DamsImpactPersistence modelling steps

Contents

[Input data 2](#_Toc151560402)

[Scripts 3](#_Toc151560403)

[SETTINGS\_Dams\_impact.R 3](#_Toc151560404)

[1. Species information 3](#_Toc151560405)

[2. Species data preparation 4](#_Toc151560406)

[3. Create interbasins 4](#_Toc151560407)

[4. MVRS – body size relationship 4](#_Toc151560408)

[5. Get the results (species and spatial units) 5](#_Toc151560409)

[6. Visualization and statistics 5](#_Toc151560410)

[Regional analysis 7](#_Toc151560411)

[What to run when? 8](#_Toc151560412)

[New dams dataset 8](#_Toc151560413)

[New MVRS-body size relationship 8](#_Toc151560414)

[Platform (local versions) 8](#_Toc151560415)

# Input data

* Species data
  + In folder Fishdata (located in main folder) > folder Fish\_hybas
    - IUCN range data referenced to hydrobasins v2022-2
* Dams data
  + In Dams\_data folder (located in main folder)
    - DAMS folder
      * GDAT (used in this study)
      * Optional: GOOD2, GRanD, FhRED.
    - Regional folder
      * Regional dam datasets
      * Regional datasets in DamsBrazil, DamsMekong and DamsUS folders
        + Dams Brazil via Couto, T. B., Messager, M. L., & Olden, J. D. (2021). Safeguarding migratory fish via strategic planning of future small hydropower in Brazil. Nature Sustainability, 4(5), 409-416.
* HydroBASINS
  + Level 12 (customized with lakes)
    - In Hybas\_lakeL12 folder in main folder
  + Level 8 (customized with lakes)
    - In Hybas\_lake folder in main folder
* Tedesco & Su datasets
  + In folder Tedesco
    - Leprieur, Fabien; Brosse, Sébastien; Grenouillet, Gael; Tedesco, Pablo A.; Beauchard, Olivier; Bigorne, Remy; et al. (2017). A global database on freshwater fish species occurrences in drainage basins. figshare. Collection. <https://doi.org/10.6084/m9.figshare.c.3739145.v1>
    - Su, G., Logez, M., Xu, J., Tao, S., Villéger, S., & Brosse, S. (2021). Human impacts on global freshwater fish biodiversity. Science, 371(6531), 835-838. <https://doi.org/10.1126/science.abd3369>

<https://figshare.com/articles/online_resource/Scripts_and_files_for_Human_Impacts_on_Global_Freshwater_Fish_Biodiversity_/13383170>

*Table 2 🡪 save as: Su\_et\_al\_Science\_table.csv*

* Connectivity Index values Valerio
  + In Fishdata folder in CI\_Valerio folder
  + Dataset SI from Barbarossa, V., Schmitt, R. J., Huijbregts, M. A., Zarfl, C., King, H., & Schipper, A. M. (2020). Impacts of current and future large dams on the geographic range connectivity of freshwater fish worldwide. *Proceedings of the National Academy of Sciences*, *117*(7), 3648-3655.
  + Renamed the xlsx to pnasCI.csv (thus saved as .csv).

# Scripts

* The scripts are in the scripts/R folder.
* To run stuff on a Linux cluster, one may use .sh files. These have approximately the same name as the scripts they run. The sh files are in the scripts/sh folder. In the sh files adapt to own needs.
* If the script is run locally, check the local\_model\_folder and change in the scripts.

Code adapted from FishSuit (indicated at top of a script) refers to <https://github.com/vbarbarossa/fishsuit>

## SETTINGS\_Dams\_impact.R

* General settings
  + Folders
  + Formulas
  + MVRS relationship
  + Put this in the main model folder

## 1. Species information

1. 01A\_IUCN\_hybas\_fishbase.R
   * Get Fishbase information for all species that are in the IUCN (referenced to HydroBASINS) range data
   * Output: Fishbase/iucn\_hybas\_fb21names.csv and Fishbase/iucn\_hybas\_fb.csv
2. 01B\_IUCN\_habitat\_list.R
   * Get IUCN red list habitat info for each species
   * Output: Fishdata/iucn\_habitat\_list/
   * Run on cluster
3. 01C\_IUCN\_habitat\_type.R
   * With the results from 01B, create a table indicating if a species is lotic, lotic\_brackish, lentic or seasonal (0,1), (based on words in habitat info).
   * Output: Fishdata/iucn\_habitat\_type.csv
4. 01D\_compile\_traits.R
   * Combines info on habitat, traits and IUCN code
   * Input is all previous
   * Output: Fishdata/species\_traits\_extra.csv
5. 01E\_IUCN\_threat\_list.R
   * Get the threats (e.g. large dams, severe …) according to IUCN.
   * Input is species\_traits\_extra.csv
   * Output: in Fishdata/iucn\_threat\_list/
   * If combining the dams impact with the FishSuit model see scrips (adapt some lines to be sure).
   * Run on cluster
6. 01F\_IUCN\_threat\_dams.R
   * Get the info on dam threats for each species
   * Input is 1D and 1E
   * Output: Fishdata/IUCN\_DamsThreat.csv
   * Run on cluster

