# Identification of Regions Most Susceptible to $CO_2$ Emissions Through Cluster Analysis

**Ajani Blackwood**
Electrical and Computer Engineering
University of Connecticut
Storrs, CT 06269
ajani.blackwood@uconn.edu
ajaniblackwood@hotmail.com

## Abstract

Climate change is a phenomenon that negatively affects ocean dynamics and may lead to the destruction of coral reefs. Anthropogenic $CO_2$ emissions cause more susceptible ocean regions to experience greater changes in sea surface salinity (SSS) and sea surface temperature (SST). Analyzing these regions allows us to have a better understanding of how marine ecosystems may fare in the future. In this work, mid-Pliocene data showing average SSS and SST changes from fourteen state of the art Earth System Models (ESMs) is used to predict the future effects on coral reefs. Results show that coral reef outplants near China, Taiwan, Cambodia, and the Philippines may experience future SST levels up to 2°C more than the climatological temperature. Regions near India, Sri Lanka, and Indonesia may experience the greatest magnitude of ocean freshening by -1.5 PPT. These results are concerning for coral reef outplants in those regions as certain species have not adapted to those changes.

### Introduction and Motivation

The motivation behind this project is to gain a better understanding of how certain oceanic ecosystems respond to increases in CO2 emissions. Changes in the ocean's environment can be quantified in terms of sea surface salinity (SSS) and sea surface temperature (SST) responses. Certain coral species have different tolerances to temperature changes, but in some cases an increase of 1°C may lead to coral death (Foo and Asner, 2020). In order to assess regional impacts from salinity and temperature changes, cluster analysis will be used to identify regionally coherent patterns of temperature and salinity changes across the globe. Based on these patterns, we will be able to predict the future health of a marine ecosystem by focusing on the effects on coral species echinopora, merulina, porites, pocillopora, and acropora. The work was constrained in this way as other marine animals rely on coral reefs for their survival, so this gives us an idea about the health of an ecosystem.

## Methodology

The datasets in use are from PlioMIP2. They are comprised of Earth System Model outputs showing the changes in sea surface temperature and sea surface salinity between the mid-Pliocene and pre-industrial era. Both variables are averaged across all fourteen models for this analysis. The data is merged such that we have the modelled average changes in temperature and salinity per latitude and longitude. Cluster analysis is conducted with the standardized variables to negate the effects of scale. The optimal number of clusters for Partitioning Around Medoids (PAM) and the Gaussian Mixture Model (GMM) method are then determined using the proper metrics. Model outputs are then compared to see which clustering solution is the most sensible given regional characteristics. The chosen solution will help identify regions showing considerable temperature and salinity changes. We may also be able to determine if a change in one parameter is fairly muted and the other is more pronounced within a given region. This can be used to help predict whether or not an ecosystem's wellbeing will be of great concern in the future given the current trends in $CO_2$ emissions.

## Results

### Optimal Cluster Choice Selection

The initial choice for regional separation for PAM was determined with the aid of gap statistics and within sum of squares error (WSSE). Gap statistics showed a possible 21 cluster solution, but this resulted in undesirable regional separation. 11 clusters was the final choice as it was supported by WSSE and led to better differentiation of regions. Similarly, for Gaussian Mixture Modelling, the Bayesian Information Criterion (BIC) was used. BIC showed a maximum value at 23 clusters, but the choice was ambiguous. 7 was a possible choice as the rate of change decreased at that point, but the transition from 11 to 12 showed a slope of 0. As a response to this uncertainty, both 7 and 11 clusters were tested for the latter portion of the analysis. Unfortunately, 7 clusters looked similar to the case with 23 clusters according to calculations for the per cluster median SSS and SST. Due to this finding as well as the 0 slope from 11 to 12 clusters, 11 clusters was deemed sensible for a comparison between PAM and GMM. Metrics are displayed in Figure A.

