

# Global Dynamic Height from Quality-Controlled Argo float profiles

Annabel Wade & Katy Christensen

April 2024

## 1 Global Argo float profiles

Global Argo float profiles were obtained from all Data Acquisition Centers (DACs) by Robert Drucker on August 29th, 2022 for the period January 2004 - August 2022, resulting in 2,383,458 profiles from 15,850 different floats. Variables within each profile include pressure (PRES), temperature (TEMP), practical salinity (PSAL), latitude (LAT), longitude (LON), datenum (DATENUM), cycle number (CYCLENUM), and float number (FLOATNUM), among others (see Table 1 for further description). The minimum and maximum latitude were 66.747°S and 76.219°N, respectively. The minimum and maximum longitude were 179.998°W and 179.998°E, respectively. Following the data modes listed in Section 1.7 of the Argo User's Manual [5], 5.27% of profiles were listed as 'R' (real-time or raw), 11.35% were listed as 'A' (real-time adjusted), and 83.38% were listed as 'D' (delayed mode). Data was extracted into the form of 80 NetCDF files (formatted as `global_profiles.nn.nc`), with each file containing data for 200 floats.

## 2 Pre-processing of data

Robert Drucker handled pre-processing of the global profiles and detailed this extensively in *Global Profiles* [1] and *Southern Ocean Profiles* [2]. During this pre-processing, PRES, TEMP or PSAL samples with \*\_ADJUSTED\_QC flags 1, 2, 5, or 8 were replaced with their adjusted values. Samples with \*\_ADJUSTED\_QC flags 3, 4, or 9 or set to FillValue were set to NaN. Samples with \*\_ADJUSTED\_QC flag 0 were left unadjusted (See Argo users manual, reference table 2). Additionally, samples with missing pressure or temperature were removed from data due to their necessity to this project. Samples with missing salinity were retained. Within each file, profiles with variable lengths were padded to the largest profile in that file. All data variables in the data files are listed in Table 1.

## 3 Further Quality Control

The pre-processed global profiles underwent further quality control, resulting in flag arrays that differentiate a variety of data completeness and usability. These flag arrays were stored in the form of 80 NetCDF files (formatted as `flag_arrays.nn.nc`, where each corresponds to the original `global_profiles.nn.nc` file). *Supplemental Tables* contain counts and percentages for all point, profile, and special area flags.

### 3.1 Point Flags

Each flag array file has a variable called `bad_pnts` that contains a flag array with the same dimensions (cycles, samples) as its corresponding data file. Only data flagged as 0 (Good data, see Table 2) in the `bad_pnts` array was able to be used in computing dynamic height. Figure 1 shows the distribution of the point flags. Since profiles with variable lengths were padded to the largest profile in that file, there are "buffer NaNs" flagged as point flag 1 (Missing or out of range pressure) so they can be identified and removed later.

Variable	Description
FLOATNUM	Float WMO ID
CYCLENUM	Profile number
DATENUM	Matlab datenum (days since 00-Jan-0000 00:00:00)
LAT	Profile latitude
LON	Profile longitude
POS_QC	Position QC flag*
POS_ADJUSTED_QC	Position adjusted QC flag*
POS_RECOVERED	Position recovered
DATA_MODE	'R', 'A', or 'D'
MAX_PRES_ADJUSTED_ERROR	Maximum value of PRES_ADJUSTED_ERROR
PI_NAME	PI name
DATA_CENTRE	Processing data center code
PRES	Pressure profile (dbar)
PRES_OUTOFRANGE	Flag for PRES out of range [0 1.2*ProfilePressure]
PRES_QC	Pressure realtime QC flag*
PRES_ADJUSTED_QC	Pressure adjusted QC flag*
TEMP	Temperature profile (deg. C)
TEMP_OUTOFRANGE	Flag for TEMP out of range [-3 30]
TEMP_QC	Temperature realtime QC flag*
TEMP_ADJUSTED_QC	Temperature adjusted QC flag*
PSAL	Salinity profile (PSU)
PSAL_OUTOFRANGE	Flag for PSAL out of range [0 42]
PSAL_QC	Salinity realtime QC flag*
PSAL_ADJUSTED_QC	Salinity adjusted QC flag*
ProfilePressure	Nominal profiling pressure max (dbar)
ProfilePressureSource	Source of ProfilePressure (see table)
ParkingPressure	Nominal parking pressure (dbar)
ParkingPressureSource	Source of ParkingPressure (see table)

\*See Argo Users manual Reference Table 2

Table 1: Saved variables in `global_profiles.nn.nc` files [2].

