

When Corals Lose Their Hue: A Study of Coral Bleaching

Introduction

"The Great Barrier Reef is in grave danger. The twin perils brought by climate change – an increase in the temperature of the ocean and in its acidity – threaten its very existence." - Sir David Attenborough

As Sir Attenborough has stated, the Great Barrier Reef, along with countless other reefs across the globe, are in danger caused by human activities that impact these delicate ecosystems, which support an estimated 25% of all known marine species.

Coral bleaching is a condition in which corals that are stressed by some external factor, such as fluctuating temperatures and changes in water chemistry, eject their symbiotic algae. These algae are critical to the health of the corals, providing a critical food source, in addition to their vibrant colors. Bleached coral can survive, but they are subject to higher rates of mortality.

In a paper by Indeever Madireddy, et al. titled "Using machine learning to develop a global coral bleaching predictor" [1] Madireddy and team used oceanic data from an earlier, simpler version of the Global Coral-Bleaching Database (GCBD) [2] to forecast which reefs should be allocated resources in order to preserve them. Our analysis uses the most recent version of the same data source as Madireddy, which has incorporated an additional 20 years of data, and it will be supplemented with fishing and fertilizer features from the World Bank's World Development Indicators (WDI) [3], to test the hypothesis that these features will help provide predictive power. Additionally, we used the Marine Ecoregions of the World (MEOW) [4] dataset to identify regional candidates for models.

Our goal is simple: Can we predict global and local bleaching patterns, and provide information to help local institutions, governments, and organizations mitigate coral bleaching? We will test fertilizer and fishing activity features to determine if they can be used to predict the health of coral reefs, along with various weather and geographic features. These results could help influence policy and decision making for local authorities.

An additional component of our analysis is to identify similar reef sites across the globe to give local communities a cluster of comparable reefs to observe long term patterns in coral bleaching and the activities which impact that pattern.

Methodology

In this section we cover the entire process that allowed us to generate our models and ultimately answer our primary question. We cover the dataset creation, cleaning and splitting which segues into the feature selection and importance determination. From there we touch on the various machine learning techniques that were explored in this project.

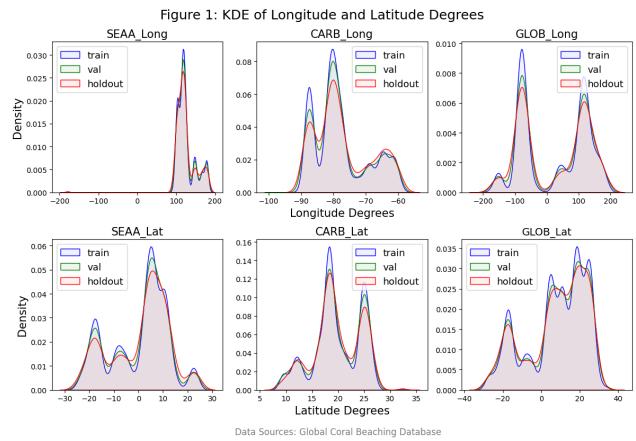
Datasets & Preparation

Our primary dataset for this project was Global Coral-Bleaching Database (GCBD), which contains measurements from reef sites across the world taken by 7 different research organizations/projects between 1980-2020. We supplemented this with data from the World Bank's World Development Indicators (WDI) to include features on fishing and agriculture, as well as data from the Marine Ecoregions of the World (MEOW) to include features on the ecological region of the reef sites. GCBD and WDI are covered by the Creative Commons Attribution 4.0 license, while MEOW is public for non-commercial use with attribution. A data dictionary of evaluated features can be found in Appendix E.

In efforts to clean our data we restricted our analysis to 1998-2020, as there appeared to be an issue of selection bias in measurements pre-1998 where the majority of reef sites surveyed showed high levels of bleaching. We also limited our investigation to hard corals, as not all soft corals exhibit the same symbiotic relationship with algae [5].

In addition to analyzing Coral Bleaching on a global scale we also wanted to take a closer look at the Caribbean and Southeast Asia and Australia (denoted as CARB and SEAA, respectively, in this report). Focusing on these regions allowed us to improve model performance by allowing the model to better learn the nuance of the regions, for example the Pacific ocean is generally colder than the Atlantic ocean.

From here we chose to split our dataset into training, validation, and holdout datasets using an 70/15/15 split to combat any data leakage issues. We verified that features like depth, location, and sampling timing by month were all equally distributed such that any one of these splits would not contain disproportionate amounts of information (Fig. 1)

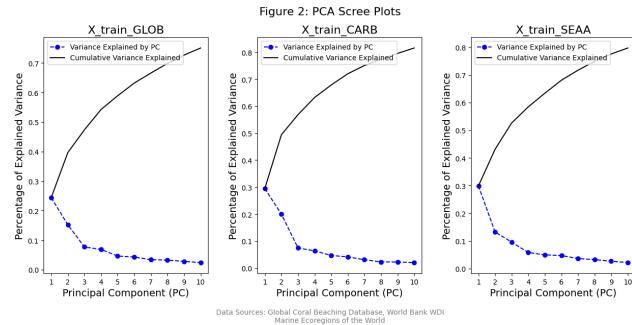


Feature Reduction & Selection

We explored various methods to select and reduce our feature set to improve our models performance including Principal Component Analysis, Variance Inflation Factor Analysis, Shapley Additive Explanations to understand interaction variables, and a robust feature selection process.

Principal Component Analysis

Since our cleaned dataset contained more than 90 features, we decided to perform a PCA analysis, thinking it might be useful in reducing the dimensionality of the dataset by providing latent features.



The “elbow” of the Scree plots (Fig. 2) all lined up with the first 3 principal components and account for approximately 50% of the total explained variance. We determined that in order to retain a majority of the explained variance, we needed to retain almost all of the original features, indicating each feature was already linearly independent. Based on this we looked to other methods to help reduce the feature space.

