Guide to LAGOS-US DEPTH v1.0: Data module of observed maximum and mean lake depths for a subset of lakes in the conterminous U.S.

Citation

Jemma Stachelek, Lauren K. Rodriguez, Jessica Díaz Vázquez, Arika Hawkins, Ellie Phillips, Allie Shoffner, Ian M. McCullough, Katelyn B.S. King, Jake Namovich, Lindsie A. Egedy, Maggie Haite, Patrick J. Hanly, Katherine E. Webster, Kendra S. Cheruvelil, Patricia A. Soranno. 2021. LAGOS-US DEPTH v1.0: Data module of observed maximum and mean lake depths for a subset of lakes in the conterminous U.S. Environmental Data Initiative. [insert data package doi]. Dataset accessed [insert date].

Data creation and documentation contribution

JS organized and led the effort; LKR linked the lake polygons in GIS to help with identifying the location of lakes; JDV, AH, EP, AS, IMM, KBSK, JN, LAE, and MH conducted online searches for lake depth data; JS, PJH, IMM, KEW, KSC, and PAS wrote and revised the User Guide; JS, MH, KEW, IMM, and PAS performed QA/QC checks on depth data; and KSC and PAS conceived of the data module, created and managed the research team, and provided oversight for the data and documentation.

Acknowledgements

Additional thanks to the Continental Limnology research team and the following individuals who provided additional help: Claire Boudreau, Noah Lottig, and Natalie Schmer.

Table of Contents

Citation.....i
Data creation and documentation contribution.....i
Acknowledgements.....i

Table of Contents..... ii

List of Figures iii

List of Tables iii

List of Appendices iii

Overview of the LAGOS-US research platform..... iv

Description of LAGOS-US DEPTH Version 1.0 1

1. Database Design for LAGOS-US DEPTH 1

1.1 Definition of terms relating to entities 1
1.1.1. Lake 1
1.1.2. Entity identifiers 2
1.2. Overview of data tables and variables..... 2
1.2.1. lake_depth table 2
1.3. Module data schema 2
1.3.1. Data table (data) 2
1.3.2. Data dictionary (metadata)..... 2

2. Data in LAGOS-US DEPTH 3

2.1. Data sources 3
2.2. Metadata tables and variables 3
2.2.1 data_dictionary_depth table..... 3
2.3. Data tables and variables 3
2.3.1. lake_depth table 3
2.4. Variable naming conventions 3
2.5. Data dictionary..... 4

3. METHODS for LAGOS-US DEPTH.....7

4. Quality Control/Quality Assurance 9

5. Data limitations.....9

6. References..... 10

Appendices 11

List of Figures

Figure i. The LAGOS-US Platform. v

List of Tables

Table 1. Select columns from the DEPTH data_dictionary_depth table. 4

Table 2. The lake_depth table data dictionary. 6

List of Appendices

Appendix 1. Number of lakes with maximum depth data by state based on the state postal abbreviation used in the depth search. 11

Appendix 2. Map of the 17,675 lakes in the conterminous U.S. with depth data in LAGOS-US DEPTH. 11

Overview of the LAGOS-US research platform

LAGOS-US provides an extensible research-ready platform to study the 479,950 lakes and reservoirs larger than or equal to 1 ha in the conterminous U.S. at multiple scales of space and time and at broad spatial extents. Although lakes are the focal unit of study, studying land-water interactions requires not only *in situ* lake water quality measurements, but also descriptions of the lakes, their watersheds, and their landscape ecological context (i.e., the land use, geologic, climatic, and hydrologic setting of lakes). Each lake's ecological context can be characterized at a variety of spatial extents (e.g., ecoregions, watersheds), which we call spatial divisions. Some of these ecological context variables are relatively static through time and are therefore characterized for a single date, whereas others are dynamic through time and are characterized at multiple time steps. Whenever possible, we include data for all lakes which can be used as the 'census' population of lakes $\geq$ 1 ha in the conterminous U.S.

