Analysis of Coral Reef Bleaching

Introduction

Corals are invertebrate animals belonging to the phylum Cnidaria, which consists of over 11,000 species of aquatic animals. While each species of coral has unique physical characteristics, all coral possess a simple stomach with a single mouth opening surrounded by stinging tentacles. Coral colonies are made up of hundreds of thousands of individual coral animals known as polyps. A colony is formed through budding, which occurs when a polyp grows copies of itself.

Coral and algae are in a symbiotic relationship, meaning that both organisms are in a mutually beneficial relationship, and they are both gaining from being in each other's vicinity. Corals emit their waste product in the form of ammonium, and the algae take it up as a nutrient. In return, the algae undergo photosynthesis, in which they emit and pass their nutrients back to the coral. This positive relationship between the algae can only continue when the coral and algae are under non-stressful conditions. Recently, global warming and other conditions brought on by climate change have created a stressful environment for the coral. One example is the SST or sea surface temperature in the sea, which has been on the rise and has caused the coral to expel the algae living on them. When algae experience heat stress, they produce harmful chemicals and are subsequently expelled from coral tissue, causing the coral to turn white. This phenomenon is known as coral bleaching which can result in the death of coral reefs.

Data

The Global Bleaching and Environmental Data (version two) is a dataset that includes information on indicators of coral bleaching, such as percent bleaching and sea surface temperature anomalies (SSTA), at various sites in the Pacific, Atlantic, and Indian Oceans from 1980-2020. The data were collected from seven coral bleaching studies and databases by researchers at the Florida Institute of Technology and the University of California-Santa Barbara in 2022. Researchers converted all site coordinates to decimal degrees and compared them to ensure there were no duplicates. Coordinates located on land or more than one kilometer away from a coral reef were discarded. We downloaded the cleaned dataset from the Biological and Chemical Oceanography Data Management Office, which had a total of 62 variables. For our exploratory analysis we analyzed the following variables: percent bleaching, temperature, distance to shore, turbidity, and cyclone frequency. For our model, bleaching percent was our dependent variable and the other 61 parameters were our independent variables. While reading from the csv file, we replaced all "nd" with NA so that the numeric columns like Temperature_Kelvin, etc. automatically got converted to dbl. For our processing from case to case basis, we also omitted all rows with NA.

Exploratory Analysis

Visual Introduction to Data

Corals are found across the world's oceans in both shallow and deep water, however, reef-forming corals are only found in shallow tropical and subtropical waters ("Corals and Coral Reefs."). Figure 1 shows a map depicting the locations of shallow water coral reefs available in our dataset.

Figure 1. A plot of the data collection sites.

Notice that almost all coral reef locations fall within 30 degrees of the equator. This is because "algae found in their tissues need light for photosynthesis and they prefer water temperatures between 70-85°F (22-29°C)" ("Corals and Coral Reefs.").

Figure 2 shows a heatmap of the coral reef locations included in the dataset. The heatmap is weighted by percent bleaching of the corals, which highlights regions of concentrated coral bleaching. Note that the Caribbean Sea, South China Sea, and Tasmanian Sea experience the highest concentrations of severe coral bleaching.

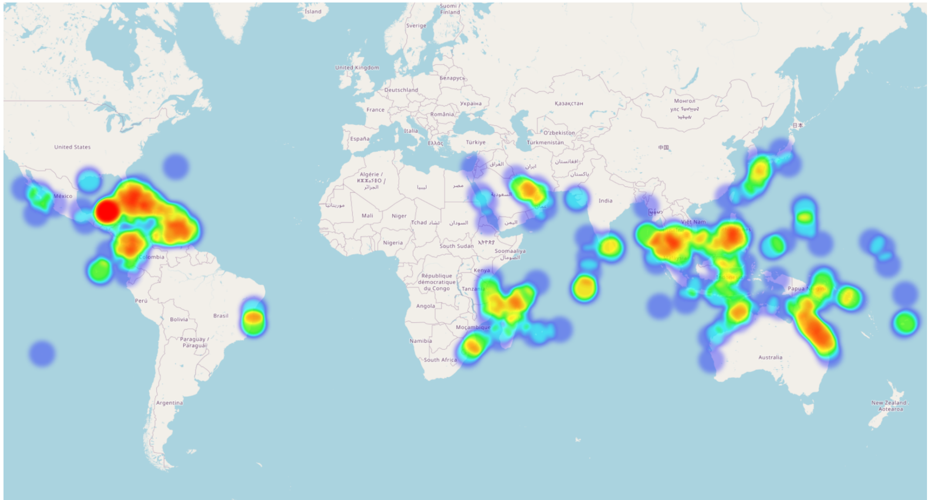

Figure 2. A heat map of the percent bleaching at data collection locations.

Upon initial inspection of our dataset, we noticed that one major limitation would likely be a lack of data in the early time span of our dataset. Figure 3 shows the distribution of the 4 primary sources that make up the data set.

