A hydrography upscaling method for scale invariant parametrization of distributed hydrological models

Dirk Eilander[1,2], Willem van Verseveld[2], Dai Yamazaki[3], Albrecht Weerts[2,4], Hessel C. Winsemius[2,5], Philip J. Ward[1]

[1]Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
5   [2]Deltares, Delft, The Netherlands
[3]Institute of Industrial Sciences, the University of Tokyo, Tokyo, Japan
[4]Hydrology and Quantitative Water Management Group, Wageningen University & Research, Wageningen, The Netherlands
[5]Dar Es Salaam Resilience Academy, Dar Es Salaam, Tanzania

Correspondence to: Dirk Eilander (dirk.eilander@deltares.nl)

10   **Abstract.** Distributed hydrological models rely on hydrography data such as flow direction, river length, slope and width. For large-scale applications, many of these models still rely on a few flow-direction datasets, which are often manually derived. We propose the Iterative Hydrography Upscaling (IHU) method to upscale high-resolution flow direction data to the typically coarser resolutions of distributed hydrological models. The IHU aims to preserve the upstream-downstream relationship of river structure, including basin boundaries, river meanders and confluences, in the D8 format, which is commonly used to 15   describe river networks in models. Additionally, it derives representative sub-grid river length and slope parameters, which are required for resolution-independent model results. We derived the multi-resolution MERIT Hydro IHU dataset at resolutions of 30 arcsec (~1km), 5 arcmin (~10 km) and 15 arcmin (~30 km) by applying IHU to the recently published 3 arcsec MERIT Hydro data. Results indicate improved accuracy of IHU at all resolutions studied compared to other often applied upscaling methods. Furthermore, we show that MERIT Hydro IHU minimizes the errors made in timing and magnitude 20   of simulated peak discharge throughout the Rhine basin compared to simulations at the native data resolutions. As the method is open source and fully automated, it can be applied to other high-resolution hydrography datasets to increase the accuracy and enhance the uptake of new datasets in distributed hydrological models in the future.

1 Introduction

Large-scale distributed hydrological and land surface models are used to provide estimates of available water resources 25   (Schewe et al., 2014; Wada et al., 2014), flood risk (Hirabayashi et al., 2013; Ward et al., 2013), drought risk (Veldkamp et al., 2017; Wanders et al., 2015) and food production (Kummu et al., 2014), amongst other applications. These models contain a routing module to simulate streamflow, i.e.: the lateral flow of water on the land surface. This is a key variable for understanding the water, energy and biogeochemical cycles and the effects of disturbances from anthropogenic climate change on these cycles (Wood et al., 2011).

30   The spatial pattern of average streamflow conditions is largely determined by the contributing area of a river segment (Quinn et al., 1991) which is imposed on a distributed model by its flow direction data, i.e.: a rasterized representation of the river network. Simulated peak streamflow is particularly sensitive to the accuracy of the flow directions and river channel and floodplain parametrization (Paiva et al., 2011; Zhao et al., 2017) and very important at river confluences (Geertsema et al., 2018; Guse et al., 2020; Metin et al., 2020) and river outlets (Couasnon et al., 2020; Eilander et al., 2020), where multiple

35   fluvial and/or coastal flood drivers may combine to modulate a flood event. Furthermore, streamflow is the only measurable integral signal of basin response and is therefore widely used for model calibration (Beven, 2012; Bouaziz et al., 2020), underlining the importance of flow direction data in distributed hydrological models.

 Over the last decade, large-scale distributed hydrological models have been applied at increasingly higher resolutions (Bierkens, 2015), which poses a challenge on the parametrization of these models to achieve similar model results independent

40   of spatial resolution (Samaniego et al., 2017; Wood et al., 2011). One particular challenge in this regard is the lack of adequate methods to scale hydrography data such as flow directions and river length and slope parameters (Imhoff et al., 2020). Furthermore, integrated models, such as hydro-ecological models (Lowe et al., 2006) or coupled hydrological-hydrodynamic flood models (Hoch et al., 2019), often require the representation of various processes at different spatial resolutions. This can be achieved by hydrological nesting of models but requires multi-resolution hydrography data for the seamless coupling of

45   flow directions at the model domain boundaries.

Flow direction datasets are developed at a fixed, high as possible, resolution, and typically described in the "deterministic eight neighbors" (D8) format, which sets the downstream direction of each cell to one of its eight neighboring cells. Well-known high-resolution (≤ 1 km) flow direction datasets include the 30 arcsec resolution hydro1k (U.S. Geological Survey, 2000) and the 3 arcsec resolution HydroSHEDS (Lehner et al., 2008) and MERIT hydro (Yamazaki et al., 2019) datasets. These datasets

50   are derived from hydrologically conditioned high-resolution elevation data based on the direction with the steepest slope (e.g. Lehner et al., 2008; Yamazaki et al., 2019). Unlike typical geospatial data, such as elevation, hydrography data cannot easily be scaled by spatial resampling techniques and extensive data processing is required to change its spatial resolution (Lehner and Grill, 2013). Therefore, specialized automated upscaling methods to describe high resolution flow directions and river parameters at coarser model resolutions (typically ≥ 1 km) are required to leverage these datasets for distributed hydrological

55   modelling and to achieve seamless integrated multi-resolution modelling.

Several D8 flow direction upscaling methods have been developed (Döll and Lehner, 2002; Fekete et al., 2001; Olivera et al., 2002; Wu et al., 2011; Yamazaki et al., 2008), but none provides a fully automated open-source method that can easily be applied on high resolution flow direction data. Most of these methods first determine which river segment to represent within each coarse-resolution cell and subsequently set the upscaled flow direction based on fine-resolution flow directions of that

60   river segment. However, to determine which river to represent within a cell in order to preserve the river network often requires more information than contained in just one cell and its direct neighbors. For instance, the commonly used DDM30 dataset, which is used by most global hydrological models within the Inter-Sectoral Impact Model Intercomparison Project (www.isimip.org), requires manual corrections after an automated initial procedure to ensure the river network is well

preserved (Döll and Lehner, 2002). To circumvent this problem the Flexible Location of Waterways (FLOW) method (Yamazaki et al., 2009) uses a format that allows a downstream cell to be located outside the eight direct neighbors. While effective, this format has not been used outside the CaMa-Flood model (Yamazaki et al., 2011) for which it was developed as most models are built to work with D8 data. The hierarchical dominant river tracing (DRT) algorithm (Wu et al., 2011) uses global information to order streams within a basin to determine which river segment to represent within a cell and reroutes rivers through neighboring cells if required. While DTR has proven successful at automatically upscaling 30 arcsec flow direction data to coarser resolutions (Wu et al., 2012), its application might be limited at higher resolutions as it requires entire basins to be loaded in memory. Furthermore, none of the D8 upscaling methods derives both river length and slope sub-grid parameters, which are required by most hydrological models.

The first objective of this paper is therefore to develop a fully automated D8 flow direction upscaling algorithm in order to derive flow direction and representative river length and slope parameters that can be applied to high-resolution (< 1 km) global hydrography data. The second objective is to evaluate how the choice of upscaling method and resolution affect peak discharge simulation. The remainder of this paper is set up as follows: Section 2 describes the newly developed Iterative Hydrography Upscaling (IHU) method and the accompanying multi-resolution MERIT hydro IHU dataset; Section 3 describes the benchmark and case study experiments; Section 4 presents the results of the benchmark of IHU at the global scale; Section 5 presents the results of a case-study in which we test the resolution independence of simulated peak discharge; The results are discussed in section 6 and conclusions based on this study are presented in section 7.

2 The Iterative Hydrography Upscaling algorithm

Any flow direction upscaling method needs to determine which river segment to represent within each coarse-resolution cell. The Iterative Hydrography Upscaling (IHU) method balances between traditional upscaling methods that only use local information contained in one coarse-resolution cell and its direct neighbors (Döll and Lehner, 2002; Fekete et al., 2001; Olivera et al., 2002), and upscaling methods that use global information about the hierarchy of streams to determine which river to represent within each coarse-resolution cell (Wu et al., 2011). IHU makes a first estimate of the representative river for coarse-resolution cells, but updates this estimate where it leads to errors in upstream-downstream relations between cell based on contextual information in an iterative process. This iterative approach, which makes use of contextual data, makes IHU effective and suitable to be applied to high resolution hydrography data. Section 2.1 provides a step-by-step description of IHU, section 2.2 describes how representative sub-grid river parameters are derived, and section 2.3 discusses the upscaled *MERIT Hydro IHU v1* dataset, which is released as part of this paper.

2.1 Flow direction upscaling

The IHU method is explained in this section and illustrated for a fictional river in Figure 1, where the used terminology is explained in the legend. For convenience, we will henceforth refer to the target coarse-resolution raster cells as *cells* and fine-

95 resolution raster cells as *pixels*. IHU requires an output cell resolution (grey dashed grid lines), which is a multiple of the input resolution, and two input maps: a fine-resolution flow direction and upstream area map (blue lines, where darker blue indicates a larger river). The goal of the upscaling method is to define the most representative flow direction for each cell (arrows). The IHU method exists of four iterations which all consist of three steps. The iterations are numbered and aimed at: (row 1) initiating upscaled flow directions; (row 2) fixing erroneous flow directions; (row 3) optimizing the in-between outlet pixel

100 distance; and (row 4) minimizing the error made by erroneous flow directions. The steps of each iteration are lettered: (column A) initiate, (column B) analyze and (column C) update. Each step is explained in detail below and refers to a panel of Figure 1. Iteration 2-4 can be repeated to improve the results.

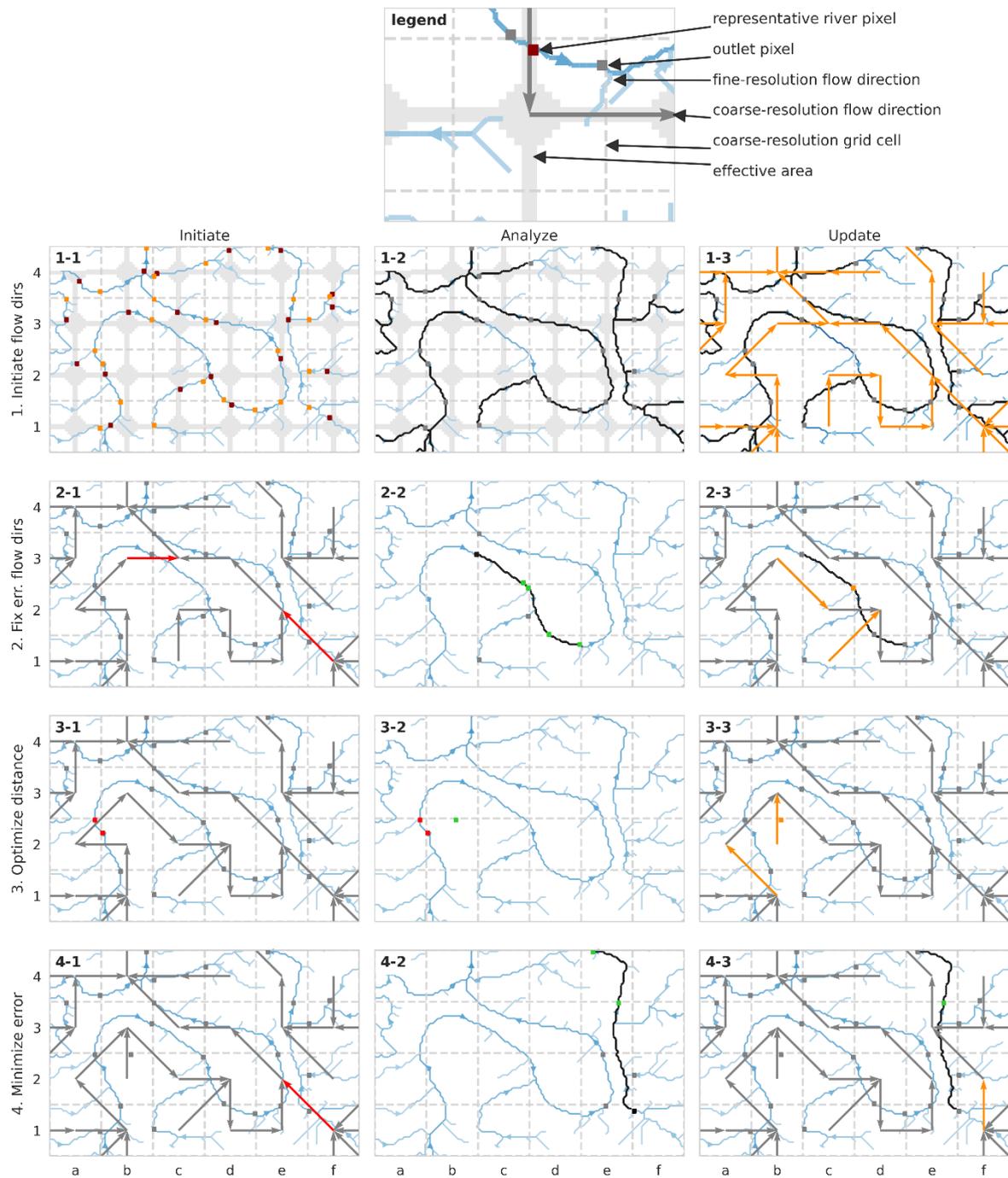



Figure 1: Illustration of the Iterative Hydrography Upscaling (IHU) method. Firstly, 1A) the representative river pixel (dark red square) inside the effective area (shaded area) and the outlet pixel (orange square) are identified for each cell based on upstream area, then 1B) the fine-resolution flow path downstream of the outlet pixel (black lines) is traced to a neighboring cell, 1C) to set the initial upscaled flow direction (orange arrow). Secondly, 2A) erroneous flow-directions (red arrows) are identified, and 2B) analyzed in context of the fine-resolution downstream flow-path (black line) with alternative outlet pixels (green squares) and tributary outlet pixels (grey squares), to 2C) fix the flow directions by relocating outlet pixels (orange square and arrows). Thirdly, 3A) outlet pixels

110 **with short in-between distance are identified (red squares), and 3B) alternative outlet pixels (green squares) with sufficient in-between distance are searched, after which 3C) outlet pixels are relocated and flow directions are updated accordingly (orange square and arrows). Fourthly, 4A) remaining erroneous flow directions are identified (red arrows), and 4B) from each neighboring cell the distance to a common downstream outlet pixel (green square) is measured, to 4C) update the flow directions (orange arrow) to the neighboring cell with the minimum distance in order to minimize upscaling errors.**

115 **Step 1-1:** The first iteration sets an initial flow direction for each cell. First, a representative river pixel is found for each cell (dark red square). This pixel is defined as the river pixel with the largest upstream area within the effective area (grey shade), as described by equation 1. Then, that pixel is traced downstream towards the outlet pixel (orange square), which is set as the most downstream pixel before leaving the cell. This first step of IHU builds on EAM as it uses the same starting point to identify an initial representative river pixel per cell. By defining the effective area for selecting the representative river pixel

120 in each cell, the method avoids selecting river segments that only pass through a corner of a cell and are unfavorable to determine flow directions (Paz et al., 2006).

$$\{(x,y) \,|\, (|x - x_0|^{0.5} + |y - y_0|^{0.5}) < R^{0.5}\}, \tag{1}$$

where, x, y are the coordinates of a pixel, x_0, y_0 are the coordinates of the center of a cell and R is half the cell length.

