Coral Beaching Prediction in India

Group 11 - Rishit Anand, Rhythm Mehra, Shrijak Kumar

coral reefs *are the barometers of ocean health – when they suffer, the entire marine ecosystem is at* **risk.**

coral bleaching

Coral bleaching is a natural process that occurs when corals expel algae from their tissues, turning them white. It's a response to stressful conditions like rising water temperatures, pollution, and changes in salinity.

effects

Loss of Marine Biodiversity

Corals provide shelter and food for 25% of all marine species. Their decline disrupts entire ecosystems.

Erosion of Coastal Protection

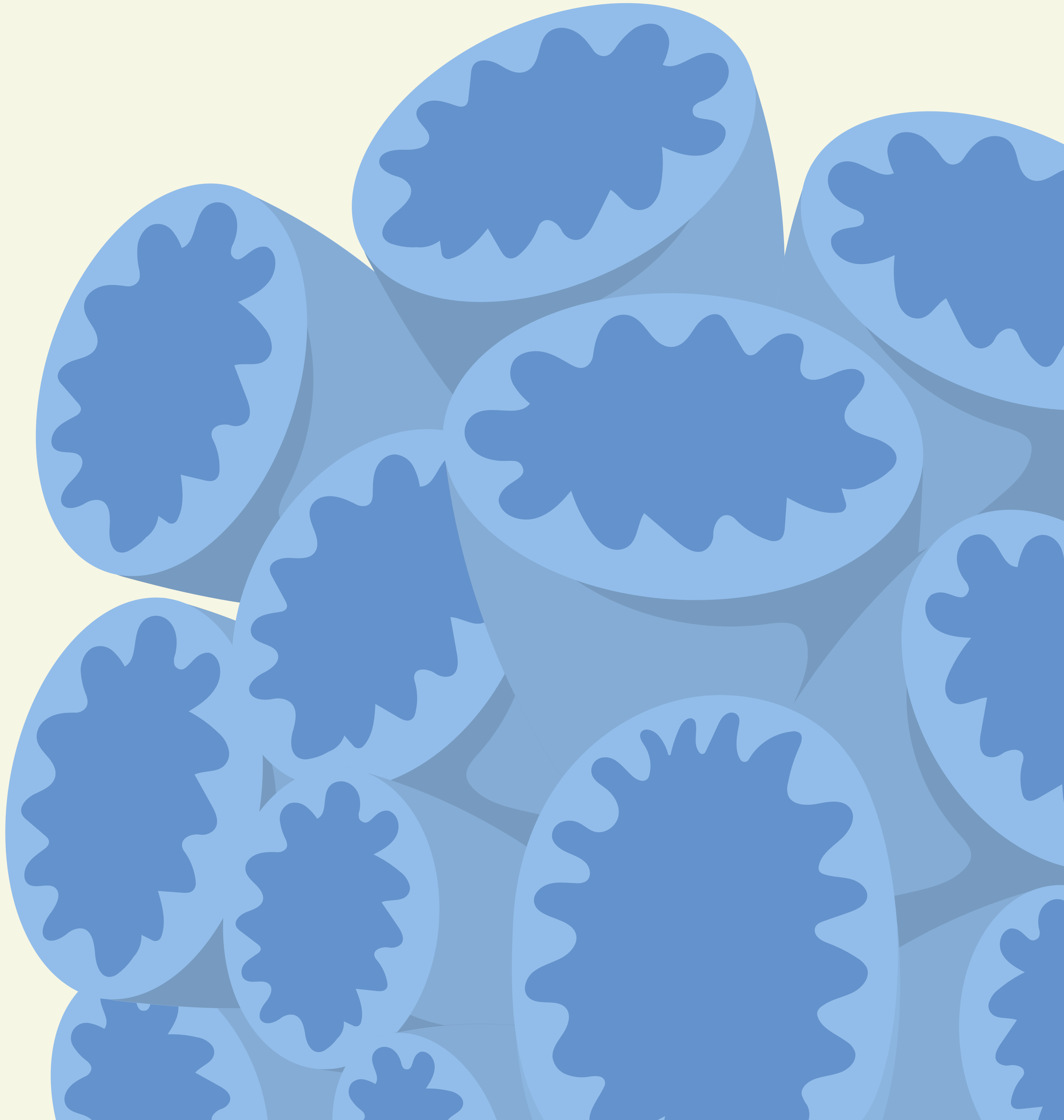
Coral reefs act as natural barriers against waves and storm surges. Their degradation leads to increased coastal erosion and flooding.