Variance Inflation Factor (VIF)

One aspect of our data that could lead to difficulties in interpretability is potential multicollinearity of the features. With highly multicollinear data, it is difficult to attribute the contributions of specific features. Initial testing with Variance Inflation Factor (VIF) indicated a massive degree of multicollinearity within the 57 continuous features that comprised the dataset, especially around the WDI data. After reducing the dataset to basic weather, fishing, and fertilizer related features, the VIF scores for the remaining 17 features were 5 or lower, indicating an acceptable degree of multicollinearity, with a loss of only 0.3 MAE when compared to the full dataset using an unoptimized XGBoost model.

Shapley Additive Explanations (SHAP)

We next used Shapley Additive Explanations against all 3 regions and 17 features to investigate complex interactions within the data. When performing summary plots against each region, the fishing and fertilizer features were consistently in the top interactions for all regions.

We created interaction variables for fishing (Number of Threatened Fish Species * Fish Captured Metric Tons), and for fertilizer (Fertilizer Kg per hectare / Km² Agricultural Land). The original features were then removed from the data in order to prevent additional multicollinearity issues, which was consistent with WDI features (Fig. 3).

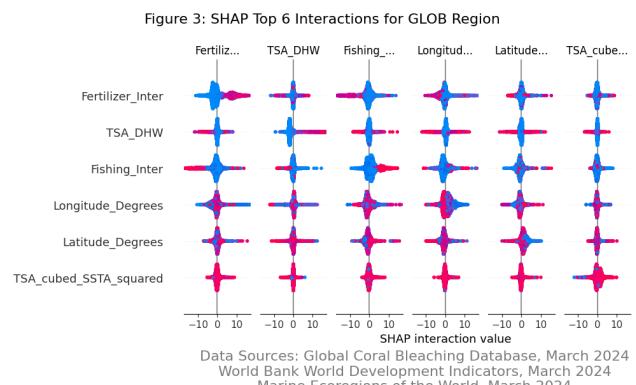
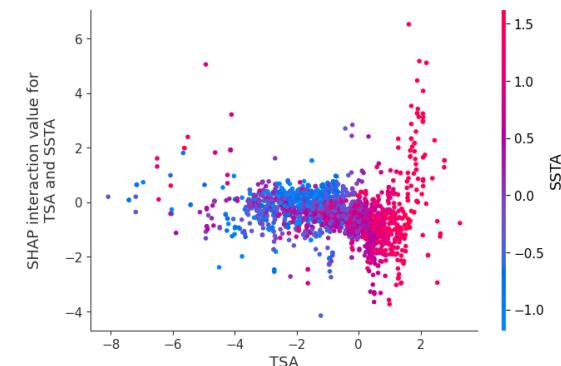


Figure 4: SHAP Dependence Plot for TSA and SSTA in SEAA Region



Data Sources: Global Coral Bleaching Database, March 2024
World Bank World Development Indicators, March 2024
Marine Ecoregions of the World, March 2024

Additional investigation into the interactions for SSTA and TSA (Sea Surface Temperature Anomaly and Thermal Stress Anomaly, respectively) indicated a potentially complex relationship (Fig. 4), prompting the creation of polynomial features, of which the TSA_cubed_SSTA_squared feature was most significant.

Considering our interest in determining if water turbidity was a contributing factor in coral bleaching, we also used domain experience to create two additional features that utilized Turbidity, measured using [Kd490nm attenuation](#). The first used the previously created fertilizer interaction feature and multiplied that by turbidity, and the second multiplied distance to shore against turbidity.

The features of interest that survived were:

Depth_m, Distance_to_Shore, Turbidity, Longitude_Degrees, Latitude_Degrees, ClimSST, Cyclone_Frequency, SSTA_Frequency, TSA_Frequency, Fertilizer_Inter, Fishing_Inter, Dist_to_Shore_and_Turbidity_Interaction, Fert_and_Turbidity_Interaction, TSA_cubed_SSTA, TSA_cubed_SSTA_squared, SSTA, TSA, SSTA_DHW, and TSA_DHW

Final Feature Selection

Given our regional approach to the data, we wanted to determine the best combination of features and models. Our criteria for selecting models included the need for regularization, the ability to handle missing data, parallelization, and the capability of handling non-linear relationships. RandomForest, HistGradientBoosting, XGBoost, and LightGBM all passed these criteria and were included in our analysis.

With models determined, we began the feature selection process. For combinations of 19 features, this meant roughly 524k models would be built for each model type and region combination. Initial testing indicated RandomForestRegressor would not be well suited to this task, given a runtime of 3 minutes 45 seconds to build 3 models (compared to 1.6 seconds for LightGBM). HistGradientBoostingRegressor was not selected for this task despite providing significantly better MAE scores than XGBoost and LightGBM, as it took 324% and 235% more time to train, respectively.

With XGBoost and LightGBM selected for customized features, we then built over 15.7 million models to test for optimal feature combinations by region, model type and using 5 cross validation folds, taking approximately 60 hours to complete. These features were then ranked and the features providing the best MAE results were selected for use in the model creation process.

Model Selection & Creation

In addition to the decision tree algorithms mentioned above, we also tested a dummy regressor and an ElasticNet regressor to ensure our predictions were at least as good as guessing (DummyRegressor) and to see how well a highly biased model would perform (ElasticNet). Hyperparameter tuning for each model type was performed using Hyperopt's Bayesian optimization with a 5-fold cross validation. Models were then tested against the validation data set and hyperparameters were tracked using Neptune.ai. By using the validation data set, we were able to ensure that no data leakage occurred in choosing our final hyperparameters, as these hyperparameters were disassociated from the holdout data set, which we evaluated after feature and model selection was finalized.

Our primary evaluation metric was Mean Absolute Error (MAE), since we were predicting a continuous value (Percent Bleached) between 0 and 100. This made the MAE metric easily understandable and explainable. We also calculated Root Mean Squared Error (RMSE) and R-squared (R²) so we could compare our model performance to the Madireddy, et al. paper.