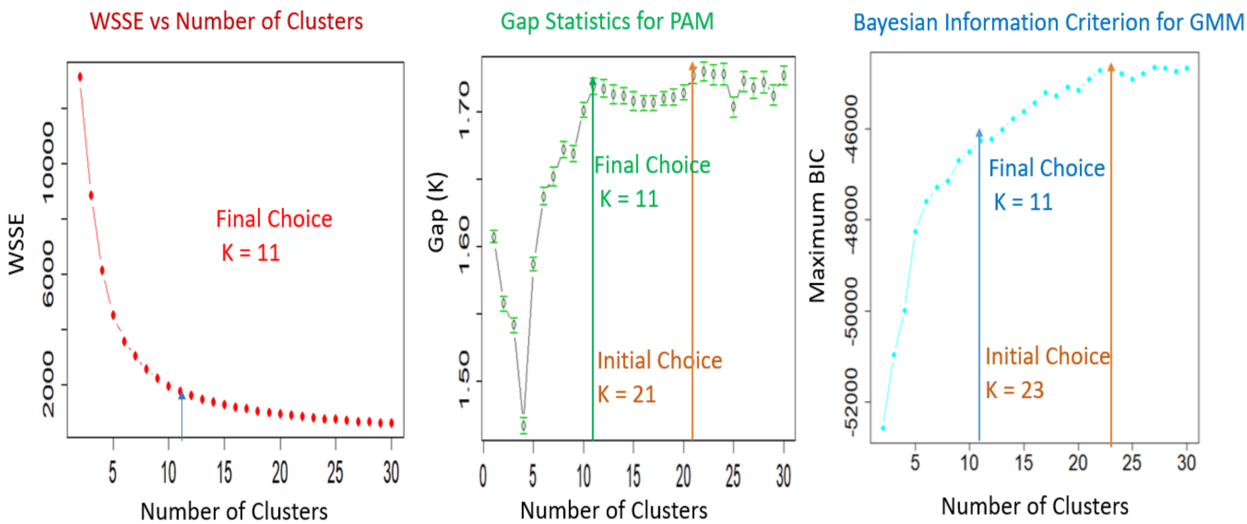


*Figure A: Clustering metrics for PAM and GMM showing Initial and final clustering choices*

**Mapped Separation & Interpretations for Final Model Choice**

The clustered regions for PAM and GMM showed great similarity in calculations for SST and SSS. This is desirable as it shows consistency in regional separation between methods. PAM however showed slightly better groupings for clusters. For this reason, the coral reef analysis will be based on the PAM results for a consistent interpretation and easier reference to the geographic locations of labelled clusters. See appendix C for the GMM cluster separation.

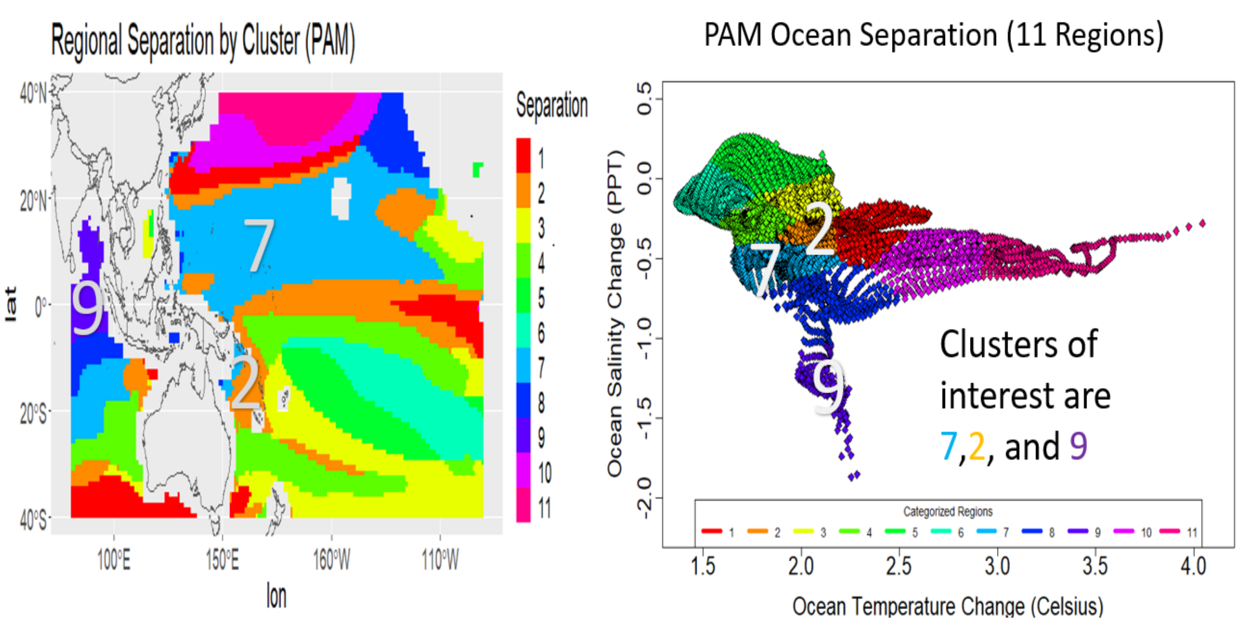


*Figure B: Clustered geographical regions and corresponding temperature and salinity ranges*

Figure B shows two complementary visualizations with identical colors. Regions 7,2 and 9 are blue, orange, and purple respectively in both images. Region 9 near Indonesia, Sri Lanka and India appears to have the strongest reductions in salinity levels. It also has temperature increases ranging from approximately 2 to 2.3°C. The general area of region 7 was chosen considering the number of coral reef outplants near China, Taiwan, Cambodia, and the Philippines. This can be verified in Figure D showing locations of coral reef outplant survivors.