Flag Digit	Description	Justification
0	Good Data	Used for further analysis; was not flagged for any of the following.
1	Missing or out of range [0, 1.2×ProfilePressure dbar] pressure (PRES)	Where the * variable was NaN or the *_OUTOFRANGE variable was True in <code>global_profiles.nn.nc</code> . Flagged since valid * values within reasonable range are required for further calculations of Dynamic Height.
2	Missing or out of range [−3, 30°C] temperature (TEMP)	
3	Missing or out of range [10, 42 PSU] salinity (PSAL)	
4	Bad temperature (TEMP) points hand-picked by Annabel Wade	See section 3.4, Hand Picked Points.
5	Bad salinity (PSAL) points hand-picked by Annabel Wade	

Table 2: Point flags and justifications

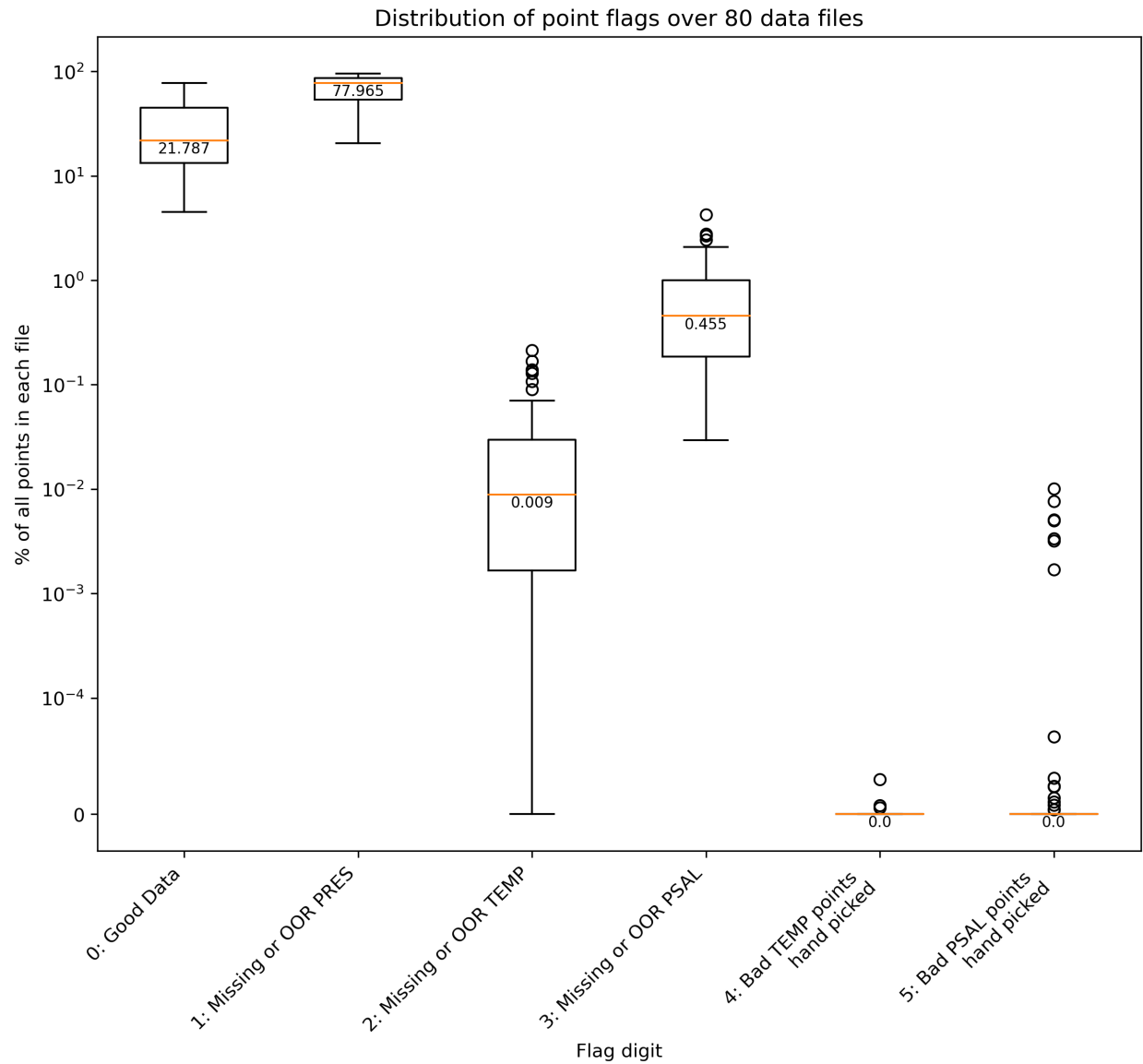


Figure 1: Box and whisker plots showing % of each point flag digit in the files. Medians for the 80 files are written below the orange median line. OOR stands for Out of Range.

