Data Users' Guide:
SWOT Ocean Cal/Val In-Situ Data and NOPP Global Internal Wave Array Data from California Current Site

Matthias Lankhorst, Scripps Institution of Oceanography

Correspondence Address:
Matthias Lankhorst
Scripps Institution of Oceanography
9500 Gilman Drive
La Jolla, CA 92093-0230
USA

Revision History

Release Date	Release Lead Author	Release Notes
19-May-2025	M. Lankhorst	Initial release. Outstanding items: Mooring velocity data; C-PIES data; Mooring GPS data.

Introduction

The SWOT satellite mission (surface water, ocean topography) launched in late 2022 and underwent a calibration and validation (cal/val) phase in 2023. As part of cal/val, an array of oceanographic instruments was deployed at a site 300 km offshore of California. This document describes the datasets from this array. The scientific objectives and results of the SWOT cal/val have been described by Wang et al. (2022 and 2025). The SWOT cal/val effort was augmented by additional instrumentation from an otherwise independent project to study internal waves, here referred to as NOPP-GIW (for National Oceanographic Partnership Program – Global Internal Waves). NOPP-GIW continued data collections at the site off California beyond the SWOT cal/val phase through late 2024, in a different configuration but with similar instrumentation. Data from NOPP-GIW are included in this collection, and described here, because they are closely related to the SWOT cal/val efforts. A user might want data merged across both projects, irrespective of which project funded which particular instrument. The instrumentation deployed consists of: CTDs (conductivity, temperature, depth), which provide temperature and salinity data; temperature-only loggers; current meters (profiling and single-depth); barometers; precision GPS receivers; and seafloor instruments called C-PIES (current- and pressure sensing inverted echo sounders) that provide currents, pressure, and water column average sound speed. Figures 1 and 2 show the instrument types as well as maps with the array locations, and table 1 lists an overview of which instruments were deployed when and on which platforms.

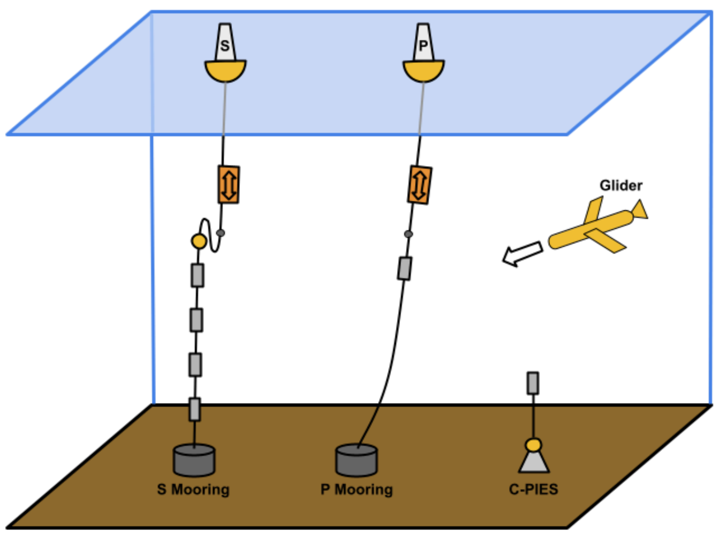

Figure 1: Schematics of the deployed platforms, not to scale. An S mooring (left) has a profiling instrument moving up and down along the riser in the upper water column (orange box with arrows) and multiple instruments fixed on the wire in the deeper water column (grey boxes, actual number of instruments is larger than shown here). A P mooring (center) has a similar profiler, but only one fixed instrument below. P mooring risers have more scope than S moorings, resulting in larger watch circle excursions. Gliders (shown on right) move freely about the array in the upper water column. C-PIES are seafloor instruments with one instrument tethered in the water column above. The SWOT cal/val array consisted of 4 S moorings, 7 P moorings, and 2 gliders. The subsequent NOPP-GIW array had 1 S mooring and 9 C-PIES.