For all regions, Figure 1C shows the climatology as well as future predictions for the temperature and salinity per cluster. For climatology, the mean was used, but for changes the median was used because of extreme values. Some observations for temperature change were 6 standard deviations above the mean and for salinity, 6 standard deviations below the mean as seen in Appendix A, Figure A-2. Results from the cluster analysis show that in general, regions will see significant increases in SST with a median increase of about 2°C. Median Changes in SSS levels show a median decrease of -0.341 PPT. These values can be seen in Appendix B, Figure B-2. Given our data and species in the analysis, we can infer the effects based on temperature. A red line has been drawn at 27°C in Figure C indicating the danger zone for the five genera: echinopora, acropora, porites, merulina, and pocillopora. Figure D shows the proportion of survivors at this temperature up to 32°C. Judging by both figures, the most troubling zones with regards to these five genera are 9 and 7. Regions 9 and 7 are the general waters around Indonesia, Malaysia, China, and the Philippines.

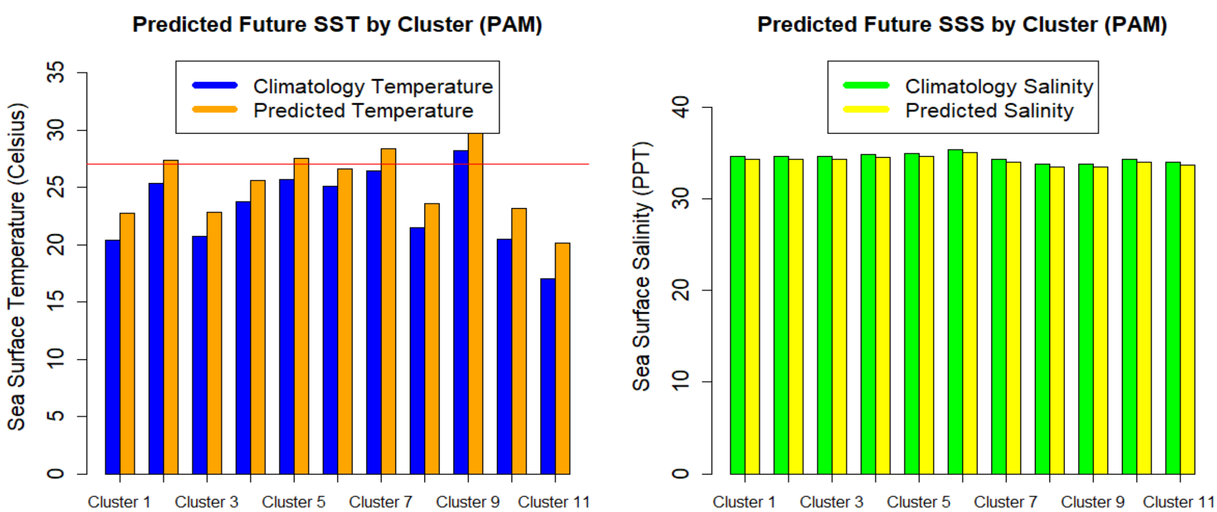


*Figure C: Projected future SST and SSS levels in clustered regions*

The map in Figure B can be compared to the one in Figure D showing the proportion of coral reef outplant survivors. As shown, regions around Vietnam, Australia, Taiwan, Indonesia, and China are locations of coral reef outplants. Red dots correspond to almost no survival, yellow is more favorable but still concerning at 33-66% survival, and green is best with a lower bound of 66% survival (Foo and Asner, 2020).

