A Study of the Sea Surface Temperature and Sea Surface Salinity Regimes in the Southern Atlantic Ocean

Jiaming Chang[1], Bolun Li[2] and Zekai Ni[3]

[1]Department of Mathematics, University of British Columbia

[2]Department of Mathematics and Applied Mathematics, Ocean University of China

[3]University of Illinois, Urbana-Champaign

Instructor: Dr. Anastasia Romanou

Columbia University

July 27, 2023

Abstract

Followed by the fast advancement of computer science and technology, studies involving massive datasets, especially studies in environmental sciences, has been more accessible for researchers. Yet, with the completion of high-resolution data about the southern Atlantic Ocean region (Aulicino, 2018), there has not been sufficient academic research focusing on this specific area. In this regard, this paper will investigate the trend of SSS and SST within the southern Atlantic region with k-means clustering technique. Several quantitative analyses were used to examine the dataset from NASA Center for Climate Simulation. While previous studies examined the other regions of all five Oceans, this study took another perspective. It is found that the southern Atlantic Ocean could be divided into four sub-regions, and each of them possesses unique SST and SSS features. Secondly, this study found that there had not been significant changes in the sub-regions from year 2010 to 2014. These findings demonstrate the characteristics and trends of the southern Atlantic Ocean, and highlight the stability of such features.

Keywords: Atlantic Ocean, k-means clustering, Sea Surface Temperature, Sea Surface Salinity

Introduction

People in the contemporary world and current generation have witnessed that people are paying more and more attention on the issue of environment. Ocean, as an area that occupies most of the Earth's surface, plays an important role in the interference of balancing the environment. To protect the environment, people need to firstly know how to protect the ocean. Ocean has different functions in different areas. To figure out methods of protecting the ocean, people should conduct a comprehensive research and analysis of the ocean. Sea Surface Temperature (SST) and Sea Surface Salinity (SSS) are two key factors in analyzing the ocean. They can affect the density and flow of seawater, thereby affecting the marine environment and ecosystems near or in the ocean: The higher the temperature, the lower the density. The higher the salinity, the higher the density. Additionally, the moving and mixing of seawater can make the salinity distribution relatively uniform. Different factors are involved in the relationship between salinity and temperature. In addition, the changes in temperature in different sea areas are also helpful for exploring the issue of global warming. It can be seen that the temperature and salinity of the ocean have an important status in the analysis of the ocean and later research on the global environment. However, how to intuitively analyze the temperature and salinity of the ocean has become a problem. With the development of technology, people have many ways to measure and detect the two factors in the ocean, thus obtaining a lot of measurement data. But having these data alone is useless. People need to integrate and analyze them through other means. As the most intelligent and fastest algorithm in the world today, programming language is the best way to analyze these data. In response to this issue, the research team prepared to study the salinity and temperature in the South Atlantic Ocean. MATLAB, as one of the common tools in coding, became our important tool. This research paper will comprehensively describe the process of this research. It will explain how the researchers use MATLAB to classify the salinity and temperature of different regions in the South Atlantic and present these abstract data clearly in charts and images.

Literature review: theoretical framework on clustering analysis and previous studies on Sea Surface Temperature and Sea Surface Salinity

Many studies have shown that data analysis, especially clustering analysis, is becoming a favorable tool for researchers in assessing oceanographical datasets (Jin, 2023; Latto, 2018). Jin (2023) found that the water masses of the south China sea could be divided into different layers with distinct characteristics in SSS and SST using the k-means clustering analysis: through this method, the research team proposed a seasonal variation model of how the south China sea interacts with other surrounding water regions. Similarly, Latto (2018) deploys clustering analysis in defining the ocean carbon cycles and patterns and examined how such research method could assist environmental science fellows in heavy data-loaded researches.

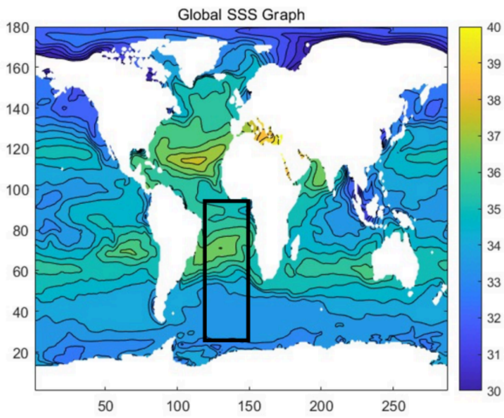
It is widely known that both Sea Surface Temperature (SST) and Sea Surface Salinity (SSS) determine many essential elements (for example, seawater density) in climate science, oceanography and other environmental studies. Knowing their importance, previous studies have been done to inspect the SST and SSS and their influence in specific regions (Bingham, 2017; Durack, 2013; Yu, 2019). Durack (2013) looked into the long-term SSS changes and proposed their conjectures for its importance to the global water cycle. Similarly, by examining the global SSS level over the range from 60°S to 60°N, Bingham (2017) suggests that instead of having a strong spacial trend, a seasonal pattern is more prevalent in the distribution of SSS in terms of global scale, which signifies the oceans' ability in regulating SSS variability. Furthermore, Yu (2019) analyzed the sea surface temperature in the south China sea, and put forth a correlation between wind trend and SST trend front. Addition to that, Yu (2019) revealed a significant correlation between the SST trend anomaly and the El Niño, suggesting the SST and its impact on surrounding region.

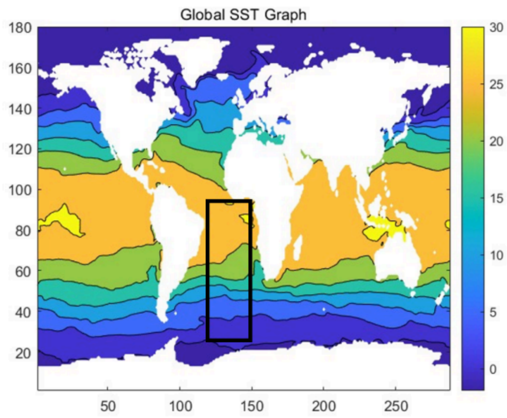
Apart from the aforementioned research, there are other researchers who investigated how SST and SSS are inter-related. The correlation coefficients of SST, SSS and other elements in the field of environmental studies are closely examined by researchers (Hidayat, 2023; Kido, 2021; Krivoguz, 2021). Hidayat (2023) had looked into the relationship between Chlorophyll-a, SSS and SST. Focusing the researching attention on area near the Bay of Bengal, a region often suffered by seasonal monsoon, the study concludes that the Chlorophyll-a has a positive correlation with SST and maintains a negative relation with SSS. Krivoguz (2021) examined the data collected in the region of black sea and concluded that various elements, such as solar radiation and fresh water influx, would impact the level of SST and SSS in a region: more specifically, the SST and SSS are in an inverse relationship due to the influences from other environmental components. Kido (2021) found that from a global scale, the covariability between SST and SSS possesses a notable spatial scale dependence: most of the positive SST-SSS correlation coefficient are situated in the extratropic region and eastern tropical pacific, suggesting that the particularity of such relationship.