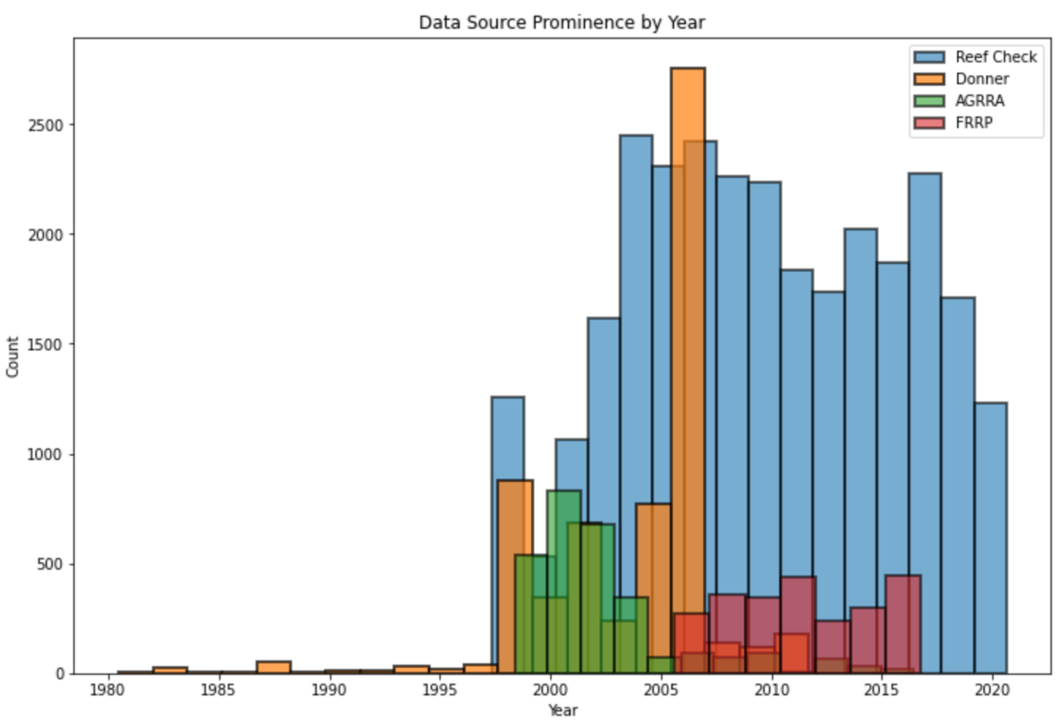

Figure 3. Histogram of available data over time.

While our dataset shows data spanning from 1980 to 2020, it is evident that the vast majority of that data is concentrated from around 1997 onwards. Due to the lack of available data from the 1980s and 1990s, making correlations over time proved to be difficult.

Hypotheses

We established a set of informed hypotheses after becoming familiar with the topic of coral bleaching and reviewing the pre-existing literature in the field. An initial exploration of the dataset led us to make the following hypotheses:

1. **Bleaching percent will increase as water temperature increases.**

   High temperatures cause the algae in coral tissue to release damaging chemicals, which means the coral must expel the algae from their tissue and turn white (National Oceanic and Atmospheric Administration, 2021).

2. **Bleaching percent will decrease as turbidity increases.**

   Turbidity is the measure of relative clarity of a liquid. Materials such as clay, silt, very tiny inorganic and organic matter, algae, dissolved colored organic compounds, and plankton cause water to become turbid. These suspended particles can shield the coral from harmful radiation that may trigger algae to release harmful chemicals and result in algae expulsion and bleaching (Carlson et al., 2022).

3. **Bleaching percent will decrease as cyclone frequency increases.**

   Cyclones absorb heat from the ocean, reducing the likelihood of heat-related coral bleaching (Davis, 2022).

4. **Corals closer to the shore will experience more bleaching than those further away.**

Agricultural runoff increases the amount of nitrogen in the water and subsequently lowers the amount of phosphorus, which makes coral more susceptible to bleaching (Lapointe et al., 2019).

Analysis of Hypotheses

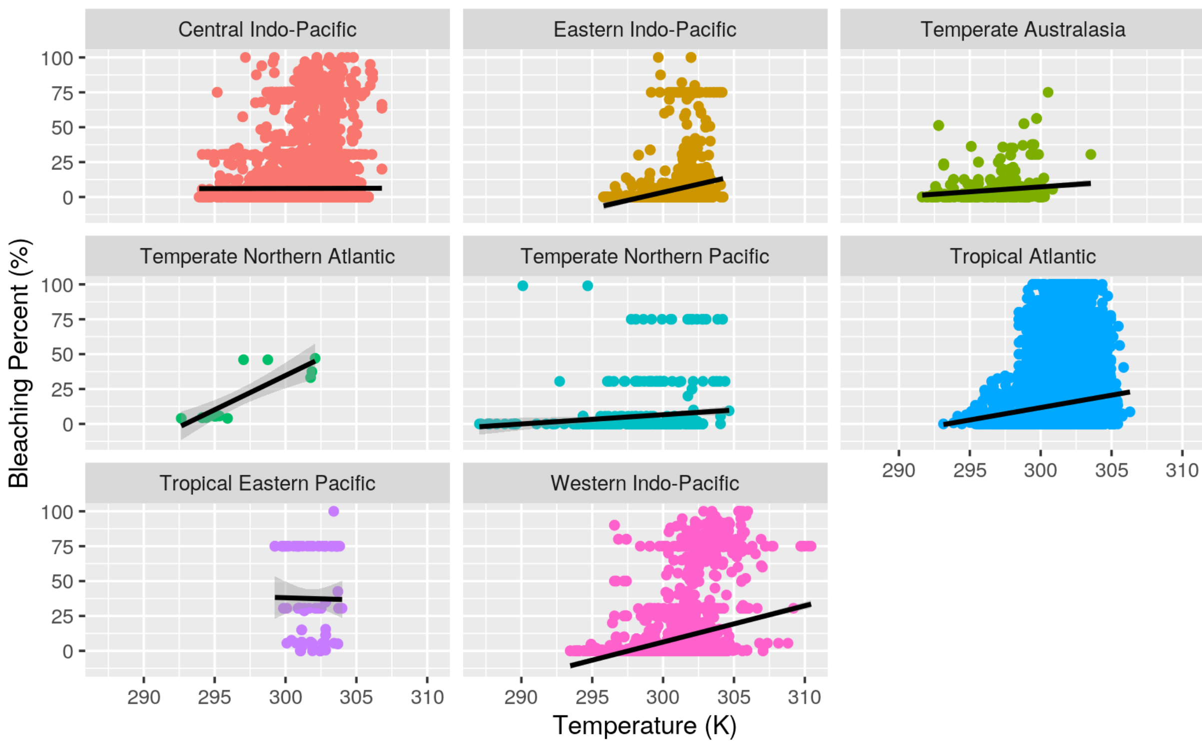

Figure 4. Plots with linear regression of Bleaching Percent vs. Temperature.

When separating for ocean realms, bleaching percent is typically positively correlated with increasing temperature. However, there were some regions, (Central Indo-Pacific & Tropical Eastern Pacific) that had a neutral or negative correlation, demonstrating that while temperature is a factor, it is not the end all be all for the graphs, and variations among the ecoregions can influence bleaching as well. But overall, these graphs seem to agree with our hypothesis concerning temperature.