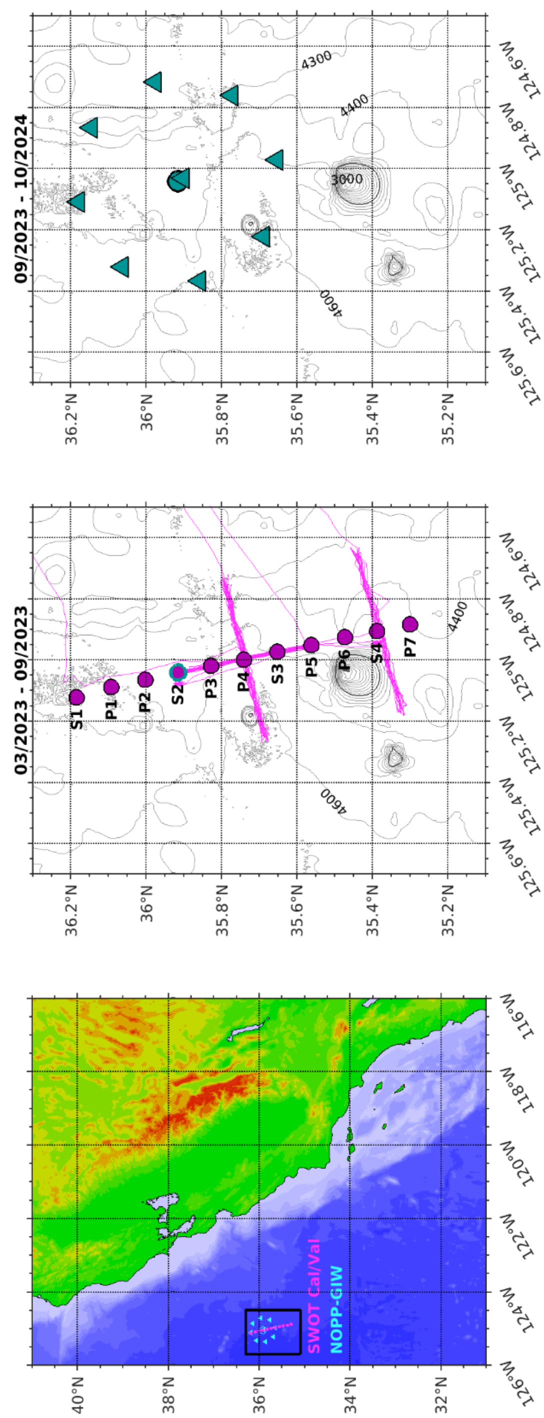

Figure 2: Maps of the array locations. The overview (bottom panel) shows the general location off the coast of California. The SWOT cal/val array (middle panel, purple color) was deployed approx. March-Sept. 2023 and consisted of 11 moorings (circles) and 2 gliders (lines). The NOPP-GIW array (top panel, teal color) was deployed for approx. one year starting Sept. 2023, and consisted of two successive deployments of one central mooring (circle, continuation of S2) and 9 C-PIES (triangles). Black numbers denote bathymetry in meters.

Table 1: List of instrumentation and data types deployed on the SWOT cal/val and NOPP-GIW arrays.

Platform / Data Type	**Deployed Feb/Mar - Sep/Oct 2023**		**Deployed Sep 2023 - Apr 2024 as part of NOPP-GIW**	**Deployed Apr - Oct 2024 as part of NOPP-GIW**
	as part of SWOT	**as part of NOPP-GIW**		
Ship-based CTD casts with water samples before/after mooring deployments	yes (to cal/val the mooring instruments)		yes (to cal/val the mooring instruments)	yes (to cal/val the mooring instruments)
Moorings with profiling and fixed CTD instruments	yes (11 moorings: P1-7, S1-4) P moorings: profiler + 1 fixed CTD S moorings: profiler + 7 fixed CTD		yes (1 mooring, repeat deployment of S2)	yes (1 mooring, repeat deployment of S2)
Additional temperature instruments on mooring	no	yes, S2 only, 23 instruments	yes, 25 instruments	yes, 25 instruments
Additional current instruments on mooring	no	yes, S2 only, 3 profiling and 4 fixed instruments	yes, 3 profiling and 4 fixed instruments	yes, 4 profiling and 4 fixed instruments
Additional barometers (air pressure) on surface buoys	yes, on P3 and P6	no	no	no
Gliders with CTD instruments	yes (2 gliders)	no	no	no
C-PIES seafloor instruments	no	yes (3, but 2 failed)	yes, array of 9 (8 C-PIES and 1 PIES)	
Additional science-quality GPS instruments on surface buoys	yes (experimental), on 7 P moorings	no	no	no

Many of the datasets of this collection use NetCDF files, a binary file format which includes metadata inside the files along with the data. Metadata fields can refer to the entire file ("global attributes") or to one variable only, thereby providing the capability to annotate certain content specifically. Typically, NetCDF files used here will follow the CF and ACDD metadata conventions. CF in particular provides a robust logic that georeferences the data, i.e. links the observations to time as well as location in latitude, longitude, and the vertical direction. Some other noteworthy features of these conventions are:

- To understand what physical property is described in a variable, read its "standard_name" attribute. The content comes from a standardized list that uniquely identifies this.
- To understand what units the data in a variable are given in, read its "units" attribute. It is common practice to use units that can be understood by the "UDUNITS" software package.
- Linkages between variables, such as from one data variable to its associated quality control flags, are made using the "ancillary_variables" attribute.
- Global attributes such as "license" and "acknowledgement" provide information about use restrictions.

For detailed information about NetCDF-compatible software and metadata conventions, the reader is referred to the documentation listed in the references for NetCDF, CF, ACDD, and UDUNITS.

Mooring-Based Temperature and Salinity Data

Water column temperature and salinity from the mooring array was the primary data used for the SWOT cal/val efforts, and therefore, these data have been processed with particular care. All of the mooring temperature and salinity data are given in NetCDF file format, with metadata that follows the CF and ACDD conventions. Data files are organized into three processing levels (L1 through L3), which mimic those generally used in satellite data processing (such as those from NASA DPL in the references) and are defined as follows:

L1 data are instrument data converted to physical units and merged with the relevant metadata. L1 data do not have adjustments to calibrations or scientific quality control applied. L1 data are meant for data provenance only, and likely not useful for outside scientific users unless they want to get involved in calibration issues. L1 data are disseminated only for the SWOT cal/val mooring data, not for the otherwise similar NOPP-GIW data.

