Mapping the world's free-flowing rivers

Dataset and attributes

Technical Documentation Version 1.0

prepared by Günther Grill & Bernhard Lehner
(guenther.grill@mail.mcgill.ca;
bernhard.lehner@mcgill.ca)

May 2019

1 Background and introduction

This documentation describes the hydrographic dataset that has been published as supplementary data for the research article by Grill et al. (2019): Mapping the world's free-flowing rivers (*Nature*) available at https://doi.org/10.1038/s41586-019-1111-9.

The dataset consists of a geometric network of the global river system and associated attributes, such as hydro-geometric properties, as well as connectivity metrics calculated using the methodologies described in the research article. Some of the dataset's attributes have been generated using source code located at https://github.com/ggrill/Free-Flowing-Rivers.

This documentation and the data set described are available for download at https://doi.org/10.6084/m9.figshare.7688801.

2 Methods and data characteristics

2.1 River reach network

A global river network delineation has been extracted from World Wildlife Fund's HydroSHEDS database (Lehner et al., 2008) at a grid resolution of 15 arc-seconds (approx. 500 m at the equator). For more information on HydroSHEDS please refer to the Technical Documentation at http://www.hydrosheds.org.

Rivers were defined to start at every pixel where the accumulated upstream catchment area exceeds 10 km^2, or where the long-term average natural discharge exceeds 100 liters per second, resulting in a total global river length of 35.9 million kilometers (excluding Antarctica). Rivers were broken into line segments at every confluence, here termed 'river reaches', representing the smallest unit of the river network and creating 8,477,883 million individual river reaches with an average reach length of 4.2 km. Each river reach is linked to an area of its contributing hydrological sub-catchment, here termed 'reach catchment', with an average extent of ~12 km^2.

In the context of the free-flowing rivers study (Grill et al., 2019), we define a **river reach** as a line segment between two confluences; a **river stretch** as two or more contiguous reaches but not a full river; and a **river** as an aggregation of river reaches that form a single-threaded, contiguous flowpath from headwater source to river outlet (Figure 1). The river outlet can represent either the river mouth at the ocean; a terminal inland depression; or the confluence with a larger river. Figure 1 further shows the relationship between rivers, river reaches, and reach catchments in grid cell format.