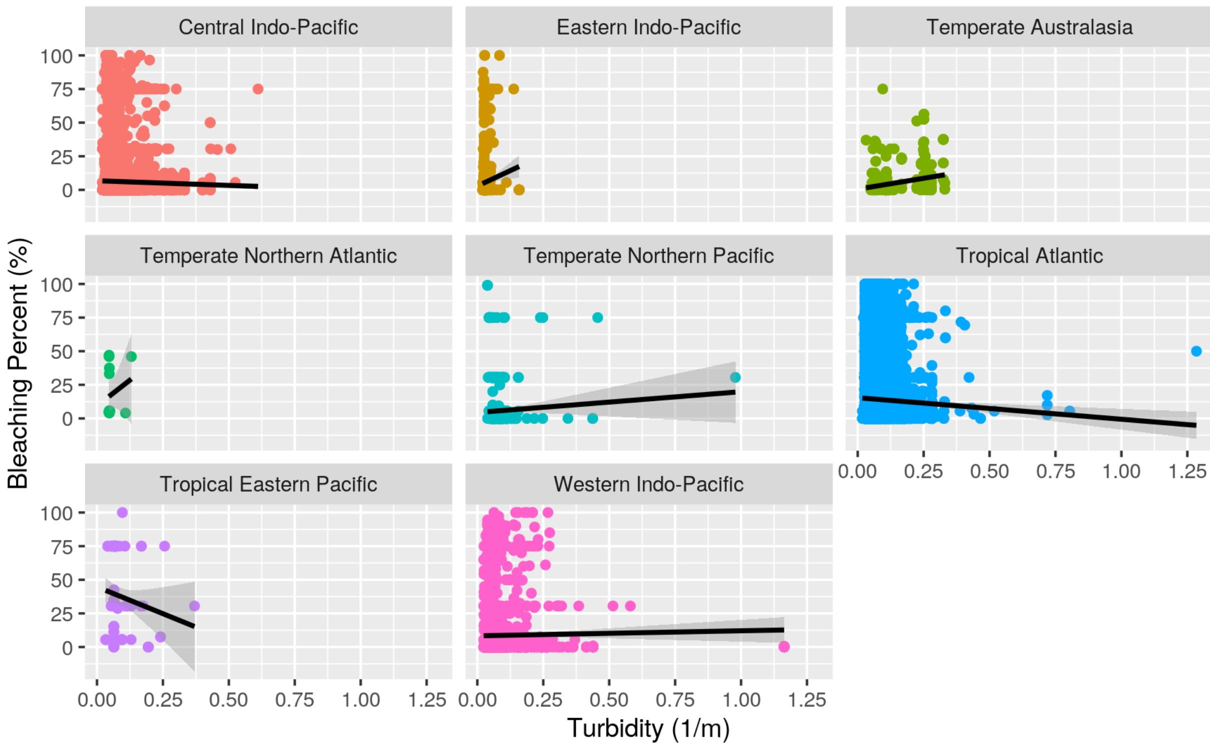

Figure 5. Plots with linear regression of Bleaching Percent vs. Turbidity.

When separating for ocean realms, bleaching percent varied across ocean realms, with correlations of positive, neutral, and negative when measured against turbidity. Turbidity is the relative clarity of the water, and can be affected by sediment levels in the water, as well as general water movement. Overall, the only ocean realms that directly support our hypothesis are the Tropical Eastern Pacific and Tropical Atlantic, while all other graphs are either neutral or positively correlated.

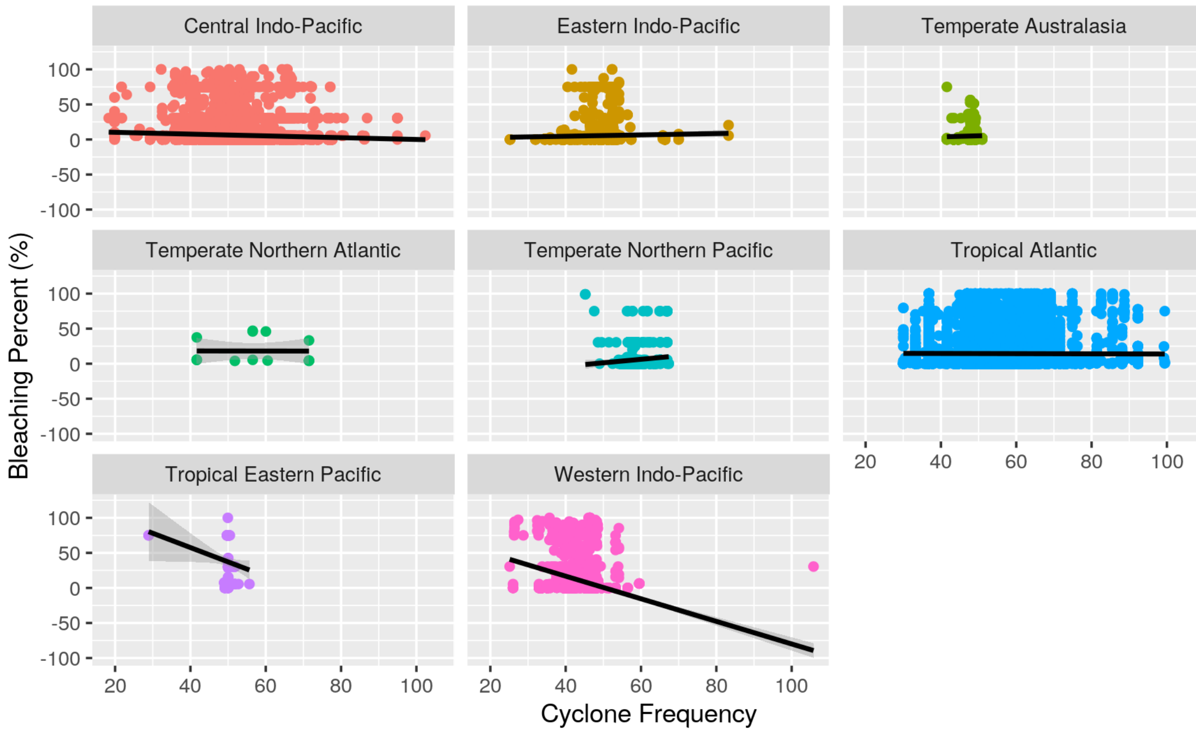

Figure 6. Plots with linear regression of Bleaching Percent vs. Cyclone Frequency.

When separating for ocean realms, bleaching percent typically shows a nearly neutral or negative correlation with Cyclone frequency of the water. Cyclone frequency can also increase turbidity, so the fact that these results are similar to the previous graph is consistent with what we read in our sources. Some regions, such as the Tropical Eastern Pacific and Western Indo-Pacific, directly support our hypothesis, while the other graphs have a neutral correlation.

