

Table 1. Column names of the individual dataset tables available in the *Global EPTO Database* and their explanations. The FRED link of each specific dataset is only included in the comprehensive tables of each order and of all orders. (FRED: IGB Freshwater Research and Environmental Database.)

Column name	Column content
order_ID	one of the E, P, T, O
dataset_ID	unique ID of the dataset of a given data source, e.g., 003
occurrence_ID	unique ID of the individual occurrence record, e.g., 000423769
species	species name (not taxonomically verified)
genus	genus name
family	family name
order	order name
longitude	longitude in decimal degrees, in WGS84
latitude	latitude in decimal degrees, in WGS84
no_decimals_published	number of decimal points of the coordinates that are published
no_decimals_available	number of decimal points of the coordinates that are available upon communication with the authors (could include more detailed coordinates)
location_river	locality or river of the observation record
country	country of the observation record
sampling_day	day of the observation
month	month of the observation
year	year of the observation
source	data source (e.g., name of original database)
FRED_link-DOI	dataset link to FRED, where all metadata is available, including dataset authors' names
notes	dataset authors' notes
other_info	additional information available upon communication with the authors, that might include physicochemical or hydrological characteristics of the sampling location, land cover, genus/species abundance or additional taxonomic information such as species and subspecies name

cen_fl	centroid flag, boolean. If '1', then the record is located in the centroid of a country that might be a placeholder for the exact location of the observation
cap_fl	capital flag, boolean. If '1', then the record is located in the capital of a country that might be a placeholder for the exact location of the observation
coun_fl	country flag, boolean. If '1', then the coordinates of the record do not match with the assigned country
snapped	boolean. If '1', then the original coordinates of the record, that fell into the sea, have been snapped to the closest sub-catchment
subcatchment_ID	unique ID of the sub-catchment where the record is located, based on the Hydrography90m (Amatulli et al., 2022)
basin_D	unique ID of the drainage basin where the record is located, based on the Hydrography90m (Amatulli et al., 2022)
elevation	elevation of the point of the record, in meters, based on the MERIT Hydro DEM (Yamazaki et al., 2019)