Prior to scoring against the holdout data, we chose to perform another iteration of tuning whereby we ablated the feature set for each model/region to see if the model performance improved. By running feature ablation we were confident that not only were our hyperparameters optimized, but our feature space was also optimized. In a few cases, due to the hyperparameter tuning, it was found that some features could be dropped to very slightly improve model performance, but the maximum value difference observed was 0.1 MAE out of 100, so we elected not to drop any features as a result of ablation.

For our unsupervised clustering model, we used a combination of Silhouette scores and an elbow plot to determine the number of clusters that would be best. We then used a K-mean clustering algorithm to cluster the reef sites worldwide and by region. Our Silhouette scores determined the number of clusters to choose was above 0.8 (Appx. C), indicating good clustering.

Results

For this project, a successful outcome would be a model that is well suited to predict coral bleaching (one that minimizes error) while also being highly explainable. Supplementary to our regression model, we wanted to use clustering methods to determine similar reef sites that can be used with our model to inform stakeholder decision making. Below is a table (Tbl. 1) that shows the model type ('model') and the region (ex: GLOB, SEAA, CARB) in which that model was used. The three columns to the right show the evaluation metrics: mae, rmse, and rsq.

Coral Bleaching Model Accuracy

Table 1: All Hyperparameter Tuned Model Results for the Holdout Data Set

	model object	region object	mae float64 7.2761 - 19.1872	rmse float64 13.7649 - 24.7044	rsq float64 0.0 - 0.5843	
	LGBM	CARB	33.3%			
	LGBM	CARB	16.7%			
	HGBM	CARB	33.3%			
	4 others	SEAA	33.3%			
0	LGBM	CARB		9.1064	16.6511	0.5457
1	HGBM	CARB		9.2632	15.9285	0.5843
2	XGB	CARB		9.3033	16.3587	0.5615
3	RF	CARB		11.4058	17.2596	0.5119
4	ENET	CARB		17.0999	22.9403	0.1377
5	Dummy	CARB		19.1872	24.7044	0
6	LGBM	GLOB		8.5803	16.6875	0.4274
7	XGB	GLOB		8.7163	15.7133	0.4923
8	HGBM	GLOB		9.0904	15.4919	0.5065
9	RF	GLOB		10.4469	16.3797	0.4483
10	ENET	GLOB		14.7235	20.9536	0.0971
11	Dummy	GLOB		16.2511	22.052	0
12	XGB	SEAA		7.2761	14.7631	0.2714
13	LGBM	SEAA		7.3996	14.1894	0.3269
14	HGBM	SEAA		7.4783	13.8646	0.3574
15	RF	SEAA		7.983	13.7649	0.3666
16	ENET	SEAA		10.6908	16.7972	0.0568
17	Dummy	SEAA		11.2259	17.2954	0

18 rows, showing 25 ▾ per page

« « Page 1 of 1 » »

↓

Referencing Tbl. 1, we can see that for the Holdout data set the best MAE values for each region were provided by XGBoost and LightGBM, and were in the range of 7-9 MAE , meaning that our models were able to make predictions with an average accuracy of 91-93%. The features that were used to train the XGBoost and LightGBM models for each region, along with their importance, can be observed in the interactive feature importance graphs below (Fig. 5 and 6). These results indicate we are successful at creating models that are capable of reliably predicting coral bleaching using the decision tree models. ElasticNet, even hypertuned, was not able to achieve a reasonable MAE score due to the aforementioned lack of linearity in the data.

Ultimately, comparing these results to the Madireddy et al. paper was not realistic. The Madireddy paper utilized an earlier version of the Global Coral Bleaching Database, which contained fewer sources of data, and fewer differences in data collection and measurement methods, which we believe created a situation of selection bias, resulting in an exceptionally high performing model (7.91 RMSE). The five weather features that were used in the Madireddy paper were tested as part of the final feature selection process with our more complete dataset, and they did not provide the best results.

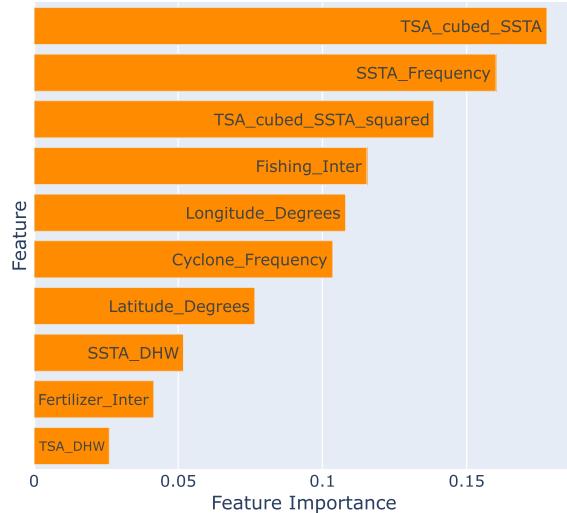
LightGBM Model Feature Importance

GLOB ▾

XGBoost Model Feature Importance

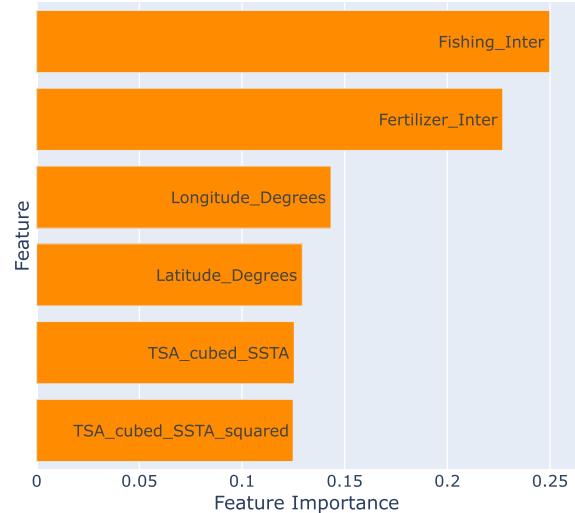
GLOB ▾

Fig. 5: Feature Importance for LGBM SEAA



Data Sources: Global Coral Bleaching Database, 2024
World Bank WDI, 2024
Marine Ecoregions of the World, 2024

Fig. 6: Feature Importance for XGB SEAA



Data Sources: Global Coral Bleaching Database, 2024
World Bank WDI, 2024
Marine Ecoregions of the World, 2024

Discussion

Interpreting the Regression Models

The differences in MAE accuracy between LightGBM and XGBoost were negligible, but model complexity tended to be higher for LightGBM than XGBoost when considering the high number of leaves and estimators and the low number of minimum child samples for LGBM. This complexity was felt during the hyperparameter tuning process, where LightGBM took 73.9 seconds per trial for the GLOB region, while XGBoost took 15.32 seconds. LightGBM's model was 0.09 MAE better than XGBoost on the validation data, but required substantially more time to tune. Our preference for future work will be to utilize XGBoost, due an acceptable tradeoff between speed and accuracy.

