

Database Infrastructure for Long-Term Data of German Lakes (LakeBase)

Lake Base Data Policy

March 2020

LakeBase was set-up within a project funded by the German Science Foundation in the programme: „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ (LIS): „Informationsstrukturen für Forschungsdaten“. Funding period: 2015-2020.

LakeBase members and affiliations

- Leibniz-Institut für Gewässerökologie und Binnenfischerei (IGB) im Forschungsverbund Berlin e.V. ,Professorin Dr. **Rita Adrian**
- Technische Universität Dresden, Institut für Hydrobiologie, Dr. **Lothar Paul**; Professor Dr. **Thomas U. Berendonk**, seit 12/2016
- Universität Potsdam, Institut für Biochemie und Biologie, Arbeitsgruppe Ökologie und Ökosystemmodellierung, Professorin Dr. **Ursula Gaedke**
- Friedrich-Schiller-Universität Jena, Fakultät für Mathematik und Informatik, Institut für Informatik, Professorin Dr. **Birgitta König-Ries**
- Universität Konstanz, Mathematisch-Naturwissenschaftliche Sektion, Fachbereich Biologie Limnologisches Institut, Professor Dr. **Karl-Otto Rothhaupt**
- Technische Universität Dresden, Institut für Hydrobiologie, Dr. **Annekatrin Wagner**

Continuous long-term records of German lakes and reservoirs are unique worldwide with respect to the length of the observational periods spanning several decades, due to their unique high temporal and taxonomic resolution (mostly phytoplankton, zooplankton, macrozoobenthos, macrophytes and fish) and due to the fact that different lake types with respect to morphometry, trophic state and catchment type are covered – see Table 1. For most lakes a full set of the most important abiotic driving forces are available for more than 30 years. An overview of the lakes and available data and the names of the data holders (contact person) are given in the IGB database infrastructure ‘**Freshwater Research and Environmental Database**’ (FRED).

Data storage:

Data will be securely stored in FRED. The IGB takes responsibility for at least a 10 years storage period till 2029. Should it become impossible to maintain the infrastructure at some point in time after 2029, every attempt to transfer the data to long-term data archives (e.g., the ones organized in the GFBio infrastructure) will be made.

Metadata: Each data set is accompanied by detailed metadata text documentation. It requires information at least about the where, when, how and who in order to be able to upload data.

This includes information on sampling, sample processing, methodology, names of data processors, and the responsible contact person. Metadata information also includes information on quality control routines that have been performed, such as the kind of raw data modifications or quality checks applied (e.g. removal of outliers according to specified criteria). All metadata will be publicly available on **FRED** (<https://fred.igb-berlin.de/Lakebase>) at each data package, regardless of any possible restrictions to access the actual data. It is, however, recommended that prospective data users contact the responsible contact person of each data set for further information.

Data use:

All data are made freely accessible to the LakeBase members (see definition below) for academic, research, educational and similar not-for-profit purposes. LakeBase members have been given a unique password to access all data. Prior to the use of data, data users (LakeBase members, students and scientists associated to the LakeBase members) contact the responsible contact person to agree on the goals and scope of the data use and discuss potential collaborations. Data providers will be given the opportunity to comment on the project plan, methods, analyses, and manuscript preparations. Co-authorship will be offered based on data provision and intellectual input in the writing process. Data must not be passed to a third party without permission of the data providers. Beyond LakeBase members data will be internationally available upon request and permission by the data providers.

In addition, LakeBase data can be made accessible to a broad public with OpenData or Creative Commons licenses.