L2 data are still at the native instrumental resolution in time and space. Data calibration adjustments have been applied, which are based on the ship-based CTD cast data and water samples. Quality control (QC) has been applied, and all metadata are merged with the data. Data from multiple instruments may be merged into a single file at L2 if the instruments sampled at the same times, otherwise they will be in separate files. The QC information is included with the data such that "bad" data is not actually removed, but instead, the data are linked with QC flags that indicate whether a data point is good or bad. Metadata fields called "flag_masks", "flag_values", and "flag_meanings" as per the CF conventions describe how to read the QC flags (their meaning can be encoded as individually defined numbers, or as a number computed bitwise from multiple true/false tests). Generally, QC flag encoding is such that values of zero denote data that have passed all tests, but the definition of all other numbers varies between files. Figure 1 shows that there are profiling instruments that move up and down the mooring wire, and ones that are fixed on the mooring wire. As the mooring itself moves about in the ambient currents, even the "fixed" instruments change depths as the whole structure tilts. For many applications, it will be necessary to use the actual depths (included in the data files as pressure records), as opposed to the nominal instrument depths. L2 data are suitable for outside scientific users that want data at the original resolution, but requires users to deal with different time axes, vertical mooring motion, and the QC flags.

L3 data are gridded data on a consistent depth-time grid, whereby "depth" is given as pressure. Multiple versions of L3 data are disseminated, for slightly different use cases: One version has one file per SWOT cal/val mooring, and the time axis is such that each profile from the profiling instrument in the upper water column constitutes one time step. Start and end times of this time step are included in the data files (linked via the "bounds" attribute, as per CF conventions). Deep data are averages from matching time intervals. This version of L3 data preserves the highest possible time resolution that can be accomplished by the upper-ocean profilers, and it is necessary to separate the data into individual files because the profilers did not sample simultaneously between moorings. The time axes are not constantly spaced in these files, because some of the profilers sampled at varying speeds. Another version of L3 data has data from all 11 SWOT cal/val moorings merged into a single file, with somewhat coarser resolution in depth and time as a compromise between the multiple moorings. This L3 data version has constantly spaced time steps. Finally, there is a dedicated L3 data version for the NOPP-GIW mooring, that merges the three successive deployments at this mooring site (the initial deployment as S2 during SWOT, and two stand-alone NOPP-GIW ones) into a single file. This latter file also incorporates data from additional temperature loggers to increase the vertical resolution at mid-depth (see table 1 for a list of instruments and deployment durations). L3 data are meant as the primary end-user datasets, and users still have to figure out which of the versions is most suitable for them. A user should consider what time resolution is needed, whether constantly spaced time steps are needed, and if data from more than one mooring at a time are needed.

Table 2 explains the file names, and how they can be used (in addition to the location of files in various folders) to identify what data and processing level is in which file. Figure 3 provides an overview of data availability from the 11 mooring sites in depth (pressure) and time. Obvious features are the continuation at the S2 site that lasts for a year beyond the SWOT cal/val phase, and the difference in depth coverage between the S and P moorings (full depth vs. upper water column only, resp.). Other differences are due to data gaps and occasional instrument malfunctions. Consistent simultaneous data from all 11 moorings exists for a time period of several weeks in May/June 2023, and of data from all 4 S moorings for several months, approximately April to September, 2023.

For a quick start, the caption for figure 3 includes an example of how to recreate one of the figure panels in MatLab with a few lines of code, based on an L3 data file.

Table 2: File naming concept for mooring temperature and salinity data. The user can understand the basic content of the files from the names alone; additional information comes from the location of the files in different folders and the global attributes inside the NetCDF files.

Project: SWOTPOSTLAUNCH or NOPP-GIW	**Proc. Level**	**Platform**	**Instrument**	**Start Date**	**End Date**	**Mode** (always DM for "delayed mode")	**Version**
`SWOTPOSTLAUNCH_L3_ALL-MOORINGS-MERGED_CTD_START20230218_END20231007_DM_VER02.nc` This file contains L3 data from CTD instruments on all 11 moorings.							
`SWOTPOSTLAUNCH_L3_MOORING-S1_CTD-MERGED_START20230303_END20230917_DM_VER03.nc` This file contains L3 data from multiple CTD instruments on the S1 mooring.							
`SWOTPOSTLAUNCH_L2_MOORING-S1_CTD-FIXED_START20230303_END20230917_DM_VER03.nc` This file contains L2 data from the CTD instruments fixed on the wire in the deeper parts of the S1 mooring.							
`SWOTPOSTLAUNCH_L1_MOORING-S1_CTD-ZNOM-2523M_START20230303_END20230917_DM_VER03.nc` This file contains L1 data from one CTD instrument on the S1 mooring (the one at nominal depth 2523 m).							
`NOPP-GIW_L2_MOORING-CA1-01_T-LOGGER-ZNOM-1153M_START20230920_END20240412_DM_VER02.nc` This file contains L2 data from a temperature instrument on one of the later NOPP mooring deployments (the instrument fixed on the wire at nominal depth 1153 m).							

