SOUTHERN OCEAN OBSERVING SYSTEM

Report Series

Report on the data management needs of Southern Ocean glider users

SOOS Report Series, #11

2020

Report on the data management needs of the Southern Ocean glider users

Authors

Sebastiaan Swart (University of Gothenburg, Sweden)

Bastien Queste (University of Gothenburg, Sweden)

Pip Bricher (SOOS International Project Office, Institute for Marine and Antarctic Studies, University of Tasamania)*

*Corresponding Author: data@soos.aw

SOOS
SOUTHERN OCEAN
OBSERVING SYSTEM

UNIVERSITY OF
GOTHENBURG

Citation

Swart, S., Queste, B., and Bricher, P. (2020): Report on the data management needs of the Southern Ocean glider users, SOOS Report Series #11.

Executive Summary

During the SOOS Data Hackathon (Incheon, Korea; May 2019), key underwater glider users identified objectives for SOOS community efforts to improve FAIR practices (Findable, Accessible, Interoperable, and Reusable) in glider data.

The glider community is relatively young, and uptake in the Southern Ocean is growing rapidly. As glider uptake increases, past datasets grow in importance. It is critical to promote user engagement with data management best practices. The absence of a centralised infrastructure has led to numerous orphan datasets and disparate data formats. Three major international efforts exist to push a standardised data/metadata format; the EGO format, the IOOS format and the IMOS format. SOOS would look to

1. communicate to the community to follow these guidelines and formats,

2. use its network (e.g. Regional Working Groups) to encourage submission of glider datasets to Data Assembly Centres (DACs) to enhance data discoverability and usability, and

3. advise international programmes on new requirements for polar datasets that may not be represented otherwise.

This report uses the framework of FAIR (Findable, Accessible, Interoperable, and Reusable) data to explore key tasks needed to improve the data management of glider data in the Southern Ocean. It is not intended to replace or override similar conversations that are occurring in the global glider data management community (e.g. https://www.oceangliders.org/taskteams/data-management/).

Data interoperability & reusability

There are a number of activities that could improve the interoperability of existing and future glider datasets collected in the Southern Ocean:

- Alignment among the major data centres on data standards and formatting issues;in particular, agreement and clear direction to scientists is needed on standards for raw data and gridded products, as well as best practice advice on metadata records and associated data processing files.

- Encouraging national data centres to adopt recognised standards or interface with existing specialised glider DACs.

- There is no established autonomous surface vehicle community at the global level and little understanding of how surface and underwater vehicle data requirements overlap. Encouraging discussion in the glider community is an important first step.

- Community agreement is needed on fundamental issues, including how do scientists currently deal with derived datasets (turbulence from varied methods, velocities uvw) and how should they do so. Decisions should be made on whether derived datasets should be published for general use or only made available by contacting their owner, due to the variety of methods for calculating many derived variables. Such an approach would limit access to these datasets but encourage more nuanced discussion about appropriate use.

Data findability and accessibility

Increasing reuse of glider data requires improvements in data curation that make it easier for researchers to find existing datasets and assess data quality. Tasks that could assist in improving the discoverability of Southern Ocean datasets include:

- Identify key underwater glider role players and data formats. Participants in the discussion are aware of numerous orphan glider datasets that could be targeted for curation at a glider DAC or other relevant data centre.

- Following the lead of data managers and scientists associated with EGO, IOOS, and IMOS, promote data discovery best practices among the Southern Ocean glider community

- SOOS Regional Working Groups could be tasked with keeping track of deployments in their region to ensure discoverability by the wider SOOS community.

Strategy

Achieving the increase in data discoverability and interoperability for gliders will require some strategic consideration. It is important to recognise that nobody is currently funded to undertake the activities outlined in this document. However, the following activities should be investigated:

- Explore options for seed money that could be leveraged with other resources.
- Explore options for funding students (e.g. SCAR fellowships) who could undertake some of the tasks outlined in this document as part of their research
- Use funding from Gothenburg University to the SOOS International Project Office for short-term work on addressing these tasks.
- Identify what information is required to improve data quality.
- Community agreement is needed on standard approaches to gridding and interpolation. Is there value in publishing derived gridded datasets in addition to raw, maximum resolution datasets?
- An ideal goal would be to develop a standardised, interoperable system that takes in raw glider data and makes it available to all relevant data discovery portals (e.g. SOOSmap, IMOs, EGO) Single, standardised system to funnel data from active platforms to the World Meteorological Organisation's Global
- A standardised data format (IOOS, IMOS or EGO) would make it feasible to have a single toolkit to QC/process platform data. Currently, the reuse of glider data is slowed by the difficulty of data processing for new users. Valuable new glider processing tools are available (notably: for Python - Gregor et al., 2019: https://doi.org/10.3389/fmars.2019.00738) and new ones are being developed fairly regularly. An online tool with a graphical interface (eg. Python/Electron + D3.js) to do basic processing of data in standard formats (cf. BYQ proposal) would considerably increase the reusability of glider data.

The discussions summarised above demonstrate that SOOS has a role to play in encouraging its community to engage with the existing glider data management system and SOOS glider users have opportunities to contribute to the global planning for glider data management.

Reference:

Gregor L, Ryan-Keogh TJ, Nicholson S-A, du Plessis M, Giddy I and Swart S (2019) GliderTools: A Python Toolbox for Processing Underwater Glider Data. Front. Mar. Sci. 6:738. doi: 10.3389/fmars.2019.00738

SOOS

SOUTHERN OCEAN
OBSERVING SYSTEM

SOOS is an initiative of the Scientific Committee on Oceanic Research and the Scientific Committee on Antarctic Research

SOOS International Project Office hosted by

UNIVERSITY of TASMANIA | IMAS INSTITUTE FOR MARINE & ANTARCTIC STUDIES