Decline in Fisheries and Food Security

Many commercial fish species depend on reefs, so coral loss impacts fisheries and global food supplies.

Economic Damage

Coral reefs generate billions in tourism and fishing revenue. Bleaching events hurt economies that rely on healthy reefs.

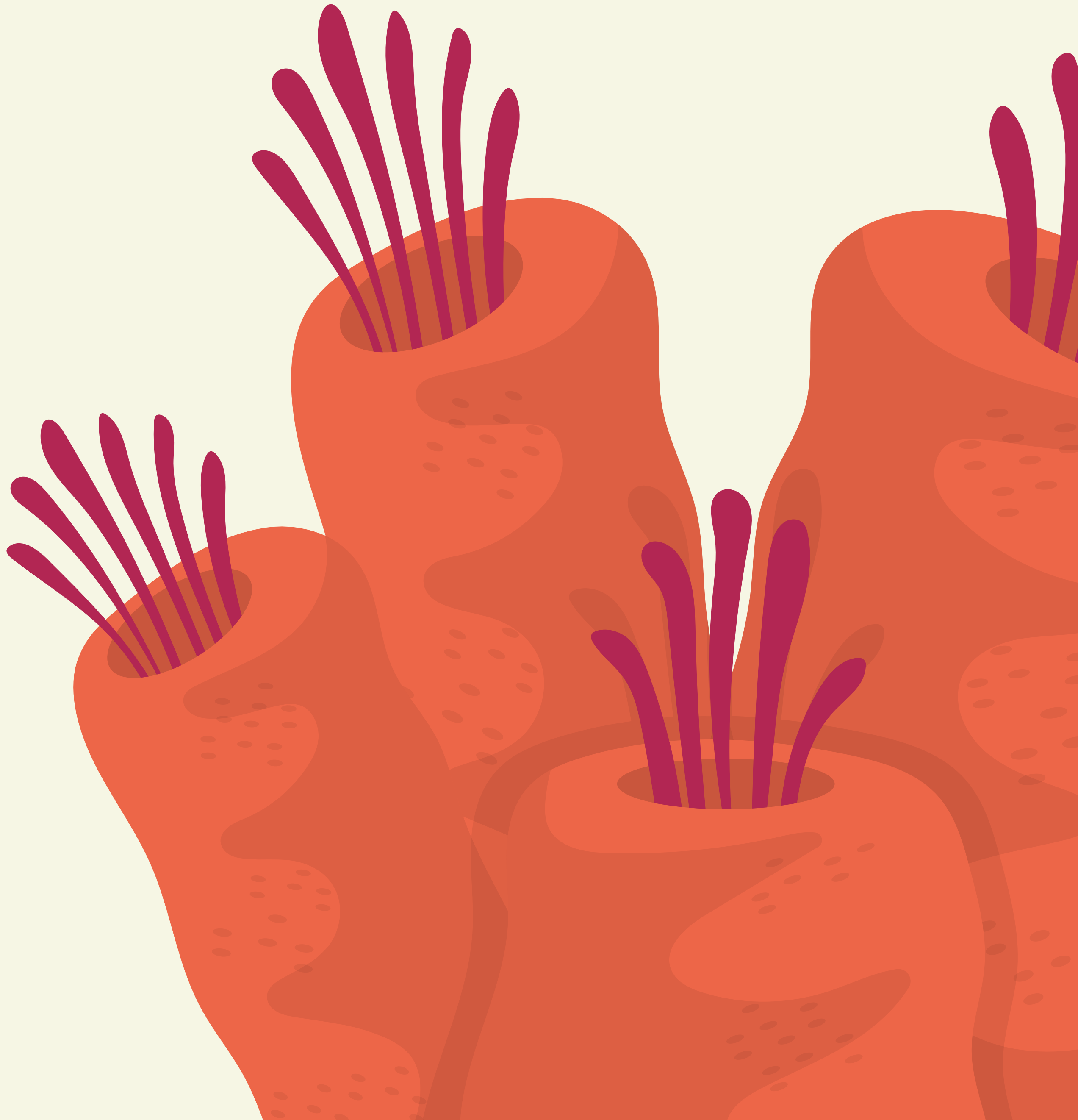


Coral Bleaching in India

The Gulf of Mannar alone hosts approximately **510** of India's 2,200 fish species, underscoring the rich marine life the reefs support.

India has 4 major reefs-
Andaman & Nicobar Islands
Lakshadweep Islands
Gulf of Mannar
Gulf of Kachchh

Reefs act as natural barriers against erosion, storms, and tsunamis, protecting **1.7 million** coastal residents.



literature review

publication	data explored	techniques	domain
https://www.researchgate.net/profile/Barbara-Brown-19/publication/226070355_Coral_bleaching_relative_to_elevated_seawater_temperature_in_the_Andaman_Sea_Indian_Ocean_over_the_last_50_years/links/	SST	Plot	Andaman and Nicobar islands
https://www.jstor.org/stable/27138240?seq=1	SST, DWH, MCB (NOAA OISST data)	Plot	Andaman, Nicobar, Lakshadweep, Gulf of Mannar and Gulf of Kachchh
https://www.sciencedirect.com/science/article/abs/pii/S0048965724031458	water parameters like pH, tCO2, fCO2, salinity, dissolved oxygen, etc	Plot	Indian Ocean
https://www.sciencedirect.com/science/article/abs/pii/S0272771424005018 (alt)	bleaching episodes, intensity, and species susceptibility	Plot	The Gulf of Kachchh, the Gulf of Mannar, Lakshadweep, and the Andaman Islands
https://www.researchgate.net/profile/Shiva-Kumar-21/publication/313404871_Coral_Bleaching_Along_Andaman_Coast_Due_to_Thermal_Stress_During_Summer_Months_of_2016_A_Geo	SST	Plot	Andaman coast
https://www.jstor.org/stable/24107307?seq=1	The thermal threshold for bleaching, degree heating month (DHM) accumulations of the sea surface temperature (SST)	Plot	Andaman, Nicobar, Lakshadweep, the Gulf of Man nar and the Gulf of Kachc
