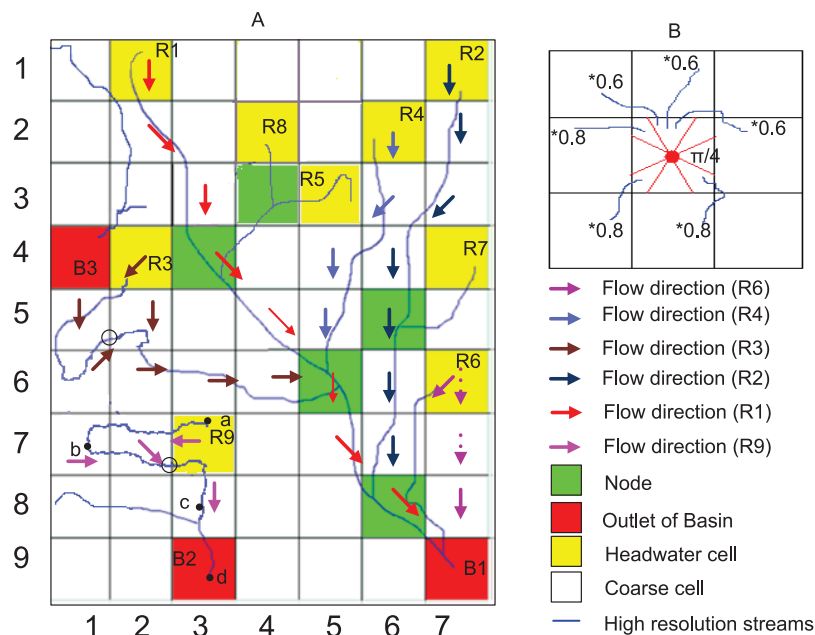


Figure 1. (a) Example schematic of the DRT algorithm for a hypothetical landscape domain and associated matrix of cells with three main basins (B1, B2, and B3) and outlet cells (in red). (b) To avoid unreasonably assigning more cardinal directions than diagonal directions, the DRT gives equal opportunity to each of eight potential flow directions (four cardinal directions and four diagonal directions) by dividing each cell into eight evenly distributed sections radiating from the cell centroid, where each section has the same central angle $\pi/4$.

the predominant underlying drainage from each cell, while preserving the overall dominant river structure of the specified domain in relation to the baseline, fine-scale hydrography inputs. Potential conflicts can arise during the upscaling process between assigning a predominant drainage direction and preserving the overall dominant river structure of the baseline hydrography; the latter is generally more important to the overall performance of an upscaling approach [Döll and Lehner, 2002]. Additional criteria for model development and upscaling include a desire to preserve drainage structure (e.g., river lengths, basin shapes, and internal drainage areas) in relation to the baseline hydrography and derive upscaled flow directions and river networks automatically, with consistent performance, at different scales and projections for all regions.

3.1. Hierarchical Dominant River Tracing Algorithm

[8] A hierarchical dominant river tracing (DRT) algorithm was developed in accordance with the above criteria for automated extraction and spatial upscaling of flow directions and river networks. A primary description of the DRT algorithms is provided in this section, while supporting information, including an algorithm flowchart, is provided in Appendix A. For clarity of text, fine-scale grid cells are hereafter referred to as pixels and upscaled grid cells as cells. The DRT prioritizes downstream rivers and flow directions by exploiting both global and local hydrographic characteristics, including overall drainage structure (two-dimension), dominant rivers and river segments (one-dimension), and FAC values (zero-dimension) from baseline, fine-scale hydrography inputs. The DRT identifies the hierarchical river structure (e.g., stream-order ranking of dominant and subdominant rivers and tributaries) of a given region by the FAC values of river mouths, sinks, and junctions. Dominant rivers are globally and hierarchically selected to assign flow directions for all intersecting cells by tracing each entire river (tributary) at the subgrid scale from headwater pixel to river mouth or sink (junction) pixel. For a given cell, the DRT selects the longest effective dominant river (LEDR) segment within the cell to determine the flow direction when it does not conflict with preserving the overall global dominant drainage structure. The LEDR of a cell is defined as the river segment (or dominant flow path when there is no river in the cell) that dominates the local drainage of the cell. During the hierarchical tracing process (i.e., larger rivers are prioritized over smaller rivers for tracing), the LEDR for a cell is generally identified as the river segment with a relatively large (not necessarily the largest) FAC relative to other river segments within the cell and exceeding a user-defined minimum length threshold within the cell (described in section 3.2.1). The LEDR generally has more tributaries, which increases the probability that the selected river segment has the longest flow path and collects the majority of runoff from pixels within the cell. In many cases, the LEDR of a cell belongs to the most dominant river of all large rivers (if any) draining through the cell. However, if two globally dominant rivers drain through the same cell and the secondary river has a longer length within the cell, and actually dominates local (within cell) drainage relative to the more dominant river, the secondary river is selected as the LEDR

for the cell if this does not lead to discontinuity of the more dominant river and degrade the global drainage structure.

[9] The DRT first identifies the dominant basin and river of the defined study area and assigns single flow directions for cells along dominant rivers beginning from headwater cells to basin mouth or sink cells. Subdominant rivers and tributary flow paths are then identified and ordered according to their respective FACs. The priority for assigning flow directions is assigned to successively higher-order rivers until all cells in the most dominant basin have assigned flow directions. The DRT then selects progressively smaller, less dominant basins and rivers and assigns flow directions in a similar manner until all cells in the given study area have been assigned flow directions. The DRT procedure is illustrated in Figure 1a for a hypothetical landscape matrix of cells with three main basins (B1, B2, and B3) and outlet cells (in red). The outlet cell of the largest basin B1 is first identified, and is designated by cell [9,7] in Figure 1a. A reverse tracing algorithm (Appendix A) is then applied to identify the globally dominant river (R1) in B1 from the outlet (cell [9,7]) to headwater (cell [1,2]). The DRT traces the river from headwater pixel in the headwater cell (cell [1,2]) to basin outlet pixel in the outlet cell (cell [9,7]) along the R1 flow path and assigns flow directions (red arrows) for all intersecting cells. All junctions (i.e., pixels with more than one tributary) in R1 are identified and appended to a junction array, which is used to record and search for subdominant rivers in B1. The outlet of the second most dominant river (R2) in B1 is identified by the junction cell [8,6] (in green), where the junction pixel corresponding to R2 has the maximum FAC value among all junctions in the junction array. R2 is then identified using the same reverse tracing algorithm starting from the junction cell [8,6]. All cells in R2 are then assigned flow directions (dark blue arrows) by tracing from the corresponding headwater cell [1,7]. All R2 junctions are identified and appended to the junction array. The junction for identifying R2 is then removed from the junction array. The remaining B1 subdominant rivers (R3, R4, R5, R6, R7, R8) are traced in a similar manner and assigned in hierarchical order by the FAC values in the junction array. All remaining cells without rivers draining through them are also assigned flow directions by tracing nonriver flow paths according to flow directions and FAC values defined from the baseline fine-scale hydrography. After all cells in B1 are assigned, other basins (B2, B3) and cells of the given region are searched and classified in a similar manner until all cells in the region are assigned flow directions.

3.2. Rules for Preserving River Structure During Spatial Upscaling

[10] Three general rules are defined to preserve baseline globally and locally dominant river structure when assigning flow directions for each cell in a given domain; these rules are designed to avoid assignments of unreasonably more cardinal than diagonal flow directions, minimize the occurrence of excluded or false river segments, and conserve important rivers under complex situations.

3.2.1. Selection of Flow Directions at the Grid Cell Level

[11] The first rule governs the selection of flow directions at the grid cell level. Each cell in the D8 flow

direction map has only four cardinal edges, which promotes the risk of assigning more cardinal than diagonal directions in upscaling algorithms using the river tracing method. To avoid unreasonably assigning more cardinal than diagonal directions for cell drainage, the DRT first divides each cell into eight evenly distributed potential flow directions (four cardinal sectors and four diagonal sectors) radiating from the cell center, where each sector has the same central angle $\pi/4$ (e.g., Figure 1b). The cell flow direction is then determined through one of the eight sectors depending on how the dominant river flows into the adjacent downstream cell. The DRT defines flow direction for each cell by tracing each river from headwater to river mouth, which identifies the downstream cell that becomes the upstream cell for the next flow direction assignment. Therefore, the LEDR of each cell needs to be identified during the cell-to-cell tracing of the entire river. To determine whether the selected dominant river is the LEDR of the downstream cell, the DRT traces the dominant river downstream an extra length following a user-defined threshold ($0.6 \times$ cell size for cardinal sectors and $0.8 \times$ cell size for diagonal sectors for this study), as shown in Figure 1b; this contrasts with the approach of *O'Donnell et al.* [1999], in which only rivers draining from diagonal sectors of a cell are traced to ensure that flow in the broader area is in a diagonal direction. However, a dominant river flowing from a cell through a diagonal (or cardinal) sector does not necessarily drain the majority of water from the cell to the adjacent diagonal (cardinal) downstream cell because the river may not maintain a sufficient length in the downstream cell. If the selected dominant river of a cell exceeds the specified length threshold in the adjacent downstream cell, it qualifies as the LEDR of the downstream cell. The flow direction for each cell is then determined as the direction of the adjacent downstream cell identified by the LEDR. Rivers draining in diagonal directions are generally closer to cell edges, have shorter lengths, and tend to be less dominant (e.g., collecting less runoff within the cell) in the immediate downstream cell than rivers draining in cardinal directions. A longer downstream length threshold is therefore specified for rivers draining through diagonal sectors relative to cardinal sectors to qualify as the LEDR of the downstream cell. The larger extra length threshold for diagonal sectors generally avoids unreasonable cardinal directions by ensuring approximately equal opportunities for selection of cardinal and diagonal directions when assigning flow directions for grid cells. The extra length thresholds for cardinal ($0.6 \times$ cell size) and diagonal ($0.8 \times$ cell size) sectors for this study were defined by model calibration in relation to the baseline hydrography over selected global basins.

