Basin90m, a new global drainage basin dataset
(https://doi.org/10.5880/GFZ.4.6.2023.004)

Chuanqi He[1,2,3], Ci-Jian Yang[4], Jens M. Turowski[1], Richard F. Ott[1,5], Jean Braun[1], Hui Tang[1], Shadi Ghantous[6], Xiaoping Yuan[3], Gaia Stucky de Quay[2]

1. *GFZ German Research Centre for Geosciences, Potsdam*
2. *Department of Earth, Atmospheric & Planetary Sciences, Massachusetts Institute of Technology, Cambridge*
3. *School of Earth Sciences, China University of Geosciences, Wuhan*
4. *Department of Geography, National Taiwan University, Taipei*
5. *Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam*
6. *Institute of Environmental Science and Geography, University of Potsdam, Potsdam*

1. Licence

2. Citation

When using the data please cite:

He, Chuanqi; Yang, Ci-Jian; Turowski, Jens Martin; Ott, Richard F.; Braun, Jean; Tang, Hui; Ghantous, Shadi; Yuan, Xiaoping; Stucky de Quay, Gaia (2023): Basin90m, a new global drainage basin dataset. GFZ Data Services. https://doi.org/10.5880/GFZ.4.6.2023.004

The data are supplementary material to:

He, Chuanqi; Yang, Ci-Jian; Turowski, Jens M.; Ott, Richard F.; Braun, Jean; Tang, Hui; et al. (2024): A global dataset of the shape of drainage systems. Earth System Science Data. 10.5194/essd-16-1151-2024

Table of contents

3. Data description

The current global dataset of drainage system shapes has a relatively low spatial resolution. We obtained a new dataset (Basin90m) by calculating the drainage basins larger than 50 km^2 globally using a 90-meter resolution Digital Elevation Model (DEM). The total number of drainage basins is 667629. For each drainage basin, we extracted the spatial distribution of the longest river channel and the sinuosity of the river. We computed fundamental geometric parameters for the drainage basins, such as area, length, width, aspect ratio, slope, and elevation. Basin90m consists of vector files (ESRI Shapefile format) containing global drainage basins and river channels. The file sizes for the basin and river data are 7.8 and 2.5 GB, respectively. All calculations were automated using a MATLAB script. For a more detailed description of Basin90m, please refer to our submitted data description article titled "A global dataset of the shape of drainage systems" (He et al., 2024).

4. Extracting global drainage systems

The global 90-meter resolution SRTM DEM (Farr et al., 2007) was divided into 130 segments based on the boundaries of large drainage basins provided by HydroBASINS (Lehner et al., 2008; 2013). This maintained the integrity of all river basins. Several standard Geographic Information System (GIS) techniques were employed, including computing flow direction and accumulation for each cell in the DEM using the D8 algorithm. The carving method ensured proper drainage connectivity despite noise in the DEM. Basins and rivers were extracted based on flow direction and accumulation. To automate these GIS analysis steps across the global 130 DEM segments, the TopoToolbox software (Schwanghart and Scherler, 2014) was utilized within a Matlab script. This allowed the calculation of metrics not easily obtained from conventional GIS platforms like ArcGIS. The script outputs vector files containing delineated basins and rivers.

The minimum basin size is 50 km^2. Only the longest river was extracted for each basin. Rivers were defined to start at drainage divides and end at river mouths. Computing was performed using the high-performance cluster infrastructure at German Research Centre for Geosciences (GFZ). An article describing the Basin90m database has been submitted to Earth System Science Data. The title of the article is "A global dataset of the shape of drainage systems". Therefore, for more detailed extraction steps of the global drainage system, please refer to that article.

5. Measuring morphological indices

After delineating basin boundaries and their longest rivers, eight morphological parameters were measured to characterize each drainage system. These metrics included basin area, length, width, aspect ratio, slope, and elevation. For rivers, length and sinuosity were computed. The parameters were automatically calculated using a script in TopoToolbox software (Schwanghart and Scherler, 2014). For example, average basin elevation was obtained by summing all pixel elevations then dividing by the total number of pixels. The values for all eight metrics and stream order were stored as attributes in the basin shapefile. Strahler stream order is a system used to categorize drainage networks by their hierarchical structure (Strahler, 1957). Small, unbranched streams are labelled as first-order. When two first-order streams merge, they create a second-order stream. The largest stream order within Basin90m is 9. It encompasses three basins: the Amazon Basin, the Nile Basin, and the Congo Basin. For a more detailed description of measuring morphological indices, please refer to our submitted data description paper titled "A global dataset of the shape of drainage systems".

