Salinity and Temperature Variability in Oceanic Regions Worldwide

1. Introduction

The Argo Ocean Dataset is a comprehensive collection of oceanographic data collected by the global Argo program. The Argo program is an international effort that uses a network of thousands of autonomous profiling floats to measure and monitor the physical and biogeochemical properties of the world's oceans in near real-time. The floats drift with ocean currents and periodically descend to depths of up to 2,000 meters, collecting data as they rise to the surface. This memorandum utilizes data from the Argo program to examine the salinity and temperature characteristics of diverse oceanic regions worldwide.

2. Data cleaning

At a specific location, each floating device recorded temperature (C°), salinity (PSU) and pressure (dbar) at different depths while descending. Only observations that included values for all four of pressure, salinity, temperature, and location were considered. To simplify the analysis, the observed values of salinity and temperature were interpolated to 100 evenly distributed pressure values ranging from 20 to 1500 decibars using the default setup of the *scipy.interpolate.interpolate1d* function. Furthermore, only observations with geographic longitudes less than 20 were included in the analysis, focusing mostly on the Atlantic Ocean region.

After data cleaning, three datasets were obtained: one for salinity, one for temperature, and one for the location at which each floating device recorded its data. Each of the salinity and temperature datasets comprised 100 rows and 22,900 columns, where each row represented one of the evenly distributed pressure values between 20 and 1500 decibars, the columns a different floating device, and the values the interpolation of salinity or temperature at each corresponding pressure level. The location dataset comprised 22,900 rows and 2 columns, with each row representing a specific floating device and the columns the corresponding latitude and longitude coordinates. The initial rows and columns of the datasets were structured as follows:

2.1 Temperature

Row Index (Corresponding Pressure)	Device 1	Device 2
20	15.737280	26.540001
34.94949495	15.737326	26.499434

49.8989899	15.740163	25.355233

2.2 Salinity

Row Index (Corresponding Pressure)	Device 1	Device 2
20	35.163133	36.310001
34.94949495	35.162152	36.307575
49.8989899	35.162080	36.423535

2.3 Location

Row Index (Device)	Latitude	Longitude
Device 1	-37.0669	2.65910
Device 2	38.7755	-62.81379
Device 3	5.0960	-17.14700

3. Support Points

Explanation

Quantization was employed as a technique to group points with similar characteristics together. Specifically, for the salinity and temperature datasets, a reduced number of columns (support points) were created to capture the overall distribution of columns. These quantization columns were generated by minimizing the following optimization problem:

$$\frac{2}{nN}\sum_{i=1}^{n}\sum_{j=1}^{N}||y_i - x_i|| - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}||x_i - x_j||$$

Where $y_1, \cdots, y_N$ are the observed columns and $x_i, \cdots, x_n$ are the constructed ones.

The following sections show the obtained support points for each dataset and their respective interpretations.

3.1 Salinity

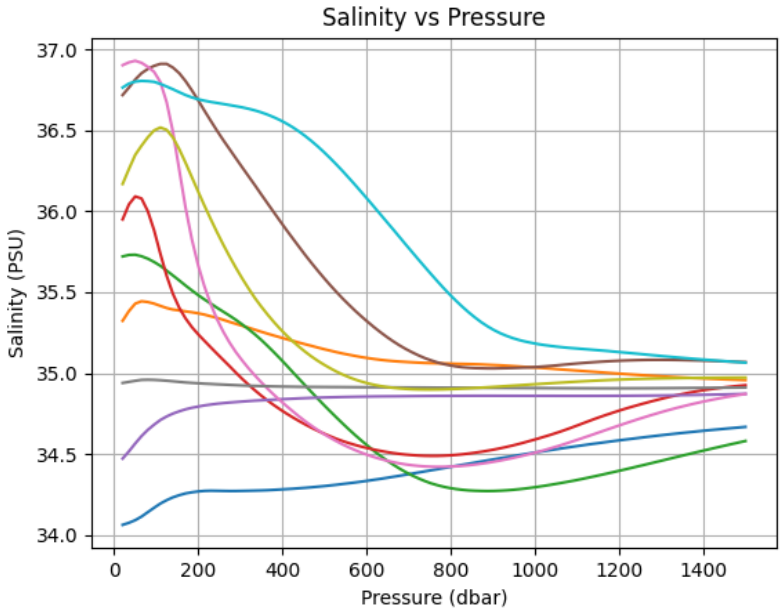
3.1.1 Support Point Statistics

10 salinity support points were obtained and the following summary statistics were calculated for each of them:

	0	1	2	3	4	5	6	7	8	9
count	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
mean	34.414186	35.122137	34.723538	34.838288	34.830004	35.530874	34.921734	34.916359	35.227100	35.803569
std	0.158869	0.147447	0.482584	0.393563	0.071583	0.656438	0.691377	0.014167	0.503229	0.676247
min	34.061291	34.956620	34.270332	34.487892	34.471208	35.028579	34.420326	34.905674	34.898209	35.064578
25%	34.279352	35.009661	34.342889	34.554239	34.837139	35.067271	34.492696	34.907477	34.929065	35.145597
50%	34.401847	35.061034	34.503715	34.736580	34.858657	35.081702	34.683605	34.910402	34.967464	35.591245
75%	34.559552	35.224201	35.100942	34.909529	34.859520	35.955097	34.868192	34.916914	35.285267	36.569212
max	34.667093	35.444113	35.731847	36.092307	34.871445	36.912045	36.931333	34.958720	36.518649	36.806759

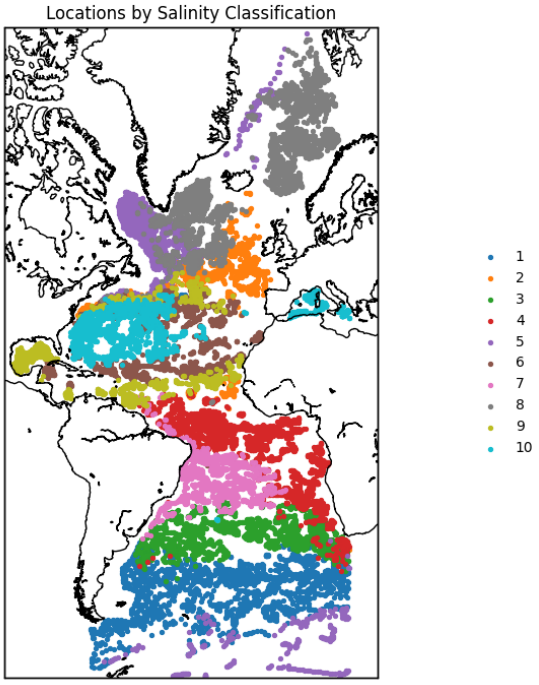
As observed in the data above, the standard deviation of salinity measurements at each pressure support point is less than 1 PSU (practical salinity unit), indicating that salinity changes are relatively small across different pressure levels.

3.1.2 Support Points vs Pressure



The above plot is a result of graphing the resulting support points against pressure. The plot above illustrates that salinity levels vary in response to changes in pressure across different support points. Some support points show an increase in salinity with increasing pressure, while others exhibit the opposite trend. Additionally, certain support points maintain a relatively constant salinity level regardless of pressure levels. Moreover, some support points show a unique pattern where salinity initially decreases as pressure decreases, but then starts to increase again.

3.1.3 Locations Classified by Closest Support Point



The plot above was created by assigning floating devices into different classifications (ranging from 1 to 10) based on their nearest salinity support point. Notably, the plot reveals that geographically close regions exhibit similar salinity characteristics. Additionally, some classifications appear in distant zones, such as the north pole and south pole areas. An interesting observation is that the classifications in the southern Atlantic are distinct and clearly separated from each other, whereas the classification zones in the northern Atlantic appear to be more mixed.

3.1.4 Support Point Comparison

By combining the two plots shown above, additional insights can be gained. The red, pink, and green classifications represent the upper part of the south Atlantic region. It is notable that in these regions, the salinity profile has a trend where salinity initially decreases with pressure, but then starts to increase at some point. On the other hand, areas closer to the poles, represented by the purple and blue classification zones, show an increase in salinity with increasing pressure. Another observation is that the gray region, corresponding to areas near the Nordic Sea, tends to have relatively constant salinity levels across different pressure levels.

