

Standard Operating Procedure (SOP): Archiving Legacy Ocean-Bottom Seismometer (OBS) and Ocean-Bottom Hydrophone (OBH) Raw Data at PANGAEA

Janine Berndt¹, Daniel Damaske², Marianne Rehage², Michael Riedel¹

¹ GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

² MARUM – Center for Environmental Sciences, University of Bremen, Bremen, Germany
(PANGAEA – Data Publisher for Earth & Environmental Science)

Keywords: NFDI₄Earth, DAM, metadata, OBS, Ocean-Bottom Seismometer, OBH, Ocean-Bottom Hydrophone, Marine Geophysics, SEG-Y, miniSEED

Changelog:

Version	Date
Version 1.0.0	15.04.2026

1 Introduction

This Standard Operating Procedure (SOP) defines the standardized process for archiving legacy Ocean-Bottom Seismometer (OBS) and Ocean-Bottom Hydrophone (OBH) raw datasets in PANGAEA. It describes in detail the required metadata, documentation standards, and submission workflow to ensure consistent, quality-controlled publication within the German marine seismic community.

Systematic archiving of legacy datasets safeguards valuable scientific data, supports long-term accessibility, and facilitates their reuse in future research. Implementing this SOP improves the findability, accessibility, interoperability, and reusability of historical marine seismic datasets through structured and standardized documentation.

2 PANGAEA Framework

2.1 General concept of data publication in PANGAEA

In PANGAEA, scientific data are published as individual datasets or as part of a data collection. A dataset represents a defined and citable unit of data (e.g., OBS/OBH recordings from a specific deployment or experiment during a research cruise) and is assigned a persistent DOI at the time of publication. A data collection may group multiple related datasets, e.g., homogeneous datasets such as time series or datasets that are not reusable independently.

The submitted metadata and data files are reviewed by a data curator prior to publication to ensure completeness and compliance with PANGAEA standards and the FAIR data principles (Wilkinson et al., 2016). If necessary, the curator will contact the data provider for further questions regarding the submission.

It is recommended to prepare all required metadata and data files in advance before starting the online submission process.

Note: For demonstration purposes, an already published dataset is used as a structural template. Submitters are expected to follow the same structure and level of detail when preparing their own dataset:

Riedel, Michael; Papenberg, Cord; Bialas, Jörg (2024): Continuous seismic raw data (Ocean Bottom Seismometer working area dataset) of RV POSEIDON during cruise POS515, Poseidon mud volcano chain, Calabrian Arc, Mediterranean Sea [dataset]. PANGAEA, <https://doi.org/10.1594/PANGAEA.967552>

All referenced weblinks and supplementary resources are listed in [Appendix 2](#).

Further information on PANGAEA and the submission process is provided in [Felden et al. \(2023\)](#) and in the PANGAEA Authors Guidelines (see [Appendix 2 \[1\]](#)).

2.2 Step-by-step submission procedure

2.2.1 *Step 1: Access the PANGAEA submission system*

Navigate to the PANGAEA website (see Appendix 2 [2]) and select "Submit your data" from the main navigation menu. Log in using your registered account credentials.

2.2.2 *Step 2: Submit your data*

After selecting "Create new submission", the submission workflow is initiated. The "Basics" page will open. Complete the following fields:

Title

Provide a concise and descriptive dataset title that clearly reflects what was measured, where, and during which expedition. The title must include:

- Type of data (raw)
- Instrument type (OBS or OBH)
- Cruise/expedition identifier
- Research vessel
- Study area (optional)

Adapt the example title cited in this SOP to reflect the specific dataset being submitted.

When referring to geographic locations, use official place names (e.g., those listed in the IHO-IOC GEBCO Gazetteer of Undersea Feature Names, see Appendix 2 [3]).

Authors

Enter all authors in the correct order.

For each author, provide:

- Family name
- Given name
- Email address
- Institutional affiliation

Ensure that the spelling of each author name matches the corresponding ORCID record. Where applicable, use the official institutional affiliation as listed in the Research Organization Registry (ROR) (see Appendix 2 [4]).

Keywords

Add relevant keywords describing the instrument type and geological or tectonic setting. Avoid overly generic terms (e.g., "data", "experiment"). Enter each keyword separately and use standardized terminology where possible.

Abstract/Describe your Data

Provide a structured description that includes:

- Type of data (e.g., continuous records of seismic raw data)
- Study area
- Instrument configuration (number of channels, type)
- Number of stations
- Number of profiles (if applicable)

The abstract must allow users to understand the dataset without additional documentation. In addition, cite the cruise report (in text citation only).

