Exploring Temperature and Salinity Variability in the Pacific Ocean Using Principal Component Analysis

The temperature and salinity of ocean water play a big role in shaping ocean currents, climate patterns, and marine life. These factors help determine how water moves and mixes, influencing everything from weather patterns to ecosystems. However, because the ocean is so vast and complex, studying these patterns can be challenging. To make sense of the data, we use a method called Principal Component Analysis (PCA), which helps identify the most important trends in large datasets. In this study, we analyze temperature and salinity measurements from different depths across the Pacific Ocean. By focusing on the first two principal components, we can better understand how these properties vary across different regions and depths. This approach allows us to see the big picture of temperature and salinity changes, providing insights into the factors that drive these variations in the ocean.

Before analyzing the data, it is important to get an overview of what the data is that we are working with. This dataset consists of paired observations of temperature and salinity, measured at a series of depths beneath the ocean surface. Each observation contains two 100-dimensional vectors, one of temperature and one of salinity, each viewed as a function of pressure. These vectors are referred to as “profiles”. For this research goal, we will use Principal Components Analysis (PCA) to better understand these profiles. In this Argo dataset, there are 50009 pressure levels, measured in dBar along with longitudinal and latitude data for these observations in the Pacific Ocean.

Principal Component Analysis is a statistical technique used for dimensionality reduction, retaining the more important patterns in the data. In this case, we will be looking at salinity and temperature throughout the longitude, latitude, and depth of the water in the Pacific Ocean. PCA transforms a dataset into a new coordinate system where the principal components, which are linear combinations of the original variables, capture the directions of maximum variance. PCA is performed by computing the eigenvectors and eigenvalues of the covariance matrix or using singular value decomposition (SVD). For this analysis, we will perform PCA by computing the eigenvectors and eigenvalues of the covariance matrix. The first few principal components preserve the majority of the data’s variation, making PCA useful for visualizing and feature extraction.

In this analysis, we will be looking majorly at the first 2 components from PCA, drawing inferences about the temperature and salinity at the various locations and depths. A variety of plots will be analyzed, including loadings by temperature and salinity across the pressure readings and maps that contain the top 20 percent and bottom 20 percent of the 2 component scores to get an idea how temperature and salinity varies with location.

After performing PCA, we can analyze temperature variations across depth and location using a series of plots, which help clarify the significance of the first two principal components. The PC1 loading for temperature, shown in Figure 1, indicates a steep initial decline. Beyond 600 dBar, the decline slows, and the standard deviation bands converge toward the mean. For PC2 loadings in Figure 2, the standard deviation bands begin with the blue above the red, crossing near 200 dBar. Salinity follows a similar trend: the PC1 loading in Figure 3 shows a gradual increase until 200 dBar, a decline until 600 dBar, and a slight increase at greater depths. Standard deviation remains relatively consistent throughout. The PC2 loading for salinity in Figure 4 exhibits an initial standard deviation where the blue line (one standard deviation below the loading pattern) is above the mean, crossing at 400 dBar and reversing positions.

To explore geographic variations, we analyze PC scores across latitudes and longitudes using Figures 5 and 6. Temperature PC1 scores (Figure 5) are highest near the equator and lowest near the poles, with the highest values concentrated on the western side of the Pacific, while no top 20% scores appear near the Americas. PC2 temperature scores show greater variation: most of the lowest scores are concentrated around the equator, while some of the highest scores appear along the Americas. The highest PC2 scores are found in southern polar waters, with some low percentile scores just below this group. For salinity in Figure 6, the highest PC1 scores are near and slightly south of the equator, while the lowest scores are concentrated near the poles. PC2 scores follow an opposite trend, with the lowest values near the equator and the highest near polar waters.