The LAGOS-US research platform includes three core data modules for lakes $\geq$ 1 ha in surface area and located in the conterminous U.S.:

1) **LAGOS-US LOCUS** for locational, identifying, and physical information of lakes and their watersheds (Smith et al. 2021, Cheruvelil et al. 2021),

2) **LAGOS-US GEO** for geospatial and temporal ecological context variables (e.g., land use, climate, hydrology) for lakes characterized at multiple spatial divisions (e.g., equidistant buffers around lakes, watersheds, ecoregions),

3) **LAGOS-US LIMNO** for the subset of lakes with *in situ* limnological physical, chemical, and biological measurements located in the 31 states and Washington D.C. not included in LAGOS-NE LIMNO (Soranno et al. 2017).

These LAGOS-US core modules were created using existing datasets from a variety of data sources, such as national-scale climate, land use/cover, and hydrology data, as well as government, tribal, and community science lake data. In building this research platform, we followed a set of three fundamental principles that are similar to those used to create LAGOS-NE, an earlier version of the database system for a subset of 17 U.S. states (Soranno et al. 2015, 2017). The first principle is that LAGOS-US should be based on a foundation of open science by which we make our data publicly available when each module is completed, error-checked, and documented. This includes providing a permanent identifier and a versioning system to facilitate future reuse of the data. Second, we document and describe the original data sources, our methods for integrating data, possible errors that may exist in the data, and we provide code for such methods, when possible (i.e. detailed User Guides). Third, we preserve the provenance of the original data as much as possible.

The LAGOS-US research platform was designed to be modular, i.e., each data module is made of variables that were derived using similar methods or data sources. This modularity facilitates documentation of the entire database and makes the data tables of manageable size. In addition, LAGOS-US was designed to be easily extensible, i.e., to allow other users to build extension modules that can be easily integrated into the LAGOS-US research platform. Future extension modules will be able to connect to any core module of LAGOS-US through common identifiers.

There are currently four LAGOS-US extension modules either published or in development by members of our research team that will connect to LAGOS-US LOCUS through common lake identifiers: RESERVOIR provides a predicted classification of all 137,465 lakes ≥ 4 ha as either a natural lake or a reservoir using a machine-learning algorithm and aerial imagery; LAKE DEPTH includes mean and/or maximum depth measurements of over 17,000 lakes ≥ 1 ha that were manually compiled from a wide range of online sources (Stachelek et al. in press); NETWORKS uses graph theory to identify 898 lake networks that include 86,511 lakes ≥ 1ha and provide quantitative surface water connectivity metrics for those networks and lakes (King et al. 2021a and b); and LANDSAT provides predicted water quality measurements for chlorophyll *a*, Secchi depth, and colored dissolved organic matter for all lakes ≥ 4 ha using machine-learning models based on atmospherically corrected Landsat imagery and LIMNO data, in addition to lakewide values of reflectance for each Landsat band and satellite overpass.