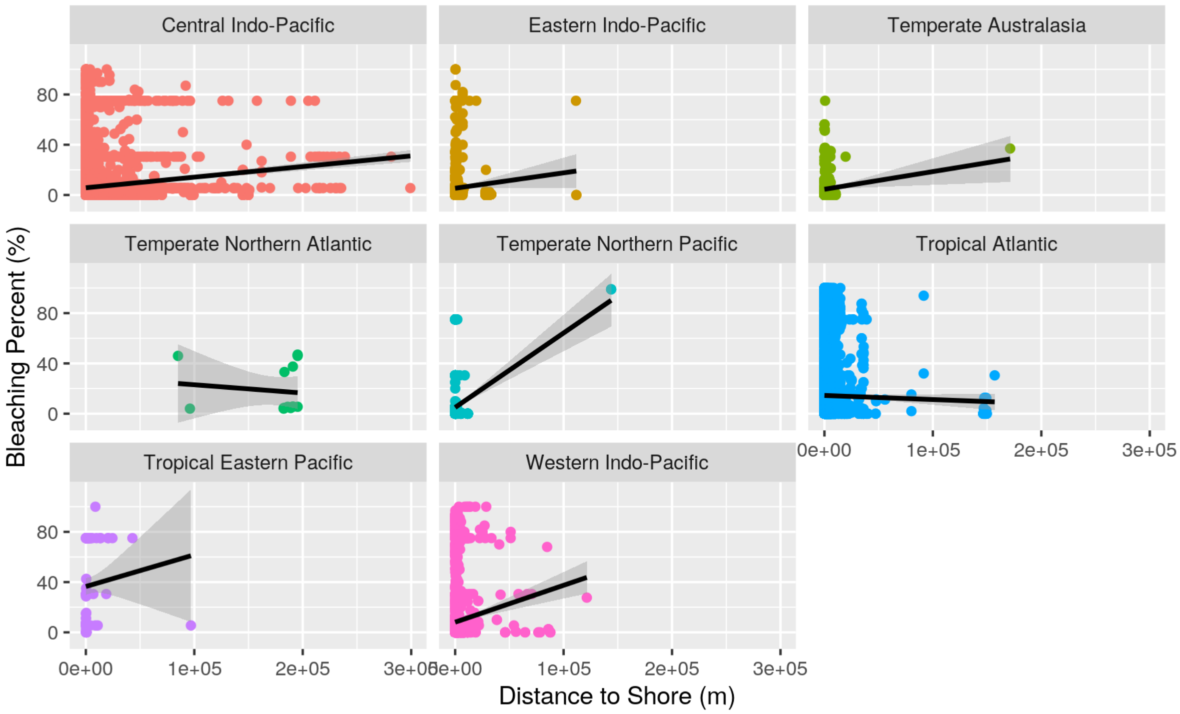

Figure 7. Plots with linear regression of Bleaching Percent vs. Distance to Shore.

When separating for ocean realms, bleaching percent typically shows a positive correlation with Distance to Shore for most ocean realms. A majority of the graphs, besides Temperate Northern Atlantic and the Tropical Atlantic, have a positive correlation. We know from our sources that farther away from the shore, water tends to be warmer in the summer months. This gives us a plausible explanation for our findings, and stands in opposition to our original hypothesis.

Table 1. R^2 values for linear regression of Bleaching Percent vs. Hypotheses Variables.

Variable	R^2
Temperature	0.0116
Turbidity	0.00198
Cyclone Frequency	0.000295
Distance to Shore	0.00163

According to the R-squared values in Table 1, the lines of best fit are very weak. We did not expect to have such low values. Splitting up the data by time intervals could increase the R-squared values, yet for our project we focused on the whole dataset from 1980-2020.

Correlation of Variables

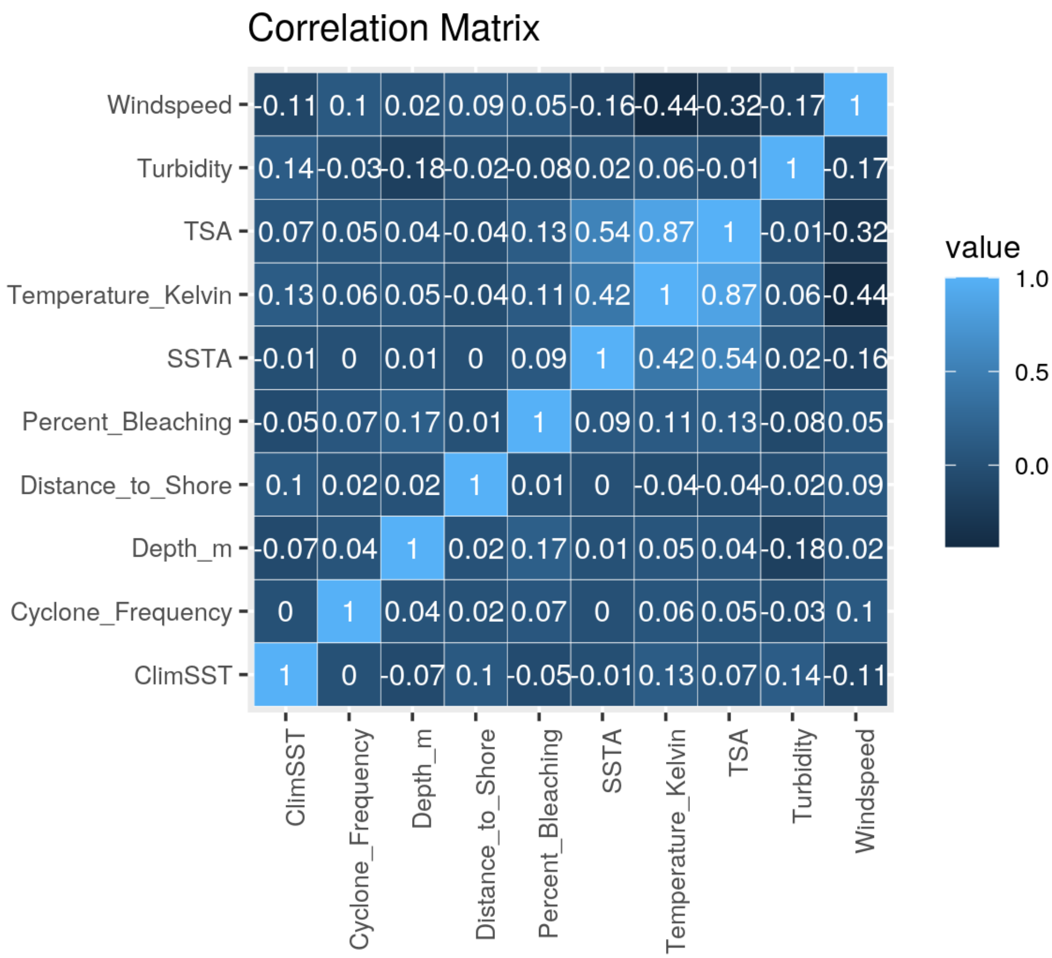

Figure 8. A correlation matrix of all the numerical variables in the dataset.