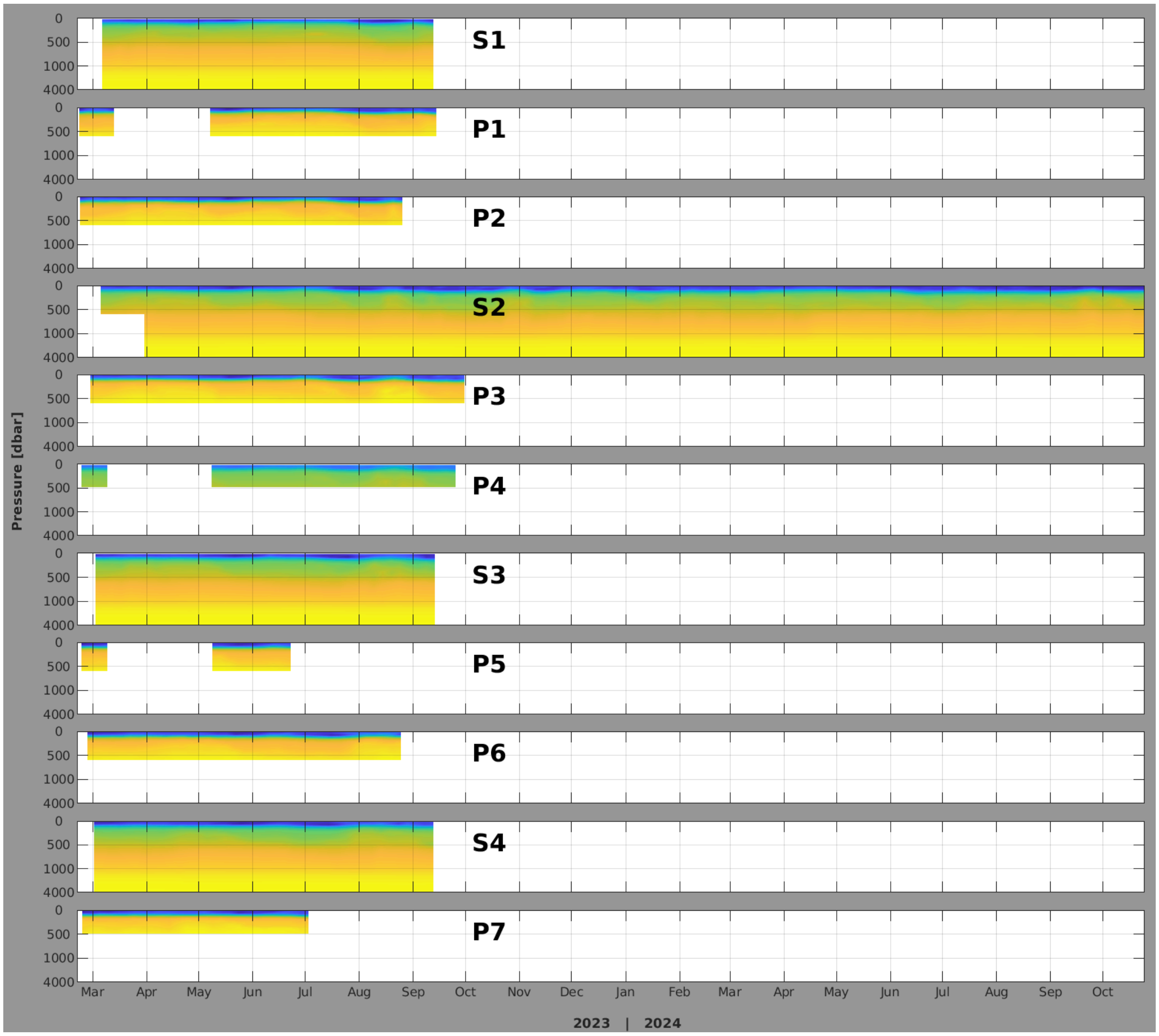

Figure 3: Depth-time plot of salinity data available in L3 products at all 11 mooring sites. Blue shading indicates low salinity near the surface, orange and yellow colors show higher salinity at depth. The purpose of this graph is to document data availability. A figure similar to the S2 panel can be produced in MatLab with the following lines of code, assuming the data file has been downloaded locally:

```
filename = ...
    'NOPP-GIW_L3_MOORING-S2-AND-CA1_MERGED_START20230303_END20241028_DM_VER02.nc';
tim = ncread(filename, 'TIME') + datenum(1950,1,1);
p = double(ncread(filename, 'PRESSURE'));
s = double(ncread(filename, 'SALINITY'));
figure;
pcolor(tim, p, s);
shading flat
set(gca, 'ydir', 'reverse');
datetick('x');
```

Glider-Based Temperature and Salinity Data

Two oceanographic gliders were deployed as part of the SWOT cal/val effort. Gliders are autonomous vehicles that move up and down in the water column by changing their buoyancy, and while doing so, generate forward propulsion with wings and attitude control much like a glider airplane. Here, the gliders were equipped with CTD sensors for temperature and salinity, and operated in the upper 1000 m of the ocean. Each up/down motion generates a CTD profile, much like a ship-based CTD system or the moored CTD profilers but with some lateral motion in addition to the vertical profiling. Figure 2 shows the glider track lines, and figure 4 additional details of each individual glider track with the time color-coded. Parts of the missions were operated on a line connecting the moorings, which allows for cross-calibration between mooring and glider sensors as well as for filling some of the mooring data gaps visible in figure 3. Other parts of the missions were operated on lines perpendicular to the mooring array, to create data coverage in this additional direction that was not available from the mooring data alone.