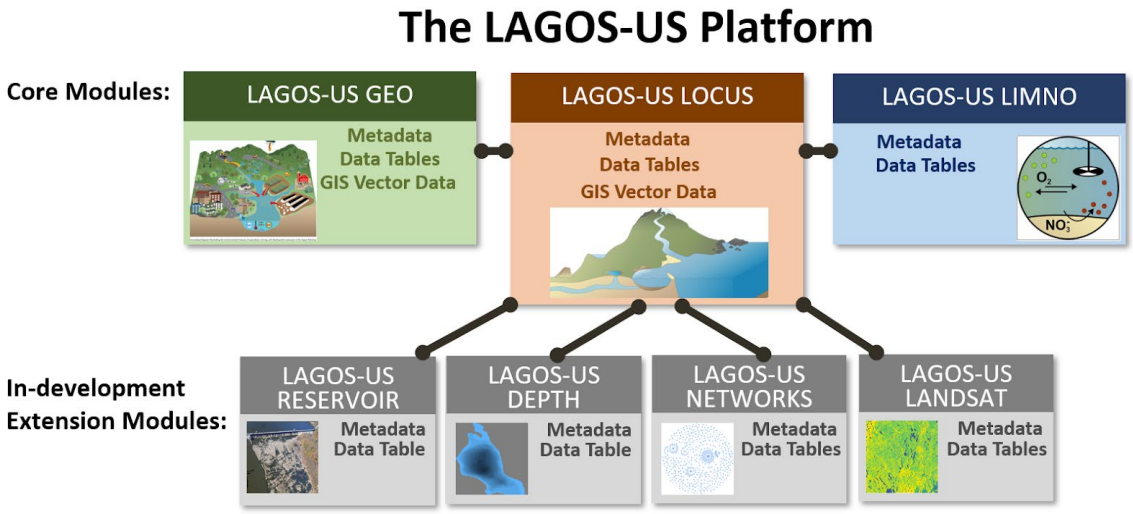


Figure i. The LAGOS-US Platform.
The platform includes three core LAGOS-US modules: GEO, LOCUS, and LIMNO. Four existing extension modules are also shown in grey. The major categories of variables within each module are shown in bold. Each represents a separate table in the database (except for observation-level flags, which are variables embedded directly in data tables). See text for the descriptions of each module. An R package, called 'LAGOSUS', is also available to facilitate access. LANDSAT image courtesy of the U.S. Geological Survey and images for GEO, LOCUS, LIMNO, and RESERVOIR are from the Integration and Application Network, University of Maryland Center for Environmental Science (https://ian.umces.edu/imagelibrary/).

Description of LAGOS-US DEPTH Version 1.0

The LAGOS-US DEPTH v1.0 module (hereafter, called DEPTH) contains in situ measurements of lake depth for a subset of all lakes (n = 17,675; Appendix Figure 1) in the conterminous US ≥ 1 ha (3.7% of 479,950) that are in the LAGOS-US LOCUS v1.0 data module (Cheruvelil et al. 2021, Smith et al. 2021). All 17,675 lakes in the DEPTH module have a maximum depth and 6,137 have a mean depth. DEPTH includes 65 sources of lake depth obtained from community, government, and university monitoring programs, as well as academic reports and commercial websites. The DEPTH module includes lake identifiers, lake location, lake area, lake depth (both maximum and mean depth when available), and source information. The unique lake identifier for all lakes, lagoslakeid, is the same one used in LAGOS-US LOCUS v1.0.

Because lake depths had to be searched for manually to build this data module, we needed to prioritize the numbers of lakes to search. Therefore, the starting population of lakes used to create this data module only included those that either had an existing water quality sampling site (and, thus, assumed to also have observed water quality data) or known water quality data. Specifically, we included lakes that had limnological data in the LAGOS-NE research platform (Soranno et al. 2017), the Water Quality Portal (WQP), the 2007 and 2012 U.S. National Lakes Assessment, and a lake bathymetric study of 9 states (Stachelek et al., In Press). Therefore, the depths reported here should not be considered a comprehensive search of depths for all conterminous U.S. lakes.

1. Database Design for LAGOS-US DEPTH

1.1 Definition of terms relating to entities

1.1.1. Lake

Definition: A 'lake' in LAGOS-US is a perennial body of relatively still water. We include lakes and reservoirs that range from being completely natural to highly modified: lake basins can be completely natural, modified natural (i.e., a water control structure on a natural lake), or a fully impounded stream or river (i.e., a reservoir, impoundment, or other). We explicitly exclude the following water bodies that are coded in the National Hydrography Dataset, High Resolution (NHD HR; U.S. Geological Survey 2017) data source as lakes: sewage treatment ponds, aquaculture ponds, or other such detention ponds that are known to contain basins that are entirely artificial and were built for high-intensity human use. The operational minimum lake size in our database is 1 ha (n = 479,950). This definition of 'lake' has been developed for the purpose of the LAGOS-US database and its applications (e.g., to study lakes at macroscales) and is based on geographic representations of lakes from the NHD HR snapshot compiled at the beginning of the project (Smith et al. 2021). Lakes are represented as polygons for almost all spatial operations leading to derived variables in LAGOS-US. A central point for the lake (the label point, rather than the centroid is used to guarantee the point falls within the lake) is used to assign exclusive membership to one state, county, or zone.