https://www.tandfonline.com/doi/abs/10.1080/01431161.2022.2161850	SST	Plot	Indo-Pacific Region
https://www.nature.com/articles/nature01987	SST	Plot	33 Indian Ocean sites
https://www.tandfonline.com/doi/abs/10.1080/10106049.2021.1886345	SST, Bleaching Threshold (BT), Positive Anomaly (PA), and Degree Heating Weeks (DHW), Dipole indices	Plot	Eastern Arabian Sea
https://eprints.cmfri.org.in/2158/	SST	Plot	Andaman Sea, Nicobar Sea, Lakshadweep Sea, Gulf of Mannar and Gulf of Kachchh
https://link.springer.com/article/10.1007/s10236-023-01562-y	sea surface temperature, salinity, in situ measurements	Plot	Lakshadweep Sea
https://link.springer.com/article/10.1007/s00338-006-0153-7	SST, past temperature anomalies, and coral community susceptibility	**regression formula**	Western Indian Ocean
https://link.springer.com/article/10.1007/s12524-021-01345-2	SST, Floating algae index (FAI), and turbidity mapping	**maximum likelihood classifier (supervised)**	Andaman and Nicobar Islands
https://www.researchgate.net/profile/Serena-Mccalla-2/publication/374340	Sea surface temperature, sea surface temperature anomalies, longitude, latitude, and coral depth below the surface	**k means, random forest + ica pca**	Global
https://www.sciencedirect.com/science/article/abs/pii/S2352485516301372	SST, Photosynthetically Active Radiation (PAR), species data	Plot	Palk Bay
https://www.sciencedirect.com/science/article/abs/pii/S0272771404003191	SST anomaly and Degree Heating Week (DHW)	Plot	Lakshadweep Sea
https://www.sciencedirect.com/science/article/pii/S2405844024142806	satellite images	svm	Red Sea
https://ieeexplore.ieee.org/document/9576973	satellite images	**Naive Bayes, Decision Tree, KNN, SVM, Random Forest and XGBoost**	mauritius
https://theacademic.in/wp-content/uploads/2025/01/83.pdf	underwater imagery	yolov8	india
https://www.ijeta.org/journals/ijdne/paper/10.18280/ijdne.170313	satellite based geospatial data	random forest, decision tree, mlp ann, radial basis function	thailand