3.2.2. Preservation of Subdominant Rivers

[12] A second rule governs the preservation of subdominant rivers during spatial upscaling. When there is more than one major river in a cell (e.g., cell [4,6] in Figure 1a), flow direction for the cell is determined by the LEDR of the most dominant river (e.g., R2 in Figure 1a) going through it. This may lead to flow discontinuity of subdominant river segments in subsequent tracing. For example, R4 will be absorbed by R2, in Figure 1a. The DRT addresses this potential error by moving the detector (see definition in Table A1) downstream along the subdominant river (R4) to locate the nearest adjacent cell that has not yet been assigned

a flow direction (e.g., cell [4,5]). If the unassigned downstream cell is an immediate neighbor, the flow direction of the given cell (e.g., cell [3,6]) is assigned in the direction of this downstream cell (i.e., cell [4,5]). Thus, the subdominant river (R4) is preserved and not absorbed by the dominant river (R2) in the cell. If a cell under consideration (e.g., cell [6,7] of R6) has an unassigned downstream cell that is not an immediate neighbor (i.e., cell [8,7]), in order to preserve the subdominant river (R6), the DRT assigns a downstream flow direction by diverting the river passing through the adjacent cell (e.g., dashed arrow in cell [7,7]) parallel along the river defined at fine scale (R6) to the unassigned downstream cell (i.e., cell [8,7]). The DRT preserves such subdominant rivers by this “diverting” algorithm only when the subdominant river drainage area exceeds the threshold FAC value (defined as the “basin area threshold” in Table 5) and there are appropriate cells available for diverting. If there are other important rivers in cells that a river is diverted through (e.g., cell [7,7] in Figure 1a) that cannot be diverted because of a relatively dense river network in the surrounding cells, the DRT allows the specified subdominant river (R6) to be absorbed by the more dominant river (R2). In these cases a multiple flow direction algorithm may be the best way to preserve subdominant rivers.

3.2.3. Preservation of Sinuous Flow Paths

[13] A third rule governs the upscaling of sinuous flow paths in a cell. Situations where a sinuous dominant river flows out of a cell, back into the same cell and out again (e.g., cell [5,1] and cell [7,2] in Figure 1a), may result in the assignment of inverse flow directions to adjacent cells (e.g., cell [5,1] and cell [6,1], and cell [7,2] and cell [7,1] in Figure 1a), resulting in these cells draining into each other or generating a circular flow path. The DRT avoids these potential errors by performing the following tasks. First, the dominant river is traced from the headwater (e.g., R3 and R9) to the outlet or junction of the river basin (subbasin). The algorithm then determines whether the river returns to the current cell (e.g., cell [5,1] in R3 and cell [7,2] in R9). The algorithm then assigns the flow direction along the dominant river according to the first and second rules above. If the dominant river returns and the assigned flow directions result in cells draining into each other (e.g., in R9, cell [7,2] and cell [7,1]) or generating a circular flow path, the DRT moves the detector downstream until the river finally leaves the cell (e.g., river pixels identified (circled) in cell [5,1] and cell [7,2] in Figure 1a), and continues to assign flow directions following the first and second rules above according to how the river finally leaves the cell. Otherwise, the algorithm continues river tracing and flow direction assignments following the above rules (e.g., cell [5,1] and cell [6,1]). The identified upscaled flow directions from this process are illustrated for R3 (brown arrows) and R9 (pink arrows) in Figure 1a. In this example the detour river segment cell [6,1] is preserved for R3, while cell [7,1] is excluded from R9. These general rules enable the DRT algorithms to conserve baseline fine-scale river structure in complex landscapes.

3.3. Four Metrics on Dominant River Tracing Performance

[14] We derived four metrics (equations (1)–(4)) to quantify DRT performance for matching hydrological

modeling needs in preserving dominant rivers and river segments, and assigning predominant drainage directions:

$$f_1 = N_d/N_t, \quad (1)$$

$$f_2 = N_m/N_t, \quad (2)$$

$$f_3 = N_{m50}/N_t, \quad (3)$$

$$f_4 = N_{m80}/N_t, \quad (4)$$

where N_d is the number of cells with flow directions assigned according to the LEDR, which is identified by the DRT to determine the predominant flow direction and drainage from a cell among all river segments (if any) within the cell; N_t is the total number of valid cells defined in the coarse-scale DEM (mask) that are assigned flow directions; N_m is the number of cells draining the most runoff in the assigned flow direction from the eight potential directions; N_{m50} and N_{m80} are the number of cells draining more than 50% and 80%, respectively, of runoff from the cell in the assigned flow direction.

[15] The locally dominant river segment (i.e., LEDR) draining the most runoff of each cell was determined from the baseline fine-scale hydrography. For the evaluation process, a similar flow direction determination algorithm (using the same three general rules) as the DRT was used to define the flow direction of each cell according to the selected dominant river (i.e., LEDR), which is called the evaluation direction. The N_d metric quantifies the number of cells where the evaluation direction (derived by local tracing) is the same as the DRT upscaled results (derived by global tracing of the entire dominant river). The f_1 metric is used to verify whether the DRT algorithm correctly identifies the LEDRs and dominant rivers on a grid cell-by-cell basis. A large f_1 value indicates that the algorithm has successfully identified more river segments of the dominant rivers generally with a higher f_2 . However, correctly assigning flow directions for cells corresponding to the $1 - f_1$ term is critical for successfully preserving all dominant rivers and the overall drainage structure in a study area; the $1 - f_1$ term refers to cells for which flow directions are not defined by the LEDR to preserve the overall drainage structure. With the overall dominant drainage structure well preserved, larger f_2 , f_3 , and f_4 values enable more accurate hydrological modeling.

4. Dominant River Tracing Results and Evaluation

[16] Flow direction is widely recognized as being difficult to project [e.g., Fekete et al., 2001; Olivera et al., 2002; Döll and Lehner, 2002], leading to difficulties in converting upscaled results between different geographic projections and coordinate systems. Flexible upscaling algorithms are therefore required to derive upscaled results from baseline hydrography under different projections or coordinate systems to meet the needs of various applications. GCMs are usually set up in a nonprojected environment, while hydrological models are implemented using various geographic projections or nonprojected environments depending on specific applications. The effect of geographic projection on the

upscaling process is well documented [Olivera et al., 2002]. The DRT is designed to allow the derivation of upscaled results from either projected or nonprojected baseline hydrography inputs, while different methods are used for basin area and river length calculations. In this study we applied the DRT in two projections, including Lambert azimuthal equal area projection and a geographic (latitude/longitude) projection referenced to datum WGS84 (hereafter referred to as Lambert and WGS84 projections, respectively). Because HydroSHEDS currently does not include high-latitude areas, the HYDRO1K still plays a critical role in global applications. For this investigation we use the HYDRO1K hydrography to define baseline 1 km resolution information for global DRT upscaling. The DRT upscaled global hydrography layers were derived at multiple spatial scales, including 1/16°, 1/8°, [1/4]°, [1/2]°, 1°, and 2° resolutions, with respective cell sizes approximating 7.5, 15, 30, 60, 120, and 240 km (referred to as "cell size" in Table 5) in the Lambert projection. The generation of DRT upscaled hydrography in the Lambert projection was facilitated by the baseline HYDRO1K format, which is also in a Lambert projection. To generate the WGS84 data set, we reprojected the original HYDRO1K DEM and stream network into the WGS84 format, and then burned the resulting stream network into the WGS84 DEM format; the WGS84 baseline fine-scale (1/120° approximating 1 km resolution) hydrography (e.g., flow direction, FAC, and stream network) layer was then prepared for DRT upscaling using standard GIS tools (e.g., Arc Hydro).

[17] The DRT results were verified against other upscaled results from the literature over regional (the contiguous U.S.) and global domains as described in sections 4.1 and 4.2. The DRT results (e.g., upscaled flow directions, river networks, basin areas, upstream drainage area of each cell, lengths of stem rivers and tributaries, length of upstream river segments of each cell, basin shapes) were also verified against the baseline fine-scale hydrography inputs using visual comparisons and quantitative methods. The four metrics (discussed in section 2.3) were also used to evaluate DRT performance in meeting the requirements for macroscale hydrological modeling. We derived modeling efficiency M_E , equivalent to the Nash-Sutcliffe coefficient [Nash and Sutcliffe, 1970], root-mean-square error (RMSE) and normalized RMSE (NRMSE), mean absolute relative error (MRE) and absolute relative error (RE) statistics to evaluate correspondence between DRT upscaled basin geometry calculations and the baseline hydrography:

$$M_E = 1.0 - \frac{\sum_{i=1}^N (S_i - O_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2}, \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - S_i)^2}. \quad (6)$$

$$NRMSE = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - S_i)^2}}{O_{\max} - O_{\min}}, \quad (7)$$

$$\text{MRE} = \frac{1}{N} \sum \frac{|O_i - S_i|}{O_i}, \quad (8)$$

$$\text{RE} = \frac{|O_i - S_i|}{O_i}, \quad (9)$$

where O_i is defined from the baseline, fine-scale hydrography; S_i is the same variable derived from the DRT upscaled results; \bar{O} is the average value of O_i ; O_{\max} is the maximum value of O_i ; O_{\min} is the minimum value of O_i , and N is the number of basins, rivers, or individual cells selected for the comparison.

4.1. Dominant River Tracing Comparisons With Other Upscaling Methods Over a Contiguous U.S. Regional Domain

[18] The DRT results were compared with other upscaling methods from the literature, including NSABE (1/12° and 1/2° resolution) and DDM30 (1/2° resolution) over a contiguous U.S. regional domain. All of the methods (except DDM30) were applied to derive flow directions at targeted spatial resolutions based on the same baseline inputs (i.e., HydroSHEDS 15 arc second resolution hydrography) without manual corrections to the results. We also performed a global comparison between the DRT and DDM30 using HYDRO1K baseline fine-scale hydrography inputs as described below. While the DDM30 was primarily generated using HYDRO1K baseline information consistent with the global DRT results, intensive postprocessing and manual corrections (for over 30% of the total cells processed) using vectorized river networks may have steered the DDM30

away from the HYDRO1K baseline. The DDM30 results were included in the contiguous U.S. regional comparison as a reference for the global DRT and DDM30 comparison (described in section 4.2), and to assess model performance against upscaled results derived from intensive manual corrections. To facilitate objective comparisons of upscaled flow directions and river networks from the different methods, the same algorithms for calculating upscaled drainage areas and river lengths defined by the DRT were applied to the other upscaling results. Individual cells were selected from all upscaled results and matched with corresponding HydroSHEDS basin areas and/or river lengths following the rule-based approach of *Döll and Lehner* [2002]. The baseline HydroSHEDS pixels with maximum values of all HydroSHEDS upstream drainage areas within individual cells were selected for the comparison. Only those cells are taken into account, for which (1) the upstream drainage areas (or river lengths) of the HydroSHEDS and corresponding upscaled results differ by less than a factor of 3, and (2) the baseline basin area (or river length) exceeds a minimum size threshold.