1.1.2. Entity identifiers

Each lake in the LAGOS-US LOCUS data product is assigned a unique, non-nullable identifier, lagoslakeid. For most lakes, lagoslakeid in the LAGOS-NE database platform is the same in LAGOS-US, and lagoslakeids have never been re-assigned to a new lake. However, because lagoslakeid is based on the high resolution NHD data product that is frequently updated, there have been some changes to a few lake lagoslakeids between LAGOS-NE and LAGOS-US due to changes in how lakes are represented in the NHD. Therefore, in LAGOS-US LOCUS, the *lake_link* table provides a crosswalk between these two datasets for those cases where they differ, along with the lake identifiers from several other data products (see Smith et al. 2021).

1.2. Overview of data tables and variables

The variables in DEPTH are organized into a single machine-readable comma separated value (CSV) data table, which includes lagoslakeid as a common identifier:

1.2.1. *lake_depth* table

Includes depth metrics for each lake (e.g., mean and max), derived metrics of lake geometry (e.g., area) and information about lake location (i.e., latitude, longitude).

1.3. Module data schema

The DEPTH module consists of a data table and a metadata table.

1.3.1. Data table (data)

Contains the observations of the variables. This is a single data table that can be linked through the common LAGOS-US LOCUS identifier lagoslakeid. This table contains only a single lake depth value *(see § 3)* for each lake.

1.3.2. Data dictionary (metadata)

Provides a definition for each variable name or 'column' of every table in the module, and includes other useful information such as units.

2. Data in LAGOS-US DEPTH

2.1. Data sources

Where available, the source for each lake maximum depth value is provided in the data table. Because there are over 65 individual sources, we do not list them here, but instead provide them in the *lake_depth* data table itself.

2.2. Metadata tables and variables

2.2.1 *data_dictionary_depth* table

Definition: A table that contains variable names, description, and units for each column in the *lake_depth* table.

2.3. Data tables and variables

2.3.1. *lake_depth* table

Definition: A table that contains maximum depth for all lakes and mean depth for a subset of lakes, source data, effort type, derived metrics of lake geometry (e.g., area) and information about lake location (i.e., latitude, longitude, state).

2.4. Variable naming conventions

Where applicable, we used similar naming conventions to those in LAGOS-US LOCUS v1.0 such as variable names being all lowercase and parts of names separated by an underscore ('_'). In addition, we use the same lake unique identifier (lagoslakeid) as in LAGOS-US LOCUS v1.0 to allow the DEPTH data to be integrated with other LAGOS-US data products.

2.5. Data dictionary

This section provides a summary of variables and their descriptions for the metadata and data tables. For more complete definitions, consult the built-in metadata table *data_dictionary_depth*.

Table 1. Select columns from the DEPTH *data_dictionary_depth* table.
This table depicts the structure of the data within the data dictionary. These variables are unitless.

variable_name	variable_description	taxonomy_type	missing_values
table_name	the LAGOS-US table name.	metadata	N
variable_name	the LAGOS-US variable name (column name).	metadata	N
variable_description	the variable description/definition.	metadata	N
variable_name_group	The variable name group across spatial divisions. This value is prefixed with the division name to generate the full variable name.	metadata	N
taxonomy_type	the LAGOS-US LOCUS/GEO variable taxonomy type defined for the variable.	metadata	N
taxonomy_division	the spatial division prefix for the variable.	metadata	N
taxonomy_main_feature	the "main feature" category for the variable name according to the LAGOS-US LOCUS/GEO variable taxonomy.	metadata	N
taxonomy_subgroup	the "subgroup" category for the variable name according to the LAGOS-US LOCUS/GEO variable taxonomy.	metadata	Y
taxonomy_units	the units suffix for the variable name.	metadata	Y
units	the full description of the units used for the variable.	metadata	Y
missing_values	indicates whether the variable has missing values.	metadata	N
in_lagosne	indicates whether an equivalent variable was included in LAGOS-NE. Variables were re-named and re-organized in between the LAGOS-NE and LAGOS-US versions.	metadata	N