From the correlation matrix we found that the percent bleaching is positively correlated with depth (m), total survey area (TSA), temperature (Kelvin), sea surface temperature anomalies (SSTA), cyclone frequency, wind-speed, and distance to shore (in order from strongest to weakest correlation). Percent bleaching is negatively correlated with turbidity and climatological sea surface temperature (ClimSST). After analyzing the results of the correlation matrix, we decided to investigate how the bleaching percentage changes over time.

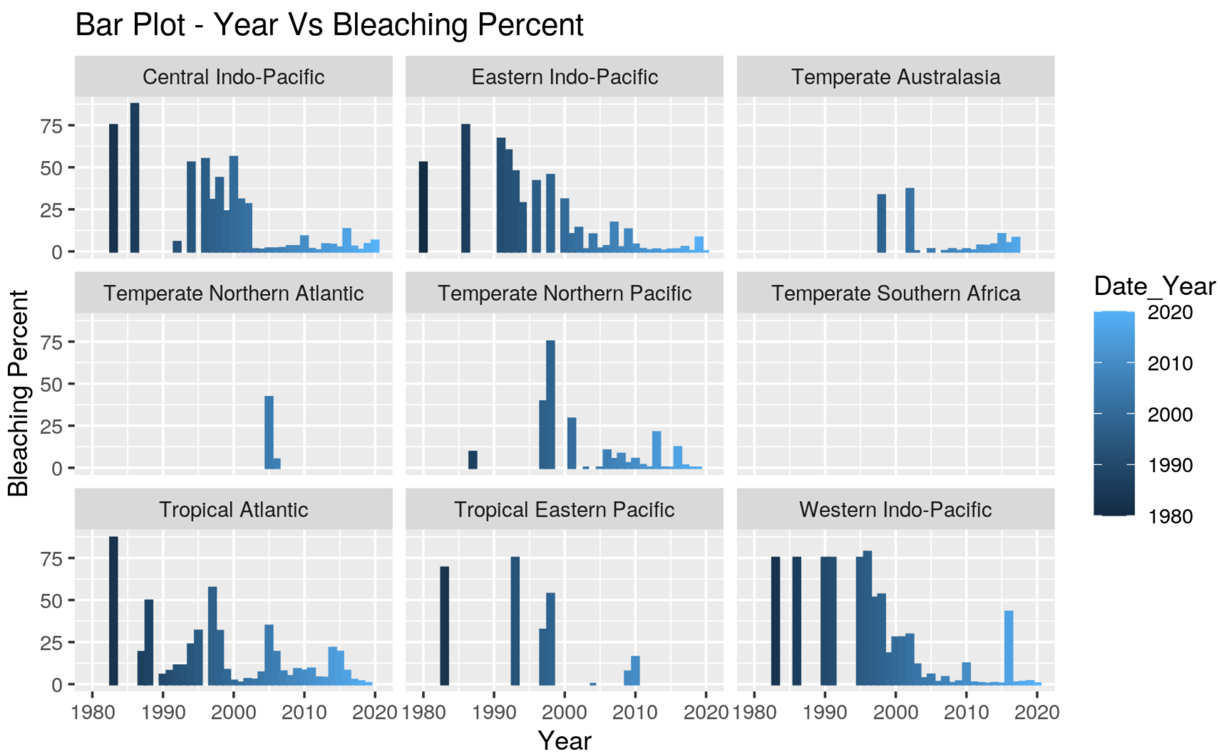

Figure 9. Year vs Bleaching Percentage.

According to Figure 9, percent bleaching decreased over time. This may indicate that the coral is becoming healthier over time; however, this trend may be due to an error in the data collection process. Researchers continually collected data from corals that had reached 100% bleaching and died.

Effect of ENSO on Coral Bleaching

To further investigate some of the unexpected findings seen when exploring bleaching percentages over time, we decided to study how the El Niño-Southern Oscillation (ENSO) cycle affects coral bleaching. ENSO refers to a major recurring climate pattern involving changes to water temperatures in the tropical regions of the Pacific Ocean. ENSO cycles occur about every three to seven years, and can have major implications to weather and climate across the globe.

The cycle is characterized by two opposing phases (El Niño/La Niña), as well as neutral phases. El Niño refers to years in which a warming of the ocean occurs, resulting in above-average sea surface temperatures (SST), while La Niña refers to periods of cooling or below-average sea surface temperatures. When SST is close to the historical average, the phase is classified as a neutral year ("What Is ENSO?"). Coral bleaching is traditionally associated with El Niño years, however, this could be changing (Ritchie et al.).

For this study, we wanted to investigate coral bleaching patterns associated with different phases of the ENSO cycle and how they have changed over time. For this reason, the dataset was split into a past subset of years (1996-2004) and a more recent subset of years (2007-2020). The range of years from 1996 to 2020 was selected because it represented the widest range for which sufficient data was available. Next, for each subset of years, each observation was classified by its ENSO phase based on the year it occurred. Those classified as strong El Niño, strong La Niña, and neutral phases were added to their respective dataset in the past years subset or modern years subset.

The resulting observations were mapped with marker sizes corresponding to bleaching severity and marker colors corresponding to ENSO phase. The results in Figure 10 represent bleaching trends across ENSO phases in the past subset of years. It can be seen that the vast majority of bleaching has historically occurred during El Niño years, while past Neutral and La Niña phases have seen significantly less.