The approach of using regional models is justified when comparing the list of features (Fig. 5 & 6) that were important to each region/model combination. Using XGBoost as our preferred model, and focusing on the difference between the Caribbean and SEAA regions, Turbidity was a factor to the Caribbean model both as a predictor and as an interaction variable, while the SEAA models did not use Turbidity in any way. The Fishing, Fertilizer, and TSA_cubed_SSTA_squared interaction features were important to both models, but items such as Depth and Distance to Shore were important to only the Caribbean model. This validates the idea that some features have predictive power in one region, but not another.

Interpreting Our Predictions

Region

Data Split

Minimum Error



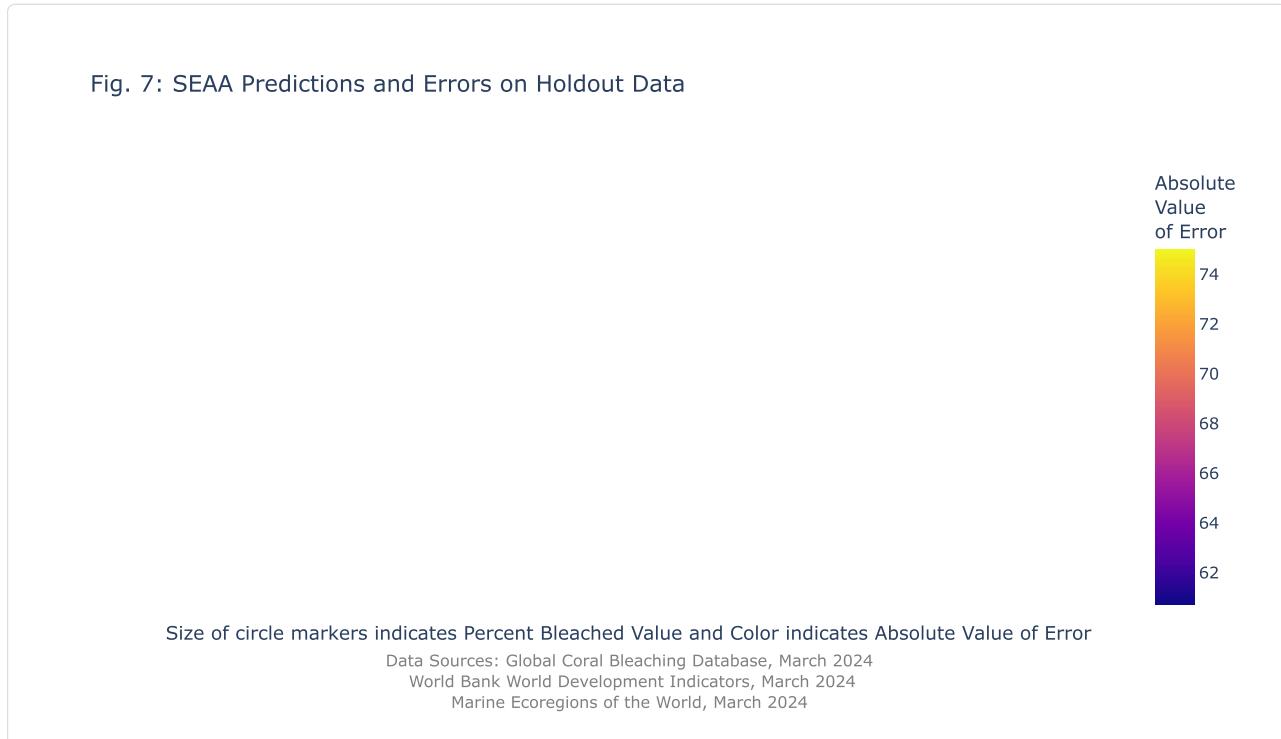
60 / 100



100 / 100

Maximum Error

Fig. 7: SEAA Predictions and Errors on Holdout Data



To evaluate the efficacy of our approach and models, we utilized the absolute value of error on the predictions on the holdout data, as indicated by Table 2. These results indicate our best performing model was the SEAA region, with the lowest mean error, standard deviation, and the lowest 95th percentile MAE (calculated as Mean Abs Val Err + 1.96*StdDev). Per Table 2, 95% of our predictions were within roughly 33 MAE or less of the actual percent bleached values, and within 1 MAE for the mean, for all models. Details on some predictions can be viewed as interactive SHAP waterfall plots in Appendix D.

Table 2 Absolute Value of Error of Predictions				
Region	Split	Mean Abs Val Error	Std Dev	95th Percentile MAE
GLOB	Holdout	0.27	16.69	32.98
CARB	Holdout	0.63	16.64	33.24
SEAA	Holdout	0.02	14.76	28.95

We then used interactive visualizations (Fig. 7) to explore the geographic distribution of the absolute value of errors. In Fig. 7, when selecting the SEAA region, the Holdout data split, and a minimum error of 50, there is a small cluster of records that have a high value of bleaching between Brunei and Philippines, but our model predicted very low numbers. This example demonstrates the presence of poorly modeled stressors in the Palawan/North Borneo ecoregion, and justifies additional regional splits.

To create regional models for this analysis, we used the Marine Ecoregions of the Word (MEOW) "Provinces" to identify analytic regions. The relatively poor performance of the Caribbean region may be explained because it is composed of a single Province, but the SEAA region is made up of 9 smaller Provinces, whose mean absolute value of error on predictions can be reviewed in Table 3.