Data are given in NetCDF files that follow the CF and ACDD metadata conventions. There is one file per profile, and a total of more than 3000 such files. The file names have the following naming convention:

```
ru38_20230601T000742Z_dbd.nc
```

The first four characters denote the glider ID (either "ru32" or "ru38", for "Rutgers" and their numeric IDs), and the subsequent date/time string denotes the mid-time of the CTD profile contained inside the file. The files are not organized further by processing level like the mooring data. Instead, all information that is equivalent to both L1 and L2 is included in each glider profile file. No glider data products equivalent to L3 were produced. There are a large number of variables inside each glider data file. Table 3 lists the most important variables that are equivalent to L1 (sensor data converted to physical units) and L2 (final calibration adjustments and quality control applied).

Table 3: Key content in glider profile files, and their equivalent data processing levels from the mooring temperature/salinity data.

L1	*Coordinate variables:*	**time, depth, latitude, longitude** (and **profile_time, profile_lat, profile_lon** for single-point values)
	Data variables:	**conductivity, temperature, pressure** (and **salinity, density**)
L2	*Coordinate variables:*	(like L1)
	Data variables:	**conductivity_adjusted, temperature, pressure** (and **salinity_adjusted, density_adjusted**)
	QC variables:	**conductivity_qartod_summary_flag, temperature_qartod_summary_flag, pressure_qartod_summary_flag** (and **salinity_qartod_summary_flag, density_qartod_summary_flag**) *Flag values of 1 generally denote data that have passed the QC tests.*

The QC flags inside the files can be used to reject bad and missing data points; their meaning is explained in the metadata attributes “flag_values” and “flag_meanings” as per the CF conventions and generally such that flag values of 1 denote good data. The adjustments to the conductivity (and derived salinity and density) data that differentiate L2 from L1 data were generated by cross-calibration between the glider and nearby mooring data (see figure 4 for track lines relative to mooring locations), such that the resulting temperature and salinity data are consistent across all gliders and moorings.

It is recommended that users start their glider data analyses with the L2-equivalent data shown in table 3, which is a subset of the data included in the files.