[19] The number of cells selected according to the above rules provides an important performance metric because cells not selected by this procedure have large (>threefold) differences between upscaled results and the baseline hydrography. As shown in Table 1, the DRT preserves the most cells and rivers for all selection constraints summarized. For example, at 1/2° resolution, the DRT preserves 84% (464) more cells than the NSABE with basin area and river length discrepancies within a factor of 3 from baseline conditions for respective drainage areas and river lengths above 3600 km² and 50 km. The DRT also preserves a similar number (1019) of cells as the DDM30 (1034). The

Table 1. Comparison of Numbers of Selected Individual Basin and River Grid Cells From NSABE, DDM30, and DRT Upscaling Methods for the Contiguous U.S. Domain^a

	Rivers/Cells	NSABE	DDM30	DRT
<i>1/2°</i>				
Area > 3600; Factor 3 (area)	Cells	3110	3535	3874
Area > 3600; RE (area) < 0.05	Cells	748	1505	1635
Area > 3600; RE (area) < 0.1	Cells	1180	2069	2226
Length > 50; Factor 3 (length)	Cells	3109	4130	4806
Length > 50; Factor 3 (length)	Rivers	1441	1777	2280
Length > 50; RE (length) < 0.05	Cells	966	1513	2304
Length > 50; RE (length) < 0.1	Cells	1195	1897	2862
Factor 3 (Area & Length); Area > 3600 & Length > 50	Cells	555	1034	1019
<i>1/12°</i>				
Area > 100; Factor 3 (area)	Cells	114,016	–	145,868
Area > 2500; Factor 3 (area)	Cells	29,586	–	36,116
250 < Area < 2500; Factor 3 (area)	Cells	51,061	–	67,895
100 < Area < 250; Factor 3 (area)	Cells	33,369	–	41,857
Area > 100; RE (area) < 0.05	Cells	38,711	–	74,046
Area > 100; RE (area) < 0.1	Cells	51,847	–	93,317
Length > 8; Factor 3 (length)	Cells	86,592	–	133,988
Length > 8; Factor 3 (length)	Rivers	27,864	–	43,679
Length > 100; Factor 3 (length)	Cells	19,004	–	25,323
50 < Length < 100; Factor 3 (length)	Cells	12,956	–	19,922
8 < Length < 50; Factor 3 (length)	Cells	54,633	–	88,744
Length > 8; RE (length) < 0.05	Cells	36,188	–	72,383
Length > 8; RE (length) < 0.1	Cells	47,205	–	93,405
Area > 3600 & Length > 8; Factor 3 (Area & Length)	Cells	45,272	–	75,629

^a“Rivers” refer to stem rivers and tributaries defined from headwater cells to junction or outlet cells; area and length units are in km² and km; “factor 3” is the discrepancy (for basin area and river length) between the upscaled results and baseline hydrography within a factor of 3 difference following the rule based approach of *Döll and Lehner* [2002].

DRT derives a greater number of cells with smaller RE (e.g., less than 5% and 10%) for drainage area and river length calculations than the NSABE and DDM30. The DRT also derives 70% (93,405 cells) of the total (133,988) selected cells with RE for river length less than 10%, while the NSABE produces a lower 55% (47,205 cells) of the total (86,592) selected cells at 1/12° resolution.

[20] The performance comparisons based on the selected cells were performed separately between the DRT and each alternative upscaling method as shown in Table 2. In each comparison, metrics were calculated from the total cells, in which each cell was derived by either the DRT or other upscaling method, with discrepancies of upstream drainage area or river length between the upscaled results and baseline hydrography less than a factor of 3. The DRT results show improved performance relative to the other methods in terms of MRE, RMSE, and M_E metrics (Table 2). All of the methods show similar and favorable performance in preserving cells with relatively large upstream drainage areas or river lengths (e.g., at 1/12° resolution), including upstream drainage areas greater than 50,000 km² ($M_E > 0.980$). However, the DRT results show relatively improved performance in preserving smaller basins and rivers. For example, at 1/12° resolution for upstream drainage areas between 250 and 2500 km², the DRT results produce favorable M_E (0.908), RMSE (172 km²), MRE (0.118), and selected cells (67,934) metrics, while the NSABE results for these areas show reduced M_E (0.783), RMSE (265 km²), MRE (0.204) and selected cells (51,084) results.

[21] The f_1 , f_2 , f_3 , and f_4 metrics described above were also calculated for the contiguous U.S. regional comparison as shown in Table 3. The DRT results showed larger f_1 , f_2 , f_3 , and f_4 metrics than the NSABE and similar metrics as the DDM30. The relative lower f_1 for the NSABE results from the NSABE, defining flow direction for each cell using the most dominant river with maximum FAC, while in many cases these criteria result in flow directions

through the corner of the cell with relatively short flow length. These results indicate that the DRT derived river network is accurate for hydrological modeling in which the D8 form single flow direction map is used to delineate routing flow paths.

4.2. Global Comparisons Between Dominant River Tracing and DDM30 Results

[22] The global 1/2° flow direction map, DDM30 [Döll and Lehner, 2002], was selected for global comparison against DRT 1/2° upscaled results derived using similar baseline fine-scale hydrography (HYDRO1K) inputs. Individual cells with upstream drainage areas exceeding 10,000 km² were selected from the DRT upscaled results and matched with corresponding HYDRO1K basin areas following the same rules described in section 4.1. There were 11,421 and 18,101 DDM30 cells selected globally to compare with corresponding DRT results in Lambert and WGS84 projections, respectively. Even without Australia (HYDRO1K river network was not available) the WGS84-based DRT results show 20% more cells selected for the comparison with HYDRO1K than the DDM30, which indicates better DRT performance. The fewer number of selected DRT cells in the Lambert projection relative to WGS84 are because of the different projection geometries. The M_E of the DDM30 is 0.989 for all selected cells, while the DRT upscaled results show higher M_E values of 0.996 and 0.997 under respective WGS84 and Lambert projections. The DDM30 has a M_E of 0.727 for all cells with drainage areas between 10,000 and 100,000 km², while the DRT has a M_E of 0.846 and 0.878 for these areas under respective WGS84 and Lambert projections. The DRT (Lambert) results show more consistent performance across regions than the DDM30, especially for cells with drainage areas between 10,000 and 100,000 km²; the DRT regional M_E results range from 0.840 (Europe) to 0.920 (South America), while the DDM30 based M_E values range from

Table 2. Comparison of Derived Upstream Drainage Areas and River Lengths by NSABE, DDM30, and DRT for the Contiguous U.S. Domain^a

	Selection Constraints	Algorithms	MRE	RMSE	M_E	
Comparison of upstream drainage areas from all selected individual cells with upstream basin area greater than 3600 km ² and 100 km ² at 1/2° and 1/12° resolution, respectively	1/2°	NSABE and DRT	NSABE	0.311	130,447	0.916
			DRT	0.166	62,810	0.978
	1/12°	DDM30 and DRT	DDM30	0.170	66,251	0.975
			DRT	0.166	21,102	0.998
	1/12°	NSABE and DRT	NSABE	0.383	97,893	0.734
			DRT	0.153	20,063	0.988
Comparison of upstream river lengths from all selected individual cells with upstream river length greater than 50 km and 8 km at 1/2° and 1/12° resolution, respectively	1/2°	NSABE and DRT	NSABE	0.445	234.9	0.807
			DRT	0.218	98.2	0.954
	1/12°	DDM30 and DRT	DDM30	0.294	159.9	0.893
			DRT	0.220	127.4	0.927
	1/12°	NSABE and DRT	NSABE	0.411	104.8	0.847
			DRT	0.153	63.6	0.942

^aThe comparisons were performed separately between the DRT and each literature method, with metrics calculated from the total cells selected by all methods where upscaled upstream drainage areas or river lengths were within a factor of 3 discrepancy from the baseline hydrography.

Table 3. Comparison of NSABE (1/2° and 1/12°), DDM30 (1/2°) and DRT (1/2° and 1/12°)^a

	Algorithms	N _t (cells)	f ₁	f ₂	f ₃	f ₄
<i>Regional (Contiguous U.S. Domain) Comparison Based on HydroSHEDS 15-s Hydrography</i>						
1/2°	NSABE	6874	0.637	0.749	0.610	0.437
1/2°	DDM30	6874	0.809	0.922	0.850	0.605
1/2°	DRT	6874	0.875	0.927	0.836	0.604
1/12°	NSABE	232,460	0.711	0.852	0.761	0.561
	DRT	232,460	0.833	0.919	0.863	0.629
<i>Global Statistics Summarizing DRT Performance Based on HYDRO1K Hydrography</i>						
2°	DRT	2399	0.893	0.912	0.791	0.594
1°	DRT	9114	0.884	0.923	0.834	0.616
1/2°	DRT	35,954	0.872	0.922	0.843	0.618
1/4°	DRT	139,525	0.876	0.924	0.863	0.625
1/8°	DRT	547,725	0.873	0.923	0.862	0.633
1/16°	DRT	2181,750	0.911	0.915	0.853	0.646

^aComparison of performance in serving hydrological models in terms of the four metrics ($f_1 = N_d/N_i$; $f_2 = N_m/N_t$; $f_3 = N_m50/N_i$; $f_4 = N_m80/N_i$) using HydroSHEDS 15-s hydrography as baseline for the contiguous U.S. domain. Global statistics summarizing DRT performance in the four metrics based on HYDRO1K hydrography across variable upscaling resolutions.

0.561 (South America) to 0.801 (Europe) (Table 4). The computed DRT (Lambert) M_E for each continent is larger than the corresponding DDM30-based M_E values (Table 4). Regional comparisons of the DRT (WGS84) results were not performed, though the global statistics (Table 4) indicate consistent DRT performance in both projections.