variable_name	variable_description	taxonomy_type	missing_values
lagosne_table_name	the table name where the equivalent variable was located in LAGOS-NE, if applicable. Variables were re-named and re-organized in between the LAGOS-NE and LAGOS-US versions.	metadata	Y
lagosne_variable_name	the variable name used for the equivalent variable in LAGOS-NE. Variables were re-named and re-organized in between the LAGOS-NE and LAGOS-US versions.	metadata	Y
variable_source_code1	the source_code used in source_table of the first or only source dataset used to derive this variable for LAGOS-US.	metadata	Y
variable_source_code2	the source_code used in source_table of the second source dataset used to derive this variable for LAGOS-US.	metadata	Y
methods_tool	the script function used to generate the variable. Functions may be found in ArcGIS (arcpy), LAGOS GIS Toolbox, or other LAGOS code. "Hand" indicates manual/non-reproducible creation and "copy" indicates values were copied from source data unaltered.	metadata	Y
source_value_code	the original source dataset value that was re-coded to name this variable in LAGOS-US. For example, the land cover category named "openwater11" in LAGOS-US was assigned the pixel value 11 in the source dataset (NLCD).	metadata	Y
data_type	the type of data (R datatype) stored in the variable. Domain values are char, factor, int, and numeric	metadata	N
precision	the number of decimal places provided for numeric variables (0 for integers), or the maximum number of characters for character (text/categorical) variables.	metadata	N
domain	the permissible values of the variable, separated by semicolons.	metadata	Y

Table 2. The *lake_depth* table data dictionary.
Where the units column is blank, variables are unitless.

Variable	Description	Units
lagoslakeid	unique lake identifier developed for LAGOS-US LOCUS v1.0	
lake_namegnis	lake name from the GNIS database by way of the LAGOS-US LOCUS v1.0 module	
lake_states	two-letter postal abbreviation(s) of state(s) intersecting the lake polygon from the LAGOS-US LOCUS v1.0 module	
lake_depth_state	two-letter postal abbreviation of the state used to search for a lake's depth	
lake_lat_decdeg	the latitude of the lake center point (NAD83) from the LAGOS-US LOCUS v1.0 module	decimal degree
lake_lon_decdeg	the longitude of the lake center point (NAD83) from the LAGOS-US LOCUS v1.0 module	decimal degree
lake_maxdepth_m	lake maximum depth	meter
lake_meandepth_m	lake mean depth	meter
lake_waterarea_ha	surface area of lake waterbody polygon from NHD (excluding islands) by way of the LAGOS-US LOCUS v1.0 module	hectare
lake_depth_sourcename	name of the source of lake depth data	
lake_depth_sourceurl	url link to the source of lake depth data	
lake_maxdepth_effort	name representing one of six different maximum depth searching efforts (used for quality assurance and de-duplication); one of LAGOSNE, LAGOSUS, bathymetry, or NLA	
lake_meandepth_effort	name representing one of four different mean depth searching efforts (used for quality assurance and de-duplication); one of LAGOSNE, LAGOSNE_Q500, LAGOSUS, LAGOSUS_Q500	

3. METHODS for LAGOS-US DEPTH

This section describes the methods used to create each variable in LAGOS-US DEPTH.

lagoslakeid was obtained from the LAGOS-US LOCUS v1.0 database. It was calculated by generating sequential integers for each lake maintained in the LAGOS-US LOCUS v1.0 lake population. LAGOS-NE identifiers were re-used when the NHD Permanent_Identifier used from the NE snapshot matched the Permanent_Identifier from the US snapshot (the most common case). LAGOS-NE identifiers, when not re-used for the same lake due to identifier changes over time, were never recycled and used for another lake (*see § 1.1*).

lake_states was obtained using the ArcGIS Spatial Join operation to identify all states that intersect the lake polygon. All intersecting states were identified, abbreviated, and merged into a single value (separated by semicolons).

lake_state is the nominal state that a lake is located in that was used to search for lake depths on the internet as part of lake_depth_effort LAGOSUS (see below).

lake_lat_decdeg, lake_lon_decdeg are coordinates calculated with the Calculate Geometry tool in ArcGIS using the NAD83 geographic coordinate system in units of decimal degrees, based on the point representation of the lakes. The point representations are always within the lakes (a central "labeling" location), and are not based on the true centroid of the lake.