Step 1-2: The outlet pixel of each cell (grey square) is traced downstream (black lines) until an outlet pixel in a neighboring

125 downstream cell is found. If no outlet pixel is found before leaving the eight neighboring cells, the trace is ended at the first pixel inside the effective area downstream of the outlet pixel, which is the default EAM procedure, see trace downstream of the outlet pixel of *cell b3* to *cell c3* in the example.

Step 1-3: The initial upscaled flow direction (orange arrows) is set for each cell in the direction of the cell where the trace in step 1-2 ends.

130 **Step 2-1**: The second IHU iteration aims to conserve fine-resolution flow directions between outlet pixels at the coarse resolution, by repairing erroneous flow directions. The flow direction of a cell is erroneous if the first outlet pixel downstream is not located in its downstream cell (i.e. where the flow direction points to). In this step erroneous flow directions (red arrows) are identified. In the example erroneous flow directions are found in *cell b3* as the next downstream outlet pixel is found in *cell d2* and not *cell c3* and in *cell f1* as the next downstream outlet pixel is found *cell e3* instead of *cell e2*. Step 2-2 and 2-3

135 are then executed for each cell with erroneous flow direction, sorted from cells with a small to large upstream area at the outlet pixel, and iterated until no more flow direction can be corrected.

Step 2-2 (illustrated for *cell b3* only): The outlet pixel of a cell with erroneous flow direction (black square) is traced downstream (black line) while potential alternative outlet pixels (green squares) are identified: these are defined as the last pixel before entering a new cell on the trace. The trace ends at the next downstream outlet pixel of a cell with correct flow

140 direction and with only one potential outlet pixel. Cells directly upstream of cells with (alternative) outlet pixels on the trace are marked as tributary cells and their outlet pixels as tributary outlet pixels (grey squares). For all tributary cells the first and last alternative outlet pixel to which a valid flow direction can be set are identified. The erroneous flow directions are then updated based on the following iterative procedure:

Starting from the most upstream outlet pixel on the trace, an outlet pixel is relocated to the most downstream alternative outlet pixel in a neighboring cell for which flow directions from the upstream and all tributary outlet pixels can be set correctly. If required to set the flow directions of tributary cells correctly, a tributary outlet pixel of a headwater cell (i.e. cells without upstream neighbors) can be relocated to an alternative outlet pixel. This is repeated until the end of the trace is reached.

If at some point no valid alternative outlet pixel is found, the position of the last relocated outlet pixel is flagged as a bottleneck and not considered in the next iteration. Note that there are no bottlenecks in the example.

This step is iterated until successful or no new bottlenecks are found.

Step 2-3: If step 2-2 is successful the flow directions are updated accordingly. In this example the outlet pixel of *cell c2* is relocated (from black to orange square), thereby changing the flow direction for *cell b3* and *cell c1* (orange arrows). The first outlet pixel downstream of *cell b3* is now located in *cell c2* where the outlet pixel is relocated to the main stream. The first outlet pixel downstream of *cell c1* is located in *cell d2*. Note that the flow direction of *cell f1* cannot be repaired because two rivers flow parallel through its downstream *cell e2* of which only one can be represented at the output resolution.

Step 3-1: The third iteration aims to optimize the distance in-between outlet pixels, measured along the fine-resolution flow directions. If this distance is short, one of the outlet pixels can potentially be removed in favor of another river segment within the same cell. A short in-between outlet pixel distance is not favorable when this distance is used to set the river segment length in routing models as it will decrease the accuracy or require smaller timesteps. In this step, outlet pixels with an in-between outlet pixel distance smaller than a threshold value are flagged (red squares). This threshold is set to 25% of the length of a cell edge resulting in flagged outlet pixels for *cell a2* and *cell b2* in the example. Then, steps 3-2 and 3-3 are executed for every cell with a flagged outlet pixel until no more outlet pixels are relocated.

Step 3-2: First, it is checked whether a flagged outlet pixel can be removed while the flow directions of its upstream neighboring cells can be set correctly. Then, alternative outlet pixels within the same cell are identified (green square). Alternative outlet pixels should have a minimum upstream area; a minimum distance to the next outlet pixel and may not be located downstream of any other outlet pixel. The minimum upstream area threshold is set to 25% of the cell area to avoid creating cells with small sub-grid area. In the example an alternative outlet pixel is found in *cell a2*.

Step 3-3: If one or more alternative outlet pixels are found for a cell in step 3B, the outlet pixel is relocated to the alternative outlet pixel with the largest upstream area, and the flow directions are updated accordingly. In the example the outlet pixel of *cell a2* is relocated (orange square), thereby changing the flow direction for *cell a2* and *cell b2* (see orange arrows). The first outlet pixel downstream of *cell a2* is now located in *cell b2* because *cell a2* now represents another stream. The first outlet pixel downstream of *cell b2* is now located in *cell b3* as the original downstream outlet pixel in *cell a2* is relocated.

Step 4-1: This iteration aims to minimize upscaling errors where erroneous flow directions cannot be repaired. This occurs mostly where multiple rivers flow parallel through the same cell while one can be represented in the D8 format. First, cells with remaining erroneous flow directions after step 2 are identified (red arrows). Then, step 4-2 and 4-3 are executed for each identified cell, sorted from cells with a large to small upstream area at the outlet pixel. In the example, the flow direction of *cell f1* is erroneous as two rivers flow parallel through its downstream *cell e2*.

Step 4-2: The fine-resolution path downstream of a cell with erroneous flow direction is traced (black line) and outlet pixels on the trace are identified (green squares). For each neighboring cell the coarse-resolution flow-direction is followed until it reaches an outlet pixel on the trace. The distance to this outlet pixel from the neighboring cell and to this outlet pixel from the cell with erroneous flow directions are measured in number of outlet pixels and summed. This yields a combined distance to a common downstream outlet pixel for each neighboring cell. If multiple neighboring cells have the same combined distance, the cell with the largest upstream are at the outlet pixel is selected as downstream pixel. Finally, if setting the flow direction to a neighboring cell yields the flow directions from two adjacent cell to cross, this cell is not considered. In the example, combined distance from the pixel of *cell f1* and neighboring cells *cell e1, f2* and *e2* to a common downstream outlet pixel are calculated.

If no downstream outlet pixel is found on the trace and the last pixel of the trace is at a river mouth or sink, that pixel is set as outlet pixel in the cell with erroneous flow direction if within 2 cells distance. Note that in this case the outlet pixel is located outside the cell where it belongs to. If iteration 2-4 are repeated, this step is only executed in the last repeat. This situation occurs if a larger river flows through the cell with the river mouth or sink pixel. This step preserves secondary rivers in cells with larger rivers or multiple river outlets or sinks.

Step 4-3: The flow direction is updated (orange arrows) to the neighboring cell with the shortest combined distance to a common downstream outlet pixel (green square). In the example the shortest combined distance from *cell f1* to the common downstream outlet pixel in *cell e3* is found in *cell f2*, changing the flow direction of *cell f1* to *cell f2* (north). While this introduces a small error in *cell f2*, the error is contained to just that cell minimizing the upscaling error.

The sensitivity of the *R* parameter to define the effective area in step 1-1 as well as the minimum length and minimum upstream area thresholds used to optimize sub-grid river length in step 3 are tested for the river Rhine basin, see appendix E. As step 2-4 are iterated, the minimum river length and upstream area thresholds may also affect the upscaling accuracy as it may provide room for improvements in the next iteration of step 2. We found that the thresholds change the accuracy of the upscaled maps at less than one percent of the output basin cells see Figure E1. A lower minimum upstream threshold generally has a positive effect on the upscaling accuracy and number of cells with small river length but increases the number of cells with small upstream area. The selected thresholds provide a balance between accuracy and cells with small river length or contributing area, but if the latter is of less importance the minimum upstream area threshold might thus be lowered for improved accuracy.

2.2 Sub-grid hydrography variables

Based on fine-resolution flow direction and so-called "outlet pixels", several sub-grid variables are derived as shown in Figure 2. The outlet pixel is the most downstream pixel of the representative river within each cell, see previous section, and used as a link between the coarse- and fine resolution data.

- The *sub-grid area* is defined by the pixels draining to the outlet pixel of a cell and is confined by upstream outlet pixels, see Figure 2B. This area is also referred to as the unit-catchment area as introduced by Yamazaki et al (2009).

- The *river length* is defined by the fine-resolution flow path found by tracing the outlet pixel of a cell either up- or downstream until the next outlet pixel, see Figure 2C-D. When tracing a pixel upstream, the upstream pixel with the largest upstream area is selected in case of multiple upstream pixels. The length is calculated along the sub-grid flow path based on the pixel size, with diagonal steps are taken to be $\sqrt{2}$ times the pixel size. Both up- and downstream river lengths are used in routing models and derived here.

- The *river slope* is based on the MERIT Hydro elevation difference between two pixels at a set distance up- and downstream of the outlet pixel. Here we used a default distance of 2 km, 1 km up- and downstream of the outlet pixel. The flow path along which the slope is derived is shown in Figure 2E.

- The *river width* is based on the MERIT Hydro width data layer at the outlet pixel. Note that this data contains gaps, i.e. not all outlet pixels have a river width in the underlying data, see Figure 2F which shows river widths in green colors. For global coverage and application in hydrological models these gaps need to be filled, which is outside the scope of this paper.

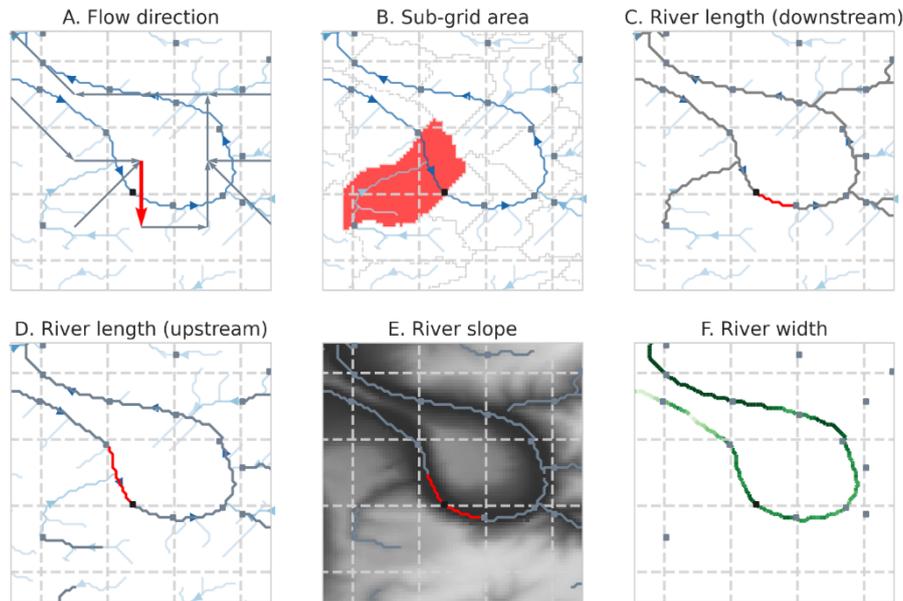



Figure 2: Output hydrography variables based on fine-resolution flow directions (blue lines in A-D; darker indicates larger upstream area) and/or outlet pixels (squares). Each variable is highlighted in red for the center cell and grey for other cells. The sub-grid area (B) is based on all pixels draining to the outlet pixel and limited by upstream outlet pixels. River length is derived based on the length of the flow path from outlet pixel to next downstream (C) or upstream (D) outlet pixel. The River slope (E) is calculated as the elevation (grey colors) difference over a flow path from a set length upstream to downstream of the outlet pixel. The river width (F) is derived based on the sub-grid river width (green colors) at the outlet pixel location.

2.3 Multi-resolution hydrography dataset

We derived the multi-resolution MERIT Hydro IHU dataset at resolutions of 30 arcsec (~1km), 5 arcmin (~10 km) and 15 arcmin (~30 km) by applying IHU to the recently published 3 arcsec MERIT Hydro data (Yamazaki et al., 2019). The original

MERIT Hydro data were near-automatically derived based on the MERIT DEM (Yamazaki et al., 2017) and several water body datasets and show good agreement with drainage areas reported by the Global Runoff Data Center (GRDC). We selected

235 this MERIT Hydro as it has a larger spatial coverage (N90 to S60) and better representation of small streams (Yamazaki et al., 2019) compared to earlier published hydrography datasets. It also provides supplementary data layers including hydrologically adjusted elevation, which is used to derive sub-grid river slope, and river channel width derived from the G1WBM permanent water body layer (Yamazaki et al., 2014), which is used to derive sub-grid river width. An overview of the layers in the upscaled *MERIT Hydro IHU* dataset is given in Table 1.

240

Table 1: Overview of hydrography and metadata layers of the MERIT Hydro IHU v1 dataset.

Parameter	Name	Unit	Description
Hydrography			
Flow direction	flwdir	-	D8 flow directions
River length	rivlen	m	Sub-grid distance between two outlet pixels along the flow path, diagonal steps are taken to be $\sqrt{2}$ times the pixel size. River length in the downstream direction has a "_ds" postfix
River slope	rivslp	mm^{-1}	Average slope based on the elevation difference between pixels at a set distance of 2 km around (1 km up- and downstream) the outlet pixel
River width	rivwth	m	Width at sub-grid outlet pixel. Note that the river width dataset contains gaps, similar to the underlying MERIT hydro width which need to be filled before application in hydrological models.
Sub-grid area	subare	m^2	Sum of sub-grid pixel areas draining to the pixel outlet confined by the upstream sub-grid pixel(s)
Upstream area	uparea	km^2	Accumulated sub-grid area
Stream order	strord	-	Strahler stream order
Elevation	elevtn	m+EGM96	Hydrologically adjusted outlet pixel elevation where all downstream cells have equal or lower elevation than its upstream neighboring cells, following the algorithm described by Yamazaki et al. (2012)
Meta data			
Erroneous flow direction	flwerr	-	Erroneous flow directions (binary), see section 2.4
Upstream area error	upaerr	km^2	Difference in upstream area error between the upscaled and native resolution river network at the outlet pixel.
Outlet pixel coordinates	outlon / outlat	-	Outlet pixel coordinates in EPSG:4326 projection.

3 Methods

Besides IHU, see Section 2.1, we use the Effective Area Method (EAM) (Yamazaki et al., 2008) and Double Maximum Method (DMM) (Olivera et al., 2002) to perform benchmark and case study experiments. DMM is selected as it is still often applied,

245 for example in the recently published multi-scale routing model (Thober et al., 2019), and EAM as it is at the basis of the IHU method. For a full description of the methods we refer the reader to Appendix A as well as the referenced papers. In section 3.1 we describe the implementation of upscaling algorithms and the application to the global MERIT hydro dataset, in section

3.2 the metrics used to evaluate the accuracy of the upscaling methods in a global benchmark and in section 3.3 the assessment of the effect of upscaling method and resolution on simulated discharge for a case study of the Rhine basin.