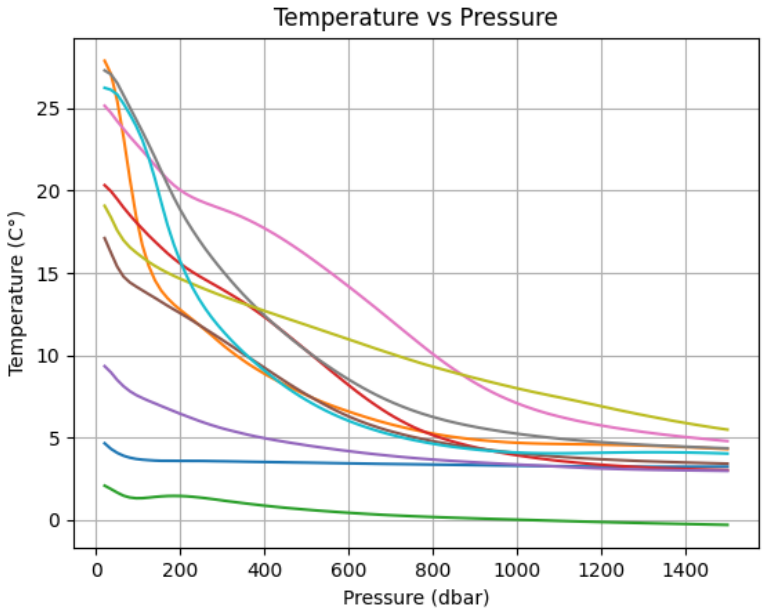
3.2 Temperature Regions

3.2.1 Support Point Statistics

	0	1	2	3	4	5	6	7	8	9
count	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
mean	3.434027	7.805263	0.443701	8.087517	4.377009	6.841474	12.168676	9.715962	10.226896	8.031110
std	0.223640	5.092786	0.618302	5.298763	1.539235	3.790965	6.258930	6.474401	3.450743	6.185380
min	3.238443	4.314019	-0.287092	3.034616	2.988764	3.431627	4.798823	4.362690	5.495982	4.042128
25%	3.265709	4.611261	-0.066978	3.519963	3.212574	3.797464	6.107339	4.874984	7.298816	4.110943
50%	3.388193	5.459114	0.230317	5.580262	3.759832	5.009393	10.894553	6.589276	9.620766	4.830928
75%	3.532692	9.034281	0.908352	12.516164	5.024838	9.417525	17.855094	12.664053	12.791969	9.285013
max	4.664456	27.888587	2.093429	20.334721	9.347203	17.120403	25.157955	27.298208	19.082153	26.247025

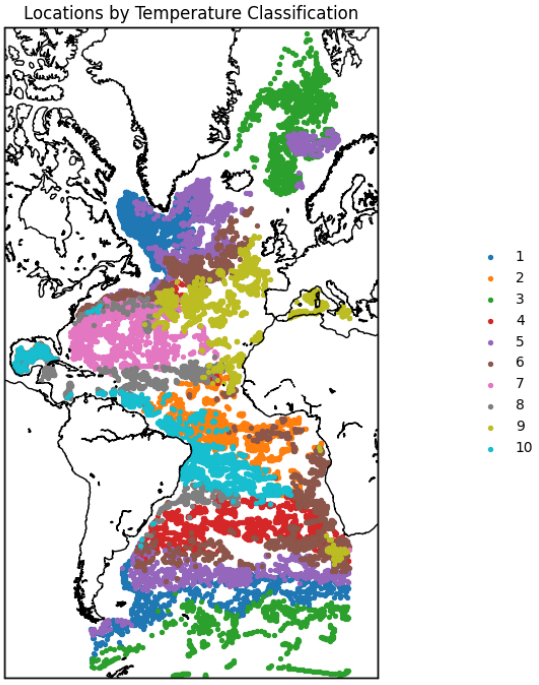
The table above indicates that the temperature support points obtained exhibit significantly greater variation compared to the salinity support points. Additionally, it is noteworthy that the maximum temperatures in each support point appear to have more variation compared to the corresponding minimum temperatures.

3.2.2 Support Points vs Pressure



The plot shown above displays the relationship between the temperature support points and their respective pressure values. The support points show a general trend where temperature tends to decrease as pressure increases. This implies that as one goes deeper into any of the analyzed regions of the ocean (pressure is higher at greater depths), the water is expected to become colder. It is important to note that the rate of temperature decrease with increasing pressure varies across support points. For example, the orange and turquoise support points show a faster rate of temperature decrease compared to the pink support point.

3.2.3 Locations Classified by Closest Support Point



Similar to the map generated by the salinity support points, the above map obtained from the temperature support points reveals that regions in close proximity to each other tend to have similar temperature characteristics. However, it is interesting to note that there are areas located far apart from each other, such as the Mediterranean area and the South of Africa region, that were classified with the same support point. Moreover, it can be observed that the classifications in the South Atlantic are more mixed compared to the classifications generated by the salinity support points.