lake_maxdepth_m, lake_meandepth_m, sourcename_depth, sourceurl_depth values were generated differently depending on the value of lake_maxdepth_effort and lake_meandepth_effort.

a) The values associated with lake_maxdepth_effort and lake_meandepth_effort LAGOSNE came from a variety of sources including internet searches and direct government agency requests. No source data are available from this effort.
b) The values for lake_maxdepth_effort and lake_meandepth_effort LAGOSUS came from public government agency databases, lake websites, and grey literature reports (see sourceurl_depth).
c) The values for lake_maxdepth_effort bathymetry came from generated bathymetric surfaces derived from publicly available state agency databases (Stachelek et al., In Press). Maximum depth was equivalent to the deepest point on the bathymetric surface.
d) The values for lake_maxdepth_effort NLA came from the National Lakes Assessment 2007 and 2012 surveys (U.S. Environmental Protection Agency, 2016).

lake_waterarea_ha was calculated using the NHD lake polygons using the Calculate Geometry tools in ArcGIS with the Albers USGS Conical Area Projection. The "water area" is equal to the areal size of the original polygon.

lake_maxdepth_effort is the name of the depth-searching effort used for quality assurance, deduplication, and provenance tracking for maximum depth; one of LAGOSNE, LAGOSNE_Q500, LAGOSUS, LAGOSUS_Q500, bathymetry, NLA.

- Depth from the LAGOSNE effort is from Soranno et al. (2017) and is limited to 17 Northeast and Midwestern states and includes data from web searches as well as being sourced directly from data providers.
- Depth from the LAGOSUS effort was collected from manual searches of web pages, government databases, and grey literature reports for lakes in the remaining 31 contiguous states (not the District of Columbia). To restrict our search only to lakes with probable depth data, search queries were conducted for all individual lakes that either: had limnological data in the LAGOS-NE research platform, had a Water Quality Portal site assigned to it, or were in the 2007 or 2012 US EPA NLA dataset (for the other 31 states). Queries were not standardized but common search engine queries included: LAKENAME + STATE + lake depth, LAKENAME + maximum depth, and LAKENAME + COUNTY + lake depth.
- Depths from the 10 deepest lakes in each of the lower 48 states were manually checked and indicated as LAGOSUS_Q500 or LAGOSNE_Q500, depending on state membership.
- Depth from the bathymetry effort was derived from bathymetric surfaces derived from application of a geometric cone model to bathymetric data extracted from 9 state-level databases (Stachelek et al., In Press; Stachelek 2021).
- The above efforts were supplemented by data from the U.S. National Lakes Assessment (NLA) 2007 and 2012 surveys (USEPA, 2016).

After quality assurance and deduplication, the number of lakes sourced from each effort were as follows: 8,316 from LAGOSNE, 3,032 from LAGOSUS, 4,990 from bathymetry, and 1,337 from NLA. Based on our search protocol, 16,012 of the 17,675 depths in LAGOS-US DEPTH have limnological data. While lakes with limnological data and depth data in the LAGOS-US research platform can be considered fully searched for during the LAGOS-US DEPTH data collection, the depths reported here should not be considered a comprehensive search of depths for all conterminous U.S. lakes. For example, reported depths are available for more lakes than the 23 lakes identified in TN in LAGOS-US DEPTH (Appendix Table 1), but many of the state's lakes were not individually searched for depth data due to the limited number of water quality data points in LAGOS-US for this state.

lake_meandepth_effort is the name of the depth-searching effort used for quality assurance, deduplication, and provenance tracking for mean depth; one of LAGOSNE, LAGOSNE_Q500, LAGOSUS, or LAGOSUS_Q500. Definitions of these efforts are described above for lake_maxdepth_effort with the exception that no depths were derived from either bathymetry or NLA for mean depth. After quality assurance and deduplication, the number of lakes with mean depth (total number = 6137) sourced from each effort were as follows: 5,215 from LAGOSNE and 922 from LAGOSUS.