3.1 Global application of flow direction upscaling methods

In this section we describe the application of the DMM, EAM, and IHU algorithms on the global MERIT Hydro dataset (Yamazaki et al., 2019). All flow direction methods described in this section, including the DMM, EAM, and IHU upscaling algorithms, are implemented in the open source python *pyflwdir* package (https://pypi.org/project/pyflwdir). For this paper *pyflwdir* v0.4.4 was used. For the reading and writing of the geospatial raster data we used the rasterio python package (https://pypi.org/project/rasterio).

First some preprocessing is required to create a unique id and delineate each basin. As we cannot fit the entire global dataset into memory, an initial estimate based on the HydroBASINS dataset (https://hydrosheds.org/images/inpages/HydroBASINS_TechDoc_v1c.pdf) was used. First, we assigned each outlet in the native-resolution MERIT Hydro dataset tile by tile to the nearest Pfafstetter Level-2 HydroBASINS basin. We then used the bounding box of each Pfafstetter level-2 HydroBASINS basin, to combine the MERIT Hydro data tiles and delineate the basins in the MERIT Hydro dataset. Within each Pfafstetter level-2 basin, all individual basins were numbered from the largest to smallest basin based on area to get a unique ID for each basin.

As the upscaling algorithms do not require information about the entire basin, the algorithms can be applied to each tile with sufficient overlap. We found that tiles of 10 by 10 degree with a buffer of 10 times the target resolution provide consistent results. For each tile the flow directions are upscaled, local flow direction errors (see next section) are calculated, sub-grid area, river length, slope and width variables are derived, and the native resolution upstream area and basin values at the outlet pixels are read to assess the upscaling accuracy (see section 3.2). After upscaling, the coarse-resolution data were again combined for each Pfafstetter level-2 basin to derive upstream area, stream order and hydrologically adjusted elevation.

3.2 Upscaled flow direction accuracy metrics

We compute the following metrics at the global scale to benchmark the IHU against DMM and EAM:

- *Erroneous flow directions*: The flow direction of a cell is erroneous if the first outlet pixel downstream of the outlet pixel of that cell is not located in its downstream cell (i.e. where the flow direction points to). Examples of erroneous flow directions are given by the red arrows in Figure 1 panel 2A. This measures the local accuracy of the upscaled flow directions, with less erroneous flow directions indicating a better representation of fine resolution confluences at coarser resolutions.

- *Upstream area error*: The difference in upstream area between the target resolution upstream area at cell i $\widehat{A_i}$ and the fine-resolution upstream area at the cells' outlet pixel A_i. This is an aggregated measure of the accuracy of all

upstream flow directions. Based on the upstream area error we define: (2) absolute error ϵ; (3) relative error ϵ_{rel}; and (4) mean relative error $\overline{\epsilon_{rel}}$:

$$\epsilon = \widehat{A_i} - A_i, \qquad (2)$$

$$\epsilon_{rel} = \frac{|\widehat{A_i} - A_i|}{A_i}, \qquad (3)$$

$$\overline{\epsilon_{rel}} = \frac{1}{N}\sum_{i=0}^{N}\frac{|\widehat{A_i} - A_i|}{A_i}, \qquad (4)$$

3.3 Case study setup

For a case study of the river Rhine in Europe, we assessed the effect of resolution and upscaling method on simulated river discharge for a synthetic runoff event. For each method we calculated the difference in simulated peak flow magnitude and timing between three upscaling methods at resolution of 30 arcsec, 5 and 15 arcmin and a simulation based on the baseline 3 arcsec resolution. We expect smaller differences for IHU compared to other upscaling methods, especially at river confluences. The Rhine basin catchment area up to the river outlet near Rotterdam, the Netherlands, has a surface area of approximately 195,000 km^2, see Figure 3. The basin has many smaller contributing sub-basins including the Meuse basin with its confluence near the river mouth after flowing parallel to the Rhine for many kilometers. Further upstream, the Moselle basin has many meanders and in the upstream Swiss sub-basins many lakes are present. These features are typically hard to represent at coarser resolutions and therefore allow for a detailed benchmark between upscaling methods. Note that in reality the river flow in the Dutch part of the Rhine is more complicated than can be captured in D8 flow directions with an important bifurcation, splitting the Rhine into the Ijssel and Waal rivers and canals between the Waal and Meuse rivers.

The synthetic runoff event is assumed to be uniformly distributed throughout the Rhine basin. The runoff event is triangular shaped with a total duration of 10 days, it starts with 1.2 mm day^{-1} and increases linearly to 6 mm day^{-1} in 5 days after which it decreases back to 1.2 mm day^{-1} in the next 5 days. This yields an initial flow of around 2,700 m^{3}s^{-1} and a peak discharge of around 10,800 m^{3}s^{-1}, which roughly correspond to average and 1-in-35 year discharge conditions respectively for the Rhine basin at Lobith (Hegnauer et al., 2014).

A routing model was setup to simulate channel flow for river cells, here defined as cells with a minimum upstream area of 10 km^2. By using this threshold, the area of headwater catchments, for which we assume instantaneous drainage, is more comparable between resolutions. Routing was based on a kinematic wave routing model, solved using the Newton-Rhapson scheme as described in Chow et al. (1988) at a fixed timestep of 15 minutes. Runoff of a headwater cell and within a river cell is instantly accumulated and fed to the channel at the outlet pixel of that cell. Channel length, slope and width at all resolutions are based on definitions in section 2.2, where for the fine-resolution baseline data every pixel is considered to be an outlet pixel. A default length of 2 km around (1 km up- and downstream of) the outlet pixel was used to calculate the slope. To fill gaps in the river width observations we fitted a power-law relation between upstream area (A), as a proxy for bank full discharge, and MERIT Hydro river width (w) according to equation 5, where a and b are fitted to be 0.15 and 0.664

respectively, for more details see Appendix B, Figure B1. Note that this is a simple approach that will not yield satisfying

310 results if applied on large scale or to actual events instead of a sensitivity analysis based on a synthetic event.

$$w = a A^b ,$$ (5)

In addition, we applied a default manning roughness coefficient of 0.03 and a minimum slope of 0.1 m/km. The default parameters are selected after a sensitivity analysis of the results to the channel slope length, width and roughness parameters, see Appendix E, Figure F1-2. While the simulated discharge peak magnitude and timing are sensitive to these parameters in

315 varying degrees, it does not greatly affect the differences between methods and does not change the conclusions based on it.

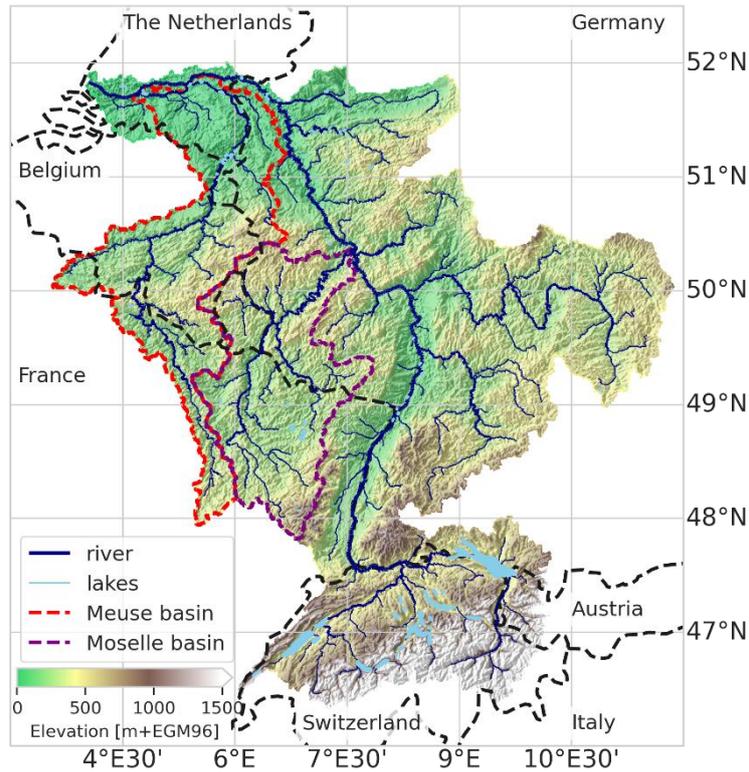



Figure 3: Study area: Rhine basin with average, elevation, rivers and basin outlines based on MERIT Hydro IHU 30 arcsec dataset (this paper); lakes are derived from the hydro Lakes dataset *(Messager et al., 2016).*

4 Accuracy of upscaling methods

320 In this section we benchmark the accuracy of IHU against the DMM and the EAM globally, based on erroneous flow direction and upstream area errors. Note that the results are presented at different spatial scales, from individual cells (Figure 4) to basins (Table 2) and 1 by 1 degree tiles (Figure 5).

First, we analyze the percentage of native resolution basins that are resolved in the upscaled flow direction maps. A basin is not resolved when it drains completely into neighboring basin(s) when upscaled and subsequently has no river outlet or pit at

the coarser resolution. At each resolution we analyze the basins with a total area larger than approximately one cell. IHU resolves more than 96.2% of the basins above the set thresholds compared to 85.7% and 87.6% for DMM and EAM at 30 arcsec resolution while a larger percentage is resolved at courser resolutions, see first row in Table 2. Only 2 basins larger than 5,000 km^2 are not resolved at 15 arcmin resolution using IHU. These are an endorheic basin in the south of the Arabian Peninsula (6996 km^2) and a small basin in Ontario, Canada (6830 km^2), see Figure C1-2. Both are merged with a larger nearby basin as the river mouth or pits runs through the same cell as the outlet or sink and it cannot be not assigned to another neighboring cell. The largest unresolved basins at 5 arcmin has an area of 3521 km^2 and at 30 arcsec an area of 40 km^2.

Next, we assess the percentage of cells with erroneous flow directions. This error is at the base of all upscaling errors discussed in this section and thus an important performance metric. The percentage of resolved basins that have less than 5% cells with erroneous flow directions is above 92.2% for IHU at all resolutions analyzed compared to 20.8% for DMM and 43.7% for EAM, see third row in Table 2, indicating that many more confluences are properly resolved at the upscaled resolution. The difference between the methods is smaller for basins with less than 1% cells with erroneous flow directions. While the second iteration of IHU successfully limits this error compared to DMM and EAM, it cannot be avoided. For cells with parallel fine-resolution flow paths it is impossible to correctly set upscaled flow direction for all cells in the D8 format.

Table 2: Percentage of resolved basins with area larger than approximately one cell that meet performance criteria based on relative basin area error, relative upstream area error (ϵ_{rel}) larger than 1% and cells with erroneous flow direction. For each criterium, the worst performance across all resolutions per method is highlighted.

| | 30 arcsec | | | 05 arcmin | | | 15 arcmin | | |
| | ~1 km^2 (510637 basins) | | | ~100 km^2 (27043 basins) | | | ~900 km^2 (7506 basins) | | |
	DMM	EAM	IHU	DMM	EAM	IHU	DMM	EAM	IHU
1. Basins resolved (percentage of total basins)	437821 **(85.7%)**	447228 **(87.6%)**	491203 **(96.2%)**	24060 (89%)	24693 (91.3%)	26537 (98.1%)	6502 (86.6%)	6758 (90%)	7336 (97.7%)
2. < 1% cells with flow dir errors	30.6%	49.4%	89.5%	**20.8%**	**42.1%**	**82.2%**	27.4%	50.1%	86.5%
3. < 5% cells with flow dir errors	30.6%	50.8%	95.7%	**20.8%**	**43.7%**	**92.2%**	27.4%	51.2%	92.7%
4. < 1% cells with ϵ_{rel} > 1%	27.0%	45.1%	85.8%	**16.9%**	**37.4%**	**79.4%**	23.1%	45.0%	84.5%
5. < 5% cells with ϵ_{rel} > 1%	69.0%	79.3%	95.1%	**59.7%**	**75.4%**	**93.9%**	69.9%	83.7%	96.3%
6. < 1% basin area error	68.3%	69.3%	97.9%	61.4%	67.1%	96.6%	**61.0%**	**66.0%**	**96.3%**
7. < 5% basin area error	75.0%	77.1%	98.3%	70.5%	76.2%	97.2%	**68.1%**	**74.9%**	**96.8%**

The relative upstream area error (ϵ_{rel}) considers the cumulative upstream error of erroneous flow directions. We find that more than 93.9% of the resolved basins have less than 5% cells exceeding the 1% upstream area error threshold for IHU compared to 59.7% for DMM and 75.4% for EAM, see fifth row in Table 2. The difference between the methods is larger for basins which have less than 5% cells exceeding the 1% upstream area error threshold. Figure 5 shows the percentage of cells within a 1 by 1 degree tile with a relative upstream area error larger than 1% for 5 arcmin resolution output maps. The differences between methods are consistent across the resolutions analyzed, see Figure D1-2. For IHU, most tiles have less than 1% cells with larger than 1% relative upstream area error. Exceptions are found in mountainous, glacierized and dry-land regions, see green areas in Figure 5. For example, in dry-land areas such as the south part of Arabian Peninsula, North part of Lake Eyre in Australia and some parts in the Sahara, where large rivers are absent, existing streams are ephemeral and flow directions extremely parallel, see for example Figure C1 and C4. In regions covered by ice sheets such as Greenland, streams are not

well depicted by the terrain elevation based on which flow direction are estimated to be extremely parallel. In such areas, even at high resolutions, flow directions are highly uncertain.

355 The basin area error is analyzed based on the relative upstream area error at the basin outlet cell, as shown with dots in Figure 5 for the 500 largest basins globally. For IHU more than 96.8% of the resolved basins have a basin area error relative to original basin area of less than 5% compared to 68.1% for DMM and 74.9% for EAM, see seventh row in Table 2. This difference between the methods is larger with a basis area error of less than 1%. While the 4th iteration of IHU, see section 2.1, successfully limits this error in comparison to DMM and EAM it cannot completely be avoided. Large basin area errors of more than 10%

360 for basins larger than 1000 km^2 occur at 360 basins at 15 arcmin, 91 at 5 arcmin and zero at 30 arcsec resolution. The largest basins with a 10% basin area error at each resolution are shown in Figure C3-5. These errors occur when sections of basins are merged with neighboring basins because a cell from one basin gets isolated between cells from another basin and none of the neighboring cells share a downstream outlet pixel.