Interpreting these results requires understanding the patterns captured by the principal component loading profiles. PC1 for temperature primarily reflects overall temperature differences: higher PC1 scores correspond to warmer waters, while lower scores indicate colder regions. This aligns with Figure 5, where higher PC1 scores are associated with warmer equatorial and subtropical regions, and lower scores with colder high-latitude waters. PC2 reveals temperature variation with depth: a high PC2 score corresponds to below-average temperatures at shallow depths and above-average temperatures at greater depths, while a low PC2 score indicates the reverse. This pattern, seen in Figure 5, suggests that high PC2 scores in the North and South Pacific correspond to deep-reaching temperature structures which could be influenced by seasonal variability, deep convection, or upwelling. Conversely, low PC2 scores in equatorial and subtropical regions indicate strong stratification, likely due to surface heating and limited vertical mixing.

The principal component analysis (PCA) of salinity reveals distinct spatial patterns in the Pacific Ocean, with PC1 capturing large-scale salinity variations and PC2 highlighting more localized differences. The first component primarily reflects the contrast between high-salinity subtropical regions and lower-salinity waters in the North Pacific and Southern Ocean, suggesting a strong influence of evaporation and precipitation patterns. In contrast, PC2 shows a different spatial pattern, with high positive loadings (top 20 percent) in the North Pacific and mid-latitudes, while strong negative loadings (bottom 20 percent) are concentrated along the equatorial Pacific and Southern Ocean. This suggests PC2 may represent changes in the vertical salinity structure, where some regions experience stronger surface salinity anomalies while others have deeper water influences. The lower PC2 score in equatorial regions aligns with areas of significant upwelling and mixing, although further investigation is needed to confirm the precise mechanisms driving these patterns.

This analysis utilized Principal Component Analysis (PCA) to examine temperature and salinity variations in the Pacific Ocean, revealing key spatial and vertical patterns. PC1 primarily captured the dominant horizontal gradients, with higher temperature and salinity values concentrated in the equatorial and subtropical regions, while lower values were found at higher latitudes. PC2 provided additional insight into vertical structure variations, highlighting differences in stratification and deep-water mixing processes. These findings enhance our understanding of the physical processes governing ocean temperature and salinity distributions, which have significant implications for ocean circulation, climate variability, and marine ecosystems.

Figures

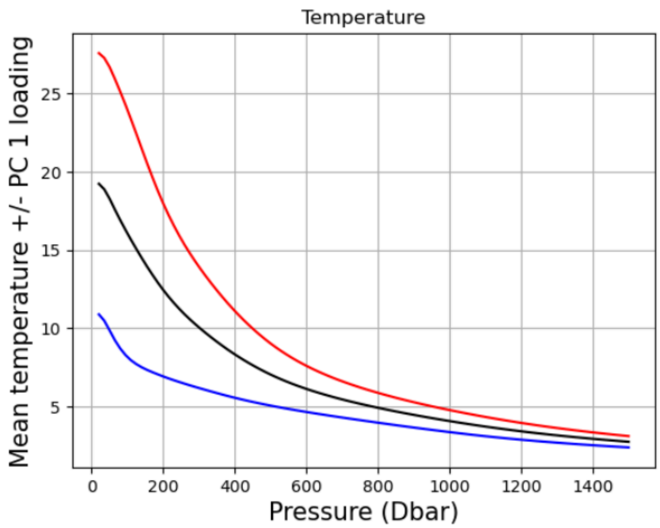

Figure 1: Mean temperature loading for PC1. The red line represents +1 standard deviation of the loading pattern while the blue line represents -1 standard deviation of the loading pattern.