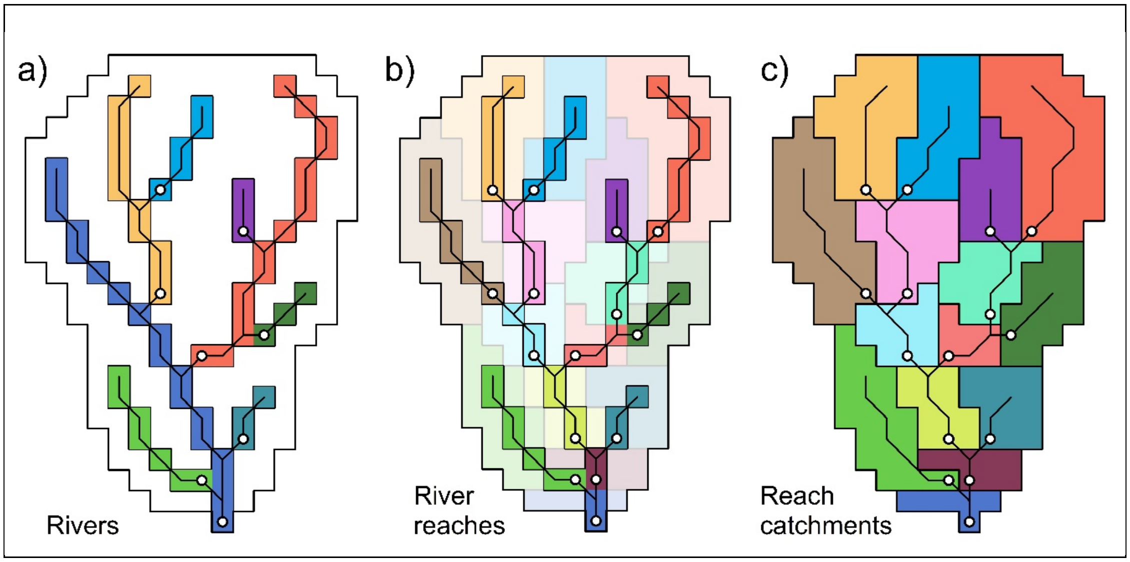

Figure 1: The panels show the grid cell delineation into (a) rivers; (b) river reaches; and (c) spatial zones that represent the contributing area of each river reach, i.e. reach catchments. Individual rivers, river reaches, and reach catchments are identified by different solid colors. Black river or river reach lines show the corresponding vector representation. The white dots represent the pour point of each river or river reach.

2.2 River reach attributes

The attributes associated with the individual river reaches of the global river network dataset combine hydrographic and geometric information as well as information related to free-flowing rivers (FFRs), including pressure indicators and final FFR status.

The primary sources of geometric attributes (e.g., length, distances, upstream area, etc.) stem from a combination of HydroSHEDS (Lehner et al., 2008), HydroBASINS (Lehner and Grill, 2013), and a beta version of HydroATLAS (Linke et al., (under review)), with discharge information provided by the global hydrological model WaterGAP (Alcamo et al., 2003; Döll et al., 2003) (v2.2 as of 2014).

Attribute information related to FFR indicators and results were generated based on methods developed in Grill et al. (2019) using underlying data such as GRanD v1.1 (Lehner et al., 2011); GOODD v1 (Mulligan et al., 2009); HydroLAKES v1.0 (Messager et al., 2016); (Borrelli et al., 2017); GRIP v3 (Meijer and Klein Goldewijk, 2009); DMSP-OLS v4 (Doll, 2008); GIEMS-D15 (Fluet-Chouinard et al., 2015), and Modis-derived urban areas (Schneider et al., 2009). River connectivity calculations are based on a river routing model called HydroROUT (Grill, 2015).

3 Data format and distribution

3.1 Data format and projection

The data layer, including attribute table, is provided in ESRI© Geodatabase format. This format can also be opened in other GIS applications, e.g. QGIS. The data is projected in a Geographic Coordinate System using the World Geodetic System 1984 (GCS_WGS_1984).

The dataset is available electronically in compressed zip file format. To use the data files, the zip file must first be decompressed. The zip file also includes a copy of this technical documentation.

3.2 Available columns and FFR attributes

The attribute table (Table 1) of the FFR river reach layer contains river network properties and identifiers (e.g. REACH_ID, GOID, NOID, etc.), geometric attributes of the river reach elements (e.g. length, upstream area) and hydrological attributes (discharge and volume). Attribute names starting with “BB_” describe rivers (also termed backbone rivers) (e.g., 'BB_ID').

Furthermore, the attribute table contains a set of pressure indicators (DOF, DOR, SED, USE, RDD, URB), as well as indices describing the connectivity status of the river network (CSI), the dominant pressure index (DOM), and the FFR status (for more explanations on these attributes see Grill et al., 2019).

Table 1: Attribute table of the FFR river reach network. See Grill et al. (2019) for additional information.