For future analysis, we should consider breaking up the SEAA region into a few, smaller regions based on Provinces, giving more granularity to the models. The Caribbean region should instead utilize the MEOW "Ecoregions", which enables the "Tropical Northwestern Atlantic" Province to be broken up into 9 smaller Ecoregions for the Caribbean. Another approach would be to leverage clustering of MEOW Provinces or Ecoregions to create non-regional models. That approach is outside the scope of this analysis, but Fig. 8 illustrates how this distribution could potentially look if we clustered by site.

Table 3: Mean Absolute Value of Error of Predictions per Province

Note: The table columns can be filtered and sorted by clicking on the column headers.

	Region object	PROVINCE object	Split object	Mean_Abs_Differ...	Count_of_Obs int64	
	GLOB	74.7%	Java Transiti...	5.3%	0.44 - 26.08	
	SEAA	22.7%	South China ...	5.3%	Validation 53.3%	
	CARB	2.7%	28 others	89.3%	Holdout 46.7%	
62	SEAA	Java Transitional	Holdout	0.44	10	
42	GLOB	Tropical Northwe...	Holdout	1.56	18	
30	GLOB	South China Sea	Holdout	2.09	55	
60	SEAA	Eastern Coral Tria...	Holdout	2.16	23	
4	GLOB	Bay of Bengal	Validation	2.16	2	
32	GLOB	South Kuroshio	Holdout	2.6	22	
38	GLOB	Tropical East Paci...	Holdout	3.08	4	
68	SEAA	South China Sea	Validation	3.4	120	
67	SEAA	South China Sea	Holdout	3.41	74	
25	GLOB	Red Sea and Gulf...	Holdout	3.71	106	

75 rows, showing 10 per page

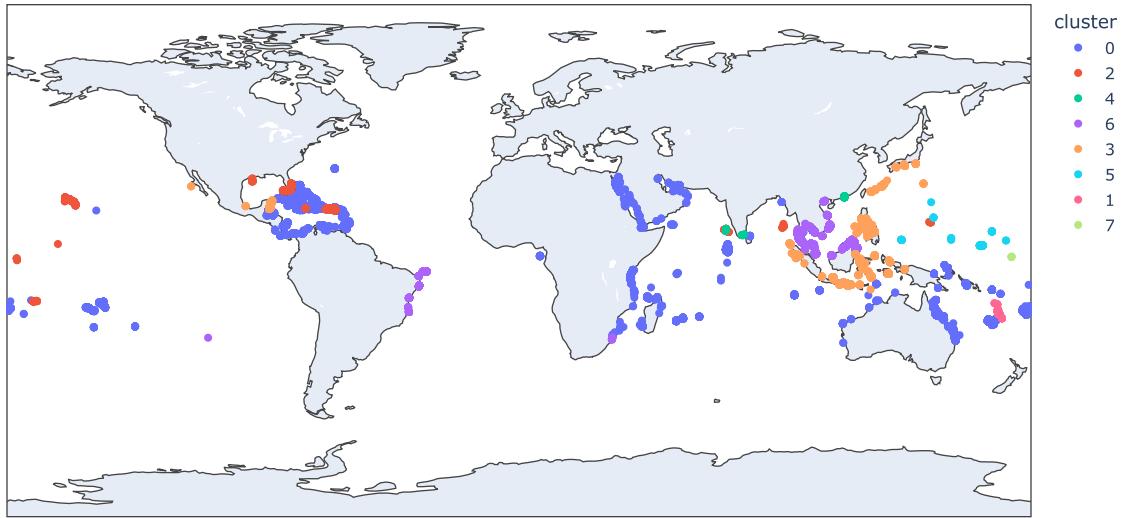
<< < Page 1 of 8 > >>

↓

Interpreting the Clustering Model

To further improve our ability to give recommendations and assist local decision makers, we performed clustering on the unlabeled data. When deciding the number of clusters to use we wanted strong clusters but also wanted to avoid having clusters that simply represented the eco-regions. Figure 8 shows the results of the Global k-means clustering, where we were able to identify 8 clusters. Seven of the eight clusters successfully had reef sites from multiple ecoregions, with the last cluster only including the Kiribati, a small island nation. These clusters can be used to inform local decision-makers about bleaching levels in similar reefs sites around the world and potential interventions to reduce coral bleaching.

Figure 8: Global Coral Bleaching Clustering Map



Broader Impacts

This project was developed to help inform stakeholders about the stressors that affect coral reefs worldwide. The features that were most successful in our models indicate a complex relationship of factors that impact coral bleaching, from climate change to fishing and the agricultural industries that leverage fertilizers. The stakeholders in conversations to mitigate coral bleaching are not just Oceanic organizations like [NOAA](#), but also the industrial, manufacturing, fishing and agricultural sectors as well.

While intending positive change, the findings could also be exploited by "bad actors". For example, a Florida Keys study from 2009-2012 showed elevated nitrogen and phosphorus levels tripled coral bleaching and doubled disease prevalence [6]. Fertilizer companies may lobby with governments to protect their interests so it is paramount to effectively communicate these multifaceted threats to all stakeholders and keep reef health in the forefront of the discussion.

Final Thoughts

In this project, we built several regression models, but preferred XGBoost as it is easily explainable, quick to tune/train, and produced accurate results. Our K-means clustering model produced reasonable site groupings to assist local decision makers as to how other similar reefs are managing this complex issue.

Regardless of region and model, there seemed to be a strong correlation to weather stresses indicating that climate change is a major contributor to coral bleaching. In further analyses it is recommended to cross reference with El Nino/Nina data to see how these global stressors impacted coral bleaching. Outside of weather, we found that fertilizer and fishing were consistently important across all models and regions. Looking at other WDI indicators like tourism may lead to some interesting iterations of this project. Oceanic features like turbidity, depth, and distance to shore were crucial for the CARB region but were not important for SEAA. Furthermore, CARB produced our worst results indicating that dividing the region up may be a fruitful approach for further interactions.

