

# Metadata for *Time span in anaerobic condition predicts biogeochemical consequences of oxygen depletion in lakes*

**Table 1.** Description of the fields needed to describe the creation of your dataset.

<b>Title of dataset</b>	<i>Dataset used in 'Time span in anaerobic condition predicts biogeochemical consequences of oxygen depletion in lakes' by LaBrie, Hupfer and Lau</i>
<b>URL of dataset</b>	<a href="https://fred.igb-berlin.de/data/package/819">https://fred.igb-berlin.de/data/package/819</a>
<b>Abstract</b>	<i>In this study we developed a novel metric, anaerobic duration, to describe anoxia in aquatic ecosystems. Anaerobic duration can be used to predict nutrients concentration (NH<sub>4</sub><sup>+</sup>, SRP, TP) and a fluorescent dissolved organic matter (FDOM) component, and likely many reduced compounds such as Fe, Mn, CH<sub>4</sub> and H<sub>2</sub>S.  We used dissolved oxygen (DO) data (profiles and loggers), nutrients (NH<sub>4</sub><sup>+</sup>, SRP and TP) and a FDOM component</i>
<b>Keywords</b>	<i>Anoxia, hypolimnion, lakes, anaerobic duration, oxygen, nutrients</i>
<b>Lead author for the dataset</b>	<i>Richard LaBrie</i>
<b>Title and position of lead author</b>	<i>Post-doctoral researcher</i>
<b>Organization and address of lead author</b>	<i>Interdisciplinary Environmental Research Centre, TU Bergakademie Freiberg, Akademiestraße 6, 09599 Freiberg, Germany</i>
<b>Email address of lead author</b>	<i>Richard.labrie90@gmail.com</i>
<b>Additional authors or contributors to the dataset</b>	<i>Michael Hupfer, Maximilian P. Lau</i>
<b>Organization associated with the data</b>	<i>Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 310, 12587 Berlin, Germany</i>
<b>Funding</b>	<i>Maximilian Lau, State Agency for Flood Protection and Water Management Saxony-Anhalt (LHW)</i>
<b>License</b>	<i>CCBY</i>
<b>Geographic location – verbal description</b>	<i>Lake Arendsee, Germany Lake Delavan, USA Lake Mendota, USA Lake Stechlin, Germany</i>
<b>Geographic coverage bounding coordinates</b>	<i>Arendsee (52°53'21"N 11°28'27"E) Delavan (42°36'18"N, 88°36'13"W) Mendota (43°06'24"N, 89°25'29"W) Stechlin (53°9'8"N, 13°1'24"E)</i>
<b>Time frame - Begin date</b>	<i>1984-01-17</i>
<b>Time frame - End date</b>	<i>2021-12-08</i>
<b>General study design</b>	<i>This study is based on multiparameter profilers monitoring data on lakes Arendsee, Delavan, Stechlin and Mendota and on DO loggers deployed in lake Arendsee at 30, 35, 40, 45 and 47m deep. Using the DO data, we developed a new metric, anaerobic duration, to describe and quantify anoxia in lakes hypolimnia. This objective was to relate this novel metric to</i>

	<i>redox sensitive compounds (NH<sub>4</sub><sup>+</sup>, SRP, TP) and with FDOM to better understand and predict the consequences of anoxia.</i>
<b>Methods description</b>	<p><i>Lake Arendsee was sampled using autonomous multiparameter profiler (YSI, daily) and continuous dissolved oxygen loggers (hourly).</i></p> <p><i>Lake Delavan was sampled using a multiparameter profiler (YSI) (Jane et al. 2021)</i></p> <p><i>Lake Mendota was sampled using a multiparameter profiler (YSI Exo2), which includes a fluorescent dissolved organic matter (FDOM) probe (Magnuson et al. 2021, <a href="https://doi.org/10.6073/pasta/5f15bf453851987fc030b2f07a110b21">https://doi.org/10.6073/pasta/5f15bf453851987fc030b2f07a110b21</a>) and its hypsography was available online (<a href="http://www.bathybase.org/Data/100-199/100/">http://www.bathybase.org/Data/100-199/100/</a>)</i></p> <p><i>Lake Stechlin was sampled using a multiparameter profiler (YSI) bi-weekly in summer and monthly in winter (<a href="https://fred.igb-berlin.de/data/package/617">https://fred.igb-berlin.de/data/package/617</a>)</i></p>
<b>Laboratory, field, or other analytical methods</b>	<p><i>Ammonium and soluble reactive phosphorus quantification (Scan++, Skalar Analytical, Netherlands) using indophenol (Bolleter et al. 1961) and molybdenum blue (Murphy and Riley 1962).</i></p> <p><i>Lake Delavan TP measurements were taken from Jane et al. 2021</i></p>
<b>Taxonomic species or groups</b>	NA
<b>Quality control</b>	<i>Most data used in this study comes from published datasets that were already quality controlled. When values were flagged as suspicious, we removed them to avoid biases.</i>
<b>Additional information</b>	