4. Quality Control/Quality Assurance

All lake maximum and mean data underwent initial QAQC to remove lakes where maximum depth was less than 1 ft and to remove mean depth values that were greater than or equal to corresponding maximum depth values. Beyond these checks, further data checking proceeded differently depending on each lake's lake_maxdepth_effort value. For data from the bathymetry effort, lakes were removed that had too few contours (or points, or polygons) or where the deepest point in the lake was located directly on the lake shoreline. The latter condition was added to catch polygon (lake-linking) mismatch issues where an NHD polygon did not encompass the entire "lake". Data from other efforts were not checked independently of the initial checking effort (see above) with the exception of the LAGOS-NE data where we removed entries that lacked a 1:1 lagoslakeid match (due to differences in NHD resolution among LAGOS-NE and LAGOS-US data products).

The next QAQC step was to compare the lake depth values of lakes with more than two data sources. Where multiple sources of data were available for maximum depth, data were selected in priority order by effort of bathymetry > LAGOSUS > LAGOSNE > NLA. For mean depth the priority was LAGOSUS > LAGOSNE.
The final QAQC step was to double-check maximum lake depth values from the deepest lakes in the dataset (i.e., well-known lakes in which users would be relatively more likely to detect errors). We performed this manual step for the deepest 10 lakes in each of the lower 48 US states. Although no significant changes were necessary for the vast majority of lakes (median absolute change = 0 m), we detected discrepancies of ≥ 10 m for lake maximum depth for approximately 12% of the manually checked lakes. Manually checked lakes and that had replaced values were designated as LAGOSUS_Q500 or LAGOSNE_Q500 under lake_maximumdepth_effort. While checking maximum depth, where lake mean depth was also checked. Only one replacement value for mean depth was found.

5. Data limitations

The depths reported here should not be considered a comprehensive search of depths for all conterminous U.S. lakes. The subset of lakes included in DEPTH was created based on the presence of water quality data availability in ~ 2017 or earlier from a range of existing data sources that include: the LAGOS-NE research platform, the Water Quality Portal (https://www.waterqualitydata.us/), the 2007 and 2012 U.S National Lakes Assessment, and an effort to compile lake bathymetry for nine states (Stachelek et al., In Press). Therefore, there are likely additional lake depths that are available through online sources that are not in this data module but that were not included due to lack of resources to search for larger numbers of lakes. Future efforts could easily fill this gap to merge with this data product.

6. References

Cheruvelil, K. S., Soranno, P. A., McCullough, I. M., Webster, K. E., Rodriguez, L. K., & Smith, N. J. 2021. LAGOS-US LOCUS v1.0: Data module of location, identifiers, and physical characteristics of lakes and their watersheds in the conterminous US. Limnology and Oceanography Letters. 6(5), 270-292. https://doi.org/10.1002/lol2.10203

King, K. B., Wang, Q., Rodriguez, L. K., & Cheruvelil, K. S. 2021. Lake networks and connectivity metrics for the conterminous US (LAGOS-US NETWORKS v1). Limnology and Oceanography Letters. 6(5), 293-307. https://doi.org/10.1002/lol2.10204

Smith, N. J., K. E. Webster, L. Rodriguez, K. S. Cheruvelil, and P. A. Soranno. 2021. LAGOS-US LOCUS v1.0: Data module of location, identifiers, and physical characteristics of lakes and their watersheds in the conterminous U.S. Environmental Data Initiative. https://doi.org/10.6073/pasta/e5c2fb8d77467d3f03de4667ac2173ca

Soranno, P.A., et al., 2017. LAGOS-NE: a multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of US lakes. GigaScience 6, 1–22. https://doi.org/10.1093/gigascience/gix101.

Stachelek, J. 2021. Bathymetry data for 5,000 lakes. Figshare. https://doi.org/10.6084/m9.figshare.12722246.v1