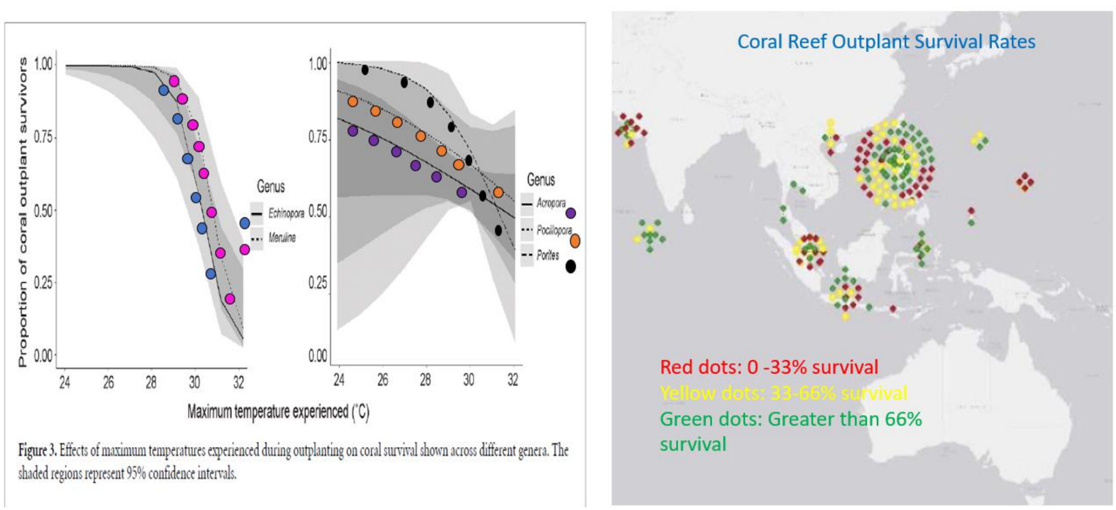


*Figure D: Coral reef outplant survivors. Adapted from* (Foo and Asner, 2020)

## Discussion

### General Effects on Outplant Survivors

Corals tend to have negative growth effects with salinity changes of +- 2 PSU (Kuanui et al., 2015). Concerning ranges for base salinity levels tend to be less than 25 PPT or greater than 45 PPT for the species in question (Kuanui et al., 2015). Corals respond unfavorably to sudden changes in salinity from high freshwater input, but in general temperature changes are more of a concern (Kuanui et al., 2015). Furthermore, the salinity range above has not been exceeded based on the ocean region clustering results. Beyond this level is when echinopora, pocillopora, porites, acropora, and merulina coral outplants suffer the most (Foo and Asner, 2020).

### Modelling Uncertainty & Possible Improvements

Regional separation with hierarchical clustering was not performed due to computing restraints. Comparison to the hierarchical model would have allowed us to see if coherent patterns in temperature and salinity changes were consistently reproduced with three models.

The algorithm used for calculating gap statistics has documentation recommending 500 Monte Carlo bootstrap samples to make a determination since this yields consistent plotting results. Unfortunately, only 20 bootstrap samples were used because of computing restraints. Considering that the goal is to see the effects on coral reef species, ocean PH balance would be a good variable to include in the future. The groupings of clustered regions would have given more insight as to ecosystem outcomes. Furthermore, this analysis concerns the effects on only five coral genera, so these interpretations may not be valid for other species.

## Conclusion

Patterns of temperature and salinity changes were verified across regions within the Pacific Ocean. Temperature increases across regions may result in dangerous levels for the five genera as some future predictions exceed the 27°C threshold. The salinity projections shown may have broader implications but do not have a strong effect on the species focused on in this analysis. Clustering confirmed that danger zones are the waters near Indonesia, Malaysia, the Philippines, China, and Vietnam for the five genera. Model improvements discussed may give better insights involving more species. Hopefully future research with similar data from the mid-Pliocene Warm Period can help us make stronger predictions and assessments.

## References

Bœuf, G. and Payan, P. (2001). How should salinity influence fish growth? *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, [online] 130(4), pp.411–423. Available at: https://pubmed.ncbi.nlm.nih.gov/11738629/#:~:text=In%20larger%20fish%2C%20salinity%20is [Accessed 19 Oct. 2020].

Chan, W.-L. and Abe-Ouchi, A. (2020). Pliocene Model Intercomparison Project (PlioMIP2) simulations using the Model for Interdisciplinary Research on Climate (MIROC4m). *Climate of the Past*, 16(4), pp.1523–1545.

Foo, S.A. and Asner, G.P. (2020). Sea surface temperature in coral reef restoration outcomes. *Environmental Research Letters*, 15(7), p.074045.

Haywood, A.M., Dowsett, H.J. and Dolan, A.M. (2016). Integrating geological archives and climate models for the mid-Pliocene warm period. *Nature Communications*, [online] 7(1), p.10646. Available at: https://www.nature.com/articles/ncomms10646 [Accessed 30 Apr. 2021].

Jenkins, C.N. and Van Houtan, K.S. (2016). Global and regional priorities for marine biodiversity protection. *Biological Conservation*, 204, pp.333–339.

Kuanui, P., Chavanich, S., Viyakarn, V., Omori, M. and Lin, C. (2015). Effects of temperature and salinity on survival rate of cultured corals and photosynthetic efficiency of zooxanthellae in coral tissues. *Ocean Science Journal*, 50(2), pp.263–268.