4.3. Visual Evaluation of Upscaled Flow Directions and River Structure

[23] The upscaled flow direction is the most critical variable for determining other hydrography layers, including FAC, upscaled river network, basin boundary and shape, and river length. Visual comparisons between upscaled flow directions (i.e., river networks) and the baseline fine-scale hydrography is the most direct method to evaluate upscaling algorithm performance. Figure 2 shows a 1° resolution DEM

of South America with DRT upscaled (red lines) and baseline HYDRO1K (blue lines) river networks, and DRT upscaled river mouths and sinks (solid circles). The upscaled river network in Figure 2 is defined using a FAC threshold value of 1 (cell) that defines the flow direction of each cell to facilitate visual inspection. Visual inspection of each flow path indicates that the upscaled flow directions track the baseline fine-scale dominant rivers and tributaries defined from HYDRO1K. The results also show that the DRT correctly identifies mouth and sink cells relative to HYDRO1K. The DRT utilizes baseline fine-scale river networks (e.g., defined by FAC thresholds of 1000 pixels from the “stream network definition” in Table 5) for spatial upscaling, while the river networks displayed in Figure 2 are defined using a larger (3000 pixels) FAC threshold for mapping and visual inspection. Similar results were found for other diverse global

Table 4. Upstream Drainage Areas of Individual Cells as Presented by the DRT and DDM30 Compared to the Respective Areas of HYDRO1K

	Drainage Area (km ²)		M _E	NRMSE	MRE	Number of Cells
Globe	>10,000(pixels)	DRT (WGS84)	0.996	0.005	0.106	18,101
	10,000 ~ 100,000(pixels)		0.846	0.098	0.131	12,881
Globe	>10,000	DDM30	0.989			14,840
		DRT_L ^a	0.997	0.004	0.103	11,421
Globe	10,000 ~ 100,000	DDM30	0.727			11,435
		DRT_L	0.878	0.086	0.123	8371
Asia	>10,000	DDM30	0.986			
		DRT_L	0.999	0.005	0.095	3464
Asia	10,000 ~ 100,000	DDM30	0.754			
		DRT_L	0.875	0.087	0.119	2241
Africa	>10,000	DDM30	0.983			
		DRT_L	0.998	0.006	0.107	2937
Africa	10,000 ~ 100,000	DDM30	0.670			
		DRT_L	0.878	0.086	0.127	2243
Europe	>10,000	DDM30	0.980			
		DRT_L	0.986	0.009	0.124	1465
Europe	10,000 ~ 100,000	DDM30	0.801			
		DRT_L	0.840	0.102	0.138	1190
N. America	>10,000	DDM30	0.996			
		DRT_L	0.998	0.005	0.098	1864
N. America	10,000 ~ 100,000	DDM30	0.784			
		DRT_L	0.882	0.085	0.114	1443
S. America	>10,000	DDM30	0.994			
		DRT_L	0.996	0.007	0.095	1691
S. America	10,000 ~ 100,000	DDM30	0.561			
		DRT_L	0.920	0.067	0.118	1254

^aDRT under Lambert projection.

regions indicating generally consistent and favorable DRT performance in preserving baseline fine-scale river structure.

4.4. Evaluation of Dominant River Tracing-Defined Basin Areas Relative to Baseline Hydrography

[24] In hydrological models (e.g., the VIC model), accurate determination of the area fraction of cells in basin boundary areas that actually contribute runoff to a given basin outlet is critical for correct identification of upscaled basin areas and accurate hydrograph simulations. The DRT first identifies all basin cells contributing to the given outlet cell (or any individual cell); a buffer (1-cell) area is

then defined outside the basin boundary and applied to calculate basin area. The area fraction of all cells within the basin boundary and buffer area is determined by tracing the fine-scale flow path while only counting pixels that actually contribute to the outlet cell. The area of each basin cell is calculated directly at coarse scale. Two different methods were used to calculate the surface drainage area for the two projections. For the Lambert projection, the area of each cell is calculated as the sum of each pixel area within the cell, because they have the same area. For the WGS84 projection, the area of each cell is calculated using the sphere surface distance and area formulas (Appendix B).

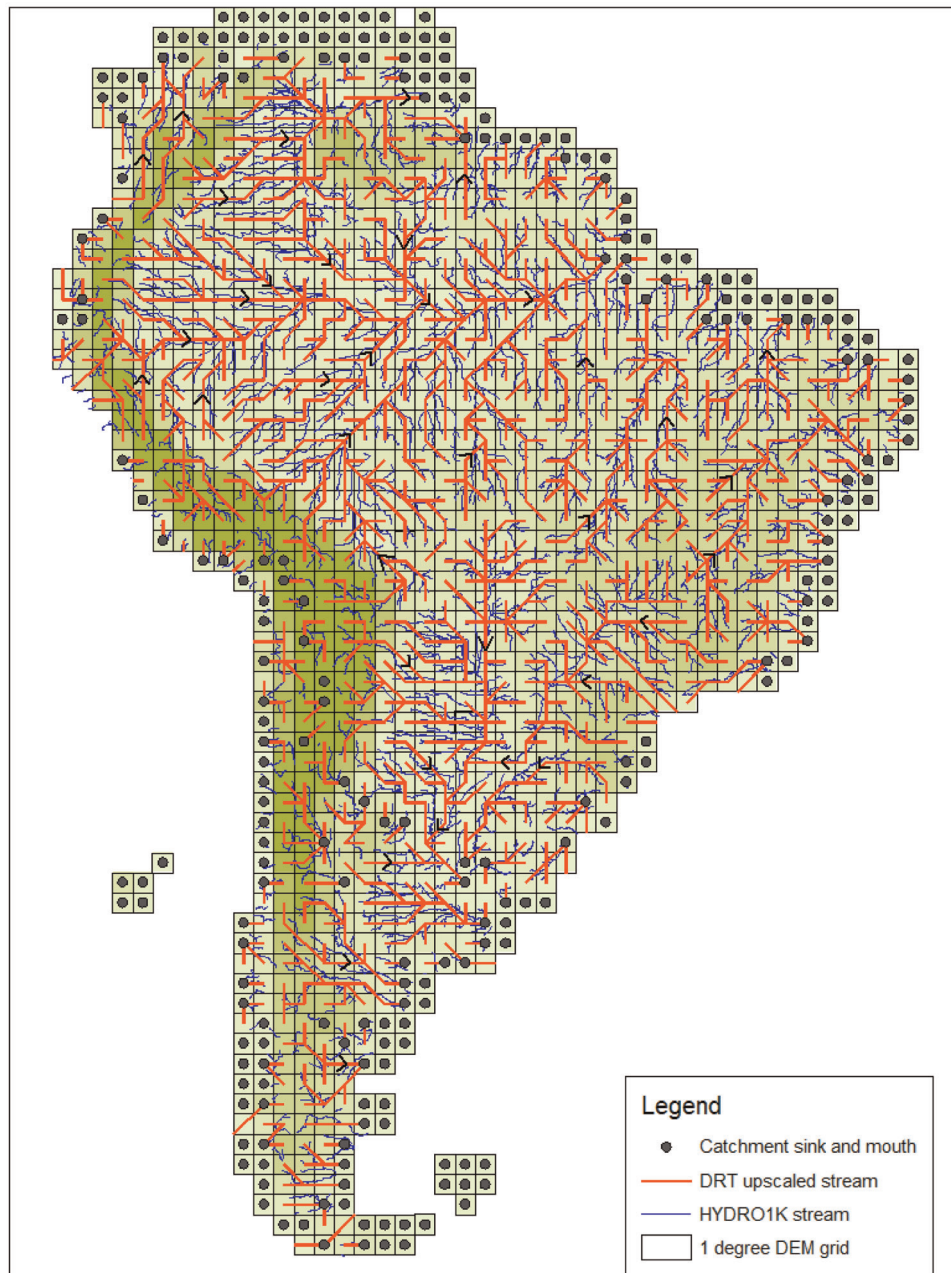


Figure 2. DRT derived flow directions and river network (red line) for South America at 1° resolution. The baseline river network (blue line) is derived from the HYDRO1K using a FAC threshold of 3000 pixels, while the DRT derived river network is defined using a FAC threshold of 1000 pixels. The 1° resolution DEM is also shown ranging from high (green) to low (light green) elevation.

Table 5. Threshold Values Used in the Upscaling Process at Different DRT Spatial Scales

	2°	1°	1/2°	1/4°	1/8°	1/16°
Cell size ^a	240	120	60	30	15	7.5
Cell area ^b	57,600	14,400	3600	900	225	56
Basin area threshold ^{b,c}	60,000	15,000	4000	1000	300	50
Stream network definition ^{d,e}	1000	1000	50	50	20	5
Basin area evaluation ^b	10,000	5000	1000	500	100	50
River length evaluation ^a	100	50	50	20	15	10

^aUnits in kilometers.

^bArea in kilometer².

^cThe threshold value of basin area which are used by the DRT to preserve basins with a larger area than the threshold by calling the “diverting” functions.

^dThe threshold values of FAC are used to define the fine-scale stream network which are employed in the DRT.

^eUnits in pixels.

[25] We conducted a global comparison of basin areas derived from the DRT upscaled results and the baseline HYDRO1K inputs (Table 6). We used the same rules from *Döll and Lehner [2002]* to select basins (individual cells) for comparison throughout this paper, with minimum basin size thresholds (see “basin area evaluation” in Table 5) of 10,000, 5000, 1000, 500, 100, and 50 km² for respective

spatial scales of 2°, 1°, [1/2]°, [1/4]°, 1/8°, and 1/16°. Hereafter, all evaluations were performed on DRT results in the Lambert projection because the extraction of baseline information (e.g., fine-scale upstream area, river length) for comparison is more convenient using the equal-area (relative to WGS84) projection format. The number of selected global basins ranged from 456 (2° resolution) to 24,374 (1/16° resolution). The resulting minimum basin size meeting the above rule criteria is only 1 cell for each upscaling level, though the DRT can derive smaller catchments with drainage areas less than the size of an individual cell.

[26] The DRT results indicate that basin areas are preserved across all spatial scales relative to the fine-scale (HYDRO1K) hydrography inputs for all upscaling scales ($R^2 = 1.00$; $p < 0.0001$). The NRMSE differences between the DRT upscaled results and baseline hydrography ranged from 0.02% (1/16° resolution) to 0.11% (2° resolution), while MRE differences ranged from 6.34% (1/16° resolution) to 1.18% (2° resolution). Both NRMSE and MRE terms vary subtly across all upscaling levels (Table 6). The relatively larger residual difference at 1/16° resolution is attributed to the very small size (i.e., 50 km²) of catchments selected for the evaluation. For a given scale, the MRE is smaller for larger basins because basin boundaries tend to

Table 6. Global Comparison of HYDRO1K Versus DRT Derived Basin Area, Lengths of Stem Rivers and Major Tributaries, and Basin Shapes^a