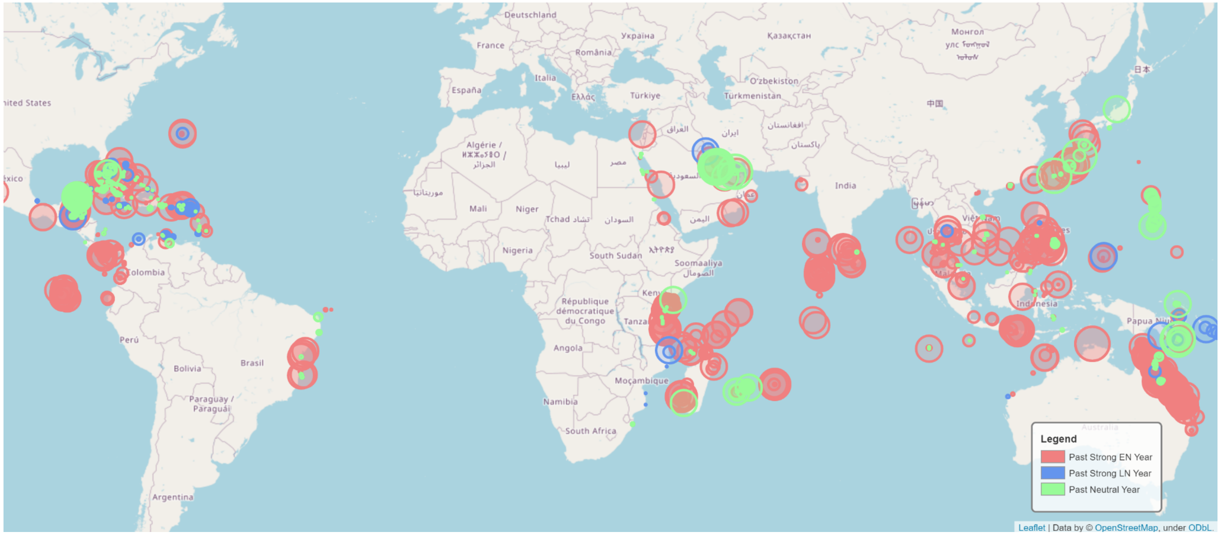

Figure 10. Map of bleaching events in past years by ENSO cycle.

This trend makes sense, as coral bleaching is generally associated with warmer ocean temperatures. This would mean that corals would have time to recover from bleaching events during the season of cooler water temperatures brought on by La Niña phases. By comparison, Figure 11, which shows the subset of observations from recent years, displays a drastic change from the previous map. In this figure we can see that bleaching events occur somewhat evenly across all phases of the ENSO cycle for recent years.

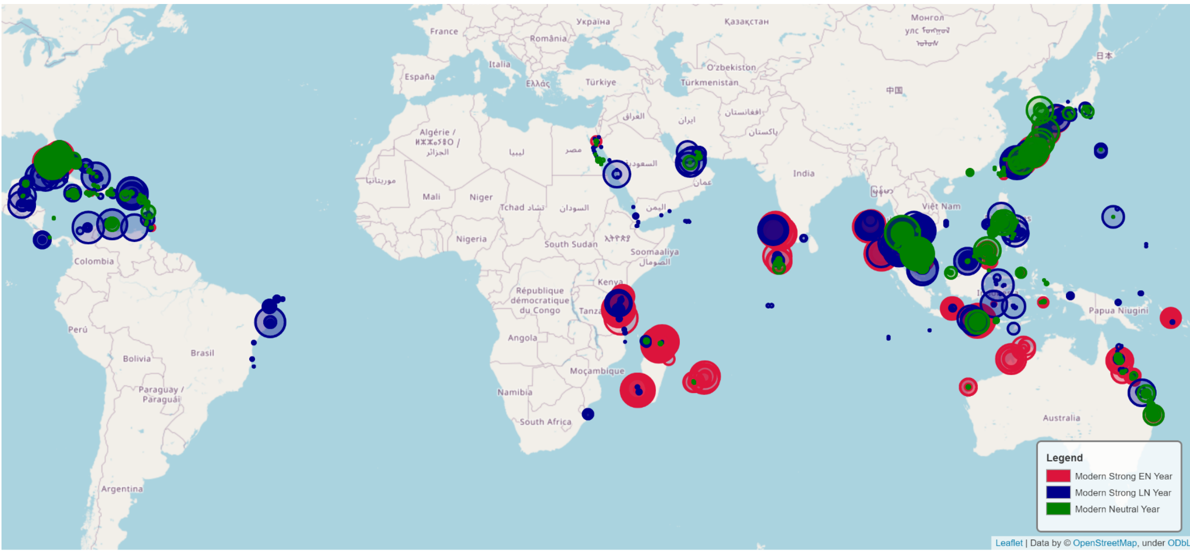

Figure 11. Map of bleaching events in recent years by ENSO cycle.

This means that bleaching has become more common in both Neutral and La Niña years. Figures 12 and 13 further emphasize this point by comparing bleaching in past and recent years of La Niña and Neutral phases, respectively.

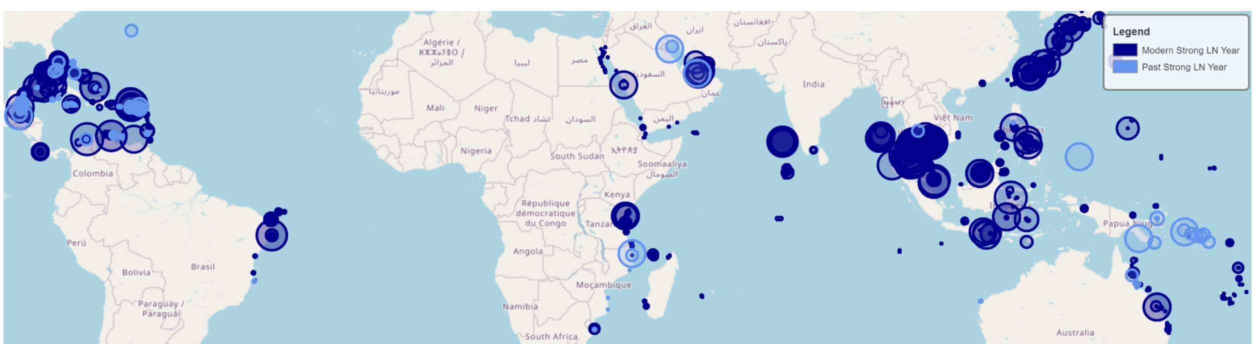

Figure 12. Map comparing bleaching events in past and recent La Niña phases.