Although past studies have investigated SST and SSS, whether in regional studies or from a global perspective, there has been little research focusing on the Southern Atlantic region, where detailed datasets have been collected but never been utilized (Aulicino, 2018). Using the method of clustering analysis, this study seeks to address the question: what is the relationship between the SSS and SST in the Southern Atlantic region, and what are its implications?

Methodology

This research is based on the method of machine learning: clustering analysis. Firstly, the research team utilized the data of sea surface temperature (SST), sea surface salinity (SSS), and the corresponding longitude, latitude every month in the past 55 years to make the contour plots about the global value of SST and SSS. Based on the two plots, the researchers chose the South Atlantic to be the research subject (latitude: 65°S to 0°, longitude: 60°W to 30°W, the area is marked within the black box in the first graph) because the values of SST and SSS change quickly in this area. To facilitate the research and avoid the burden of massive calculation, the team chose to analyze the data within the last five years in the datasets (2010-2014).

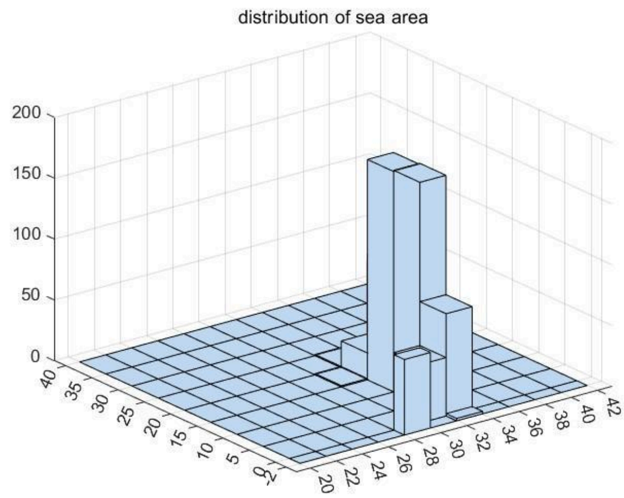




Secondly, the datasets were processed. The researchers applied two different methods to accomplish this step:

Method 1: the team directly reshaped the form of the data of SSS and SST to reduce the dimensions and transform the 4-dimensional datasets into 2-dimensional matrix forms. Then, the team utilized the k-means clustering to divide the datasets into several clusters and plotted the images of every cluster.

Method 2: the data of SSS and SST were divided into several intervals, and reshaped into a matrix after counting how many sea areas fit each region. This enabled the researchers to turn the data into a histogram density graph. And k-means clustering was applied to the density graph.



Note that, to determine the number of the clusters, the team started with an initial guess (denoted as k), and then change the numbers of the clusters based on the performance of the results: Firstly, the researchers were required to observe the images of every cluster. Then, if two or more images are similar, the number of clusters were to be reduced in to $k-1$; if no images are similar, the number of clusters were to be increased to $k+1$ to make a further guess of a finer result. Finally, the steps were repeated until a suitable number of clusters was found.

Compared with Method 1, the images Method 2 produced reflected the centroids of clusters. By contrasting the location of centroids, researchers were able to adjust the algorithm appropriately. On the contrary, the Method 1 only reported the clusters, thus, the researchers did know the optimum number of clusters and could only modify the code with trial and error.

Lastly, through analyzing these plots, the team produced results pertaining to variation of SST and SSS during a period time and some features of South Atlantic seawater.

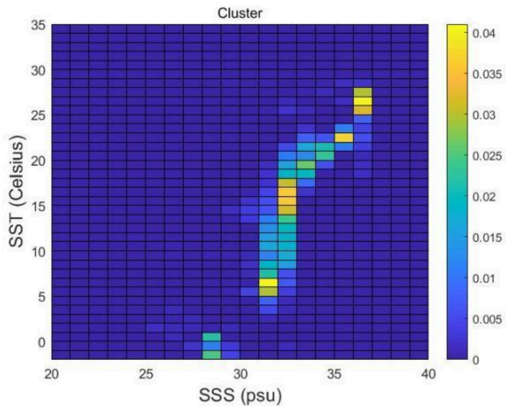
Result

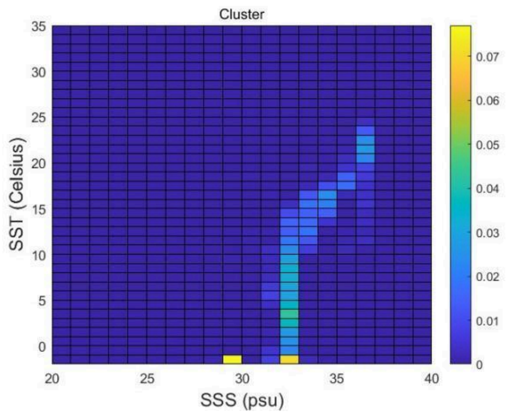
In this study, analysis on those SSS and SST data showed that the southern Atlantic Ocean tends to have a spatial regionality based on their unique SSS and SST features, and the regions within the southern Atlantic Ocean underwent no significant changes from year 2010 to year 2014. The details are elaborated as follows.