### 3.2 Profile Flags

Each flag array file has a variable called `bad_prof` that contains a flag array with corresponding values for every profile in its corresponding data file. Only profiles flagged as 0-3 (see Table 3 and Section 3.5) in the `bad_prof` array were able to be used in computing dynamic height. 88.09% of profiles across all files were flagged as good data (flag 0) or correctable (flags 1-3). Figure 2 displays the distribution of these profile flags among all 80 `global_profiles.nn.nc` files.

Flag Digit	Description	Justification
0	Good Data	Used for further analysis; was not flagged for any poorly structured or unusable data.
1	Repeated or out of order pressure	The profile data needs to be in pressure-ascending order and without repeat pressures for later computations.
2	Salinity gradient (larger than $0.002 \frac{PSU}{dbar}$ below 1000 decibars)	The last point in the profile was taken as the float was beginning its ascent and can throw off Dynamic Height calculations.
3	Both flag 1 and 2	Profiles with both flags, repeated or out of order pressure and salinity gradient $> 0.002 \frac{PSU}{dbar}$ below 1000 decibars, must be appropriately corrected.
4	Four or fewer measurements in profile	Interpolation during later analysis requires at least 5 measurements.
5	Out of bounds time range (up to 2022) or no time data	Data after 2022 is not being used since it may not have initial Argo QC done. Missing time data is flagged since it is required for later analysis.
6	Missing geo-coordinates or Argo QC Position flag $> 2$	Geo-coordinates were required for later analysis. Positions with QC flag $> 2$ were flagged to avoid using bad position data and because some of the interpolated positions (according to data variables <code>POS_RECOVERED_BY_ME</code> and <code>POS_RECOVERED_BY_DAC</code> ) appeared to be incorrect and unusable (i.e. location over a continent).