6. Removing basins from lakes and deserts

The goal of Basin90m is to provide drainage systems formed by surface water processes. While related to water flow, basins entirely within lakes were inappropriate. The HydroLAKES global lake dataset (Messager et al., 2016) was used to remove drainage systems. Sandy deserts were also excluded as their high permeability prevents surface water networks. We used global aridity index data (Zomer et al., 2022) to identify desert locations. Basins with over half their area in lakes or deserts were deleted, along with related rivers. Additionally, drainage systems intersecting the 60°N latitude was manually removed. In total, out of 840000 originally extracted basins, around 170000 were excluded, leaving 667629 basins and their longest rivers in Basin90m.

7. File description

The Basin90m dataset includes data in four sections. The first section comprises drainage basins globally with an area larger than 50 km². The data format is ESRI Shapefile. Eight morphometric indices of the drainage system are stored in the attribute table of the basin shapefile. Table 1 shows the definitions and units of each parameter for each drainage basin. For the calculation methods of each parameter, please refer to our submitted data description article titled "A global dataset of the shape of drainage systems".

The "**2023-004_He-et-al_Basins**" folder contains six subfolders, each representing a continent. Each continent's subfolder contains all the basins in that continent, categorized by different stream orders. For instance, the "South America" subfolder contains nine shapefile files corresponding to stream orders 1-9. The names of the shapefile files include their continent and stream order information. For example, "South_America_Basin_8.shp" represents all basins in South America with a stream order of 8.

The second part of the Basin90m data consists of global main river channels. The longest river channel of each basin is stored in a folder named "**2023-004_He-et-al_Rivers**". The internal structure of this folder is the same as the "Basins" folder. For instance, "South_America_River_8.shp" represents the main river channels in South America with a stream order of 8.

The third part of Basin90m data is an Excel file named "**2023-004_He-et-al_Basin90m**". This file contains eight morphometric parameters for all the basins. It includes both a globally merged sheet and sheets distinguishing different stream orders. The fourth part of Basin90m data is a folder named "Matlab_code", which contains Matlab code for the automated extraction of drainage systems and their morphometric parameters.

Table 1. The attribute table of basin shapefile in Basin90m

Headers in Basin Shapefile	Descriptions	Units
FID	Identifier for each drainage basin	N/A
Shape	The shape of all drainage basins is ‘Polygon’	N/A
Basin_Area	Basin area: flow accumulation at river month multiplied by pixel area	km²
Basin_L	Basin length: straight-line distance between divide and river mouth	km
Basin_W	Basin width: maximum width perpendicular to basin length	km
Basin_AR	Basin aspect ratio: ratio of basin length to width	N/A
River_L	River length: along-channel distance between divide and river mouth	km
Sinuosity	River sinuosity: ratio of river length to basin length	N/A
Basin_Slop	Basin mean slope: average topographic gradient within the catchment	Degree (°)
Basin_Ele	Basin mean elevation: average elevation within the catchment	meter

8. References

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, Rev. Geophys., 45, RG2004, 2007. https://doi.org/10.1029/2005RG000183

He, Chuanqi; Yang, Ci-Jian; Turowski, Jens M.; Ott, Richard F.; Braun, Jean; Tang, Hui; et al. (2024): A global dataset of the shape of drainage systems. Earth System Science Data. 10.5194/essd-16-1151-2024

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos, Trans., Am. Geophys. Union, 89, 93-104, 2008. https://doi.org/10.1029/2008EO100001

Lehner, B. and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, Hydrol. Processes, 27, 2171-2186, 2013. https://doi.org/10.1002/hyp.9740

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

Schwanghart, W. and Scherler, D.: Short communication: TopoToolbox 2 - MATLAB-based software for topographic analysis and modeling in Earth surface sciences, Earth Surf. Dyn., 2, 1-7, 2014. https://doi.org/10.5194/esurf-2-1-2014

Strahler, A. N.: Quantitative analysis of watershed geomorphology, Eos, Trans., Am. Geophys. Union, 38, 913-920, 1957. https://doi.org/10.1029/TR038i006p00913

Zomer, R. J., Xu, J., and Trabucco, A.: Version 3 of the Global Aridity Index and Potential Evapotranspiration Database, Sci. Data, 9, 409, 2022. https://doi.org/10.1038/s41597-022-01493-1