The absolute upstream area error for all cells shows an improvement in performance for IHU compared to DMM and EAM,

365 see Figure 4. While the error increases slightly with larger resolutions it is consistently lower at all resolutions for IHU. At the 5 arcmin resolution 2.5% of the cells have a positive and 0.7% a negative upstream area error compared to 9.9% positive and 30.2% negative for DMM and 15.3% positive and 5.8% negative for EAM. In general, DMM shows a large percentage of negative upstream area errors while EAM and IHU tend to be skewed towards positive errors. Negative errors typically result from upscaled flow directions that cause a shortcut in a meandering section of a stream. The cells between the start and end of

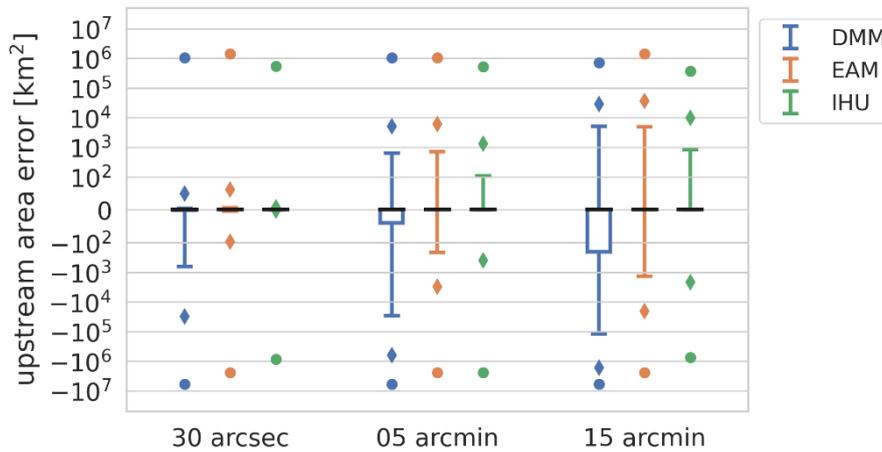
370 the shortcut then become a new branch in the upscaled flow direction map with smaller upstream area. Both positive and negative errors occur when a confluence in the upscaled flow directions is erroneously located upstream from the actual confluence, thereby increasing the upstream area in one branch while decreasing it in the other branch where the number of cells with a positive or negative error depends on the length and the number of outlet pixels on each branch, see example in Figure C6.

375

Figure 4: Absolute upstream area for DMM (blue), EAM (orange) and IHU (green) at three different resolution from 30 arcsec (~1km; left column) to 15 arcmin (~30 km; right column), where the black lines indicate the median error, the box the 25 -75 percentiles, the whiskers the 1-99 percentiles, the diamonds the 0.1-99.9 percentiles and the dots the min and max errors.

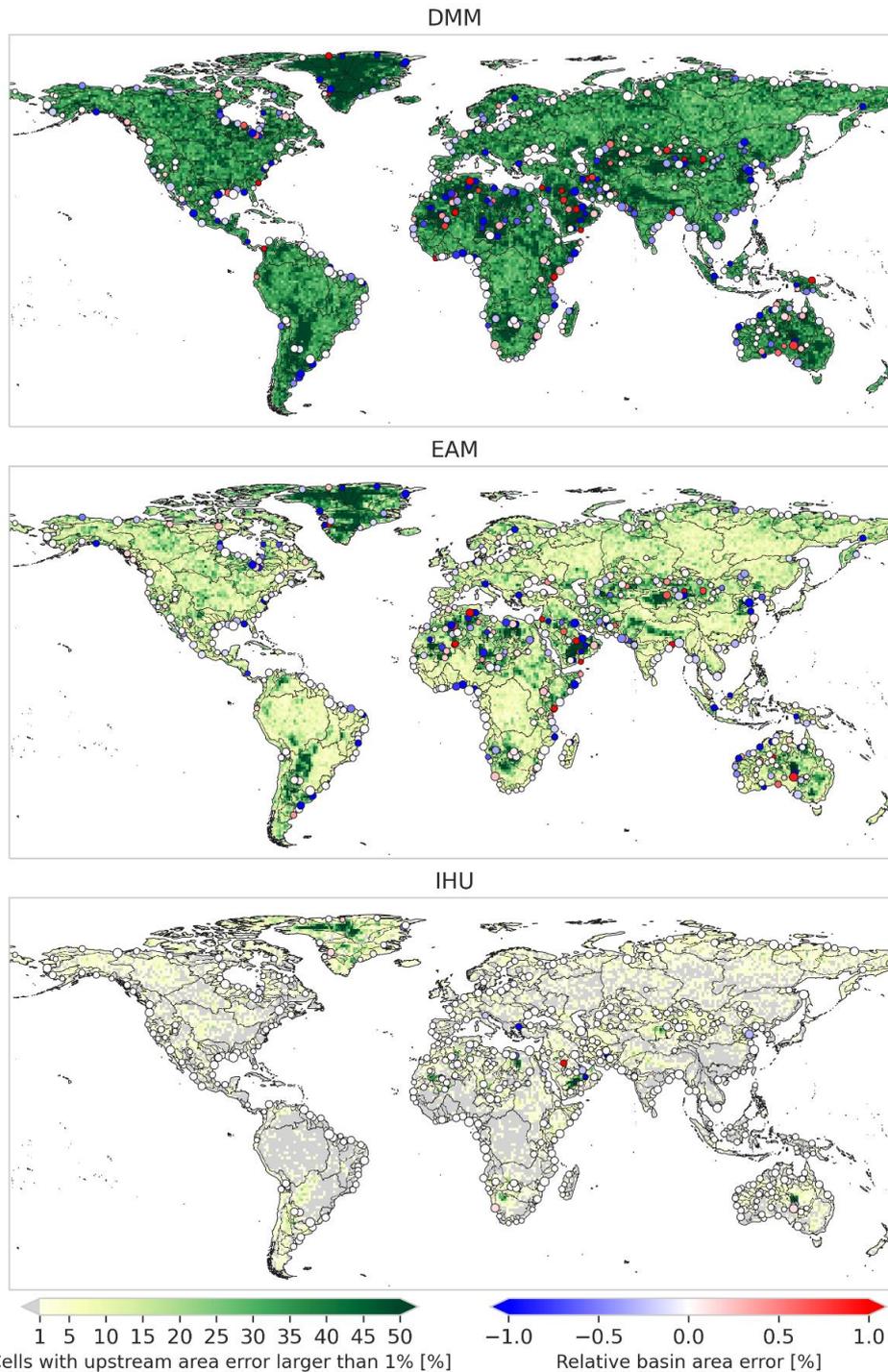



Figure 5: Relative upstream area error (ϵ_{rel}) at a 5 arcmin resolution (~10km) for DMM (upper), EAM (middle) and IHU (bottom). The background colors show the percentage of cells per 1x1 degree tile with a relative upstream area error of more than 1%, while the markers show the relative upstream area error at the outlets or sinks of the 500 largest basins globally (black lines).

5 Effect of upscaling method on simulated discharge

385   In this section we assess the effect of the flow direction upscaling method on simulated discharge for a case study of the river Rhine basin. First, we discuss the upscaled flow direction maps with resulting upstream area error as shown in Figure 6. Compared to DMM and EAM, the upstream area error for the Rhine basin based on IHU is smaller (negligible at 30 arcsec) and more localized. A clear error that occurs at 5 and 15 arcmin resolutions with DMM and EAM is the erroneous confluence of the Meuse river which is merged in the main stem upstream from the actual confluence, see Figure 6. Furthermore, at 30

390   arcsec and 5 arcmin resolution many meanders in the Moselle basin are not correctly resolved with DMM and EAM. For IHU at 15 arcmin a small error in the total basin area is made as a small stream near the outlet is erroneously merged into the Rhine basin yielding a small error of 55 km^2 (0.02%) at the outlet, see Figure 6I.

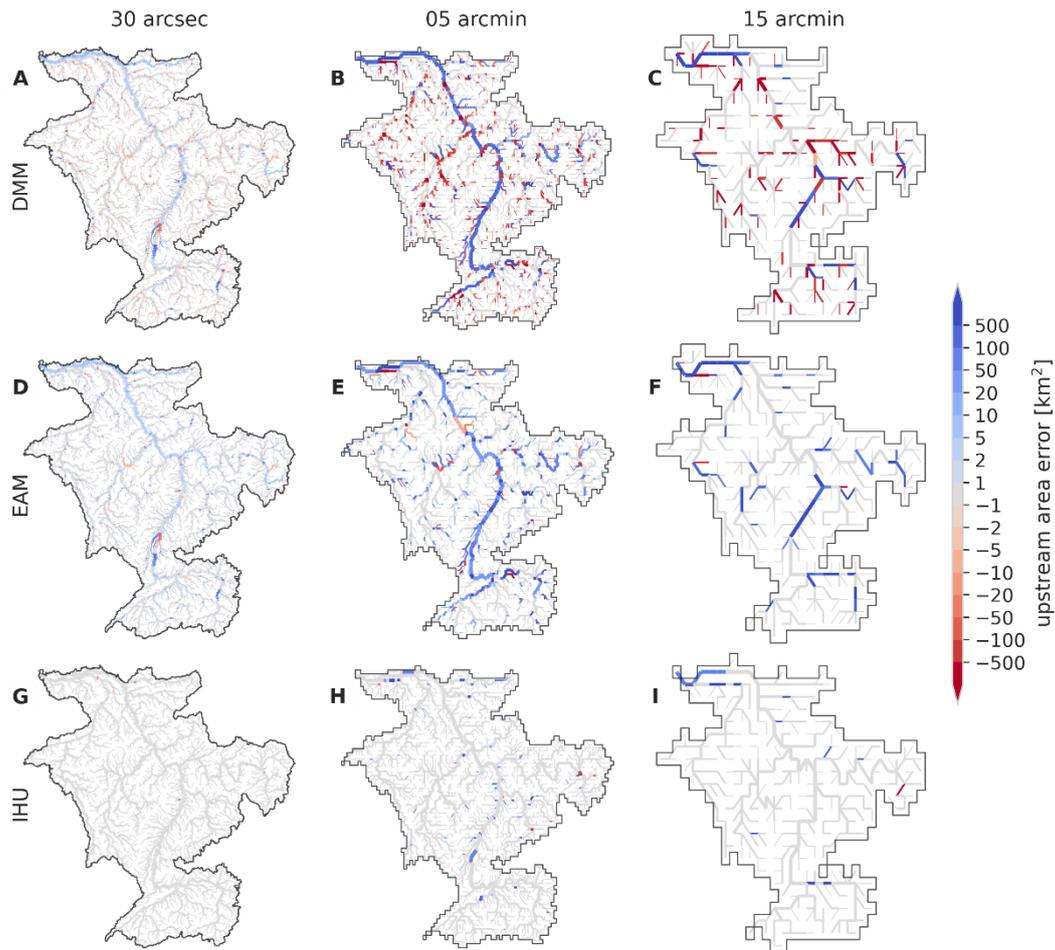



Figure 6: Upscaled MERIT Hydro flow direction network for the Rhine river at resolutions of 30 arcsec (left column), 5 arcmin
395   **(center column) and 15 arcmin (right column) as derived with DMM (first row), EAM (second row) and IHU (third row), where red colors indicate a negative and blue colors a positive upstream error. The line thickness is scaled with the upstream area.**

These flow direction maps together with sub-grid drainage area and river map are used to setup a distributed routing model to simulate discharge as the result of a synthetic runoff event that is uniformly distributed throughout the catchment. We analyze the difference between simulated discharge in the upscaled model compared to the original 3 arcsec model. Figure 7 shows the runoff event (right y-axis) and resulting flood peak wave at the river outlet (left y-axis) for all methods and resolutions. It is directly clear that models based on EAM (orange) and IHU (green) perform much better, i.e. show better similarity to the model at the original 3 arcsec resolution, than the models based on DMM (blue). The largest error in flood magnitude (+1024 m^3s^{-1}) and largest error in flood peak timing (-34 hours) are found for DMM at 5 and 15 arcmin resolution. These errors are likely due to a positive total area error in combination many shortcuts in the upscaled river network, see Figure 6A-C. The largest errors in flood peak magnitude for EAM (-252 m^3s^{-1}) and IHU (-287 m^3s^{-1}) are found at 15 arcmin resolution and the largest error in flood peak timing for EAM (+1 hours) and IHU (+2 hours) are found at 5 arcmin resolution. Both errors for both models have opposite sign and are very small compared to models based on DMM. While there are clear differences in the upstream area error between EAM and IHU, see Figure 6, the differences in simulated flood peak at the river outlet between EAM and IHU are small. This is likely because the effect of upstream area errors on simulated discharge cancel out at the river outlet, see Figure 6D-I, resulting in a net similar effect on the simulated discharge.

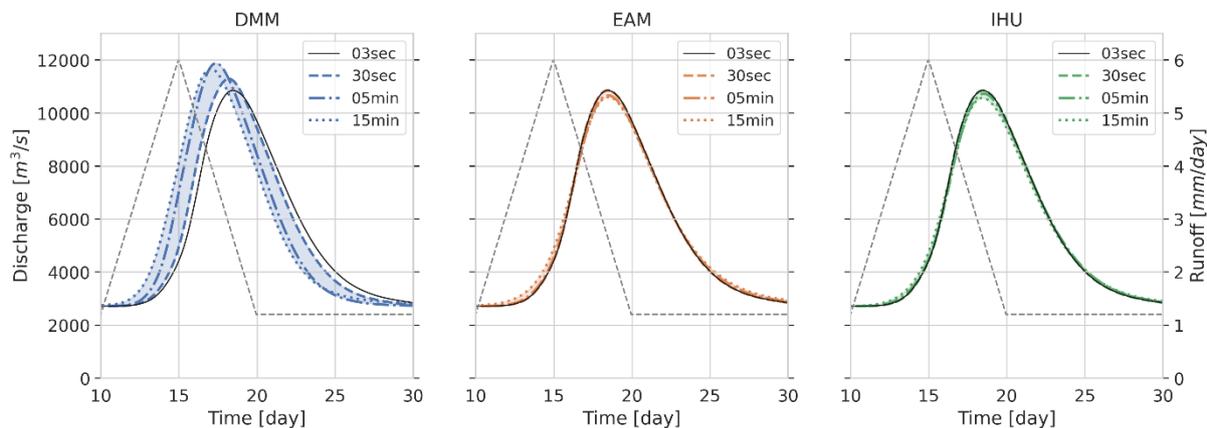



Figure 7: Simulated discharge at the river mouth of the Rhine river near Rotterdam for a synthetic accumulated runoff event (grey line) and based on models with native 3 arcsec resolution flow directions (black line) in comparison to upscaled flow directions based on DMM (blue; left), EAM (orange; center) and IHU (green; right) at 30 arcsec (dashed line), 5 arcmin (dash-dotted line) and 15 arcmin (dotted line) resolution. Note that the simulated discharge for the 30 arcsec EAM and IHU runs largely overlap with the native 3 arcsec run.