Table 3: Profile flags and justifications

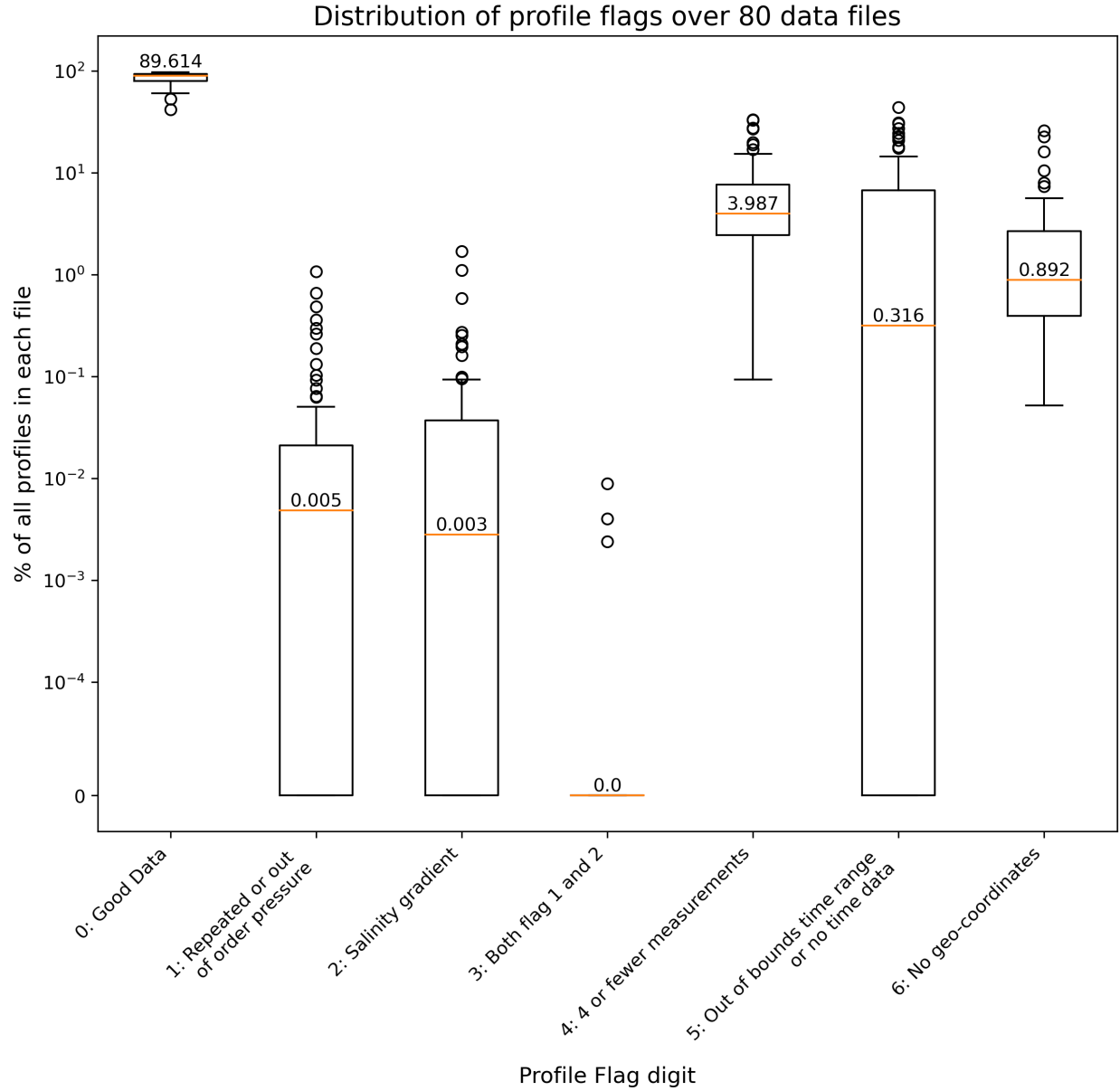


Figure 2: Box and whisker plots showing % of each profile flag digit in the files. Medians for the 80 files are written above the orange median line.

### 3.3 Special Areas Flag

Geo-coordinates in special regions were flagged for those data to be computed separately. 94.16% of all profiles were non-special, flag 0. A map of the special regions is pictured in Figure 3 and the distribution of the flags is shown in Figure 4.

Flag Digit	Area	Justification
0	Non-special area	Used for further analysis; not flagged for being in region with special conditions
1	Mediterranean and surrounding seas	Due to high salinity and temperature, must be computed separately due to their significantly different ocean dynamics.
2	Baltic Sea	Being very fresh and cold, this region must be computed separately due to its significantly different ocean dynamics.
3	Coastal Western and Indo-Pacific	Due to shallow waters and spatially limited profiles. A high spatial resolution of profiles is necessary for later computations.
4	Gulf of Mexico and Caribbean Sea	Due to shallow waters and dynamics distinct to the ocean.
5	Coastal Pakistan	This region was chosen due to a lack of float trajectory-based velocity data in this region. Having sufficient distribution of both data is necessary for our use of the Dynamic Height data.

