The Journal of Open Source Software

Lenapy: Enhancing xarray for multidimensional geophysical data analysis

Hugo Lecomte [1], **Sebastien Fourest** [1,2], and **Alejandro Blazquez** [1,2]

1 Université de Toulouse, LEGOS (CNES/CNRS/IRD/UT3), France **2** Centre national d'Études Spatiales (CNES), France

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Jayaram Hariharan

Reviewers:

- @ThomasMGeo
- @miniufo

Submitted: 17 July 2025
Published: unpublished

Summary

Lenapy is a Python library designed to facilitate the processing and analysis of geophysical and climate datasets, such as those used in oceanography, geodesy, and Earth observation in general. Built on top of `xarray` and fully compatible with `dask`, it enables scalable workflows on multidimensional datasets using community standards for data formats (NetCDF) and metadata conventions (CF).

Lenapy provides high-level accessors that extend `xarray.Dataset` and `xarray.DataArray` objects, allowing application of specialized methods.

One of `Lenapy`'s key features is its unified approach for Global Mean Sea Level estimation, enabling consistent computation of its components.

Lenapy is intended to support geophysicists, oceanographers, and climate service providers with coherent tools through an extensible interface.

Statement of need

Geophysical and climate datasets are intrinsically multidimensional, often encompassing spatial (longitude, latitude, depth/height) and temporal dimensions. Analyzing these data requires specialized operations such as harmonic analysis, climatology, thermodynamics, and geodetic corrections. For the oceanography part, the accurate computation of seawater properties is essential in physical oceanography. The Thermodynamic Equation of Seawater (TEOS-10) framework provides consistent definitions and algorithms for quantities such as potential temperature, conservative temperature, and density (McDougall & Barker, 2011).

Existing libraries address specific aspects of these requirements:

- `pyshtools` (Wieczorek & Meschede, 2018) is a comprehensive package for spherical harmonic transforms and spectral analysis, particularly in geophysics. Yet, it operates on standalone arrays and does not natively support xarray, limiting its compatibility with NetCDF-based workflows.
- `grates` (Kvas, 2024) provides object-oriented tools for spherical harmonics, but similarly only operates on standalone arrays.
- `gsw-xarray` (Caneill & Barna, 2024) provides implementations of the TEOS-10, facilitating oceanographic computations. It offers a wrapper around `GSW-Python` for xarray objects, but without Dask support.

While `gsw-xarray` is more complete than our GSW wrapper in `Lenapy`, our library propose complementary geodetic tools for spatial and spherical harmonics operations.

To our knowledge, no other existing Python library offers a coherent suite of oceanographic and geophysical operations (detailed below) within a unified, xarray-native framework supporting both scalability (via Dask (Team, 2016)) and labeled, multidimensional arrays. Moreover, critical geospatial utilities (such as surface-aware averaging, weighted statistics, spherical distance computation, and climatology fitting) remain fragmented across ecosystems or require custom implementations.

Lenapy addresses this gap by providing a modular Python package built on xarray and Dask, exposing accessors (.lngeo, .lnharmo, .lnocean, .lntime) for direct and simple application of domain-specific methods to xarray.Dataset or xarray.DataArray objects. For example, users can compute area-weighted means via `ds.lngeo.mean()` or extract the global ocean heat content using `ds.lnocean.gohc()`.

Lenapy is designed for Earth scientists, oceanographers, climate researchers, and geodesists who routinely manipulate global or regional gridded datasets and require specific processing workflows. Lenapy aims to maintain full compatibility with the PyData ecosystem.

Furthermore, Lenapy offers a unified approach for calculating Global Mean Sea Level by integrating both steric and manometric components, as well as relative sea-level changes (Gregory et al., 2019). Lenapy facilitates this decomposition with a unique Python library that compute these components directly from xarray-based datasets, enabling researchers to analyze sea-level changes comprehensively within a consistent framework.

Key Features

Spatial operations (`.lngeo`)

The `lngeo` accessor provides geodetic tools designed for gridded data on spherical or ellipsoidal Earth models. It includes:

- Geodetic estimation of grid cell surface areas and distances and geographical weighted operations (e.g., mean or sum);
- Isosurface computation;
- A wrapper to the `xESMF` regridding library (Zhuang et al., 2025).

Spherical harmonics operations and gravity field processing (`.lnharmo`)

The `lnharmo` accessor offers dedicated methods for working with spherical harmonic representations, particularly in the context of Earth gravity field modeling. It includes:

- Reading, handling, and manipulating datasets containing spherical harmonic coefficients (variables `clm`, `slm`), with options to change reference frames;
- Converting spherical harmonic representations into gridded spatial fields and inverse transformation.

Oceanography (`.lnocean`)

The `lnocean` accessor provides a lightweight wrapper around selected GSW (TEOS-10) routines (McDougall & Barker, 2011), exposing them as native xarray methods for oceanographic datasets. Based on any dataset containing any temperature or any salinity with depth coordinate it provides integrated values over all or part of the water column, including :

- Ocean heat content;
- Steric sea levels;
- Density and dynamic height.