McNeil, B.I., Matear, R.J. and Barnes, D.J. (2004). Coral reef calcification and climate change: The effect of ocean warming. *Geophysical Research Letters*, [online] 31(22). Available at: https://www.esrl.noaa.gov/gmd/icdc7/proceedings/abstracts/mcneil1HI75.pdf.

Ortiz, R.M. (2001). Osmoregulation in Marine Mammals. Journal of Experimental Biology, [online] 204(11), pp.1831–1844. Available at: https://journals.biologists.com/jeb/article/204/11/1831/32941/Osmoregulation-in-Marine-Mammals.

Velasco, J., Gutiérrez-Cánovas, C., Botella-Cruz, M., Sánchez-Fernández, D., Arribas, P., Carbonell, J.A., Millán, A. and Pallarés, S. (2018). Effects of salinity changes on aquatic organisms in a multiple stressor context. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1764), p.20180011.

www.rdocumentation.org. (n.d.). *clusGap function - RDocumentation*. [online] Available at: https://www.rdocumentation.org/packages/cluster/versions/2.1.2/topics/clusGap [Accessed 30 Apr. 2021].

## Appendix A Comparison of Non-standardized and Standardized Parameters

Figure A-1 below shows the non standardized parameters in histogram form. Figure A-2 shows the standardized parameters used to perform the cluster analysis. The existence of extreme values for both SSS and SST led to the use of the median as opposed to the mean for further portions of the analysis.

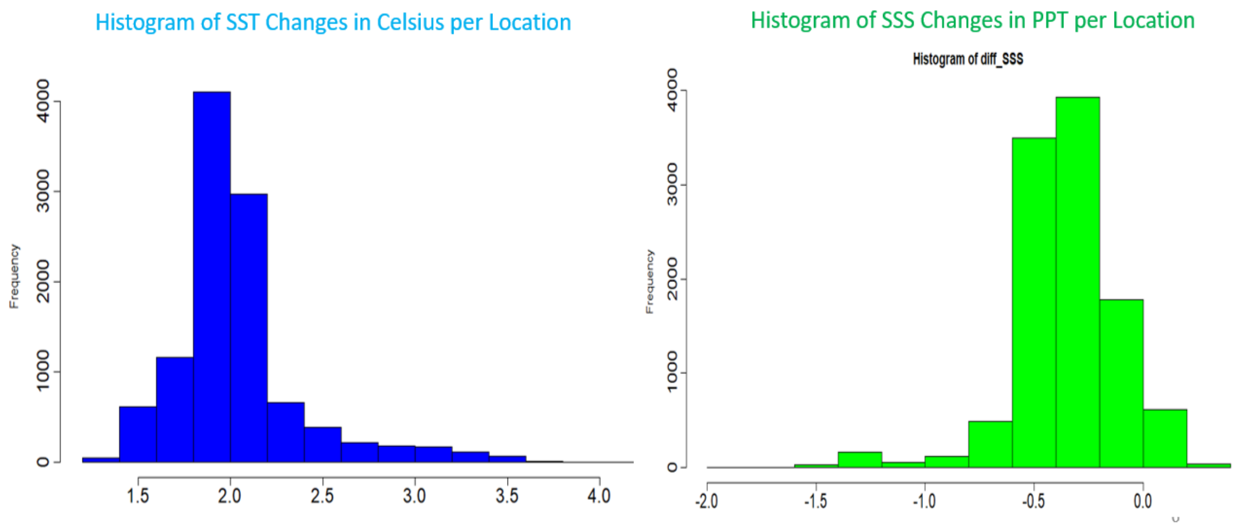


*Figure A-1: Histograms showing distributions of SSS and SST changes (not used in cluster analysis)*