	2°	1°	1/2°	1/4°	1/8°	1/16°
<i>Global Comparison of HYDRO1K Versus DRT Derived Basin Areas</i>						
Basins with variable sizes						
Basin size	>10,000 km ²	>5,000 km ²	>1,000 km ²	>500 km ²	>100 km ²	>50 km ²
Number of basins	456	1016	2719	5372	13,298	24,374
NRMSE	0.11%	0.09%	0.03%	0.03%	0.02%	0.02%
MRE	1.18%	2.19%	1.75%	2.30%	2.20%	6.34%
ME	1.000	1.000	1.000	1.000	1.000	1.000
Basins with drainage area between 5000 and 50,000 km ²						
Number of basins	304	752	1098	1169	1165	1181
NRMSE	1.02%	1.47%	1.21%	1.28%	1.12%	1.01%
MRE	0.38%	0.81%	1.02%	0.92%	1.00%	1.14%
ME	0.998	0.996	0.997	0.997	0.997	0.998
<i>Global Comparison of HYDRO1K Versus DRT Derived River Lengths</i>						
Rivers (major tributaries) with variable lengths						
River length	>100 km	>50 km	>50 km	>20 km	>15 km	>10 km
Number of rivers	1036	3202	9834	37,907	103,382	174,265
NRMSE	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%
MRE	0.24%	0.45%	0.85%	2.04%	3.80%	5.91%
ME	1.000	1.000	1.000	1.000	1.000	1.000
Rivers (major tributaries) with length between 20 ~ 200 km						
Number of rivers	247	1485	8692	34,684	82,529	92,633
NRMSE	0.47%	0.54%	0.64%	0.69%	0.71%	0.71%
MRE	0.54%	0.79%	1.25%	2.20%	3.30%	3.50%
ME	0.999	0.999	0.999	0.999	0.999	0.998
<i>Global Comparison of HYDRO1K Versus DRT Derived Basin Shapes</i>						
Basins with drainage area greater than 1000 km ²						
Number of basins	604	1422	2618	3799	4136	4192
MRE	2.44%	2.41%	3.46%	4.63%	3.67%	2.43%
ME	0.881	0.928	0.891	0.866	0.914	0.990

^aRivers are all major tributaries.

be preserved to a greater degree for larger basin polygons than smaller polygons where the size of single grid cell approaches the size of the basin polygon [Michael, 1999]. At 1/16° resolution, a bias of 1 pixel (i.e., 1 km²) represents 2% of a catchment with a drainage area of 50 km². From the global statistics, 7756 of the 24,374 selected basins (32%) are smaller than 100 km² at 1/16° resolution and show a MRE of 9.12%. In contrast, 4177 basins are larger than 1000 km², account for 17% of the total selected basins and have a MRE of 2.06%. Further comparisons between DRT upscaled and HYDRO1K basin areas between 5000 and 50,000 km² across all scales show the largest MRE (1.14%) at 1/16° resolution, and the lowest MRE (0.38%) at 2° resolution (Table 6). The difference between the largest and lowest MRE is smaller than 1%.

4.5. Validation of Dominant River Tracing-Derived River Lengths

[27] Instead of using a meandering factor to estimate flow distance within each cell [e.g., Oki and Sud, 1998; Fekete et al., 2001; Olivera and Raina, 2003], the DRT preserves river lengths by tracing each fine-scale river from headwater pixel to river mouth pixel. As the entire river tracing progresses, the river length between the two centroids of adjacent upstream and downstream cells is calculated and saved for each cell with flow direction assigned by the river. The centroid (e.g., dark point (pixel) a–d in Figure 1a) of a cell is defined as the middle pixel of the dominant river segment (on the river being traced) of the cell, while headwater and outlet pixels are designated as the centroids for respective headwater and outlet cells.

[28] We conducted global comparisons of the DRT-derived river lengths not only for basin stem rivers, but also for major tributaries in each selected basin and across all scales relative to the baseline HYDRO1K inputs. To extract lengths of rivers and major tributaries for comparison, the same rules from the basin area selections were applied while using different threshold values for selecting rivers with lengths greater than 100, 50, 20, 15, and 10 km (see “river length evaluation” in Table 5) for respective 2°, 1°, [1/2]°, [1/4]°, 1/8°, and 1/16° spatial scales. The number of selected stem rivers and tributaries ranged from 1036 (2° resolution) to 174,265 (1/16° resolution). The global comparison (Table 6) indicates that the total lengths of rivers are well preserved across all spatial scales relative to the fine-scale (HYDRO1K) hydrography ($R^2 = 1.00$; $p < 0.0001$), with the same NRMSE difference of 0.02% for all upscaling levels, while MRE differences range from 0.24% to 5.91%. Similar to basin area, the relatively larger residual difference of river lengths at 1/16° resolution is because of the large number of very short (i.e., 10 km) rivers selected for the evaluation. At 1/16° resolution, a bias of 1 (1 km) pixel represents 10% of a 10 km length river. From the global statistics, 83,180 of the 174,265 selected rivers (48%) are between 10 and 20 km in length at 1/16° resolution and show a MRE of 8.83%. In contrast, 10,716 rivers are longer than 100 km, account for 6% of the total selected rivers, and have a MRE of 0.76%. Further comparisons between DRT and HYDRO1K river lengths between 20 and 200 km indicate consistent DRT performance across all spatial scales in this size category, with the largest MRE (3.50%) at 1/16° resolution from 92,633 rivers selected, and the lowest MRE (0.54%) at 2° resolution from 247 rivers selected (Table 6).

The difference between the largest and lowest MRE is less than 3%. The MRE difference is smaller for larger rivers with lengths between 100 and 500 km, ranging from 0.25% at 2° resolution from 576 rivers selected to 0.80% at 1/16° resolution from 10,093 rivers selected, with differences of less than 1% between the largest and lowest MRE values across all spatial scales in this size category. These results (Table 6) indicate that the residual differences in basin area and river length calculations between the DRT and baseline hydrography are very small at all upscaling levels, while model performance is largely independent of spatial scale.

4.6. Verification of Dominant River Tracing-Derived Basin Shapes

[29] Six global basins were selected for visual inspection of DRT-derived basin shapes against baseline HYDRO1K results. These results are illustrated in Figure 3 for a range of global basins including the Amazon, Yangtze, Danube, Hai, Salmon, and a randomly selected smaller basin (i.e., Timbo river basin, Liberia). The selected basins ranged in area from 9112 to 5,880,712 km² at all upscaling levels. The DRT shows generally favorable performance in preserving basin shape (relative to the fine-scale hydrography inputs) for different spatial scales and basin sizes. As spatial resolution of the DRT derived hydrography increases, the upscaled basin shape becomes smoother along boundary edges and closer to the original fine-scale basin shape. These results reflect the limitations of the raster model grid to represent basin edges at coarser spatial scales, rather than any change in DRT algorithm performance.

[30] We applied a shape metric [Vörösmarty et al., 2000] to quantitatively evaluate DRT performance in preserving basin shapes, as

$$S = L/\sqrt{A}, \quad (10)$$

where S is the basin shape metric, L is the stem river length [kilometer], and A is the basin area [km²]. Basin shape indices were calculated for all DRT spatial scales for the selected basins and compared with baseline HYDRO1K results. These results (Table 6) indicate favorable DRT performance in preserving basin shapes for basins with drainage areas greater than 1000 km² for all spatial scales, with MRE differences ranging from 2.41% at 1° resolution to 4.63% at 1/4° resolution. On a global basis across all spatial scales, the MRE of basin shapes is >20% for 4.3%, <10% for 95.1%, and <5% for 93.7% of basins with drainage areas greater than 1000 km².

4.7. Validation of Dominant River Tracing Internal Drainage Structure of a Specific River Basin

[31] Macroscale hydrological modeling studies generally require more detailed measures of drainage area and river length within a basin. The correct derivation of total basin area, stem river length, and basin shape does not necessarily mean that the river network is properly represented within the basin. Accurate depictions of basin internal drainage structure, including upstream area, river segments of each cell along the river flow path and local terrain characteristics, are essential for determining runoff lag time, which is critical for many model processes, including streamflow and stream temperature simulations. Döll and Lehner [2002]

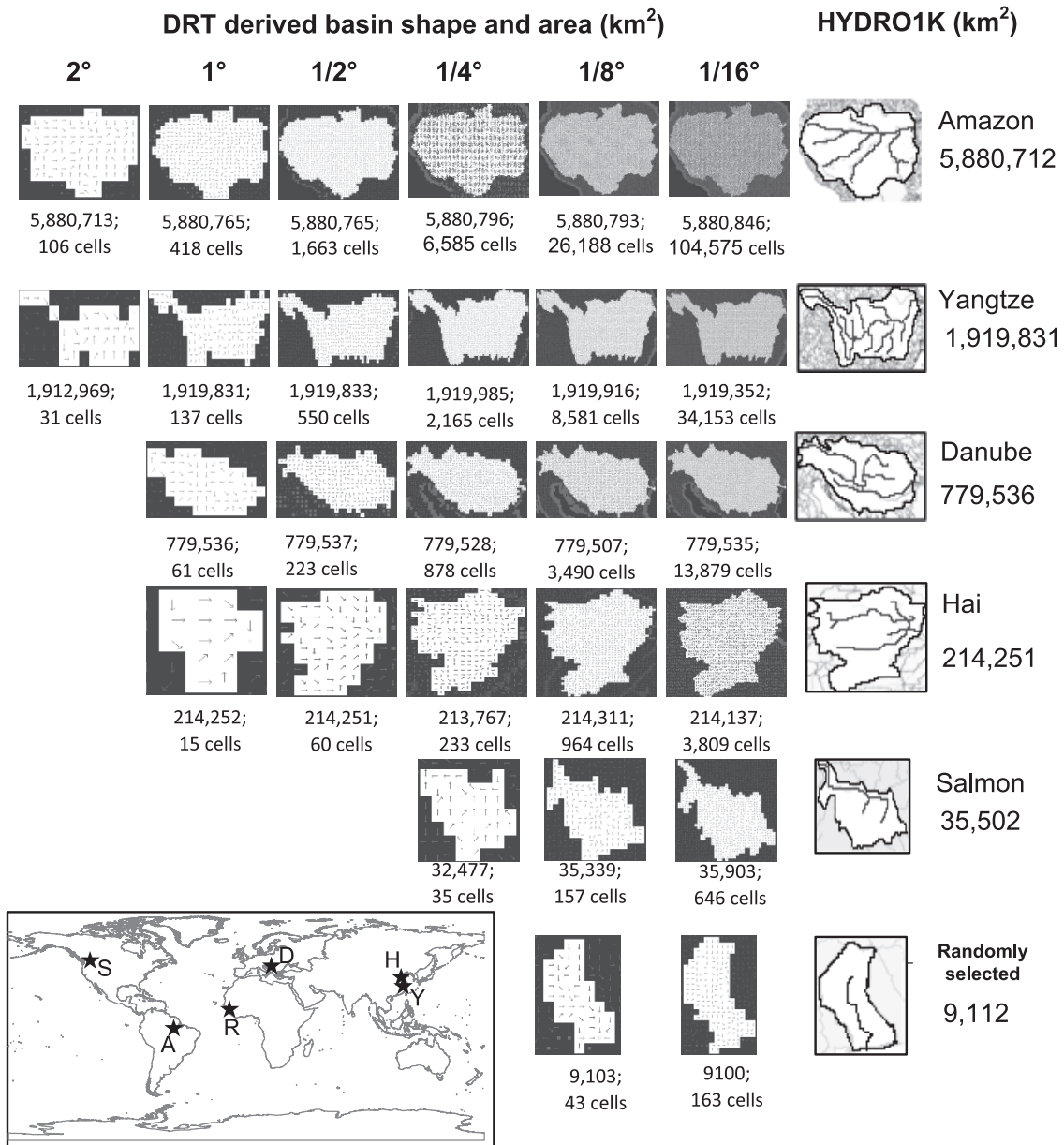


Figure 3. Comparison of DRT upscaled basin shapes in relation to the HYDRO1K defined baseline for a range of selected spatial scales and basin sizes. The HYDRO1K basins were delineated in ESRI ArcGIS software by merging subbasins according to the fine-scale river networks. Each DRT derived basin shape figure shows the respective area (km²) and the number of cells in the basin at the corresponding upscaled resolution. The numbers along the right side of the figure are the basin areas defined from baseline HYDRO1K information. The inset map shows the geographic locations of the selected basins denoted by the first letter of each basin name.

