|  |  |  |
| --- | --- | --- |
|  | Files | Description |
|  | Master\_Thesis.Rproj | R Project file |
|  | mth\_functions.R | R-file with self-written R functions in the framework of the master’s thesis |
|  | .gitignore | Gitignore file in which all folders are named which are to be ignored to upload on github |
|  | **Folder: Pre\_Processing** | **Description** |
| Quarto files | basement\_simulations.qmd | Quarto-file simulating all discharge scenarios of eight river sections using a self-written function from “mth\_functions.R” |
| hydrodynamic\_model\_results.qmd | Quarto-file reading all basement simulation results and converting them into single attribute rasters (in the raster\_files folder) and stacked rasters (in the rasters\_stacked folder) using a self-written function from “mth\_functions.R” |
| reconstruct\_discharges.qmd | Quarto-file reconstructing hydraulic conditions on the days of macroinvertebrate sampling |
| prepro\_data.qmd | Quarto-file reconstructing and joining abiotic, GPS and macroinvertebrate sampling data into one .csv file |
| index\_calculation.qmd | Quarto-file calculating all macroinvertebrate indices and constructing/combining the calculated indices with all environmental predictors needed for the machine learning models (one data frame per environmental predictor set used in the machine learning models) |
| added\_variables\_hdm.qmd | Quarto-file constructing variables selected based on literature (altitude, ecomorphological class, bioregion, phosphor, nitrogen, velocity and water depth) as raster files of eight river sections, masking each raster for each discharge scenario considered (in the raster\_files folder) and stacking all variables into a raster stack (in the rasters\_stacked folder) using self-written functions from “mth\_functions.R” |
| add\_parameters\_hdm.qmd | Quarto-file calculating additional parameters (Froude, stream power and shear stress) and saving them as raster files (in the raster\_files folder) and stacked rasters (in the rasters\_stacked folder) |
| Folders 🡪 *Hidden in github* | abiotic\_mi\_sampling | Contains all .csv and .txt files used in the Quartos “index\_calculation.qmd” and “prepro\_data.qmd” and result .csv files from the Quartos “index\_calculation.qmd”, “prepro\_data.qmd” and “reconstruct\_discharges.qmd” |
| hdm\_models | Contains all hydrodynamic model files of eight river sections used and resulting from the Quarto “basement\_simulations.qmd” and used in “hydrodynamic\_model\_results.qmd” |
| raster\_files | Contains all raster files represented by a single attribute (velocity, water depth, Froude, altitude etc.) for each discharge of a total of eight river sections |
| rasters\_stacked | Contains all stacked raster files represented by a bundle of attributes (velocity, water depth, Froude, altitude etc.) combined in one raster stack for each discharge of a total of eight river sections |
|  | **Folder: HSC\_model** | **Description** |
| Quarto files | foen\_regr\_data.qmd | Quarto-file obtaining observed and simulated values for a regression analysis in Stat\_Analysis, using a self-written function from “mth\_functions.R” |
| hsc\_plot.qmd | Quarto-file constructing a plot of the HSC for the masters thesis |
| metric\_results.qmd | Quarto-file calculating all evaluation metrics also used for machine learning models |
| habitat\_model\_foen\_qmd | Quarto-file calculating raster files using an univariate HSC from “HSC\_MI”, using a self-written function from “mth\_functions.R” |
| Folders 🡪 *Hidden in github* | HSC\_MI | Contains the .csv file of the HSC and a plot of the HSC used in the masters thesis |
| results\_foen | Contains resulting raster and .csv files from “foen\_regr\_data.qmd” and “habitat\_model\_foen.qmd” |
|  | **Folder: ML\_model** | **Description** |
| Quarto-files | BRT\_model.qmd | Quarto-file containing the whole procedure described in the masters thesis modelling MI indices using BRT algorithms |
| RF\_model.qmd | Quarto-file containing the whole procedure described in the masters thesis modelling MI indices using RF algorithms |
| BRT\_model\_add.qmd | Quarto-file containing the whole procedure described in the masters thesis modelling MI indices using BRT algorithms and added parameters which were excluded due to the principle of parsimony |
| RF\_model\_add.qmd | Quarto-file containing the whole procedure described in the masters thesis modelling MI indices using RF algorithms and added parameters which were excluded due to the principle of parsimony |
| pdp\_plots.qmd | Quarto-file obtaining all PDPs from the best-performing ML model obtained from the workflow |
| Folders 🡪 *Hidden in github* | plots | Contains all PDPs from the best-performing ML model constructed in “pdp\_plots.qmd” |
| var\_imp | Contains all .csv files with variable importance from the best-performing ML model from the workflow |
| pdp\_data | Contains all .csv files to construct PDPs in “pdp\_plots.qmd” from the best-performing ML model from the workflow |
| scenarios | Contains all rasters of predicted MI response variable for each discharge scenario using the best-performing ML model. Additionally, this folder contains the QGIS project constructing spatial maps for the thesis |
|  | **Folder: Stat\_Analysis** | **Description** |
| Quarto-files | hdm\_mi\_analysis.qmd | Quarto-file analysing reconstructed hydraulic conditions using regression analysis and calculating evaluation metrics |
| statistical\_analysis.qmd | Quarto-file conducting regression analyses of all ML models and obtained regression plots |
| plots.qmd | Quarto-file obtaining all plots needed within the Stat\_Analysis folder |
| Folders 🡪 *Hidden in github* | regr\_plots | All regression plots obtained in “hdm\_mi\_analysis.qmd” and “statistical\_analysis.qmd” |
| stat\_analysis\_data | All csv. files resulting from the ML models, “hdm\_mi\_analysis.qmd” and “statistical\_analysis.qmd” |
| plots | All other plots obtained in “plots.qmd” |
| data | All csv. files obtained from raster files and used to obtain boxplots and conduct statistical analysis |