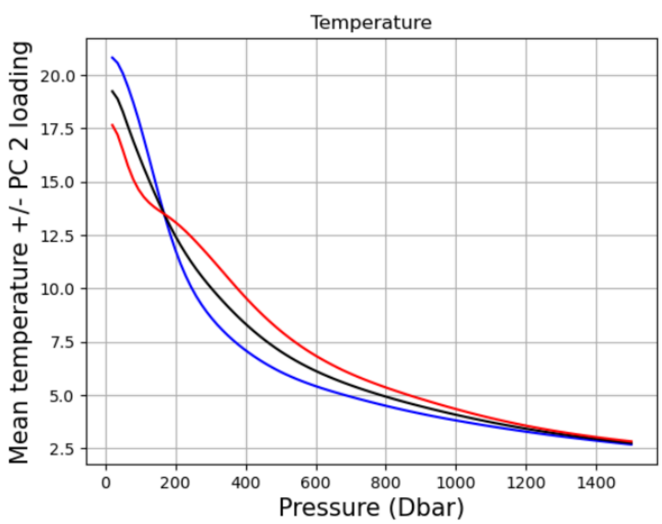

Figure 2: Mean temperature loading for PC2. The red line represents +1 standard deviation of the loading pattern while the blue line represents -1 standard deviation of the loading pattern.

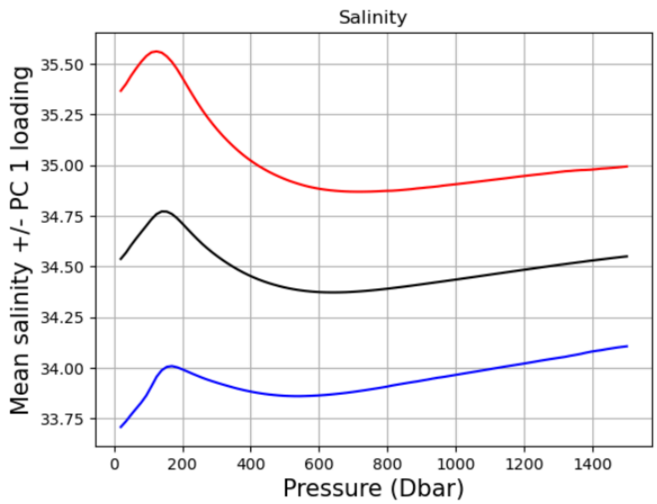

Figure 3: Mean salinity loading for PC1. The red line represents +1 standard deviation of the loading pattern while the blue line represents -1 standard deviation of the loading pattern.

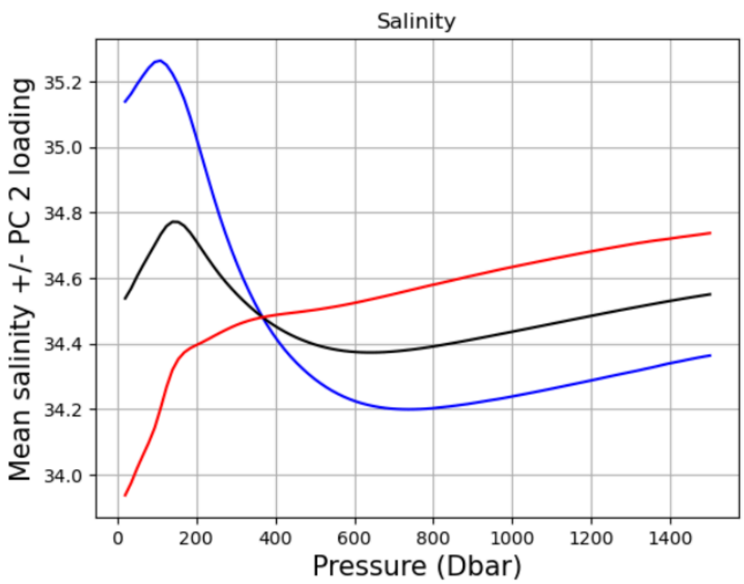

Figure 4: Mean salinity loading for PC2. The red line represents +1 standard deviation of the loading pattern while the blue line represents -1 standard deviation of the loading pattern.