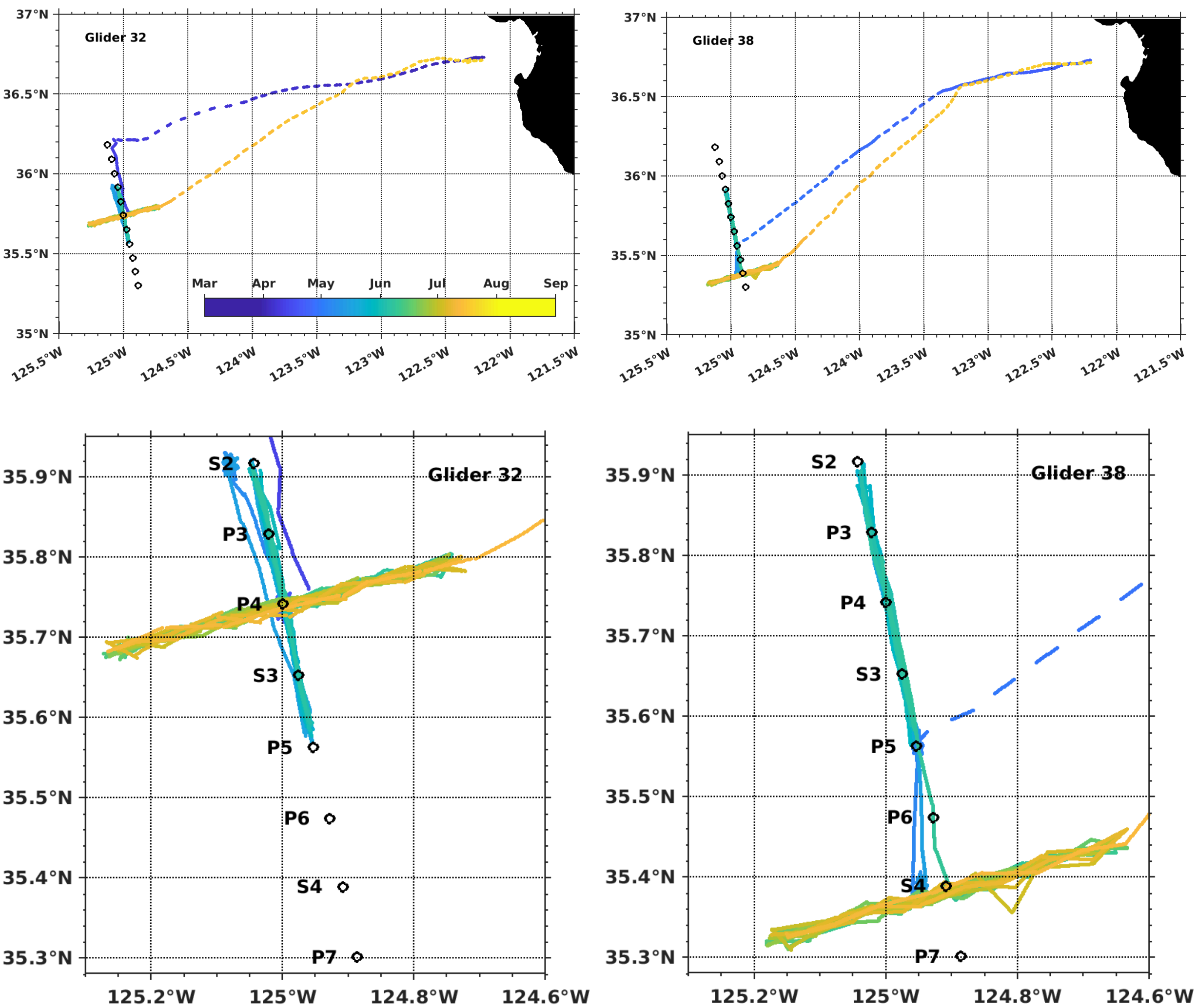

Figure 4: Tracks of the two gliders, ID 32 and 38, with the time (in 2023) color-coded. Black circles show the SWOT cal/val moorings. Both gliders transited to/from the mooring array via Monterey Bay (black land contour in top panels). Early parts of the missions (blue/turquoise) were spent in a flight pattern parallel to the moorings, and later parts (yellow/orange) were in a pattern perpendicular to the mooring array (intersecting at mooring sites P4 and S4 for the two gliders, resp.).

Mooring-Based Air Pressure Data

Two of the SWOT cal/val moorings (P3 and P6) carried barometers on the surface buoys, to collect measurements of air pressure. The data are presented in NetCDF files, named and organized into processing levels L1, L2, and L3 much like the mooring temperature and salinity data (see above): L1 files contain the basic instrumental records with metadata. L2 files contain additional QC flags. There is one L3 file that contains data from both moorings on a common time axis, based on only “good” quality L2 data. The file names are as per table 2, and contain the word “BAROMETER” as the instrument ID. No in-depth calibration adjustments or quality control were done at L2 processing, other than basic range checks and comparisons against expected pressure from weather charts. This showed periods with suspect data from the sensor on P3, which are therefore flagged in its L2 data file and cause gaps in the L3 file. The two resulting time series of air pressure, together with examples of weather charts that confirm the extreme values seen in the data, are shown in figure 5.