Nr.	Field	Description	Methods and data source
1	OBJECTID	Object Identifier	The attribute is automatically assigned by ESRI
2	Shape	Contains the geometry of the river reach	The attribute is automatically assigned by ESRI
3	REACH_ID	Reach Identifier. The reach identifier can be used to link this dataset to the HydroATLAS database.	Beta version of HydroATLAS (Linke et al., under review)
4	GOID	Global Object Identifier. The GOID is used to identify the river reach across different FFR assessments. Also used to link dams to river reaches.	HydroROUT (Grill, 2015)
5	NOID	Network Object Identifier. Same purpose as GOID	HydroROUT (Grill, 2015)
6	NUOID	Identifies the NOIDs of the next upstream river reach. If there is 'NoData' given, the reach is a headwater reach. Otherwise, the field holds 2 or more NOIDs. In the case of multiple NOIDs, the NOIDs are separated by an underscore.	HydroROUT (Grill, 2015)
7	NDOID	Identifies the NOID of the next downstream river reach. If value=0, the river reach represents a terminal reach (ocean, inland sink).	HydroROUT (Grill, 2015)
8	CON_ID	Identifier for the continent to which the reach belongs. Continent boundaries are delineated based on HydroBASINS. 1 = North America 2 = South America 3 = Europe 4 = Africa 5 = Asia 6 = Australia	HydroBASINS (Lehner and Grill, 2013)
9	CONTINENT	Name of the continent. See CON_ID.	HydroBASINS (Lehner and Grill, 2013)
10	COUNTRY	Country name	Beta version of HydroATLAS (Linke et al., under review)
11	BAS_ID	Basin Identifier. Identifies the hydrological river basin according to the HydroSHEDS framework.	Beta version of HydroATLAS (Linke et al., under review)
12	BAS_NAME	Basin Name (if available). Based on HydroSHEDS original basins and other sources.	Beta version of HydroATLAS (Linke et al., under review); multiple sources
13	LENGTH_KM	Length of the river reach in kilometers.	Beta version of HydroATLAS (Linke et al., under review)
14	VOLUME_TCM	Volume of the reach channel in thousand cubic meters (TCM). Calculated using width, length and depth of river channel.	Beta version of HydroATLAS (Linke et al., under review)

15	UPLAND_SKM	Upstream watershed area of river reach in square kilometers (SKM).	HydroSHEDS (Lehner et al., 2008)
16	DIS_AV_CMS	Average long-term (1971-2000) naturalized discharge in cubic meters per second (CMS).	Global hydrological model WaterGAP v2.2 as of 2014 (Alcamo et al., 2003; Döll et al., 2003)
17	RIV_ORD	River order. River order is here defined and calculated based on the long-term average discharge (DIS_AV_CMS) using logarithmic progression: 1 = > 100000 2 = 10000 – 100000 3 = 1000 – 10000 4 = 100 – 1000 5 = 10 – 100 6 = 1 - 10 7 = 0.1 - 1 8 = 0.01 - 0.1 9 = 0.001 - 0.01 10 = < 0.001	HydroATLAS (Linke et al., (under review))
18	ERO_YLD_TON	Sum of erosion in tons per year per river reach. Calculated as the sum of sediment erosion within the river reach catchment (i.e., sediment erosion is not accumulated along the river network).	Borrelli et al. (2017)
19	HYFALL	Indicates the presence (1) or absence (0) of one or more waterfalls along the river reach.	Global waterfalls database (Lehner et al., 2016)
20	BB_ID	Backbone River Identifier. Represents the contiguous river unit (defined as flow path from source/headwater to sink/terminus).	HydroROUT (Grill, 2015)
21	BB_NAME	Backbone river name.	HydroROUT (Grill, 2015); multiple sources
22	BB_LEN_KM	Backbone river length. Sum of length (LENGTH_KM) of the river reaches of the backbone river (BB_ID).	HydroROUT (Grill, 2015)
23	BB_DIS_ORD	Backbone river discharge order. River Order (RIV_ORD) of the most downstream reach of the backbone river (BB_ID).	HydroROUT (Grill, 2015)
24	BB_VOL_TCM	Backbone river volume. Sum of volume (VOLUME_TCM) of the river reaches of the backbone river (BB_ID).	HydroROUT (Grill, 2015)
25	BB_OCEAN	Ocean connectivity. Determines if river reach is part of a backbone river (BB_ID) that is directly connected to the ocean (value=1) or not (value=0). Used to summarize statistics based on connectivity to ocean.	HydroROUT (Grill, 2015)
26	INC	Filter field. In Grill et al. (2019), we considered all river reaches for routing purposes, but only analyzed and produced results for a subset of river reaches (INC = 1).	FFR assessment (Grill et al., 2019)