To better understand the effect of flow direction upscaling on river routing we therefore extend this numerical experiment to many locations across the Rhine basin, see Figure 8. The locations are selected based on the outlet pixels at 15 arcmin resolution that are at the same location or near outlet pixels a higher resolution. We use a maximum relative upstream area error of 1% to select nearby outlet pixels at higher resolutions on the same stream, which resulted in 214 location for EAM compared to 343 for DMM and 344 for IHU. The relatively small number of comparable locations for EAM can be explained by the definition of the outlet pixels which are selected within the effective area rather than at the edge of a target cell and therefore less likely to be the same between resolutions. Generally, simulated discharge based on the IHU upscaled river network yields

smaller flood peak magnitude and timing differences compared to DMM and EAM. In general, the IHU models show better similarity to the native resolution model across the full distribution of locations for both flood peak magnitude and timing. The maximum relative difference in flood peak magnitude between resolution for 50% (95%) of the locations is smaller or equal to 1.8% (6.3%) for IHU compared to 2.8% (24.2%) for EAM and 14.2% (98.5%) for DMM across all resolutions, see Figure 8D. Similar results are found in terms of flood peak timing, where for 50% (95%) of the locations the maximum difference between resolutions is smaller or equal to 2 (11) hours for IHU, compared to 3 (13) hours for EAM and 5 (29) hours for DMM, see Figure 8H. The differences in simulated discharge are caused by (local) upscaling errors in the river network. In both the DMM and EAM low resolution river networks the Meuse River (most downstream and largest tributary with a size of about 16% of the total basin) is merged into the main stem upstream from the actual confluence, causing large differences in both flood peak timing and magnitude in the Meuse section downstream of the erroneous confluence, see Figure 8A-B and E-F. Differences in flood peak simulation between upscaled and native resolution models are not only due to upscaling errors, also the upscaling of river width and slope are crucial for scale invariant discharge simulations. For instance, differences in river width between resolutions in the upstream part of the Rhine Basin cause a double peak at high resolutions ($\leq$ 30 arcsec) to smoothen out into a single peak at lower resolutions, yielding relatively large differences in flood peak timing, see Figure 8F and G. In the model, lakes and reservoirs are implicitly represented by larger channel widths from the MERIT Hydro dataset, introducing a buffering capacity in the river system. As the river width at lower resolutions is represented by the width at the outlet pixel of each cell the buffering capacity is different. Averaging the river width would change the buffering capacity as this results in a smoothened river width. The sensitivity of the simulated flood peak timing to river width is shown in the first row of Figure F2, where simulated river width is used for gaps in the river width data only (default); for lakes and data gaps or for all cells (for details see Appendix B). Using simulated river widths at lakes and reservoirs, minimized the timing error in the (upstream) Rhine. Besides width, the error in flood peak timing is also sensitive to the length along which the river slope is calculated, see Figure F2. If this length is varied with the model resolution, large errors in flood peak timing are introduced. However, the flood peak timing error is less sensitive to the precise length within a range of 1 to 5 km.

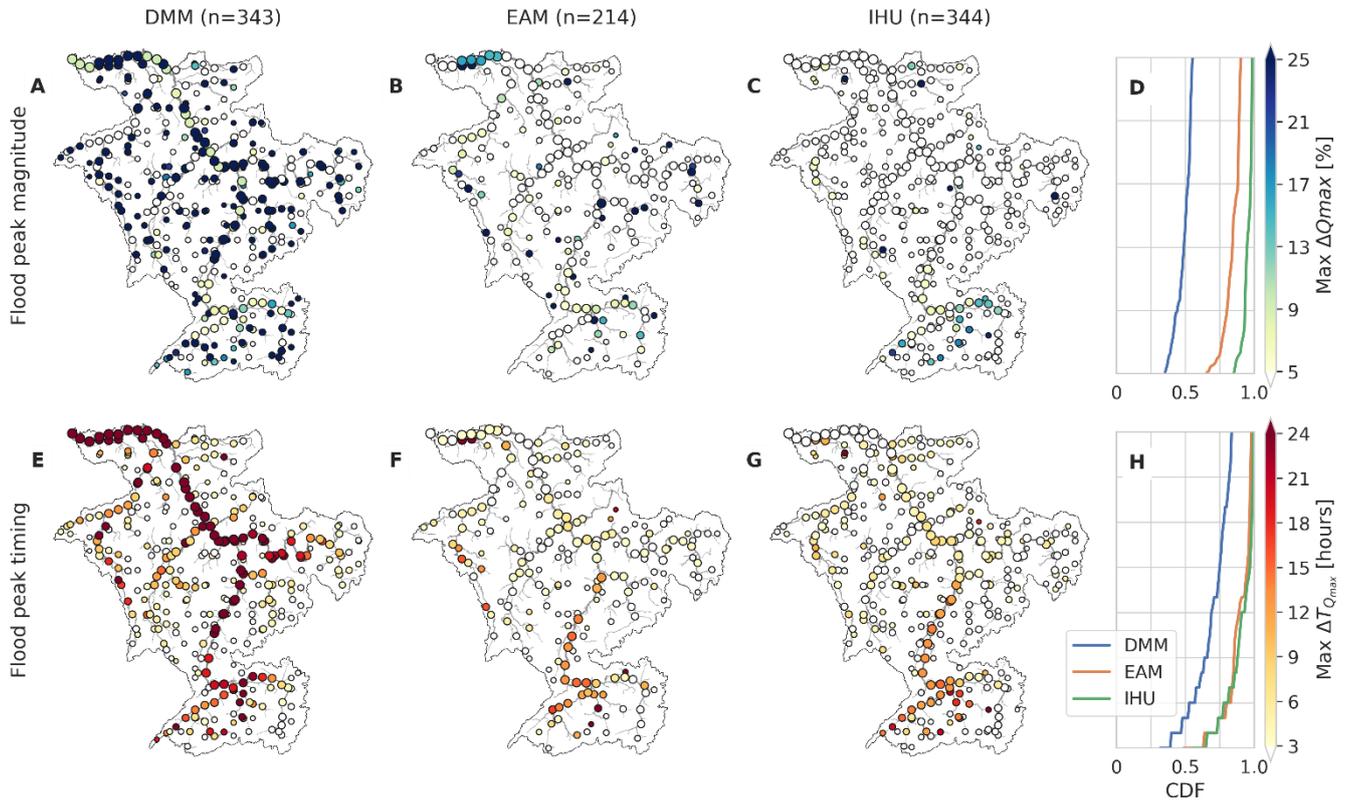



Figure 8: Maximum absolute difference in simulated flood peak magnitude (ΔQ_{max} - top row) and flood peak timing ($\Delta T_{Q_{max}}$ – bottom row) between upscaled (30 arcsec, 5 arcmin and 15 arcmin) and the baseline (3 arcsec) resolutions in the Rhine river basin based on DMM (first row), EAM (second row) and IHU (third row) upscaling methods. The right-most panels show the CDF curve for each method. with N denotes the number of comparable outlet pixels locations across all resolutions which are selected based on a maximum 1% upstream area difference. The marker size is scaled with upstream area and markers are plotted in order with increasing upstream area.

6 Discussion

The proposed IHU upscaling method was shown to successfully upscale flow direction data from 3 arcsec data to resolutions up to 15 arcmin. IHU balances between traditional methods such as DMM and EAM which only use local information and DTR which uses global information about the hierarchy of streams to determine which sub-grid stream to represent in each cell. IHU makes a first estimate of the representative sub-grid stream, but updates this for cells with erroneous flow directions based on contextual information. This makes IHU better suitable to be applied to high resolution hydrography data.

Compared to EAM and DMM, flow directions are better resolved, specifically near confluences. This is important for correctly modelling flood peak propagation downstream of confluences, especially when flood peaks in nearby (sub-)basins are correlated (Berghuijs et al., 2019). Erroneous IHU upscaled flow directions are still found in dry-land and ice-covered areas where the actual flow directions are also highly uncertain. These upscaling errors are mainly caused by many parallel flow

paths in the fine resolution hydrography data. This is partly a limitation inherent to the D8 format, which cannot represent
465   multiple rivers that run parallel through a cell, making it impossible to upscale flow direction data without any errors. While a
large majority of the basins has very small total area errors up to 15 arcmin resolution if upscaled with IHU, the small number
of large basins with more than 10% area error does increase with decreasing resolution. The exact upper limit of tolerable
upscaling errors and thus upscaling resolution, depends on its application and region of interest. We believe that the 15 arcmin
maps are suitable for many global scale applications but results in reported areas with lower accuracy should carefully be
470   interpreted. To guide the user on the quality of the upscaled MERIT hydro IHU dataset we therefore provide qualitative
metadata of erroneous flow direction and upstream area error. The DTR method by Wu et al. (2011) tries to solve the parallel
flow path issue by allowing for rivers to be diverted through adjacent cells in favor of smaller rivers within that cell. Potentially,
a stepwise upscaling procedure would even better preserve the largest basins and could be an interesting avenue for further
research.

475   Besides flow directions, IHU derives additional layers of sub-grid drainage area, river length and slope data, river width
estimates for large rivers, and hydrologically adjusted elevation. While these layers cover most parameters required in the
routing modules of many hydrological models, for a complete river parameter dataset a full-coverage river width layer is
required as well as river bed roughness. For more advanced routing models a river bed level and river bank-full depth might
also be required (Yamazaki et al., 2011). Several studies have shown that it is very hard to calculate reliable riverbed slopes
480   from global DEMs (LeFavour and Alsdorf, 2005), while river routing based on a kinematic wave solution, as commonly used
in large scale hydrological models, is very sensitive to the river slope (Thober et al., 2019; Yamazaki et al., 2011). We found
that a constant length across all resolutions of at least 1 km (500 m up- and downstream of the outlet pixels) is required to
provide a relatively scale invariant estimate of river slope as applied in the case study, see Figure F1-2. To achieve complete
coverage of river width, the sub-grid river width data require to be interpolated for data gaps and lake and reservoir areas if
485   these were to be modelled explicitly in the routing model. Here, we used a strongly simplified power-law relationship between
width and upstream area. For applications for real events instead of sensitivity analysis with synthetic data, this estimate should
be improved using the well-established geomorphic relationships between bank-full discharge and river depth as proposed in
the downstream hydraulic geometry framework (Leopold and Maddock, 1953; Savenije, 2003), for instance by the clustering
approach proposed by Andreadis et al. (2013) or additional river width data from e.g. Allen and Pavelsky (2018).

490   In this study we benchmarked several upscaling methods based on the same baseline hydrography data. This choice was made
to focus the paper on differences in upscaling methods, where we assume the underlying high-resolution data are correct. For
future studies it would be interesting to also compare the MERIT Hydro IHU dataset with often used hydrography datasets
such as hydroSHEDS (Lehner et al., 2008) and hydro1k (U.S. Geological Survey, 2000), both in terms of accuracy of the river
network and effect on simulated discharge for actual events.

495   Most multi-resolution routing models use data at pre-processed resolutions (Li et al., 2013; Wu et al., 2014; Yamazaki et al.,
2011). However, Thober et al (2019) recently presented a multi-scale routing model that includes the upscaling of flow
direction data based on DMM. Based on the presented results, using a multi-resolution modelling approach with hydrography

based on IHU would likely improve the model's capability to simulate similar fluxes independent of the model resolution in studies like Imhoff et al (2020). Furthermore, using a hydrological nesting approach, the model resolution could be varied within the model domain to add resolution where required for a correct representation of specific processes.

500

7 Conclusions

To describe flow directions and sub-grid river parameters in distributed hydrological models of different resolutions based on increasingly higher resolution hydrography datasets, automatic upscaling methods are required. The Iterative Hydrography Upscaling (IHU) method takes high resolution flow direction data and upstream area data as input to derive flow directions at

505 a coarser resolution while preserving the river structure. IHU was successfully applied to the 3 arcsec MERIT Hydro dataset to derive the MERIT hydro IHU multi-resolution hydrography dataset at resolutions of 30 arcsec (~1km), 5 arcmin (~10 km) and 15 arcmin (~30 km). Additional layers of sub-grid drainage area, river length, slope and width parameters are derived based on fine-resolution elevation and river width data.

Compared to other often used upscaling methods, upscaled flow direction maps with IHU show improved accuracy on all

510 metrics presented globally. For a case study of the Rhine basin we show that using IHU based hydrography maps allows to use lower resolution routing models with similar result for the entire the basin. Besides the upscaled flow direction data, the model similarity is also sensitive to how sub-grid river slope and width variables are derived. Because IHU is completely automated and yet accurate, it allows for a rapid uptake of new high-resolution flow direction data in distributed hydrological models at different resolutions.

515

Appendix A: Upscaling methods used for benchmarking

The EAM, DMM and IHU have three steps in common. First, a pixel is selected for every cell that determines the representative river within the cell. Second, the pixel is traced downstream until a certain criterium in a neighboring cell is met. Third, the upscaled flow direction is set towards that neighboring cell. The differences between the methods are based on how the first two steps are implemented. DMM and EAM are illustrated in Figure A1 and briefly describe below, more detailed descriptions can be found in the papers referenced.

Double maximum method:

- Step 1-1: The outlet pixel is defined as the pixel with the largest upstream area that is either a basin outlet pixel within the cell or located at the edge of the cell (grey squares).

- Step 1-2: The cell is offset half the cell size in the direction of the cell quadrant in which the outlet pixel is found (dashed grid lines). The outlet pixel is then traced downstream until it leaves the offset cell (black lines).

- Step 1-3: The upscaled flow direction is set to the cell where the trace in step 1-2 ends (orange arrows).

For a detailed description of the method we refer to Olivera *et al* (2002).

Effective Area Method:

- Step 2-1: The representative pixel (dark red squares), is defined as the pixel with the largest upstream area which is located within the effective area (shaded area) defined by equation 1, see section 2.1.

- Step 2-2: The representative pixel is then traced until the first downstream effective area is reached, which by definition is in a neighboring cell (black lines).

- Step 2-3: The upscaled flow direction is set to the cell where the trace in step 2-2 ends (orange arrows).

For a detailed description of the method we refer to Yamazaki *et al* (2008).

540

Appendix B: River width interpolation

To fill gaps in the river width observations we fitted a power-law relation between upstream area (A), as a proxy for bank full discharge, and MERIT Hydro river width (w), see equation 4.The MERIT Hydro river width was taken for all river cells within

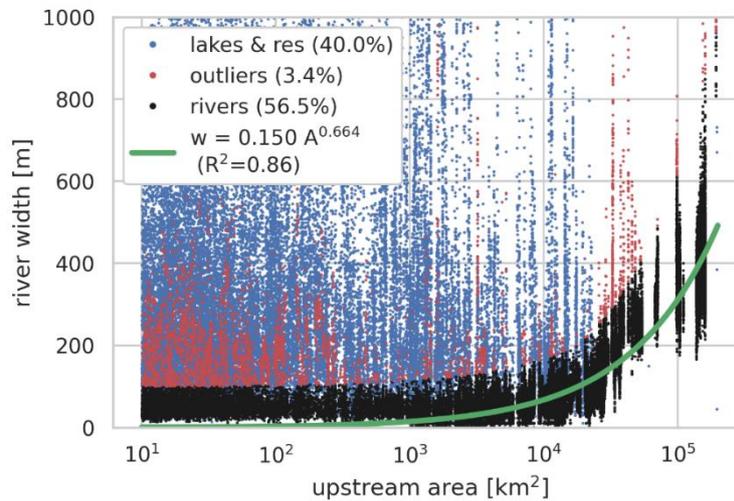
545   the Rhine catchment excluding cells within lakes and reservoirs based footprints from the hydro lakes (Messager et al., 2016) and GRAND databases (Lehner et al., 2011). We used a least squared error fitting algorithm from the python scipy.optimize package (Virtanen et al., 2020) which was iteratively fitted to the sample after removing outliers based on the best fit. Outliers are defined based on the difference between observed and simulated width of at least 200 m (for small widths) and the simulated width (for widths larger than 100 m).