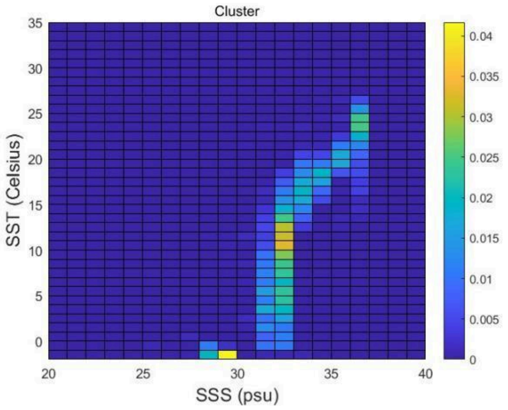
1. The variation of SSS and SST within five years

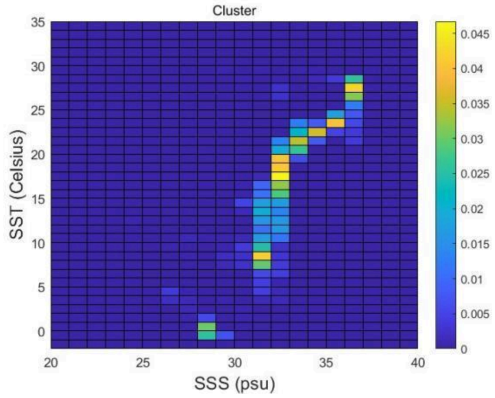
The researchers chose datasets of 5 years (from 2010 to 2014) to analyze the variation of SSS and SST.

Utilizing method 2 to process data, the researchers divided SSS and SST into narrow intervals to make it easier to detect the effect of clustering. The plots are as follows (year 2010):









The research team chose the grids with the highest ocean area values (the yellow grids in the plots) and calculated their average temperature and salinity (the middle point of every grid represents its temperature and salinity).

Through the calculation of every cluster of every year, the data reads as follows:

year	Cluster 1	Cluster 2	Cluster 3	Cluster 4
2011	(31, -1.5)	(31.75, 8.25)	(33.625, 19.875)	(33.8, 17.5)
2012	(29.5, -1.5)	(31, -1.5)	(33, 12.67)	(33.375, 18.75)
2013	(31, -1.5)	(31.75, 8.25)	(33.625, 19.875)	(34.055, 17.5)
2014	(30.75, -1.5)	(31, 4.5)	(33.64, 19.64)	(34, 19)
2015	(31, -1.5)	(31, 4)	(33.64, 22.8)	(34, 18.8)

$$(x, y)\ represent\ (\overline{SSS},\ \overline{SST}).$$

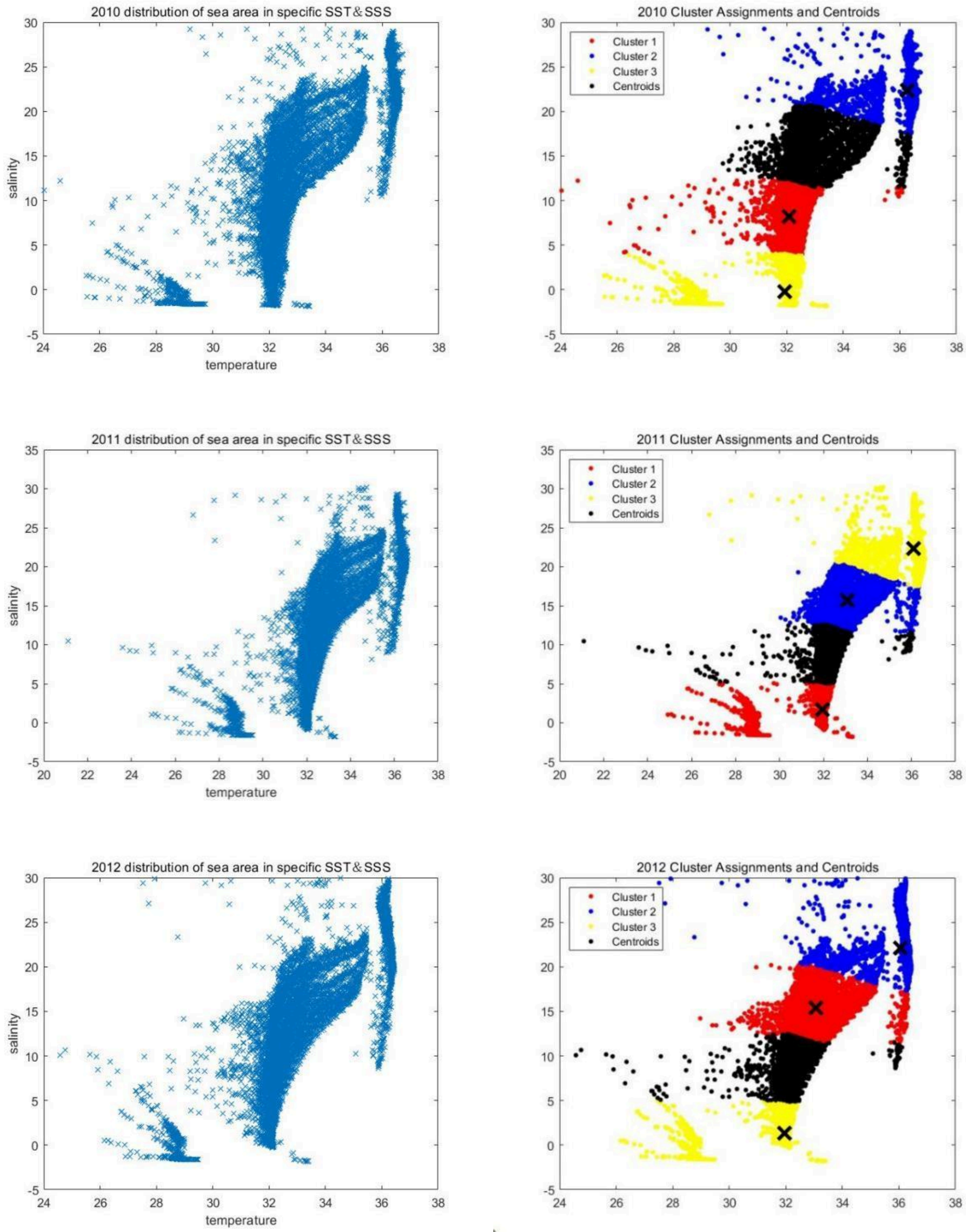
$\overline{SSS}$: the mean value of SSS, $\overline{SST}$: the mean value of SST.

