# Documentation GloRiSe – A global database on riverine sediment composition

By Gerrit Müller, Jack J. Middelburg and Appy Sluijs,   
Department of Earth Sciences, Utrecht University

Contact: [g.muller@uu.nl](mailto:g.muller@uu.nl)

*December 2020*

This document summarizes the files belonging to the database, and explains their content and structure. The meaning of abbreviations used in each of the variables and their units within the files can found in the corresponding second sheet of each file. Data is provided in Excel-format (.xlsx) and as a Matlab datafile (.mat) for easy import and processing.

## References

File: “*SedimentDatabase\_ref.xlsx*”

All information on the references of the included observation can be found here. This includes a Source\_ID, that is a 3 digit abbreviation also used in the Sample\_ID (see *ID*), the type of the source (paper, report or book chapter), if the specific reference was peer-review (YES/NO), publication year and access date along with the short (= in-text) citation. Full citation are appended as a bibtex and endnote files (“References\_GloRiSe”).

## Location

File: “*SedimentDatabase\_Location.xlsx*”

The Location sheet summarizes all information on sampling locations, such as geographic coordinates (WGS 84), elevation (where given), country and region/site of the measurement. Where no coordinates were given, all available information in the sources was used to estimate coordinates via © GoogleEarth, which usually can be trusted to precision of at least 20 km. However, this precision finally depends on the quality of the information used, which is the quality of the maps provided for most papers/reports. For spatial averages, the central position is given. A unique 12-digit identifier of a location “Location\_ID”, which is composed of the DIN-ISO-alpha-3 country code (3 digit) indicating the country in which the river mouth is located , a 3-digit abbreviation for the main river basin (here defined as the most downstream river that is a tributary only to the ocean or big inland water masses for endorheic drainage), and a running 6-digit number starting from 111111, can be used to trace all samples to a certain location. Under “Citation” all papers are listed, that provide observation from that same location.   
Each Location was assigned a Basin\_ID, which is a unique numerical identifier of the main river basin. The Basin\_ID is a 4-digit number, where the first digit represents the geographical area, while the other 3-digits represent running numbers. Geographical areas are here loosely and functionally defined as follows: 1) South- and Central America, 2) North America (including parts of southern Canada), 3) Europe (extending from Portugal to the Black Sea and from the Mediterranean to the Scandinavia), 4) East and South-East Asia, including island states, 5) Africa, 6) Australia and New Zealand, 7) Antarctica, and 8) the Arctic region (Northern Canada & Alaska, Greenland, Scandinavia to far-east Siberia (north of the Amur river). Antarctica is still empty, but may be filled in future. Basin\_IDs and main river basin abbreviations are listed in the second sheet of the file. Number of samples per location and the type of data provided are listed in the columns Observations and Content, respectively.

## ID

File: “SedimentDatabase\_ID.xlsx”

All information necessary to identify each unique observation in the database is given in this sheet. Each observation has a unique identifier (“Sample\_ID”), that is composed of the Location\_ID plus the Source\_ID and a running number. Citations and locations are linked to the Sample\_ID in this sheet along with the Original ID (the one used in the original data source), sampling date and time (as precise as possible), notes on the type of observation (from single measurement to spatial and annual averages, where original data was not accessible) and if it represents export to the ocean (“Rep\_ID”). The later provides the possibility to easily isolate samples that can be used to estimate riverine particulate fluxes to the ocean, because it excludes tributaries, far-upstream measurement and such measurement that may be already influence by the marine realm. Date and time of sampling are also available from this table. The closest information on sampling date is given as a string, and day, month and year of these sampling dates are additionally given as separate numerical variables.

## Major Elements, Minor Elements and Nutrients

File: “SedimentDatabase\_ME\_Nut.xlsx”

All measurements of major and minor elements (expressed as Oxide-concentration) and nutrients (C and N in elemental concentration) along with hydro-chemical information (where given) are listed in this file, and are linked to Location\_ID and Sample\_ID. Where given, instantaneous water discharge and suspended sediment concentrations, grain size distribution in %, as well as average grain size (in µm), filter size (as a lower limit of grain size in µm) and sieve size (as an upper limit for grain sizes in µm) are listed, further describing the sample. For traceability, original units from which concentrations were converted and sample type (suspended sediment (SS) or bed sediment including active riverbank deposits (BS)), methods of sample treatment and measurements are given. If multiple methods were used all of them are listed, separated by a “/”. Similarly, multiple units are separated by “,”. Type of observation (single or average), Basin\_ID and Rep\_ID are repeated in this sheet, because they are considered to be the most important grouping variables and may be used for simple calculations without merging all datasets before.

## Trace Elements

File: “SedimentDatabase\_TE.xlsx”

This file contains all information on trace element concentrations (elemental concentrations in ppm) along with Sample\_ID, Location\_ID, Basin\_ID and metadata, i.e. observation type, filter and sieve sizes, sample treatment and methods of measurement, instantaneous water discharge and suspended sediment concentration.

## Minerals

File: “SedimentDatabase\_Minerals.xlsx”

This file lists are information on the mineral (= modal) composition of the sediment samples. Units differ with the method of measurement, i.e. quantitative XRD measurement will be given in wt%, while semi-quantitative (older) XRD measurements will be given in terms of intensity ratios and results from light microscopic investigations are given in area %. As sediments are often composed of lithic fragments to a large fraction, these were also included and grouped. E.g. the variable “L” is the sum of all lithic components, while Lv is the sum of all Lithic clasts that stem from volcanic rocks and Lvf only contains lithic clasts derived from felsic volcanic rocks. The last section (tot(HM), the total Heavy Mineral content, and the following variables) represent heavy mineral contents expressed in % of the total heavy mineral fraction. Thus, the user needs to assure comparability of the used samples and variables using metadata given in the first columns. Similar to the chemical data, SeaCat, observation type, sample type, Basin\_ID, Sample\_ID, and Location\_ID are given along with methods of sample treatment and measurement, and their original unit (which corresponds to the reported unit here).

Feedback and suggestions are welcome, as are potential future contributions!   
Please contact [g.muller@uu.nl](mailto:g.muller@uu.nl)