If no cruise report is available, the Cruise Summary Report (CSR) via SeaDataNet should be cited where applicable (see Appendix 2 [5]). CSR entries provide standardized metadata describing the cruise.

Cite the CSR as follows:

Riedel, M. (2017). Cruise summary report: POS515 (RV Poseidon, Germany, June 16–July 07, 2017) [Data set]. SeaDataNet. <https://csr.seadatanet.org>

License

Select the appropriate publication license (e.g., CC-BY 4.0). Unless otherwise required, CC-BY 4.0 is recommended.

2.2.3 Step 3: Provide related references and contextual information

The References section allows you to link the cruise report (mandatory where applicable), publications, and other related datasets including their DOIs, which provide the scientific context for the dataset.

2.2.4 Step 4: Provide funding references

On the "Projects and Grants" tab, enter all relevant projects and funding sources associated with the dataset.

For each project or grant, provide:

- Project title
- Funding agency (see Appendix 2 [6] or [7])
- Grant or project number
- Project website (if available)

Funding information must be entered consistently and completely to ensure accurate attribution and compliance with funding requirements. Where available, include a direct link to the official project website.

If multiple grants contributed to the dataset, add each grant as a separate entry.

Use the official project title and grant number as provided by the funding agency.

2.2.5 Step 5: Upload your data

On the "Upload" tab, provide the data files associated with the dataset. However, OBS/OBH datasets typically exceed standard browser upload limits. Therefore, request an upload link by clicking the checkbox before uploading your data.

Data level selection

Each dataset must represent a single, clearly defined data level to ensure transparent DOI assignment, unambiguous citation, and long-term FAIR compliance.

Submit only:

- **Raw data** (strictly acquisition-level data without processing)

Raw and processed data constitute distinct data levels and must therefore be archived as separate datasets.

If multiple data levels are to be published, create a separate dataset entry for each level within the submission workflow.

The dataset title must clearly indicate whether the submitted data are raw or processed.

File preparation requirements

Before uploading, ensure that:

- Files are complete and finalized
- File names follow a consistent and documented naming convention
- The file format is clearly specified (e.g., SEG-Y, PASSCAL SEG-Y, miniSEED)
- No temporary, duplicate, or intermediate files are included
- Do not use a standard archive format (e.g., ZIP or TAR.GZ)

Documentation of the file structure

The file naming convention and internal data structure must be described in the "Abstract/Describe your Data" section (see Section 2.2.2, Step 2).

Where applicable, clearly document:

- Channel configuration
- Sampling rate
- Trace length
- Header information relevant for interpretation

"File descriptions"

Provide a structured description of the uploaded files in the "File descriptions" field of the submission interface. The description should include:

- File format
- Naming convention
- Survey or station identifiers
- Number of files
- Total data volume
- Relevant acquisition or processing remarks

This information supports verification of data completeness and integrity during editorial review.

The following example defines the expected structure and level of detail for the file description. The same structure should be applied and adapted to the respective dataset.

Comment:

Data are stored for each OBS location separately. Data comprise the entire recording period. We are using the Julian Day numbering convention. There are 12 OBS stations in total (P01 - P12). Julian days are from day 183 (= July 2, 2017) until day 186 (= July 5, 2017).

Physical location of OBS drop position can be found in the associated cruise report (Riedel et al., 2017).

A data-file for any given OBS has a name following this format:

e.g. P.01.17.183.16.22.17.0504.1

P.*.*...**. Poseidon mud volcano chain: P

.01.*.*...**. OBS station number: 01

..17.*.*...**. Year: 2017

...183.*.*...**. Julian Day of the start of the record: 183

...*.16.*.*...**. Hour of the start of the record: 16

...*.22.*.*...**. Minute of the start of the record: 22

...*.*.17.*.*...**. Second of the start of the record: 17

...*.*.0504. Serial number of the MBS-Recorder: 0504

...*.*.*.*.1 This value is the recorded channel (1= hydrophone, 2= horizontal 1, 3 = horizontal 2, 4 = vertical)

Sampling rate of the recordings was at 1000 Hz.