License name	Brief description	License web address
Open Data Commons Public Domain Dedication and Licence (PDDL)	All rights waived, this means there is no restrictions for the usage of the data.	https://opendatacommons.org/licenses/pddl/1.0/index.html
Open Data Commons Attribution License (ODC-By)	Some rights reserved: Attribution. This means the creator must be mentioned.	https://opendatacommons.org/licenses/by/1.0/index.html
Open Data Commons Open	Some rights reserved: Attribution-ShareAlike for data. This means the	https://opendatacommons.org/licenses/

License name	Brief description	License web address
Database License (ODbL)	creator must be mentioned and derived works must be shared under the same conditions.	odbl/summary/index.html
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)	Some rights reserved: Attribution-NonCommercial-ShareAlike. This means the creator must be mentioned and derived works must be shared under the same conditions. Commercial usage is prohibited.	https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode
Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)	Some rights reserved: Attribution-NonCommercial-NoDerivatives. This means the creator must be mentioned. Redistribution and sharing copies in any medium or format is allowed. Distributions of transformations, remixes or build upon works and commercial usage are prohibited.	https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode

DOI:

Lakebase data are provided with a DOI, the allocation and maintenance of the versions is the responsibility of the Lakebase members.

Disclaimer:

While substantial efforts are made to ensure the accuracy of all data and the metadata documentation, complete accuracy cannot be guaranteed. It is the responsibility of each data user to assess the data quality. All data are made available "as is." Data users should also be aware that data sets are periodically updated. Data providers shall not be liable for any damages or other consequences resulting from any use or interpretation of the data sets.

In this policy the following definitions are used:

(Primary) Data:

All measurements and observations obtained for the lakes within LakeBase

Metadata:

Descriptive information about primary data, necessary to interpret, use, and disclose data

LakeBase members

Rita Adrian, Ursula Gaedke, Mark Gessner, Peter Kasprzak, Lothar Paul, Thomas U. Behrendonk, Annkatrin Wagner, Karl-Otto Rothhaupt.

Data owner:

The data owner is the University or research institution for which the data are collected, or in some cases the contact person, who is mentioned related to each data package. The data owners are represented by the LakeBase members who are listed as the contact person in the database. The data owner has to ensure that other people associated with the data (e.g. PhD-students or scientists who work with the data) are informed about the data policy rules with LakeBase.

Data provider:

The person uploading data to the LakeBase database. In our case, the LakeBase members.

Data manager:

Eric Wündisch, Simone Frenzel (employed through the DFG LakeBase Project; located at IGB) and Daniel Langenhaun (Data Manager at IGB)

Variables Lake (trophic state)	Phy	Nut	Phyto	Zoo	MEV	Fish	Map h	Catchment type	Time period	Contact Person
Stechlinsee (mesotrophic)	x	x	x	x	(x)	(x)	(x)	forest	1958/70 ongoing	Mark Gessner HP Kasprzak (IGB)
Müggelsee (eutrophic)	x	X	x	x	(x)	-	(x)	Agric., forest, urban	1975-ongoing	Rita Adrian (IGB)
Talsperre Neunzehnhain (oligotrophic)	x	x	x	x	-	-	-	Forest, agric.	1992-ongoing	Lothar Paul (TU-Dresden)
Saidenbach Talsperre (mesotrophic)	X	X	X	X	(x)	(x)	-	Agric., forest, urban	1975-ongoing	Lothar Paul (TU-Dresden)
Talsperre Bautzen (eu-/hypertrophic)	X	X	X	X	x	(x)	-	Agric., urban	1977-1999	Annekatrin Wagner (TU-Dresden)
Bodensee (eu/mesotrophic)	x	x	x	x				Alpine, forest, agric.	1979-1998	U. Gaedke (Uni Potsdam) KO. Rothhaupt (Uni Konstanz)

Table 1: Available decadal time series of LakeBase lakes include lakes of different morphometry, trophic state and catchment properties. The most common measured variables are: Physical variables (water temperature, ice, oxygen, pH, conductivity; Phy), nutrients (Nut) (combined as abiotic in data volume), phyto- and zooplankton (Phyto, Zoo), macro-invertebrates (MEV), macrophytes (Macroph) fish. X: continuous time series; (x): multiple single observations; - no observations. Data providers are given under 'Contact person'.