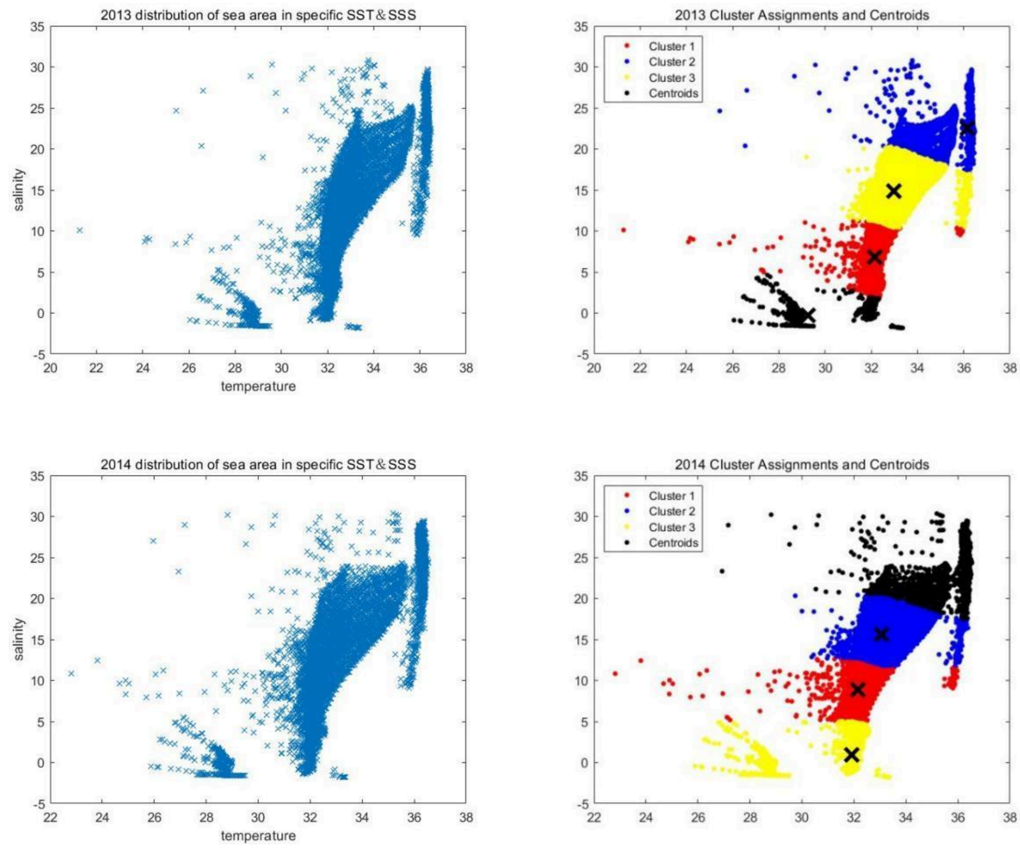
$\overline{SSS}\ and\ \overline{SST}$ in the same cluster showed no obvious fluctuation in the five years.

According to the above analysis, the team concluded that within five years, the sea surface temperature and sea surface salinity did not change and tended to be stable.

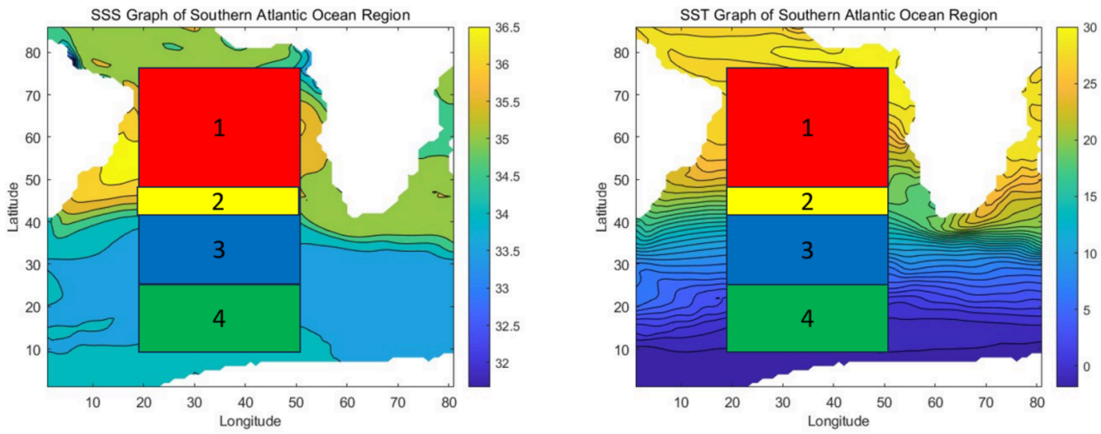
2. **Classify the south Atlantic Ocean**

Utilizing method 1 to cluster the data, the researching team calculated the average range of SSS and SST of every cluster in the five years.





Based on the average SSS and SST range of each cluster, the corresponding regions were colored in the SSS and SST regional map.



Finally, the team classified the South Atlantic Ocean into four clusters according to five-year datasets. Each region has different features. Region 1 (marked in red) has high temperature and high salinity. Region 2 (marked in yellow) has warm temperature and high salinity. The third region (marked in blue) has moderate temperature and low salinity. Region 4 (marked in green) has the lowest temperature and mostly freshwater.

Conclusion

Overall, this study tried to address the connection between SSS and SST in the southern Atlantic Ocean and its implications. This study had identified four different oceanographical regions in the aforementioned ocean area through the k-means clustering technique. Moreover, through examining the changes of the regional clusters from 2010 to 2014, the research team concluded that their characteristics underwent no significant changes during the last five years.

Discussion

This study is limited by the researchers' coding ability. The lack of proficiency in MATLAB had become a hurdle for the team to conduct the research. Equipped with better knowledge in coding, this study could have made a conclusion based on a more detailed and improved result. The time also sets up a limitation to the research. With only two weeks of time, the researchers decided the research question in a hurry. Once the result came out, the research team did not have sufficient time to assess the quality of clustering, such as using the silhouette function in the MATLAB. Additionally, the research is built only on one of the datasets from the NASA database. more datasets could provide a broader insight into the relationship of SST and SSS in the Southern Atlantic Region.

Lastly, this study could inspire more research topics pertaining to the region. There could be research focusing on the rivers or glaciers that provide freshwater influx into the region and how that influences the SSS and SST. The Southern Atlantic region is also known as a place without the Inter-Tropical Convergence Zone, which often engenders tropical cyclones that might evolve into hurricane, monsoon and typhoon. Research could be done on examining such characteristics and its correlation with the SST/SSS trends in that region. Further research could also investigate the change in SST/SSS trend in a longer time span, because this research only looked at data from 2010 to 2014.