Clock time drift (skew) with respect to GMT +0

OBS P01: Clock start recording Julian Day 183, 16:20:00, End recording Julian Day 187, 12:10:00, drift +9 ms

OBS P02: Clock start recording Julian Day 183, 16:41:00, End recording Julian Day 187, 12:46:00, drift -28 ms

OBS P03: Clock start recording Julian Day 183, 15:50:00, End recording Julian Day 187, 12:29:00, drift -24 ms

OBS P04: Clock start recording Julian Day 183, 17:00:00, End recording Julian Day 187, 16:25:00, drift -26 ms

OBS P05: Clock start recording Julian Day 183, 17:21:00, End recording Julian Day 187, 14:36:00, drift +35 ms

OBS P06: Clock start recording Julian Day 183, 17:27:00, End recording Julian Day 187, 13:34:00, drift -153 ms

OBS P07: Clock start recording Julian Day 183, 17:34:00, End recording Julian Day 187, 13:43:00, drift -2 ms

OBS P08: Clock start recording Julian Day 183, 17:43:00, End recording Julian Day 187, 17:02:00, drift -131 ms

OBS P09: Clock start recording Julian Day 183, 17:52:00, End recording Julian Day 187, 16:46:00, drift +22 ms

OBS P10: Clock start recording Julian Day 183, 18:10:00, End recording Julian Day 187, 15:31:00, drift +24 ms

OBS P11: Clock start recording Julian Day 183, 18:21:00, End recording Julian Day 187, 15:04:00, drift +1 ms

OBS P12: Clock start recording Julian Day 183, 18:28:00, End recording Julian Day 187, 14:02:00, drift -252 ms

Julian Day 183 = July 2, 2017

2.2.6 Step 6: Review and submit the dataset

On the "Submit" tab, before final submission, verify that:

- Only one data level (raw) is included
- The dataset title clearly indicates the data level
- All mandatory metadata fields are completed
- Authors are entered in the correct order and institutional format
- The abstract contains a structured dataset description
- The file naming convention and internal data structure are documented
- No temporary, duplicate, or intermediate files are included
- Funding information includes project title, funding agency, grant number, and website (if available)
- Relevant references and cruise report are cited
- An upload link has been requested, if required
- All uploaded files are complete and finalized

State the total number of uploaded files and the overall data volume in the "Comment" field to support verification of dataset completeness.

If a moratorium is required, tick the corresponding box on the "Submit" tab and specify the requested duration.

Confirm acceptance of the terms of use before final submission. After final review, submit the dataset for editorial processing. The dataset will undergo editorial review prior to publication and DOI registration.

Appendix 1 – Abbreviations

DOI Digital Object Identifier

FAIR

F_{indable} A_{ccessible} I_{nteroperable} R_{eusable}



OBH Ocean-Bottom Hydrophone

OBS Ocean-Bottom Seismometer

ROR Research Organization Registry

CSR Cruise Summary Report

SEG Society of Exploration Geophysicists

SEG-Y Standardized file format defined by the SEG for storing and exchanging seismic data

SOP Standard Operating Procedure

Appendix 2 – References/Weblinks

Felden, J., Möller, L., Schindler, U., Huber, R., Schumacher, S., Koppe, R., Diepenbroek, M., & Glöckner, F. O. (2023).

PANGAEA—Data publisher for earth and environmental science. *Scientific Data*, 10, Article 397.

<https://doi.org/10.1038/s41597-023-02269-x>

Wilkinson, M., Dumontier, M., Aalbersberg, I., Appleton, G., Axton, M., Baak, A., Blomberg, N., & Mons, B. (2016).

The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, Article 160018.

<https://doi.org/10.1038/sdata.2016.18>

[1] [PANGAEA authors guideline]

https://wiki.pangaea.de/wiki/Authors_Guides

[2] [PANGAEA submission landing page]

<https://www.pangaea.de/submit/>

[3] [IHO-IOC GEBCO Gazetteer Undersea Features Search]

<https://www.ngdc.noaa.gov/gazetteer/view/home>

[4] [ROR]

<https://ror.org/>

[5] [CSR - SeaDataNet]

<https://csr.seadatanet.org/>

Additional resources

Funding agency listing →

[6] <https://gepris.dfg.de/gepris/>

or

[7] <https://foerderportal.bund.de>

Datasets

The following published datasets illustrate the recommended structure and level of detail:

- Riedel, Michael; Papenberg, Cord; Bialas, Jörg (2024): Continuous seismic raw data (Ocean Bottom Seismometer working area dataset) of RV POSEIDON during cruise POS515, Poseidon mud volcano chain, Calabrian Arc, Mediterranean Sea [dataset]. PANGAEA, <https://doi.org/10.1594/PANGAEA.967552>
- Riedel, Michael; Papenberg, Cord; Bialas, Jörg (2024): Continuous seismic raw data (Ocean Bottom Seismometer working area dataset) of RV POSEIDON during cruise POS515, Venere mud volcano, Calabrian Arc, Mediterranean Sea [dataset]. PANGAEA, <https://doi.org/10.1594/PANGAEA.967530>

Appendix 3 – Changelog (extended version)

Version	Date	Comment
Version 1.0.0	15.04.2026	Initial release – OBS/OBH raw datasets