## Table 2. Data dictionary

### Oxygen files

Dataset filename: Arendsee.DO-YSI-2017.csv

Dataset description: Oxygen data from profiler casts (excerpt from longer dataset)

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth_m	Depth where oxygen measurements were taken	m		values	No missing values
All other columns, written as dates	Oxygen values	mg L <sup>-1</sup>	Column name explanation: yyyy-mm-dd-HH-mm-ss	values	NA

Dataset filename: Arendsee.DO-YSI-2018.csv

Dataset description: Oxygen data from profiler casts (excerpt from longer dataset)

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth_m	Depth where oxygen measurements were taken	m		values	No missing values
All other columns, written as dates	Oxygen values	mg L <sup>-1</sup>	Column name explanation: yyyy-mm-dd-HH-mm-ss	values	NA

Dataset filename: Arendsee.DO-YSI-2019.csv

Dataset description: Oxygen data from profiler casts (excerpt from longer dataset)

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth_m	Depth where oxygen measurements were taken	m		values	No missing values
All other columns, written as dates	Oxygen values	mg L <sup>-1</sup>	Column name explanation: yyyy-mm-dd-HH-mm-ss	values	NA

Dataset filename: Arendsee.DO-YSI-2020.csv

Dataset description: Oxygen data from profiler casts (excerpt from longer dataset)

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth_m	Depth where oxygen measurements were taken	m		values	No missing values
All other columns, written as dates	Oxygen values	mg L <sup>-1</sup>	Column name explanation: yyyy-mm-dd-HH-mm-ss	values	NA

Dataset filename: Arendsee.DO-YSI-2021.csv

Dataset description: Oxygen data from profiler casts (excerpt from longer dataset)

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth_m	Depth where oxygen measurements were taken	m		values	No missing values
All other columns, written as dates	Oxygen values	mg L <sup>-1</sup>	Column name explanation: yyyy-mm-dd-HH-mm-ss	values	NA

Dataset filename: Arendsee.DO-logger.csv

Dataset description: Oxygen data from loggers, long format (excerpt from longer dataset)

Column name	Description	Units	Code explanation	Data format	Missing data code
UTC	Date and time of measurement			yyyy-mm-dd-HH:mm:ss	
CET	Date and time of measurement			yyyy-mm-dd-HH:mm:ss	
value	parameter value	mg L <sup>-1</sup> or °C	Column name explanation: yyyy-mm-dd-HH-mm-ss		
parameter	oxygen concentration or temperature				
depth	Loggers depths	m			
year	Year of measurement			YYYY	
mm_dd	Month and day of measurement			mm-dd	
dd	Day of measurement			dd	

Dataset filename: DO\_FDOM.Mendota.csv

Dataset description: Oxygen data from profiler casts and FDOM data

Column name	Description	Units	Code explanation	Data format	Missing data code
Sampledate	Date of measurement	NA		yyyy-mm-ss	
sampletime	Time of measurement	NA		HH:mm:ss	
depth	Depth of measurement	m			
do_raw	Oxygen concentration	mg L <sup>-1</sup>			blank
do_sat	Oxygen saturation	%			blank
fdom	FDOM component at	QSU			blank

	excitation = 365 nm and emission = 480 nm				
flag_do_raw	Flags of bad oxygen data		C is used to flag missing data. These will be removed in the R script		

Dataset filename: Delavan.DO.csv

Dataset description: Oxygen data from profiler casts

Column name	Description	Units	Code explanation	Data format	Missing data code
row_id	Unique identifier from source file	NA			
lake_id	Unique identifier from source file	NA			
date	Date of measurement	NA		yyyy-mm-dd	
depth	Depth of measurement	m			
do_con	Dissolved oxygen concentration	mg L <sup>-1</sup>			
year	Year of sampling	NA		YYYY	

Dataset filename: Stechlin.DO.csv

Dataset description: Oxygen data from profiler casts (excerpt from longer dataset)

Column name	Description	Units	Code explanation	Data format	Missing data code
Date	Date of measurement	NA		yyyy-mm-dd	
Depth	Depth of measurement	m			
DO_mgL	Dissolved oxygen concentration	mg L <sup>-1</sup>			

## Bathymetry and alpha(z) files

Dataset filename: Arendsee.alpha.csv

Dataset description: Sediment to volume ratio values for lake Mendota

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth_1m	Depth	m			
alpha	Sediment to volume ratio	m <sup>-1</sup>			

Dataset filename: Hypso\_Mendota.tsv

Dataset description: Hypsography taken from bathybase.org

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth (m)	Depth	m			
Area (m <sup>2</sup> )	Area of the stratum at	m <sup>2</sup>			

	specified depth				
--	-----------------	--	--	--	--

Dataset filename: Mendota.alpha.csv

Dataset description: Sediment to volume ratio values for lake Mendota

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth	Depth	m			
alpha	Sediment to volume ratio	m <sup>-1</sup>			