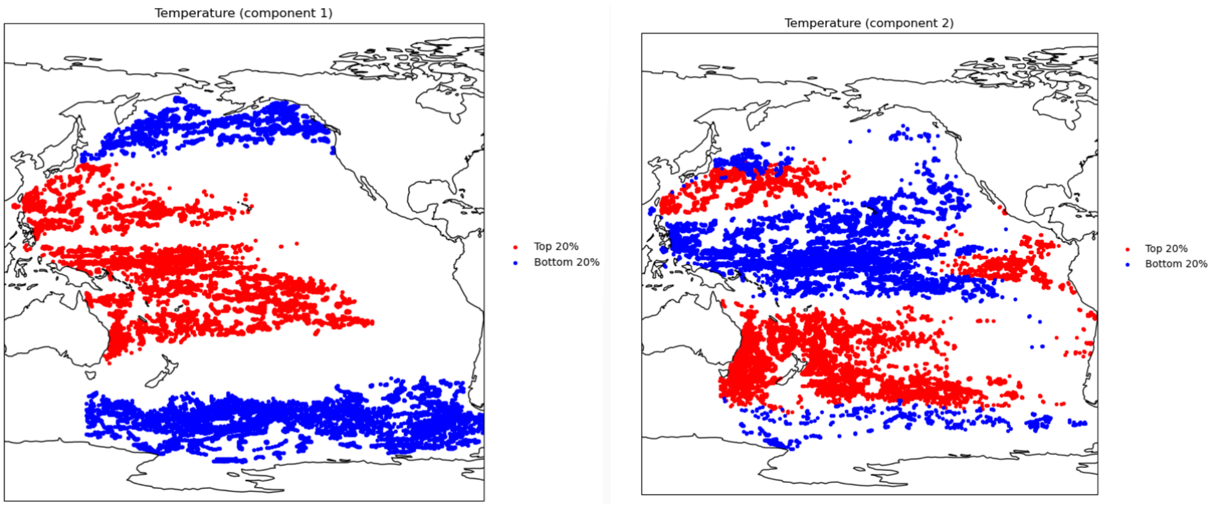

Figure 5: Top 20% and bottom 20% of PC1 scores and PC2 scores for temperature in the Pacific Ocean.

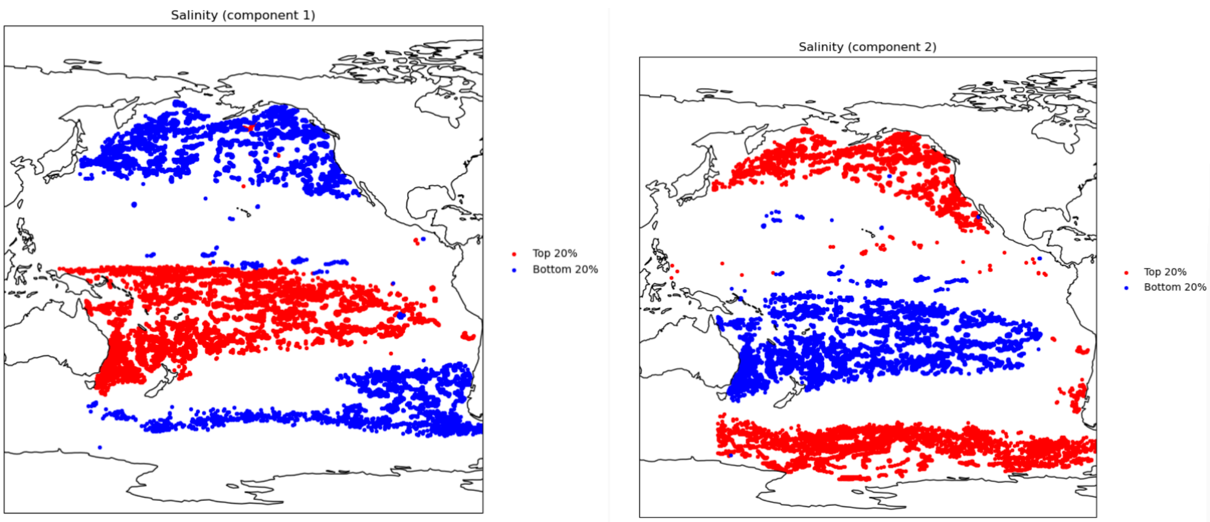

Figure 5: Top 20% and bottom 20% of PC1 scores and PC2 scores for salinity in the Pacific Ocean.

Focus	
Methods	
Writing	
Findings	