Using machine learning to develop a global coral bleaching predictor

Methodology

The study identified that SST, SSTA, longitude, latitude, and depth were the features most correlated with coral bleaching events.

The final prediction system uses the random forest regressor to input ocean parameters and calculate both the percentage of coral colonies at risk for bleaching at a given location and categorize the risk level as high or low.

Accuracy

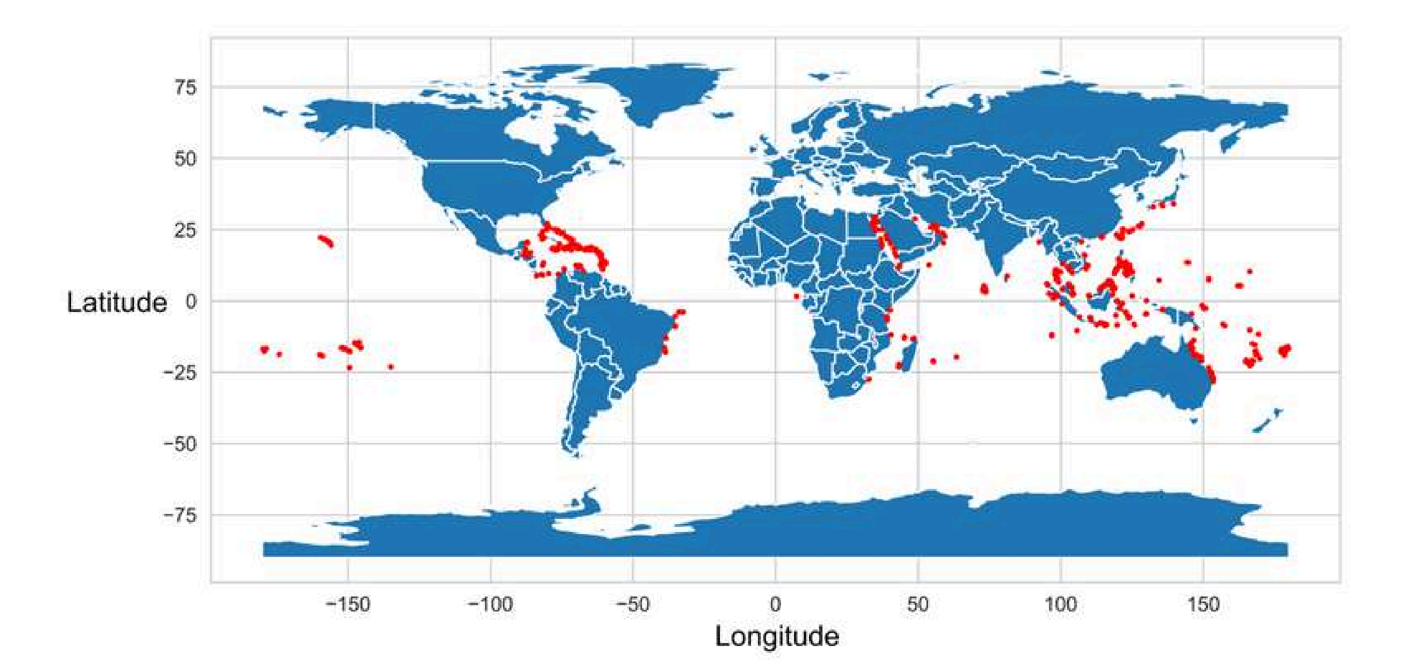
R-squared value of 0.25

Evaluation

The model is not as accurate of a classifier as it is trained on global data, and would benefit from being trained on local data to help better explain the effect of local parameters.

Journal of Emerging Investigators 2023
Serena McCalla, et. al

Model	R-Squared	RMSE	Time Taken (s)
RandomForestRegressor	0.25	7.91	2.08
ExtraTreesRegressor	0.24	7.96	0.75
HistGradientBoostingRegressor	0.22	8.06	1.13
GBMRegressor	0.22	8.07	0.09
XGBRegressor	0.19	8.20	0.34
LinearSVR	-0.08	9.52	0.08
PassiveAggressiveRegressor	-0.20	10.03	0.02
DecisionTreeRegressor	-0.21	10.07	0.05
ExtraTreeRegressor	-0.33	10.56	0.02
AdaBoostRegressor	-0.44	10.97	0.10



Global Map of Coral Reef Bleaching and non-Bleaching Events Included in Dataset

Spatio-temporal Mapping to Investigate Coral Bleaching in Andaman and Nicobar Islands, India Using Geoinformatics

Methodology

Satellite images to map coral reef patterns and bleaching extents between 2010 and 2019 in selected islands of the Andaman and Nicobar archipelago analysed using Maximum Likelihood Classification.