used the upstream areas of all cells along the stem river to represent the internal river network of a basin. We applied a similar approach to assess DRT preservation of the internal river network of a selected basin (i.e., Columbia basin, North America) at 1/16° resolution. We compared the DRT derived upstream drainage area of all cells along Columbia basin rivers relative to coincident area calculations derived from the baseline (HYDRO1K) hydrography. We also conducted a more intensive validation of the DRT derived internal river network relative to HYDRO1K for the Columbia basin by comparing both basin stem river cells and all indi-

vidual basin cells with drainage areas greater than 500 km² (~10 cells defined at 1/16° resolution). River lengths from headwater cell to each cell along the river flow path, and the corresponding upstream drainage area of each cell were compared between the DRT results and HYDRO1K, as well as the shape of the upstream drainage area of each cell.

[32] There were 2374 cells (with drainage areas greater than 500 km²) selected for evaluating upstream drainage areas, river lengths and basin shapes from the DRT (1/16°) results and HYDRO1K (Figure 4) for the Columbia basin. The DRT generally preserves fine-scale upstream drainage

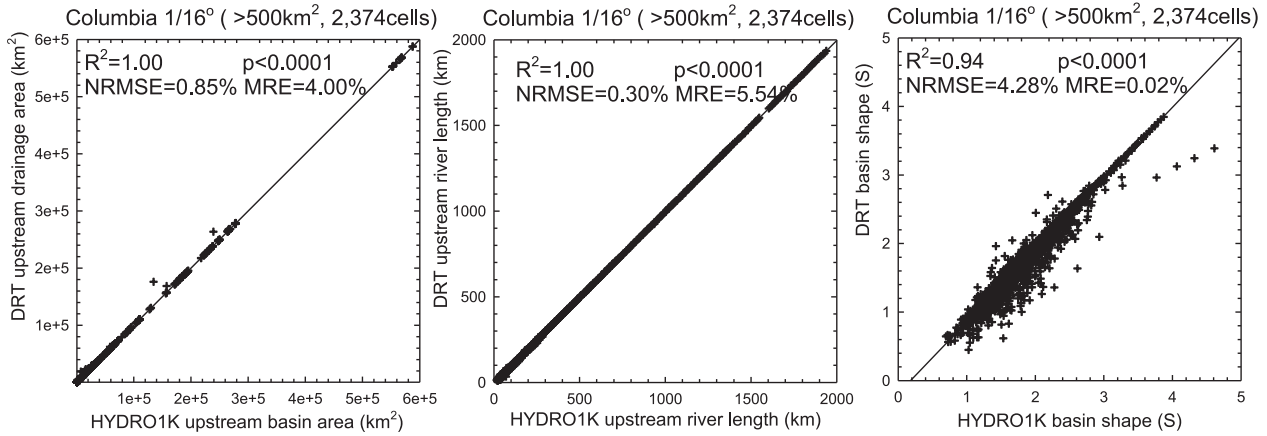


Figure 4. Comparison of HYDRO1K versus DRT derived 1/16° resolution upstream drainage areas, upstream river length, and upstream basin shapes for the Columbia river basin of North America. The upstream river length is defined as the river length from headwater to each cell along the defined river flow paths.

areas, river lengths and basin shapes for all selected cells, as indicated by strong agreement with the HYDRO1K results (Figure 4). The NRMSE and MRE were 1.38% and 1.52%, respectively, between DRT and HYDRO1K upstream drainage areas for individual cells (912) with drainage areas larger than 5000 km². All cells with upstream river lengths greater than 100, 200 and 500 km show respective NRMSE differences of 0.28%, 0.27% and 0.32%, and MRE differences of 1.61%, 0.89%, and 0.43%. Approximately 83% and 91% of selected cells had MRE differences for upstream drainage

areas within 5% and 10% of the baseline fine-scale results, respectively; while 72% and 93% of cells, with upstream river lengths larger than 50 km, had MRE differences for upstream river lengths within 5% and 10% of HYDRO1K results; 98% of cells with upstream river lengths larger than 100 km had MRE differences within 10% of the baseline fine-scale results.

[33] The upstream area-river length relationship of each DRT derived 1/16° cell along the Columbia stem river and three main tributaries (i.e., Snake, upper Columbia, and

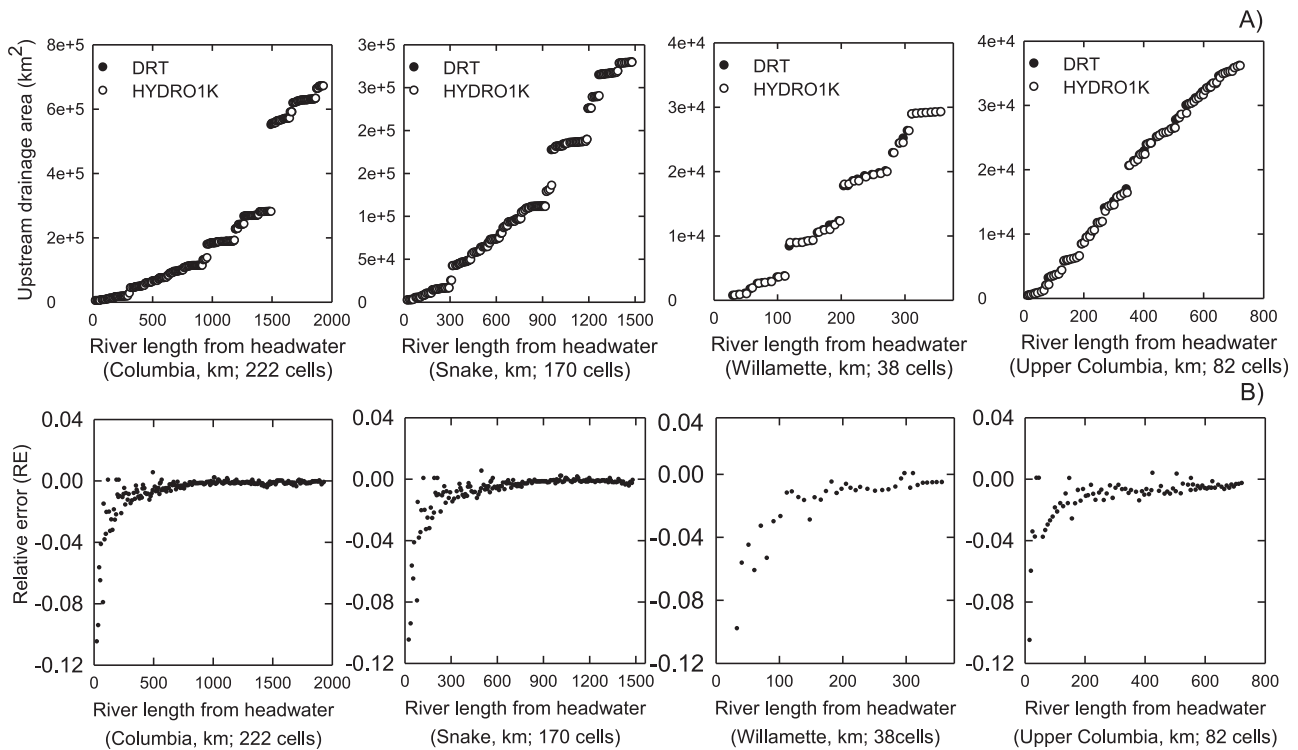


Figure 5. (a) Comparison of HYDRO1K versus DRT derived upstream area and corresponding river length relationships of DRT 1/16° cells along the Columbia basin stem river and its three main tributaries, i.e., Snake, upper Columbia, and Willamette rivers; (b) the relative errors of estimated river lengths of cells along the river flow paths.

Willamette rivers) also showed favorable performance relative to HYDRO1K indicating a consistency in river segments along basin flow paths (Figure 5a). In Figure 5a, the numbers of respective upstream river cells for the Columbia, Snake, Willamette, and upper Columbia rivers are 222, 170, 38, and 82, corresponding to 1935, 1484, 355, and 723 km of DRT river lengths; the MRE values of all river segment lengths of the four rivers are 0.67%, 0.82%, 1.84%, and 0.74%, respectively. These results indicate that the DRT preserves the upstream area-river length relationship defined from the baseline hydrography (Figure 5a). Abrupt increases in upstream drainage area with increasing river length (Figure 5a) occur where subbasin drainage areas merge in downstream portions of the basin, while river lengths increase conservatively. These results also indicate that river confluences and the basin internal drainage structure are preserved during spatial upscaling. Relative errors in river length calculations generally decrease toward the river mouths of all four rivers (Figure 5b), indicating that errors primarily occur in upstream areas with decreasing impacts on longer downstream river lengths; these results also explain why relative error tends to decrease for longer rivers. Overall, the DRT preserves the internal river network of a basin, including upstream drainage area, river length and basin shape of each grid cell.

4.8. Dominant River Tracing Performance in Serving Hydrological Models

[34] As shown in Table 3, the values of f_1 conservatively range from 0.873 to 0.911 for all upscaling levels based on the global statistics. Considering that the LEDR is not always used to determine flow direction in order to preserve the globally dominant river structure and the $1 - f_1$ term must be greater than 0, the DRT derived f_1 value should be very close to its upper limit.