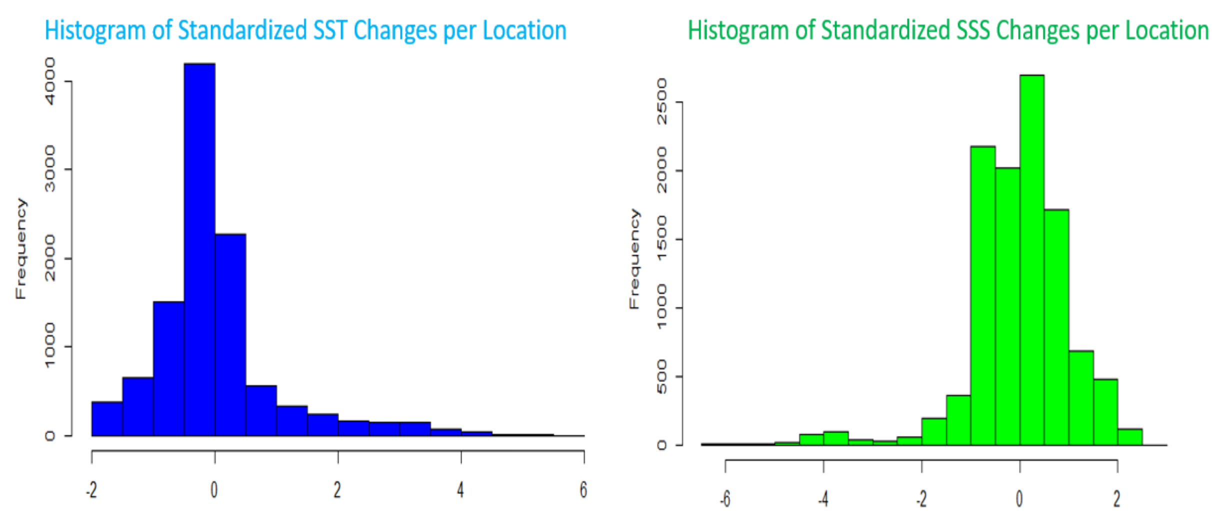


*Figure A-2: Standardized SST and SSS used for clustering*

## Appendix B Calculated per Cluster Median Parameter Changes

Figures B-1 and B-2 below are supplementary for PAM. B-1 shows the median sea surface salinity change for each cluster in the left panel, and median sea surface temperature in the right panel. Figure B-2 shows the outcome of calculations for the absolute unsigned difference from each cluster to the overall median. The unsigned difference was used just to infer how much change was present relative to other clusters regardless of direction.