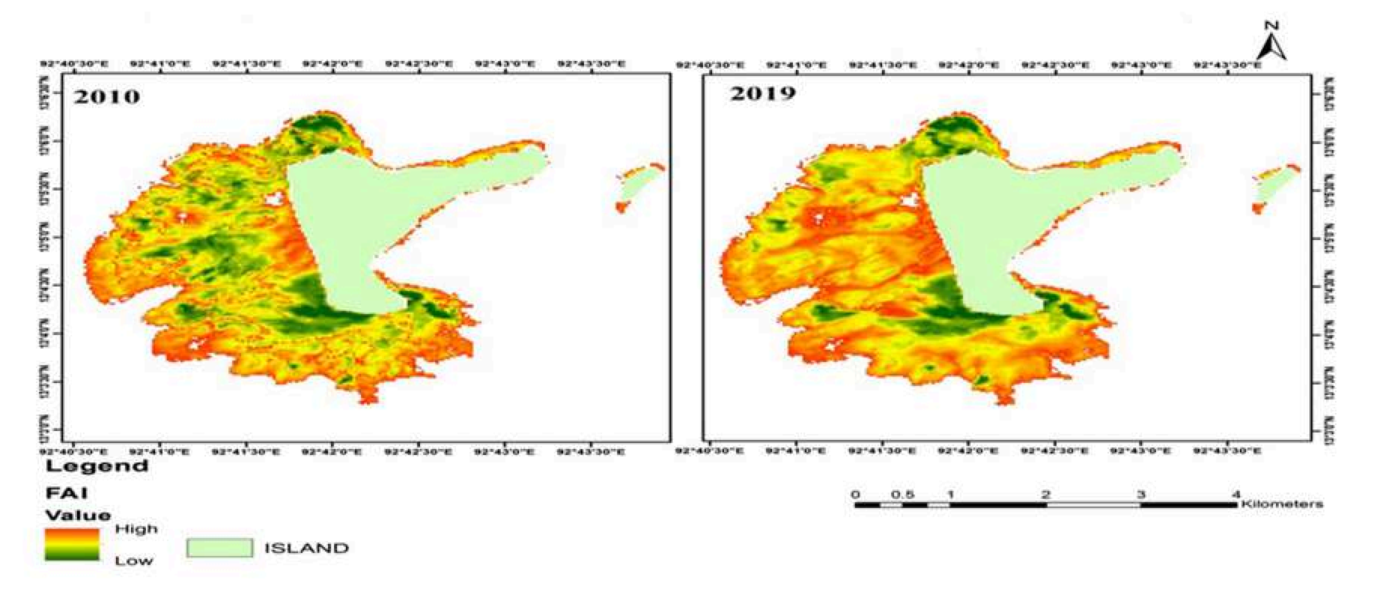
Accuracy

The classification accuracy of the coral maps was evaluated, yielding overall accuracies of 86% (2010) and 84% (2019) for North Reef Island.