References

Aulicino, G., Cotroneo, Y., Ansorge, I., van den Berg, M., Cesarano, C., Maria Belmonte Rivas, & Estrella Olmedo Casal. (2018). Sea surface salinity and temperature in the southern atlantic ocean from south african icebreakers, 2010–2017. *Earth System Science Data, 10*(3), 1227-1236. https://doi.org/10.5194/essd-10-1227-2018

Bingham, F. M., & Lee, T. (2017). Space and time scales of sea surface salinity and freshwater forcing variability in the global ocean (60 degrees S-60 degrees N). *Journal of Geophysical Research. Oceans, 122*(4), 2909-2922. https://doi.org/10.1002/2016JC012216

Durack, P. J., Wijffels, S. E., & Boyer, T. P. (2013). Long-term salinity changes and implications for the global water cycle. *International geophysics* (pp. 727-757). Elsevier Science & Technology. https://doi.org/10.1016/B978-0-12-391851-2.00028-3

Hidayat, M., Wafdan, R., Ramli, M., Muchlisin, Z., & Rizal, S. (2023). Relationship between chlorophyll-a, sea surface temperature, and sea surface salinity. *Global Journal of Environmental Science and Management, 9*(3), 389-402. https://doi.org/10.22034/gjesm.2023.03.03

Jin, S., Nie, X., Wang, G., Teng, F., & Xu, T. (2023). Analysis of the distribution and seasonal variability of the south china sea water masses based on the K-means cluster method. *Journal of Marine Science and Engineering, 11*(3), 485. https://doi.org/10.3390/jmse11030485

Kido, S., Nonaka, M., & Tanimoto, Y. (2021). Sea surface Temperature–Salinity covariability and its Scale‐Dependent characteristics. *Geophysical Research Letters, 48*(24), n/a. https://doi.org/10.1029/2021GL096010

Krivoguz, D., Semenova, A., & S, M. (2021). Spatial analysis of seasonal patterns in sea surface temperature and salinity distribution in the black sea (1992-2017). *IOP Conference Series.Earth and Environmental Science, 937*(3) https://doi.org/10.1088/1755-1315/937/3/032013

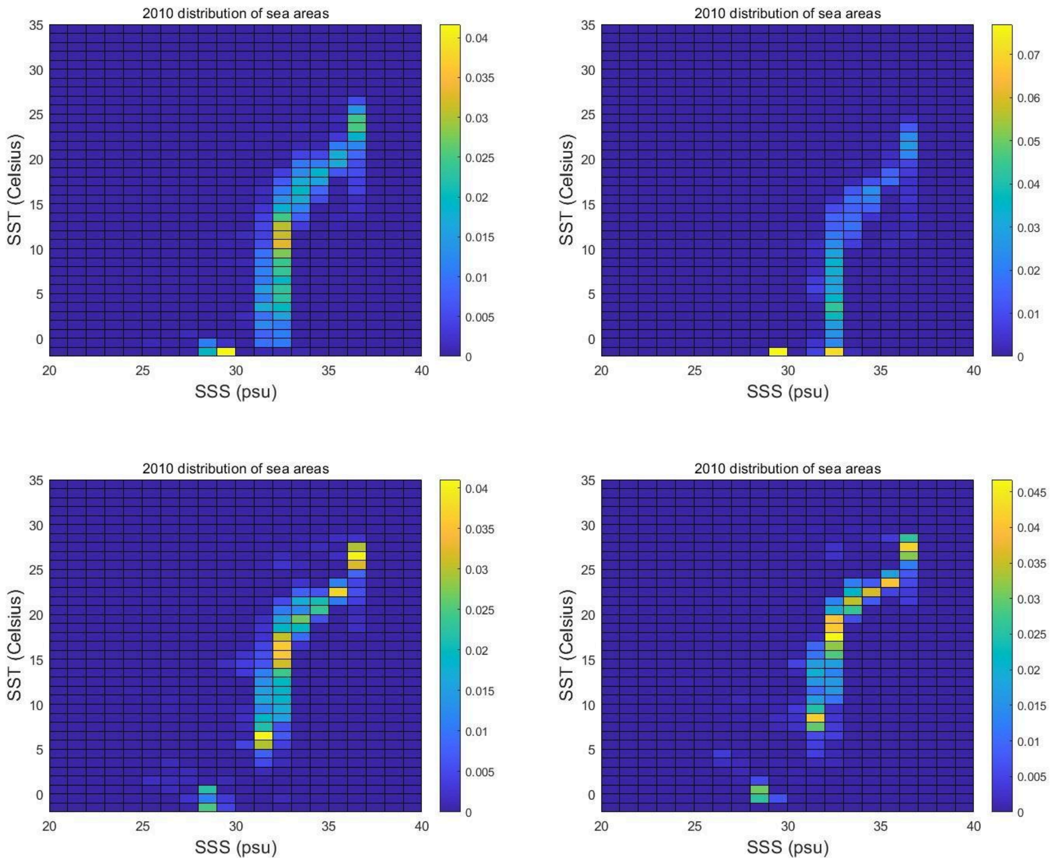
Latto, R., & Romanou, A. (2018). The ocean carbon states database: A proof-of-concept application of cluster analysis in the ocean carbon cycle. *Earth System Science Data, 10*(1), 609-626. https://doi.org/10.5194/essd-10-609-2018

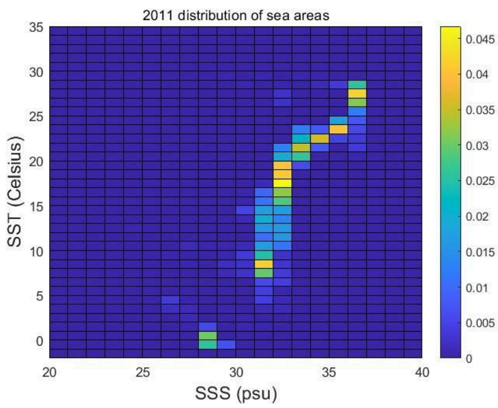
NASA, portal.nccs.nasa.gov/datashare/modelE_ocean/CMIP6output/. Accessed 28 July 2023.