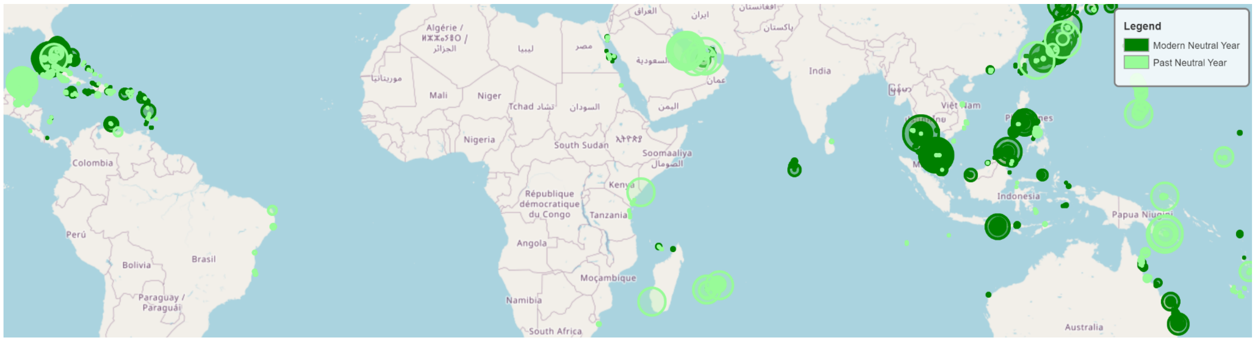

Figure 13. Map comparing bleaching events in past and recent Neutral phases.

After studying the degree to which ENSO phases are relevant to the dataset, we were able to better understand some of the temporal trends and cyclical bleaching observed. We can conclude that, while severe bleaching has occurred throughout time, it has historically been limited to El Niño phases, whereas it is now seen across all ENSO cycles. Thus, bleaching events are becoming more common, denying corals time to recover.

Modeling

Random Forest Classifier

A random forest is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting.

If we want to quantify how much each feature contributes to the specific probability score, we use SHapley Additive Explanation method (SHAP) to answer this question. We can train a random forest classifier and then construct a SHAP object to explain the prediction

.

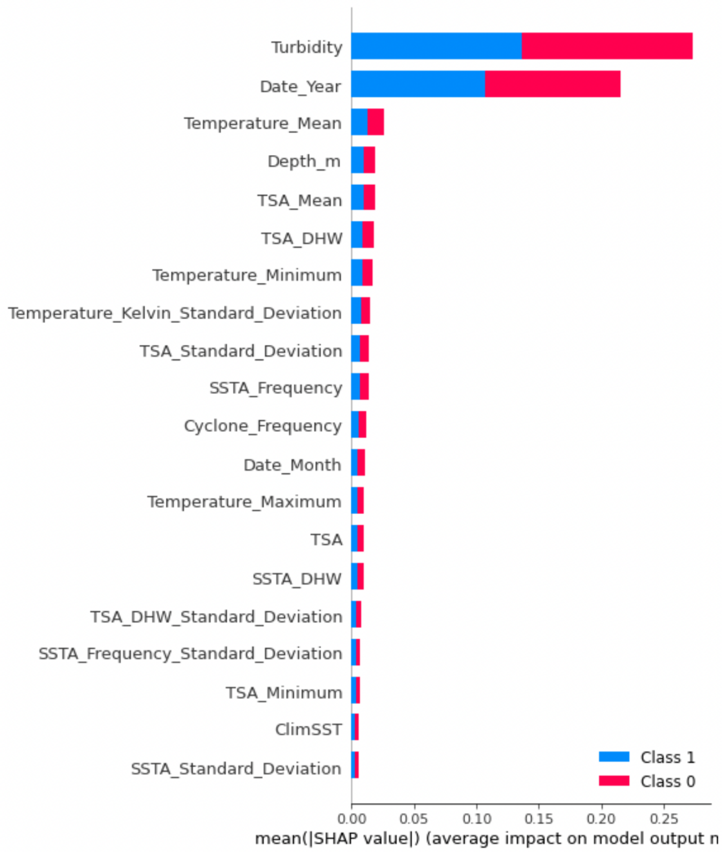

Figure 14. A SHAP summary plot of the variables that affect Bleaching Percent the most.

In the above graph, a variable belongs to class 1 if Percent Bleaching >= 50.0%, otherwise for Percent Bleaching < 50.0% it belongs to class 0. In our analysis, we will assume that class 1 signifies dead coral and class 0 represents alive corals.

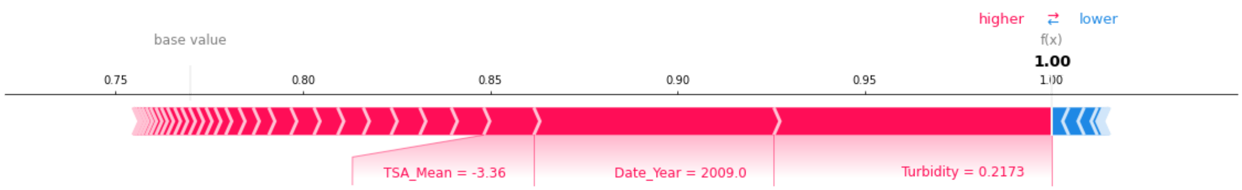

Figure 15. A SHAP force plot that displays how much each variable affects Percent Bleaching.

Figure 15 is a force plot which shows the influence of each feature on the current prediction on object ID 55. The base value is averaged over the predicted probability across all samples. The red arrow represents the features that have a positive influence on the prediction. The blue arrows represent the features that have a negative influence on the prediction. Since the probability that this data point belongs to class 0 is 0.99 and since the base probability for class 0 is 0.77, we can say that this data point belongs to class 0 and is alive, (Percent Bleaching < 50.0%). The force plot also gives us the features like year built, residential building value, longitude and land value for object 55.

Thus, from the model we can infer that turbidity, date (year), temperature (mean) and depth (m) respectively had the most effect on the bleaching percentage of the corals in the study. One explanation for the year being significant is due to the ENSO cycles, which were explained earlier in this report.

Discussion

Model Performance

Table 2. Random forest classifier evaluation.