## 2. Species data preparation

* 02\_IUCNHB08L\_to\_HB12L.R
* IUCN species range data from level 8 to level 12 HydroBASINS
* Input: HydroBASINS L8 & L12, IUCN range data
* Output: csv’s per continent in Fishdata/Fish\_hybas/HBlevel12withlakes
* Run on cluster

## 3. Create interbasins

1. 03A\_DI\_dams2HB.R
   * Refer dams to HydroBASINS subbasins (level 12)
   * Input: dams dataset (choose in SETTINGS\_Dams\_impact.R)
   * Output: Dams\_impact/proc/dams\_gdat\_all\_hydrobasins.rds
2. 03B\_DI\_interbasins.R
   * Create interbasins between sink & dams, dams & dams or dam & basin boundary
   * Input: 03A output and HydroBASINS subbasins (level 12), also functions\_connectivity.R.
   * Output: Dams\_impact/proc/InterBasins\_tab\_GDATall\_',cont,'.rds
   * Run on cluster

## 4. MVRS – body size relationship

1. 04A\_Tedesco\_fb21\_names.R
   * Tedesco name validation using Fishbase
   * Input: IUCN fish ranges, Tedesco & Su datasets
   * Output: Tedesco/Occurrence\_table\_fb21names.csv
2. 04B\_Select\_tedescoSU\_IUCN.R
   * Filter the fish ranges data based on Tedesco and Su datasets
   * Input: IUCN fish ranges, HydroBASINS L8, Tedesco & Su datasets, 4A
   * Output: Tedesco/comp/Select\_TedescoSu\_IUCNspec\_",cont,".rds
   * May be executed on cluster (change to local = FALSE)
3. 04C\_decideMVRS.R
   * Get the MVRS – body size relationship. There are options to choose (e.g. what to include/exclude).
   * Input: choices (CHECK), file names (CHECK), IUCN ranges or 4B, HydroBASINS L8
   * Output: Dams\_impact/figs/MVRSbodysize/foldername -> figure, stats, points
   * When this is done, filled in the resulting equations in SETTINGS\_Dams\_impact.R (CHECK)

## 5. Get the results (species and spatial units)

1. 05A\_DI\_interb\_speciesranges.R
   * For each species, determine if the range within each interbasin is enough (>MVRS)
   * TRUE = large enough, FALSE = not large enough (<MVRS)
   * Input: 2, 3B, HydroBASINS L12, 1D
   * Output: for each species for each subbasin T/F.
   * Dams\_impact/proc/mvrstype\_damsused/model\_DI\_occurrence/
   * Run on cluster
2. 05B\_prep\_DI\_create\_CRL\_PAF\_PLR\_tabs.R
   * Prepare (almost) empty tables to fill in in 5C.
   * Input: 3, HydroBASINS L12
3. 05C\_DI\_create\_CRL\_PAF\_PLR\_tabs.R
   * Inputs: 5A, 5B, HydroBASINS L12
   * Create tables
     + Overview of the PLR for each species
       - Dams\_impact/proc/mvrstype\_damsused/model\_DI\_occurrence/tabs/PLRfragm\_tabDI\_fwonly.csv
     + Per subbasin PAF
       - RESsubbasins\_tabDI\_fwonly.rds and gpkg
     + Per interbasin CRL
       - RESinterbasins\_cur\_fwonly.rds and gpkg
     + Per hydrological basin PAF & CRL
       - mbres\_fwonly.rds and gpkg
   * Run on cluster

## 6. Visualization and statistics

* N.b. scripts need to be adapted when using other dam datasets than GDAT

1. 06A\_DI\_figures\_GDAT.R
   * Create all figures and tables regarding species specific results
   * Input: results of 1-6
   * Output:
     + In Dams\_impact/figs/Global/mvrstype\_damsused
     + Tables & all graph figures