550

Figure B1: Fitted relationship between river width and upstream area for the river Rhine basin

Appendix C: Examples of upscaling errors

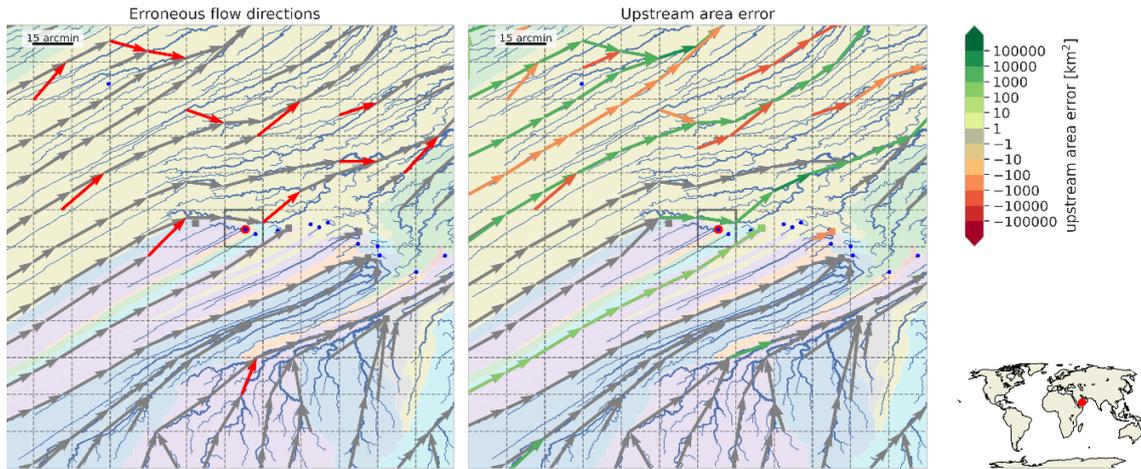



Figure C1: Largest endorheic basin (6996 km²) which is not resolved at 15 arcmin resolution indicated with highlighted basin pit. The blue lines show the fine resolution river, the background colors show basin boundaries, the dash lines the coarse resolution grid and the arrows the upscaled flow directions pointing from the outlet pixel or the original cell to outlet pixel of the destination cell. Flow directions are red if erroneous (left) and green for a positive- and red for a negative upstream area error (right).

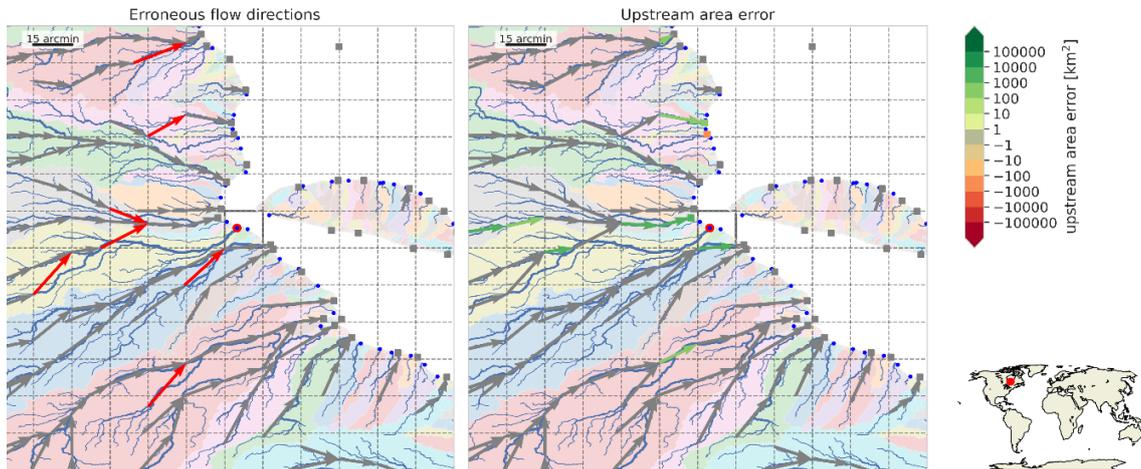



Figure C2: Largest exorheic basin (6830 km²) which is not resolved at 15 arcmin resolution indicated with highlighted basin outlet. See caption of Figure C1 for a full description.

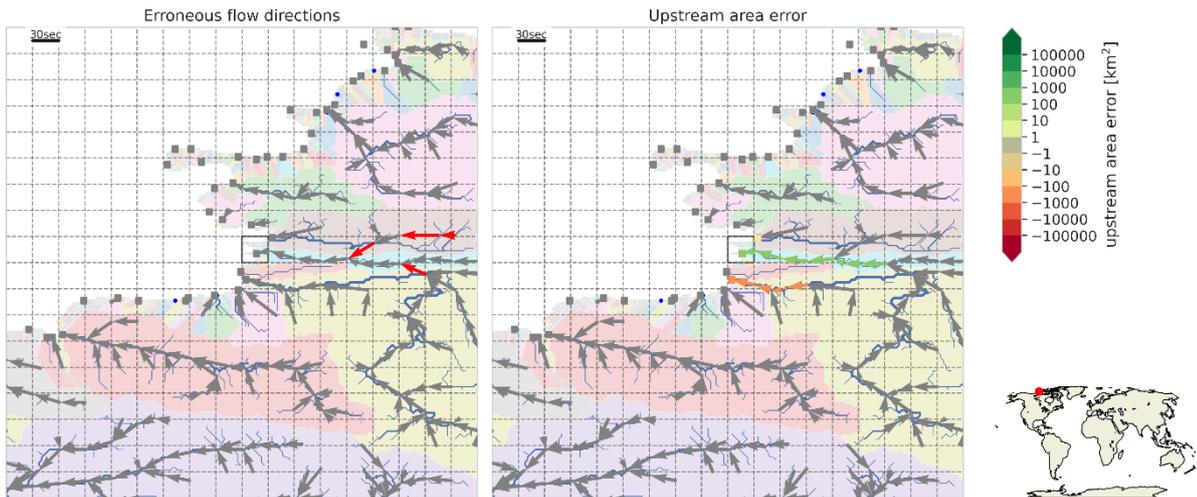
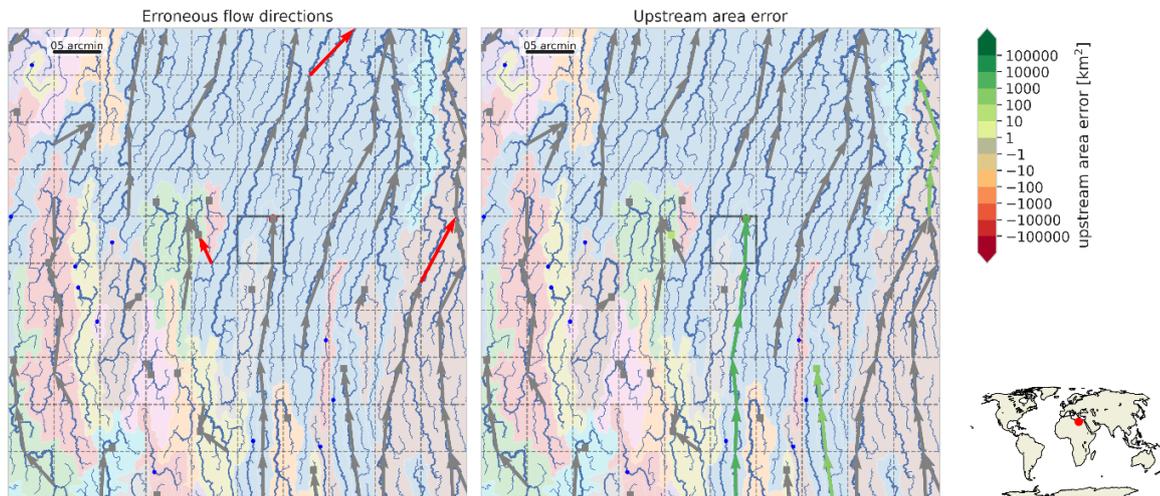



Figure C3: Largest basin (914 km²) with a relative upstream area error of more than 10% at 30 arcsec resolution indicated with highlighted cell. See caption of Figure C1 for a full description.

565 **Figure C4: Largest basin (16717 km²) with a relative upstream area error of more than 10% at 5 arcmin resolution indicated with highlighted cell. See caption of Figure C1 for a full description.**

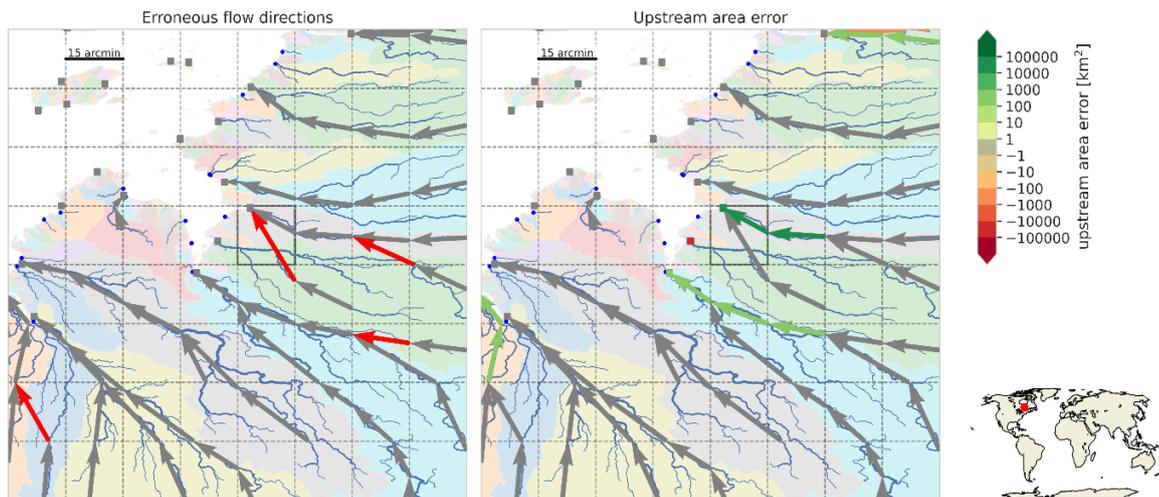



Figure C5: Largest basin (42017 km^2) with a relative upstream area error of more than 10% at 15 arcmin resolution indicated with highlighted cell. See caption of Figure C1 for a full description.

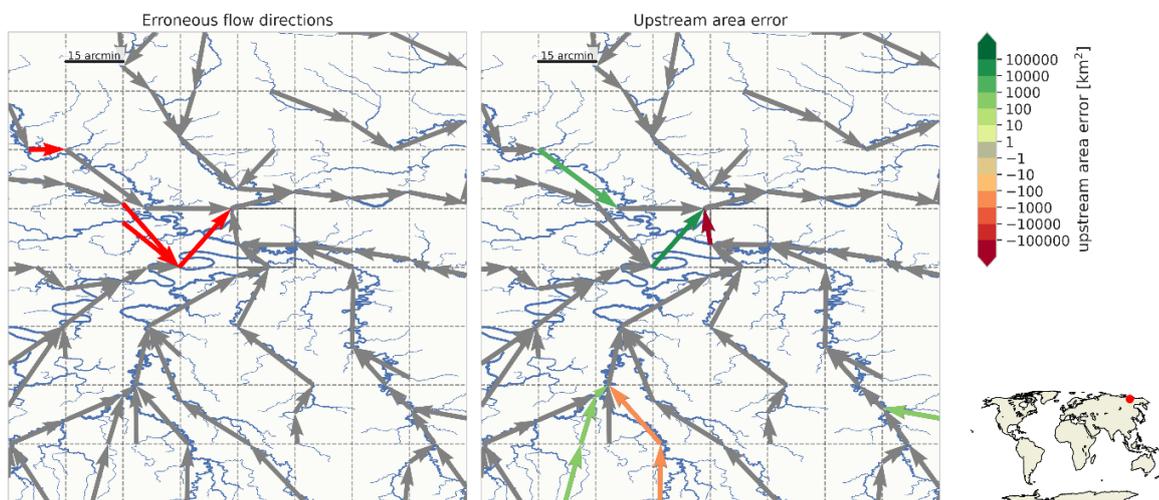



Figure C6: Second largest local upstream area error (-220876 km^2) at 15 arcmin resolution indicated with highlighted cell. See caption of Figure C1 for a full description

Appendix D: Accuracy benchmark of upscaling methods

This section shows maps of the relative upstream area error at resolutions of 30 arcsec and 15 arcmin in addition to the map at 5 arcmin as presented in section 3.