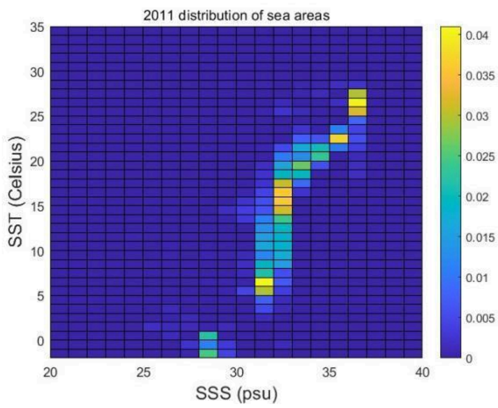
Yu, Y., Zhang, H., Jin, J., & Wang, Y. (2019). Trends of sea surface temperature and sea surface temperature fronts in the south china sea during 2003–2017. *Acta Oceanologica Sinica, 38*(4), 106-115. https://doi.org/10.1007/s13131-019-1416-4

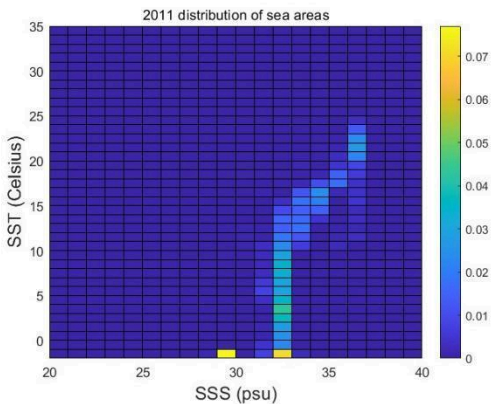
Appendix

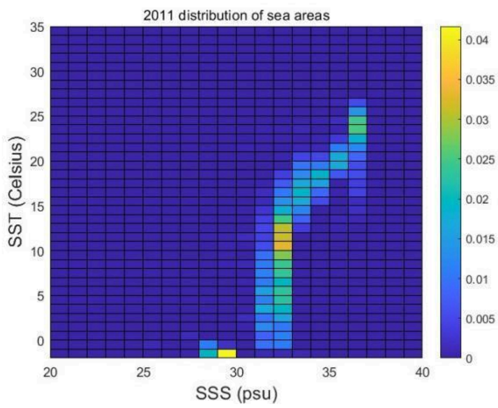
Distribution of sea areas:

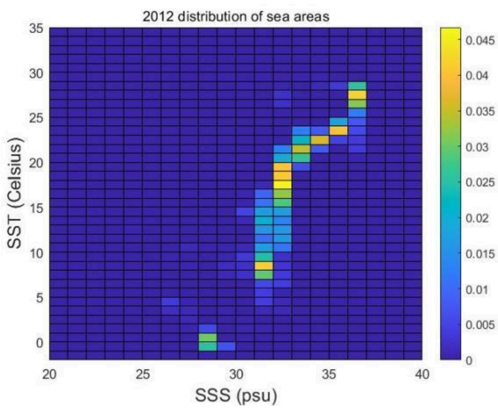


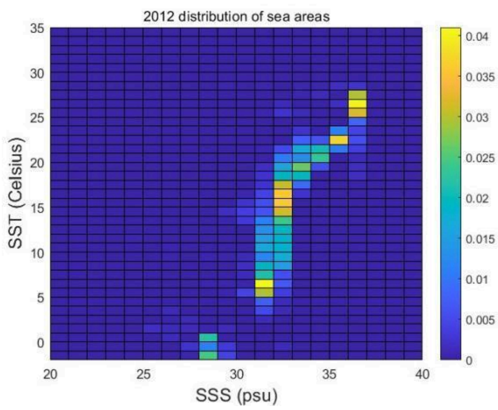

2011 distribution of sea areas
SST (Celsius)
SSS (psu)

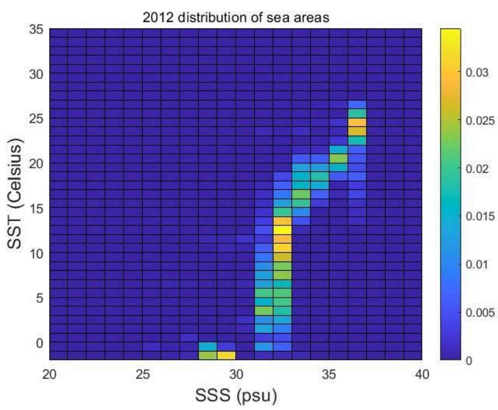

2011 distribution of sea areas
SST (Celsius)
SSS (psu)

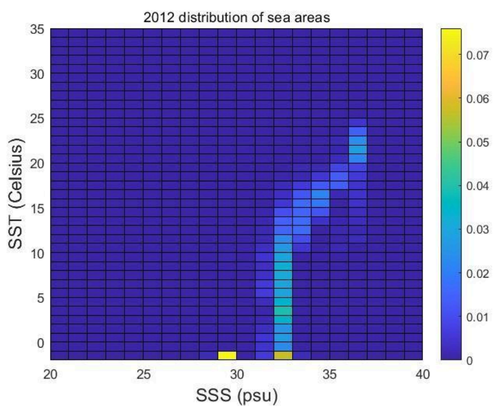

2011 distribution of sea areas
SST (Celsius)
SSS (psu)

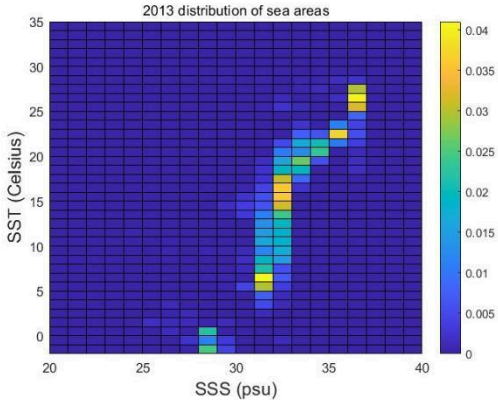

2011 distribution of sea areas
SST (Celsius)
SSS (psu)

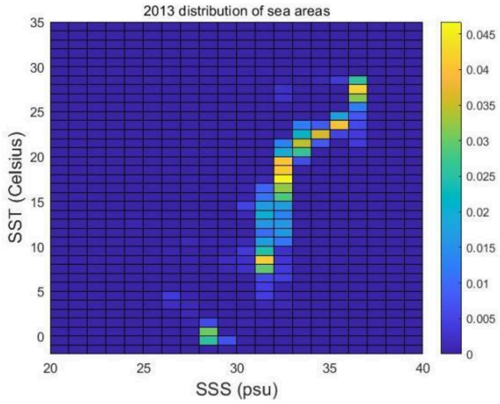

2012 distribution of sea areas
SST (Celsius)
SSS (psu)