We feel that this analysis successfully enables stakeholders to understand and forecast coral bleaching in reefs nearest them. We hope that this information allows for a jumping off point for future projects, modeling, and results that can help manage coral bleaching for years to come.

Special Thanks

We would like to acknowledge our mentor, Kyle Balag, for helping refine multiple aspects of this project. He provided assistance in identifying resources, as well as direction in focusing on a regression analysis and investigating a regional approach, which significantly improved the project.

Statement of Work

This project was developed by James Conner, Cody French, and Michael Vizzini as part of our Capstone Project for the University of Michigan Master of Applied Data Science program. Below shows a highlight of the contributions for each member of the team.

James Conner:	Cody French:	Michael Vizzini:
<ul style="list-style-type: none">Interactive VisualizationsData Wrangling/CleaningLead Modeling EngineerLead Analytics Engineer	<ul style="list-style-type: none">Data CleaningLead Data VisualizerModeling EngineerLead Unsupervised Learning Engineer	<ul style="list-style-type: none">Project ManagerData CleaningModeling EngineerData Visualizer

Citations

[1] Indeever Madireddy, et al. "Using Machine Learning to Develop a Global Coral Bleaching Predictor." *Journal of Emerging Investigators*, vol. 6, no. 8, 21 Feb. 2023, emerginginvestigators.org/articles/22-056/pdf, <https://doi.org/10.59720/22-056>. Accessed 2 Apr. 2024.

[2] van Woesik, Robert; Kratochwill, Chelsey (2022). *Global Coral Bleaching Database*. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.17076290>. Accessed 21 Feb. 2024.

[3] World Bank, World Development Indicators. (2024). *World Development Indicators / DataBank*. Dataset. <https://databank.worldbank.org/source/world-development-indicators>

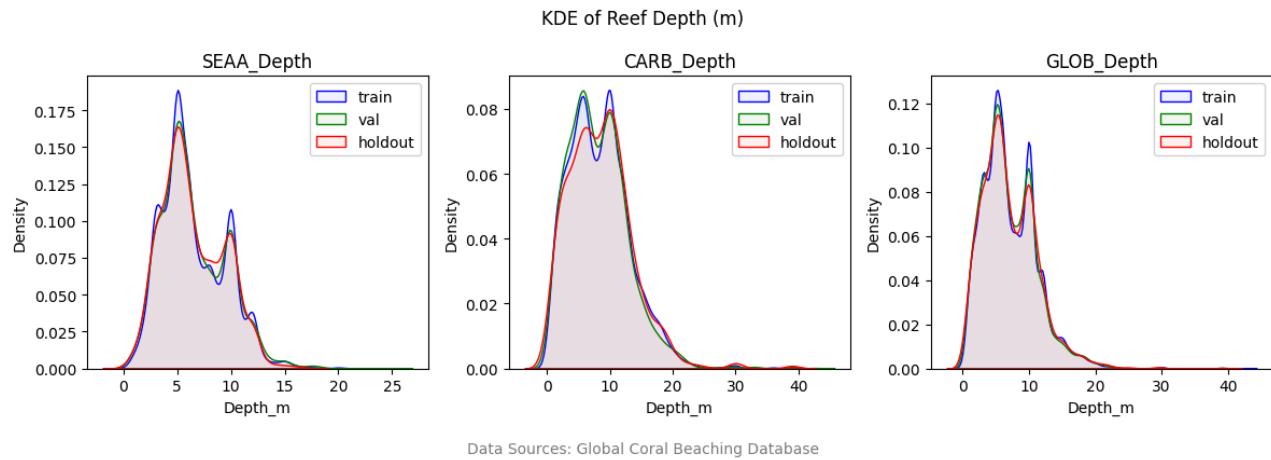
[4] Geospatial Conservation Atlas, Marine Ecoregions Of the World. (2019). *Marine Ecoregions Of the World (MEOW) / Geospatial Conservation Atlas*. Dataset. https://geospatial.tnc.org/datasets/ed2be4cf8b7a451f84fd093c2e7660e3_0/about

[5] Difference between hard and soft coral, situation of coral in Kenya. Conservation Education Society. (n.d.). <https://ceskenya.org/difference-between-hard-and-soft-coral-situation-of-coral-in-kenya/> Accessed 5 Apr. 2024.