3.2.4 Support Point Comparison

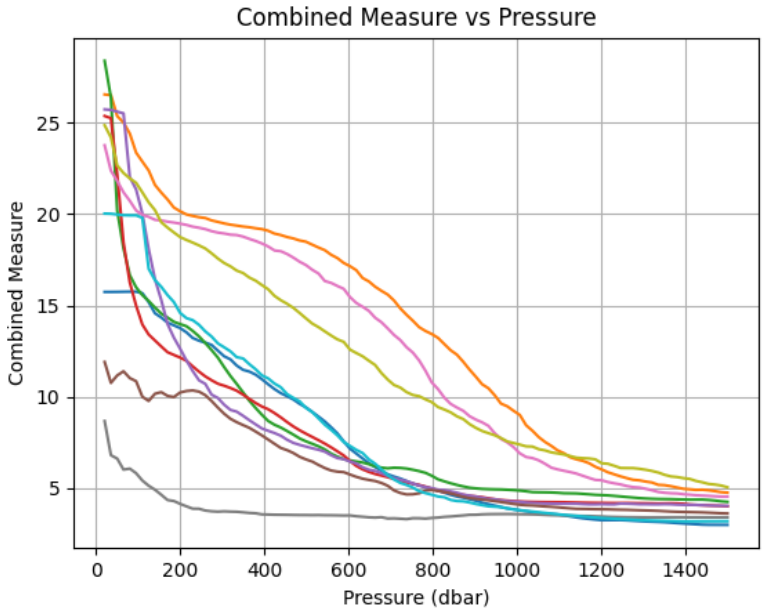
By combining the information from the two plots above, we can glean further insights. For instance, the green areas, which are close to the poles, show colder temperatures at any given pressure level and tend to decrease in temperature slowly. Furthermore, regions to the east of the United States show a gradual decrease in temperature as pressure increases. Moreover, areas in the upper part of the southern Atlantic show a fast decrease in temperature and tend to stabilize around 5°C.

3.3 Salinity/Temperature Regions

In this section, a similar analysis is conducted as in the previous sections, but with support points obtained from combining the salinity and temperature datasets. This combined dataset is expected to provide a more generalized classification of the overall data, as it represents both

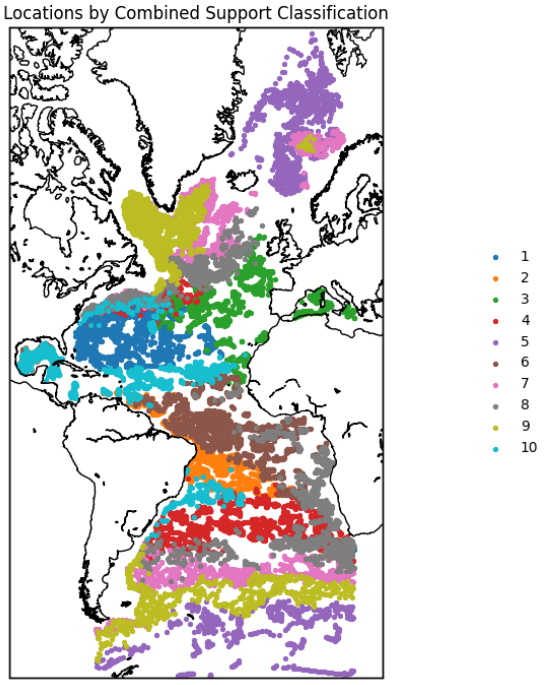
salinity and temperature simultaneously. By integrating both datasets, a more comprehensive understanding of the ocean characteristics can be gained.

3.3.1 Support Points vs Pressure



The plot above was generated by horizontally combining the salinity and temperature datasets, resulting in a dataset with 100 rows and 45800 columns. From this combined dataset, 10 support points were calculated to represent the overall data. The plot shows the resulting support points plotted against the respective pressure values, providing insights into the relationship between salinity, temperature, and pressure in the ocean. The plot above shows that the generated measure decreases as pressure increases. It is also seen that different support points decrease at different rates with respect to pressure.

3.3.2 Locations Classified by Closest Support Point



The plot displayed above is generated by vertically combining the salinity and temperature datasets, resulting in a dataset with 200 rows and 22900 columns. From this combined dataset, 10 support points were calculated to represent the overall data in a column-wise manner. The plot shows the classification of each floating device based on its closest support point. Notably, regions that are geographically close to each other tend to be classified similarly, as observed in previous plots. However, it's worth mentioning that the classification regions obtained by combining salinity and temperature differ from those obtained by considering them individually.

4. Conclusion

In conclusion, by examining the salinity and temperature datasets and calculating support points, we were able to generate plots that showcased the characteristics of different regions in the ocean. These plots revealed trends such as the relationship between temperature, salinity, and pressure, and how different regions tend to behave differently in terms of these variables. It was observed that regions close to each other tend to have similar characteristics in terms of support points, while distant regions may exhibit similar support points despite being geographically far apart. The combined analysis of salinity and temperature datasets also revealed unique classification regions that differed from those obtained by considering them individually.

5. Code

The code used for this analysis can be found here.