# EDI Metadata Template (2016)[[1]](#footnote-1)

## Dataset Title

**LAGOS-NELIMNO v1.087.1:** A module for LAGOS-NE, a multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of U.S. lakes

## Short name or nickname you use to refer to this dataset:

**LAGOS-NELIMNO v1.087.1**

## Data Use Policy

**CC BY.** This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation.

## Abstract

This data package, LAGOS-NELIMNO v1.087.1, is 1 of 5 data packages associated with the LAGOS-NE database-- the LAke multi-scaled GeOSpatial and temporal database. Three of the data packages each contain different types of data for 51,101 lakes and reservoirs larger than 4 ha in 17 lake-rich U.S. states to support research on thousands of lakes. These three package are: (1) LAGOS-NELOCUS v1.01: lake location and physical characteristics for all lakes. (2) LAGOS-NEGEO v1.05: ecological context (i.e., the land use, geologic, climatic, and hydrologic setting of lakes) for all lakes. These geospatial data were created by processing national-scale and publicly-accessible datasets to quantify numerous metrics at multiple spatial resolutions. And, (3) LAGOS-NELIMNO v1.087.1: in-situ measurements of lake water quality from the past three decades for approximately 2,600-12,000 lakes, depending on the variable. This module was created by harmonizing 87 water quality datasets from federal, state, tribal, and non-profit agencies, university researchers, and citizen scientists. The other two data packages contain supporting data for the LAGOS-NE database: (4) LAGOS-NE-GIS v1.0: the GIS data layers for lakes, wetlands, and streams, as well as the spatial resolutions that were used to create the LAGOS-NEGEO module. (5) LAGOS-NE-RAWDATA: the original 87 datasets of lake water quality prior to processing, the R code that converts the original data formats into LAGOS-NE data format, and the log file from this procedure to create LAGOS-NE. This latter data package supports the reproducibility of LAGOS-NELIMNO.

The LAGOS-NELIMNO v1.087.1 module includes in situ measurements of lake water quality. We included variables that are most commonly measured by state agencies and researchers for studying eutrophication. For each water quality data value, we also include metadata related to the sampling program, methods, qualifiers with data flags from the original program (qual, not standardized for LAGOS-NE), censor codes from our quality control procedures (censorcode, standardized for LAGOS-NE), and the date of each sample.

Citation for the full documentation of this database:

Soranno, P.A., E.G. Bissell, K.S. Cheruvelil, S.T. Christel, S.M. Collins, C.E. Fergus, C.T. Filstrup, J.F. Lapierre, N.R. Lottig, S.K. Oliver, C.E. Scott, N.J. Smith, S. Stopyak, S. Yuan, M.T. Bremigan, J.A. Downing, C. Gries, E.N. Henry, N.K. Skaff, E.H. Stanley, C.A. Stow, P.-N. Tan, T. Wagner, K.E. Webster. 2015. Building a multi-scaled geospatial temporal ecology database from disparate data sources: Fostering open science and data reuse. GigaScience 4:28 doi:10.1186/s13742-015-0067-4

Citation for the data paper for this database:

Soranno, P.A., L.C. Bacon, M. Beauchene, K.E. Bednar, E.G. Bissell, C.K. Boudreau, M.G. Boyer, M.T. Bremigan, S.R. Carpenter, J.W. Carr, K.S. Cheruvelil, S.T. Christel, M. Claucherty, S.M.Collins, J.D. Conroy, J.A. Downing, J. Dukett, C.E. Fergus, C.T. Filstrup, C. Funk, M.J. Gonzalez, L.T. Green, C. Gries, J.D. Halfman, S.K. Hamilton, P.C. Hanson, E.N. Henry, E.M. Herron, C. Hockings, J.R. Jackson, K. Jacobson-Hedin, L.L. Janus, W.W. Jones, J.R. Jones, C.M. Keson, K.B.S. King, S.A. Kishbaugh, J.-F. Lapierre, B. Lathrop, J.A. Latimore, Y. Lee, N.R. Lottig, J.A. Lynch, L.J. Matthews, W.H. McDowell, K.E.B. Moore, B.P. Neff, S.J. Nelson, S.K. Oliver, M.L. Pace, D.C. Pierson, A.C. Poisson, A.I. Pollard, D.M. Post, P.O. Reyes, D.O. Rosenberry, K.M. Roy, L.G. Rudstam, O. Sarnelle, N.J. Schuldt, C.E. Scott, N.K. Skaff, N.J. Smith, N.R. Spinelli, J.J. Stachelek, E.H. Stanley, J.L. Stoddard, S.B. Stopyak, C.A. Stow, J.M. Tallant, P.-N. Tan, A.P. Thorpe, M.J. Vanni, T. Wagner, G. Watkins, K.C. Weathers, K.E. Webster, J.D. White, M.K. Wilmes, S. Yuan. *In Review.* LAGOS-NE: A multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of U.S. lakes. In Review at GigaScience. Submitted April 2017.