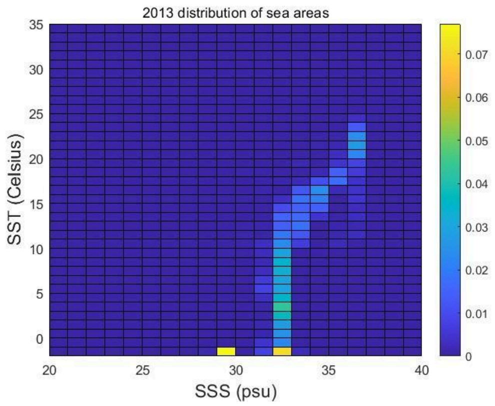

2012 distribution of sea areas
SST (Celsius)
SSS (psu)

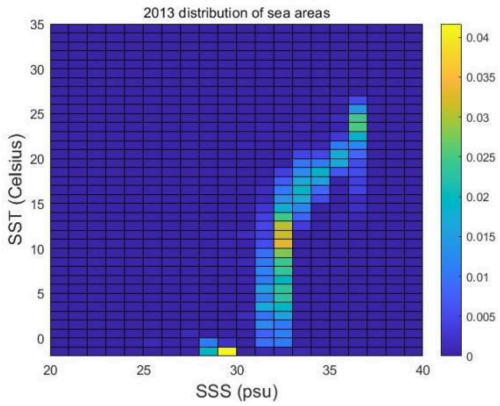

2012 distribution of sea areas
SST (Celsius)
SSS (psu)

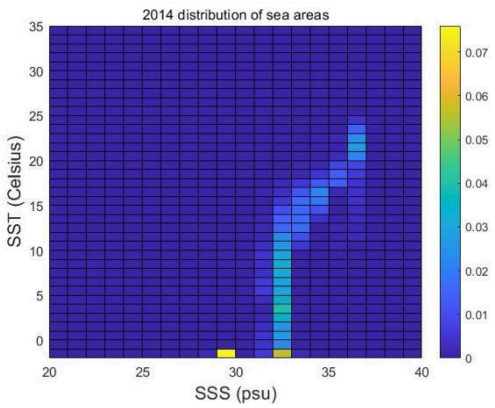

2012 distribution of sea areas
SST (Celsius)
SSS (psu)

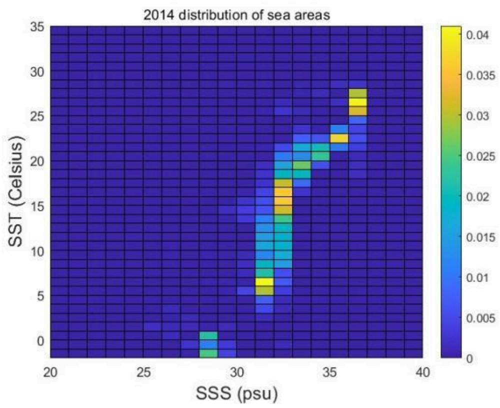

2013 distribution of sea areas
SST (Celsius)
SSS (psu)

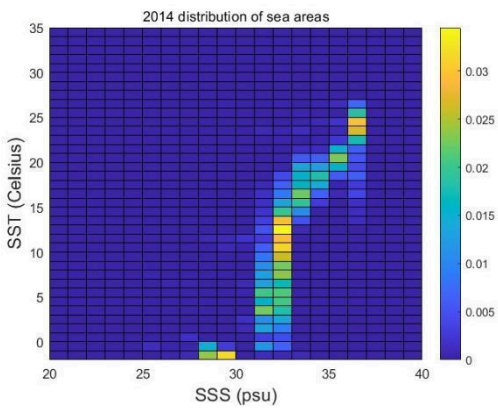

2013 distribution of sea areas
SST (Celsius)
SSS (psu)

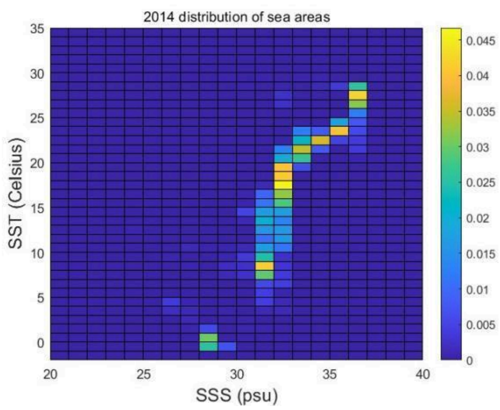

2013 distribution of sea areas
SST (Celsius)
SSS (psu)


2013 distribution of sea areas
SST (Celsius)
SSS (psu)

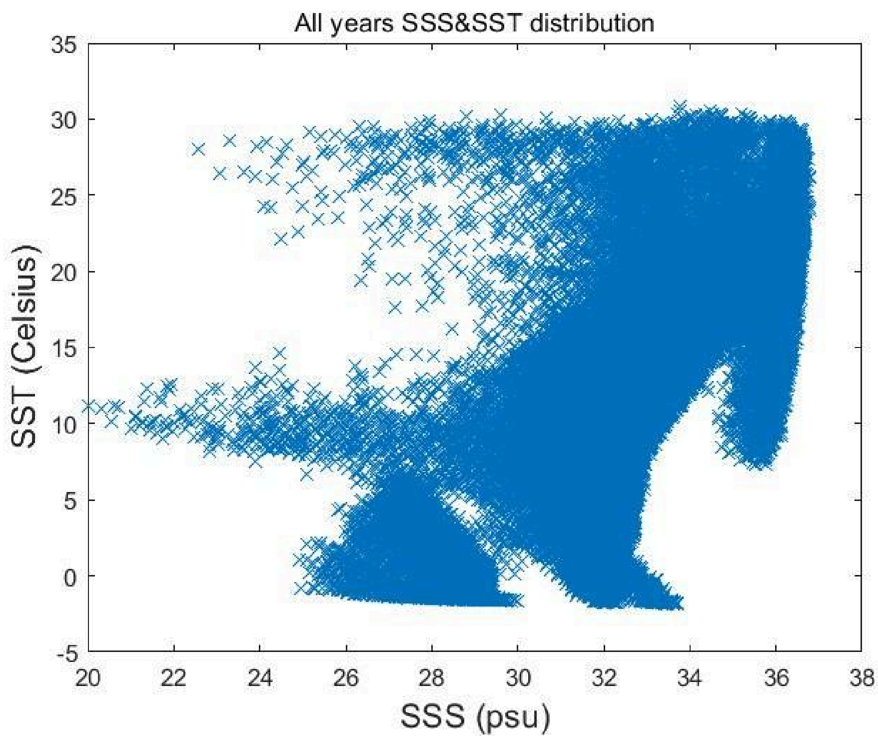

2014 distribution of sea areas
SST (Celsius)
SSS (psu)