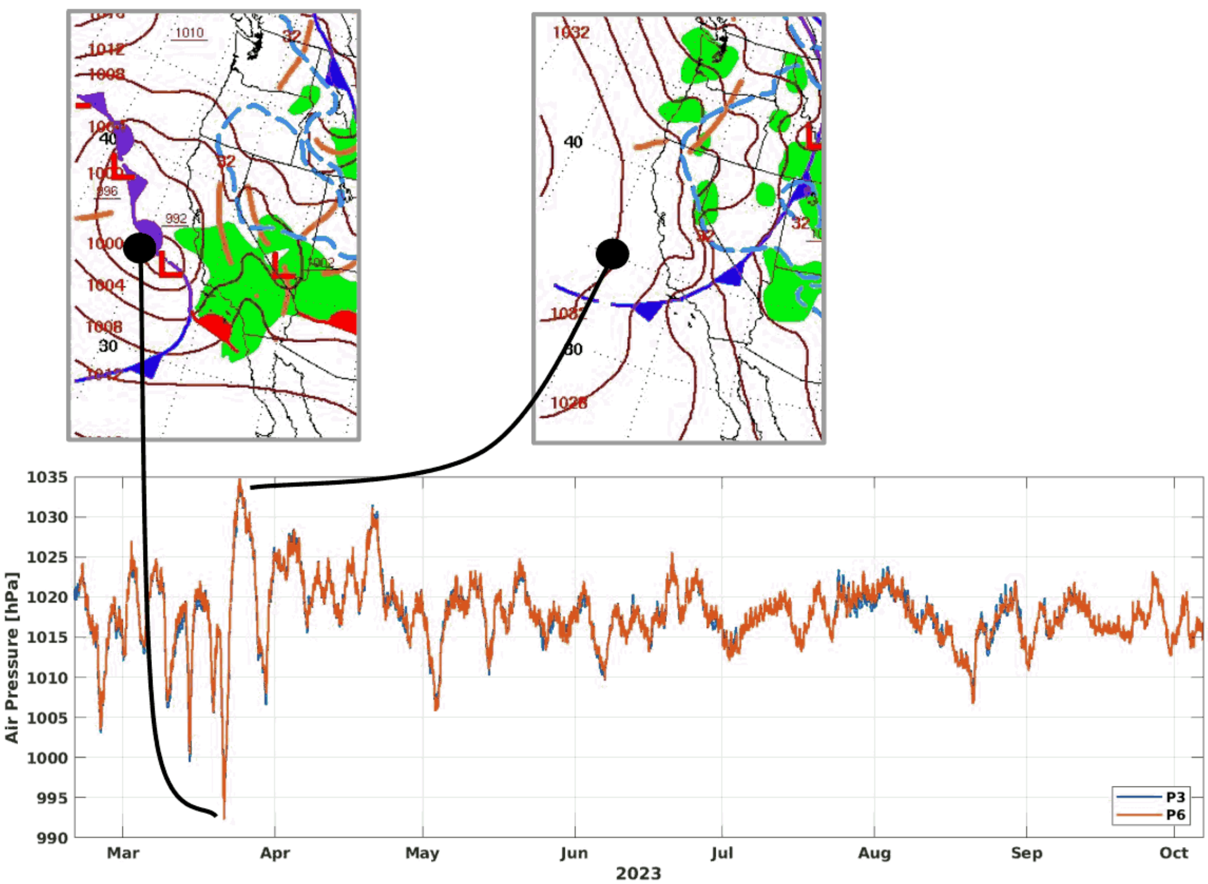

Figure 5: Time series of air pressure data from the P3 and P6 moorings (bottom panel), and comparison with two weather charts from March 2023 (top panels) to validate the reported extreme values. Weather charts are from NOAA NCEP.

Ship-Based Temperature and Salinity Data

Four of the mooring deployment/recovery cruises collected CTD cast data together with water samples of salinity. The purpose of these was to adjust the calibrations of the mooring CTD sensors to a common and well-calibrated reference, so that the moorings would obtain the best possible data to calculate steric height for the SWOT cal/val analyses. Most, but not all of the CTD casts were collected at the actual mooring sites. For some casts, mooring instruments were attached temporarily to the ship CTD system for cross-calibration, and these may have been done anywhere en route to/from the mooring sites. Due to the varying depth ratings of the mooring instruments thus attached, not all CTD casts covered the full water column. The resulting CTD data collection is an irregular pattern of sampling locations and depths, which includes a number of full-depth casts in the vicinity of the moorings. Locations where CTD casts were done, as well as cruise details and data file names, are shown in figure 6 and table 4.

Water sample data are given in NetCDF files that contain each sample result and metadata. The file structure is based on a template from NOAA NCEI (their "profile, incomplete" template; see NCEI-NetCDF in the references), which follows the CF metadata conventions, and there is one such file per cruise.

CTD cast data are also given in NetCDF files, with metadata included. The data were vertically averaged to the final resolution. Conductivity (and thereby salinity) data were adjusted to best match the salinity water samples. If the ship CTD system had sensors in addition to temperature and salinity (such as oxygen or chlorophyll concentrations), these data were included in the files. Metadata indicates what calibrations, if any, were done to those data. The file structure is based on a template from NOAA NCEI (their "profile, orthogonal" template; see NCEI-NetCDF in the references), which follows the CF metadata conventions, and there is one such file per cruise.

Table 4: List of cruises that collected CTD cast and water sample data, and their dates and associated data file names. Additional cruise-related data for the three Sally Ride cruises are available at the links listed in the references for SR2321, SR2405, and SR2420.

Ship ***Cruise ID***	**Cruise Dates**	**Number of CTD Casts**	**Data File Names**
Bold Horizon	2023-02-23 – 2023-03-05	8	Bold-Horizon_SWOT_Feb-Mar-2023_CTD-Data_Processed.nc Bold-Horizon_SWOT_Feb-Mar-2023_Water-Sample-Data_Processed.nc
Sally Ride *SR2321*	2023-09-13 – 2023-09-25	43	SR2321_CTD_Data_Processed.nc SR2321_Water_Samples.nc
Sally Ride *SR2405*	2024-04-10 – 2024-04-18	10	SR2405_CTD_Data_Processed.nc SR2405_Water_Samples.nc
Sally Ride *SR2420*	2024-10-26 – 2024-11-02	8	SR2420_CTD_Data_Processed.nc SR2420_Water_Samples.nc