[35] The summary statistics in Table 3 indicate favorable DRT performance, with more than 91% of cells draining the most runoff in the assigned flow directions at all upscaling levels. As more major rivers tend to occupy a given cell at coarser spatial scales, the proportion of total runoff from the cell draining through the assigned flow direction tends to be smaller. The f_3 and f_4 values therefore decrease conservatively with increasing upscaling levels from 0.853 and 0.646 at $1/16^\circ$ resolution to 0.791 and 0.594 at 2° resolution, respectively. The small ranges of f_2 , f_3 , and f_4 indicate consistent DRT performance in assigning flow directions in which the majority of water drains from all cells regardless of spatial scale. The upper limit of f_3 and f_4 is f_2 . From Table 3, using the DRT, f_1 is close to f_2 , which indicates that the assigned drainage direction generally matches the path of the dominant river segment in a cell, which validates our assumption that the LEDR collects the most runoff within a cell. For all upscaling levels, f_2 is slightly greater or equivalent to f_1 , which is probably caused by secondary river segments and flow paths draining in the assigned flow direction leading to some $1 - f_1$ cells having the most runoff directions. The DRT derived f_1 , f_2 , f_3 , and f_4 metrics in Table 3 indicate that the DRT identifies the dominant rivers in a region and preserves the dominant rivers and segments with appropriate flow directions largely independent of the upscaling level; this is a prerequisite for preserving basin shape and area, river shape, and length during spatial upscal-

ing. These results therefore indicate favorable DRT performance in meeting the requirements of hydrological models using the D8 form single flow direction map.

5. Discussion and Conclusions

[36] A hierarchical dominant river tracing (DRT) algorithm was developed for automated extraction and spatial upscaling of river network and flow directions from relatively fine-scale hydrography inputs. The DRT algorithm uses multidimensional (i.e., two-, one-, and zero-dimension) information (from both global and local drainage patterns) defined from baseline fine-scale hydrography to determine upscaled flow directions. This approach contrasts with many traditional methods that utilize zero- or quasi one-dimensional information for spatial upscaling. The DRT also uses additional constraints to minimize the occurrence of excluded or false river segments. River channels are prioritized by FACs and the effective lengths of dominant rivers. By fully exploiting baseline fine-scale hydrography information, the algorithm maintains consistency in basin shape and area calculations by minimizing absorption of smaller basins by larger basins at coarser spatial scales while generally preserving fine-scale river and segment lengths in the upscaled outputs. The DRT was applied using baseline (HYDRO1K) fine-scale hydrography inputs to produce a range of upscaled global river networks ranging from $1/16^\circ$ to 2° resolution in Lambert azimuthal equal area and WGS84 projections. A regional comparison was also conducted over a continental U.S. domain between the DRT results and other upscaling results derived from alternative (NSABE) methods using consistent baseline (HydroSHEDS) fine-scale hydrography inputs; these results were also compared against regional DDM30 results derived from HYDRO1K baseline information and intensive manual corrections. The regional comparisons show improved DRT upscaling performance in preserving baseline fine-scale river network information. These results also show robust DRT performance relative to the baseline hydrography, including (1) assigning as many upscaled flow directions as possible to drain a majority of runoff to immediate downstream cells, while preserving the overall dominant drainage structure of the fine-scale hydrography; (2) preserving river shape and length, basin shape and area, and internal drainage structure for all considerable river basins of a given region; and (3) deriving upscaled flow directions and river networks automatically and with globally consistent performance across different scales and projections. Both visual and quantitative evaluations indicate that the DRT is a robust and accurate method for automated upscaling and extraction of coarse flow directions and river networks from fine-scale hydrography information. The DRT results showed similar or better performance than the DDM30 results derived from intensive manual correction, while the DRT is fully automated and capable of producing accurate results without manual intervention and correction. The DRT upscaling process also generates other products useful for hydrological modeling, including flow distance, upstream drainage area, channel gradient, and fractional area of basin boundary cells.

[37] The DRT algorithm results are consistent with and directly traceable to the baseline fine-scale hydrography inputs. The upscaled flow direction and river network results

approach the accuracy limits of the D8 single direction flow method in representing coarse-scale flow directions and river networks, while the DRT river length and basin area calculations are generally consistent with the baseline fine-scale hydrography inputs, and are appropriate for most GCM- and MHM-based modeling studies. In some cases where there is no clearly defined single dominant flow path, a multiple flow direction (MFD) approach may provide a better solution than the D8 single direction method by allowing multiple flow directions for each grid cell. However, when the MFD method is employed in a hydrological model, fractional runoff calculations for all flow directions of each grid cell are required, which can be difficult to quantify and may require significant hydrological model adaptations. *Orlandini and Moretti* [2009] also reported that the MFD introduces artificial dispersion in large-scale hydrological applications. Alternatively, finer resolution D8 hydrography and hydrological model simulations can be used for improved spatial resolution and accuracy, but with increasing computational costs. With appropriate modifications, the DRT can be adapted to the MFD approach for further investigation. Recently, *Yamazaki et al.* [2009] derived an upscaled global river network map by dominant river segment tracing. Their method is designed for greater flexibility in assigning downstream cells during spatial upscaling, which may be more appropriately regarded as a special type of MFD.

[38] The upscaled results from this study were developed using fine-scale global (regional) hydrography inputs from HYDRO1K (HydroSHEDS), while the DRT algorithm can be applied using any baseline hydrography information. In some cases, inability of the DEM to represent fine-scale terrain characteristics or errors introduced during the depitting

process leads to incorrect delineation of the fine-scale river network [*O'Donnell et al.*, 1999], and coarse river network errors that are independent of the upscaling process. The HYDRO1K database has well documented limitations over lowlands and other areas with low topographic relief [e.g., *Fekete et al.*, 2001; *Mayorga et al.*, 2005]. Some methods including D8-LTD [*Orlandini et al.*, 2003], “stream burning,” and “fencing” have been used to improve the accuracy of fine-scale hydrography information, which could be available for many regional applications. Improved baseline hydrography inputs should enable greater accuracy in DRT upscaled river networks for any region or spatial scale to meet the needs of regional and macroscale hydrological modeling studies.

Appendix A

[39] The DRT algorithm flowchart is shown in Figure A1. The meanings of expressions in Figure A1 are explained in Table A1. The DRT inputs include a relatively fine-scale river network, FAC, and flow direction information, and a mask file at the targeted coarse spatial resolution. The algorithm implementation consists of six general steps, which are summarized below.

[40] Step 1: Assign sink and river mouth cells at the beginning of the algorithm according to the “basin area threshold” in Table 5. All cells with river mouth pixel(s) are first identified; if the largest FAC value of the river mouth pixels in the cell is greater than the threshold, the cell is identified as a river mouth or sink cell representing a basin that will be conserved in the upscaling process.

[41] Step 2: Identify the outlet of the predominant basin from the remaining unassigned area of the given region of

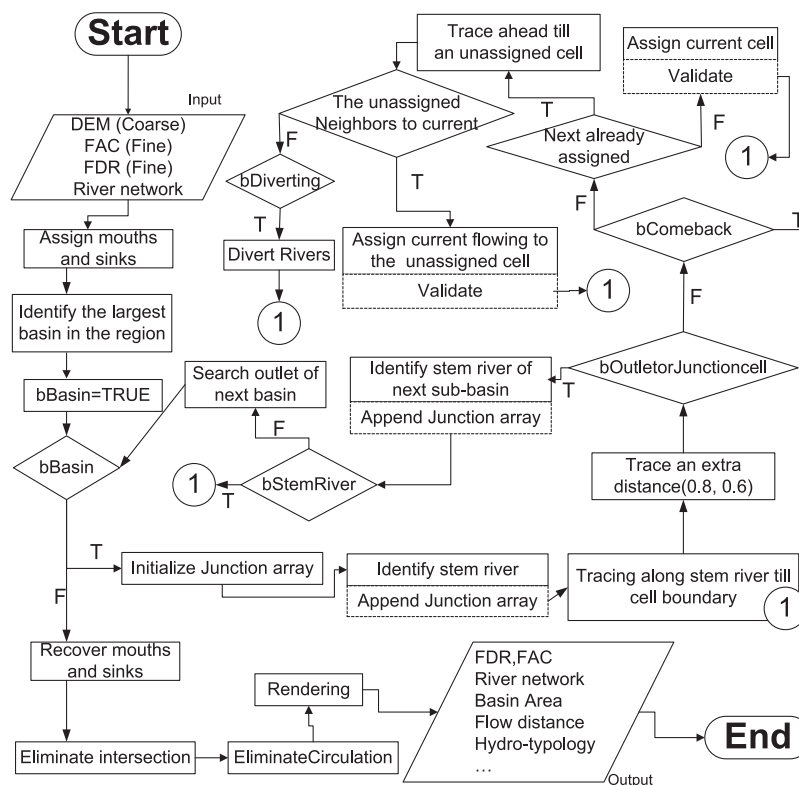


Figure A1. DRT algorithm flowchart.

Table A1. Meaning of the Expressions in Figure A1

Expressions	Meanings
T	TRUE
F	FALSE
Detector	A detector is defined as a pixel where the tracing is progressing when the DRT traces each river or flow path at the fine scale. The detector records the location (i.e., row and column number) of each pixel during the tracing process, e.g., where the DRT tracing starts or stops.
bBasin	If bBasin is TRUE, it means there is a basin which has not been assigned yet and the DRT will continue the assigning process; if bBasin is FALSE, it means all basins in the given domain have been assigned.
bStemRiver	If bStemRiver is TRUE, it means stem river of a basin or sub-basin hasn't been assigned yet and the DRT will continue the assigning process; if the bStemRiver is FALSE, it means all rivers/tributaries in a basin has been assigned and the DRT will identify a next basin to assign flow directions.
bDiverting	If bDiverting is TRUE, it means the current river should be conserved by diverting; if bDiverting is FALSE, the Diverting function will not be called.
bComeback	If bComeback is TRUE, it means a river flows out of a cell, back into the same cell and out again.
bOutletorJunction	If bOutletorJunctioncell is TRUE, it means the detector reaches an outlet cell or Junction cell.

interest. The basin outlet is identified as the pixel with the largest FAC value.

[42] Step 3: Identify the predominant (stem) river for the basin with the outlet obtained in step 2.

A1. Identify the Stem River (Reverse Tracing)

[43] Given a basin or sub-basin with a dendritic river network, the stem river is the one with the larger FAC. To identify the stem river on a cell-by-cell basis, the DRT conducts reverse tracing along the river path starting from the basin outlet and proceeding to the headwater. During reverse tracing, once the detector (Table A1) reaches a junction pixel, the DRT always traces ahead along the tributary with the larger FAC until the detector reaches the upstream end of the river (i.e., the headwater pixel without any upstream river pixels).

[44] Step 4: Assign flow directions for all intersecting cells of the stem river by tracing its entire length beginning from the headwater to the outlet.

A2. Validate

[45] The DRT algorithm verifies the flow direction for each cell once it is assigned to ensure that each cell is assigned to the LEDR. If a cell is not assigned to the LEDR, the assigned flow direction will be canceled and the cell will be assigned at a later step when the cancellation does not lead to discontinuity of the selected river.