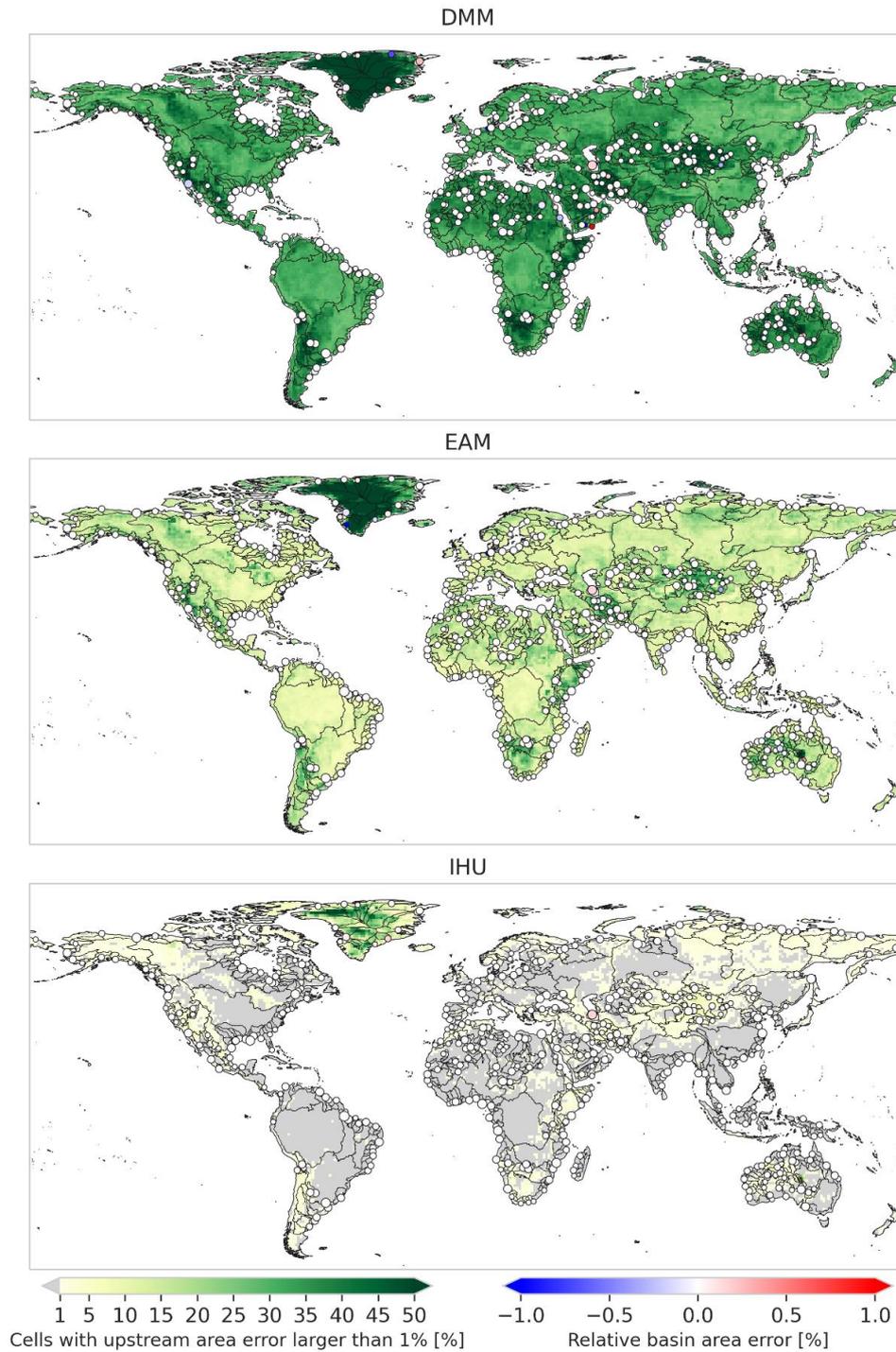



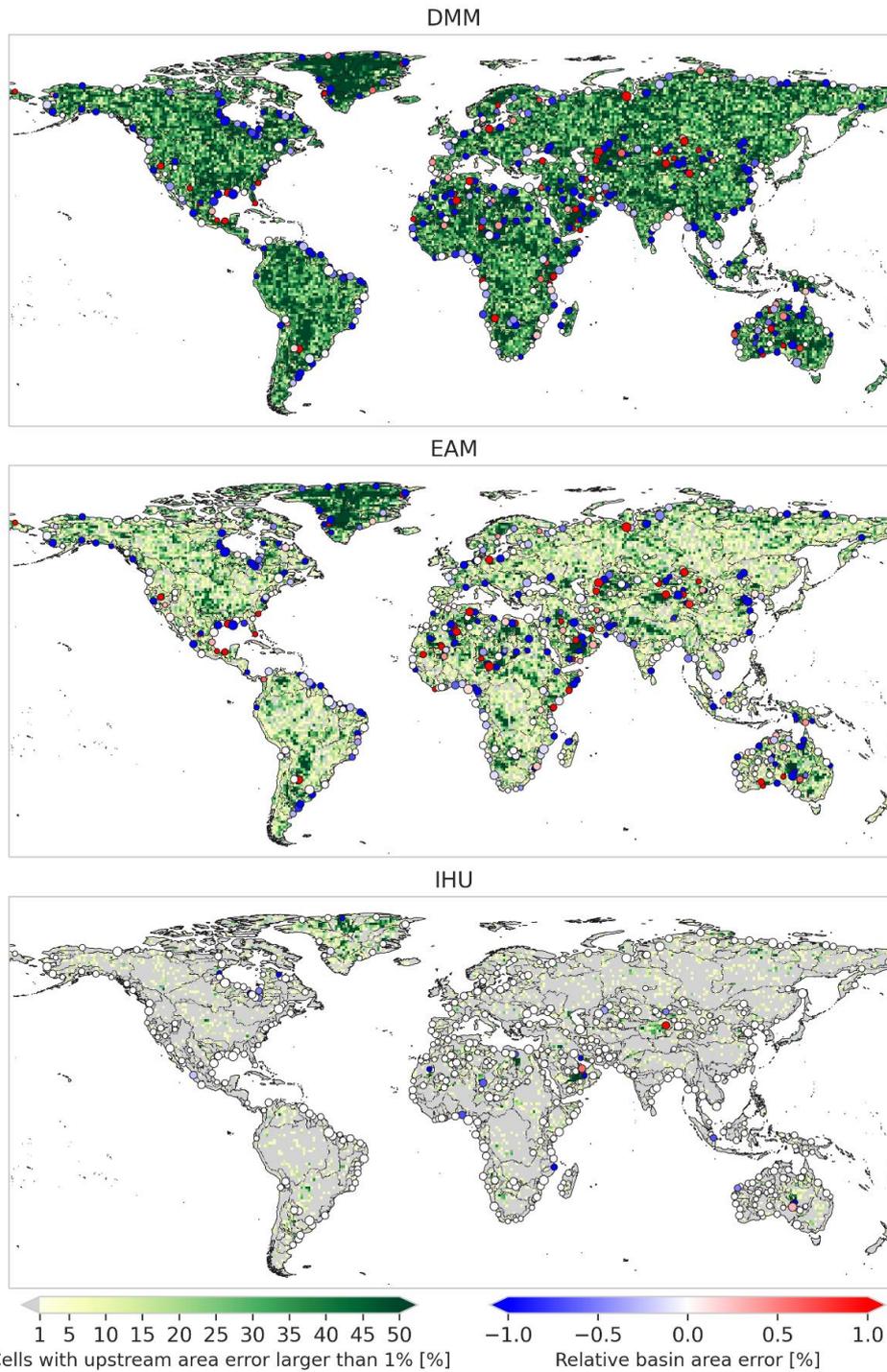
Figure D1: Percentage of cells at a 30 arcsec (~1km) resolution per 1x1 degree tile with an absolute relative upstream area error of more than 1%, while the markers show the upstream area error at basin outlet and the black lines the outlines of the 200 largest basins globally.

580

Figure D2: Percentage of cells at a 15 arcmin (~30km) resolution per 1x1 degree tile with an absolute relative upstream area error of more than 1%, while the markers show the upstream area error at basin outlet and the black lines the outlines of the 200 largest basins globally.

Appendix E: Sensitivity analysis IHU parameters

585 The sensitivity of the *R* parameter to define the effective area in step 1-1 (see equation 1) as well as the minimum length and minimum upstream area thresholds used to optimize sub-grid river length in step 3, see section 2.1, are tested for the river Rhine basin.

We tested the sensitivity to these thresholds based on four metrics expressed as percentage of the basin cells at various resolutions. Two upscaling accuracy metrics: flow direction error, relative upstream area > 1% as explained in section 3.2, as
590 well as a metric to assess the number of cells with small subgrid cell area (i.e.: < 25% of cell area) and small subgrid river length (i.e. < 25% of cell length).

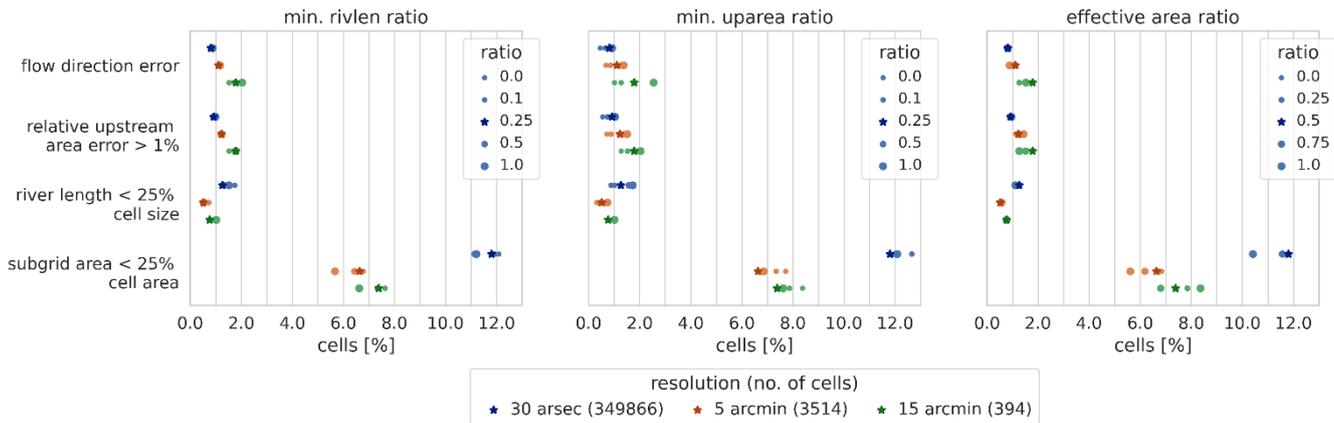



Figure E1: Sensitivity analysis for the minimum river length threshold (left panel), minimum upstream area threshold (center panel)
595 **and effective area definition (right panel), where the minimum river length threshold and R parameter in the effective area definition are expressed as a ratio of cell size and the minimum upstream area threshold as ratio of cell area. The sensitivity is tested for four metrics (y-axis) and expressed as a percentage of the total basin cells (x-axis). The default ratio is shown with a start and alternative ratios with dots for which the size is scaled with the ratio value.**

Appendix F: Sensitivity analysis runoff experiment

600 The sensitivity of the model similarity between upscaled and baseline resolution models to three key model variables is presented in this section. The similarity is expressed as difference in flood peak timing and magnitude between the upscaled and native resolution model.

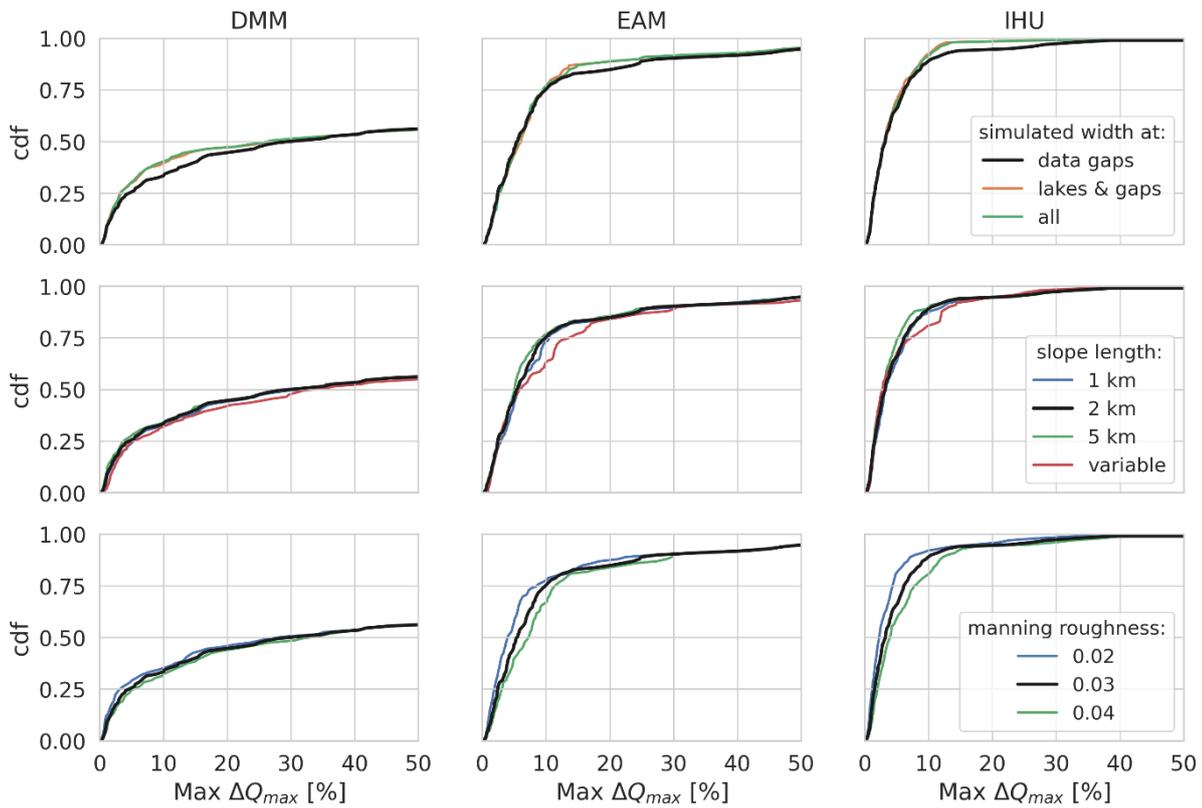



Figure F1: Sensitivity analysis of relative difference in simulated peak magnitude for a case study in the Rhine basin to three parameters (rows) and for three upscaling methods (columns). Each plot shows the CDF of all output location on the y-axis and the relative difference in simulated peak magnitude on the x-axis. The first row shows the sensitivity to the average MERIT Hydro vs power-law based channel width estimates; the second row shows the sensitivity to the minimum channel length over which the channel slope is estimated, and the third row shows the sensitivity to the manning roughness coefficient. The black line is the default case in all plots.

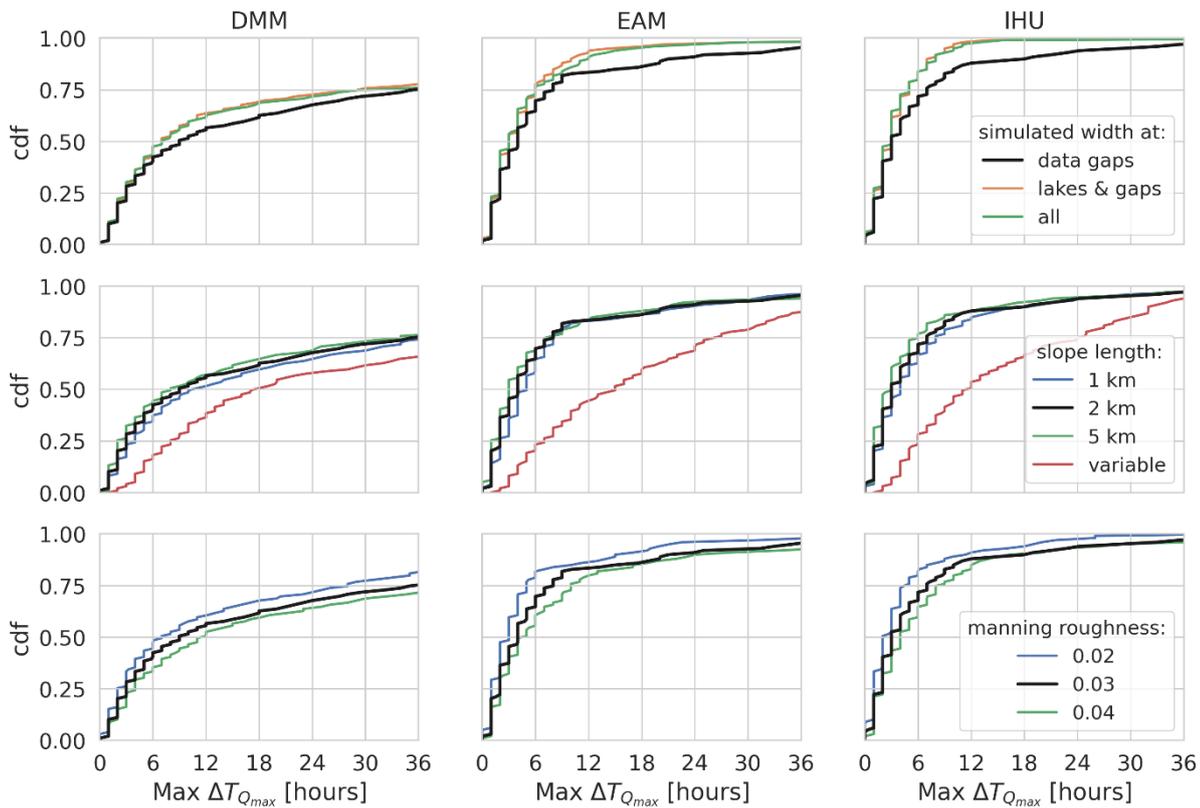



Figure F2: Sensitivity analysis of relative difference in simulated peak timing for a case study in the Rhine basin to three parameters (rows) and for three upscaling methods (columns). Each plot shows the CDF of all output location on the y-axis and the relative difference in simulated peak magnitude on the x-axis. The first row shows the sensitivity to the average MERIT Hydro vs power-law based channel width estimates; the second row shows the sensitivity to the minimum channel length over which the channel slope is estimated, and the third row shows the sensitivity to the manning roughness coefficient. The black line is the default case in all plots.