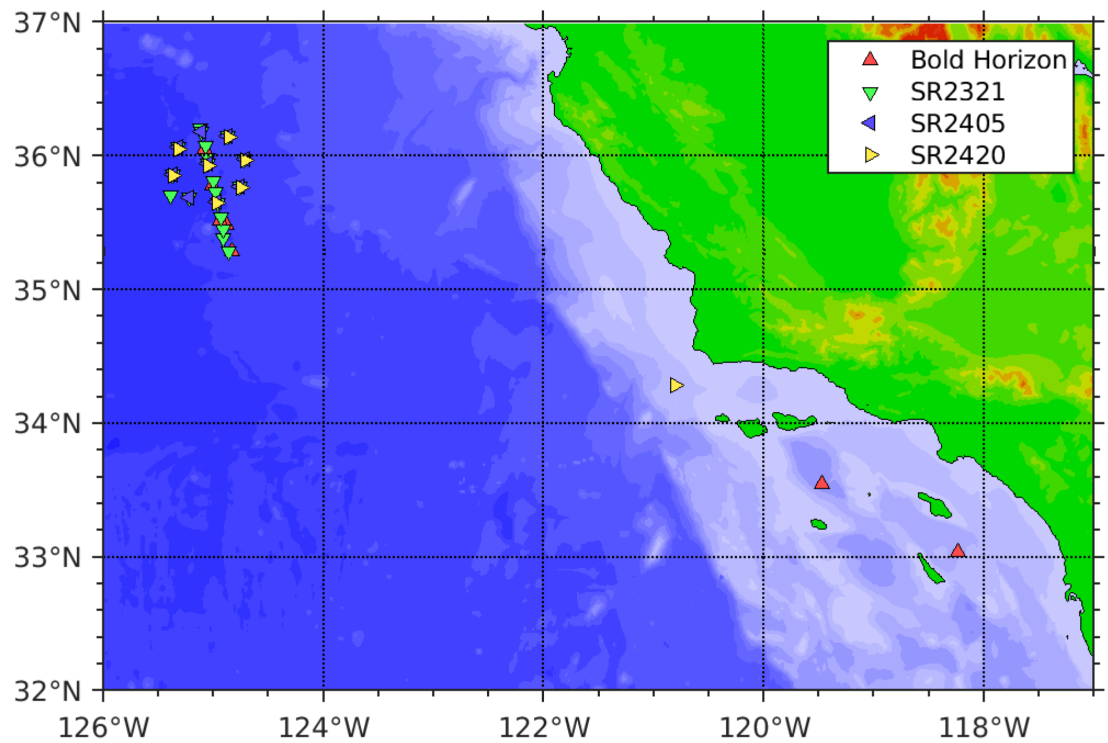

Figure 6: Map of CTD stations. The four different symbols refer to the four cruises outlined in table 4. Note a small number of stations during transit to/from the mooring array sites, in the Southern California Bight region.

Mooring-Based Ocean Current Data

[Pending release, contact: A. Waterhouse, Scripps Institution of Oceanography]

Seafloor Pressure, Current, and Inverted Echo Sounder Data

[Pending release, contact: M. Andres, Woods Hole Oceanographic Institution]

Mooring-Based Sea Surface Height (GPS) Data

[Pending release, contact: B. Haines, NASA Jet Propulsion Laboratory]

Acknowledgements

SWOT cal/val data collections were part of the SWOT mission at NASA's Jet Propulsion Laboratory, who assembled a consortium from Scripps Institution of Oceanography (PIs: A. Lucas and U. Send), Rutgers University (PI: O. Schofield), and NOAA's Pacific Marine Environmental Laboratory (PI: S. Stalin) for the field work.

NOPP-GIW data collections were funded by the U.S. Office of Naval Research (ONR) through awards N000142212215 and N000142212575, and carried out by Scripps Institution of Oceanography (PIs: A. Waterhouse, A. Lucas, and U. Send) as well as Woods Hole Oceanographic Institution (PIs: M. Andres and J. T. Farrar).

This project took advantage of software developed by UCAR/NSF Unidata (NetCDF and UDUNITS).

References

[ACDD:] Milan, A., and Graybeal, J. (committee chairs): **Attribute Convention for Data Discovery**. Published online, ESIP Wiki, 24 Aug. 2020. Accessed 9 December 2021. https://wiki.esipfed.org/Attribute_Convention_for_Data_Discovery

[CF:] Eaton, B., et al.: **Climate and Forecast Metadata Conventions**. Published online, 24 Aug. 2021, Accessed 9 December 2021. http://cfconventions.org/

[NASA DPL:] Data Processing Levels. NASA Earthdata. Published online, 13 July 2021. Accessed 9 December 2021. https://earthdata.nasa.gov/collaborate/open-data-services-and-software/data-information-policy/data-levels

[NCEI-NetCDF:] NCEI NetCDF Templates v2.0. Published online, 07 Dec. 2015. Accessed 6 May 2025. https://www.ncei.noaa.gov/data/oceans/ncei/formats/netcdf/v2.0/index.html

[NetCDF:] NSF Unidata: **Network Common Data Form (NetCDF)**. Software. Boulder, CO, USA: UCAR/NSF Unidata Program Center. DOI:10.5065/D6H70CW6 ; list of compatible software available at: https://www.unidata.ucar.edu/software/netcdf/software.html

[UDUNITS:] NSF Unidata: **UDUNITS**. Software. Boulder, CO, USA: UCAR/NSF Unidata Program Center. DOI:10.5065/D6KD1WN0

[SR2321:] Rolling Deck to Repository (R2R) data collection for: **Sally Ride, cruise SR2321**. DOI:10.7284/910183

[SR2405:] Rolling Deck to Repository (R2R) data collection for: **Sally Ride, cruise SR2405**. DOI:10.7284/910470

[SR2420:] Rolling Deck to Repository (R2R) data collection for: **Sally Ride, cruise SR2420**. DOI:10.7284/910633

[Wang et al., 2022:] J. Wang, L.-L. Fu, B. Haines, M. Lankhorst, A. J. Lucas, J. T. Farrar, U. Send, C. Meinig, O. Schofield, R. Ray, M. Archer, D. Aragon, S. Bigorre, Y. Chao, J. Kerfoot, R. Pinkel, D. Sandwell, and S. Stalin: **On the development of SWOT in-situ Calibration/Validation for short-wavelength ocean topography**. *Journal of Atmospheric and Oceanic Technology*, 2022. doi:10.1175/JTECH-D-21-0039.1.

[Wang et al., 2025:] J. Wang, A. J. Lucas , S. Stalin , M. Lankhorst , U. Send, O. Schofield, L. Kachelein, B. Haines, C. Meinig, R. Pinkel, J. T. Farrar, L.-L. Fu: **SWOT Mission Validation of Sea Surface Height Measurements at Sub-100 km Scales**. Accepted to *Geophysical Research Letters*, 2025.