2014 distribution of sea areas
SST (Celsius)
SSS (psu)


2014 distribution of sea areas
SST (Celsius)
SSS (psu)


2014 distribution of sea areas
SST (Celsius)
SSS (psu)

All data points within 55 years:



Matlab code of method 1:

```
filename='E213_hista.nc';
sst= ncread(filename,'sst_historical');
sss= ncread(filename,'sss_historical');
yrange = '[25:90]';
xrange = '[120:150]';

arr_sst = [-2 0 5 10 15 20 25 30 35]; %data range and interval based on pdd (sst)
arr_sss = [20 22 24 26 28 30 32 34 36 38 40]; %data range and interval based
                                                                        %on pdd
(sss)

sssx = sss([str2num(yrange)],[str2num(xrange)],:); %constrain to basin
sstx = sst([str2num(yrange)],[str2num(xrange)],:);

sizesss = size(sssx);
sizesst = size(sstx);

sss_reshaped = reshape(sssx,sizesss(1)*sizesss(2),660); %2091 by 660
sst_reshaped = reshape(sstx,sizesst(1)*sizesst(2),660); %2091 by 660
```

```
Xs = sss_reshaped(:,601:660);
Xt = sst_reshaped(:,601:660);
for k = 1:5
    xs(:,k) = reshape(Xs(:,(12*(k-1)+1):12*(k-1)+12),12*2046,1);
    xt(:,k) = reshape(Xt(:,(12*(k-1)+1):12*(k-1)+12),12*2046,1);
end

for iyear = 1:5
    figure
    plot(xs(:,iyear),xt(:,iyear),'x')
    xlabel('temperature');
    ylabel('salinity');
    title('distribution of sea area in specific SST & SSS');
end

opts = statset('Display','final');
k=4;
for iyear = 1:5
    X = [xs(:,iyear),xt(:,iyear)];
    [idx,C] = kmeans(X,k,'Distance','cityblock',...
    'Replicates',5,'Options',opts);
    %kmeans function

    figure;
    plot(X(idx==1,1),X(idx==1,2),'r.','MarkerSize',12)
    hold on
    plot(X(idx==2,1),X(idx==2,2),'b.','MarkerSize',12)
    plot(X(idx==3,1),X(idx==3,2),'y.','MarkerSize',12)
    plot(X(idx==4,1),X(idx==4,2),'k.','MarkerSize',12)
    plot(C(:,1),C(:,2),'kx','MarkerSize',15,'LineWidth',3)
    legend('Cluster 1','Cluster 2','Cluster 3','Centroids','Location','NW')
    title 'Cluster Assignments and Centroids'
    hold off
end
```

Matlab code of method 2:

```
filename = 'E213_hista.nc';
sst = ncread(filename,'sst_historical');
sss = ncread(filename,'sss_historical');
figure Name 'sst_graph'
contourf(sst(:,:,1,1)');colorbar
title 'Global SST Graph'
figure Name 'sss_graph'
contourf(sss(:,:,1,1)',[30:0.5:40]);colorbar
title 'Global SSS Graph'
%region focus on [] by []
yrange = '[25:90]';
xrange = '[120:150]';
```

```
arr_sst = [-2:1:35]; % data range and interval based on sst data
arr_sss = [20:1:40]; % data range and interval based on sss data
sssx = sss([str2num(yrange)],[str2num(xrange)],:);
sstx = sst([str2num(yrange)],[str2num(xrange)],:);
sizesss = size(sssx);
sizesst = size(sstx);
sss_reshaped = reshape(sssx,sizesss(1)*sizesss(2),660);
sst_reshaped = reshape(sstx,sizesst(1)*sizesst(2),660);
itime = 601; % change the value of itime to acquire data of year 2010-2014
                         % itime = 601,613,625,637,649
X = [sss_reshaped(:,itime),sst_reshaped(:,itime);
sss_reshaped(:, itime+1), sst_reshaped(:,itime+1);
sss_reshaped(:, itime+2), sst_reshaped(:,itime+2);
sss_reshaped(:, itime+3), sst_reshaped(:,itime+3);
sss_reshaped(:, itime+4), sst_reshaped(:,itime+4);
sss_reshaped(:, itime+5), sst_reshaped(:,itime+5);
sss_reshaped(:, itime+6), sst_reshaped(:,itime+6);
sss_reshaped(:, itime+7), sst_reshaped(:,itime+7);
sss_reshaped(:, itime+8), sst_reshaped(:,itime+8);
sss_reshaped(:, itime+9), sst_reshaped(:,itime+9);
sss_reshaped(:, itime+10), sst_reshaped(:,itime+10);
sss_reshaped(:, itime+11), sst_reshaped(:,itime+11)];
%X is now a matrix with SSS and SST of year 2010
figure Name '2010plot';
plot(X(:,1), X(:,2),'x')
ylabel('SST (Celsius)', 'FontSize',14);
xlabel('SSS (psu)', 'FontSize', 14);
title 'year 2010 SSS&SST'
for itime = 1:12
    X = [sss_reshaped(:,itime+598),sst_reshaped(:,itime+598)];
    %X is the matrix that stores SST&SSS value in itime-th month of year 2010
    % hist3(X,'Edges',{arr_sss arr_sst});
    n = hist3(X,'Edges',{arr_sss arr_sst});
    nf = size(n,1);
    nc = size(n,2);
    hist2d(:,itime) = reshape(n,nf*nc,1);%hist2d is the histogram of year 2010
end

Iterations = 10;
k=4;% 4 groups
[idx,C,sumD, D] = kmeans(hist2d',k,'Replicates',Iterations); %kmeans function

for ic=1:k
    clust(:,:,ic)=reshape(C(ic,:),nf,nc);
end

for ic=1:k
    figure
    pcolor(arr_sss,arr_sst,clust(:,:,ic)'/(sum(sum(clust(:,:,ic)))));
    c = colorbar;
```

```
    xlabel('SSS (psu)','FontSize',14);
    ylabel('SST (Celsius)','FontSize',14);
    title '2010 distribution of sea areas'
end
figure Name 'all data plot ';
x1=reshape(sst_reshaped,2046*660,1); % 1350360 different data points
x2=reshape(sss_reshaped,2046*660,1); % same
plot(x2,x1,'x')
ylabel('SST (Celsius)', 'FontSize',14);
xlabel('SSS (psu)', 'FontSize', 14);
title 'All years SSS&SST distribution'
```