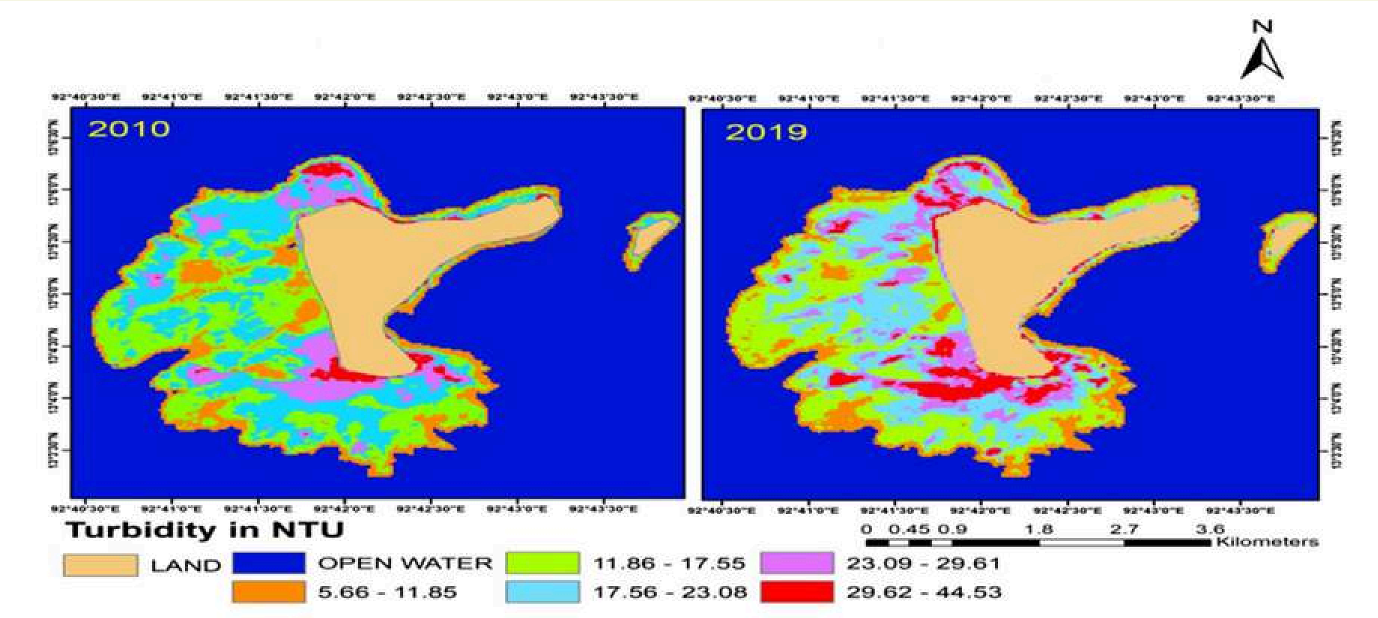
Evaluation

The data is specific to the Andaman and Nicobar islands. Furthermore, it only makes use of satellite images, making it dependent on a single data type/source.