Time series and climatology tools (`.lntime`)

The `lntime` accessor enables common temporal operations on geophysical time series, including:

- Climatological and polynomial signal extraction and fitting (e.g., seasonal decomposition);
- Filtering of time series;
- Detrending, interpolation, derivation, evaluation of missing values, …

Input/Output utilities

Lenapy includes I/O helpers to support multiple data formats used in Earth observation and geoscience:

- Readers for gridded ocean temperature/salinity products from various sources;
- Readers for spherical harmonics formats, including Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On Science Data System files and ICGEM-format files `.gfc`;
- Writers for ICGEM-format, enabling export of custom spherical harmonics to interoperable standards.

Projects using Lenapy

Lenapy is actively used in several international research projects and operational workflows focused on Earth system science and climate monitoring.

Notable projects include:

- **Sea Level Budget Closure CCI+ (SLBC_CCI+)**, funded by the European Space Agency, uses Lenapy in several work packages for computing steric and manometric contributions to sea level and closing the global sea level budget.
- **ERC Synergy GRACEFUL**, a European Research Council project dedicated to improving the understanding of Earth's interior using gravimetric data.

In addition, Lenapy has been employed in related research studies (Bouih et al., 2025) for consistent and reproducible treatment of steric, manometric, or relative sea level. Lenapy gravity field and spherical harmonics operations are used at LEGOS for the processing GRACE (& Follow-On) Level 2 dataset to create a Level 3 gravity solution (A. Blazquez et al., 2018).

Origin and research context

Lenapy originated from geophysical tools developed within the LEGOS research laboratory during PhD works focused on the variability of sea level and ocean–continent water exchanges (Alejandro Blazquez, 2020; Dieng, 2017; Meyssignac, 2012). These early tools have been generalized and integrated into a modern Python framework, making them accessible and reusable by the broader Earth science community.

Acknowledgements

Lenapy benefited of CNES support to improve its maintainability and its documentation.

We thank Robin Guillaume-Castel and Arthur Vincent for their help with the early code.

References

Blazquez, Alejandro. (2020). *Caractérisation par satellite des Échanges d'eau entre l'océan et les continents aux Échelles interannuelles à décennales* [PhD thesis]. Université de Toulouse.

Blazquez, A., Meyssignac, B., Lemoine Jean-Michel, J. M., Berthier, E., Ribes, A., & Cazenave, A. (2018). Exploring the uncertainty in GRACE estimates of the mass redistributions at the Earth surface: Implications for the global water and sea level budgets. *Geophysical Journal International*, *215*(1), 415–430. https://doi.org/10.1093/gji/ggy293

Bouih, M., Barnoud, A., Yang, C., Storto, A., Blazquez, A., Llovel, W., Fraudeau, R., & Cazenave, A. (2025). Regional sea level budget over 2004-2022. *EGUsphere*, 1–30. https://doi.org/10.5194/egusphere-2024-3945

Caneill, R., & Barna, A. (2024). *Gsw-xarray*. Zenodo. https://doi.org/10.5281/zenodo.11382921

Dieng, H. B. (2017). *Variations actuelles du niveau de la mer* [PhD thesis]. Université Toulouse 3.

Gregory, J. M., Griffies, S. M., Hughes, C. W., Lowe, J. A., Church, J. A., Fukimori, I., Gomez, N., Kopp, R. E., Landerer, F., Cozannet, G. L., Ponte, R. M., Stammer, D., Tamisiea, M. E., & van de Wal, R. S. W. (2019). Concepts and Terminology for Sea Level: Mean, Variability and Change, Both Local and Global. *Surveys in Geophysics*, *40*(6), 1251–1289. https://doi.org/10.1007/s10712-019-09525-z

Kvas, A. (2024). Grates. In *GitHub repository*. GitHub. https://https://github.com/akvas/grates

McDougall, T. J., & Barker, P. M. (2011). *Getting started with TEOS-10 and the gibbs seawater (GSW) oceanographic toolbox* (p. 28). SCOR/IAPSO WG127. ISBN: 978-0-646-55621-5

Meyssignac, B. (2012). *La variabilité régionale du niveau de la mer* [PhD thesis]. Université Toulouse 3.

Team, D. D. (2016). *Dask: Library for dynamic task scheduling*. https://dask.org

Wieczorek, M. A., & Meschede, M. (2018). SHTools: Tools for Working with Spherical Harmonics. *Geochemistry, Geophysics, Geosystems*, *19*(8), 2574–2592. https://doi.org/10.1029/2018GC007529

Zhuang, J., dussin, raphael, Huard, D., Bourgault, P., Banihirwe, A., Raynaud, S., Malevich, B., Schupfner, M., Filipe, Gauthier, C., Levang, S., Jüling, A., Almansi, M., RichardScottOZ, RondeauG, Rasp, S., Smith, T. J., Mares, B., Stachelek, J., … Li, X. (2025). *Pangeo-data/xESMF: V0.8.10*. Zenodo. https://doi.org/10.5281/zenodo.15304267