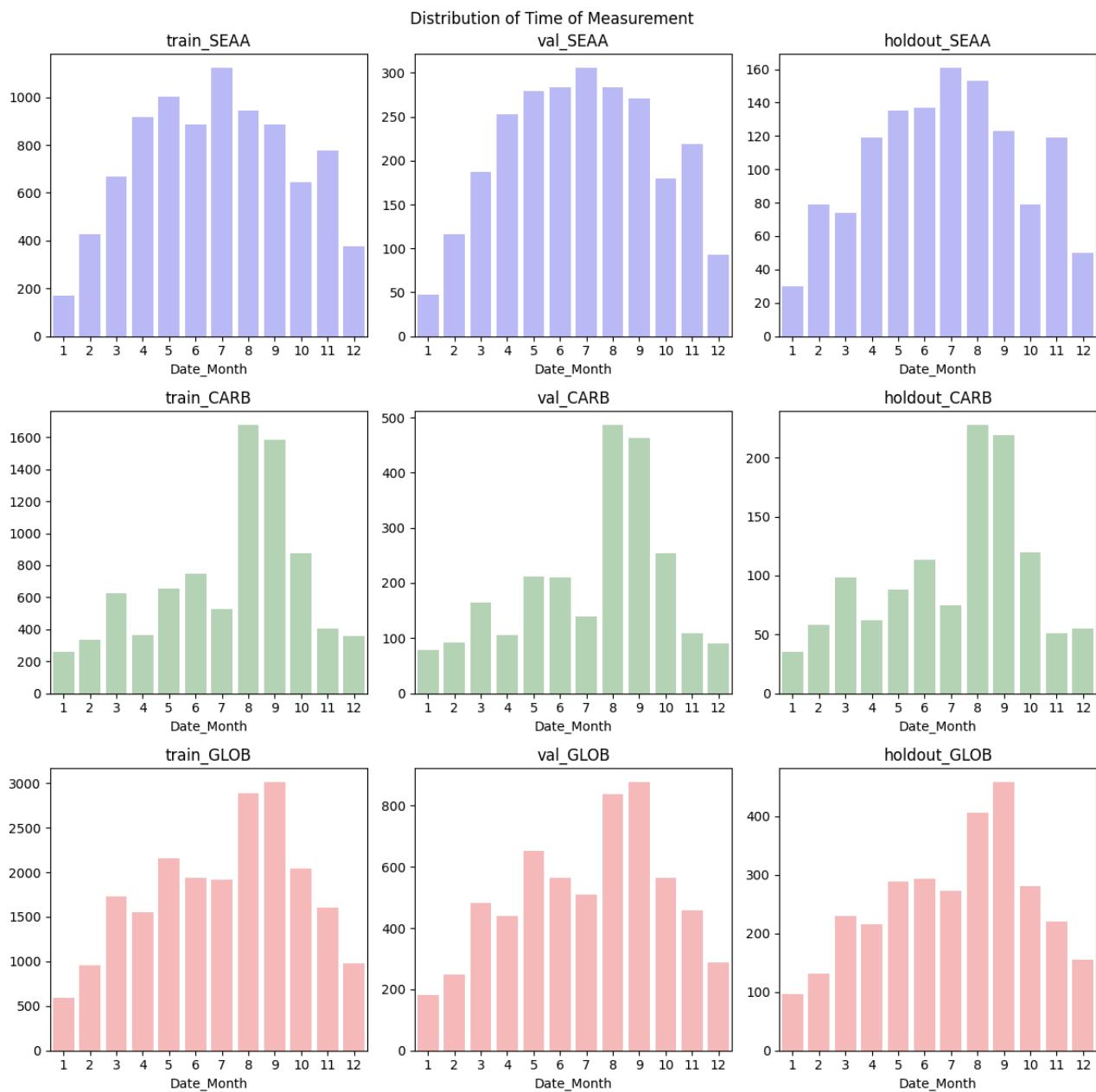
[6] Vega-Thurber, Rebecca. "Large Study Shows Pollution Impact on Coral Reefs - and Offers Solution." Oregon State University, Oregon State University, 25 Nov. 2013, today.oregonstate.edu/archives/2013/nov/large-study-shows-pollution-impact-coral-reefs-%E2%80%93-and-offers-solution. Accessed 16 Apr. 2024.

Appendix

Appendix A: Data Split Validation, Probability Density Plot on Reef Depth (m)



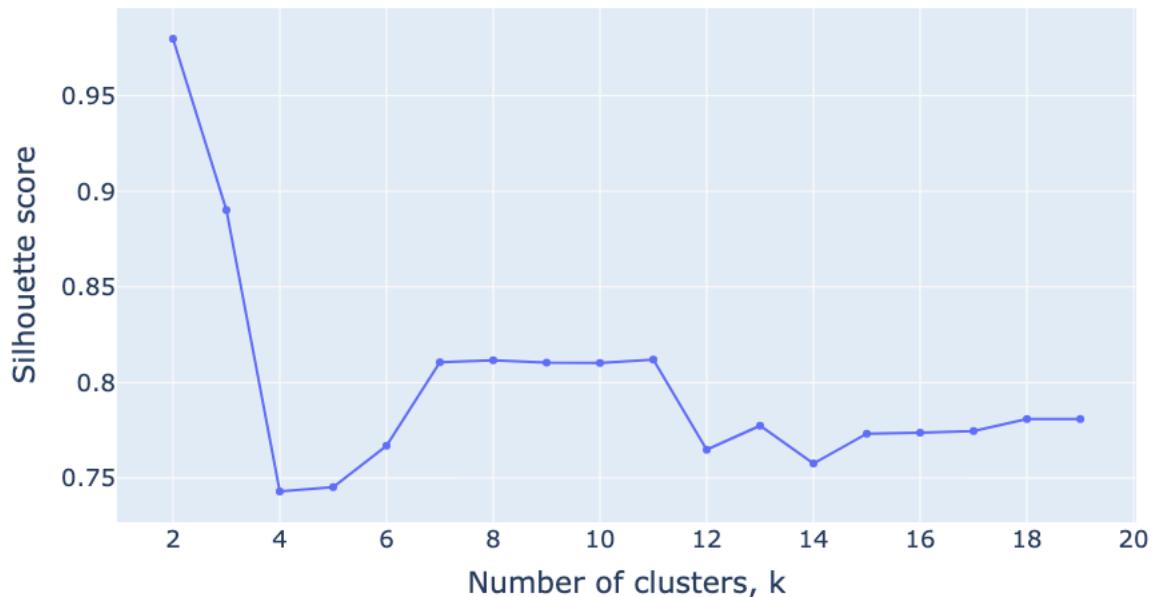
Appendix B: Data Split Validation, Histogram of Month of Observation



Data Sources: Global Coral Beaching Database

Appendix C: Silhouette Scores for K-means Cluster on Global Region

Silhouette score vs. Number of clusters (Global)