Indian Society of Remote Sensing 2021
Arvind Chandra Pandey, et. al
10.1007/s12524-021-01345-2



Floating Algae Index map of North reef of India



Turbidity of North reef of India

literature review

Monitoring Coral Reefs Death Causes with Artificial Intelligence

Methodology

Innovative approach for reef monitoring is proposed based on Machine Learning methods like Naïve Bayes, Decision Tree, KNN, SVM, Random Forest and XGBoost to automatically classify corals into varying bleaching severities by training on past bleaching events.

Accuracy

XGBoost- 80.11% , Random Forest- 77.8%

Evaluation

This research paper focuses on Mauritius and does not account for a more generalised model. Furthermore, it cannot be applied to all areas as it requires underwater imagery along with other factors, which might require further infrastructure investment.

IEEE - IST-Africa Conference (IST-Africa) 2021
Nabeelah Pooloo, et. al
10.1109/ICECET52533.2021.9576973

GAP
under the sea

Existing models, those based on **sea surface temperature (SST)** provide some predictive ability, they typically depend on **single-variable** thresholds and lack the capacity to model complex, nonlinear relationships among environmental factors.

Only a few models incorporate parameters like **turbidity or salinity**, and even then, such models are often developed for **non-Indian reef systems**. These models are not designed for proactive management they indicate bleaching likelihood retrospectively or near onset, but rarely support early intervention.

Hence, there is a need for a multivariate, **India-specific predictive framework** that can offer more timely and ecologically meaningful insights.

model impact

Environmental Impact

Helps protect marine biodiversity by enabling timely intervention and reef restoration.

Scientific Advancements

Provides researchers with accurate and data-driven insights into coral health and climate change effects.

Economic Benefits

Preserves industries reliant on coral reefs, such as fisheries and ecotourism, by mitigating bleaching-related losses.

Global Conservation Efforts

Supports organizations like the United Nations, NOAA, and WWF in their marine conservation initiatives.

data collection

National Environmental Satellite, Data, and Information Service

Name:
Gulf of Kutch, India

Polygon Middle Longitude:
69.5750

Polygon Middle Latitude:
22.2250

Averaged Maximum Monthly Mean:
29.0468

Averaged Monthly Mean (Jan-Dec):
22.3061 22.4843 23.9364 26.2399 28.3899 29.0468 28.3354 27.4208 27.7638 27.8388 26.4472 24.1583

First Valid DHW Date:
1985 25 03

First Valid BAA Date:
1985 31 03

YYYY	MM	DD	SST_MIN	SST_MAX	SST@90th_HS	SSTA@90th_HS	90th_HS>0	DHW_from_90th_HS>1	BAA_7day_max
1985	01	01	23.9600	24.5600	24.4600	0.2116	0.0000	0.0000	0
1985	01	02	23.8000	24.4900	24.3200	0.2684	0.0000	0.0000	0
1985	01	03	23.6500	24.3600	24.2300	0.1594	0.0000	0.0000	0
1985	01	04	23.4800	24.2000	24.0600	0.0377	0.0000	0.0000	0
1985	01	05	21.9900	23.1100	22.7500	-1.0803	0.0000	0.0000	0
1985	01	06	21.6800	23.0200	22.8600	-0.9003	0.0000	0.0000	0
1985	01	07	21.3700	22.6900	22.4500	-1.5342	0.0000	0.0000	0
1985	01	08	20.8400	22.0400	21.7500	-2.0787	0.0000	0.0000	0
1985	01	09	20.4900	21.8500	21.5000	-2.1342	0.0000	0.0000	0
1985	01	10	19.8800	21.5500	21.2400	-2.3503	0.0000	0.0000	0
1985	01	11	20.0800	21.6900	21.4900	-2.2261	0.0000	0.0000	0
1985	01	12	20.2800	21.8900	21.6900	-1.9894	0.0000	0.0000	0
1985	01	13	20.3800	21.9900	21.7500	-1.7648	0.0000	0.0000	0
1985	01	14	20.4500	22.0800	21.8900	-1.7181	0.0000	0.0000	0