27	DOF	Degree of Fragmentation. Index from 0 to 100% (see extended Data figure 5a of manuscript).	FFR assessment (Grill et al., 2019); HydroSHEDS (Lehner et al., 2008); GRanD v1.1 (Lehner et al., 2011); GOODD v1 (Mulligan et al., 2009)
28	DOR	Degree of Regulation. Index from 0 to 100% (see extended Data figure 5b of manuscript).	FFR assessment (Grill et al., 2019); HydroSHEDS (Lehner et al., 2008); GRanD v1.1 (Lehner et al., 2011); GOODD v1 (Mulligan et al., 2009); HydroLAKES, v1.0 (Messager et al., 2016)
29	SED	Sediment trapping. Index from 0 to 100% (see extended Data figure 5c of manuscript).	FFR assessment (Grill et al., 2019); HydroSHEDS (Lehner et al., 2008); GRanD v1.1 (Lehner et al., 2011); GOODD v1 (Mulligan et al., 2009); HydroLAKES, v1.0 (Messager et al., 2016); Erosion map; Borrelli et al. (2017)
30	USE	Water consumption. Index from 0 to 100% (see extended Data figure 5d of manuscript).	HydroSHEDS (Lehner et al., 2008); WaterGAP (Alcamo et al., 2003; Döll et al., 2003) (v2.2 as of 2014)
31	RDD	Road construction. Index from 0 to 100% (see extended Data figure 5e of manuscript).	Grip v3 (Meijer and Klein Goldewijk, 2009); HydroROUT (Grill, 2015)
32	URB	Urbanization. Index from 0 to 100% (see extended Data figure 5f of manuscript).	DMSP-OLS v4 Doll (2008); Modis-derived urban areas (Schneider et al., 2009)
33	FLD	Inundation (floodplain) extent in river reach catchment (%).	Global inundation map GIEMS-D15 (Fluet-Chouinard et al., 2015)
34	CSI	Connectivity Status. Index from 0 to 100%; 100% = full connectivity; 0% = no connectivity.	FFR assessment (Grill et al., 2019)
35	CSI_D	Dominant pressure factor (DOM). Possible field values are: DOF; DOR; SED; USE; RDD; URB.	FFR assessment (Grill et al., 2019)
36	CSI_FF	CSI above or below free-flowing threshold. Indicates if the CSI value of a river reach is below (value = 0) or above (value = 1) the threshold of 95%. The attribute is used to calculate the free-flowing status of the river (see CSI_FF1 and CSI_FF2).	FFR assessment (Grill et al., 2019)
37	CSI_FF1	Free-flowing status (two categories). Indicates river reaches that belong to a river with “free-flowing” status (value = 1) or “non-free-flowing” status (value = 3). Note that the value 2 is reserved for river stretches with “good connectivity” status (see CSI_FF2).	FFR assessment (Grill et al., 2019); HydroROUT (Grill, 2015)
38	CSI_FF2	Free-flowing status (three categories). Indicates river reaches that belong to a river with "free-	FFR assessment (Grill et al., 2019); HydroROUT (Grill,

		flowing" status (value = 1), or a river stretch with "good connectivity" status (value = 2) or a river or river stretch with "impacted" status (value = 3).	2015)
39	CSI_FFID	River stretch identifier. Additional identifier to distinguish contiguous river stretches.	FFR assessment (Grill et al., 2019); HydroROUT (Grill, 2015)

4 License and citations

a) License agreement

b) Acknowledgement and citations

Citations and acknowledgements of this dataset should be made as follows:

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., Babu, S., Borrelli, P., Cheng, L., Crochetiere, H., Ehalt Macedo, H., Filgueiras, R., Goichot, M., Higgins, J., Hogan, Z., Lip, B., McClain, M., Meng, J.., Mulligan, M., Nilsson, C., Olden, J.D., Opperman, J., Petry, P., Reidy Liermann, C., Saenz, L., Salinas-Rodríguez, S., Schelle, P., Schmitt, R.J.P., Snider, J., Tan, F., Tockner, K., Valdujo, P.H., van Soesbergen, A., Zarfl, C. (*Nature*) Mapping the world's free-flowing rivers: datasets and materials. *figshare*
https://doi.org/10.6084/m9.figshare.7688801 (2019)

We kindly ask users to cite this dataset in any published material produced using it. If possible, online links to the dataset (on figshare) should be provided.