Table 4: Special area flags and justifications

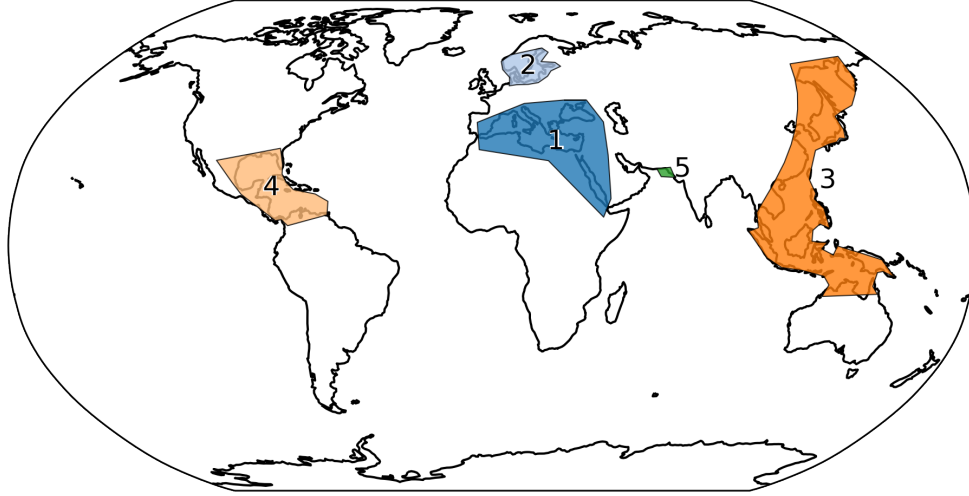


Figure 3: Special Areas of Argo float data

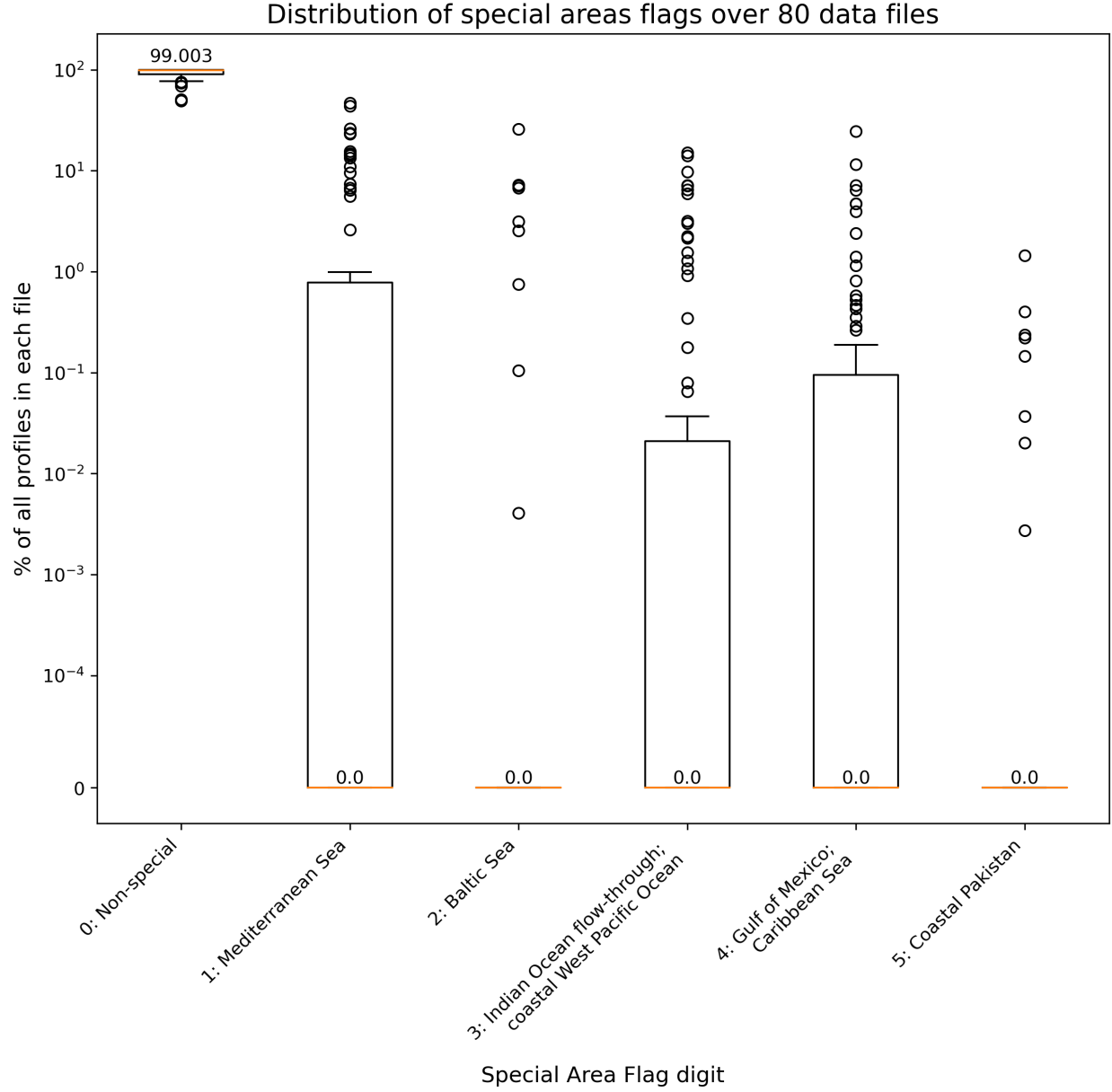


Figure 4: Box and whisker plots showing % of each special area flag digit in the files. Medians for the 80 files are written above the orange median line.