Dataset filename: Delavan.alpha.csv

Dataset description: Sediment to volume ratio values for lake Delavan

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth	Depth	m			
alpha	Sediment to volume ratio	m <sup>-1</sup>			

Dataset filename: Stechlin.alpha.csv

Dataset description: Sediment to volume ratio values for lake Stechlin

Column name	Description	Units	Code explanation	Data format	Missing data code
Depth	Depth	m			
alpha	Sediment to volume ratio	m <sup>-1</sup>			

## Nutrient files

Dataset filename: SRP-NH4.Arendsee.2017-2019.csv

Dataset description: Soluble reactive phosphorus and ammonium file for lake Arendsee, 2017 and 2019

Column name	Description	Units	Code explanation	Data format	Missing data code
Date	Date of measurements			yyyy-mm-dd	
Depth_m	Depth of measurements	m			
SRP_mgL	Soluble reactive phosphorus concentration	mg L <sup>-1</sup>			NA
NH4-N_mgL	Ammonium concentration	mg L <sup>-1</sup>			NA

Dataset filename: SRP-NH4.Arendsee.2020.csv

Dataset description: Soluble reactive phosphorus and ammonium file for lake Arendsee, 2020

Column name	Description	Units	Code explanation	Data format	Missing data code
date	Date of measurements			yyyy-mm-dd	
depth	Depth of	m			

	measurements				
SRP_mgL	Soluble reactive phosphorus concentration	mg L <sup>-1</sup>			NA
NH4_mgL	Ammonium concentration	mg L <sup>-1</sup>			NA

Dataset filename: TP.Delavan.csv

Dataset description: Nutrient data for lake Delavan

Column name	Description	Units	Code explanation	Data format	Missing data code
row_id	Unique identifier from source file	NA			
lake_id	Unique identifier from source file	NA			
date	Date of measurement	NA		yyyy-mm-dd	
depth	Depth of measurement	m			
tp_ugl	Total phosphorus concentration	µg L <sup>-1</sup>			NA

Dataset filename: SRP-NH4.Stechlin.csv

Dataset description: Nutrient data for lake Stechlin (excerpt from longer dataset)

Column name	Description	Units	Code explanation	Data format	Missing data code
Date	Date of measurement	NA		yyyy-mm-dd	
Depth	Depth of measurement	m			
SRP_mgL	Soluble reactive phosphorus concentration	mg L <sup>-1</sup>			
NH4_mgL	Ammonium concentration	mg L <sup>-1</sup>			

### Table 3. Data provenance

If you used data derived from other sources, provide the information here so future users know where the data came from.

Dataset title	Dataset DOI or URL	Creator (name & email)	Contact (name & email)
Hypso_Mendota.tsv	<a href="http://www.bathybase.org/Data/100-199/100/">http://www.bathybase.org/Data/100-199/100/</a> <a href="https://dnr.wi.gov/lakes/Documents/LakeMaps.aspx">https://dnr.wi.gov/lakes/Documents/LakeMaps.aspx</a>	NA	NA
DO_FDOM.Mendota.csv	<a href="https://lter.limnology.wisc.edu/lake_mendota_multiparameter_so">https://lter.limnology.wisc.edu/lake_mendota_multiparameter_so</a>	J.J. Magnuson S.R. Carpenter	John J. Magnuson john.magnuson@wisc.

	nde_profiles	E.H. Stanley	edu
Delavan.DO.csv TP.Delavan.csv	<a href="https://www.nature.com/articles/s41586-021-03550-y#Sec20">https://www.nature.com/articles/s41586-021-03550-y#Sec20</a>	Stephen F. Jane 43 co-authors	Kevin C. Rose (rosek4@rpi.edu)
Stechlin.DO.csv SRP-NH4.Stechlin.csv	excerpts from long term dataset: <a href="https://fred.igb-berlin.de/data/package/617">https://fred.igb-berlin.de/data/package/617</a>	Sabine Wollrab	sabine.wollrab@igb-berlin.de
Arendsee.DO-YSI-YYYY.csv	excerpts from long term dataset: <a href="https://fred.igb-berlin.de/data/package/168">https://fred.igb-berlin.de/data/package/168</a>	Sylvia Jordan	sylvia.jordan@igb-berlin.de
Arendsee.DO-logger.csv	excerpts from long term dataset: <a href="https://fred.igb-berlin.de/data/package/628">https://fred.igb-berlin.de/data/package/628</a>	Sylvia Jordan	sylvia.jordan@igb-berlin.de

## Scripts/code (software)

File name	Description	Scripting language
Main.Rmd	Main script used in this study. Will provide most of the figures, but not all	R
O2_DecayRate.R	Internal function to calculate oxygen consumption rates per stratum. Used in main script	R
Matrix_Inter.R	Internal function to interpolate oxygen concentration when dates are missing	R
O2_Model.R	Internal function to calculate oxygen consumption rates per stratum. Used in main script	R

Note. The scripts provided here will only reproduce figures for lake Arendsee.