If users rely on this dataset as the main data source for their studies, or if their findings depend on the dataset in a fundamental way, we encourage reaching out to the creators of this dataset for possible collaboration.

5 References

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., Siebert, S. (2003) Development and testing of the WaterGAP 2 global model of water use and availability. Hydrological Sciences Journal 48, 317-337.

Borrelli, P., Robinson, D.A., Fleischer, L.R., Lugato, E., Ballabio, C., Alewell, C., Meusburger, K., Modugno, S., Schutt, B., Ferro, V., Bagarello, V., Oost, K.V., Montanarella, L., Panagos, P. (2017) An assessment of the global impact of 21st century land use change on soil erosion. Nature Communications 8, 2013.

Doll, C.N., (2008) CIESIN thematic guide to night-time light remote sensing and its applications (Data downloaded Nov. 2016 from http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html#AXP). Center for International Earth Science Information Network of Columbia University, Palisades, NY.

Döll, P., Kaspar, F., Lehner, B. (2003) A global hydrological model for deriving water availability indicators: model tuning and validation. Journal of Hydrology 270, 105-134.

Fluet-Chouinard, E., Lehner, B., Rebelo, L.M., Papa, F., Hamilton, S.K. (2015) Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. Remote Sensing of Environment 158, 348-361.

Grill, G., (2015) Challenges and opportunities in large scale river routing: Development and application of a hyper-resolution river routing model to assess anthropogenic impacts on freshwater ecosystems, Ph.D. thesis. Department of Geography. McGill University, Montreal, Canada.

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., Babu, S., Borrelli, P., Cheng, L., Crochetiere, H., Ehalt Macedo, H., Filgueiras, R., Goichot, M., Higgins, J., Hogan, Z., Lip, B., McClain, M., Meng, J., Mulligan, M., Nilsson, C., Olden, J.D., Opperman, J., Petry, P., Reidy Liermann, C., Saenz, L., Salinas-Rodríguez, S., Schelle, P., Schmitt, R.J.P., Snider, J., Tan, F., Tockner, K., Valdujo, P.H., van Soesbergen, A., Zarfl, C. (2019) Mapping the world's free-flowing rivers. Nature 568.

Lehner, B., Ariwi, J., Grill, G., (2016) HydroFALLS: a global waterfall database. Available at http://wp.geog.mcgill.ca/hydrolab/, Montreal, Canada.

Lehner, B., Grill, G. (2013) Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrological Processes 27, 2171-2186.

Lehner, B., Liermann, C.R., Revenga, C., Vorosmarty, C., Fekete, B., Crouzet, P., Doll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rodel, R., Sindorf, N., Wisser, D. (2011) High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment 9, 494-502.

Lehner, B., Verdin, K., Jarvis, A. (2008) New global hydrography derived from spaceborne elevation data. EOS, Transactions of the American Geophysical Union 89, 93.

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., Thieme, M. (under review) Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. Scientific Data.

Meijer, J., Klein Goldewijk, K., (2009) Global roads inventory project (GRIP), http://mapserver.mnp.nl/geonetwork. Netherlands Environmental Assessment Agency, Bilthoven, The Netherlands.

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

Mulligan, M., Saenz-Cruz , L., van Soesbergen, A., Smith, V.T., Zurita, L., (2009) Global dams database and Geowiki. Version 1. http://globaldamwatch.org/.

Schneider, A., Friedl, M.A., Potere, D. (2009) A new map of global urban extent from MODIS satellite data. Environmental Research Letters 4, 044003.