### 3.4 Hand Picked Points

Due to QC algorithms not being perfect, remaining bad points were hand-picked to ensure the best data quality. Methods to hand pick points included selecting points beyond a certain threshold away from the mean of that data beyond a certain pressure (targeting unusual variation in a pressure region), looking at the potential density profile of that data for additional reasoning, and using a custom point picker in Python to locate them within the profiles. Hand-picked points were reasoned and confirmed by Annabel Wade and Katy Christensen.

Figure 5 shows hand-picked temperature points, where panels A and B show profiles where the first

sample was too low to be a reasonable ocean temperature, and panel C shows a point chosen based on the criteria of being over 5°C from the mean of that profile’s temperature data past 750 decibars.

Figure 5 shows the hand picked temperature and salinity points. Panels D, F, G, H, I, K, J, O, P, Q, and R show profiles with hand picked points selected due to appearing unreasonable relative to the rest of the profile. Profiles in panels E and M were selected on a similar basis, being unreasonable compared to surrounding points or having no surrounding points. Profiles 4878 and 5806 in panel H, file 15 were flagged completely due to their almost constant, low-salinity values. Profiles in panels J and N yielded unreasonable dynamic height profiles, so these were flagged in this earlier stage of QC.

### 3.5 Correction of profile flags 1, 2, and 3

Profile flag 1, repeated or out of order pressure, was corrected by sorting the profiles’ samples by pressure. If repeated pressures remained in the profile, then the samples at the repeated pressures were averaged, so that only one sample remained at that pressure. Profile flag 2 was corrected by removing the last sample because the large salinity gradient at the end of the profile occurs when there is a delay in the start of the float’s water pump which records the salinity of water collected at a different pressure level. Profile flag 3 was corrected by performing both correction methods from flag 1 and flag 2 on the profile sequentially. These corrections allowed the profiles flagged 1-3 to be used in later analysis. After correcting 2727 profiles for flags 1-3, 2,099,489 profiles were considered good to use for further analysis and calculating dynamic height.

## 4 Dynamic Height

Variable	Description
DH	Dynamic height at each level ( $m^2/s^2$ )
RHO	Density at each level ( $kg/m^3$ )
LEVEL_PRES	Pressure levels (dbar) = [5, 10, 20, 30, 50, 75, 100, 125, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000]
LAT	Profile latitude
LON	Profile longitude
DATENUM	Matlab datenum (days since 00-Jan-0000 00:00:00)
FLOATNUM	WMO Float Number
CYCLENUM	Cycle Number
SPECIAL_AREAS	Flags for areas in need of special consideration (see Section 4)

Table 5: Variables of dynamic height files

Absolute salinity and conservative temperature were computed for each data point in the final 2,099,489 profiles (after correcting for profile flags 1, 2, and 3 as detailed in Section 3.5) using the Gibbs-SeaWater Oceanographic Toolbox routines [4]. Profiles of specific volume anomaly and density(RHO) were obtained using absolute salinity and conservative temperature in addition to the pressure data. Each profile was then



interpolated at a set of pre-determined pressure levels (`LEVEL_PRES`) using a `pchip` function [3]. The levels range from 5 to 2000 dbar and were selected with higher vertical resolution towards the surface and roughly 100m resolution at depth.

Dynamic height was computed at each level by integrating the interpolated specific volume anomaly function over a bin bounded by a single pressure level and the level directly below. Only levels with data occurring between the integration bounds, as well as data in one bin above and one bin below, were used; pressure levels that did not meet the criteria were set as NaN. This procedure avoids extrapolating data where there are large gaps in the profiles or missing surface/deep data. Profiles that resulted in only NaN values were removed from the data, resulting in a total of 2,095,609 profiles of dynamic height for 29 levels.

All of the dynamic height, density, and pressure levels were saved in addition to latitude, longitude, date, float number, and cycle number with file format `dynamic_height.nn.nc` for each corresponding profile data file. Finally, each of the 80 dynamic height files were compiled into a single file with file format `global_DH.yyyymmdd.nc` where `yyymmdd` corresponds to the date that they data were compiled. Descriptions of variables included in the final format are presented in Table 5.