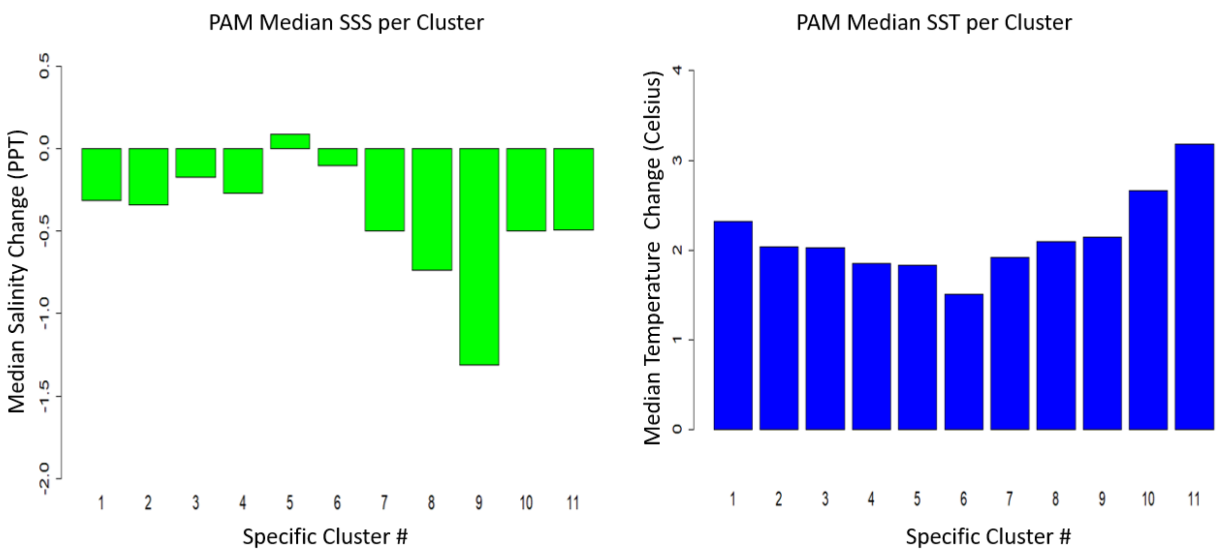


*Figure B-1: PAM Cluster centroid median change for SST and SSS*

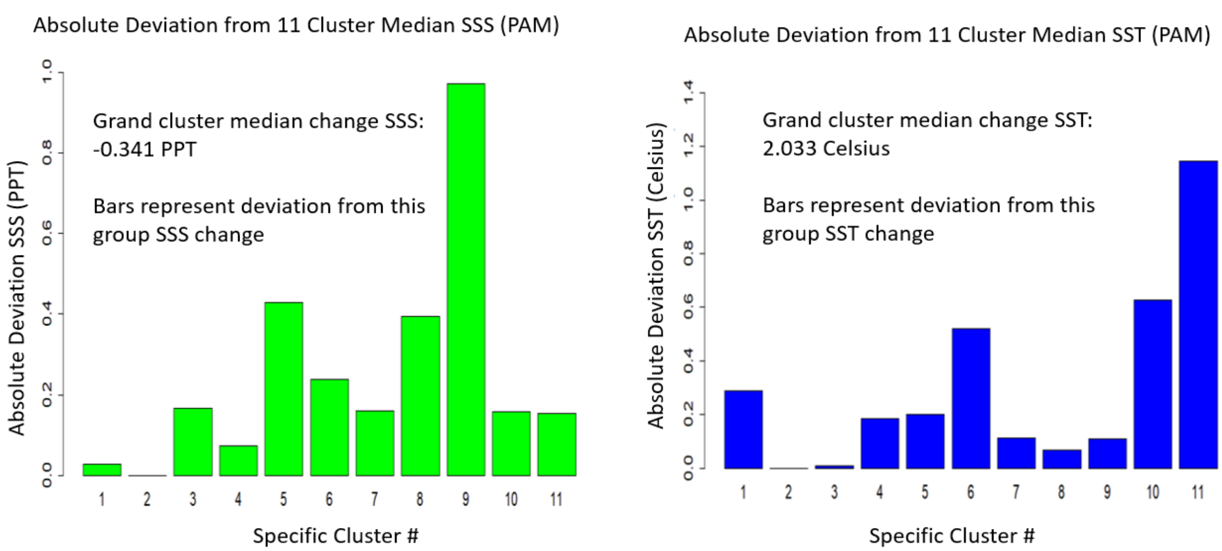


*Figure B-2: PAM Deviation from clustered grand median for SST and SSS*

## *Appendix C GMM Comparisons*

Appendix C shows the same structure of information as Appendix B, except that the information is specifically for GMM. Figure C-1 serves to illustrate that the cluster analysis for GMM showed similar results to PAM, but the labels for many of the same geographical regions are shuffled. Regions 1,5, and 10 in GMM highlight the same areas as regions 2,7, and 9 in PAM. This is not in the main report as the different labels per cluster would lead to confusion.

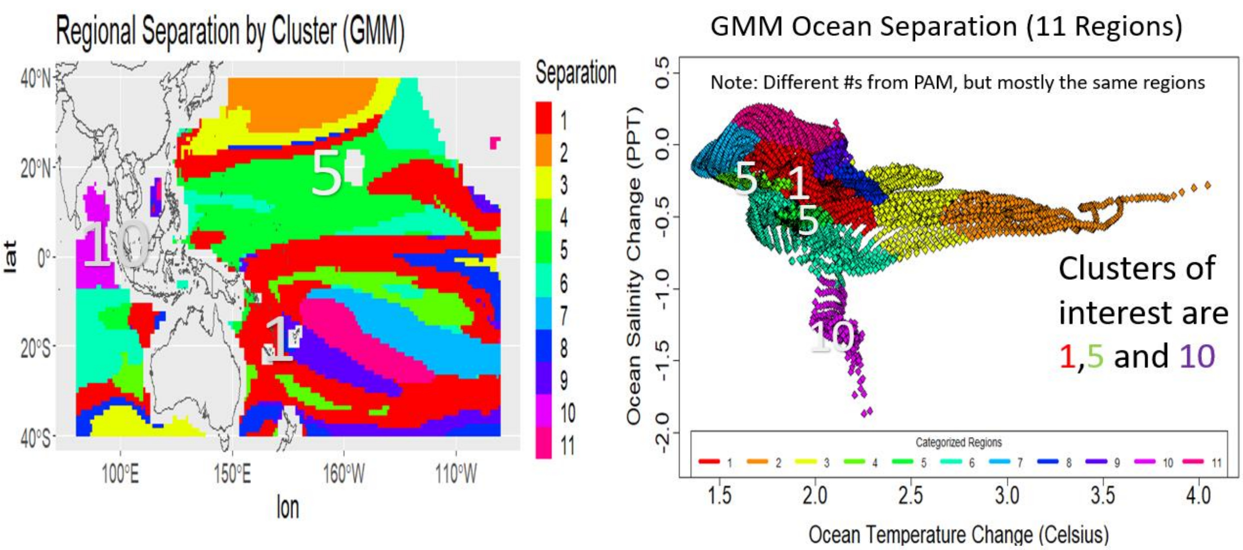


*Figure C-1 GMM mapped cluster separation showing similar regional differentiation to PAM*