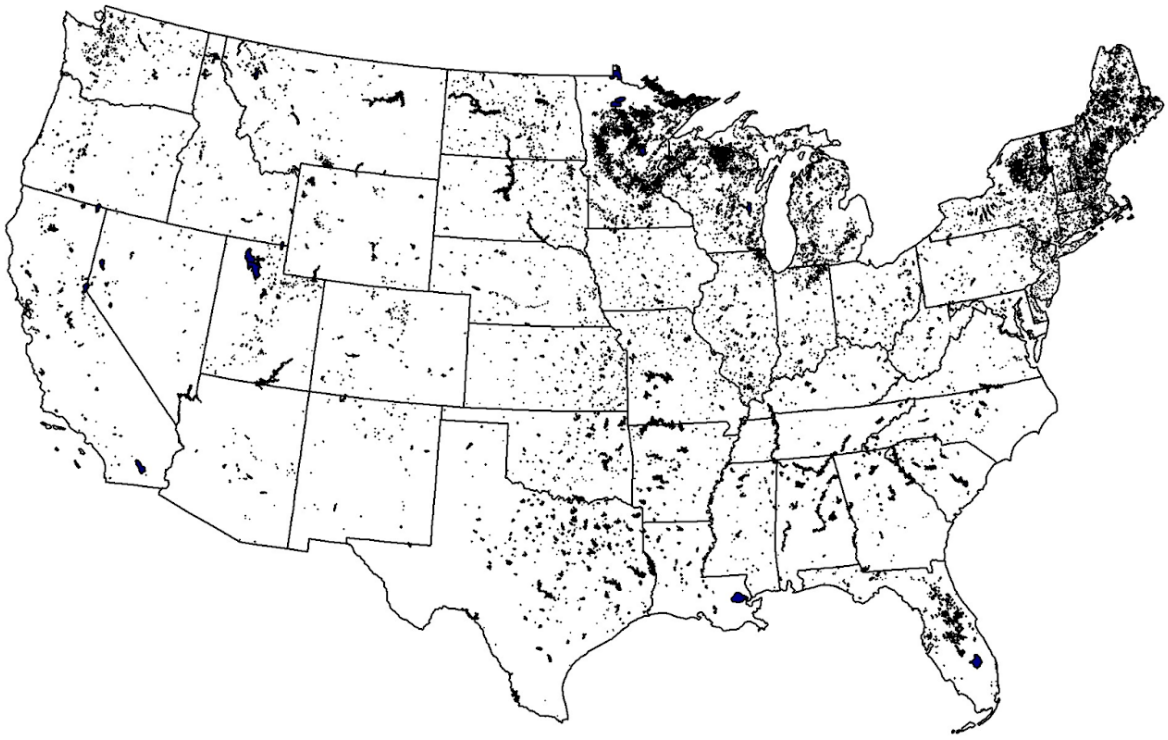
Stachelek, J., Hanly, P., Soranno, P.A. In Press. Imperfect slope measurements drive overestimation in geometric cone model of lake and reservoir depth. Inland Waters. doi:10.1080/20442041.2021.2006553

U.S. Environmental Protection Agency. 2016. National Aquatic Resource Surveys. National Lakes Assessment 2012. Available from the U.S. EPA web page: https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys. Date accessed: 2018-06-01.

U.S. Geological Survey. 2017. National Hydrography Dataset (ver. USGS National Hydrography Dataset Best Resolution (NHD) for Hydrologic Unit (HU) 4 - 2001 (published 20170105)). Accessed January 5, 2017 at URL https://www.usgs.gov/core-science-systems/ngp/national-hydrography/access-national-hydrography-products.

Appendices

Appendix 1. Number of lakes with maximum depth data by state based on the state postal abbreviation used in the depth search.

AL	AR	AZ	CA	CO	CT	DE	FL	GA	IA	ID	IL
39	84	49	149	135	155	44	676	48	152	83	482
IN	**KS**	**KY**	**LA**	**MA**	**MD**	**ME**	**MI**	**MN**	**MO**	**MS**	**MT**
364	207	40	44	322	19	1719	1099	3570	182	66	361
NC	**ND**	**NE**	**NH**	**NJ**	**NM**	**NV**	**NY**	**OH**	**OK**	**OR**	**PA**
53	252	188	788	213	64	54	1652	98	106	115	112
RI	**SC**	**SD**	**TN**	**TX**	**UT**	**VA**	**VT**	**WA**	**WI**	**WV**	**WY**
114	39	161	23	285	195	42	282	367	2238	64	81

Appendix 2. Map of the 17,675 lakes in the conterminous U.S. with depth data in LAGOS-US DEPTH.