Appendix D: Waterfall Plots for Worst and Best Predictions

Region for Poor Predictions

GLOB

Region for Good Predictions

SEAA

Data Split for Poor Predictions

Validation

Data Split for Good Predictions

Holdout

Poor Prediction Records

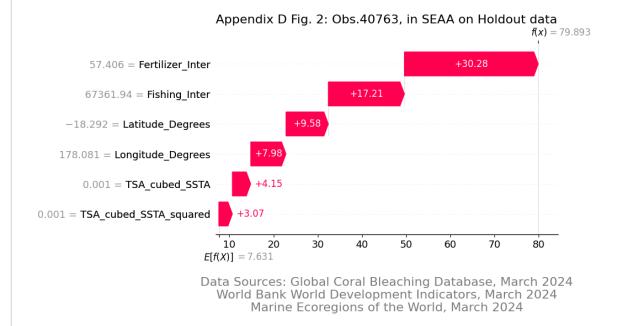
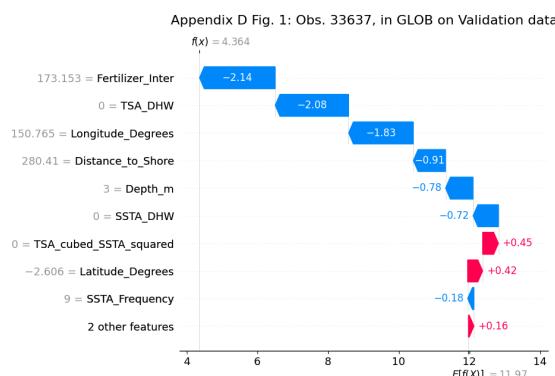
33637

Good Prediction Records

40763

Actual Bleached Value: 100.0
Prediction: 4.36
Absolute Value of Error: 95.64

Actual Bleached Value: 80.0
Prediction: 79.89
Absolute Value of Error: 0.11



Data Sources: Global Coral Bleaching Database, March 2024
World Bank World Development Indicators, March 2024
Marine Ecoregions of the World, March 2024

Appendix E: Data Dictionary

Derived Features:	
Name	Description
Fishing_Inter	Fishing Interactions ... WDI's EN.FSH.THRD.NO * ER.FSH.CAPT.MT
Fertilizer_Inter	Fertilizer Interactions ... WDI's AG.LND.AGRI.K2 / 'AG.CON.FERT.ZS'
Fert_and_Turbidity_Interaction	Fertilizer_Inter * Turbidity
Dist_to_Shore_and_Turbidity_Interaction	Dist_to_Shore * Turbidity
TSA_cubed_SSTA_squared	TSA^3 * SSTA^2
TSA_cubed_SSTA	TSA^3 * SSTA
Percent_Bleached_Value	Normalized calculation of Reefcheck and Donner measurement methods, using Percent_Bleached, S1, S2, S3, S4, and Severity_Code

Global Coral Bleaching Database:	
Latitude_Degrees	Latitude coordinates in decimal degrees.
Longitude_Degrees	Longitude coordinates in decimal degrees.
Realm_Name	Identification of realm as defined by the Marine Ecoregions of the World (MEOW).
Ecoregion_Name	Identification of the Ecoregions as defined by Veron et al.
Country_Name	The country where sampling took place.
Distance_to_Shore	The distance (m) of the sampling site from the nearest land.
Turbidity	kd490 with a 100-km buffer.
Cyclone_Frequency	Number of cyclone events from 1964 to 2014.
Depth	Depth (m) of the sampling site.
Substrate_Type	Substrate type ID field from Substrate_LUT.
S1, S2, S3, S4	Point data from different segments of transects.
S1, S2, S3, S4	Percent bleaching from different segments of transects.
Severity_Code	Coded range of bleaching severity from Donner et al.
Percent_Bleached	Percent of coral bleaching.
ClimSST	CoRTAD. Climatological Sea-Surface Temperature (SST) based on weekly SSTs for the study time frame, created using a harmonics approach.
Windspeed	CoRTAD. Meters per hour.
SSTA	CoRTAD. Sea-Surface Temperature Anomaly, weekly SST minus weekly climatological SST.
SSTA_Frequency	CoRTAD. Sea Surface Temperature Anomaly Frequency, the number of times over the previous 52 weeks that SSTA \geq 1 degree Celsius.
SSTA_DHW	CoRTAD. Sea Surface Temperature Degree Heating Weeks, the sum of the previous 12 weeks when SSTA \geq 1 degree Celsius.
TSA	CoRTAD. Thermal Stress Anomaly, weekly SSTs minus the maximum of weekly climatological SSTs in degrees Celsius.
TSA_Frequency	CoRTAD. The number of times over the previous 52 weeks that TSA \geq 1 degree Celsius.
TSA_DHW	CoRTAD. Thermal Stress Anomaly Degree Heating Weeks, the sum of the previous 12 weeks when TSA \geq 1 degree Celsius.

Marine Ecoregions of the World:	
Name	Description
ECOREGION	Marine Ecoregions of the World
PROVINCE	Marine Provinces
REALM	Marine Realms
geometry	Shape data required to draw polygons on maps

World Bank World Development Indicators:	
Code	Description
AG.CON.FERT.PT.ZS	Fertilizer consumption (% of fertilizer production)
AG.CON.FERT.ZS	Fertilizer consumption (kilograms per hectare of arable land)
AG.LND.AGRI.K2	Agricultural land (sq. km)
AG.LND.AGRI.ZS	Agricultural land (% of land area)
AG.LND.FRST.K2	Forest area (sq. km)

World Bank World Development Indicators:	
AG.LND.FRST.ZS	Forest area (% of land area)
AG.LND.IRIG.AG.ZS	Agricultural irrigated land (% of total agricultural land)
EN.FSH.THRD.NO	Fish species, threatened
EN.POP.DNST	Population density (people per sq. km of land area)
ER.FSH.AQUA.MT	Aquaculture production (metric tons)
ER.FSH.CAPT.MT	Capture fisheries production (metric tons)
ER.FSH.PROD.MT	Total fisheries production (metric tons)
ER.MRN.PTMR.ZS	Marine protected areas (% of territorial waters)
ER.PTD.TOTL.ZS	Terrestrial and marine protected areas (% of total territorial area)
NV.AGR.TOTL.CD	Agriculture, forestry, and fishing, value added (current US\$)
NV.AGR.TOTL.KD.ZG	Agriculture, forestry, and fishing, value added (annual % growth)
NV.AGR.TOTL.ZS	Agriculture, forestry, and fishing, value added (% of GDP)
NY.ADJ.DFOR.CD	Adjusted savings: net forest depletion (current US\$)
NY.ADJ.DFOR.GN.ZS	Adjusted savings: net forest depletion (% of GNI)
NY.GDP.FRST.RT.ZS	Forest rents (% of GDP)