## Investigators

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| First Name | Last Name | Organization | e-mail address | ORCID ID (optional) |
| Patricia A. | Soranno | Michigan State University | [soranno@msu.edu](mailto:soranno@msu.edu) | 0000-0003-1668-9271 |
| Kendra S. | Cheruvelil | Michigan State University | [ksc@msu.edu](mailto:ksc@msu.edu) | 0000-0003-1880-2880 |
|  |  |  |  |  |

## Other personnel names and roles

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| First Name | Last Name | Organization | e-mail address | ORCID ID (optional) | Role in project |
| Edward G. | Bissell | Michigan State University | [Bissell3@msu.edu](mailto:Bissell3@msu.edu) |  | Database administrator |
| Nicole J. | Smith | Michigan State University | [Nicole.j.smith@gmail.com](mailto:Nicole.j.smith@gmail.com) |  | GIS analyst |

## Keywords

LAGOS-NE, Lakes, nutrients, water quality, water clarity, chlorophyll, ecological context, lake trophic state, eutrophication

## Funding of this work:

Title of grant: **The effect of cross-scale interactions on freshwater ecosystem state across space and time**

Principal Investigator: **Patricia A. Soranno**

Co-investigators: **Kendra S. Cheruvelil, Emily H. Stanley, Noah Lottig, John A. Downing, Pang-Ning Tan**

Granting agency: **National Science Foundation**

Grant identification number: **EF-1065786, EF-1065818, EF-1065649**

## Timeframe

* Begin date: **May 2011**
* End date: **May 2017**
* Data collection ongoing/completed: **Completed**

## Geographic location

* Verbal description: Northeastern and upper Midwestern U.S. states, including: Minnesota, Iowa, Wisconsin, Illinois, Indiana, Missouri, Michigan, Ohio, Pennsylvania, New York, Connecticut, Rhode Island, New Hampshire, Vermont, Maine, New Jersey, New YOrk
* North bounding coordinates (decimals) – PLEASE FILL IN
* South bounding coordinates (decimals) – PLEASE FILL IN
* East bounding coordinates (decimals) – PLEASE FILL IN
* West bounding coordinates (decimals) – PLEASE FILL IN

## Taxonomic species or groups

NA

## Methods

Please see the following manuscript for detailed methods used to harmonize the datasets in LAGOS-NE; in particular, many of the specific methods and procedures are found in the online ‘Additional Files’ available at: <https://gigascience.biomedcentral.com/articles/10.1186/s13742-015-0067-4>

Soranno, P.A., E.G. Bissell, K.S. Cheruvelil, S.T. Christel, S.M. Collins, C.E. Fergus, C.T. Filstrup, J.F. Lapierre, N.R. Lottig, S.K. Oliver, C.E. Scott, N.J. Smith, S. Stopyak, S. Yuan, M.T. Bremigan, J.A. Downing, C. Gries, E.N. Henry, N.K. Skaff, E.H. Stanley, C.A. Stow, P.-N. Tan, T. Wagner, K.E. Webster. 2015. Building a multi-scaled geospatial temporal ecology database from disparate data sources: Fostering open science and data reuse. GigaScience 4:28 doi:10.1186/s13742-015-0067-4

We acquired water quality datasets for LAGOS-NELIMNO v1.087.1 by contacting individuals at each of the 17 state and 5 tribal agencies. These contacts helped us to identify the state-agency collected dataset required by the Clean Water Act and which is most likely to be in the public domain. In this way, we were able to acquire at least one (and typically more) dataset from each of the 17 states. Because states vary in sampling approach and intensity (see below for details), we sought to supplement these datasets with other known sources of water quality data, including university researchers, federal agencies, and non-profit groups. The full list of data sources acquired is in Soranno et al. (2015) in ‘Additional File 17’. Based on resources and effort required to include them, we incorporated a subset of these datasets in LAGOS-NELIMNO v1.087.1.

All methods to create this module are described in Soranno et al. (2015). Briefly, for each dataset acquired, we authored LAGOS-NE metadata in EML to aid in data provenance. We also incorporated key metadata features (e.g., methods used, censor codes (if applicable)), and sampling program information) into the database so that future users could easily identify these important attributes. Because each dataset was unique in structure, file format, and naming conventions, we manually processed each dataset and its metadata so that they could be translated into the standard LAGOS-NE vocabulary and data model. Although labor-intensive, we created customized R scripts to process and load each dataset separately, which are available in LAGOS-NE-RAWDATA data module.

## Data Table

* INSERT LAGOS-NE-LIMNO Metadata tables

1. This document liberally borrows from similar documents at SBC and GCE [↑](#footnote-ref-1)