[46] Step 5: Identify the next dominant river of the remaining unassigned area in the current basin (i.e., identify the stem river of the next subbasin). All junctions on all traced dominant rivers are identified and appended to a junction array. The tributary with the maximum FAC is selected as the next dominant river according to the junction array. This process (beginning with step 4) is continued until there are no dominant rivers with unsigned cells in the current basin.

[47] Step 6: Go to step 2 until there are no basins with unassigned cells in the given region.

[48] Step 7: Recover river mouth and sinks. In many cases, especially at coarser spatial resolutions, multiple large rivers (with FAC value of the river outlet pixel greater than the basin area threshold in Table 5) may end in the same cell, which can lead to incorrect merging of a larger basin. The DRT assigns the outlet cell to the largest river,

which ends in the cell, and recovers the outlet cells for secondary rivers by reverse tracing the secondary rivers to their immediate upstream cells, which are assigned as outlet cells of the secondary rivers.

[49] Step 8: The DRT checks each flow path to ensure that there are no circular flow paths in the derived flow network. If a circular flow path occurs, the DRT algorithm will break the circulation path by assigning the cell with the largest FAC as a sink cell with direction "0." An intersection occurs where flow directions of two neighboring cells meet in the corners of the two cells. When trying to preserve diagonal flow directions, intersections of flow paths can occur in some situations. Where intersections occur, the direction of one cell is altered to assure consistent D8 flow directions.

Appendix B

[50] The surface area of a cell in a geographic (WGS84) projection is calculated as the sum of the two surface triangles that make up the rectangular grid cell. The area of each triangle (S_{Δ}) is calculated using the sphere surface area formulas in (B1) and (B2)

$$S_{\Delta} = (A + B + C - \pi) * R^2 \text{ if } A + B + C \geq \pi, \quad (\text{B1})$$

$$S_{\Delta} = \sqrt{P * (P - a) * (P - b) * (P - c)} \text{ if } A + B + C < \pi, \quad (\text{B2})$$

where R is the earth radius [m]; A, B, C are the interior angles of the surface triangle in radians, calculated from the law of cosines

$$\begin{aligned} \sin b \sin c \cos A &= \cos a - \cos b \cos c \\ \sin c \sin a \cos B &= \cos b - \cos c \cos a \\ \sin a \sin b \cos C &= \cos c - \cos a \cos b \end{aligned} \quad (\text{B3})$$

a, b, c are the lengths of each edge of the surface triangle in radians, calculated by

$$\begin{aligned} a &= L_a/R \\ b &= L_b/R \\ c &= L_c/R \end{aligned} \quad (\text{B4})$$

L_a , L_b , and L_c are the lengths of each edge of the surface triangle [m], calculated by

$$L = R * \arccos(\sin \alpha_1 \sin \alpha_2 + \cos \alpha_1 \cos \alpha_2 * \cos(\beta_1 - \beta_2)). \quad (\text{B5})$$

α , β are the latitude and longitude of a surface point location and P is the half circumference of the surface triangle

$$P = (L_a + L_b + L_c)/2. \quad (\text{B6})$$

[51] Equation (B5) is also used to calculate river lengths from the fine-scale hydrography.

[52] **Acknowledgments.** This work was conducted at the University of MT (UMT) with financial support from the Gordon and Betty Moore Foundation. The authors would like to thank Balázs M. Fekete, Bernhard Lehner, and Stefano Orlandini for their constructive comments. The DRT upscaled global hydrography data sets generated from this study are available through the UMT online data archives at <ftp://ftp.ntsg.umd.edu/pub/data/DRT/>.

References

- Arora, V. K. and G. J. Boer (1999), A variable velocity flow routing algorithm for GCMs, *J. Geophys. Res.*, *104*, 30,965–30,979.
- Arora, V. K., and S. Harrison (2007), Upscaling river network for use in climate models, *Geophys. Res. Lett.*, *34*, L21407, doi:10.1029/2007GL031865.
- Arora, V. K., F. Seglenieks, N. Kouwen, and E. Soulis (2001), Scaling aspects of river flow routing, *Hydrol. Processes*, *15*, 461–477.
- Comanor, K., B. Nijssen, and D. P. Lettenmaier (2000), Comment on “Five-minute, 1/2°, and 1° data sets of continental watersheds and river network for use in regional and global hydrologic and climate modeling studies” by Graham et al., *Water Resour. Res.*, *36*, 3117–3120.
- Costa, M. H., and J. A. Foley (1997), The water balance of the Amazon basin: Dependence on vegetation cover and canopy conductance, *J. Geophys. Res.*, *102*, 973–989.
- Dirmeyer, P. A. (1995), An objective river routing scheme for runoff models, *Proc. Conf. on Hydrology*, Dallas, Tex., Am. Meteorol. Soc. 91–92.
- Döll, P., and B. Lehner (2002), Validation of a new global 30-min drainage direction map, *J. Hydrol.*, *258*, 214–231.
- Fekete, B. M., C. J. Vörösmarty, and R. B. Lammers (2001), Scaling gridded river network for macro-scale hydrology: development and analysis, *Water Resour. Res.*, *37*, 1955–1967.
- Gesch, D. B., K. L. Verdin, and S. K. Greenlee (1999), New land surface digital elevation model covers the Earth, *Eos, Trans., AGU*, *80*, 69–70.
- Graham, S. T., J. S. Famiglietti, and D. R. Maidment (1999), Five-minute, 1/2°, and 1° data sets of continental watersheds and river network for use in regional and global hydrologic and climate system modeling studies, *Water Resour. Res.*, *35*, 583–587.
- Jenson, S. K., and J. O. Domingue (1988), Extracting topographic structure from digital elevation data for geographic information system analysis, *Photogrammic Eng. Remote Sens.*, *11*, 1593–1600.
- Lehner, B., K. Verdin, and A. Jarvis (2008), New global hydrography derived from spaceborne elevation data, *Eos Trans., AGU*, *89*, 93–94.
- Liston, G. E., Y. C. Sud, and E. F. Wood (1994), Evaluating GCM land surface hydrology parameterizations by computing river discharges using a runoff routing model: Application to the Mississippi Basin, *J. Appl. Meteorol.*, *33*, 394–405.
- Maidment, D. R. (1996), GIS and hydrological modeling: An assessment of progress, *Third International Conference on GIS and Environmental Modeling*, Santa Fe, New Mexico, 20–25 January.
- Marengo, J. A., J. R. Miller, G. L. Russell, C. E. Rosenzweig, and F. Abrampoulos (1994), Calculations of river-runoff in the GISS GCM: Impact of a new land-surface parameterization and runoff routing model on the hydrology of the Amazon River, *Clim. Dyn.*, *10*, 349–361.
- Mayorga, E., M. G. Logsdon, M. Victoria, and R. Ballester (2005), Estimating cell-to-cell land surface drainage paths from digital channel networks, with an application to the Amazon basin, *J. Hydrol.*, *315*, 167–182.
- Michael, Z. (1999), *Modeling Our World: The ESRI Guide to Geodatabase Design*, 197 pp., ESRI Press, Redlands, Calif.
- Miller, J. R., G. L. Russell, and G. Caliri (1994), Continental-scale river flow in climate models, *J. Clim.*, *7*, 914–928.
- Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, *J. Hydrol.*, *10*, 282–290.
- Nijssen, B., D. P. Lettenmaier, X. Liang, S. W. Wetzel, and E. F. Wood (1997), Stream flow simulation for continental-scale river basins, *Water Resour. Res.*, *33*, 711–724.
- O'Donnell, G., B. Nijssen and D. P. Lettenmaier (1999), Simple algorithm for generating streamriver network for grid-based, macroscale hydrological models, *Hydrol. Processes*, *13*, 1269–1275.
- Oki, T. and Y. C. Sud (1998), Design of total runoff integrating pathways (TRIP)—A global river channel network, *Earth Interact.*, *2*, 1–37.
- Olivera, F. and R. Raina (2003), Development of large scale gridded river network from vector stream data, *J. Am. Water Resour. Assoc.*, *39*, 1235.
- Olivera, F., J. S. Famiglietti, and K. Asante (2000), Global-scale flow routing using a source-to-sink algorithm, *Water Resour. Res.*, *36*, 2197–2207.
- Olivera, F., M. S. Lear, J. S. Famiglietti, and K. Asante (2002), Extracting low-scale river network from high-scale digital elevation models, *Water Resour. Res.*, *38*(11), 1231, doi:10.1029/2001WR000726.
- Orlandini, S., and G. Moretti (2009), Determination of surface flow paths from gridded elevation data, *Water Resour. Res.*, *45*, W03417, doi:10.1029/2008WR007099.
- Orlandini, S., G. Moretti, M. Franchini, B. Aldighieri, and B. Testa (2003), Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevation models, *Water Resour. Res.*, *39*(6), 1144, doi:10.1029/2002WR001639.
- Renssen, H., and J. Knoop (2000), A global river routing network for use in hydrological modeling, *J. Hydrol.*, *230*, 230–243.
- Shaw, D., L. W. Martz, and A. Pietroniro (2005), Flow routing in large-scale models using vector addition, *J. Hydrol.*, *307*, 38–47.
- Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe (1991), The analysis of river basins and channel network using digital terrain data, *Hydrol. Processes*, *5*, 81–100.
- U.S. Geological Survey (2000), HYDRO1K Elevation Derivative Database, Cent. for Earth Resour. Obs. and Sci., Sioux Falls, S.D. (available at <http://edc.usgs.gov/products/elevation/gtopo30/hydro>).
- Vörösmarty, C. J., III, B. Moore, M. Gildea, B. Peterson, J. Melillo, D. Kicklighter, J. Raich, E. Rastetter, and P. Steudler (1989), A continental-scale model of water balance and fluvial transport: Application to South America, *Global Biogeochem. Cycles*, *3*, 241–265.
- Vörösmarty, C. J., B. M. Fekete, M. Meybeck, and R. B. Lammers (2000), Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution, *Hydrol. Processes*, *237*, 17–39.
- Yamazaki, D., T. Oki, and S. Kanae (2009), Deriving a global river network map and its sub-grid topographic characteristics from a fine-resolution flow direction map, *Hydrol. Earth Syst. Sci.*, *13*, 2241–2251.

J. S. Kimball, J. Stanford, and H. Wu, Flathead Lake Biological Station, Division of Biological Sciences, University of Montana, 32125 Bio Station Ln., Polson, MT 59860, USA. (huanwu@umd.edu).

N. Mantua, Climate Impacts Group, Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA 98195-5672, USA.