|  |  |  |
| --- | --- | --- |
| **Name** | **Type** | **Description** |
| species\_fully\_out.csv | Table | The species that are removed entirely from the analysis due to their range area already being below MVRS |
| SI\_Results\_species\_specific.csv | Table | Nice table for SI material. Per species:  Binomial = species names according to IUCN  Range\_total\_km = total range size (excluding areas< MVRS without dams)  PLR\_km = absolute PLR  PLR\_perc = PLR of range total |
| paper/figure2.jpg | Figure | Figure1 paper species specific results |
| paper/figure2\_densityplotright.jpg | Figure | Figure1 but B is a density plot |
| paper/extra/CIcomparison.jpg | Figure | SI Figure: Comparing species specific PLR (%) and CI (Barbarossa et al. 2020) values |
| paper/extra/PLR\_SpecDamsThreat\_inclPLR0.jpg | Figure | Extra, PLR values across IUCN scope,severity and score of dam threat (large & unknown size) |
| paper/extra/PLR\_SpecTraits\_inclPLR0.jpg | Figure | Extra, PLR values across species traits (commercial importance, trophic level) |
| paper/IUCNcomparison\_incl0.jpg | Figure | SI Figure: comparison PLR results and IUCN dams threat |
| paper/extra/PAF\_vs\_FragmentSize.jpg | Figure | SI Figure: comparing PAF and fragment size. PAF= (occ-SR)/occ |
| paper/extra/PAF\_vs\_FragmentSize\_dots.jpg | Figure | Same as above but with dots |
| PLR\_perc\_top.csv | Table | Species highest PLR% top |
| PLR\_km\_top.csv | Table | Species highest PLR in km top |
| stats\_IUCN\_comp.txt | Text file | Numbers regarding SI figure IUCN comparison |
| stats.txt | Text file | Statistics concerning all species specific results and comparisons |

1. 06B\_DI\_maps\_GDAT.R
   * Create all maps and tables regarding spatial results
   * Input: results of 1-5
   * Output:
     + In Dams\_impact/figs/Global/mvrstype\_damsused/
     + Tables & all graph figures

|  |  |  |
| --- | --- | --- |
| **Name** | **Type** | **Description** |
| paper/maps/figure3.jpg | Figure | Figure 3: Map CRL per basin across the globe |
| paper/maps/Pextinct\_spec.jpg | Figure | SI Figure: Map showing in which basins and how many species lose their entire range within that basin due to dams |
| stats\_PAFrange\_MBtop100.csv | Table | Table with the top 100 basins with highest CRL values |
| SI\_Results\_basin\_specific.csv | Table | Nice table for SI material. Per basin:  Basin = MAIN\_BAS as in HydroBASINS (customized with lakes) Level 12  Cum\_range = cumulative range size (BGRS) of all species (excluding areas< MVRS without dams)  Cum\_PLR = cumulative PLR of all species  CRL= Cum\_PLR/Cum\_range\*100  Species\_PRL100 = number of species with PLR=100% within the basin (PLR=BGRS) |
| paper/maps/extra/… | Figures | Maps of other variables and/or with different scales |

1. 06C\_SI\_map\_Tedesco.R
   * Figure S1
   * Input: Tedesco data, IUCN range data, HydroBASINS level 8

## Regional analysis

For the regional analysis we use the same approach as global, but use different scripts from step 3 onward. BRA = Brazil, MEK = greater Mekong region, USA = United States

* Step 3A is performed in the scripts REG3A\_DI\_dams2HB\_XXX.R where XXX is the region.
* Steps 3B up to and including step 5 are performed in the scripts REG3B5\_DI\_XXX.R where XXX is the region. Run this on a cluster.
* For the visualization (step 6) all regions are dealt with simultaneously in the scripts REG6Astats\_DI\_figures.R, REG6Aonefig\_DI\_figures, REG6B\_DI\_maps \_SIfig.R and REG6B\_DI\_maps\_dams\_Regions.
  + The maps from 6B are in the BOTH/mvrs\_type/paper/maps folder of the region.

|  |  |  |
| --- | --- | --- |
| **Name** | **Type** | **Description** |
| PLR\_perc\_top.csv | Table | Species highest PLR% top depending on the dams data used (different folders in figs/REG/…) |
| PLR\_km\_top.csv | Table | Species highest PLR in km top depending on the dams data used (different folders in figs/REG/…) |
| stats.txt | Text file | Stats on PLR results depending on the dams data used (different folders in figs/REG/…) |
| Dams\_impact/figs/REGIONAL\_compare  \_largesmalldams\_perckm2.jpg | Figure | Figure 4: comparing regional results large vs large+small dams |
| REG\_PAF\_subbasins\_combined.jpg | Figure | SI Figure: per region PAf per subbasin for large or large+small dams  In the BOTH/mvrs\_type/paper/maps folder of the region. |
| Dams location figures | Figures | SI Figure: Dams\_BOTH\_size.jpg  Figures on dam locations: only small dams, only large dams, large and small dams (with and without legend)  In the BOTH/mvrs\_type/paper/maps folder of the region. |

# What to run when?

## New dams dataset

* 3, 5, 6AB

## New MVRS-body size relationship

* 4(C), 5AC, 6AB

# Platform (local versions)

A screen shot of a computer

Description automatically generated