Metric	Score
Accuracy	0.985
Precision	0.982
Recall	0.951

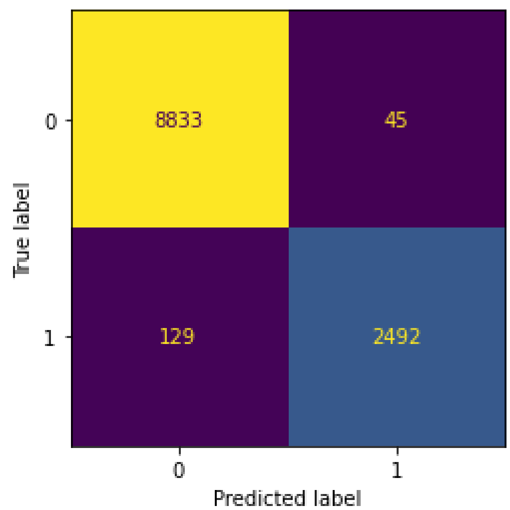

Figure 16. A confusion matrix for the Random Forest Classifier.

From the above Table 2, the accuracy of the model is 0.985. This can be explained with the equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Essentially, this means that the number of correct predictions are divided by the total number of predictions. With our data, that means all the correctly predicted Alive and Dead coral reefs are divided by the total number of predictions.

Table 2 also presents the precision of the model, being 0.982. This metric can be explained by the equation:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

This metric demonstrates the proportion of positive identifications which were correct, with ours being very high, which is an indicator of the good fitting model.

Lastly, Table 2 shows the recall of the model with the equation being:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

This metric is finding what proportion of actual positives were identified correctly. Again, at a value of 0.951, our model is indicated to be a good fit.

Figure 16 is a confusion matrix, which shows the True Positives, False Positives, False Negatives, and True Negatives. All of these confusion matrix values can be plugged into the above 3 equations to reinforce the model's accuracy, precision, and recall.

Limitations

This study was subject to several limitations. Firstly, there is insufficient data collected from 1980 to 1997, which negatively affected the data analysis process and made it difficult to detect trends in coral bleaching overtime. Secondly, as mentioned above, data was collected from dead coral reefs that had reached 100% bleaching. Moreover, the dataset does not include other important variables that are associated with coral bleaching, such as tide frequency and pollution level, which could lead to omitted-variable bias. Therefore, we recommend further research to take these limitations into account for more accurate results.

Conclusion

Main Findings

According to Figure 4, bleaching percent increases as water temperature increases in most ocean realms, meaning our first hypothesis is true. According to Figure 5, bleaching percent decreases as turbidity increases, indicating that turbid environments protect coral from bleaching. Our second hypothesis is true; however, extremely high turbidity levels may harm coral by preventing photosynthesis, so further studies are required to determine the healthy range of turbidity that corals can thrive in. For some ocean realms, bleaching percent increases as cyclone frequency increases, according to Figure 6. This means our third hypothesis is false. This correlation may be due to indirect impacts of cyclones including intense rainfall and high river loads, which increase turbidity to a harmful level and decrease salinity of reefs, resulting in bleaching. According to the correlation matrix and Figure 7, bleaching increases as distance to the shoreline increases, meaning our fourth hypothesis is false. However, most data collection sites were in the deep ocean zone, which may be the reason why there is a strong positive correlation between bleaching and distance to shoreline. More studies are required to examine the impact of distance to shoreline on bleaching.

Next Steps

A different approach to the data analysis could be splitting up the data into time intervals, such as 2015-2017. We could also focus on only one ocean realm to introduce more consistency in the data collection. Additionally, using different types of models, such as the KNeighborsClassifier, would diversify the findings. With the introduction of these different models, we could compare which one has the highest metrics and use that model instead. Finally, to reduce the error of recollecting data from already dead coral reefs, we could use an algorithm to account for the negative trend in percent bleaching to balance the data to have a more level high and low percentages when comparing said percent bleaching vs. time.

Bibliography

Carlson RR, Li J, Crowder LB and Asner GP (2022). Large-scale effects of turbidity on coral bleaching in the Hawaiian islands. *Front. Mar. Sci.* 9:969472. doi: 10.3389/fmars.2022.969472.

Davies, R. (2022, September 30). *Climate*. Britannica. https://www.britannica.com/science/climate-meteorology/Formation-of-tropical-cyclones

Donner, Simon D., et al. "A New, High-Resolution Global Mass Coral Bleaching Database." *PLOS ONE*, vol. 12, no. 4, 2017, https://doi.org/10.1371/journal.pone.0175490.

ICRI. (2020). *What are corals?* . Retrieved December 6, 2022, from https://icriforum.org/about-coral-reefs/what-are-corals/.

Kumagai, N. H., & Yamano, H. "High-Resolution Modeling of Thermal Thresholds and Environmental Influences on Coral Bleaching for Local and Regional Reef Management." *PeerJ*, vol. 6, 2018, https://doi.org/10.7717/peerj.4382.

Lapointe, B.E., Brewton, R.A., Herren, L.W. et al. (2019). Nitrogen enrichment, altered stoichiometry, and coral reef decline at Looe Key, Florida Keys, USA: a 3-decade study. *Mar Biol* 166(108) 39-46. https://doi.org/10.1007/s00227-019-3538-9.

National Oceanic and Atmospheric Administration. (2021, December 1). *What is coral bleaching?* National Ocean Service. https://oceanservice.noaa.gov/facts/coral_bleach.html

National Weather Service, NOAA's National Weather Service (2016, July 19). *What Is ENSO?*, https://www.weather.gov/mhx/ensowhat.

Nancy Knowlton, Smithsonian Ocean, Smithsonian Institution (2022). Corals and Coral Reefs.https://ocean.si.edu/ocean-life/invertebrates/corals-and-coral-reefs

Ritchie, H., Spooner, F., & Roser, M. (2021, April 15). Biodiversity. Retrieved December 6, 2022, from https://ourworldindata.org/biodiversity

Sully, S., et al. "A Global Analysis of Coral Bleaching over the Past Two Decades." *Nature Communications*, vol. 10, no. 1, 2019, https://doi.org/10.1038/s41467-019-09238-2.

U.S. Department of Commerce, N. O. and A. A. (2002, April 16). *Science education - does water temperature differ between near-shore and offshore sites? why or why not? - answer*. NDBC. Retrieved November 30, 2022, from https://www.ndbc.noaa.gov/educate/seatemp_ans.shtml.

U.S. Geological Survey. (n.d.). *Turbidity and water completed*. Turbidity and Water |. Retrieved December 6, 2022, from https://www.usgs.gov/special-topics/water-science-school/science/turbidity-and-water

van Woesik, R., Burkepile, D., & Kratochwil, C. (2022). *Global Bleaching and Environmental Data* (Version 2) [Dataset]. https://www.bco-dmo.org/dataset/773466.