Author contribution

The IHU algorithm was developed implemented in pyflwdir by D.E. in close collaboration with W.V. and D.Y. The experiment was designed by D.E., P.W., H.W. and A.W and executed by D.E. All authors contributed to the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Code and data availability

All upscaling algorithms used in this study are implemented in the open-source python *pyflwdir v0.4.4* package (https://pypi.org/project/pyflwdir/0.4.4/). The multi-resolution *MERIT hydro IHU* dataset is available for download from zenodo.org (https://doi.org/10.5281/zenodo.4138776), under CC-BY NC or ODbL v1.0 license.

Acknowledgments

The research leading to these results received funding from the Dutch Research Council (NWO) in the form of a VIDI grant (grant no. 016.161.324), the IMPREX project, as funded by the European Commission under the Horizon 2020 framework Program (Grant 641811) and by the Dutch Ministry of Infrastructure and the Environment. The contribution by D.Y. is funded through "Cross-ministerial Strategic Innovation Promotion Program" of Japan.

References

Allen, G. H. and Pavelsky, T. M.: Global extent of rivers and streams, Science, 361(6402), 585–588, doi:10.1126/science.aat0636, 2018.

Andreadis, K. M., Schumann, G. J.-P. and Pavelsky, T. M.: A simple global river bankfull width and depth database, Water Resour. Res., 49(10), 7164–7168, doi:10.1002/wrcr.20440, 2013.

Berghuijs, W. R., Allen, S. T., Harrigan, S. and Kirchner, J. W.: Growing spatial scales of synchronous river flooding in Europe, Geophys. Res. Lett., 2018GL081883, doi:10.1029/2018GL081883, 2019.

Beven, K.: Rainfall-runoff modelling, John Wiley & Sons, Ltd, Chichester, UK., 2012.

Bierkens, M. F. P.: Global hydrology 2015: State, trends, and directions, Water Resour. Res., 51(7), 4923–4947, doi:10.1002/2015wr017173, 2015.

Bouaziz, L. J. E., Thirel, G., de Boer-Euser, T., Melsen, L. A., Buitink, J., Brauer, C. C., De Niel, J., Moustakas, S., Willems, P., Grelier, B., Drogue, G., Fenicia, F., Nossent, J., Pereira, F., Sprokkereef, E., Stam, J., Dewals, B. J., Weerts, A. H., Savenije, H. H. G. and Hrachowitz, M.: Behind the scenes of streamflow model performance, Hydrol. Earth Syst. Sci. Discuss., 2020, 1–38, doi:10.5194/hess-2020-176, 2020.

Chow, V. T., Maidment, D. R. and Mays, L. W.: Applied hydrology, McGraw-Hill series in water resources and environmental engineering., 1988.

Couasnon, A., Eilander, D., Muis, S., Veldkamp, T. I. E., Haigh, I. D., Wahl, T., Winsemius, H. C. and Ward, P. J.: Measuring compound flood potential from river discharge and storm surge extremes at the global scale, Nat. Hazards Earth Syst. Sci., 20(2), 489–504, doi:10.5194/nhess-20-489-2020, 2020.

Döll, P. and Lehner, B.: Validation of a new global 30-min drainage direction map, J. Hydrol., 258(1–4), 214–231, doi:10.1016/S0022-1694(01)00565-0, 2002.

Eilander, D., Couasnon, A., Ikeuchi, H., Muis, S., Yamazaki, D., Winsemius, H. C. and Ward, P. J.: The effect of surge on riverine flood hazard and impact in deltas globally, Environ. Res. Lett., 15(10), 104007, doi:10.1088/1748-9326/ab8ca6, 2020.

Fekete, B. M., Vörösmarty, C. J. and Lammers, R. B.: Scaling gridded river networks for macroscale hydrology: Development, analysis, and control of error, Water Resour. Res., 37(7), 1955–1967, doi:10.1029/2001WR900024, 2001.

Geertsema, T. J., Teuling, A. J., Uijlenhoet, R., Torfs, P. J. J. F. and Hoitink, T. J. F.: Anatomy of simultaneous flood peaks at a lowland confluence, Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2018-142, 2018.

Guse, B., Merz, B., Wietzke, L., Ullrich, S., Viglione, A. and Vorogushyn, S.: The role of flood wave superposition in the severity of large floods, Hydrol. Earth Syst. Sci., 24(4), 1633–1648, doi:10.5194/hess-24-1633-2020, 2020.

Hegnauer, M., Beersma, J. J., Van den Boogaard, H. F. P., Buishand, T. A. and Passchier, R. H.: Generator of rainfall and discharge extremes (GRADE) for the Rhine and Meuse basins, Final report of GRADE, 2, 1209424–1209004, 2014.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S.: Global flood risk under climate change, Nat. Clim. Chang., 3(9), 816–821, doi:10.1038/nclimate1911, 2013.

Hoch, J. M., Eilander, D., Ikeuchi, H., Baart, F. and Winsemius, H. C.: Evaluating the impact of model complexity on flood wave propagation and inundation extent with a hydrologic-hydrodynamic model coupling framework, Nat. Hazards Earth Syst. Sci., 19(8), 1723–1735, doi:10.5194/nhess-19-1723-2019, 2019.

Imhoff, R. O., van Verseveld, W. J., van Osnabrugge, B. and Weerts, A. H.: Scaling Point-Scale (Pedo)transfer Functions to Seamless Large-Domain Parameter Estimates for High-Resolution Distributed Hydrologic Modeling: An Example for the Rhine River, Water Resour. Res., 56(4), 1–28, doi:10.1029/2019WR026807, 2020.

Kummu, M., Gerten, D., Heinke, J., Konzmann, M. and Varis, O.: Climate-driven interannual variability of water scarcity in food production potential: A global analysis, Hydrol. Earth Syst. Sci., doi:10.5194/hess-18-447-2014, 2014.

LeFavour, G. and Alsdorf, D.: Water slope and discharge in the Amazon River estimated using the shuttle radar topography mission digital elevation model, Geophys. Res. Lett., 32(17), doi:10.1029/2005GL023836, 2005.

Lehner, B. and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems: GLOBAL RIVER HYDROGRAPHY AND NETWORK ROUTING, Hydrol. Process., 27(15), 2171–2186, doi:10.1002/hyp.9740, 2013.

Lehner, B., Verdin, K. and Jarvis, A.: New Global Hydrography Derived From Spaceborne Elevation Data, Eos Trans. Amer. Geophys. Union, 89(10), 93, doi:10.1029/2008EO100001, 2008.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N. and Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Front. Ecol. Environ., 9(9), 494–502, doi:10.1890/100125, 2011.

Leopold, L. B. and Maddock, T. J.: The Hydraulic Geometry of Stream Channels and Some Physiographic Implications., 1953.

Li, H., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M. and Leung, L. R.: A Physically Based Runoff Routing Model for Land Surface and Earth System Models, J. Hydrometeorol., 14(3), 808–828, doi:10.1175/JHM-D-12-015.1, 2013.

Lowe, W. H., Likens, G. E. and Power, M. E.: Linking Scales in Stream Ecology, Bioscience, 56(7), 591–597, doi:10.1641/0006-3568(2006)56[591:LSISE]2.0.CO;2, 2006.

Metin, A. D., Dung, N. V., Schröter, K., Vorogushyn, S., Guse, B., Kreibich, H. and Merz, B.: The role of spatial dependence for large-scale flood risk estimation, Nat. Hazards Earth Syst. Sci., 20(4), 967–979, doi:10.5194/nhess-20-967-2020, 2020.

Olivera, F., Lear, M. S., Famiglietti, J. S. and Asante, K.: Extracting low-resolution river networks from high-resolution digital elevation models, Water Resour. Res., 38(11), 13-1-13–8, doi:10.1029/2001WR000726, 2002.

Paiva, R. C. D., Collischonn, W. and Tucci, C. E. M.: Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach, J. Hydrol., 406(3–4), 170–181, doi:10.1016/j.jhydrol.2011.06.007, 2011.

Paz, A. R., Collischonn, W. and Lopes da Silveira, A. L.: Improvements in large-scale drainage networks derived from digital elevation models, Water Resour. Res., 42(8), doi:10.1029/2005WR004544, 2006.

Quinn, P., Beven, K., Chevallier, P. and Planchon, O.: The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models, Hydrol. Process., 5(1), 59–79, doi:10.1002/hyp.3360050106, 1991.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Müller Schmied, H., Sutanudjaja, E. H., Warrach-Sagi, K. and Attinger, S.: Toward seamless hydrologic predictions across spatial scales, Hydrol. Earth Syst. Sci., 21(9), 4323–4346, doi:10.5194/hess-21-4323-2017, 2017.

Savenije, H. H. G.: The width of a bankfull channel; Lacey's formula explained, J. Hydrol., 276(1–4), 176–183, doi:10.1016/S0022-1694(03)00069-6, 2003.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B. M., Colón-González, F. J., Gosling, S. N., Kim, H., Liu, X., Masaki, Y., Portmann, F. T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L. and Kabat, P.: Multimodel assessment of water scarcity under climate change, Proceedings of the National Academy of Sciences, 111(9), 3245–3250, doi:10.1073/pnas.1222460110, 2014.

Thober, S., Cuntz, M., Kelbling, M., Kumar, R., Mai, J. and Samaniego, L.: The multiscale routing model mRM v1.0: simple river routing at resolutions from 1 to 50 km, Geoscientific Model Development, 12(6), 2501–2521, doi:10.5194/gmd-12-2501-2019, 2019.

U.S. Geological Survey: HYDRO1K Elevation Derivative Database, [online] Available from: http://edc.usgs.gov/products/elevation/gtopo30/hydro, 2000.

Veldkamp, T. I. E., Wada, Y., Aerts, J. C. J. H., Döll, P., Gosling, S. N., Liu, J., Masaki, Y., Oki, T., Ostberg, S., Pokhrel, Y., Satoh, Y., Kim, H. and Ward, P. J.: Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century, Nat. Commun., 8, 15697, doi:10.1038/ncomms15697, 2017.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, \. Ilhan, Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P. and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nat. Methods, 17, 261–272, doi:10.1038/s41592-019-0686-2, 2020.

Wada, Y., Wisser, D. and Bierkens, M. F. P. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth System Dynamics, 5(1), 15–40, doi:10.5194/esd-5-15-2014, 2014.

Wanders, N., Wada, Y. and Van Lanen, H. A. J.: Global hydrological droughts in the 21st century under a changing hydrological regime, Earth Syst. Dyn., 6(1), 1–15, doi:10.5194/esd-6-1-2015, 2015.

Ward, P. J., Jongman, B., Sperna Weiland, F., Bouwman, A., van Beek, R., Bierkens, M. F. P., Ligtvoet, W. and Winsemius, H. C.: Assessing flood risk at the global scale: Model setup, results, and sensitivity, Environ. Res. Lett., 8(4), 44019, doi:10.1088/1748-9326/8/4/044019, 2013.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A. and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resour. Res., 47(5), doi:10.1029/2010WR010090, 2011.

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

Wu, H., Adler, R. F., Tian, Y., Huffman, G. J., Li, H. and Wang, J.: Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model, Water Resour. Res., 50(3), 2693–2717, doi:10.1002/2013wr014710, 2014.

Yamazaki, D., Masutomi, Y., Oki, T. and Kanae, S.: An Improved Upscaling Method to Construct a Global River Map, in Proceedings of the 4th Asia-Pacific Hydrology and Water Resources (APHW) Conference, Beijing., 2008.

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

Yamazaki, D., Kanae, S., Kim, H. and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, Water Resour. Res., 47(4), 1–21, doi:10.1029/2010WR009726, 2011.

Yamazaki, D., Baugh, C. A., Bates, P. D., Kanae, S., Alsdorf, D. E. and Oki, T.: Adjustment of a spaceborne DEM for use in floodplain hydrodynamic modeling, J. Hydrol., 436–437, 81–91, doi:10.1016/j.jhydrol.2012.02.045, 2012.

Yamazaki, D., Sato, T., Kanae, S., Hirabayashi, Y. and Bates, P. D.: Regional flood dynamics in a bifurcating mega delta simulated in a global river model, Geophys. Res. Lett., 41(9), 3127–3135, doi:10.1002/2014GL059744, 2014.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S. and Bates, P. D.: A high accuracy map of global terrain elevations, Geophys. Res. Lett., doi:10.1002/2017GL072874, 2017.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H. and Pavelsky, T. M.: MERIT hydro: A high-resolution global hydrography map based on latest topography dataset, Water Resour. Res., 55(6), 5053–5073, doi:10.1029/2019wr024873, 2019.

Zhao, F., Veldkamp, T. I. E., Frieler, K., Schewe, J., Ostberg, S., Willner, S., Schauberger, B., Gosling, S. N., Schmied, H. M., Portmann, F. T., Leng, G., Huang, M., Liu, X., Tang, Q., Hanasaki, N., Biemans, H., Gerten, D., Satoh, Y., Pokhrel, Y., Stacke, T., Ciais, P., Chang, J., Ducharne, A., Guimberteau, M., Wada, Y., Kim, H. and Yamazaki, D.: The critical role of the routing scheme in simulating peak river discharge in global hydrological models, Environ. Res. Lett., 12(7), 075003, doi:10.1088/1748-9326/aa7250, 2017.