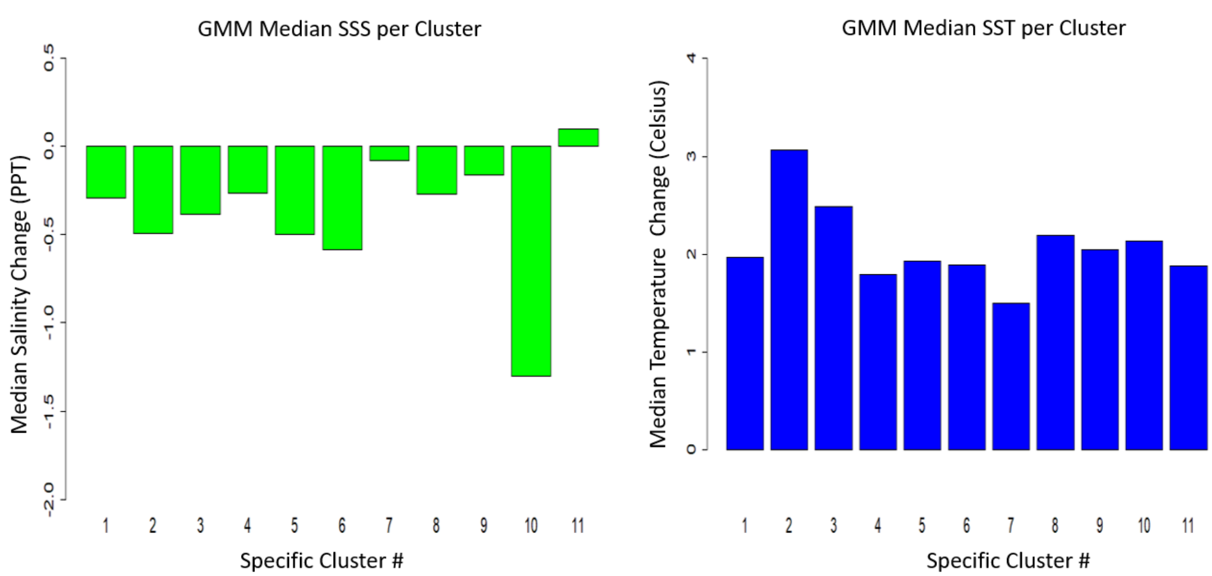


*Figure C-2: GMM Cluster centroid median change for SST and SSS*

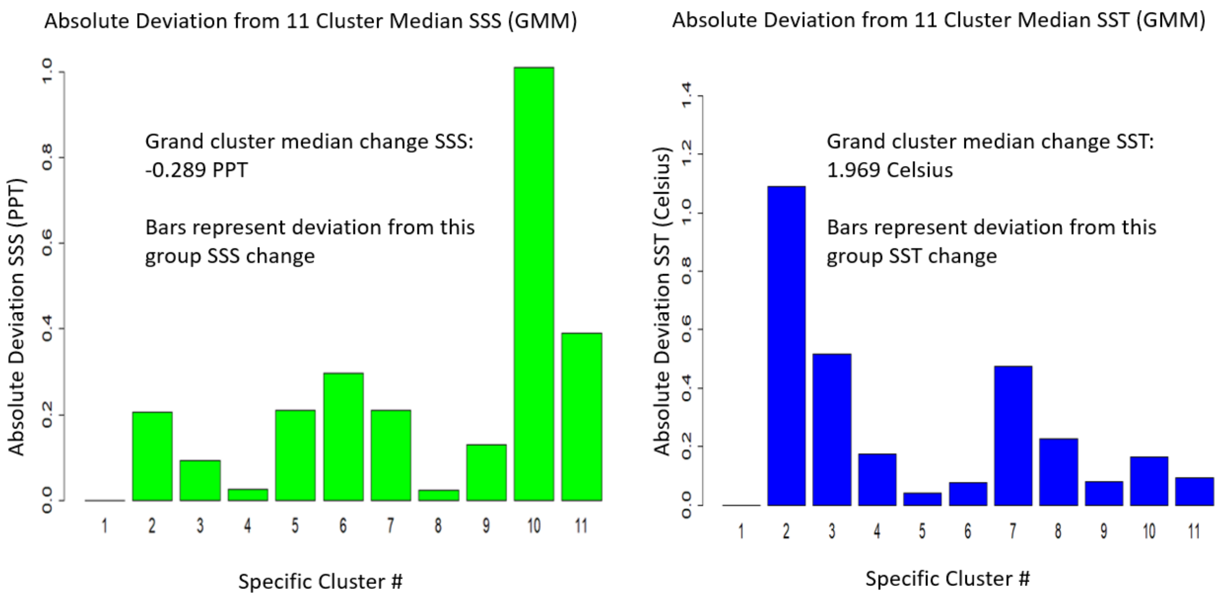


*Figure C-3: GMM Deviation from clustered grand median for SST and SSS*

We can see that the numbers for GMM were slightly different than PAM, especially in regards to sea surface salinity. The median change for temperature was closer to PAM in comparison.