## References

- [1] Robert Drucker. *Global Profiles*. 2020.
- [2] Robert Drucker. *Southern Ocean profiles*. 2019.
- [3] Frederick N Fritsch and Ralph E Carlson. “Monotone piecewise cubic interpolation”. In: *SIAM Journal on Numerical Analysis* 17.2 (1980), pp. 238–246.
- [4] Trevor J McDougall and Paul M Barker. “Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox”. In: *Scor/iapso WG* 127.532 (2011), pp. 1–28.
- [5] Argo Data Management Team. *Argo user’s manual*. 2022. DOI: <https://doi.org/10.13155/29825>.

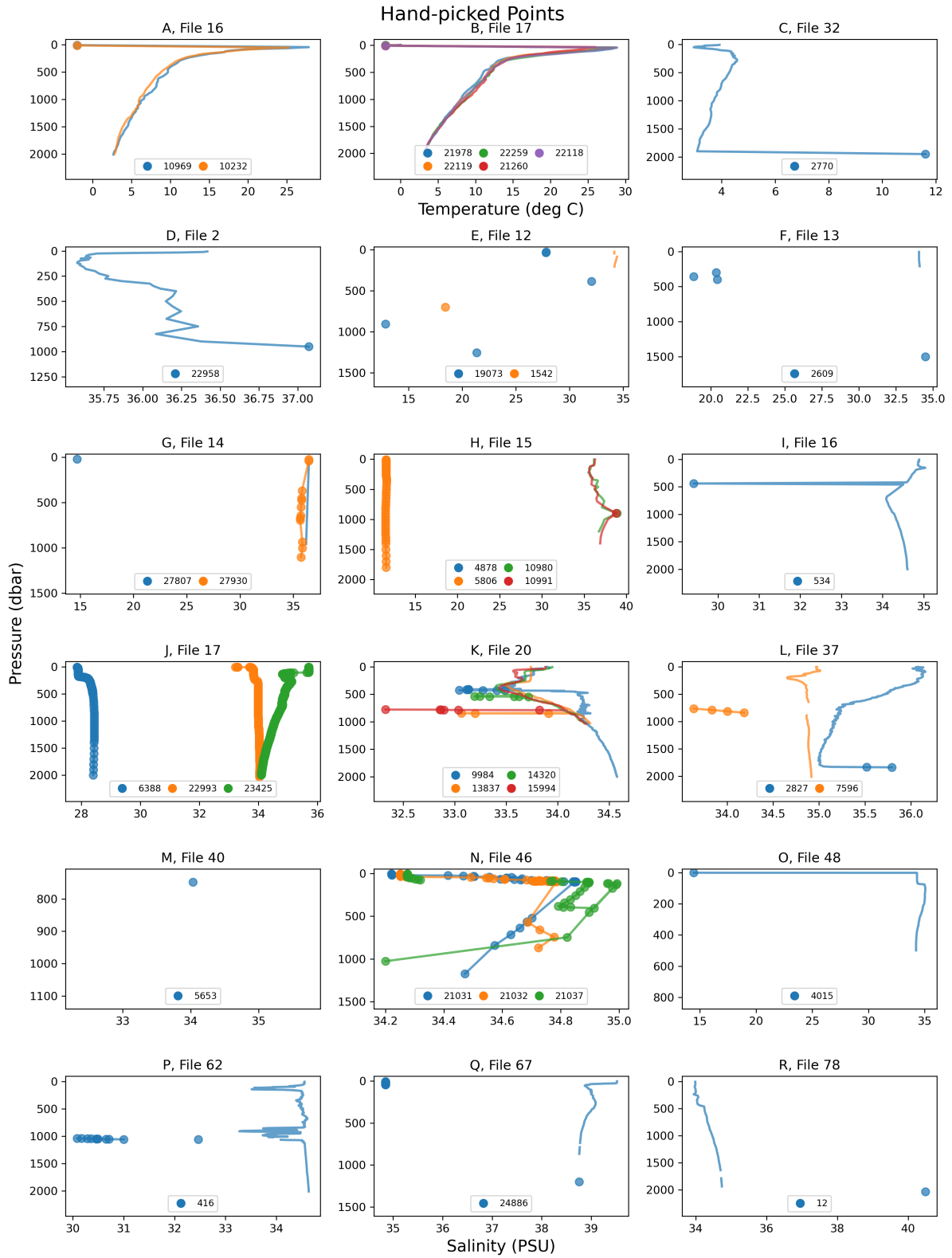


Figure 5: Temperature and Salinity profiles with hand-picked points marked as circles. Profile indices within each file are labeled in the legends of each panel.