.txt file

National Centers for Environmental Information
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

.nc file

OceanSODA-ETHZ: A global gridded dataset of the surface ocean carbonate system for seasonal to decadal studies of ocean acidification (v2023) (NCEI Accession 0220059)

1. Time (Time)
2. Latitude (lat)
3. Longitude (lon)
4. Surface Partial Pressure of CO_2 (pCO2)
5. Standard Deviation of pCO_2 (pCO2std)
6. Total Alkalinity (TAlk)
7. Standard Deviation of Total Alkalinity (TAstd)
8. Dissolved Inorganic Carbon (DIC)
9. pH on Total Scale (pH)
10. Bicarbonate Concentration (HCO3)
11. Aragonite Saturation State (omegaAR)
12. Calcite Saturation State (omegaCA)
13. Sea Water Salinity (Practical Salinity) (salinity)
14. Sea Surface Foundation Temperature (temperature)

data collection

CORAL REEFS IN INDIA

K. Venkataraman
National Biodiversity Authority
Chennai - 600 041

research papers

History of recurrent short and long-term coral bleaching events in Indian coral reefs: a systematic review of contrasting bleaching patterns, lessons learned, and future directions

Thinesh Thangadurai[1,2*], Kalyan De[3*], Sobana Murugesan[1], Sivagurunathan P[4], Riana Peter[4], Ramasamy Pasiyappazham[5], Joseph Selvin[1], Polpass Arul Jose[6], Anthony Bellantuono[7]

BCO-DMO

.csv file

Dataset: Bleaching and environmental data for global coral reef sites from 1980-2020

data on the presence and absence of coral bleaching, allowing comparative analyses and the determination of geographical bleaching thresholds, together with site exposure, distance to land, mean turbidity, cyclone frequency, and a suite of sea-surface temperature metrics at the times of survey.

Jet Propulsion Laboratory
California Institute of Technology


```
{"items":[{"id":164444,"x":"1992.7978142","y":"-1.366068","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164447,"x":"1992.8169399","y":"-1.495621","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164450,"x":"1992.8360656","y":"-1.077333","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164453,"x":"1992.8551913","y":"-1.099429","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164456,"x":"1992.8743169","y":"-1.073329","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164459,"x":"1992.8934426","y":"-1.082175","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164462,"x":"1992.9125683","y":"-1.107599","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164465,"x":"1992.931694","y":"-0.997316","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164468,"x":"1992.9508197","y":"-1.024861","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164471,"x":"1992.9699454","y":"-0.798204","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164474,"x":"1992.989071","y":"-0.405218","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164477,"x":"1993.0082192","y":"-0.339384","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
{"id":164480,"x":"1993.0273973","y":"0.036703","error_margin":null,"display_on_rollover":1,"y_margin_max":null,"y_margin_min":null,"units":null,"category_titles":["indian_ocean_dipole"]},
```


.json

preprocessing

Data Collection & Cleaning:

1. SST: Downloaded, converted to DataFrame/CSV, and combined data from four reefs (58,702 rows).
2. Water Chemistry (Salinity, fCO_2, pCO_2, pH): Extracted **.nc** files using **netCDF**, filtered by reef coordinates, removed missing rows (21,156 rows).
3. Bleaching Instances: Processed Excel data for Indian reefs, classified bleaching intensity (binary), and merged (55 rows).
4. Genera: Extracted reef-specific genera from literature and exported to CSV.
5. IOD/ENSO: Processed JSON data with date conversion.

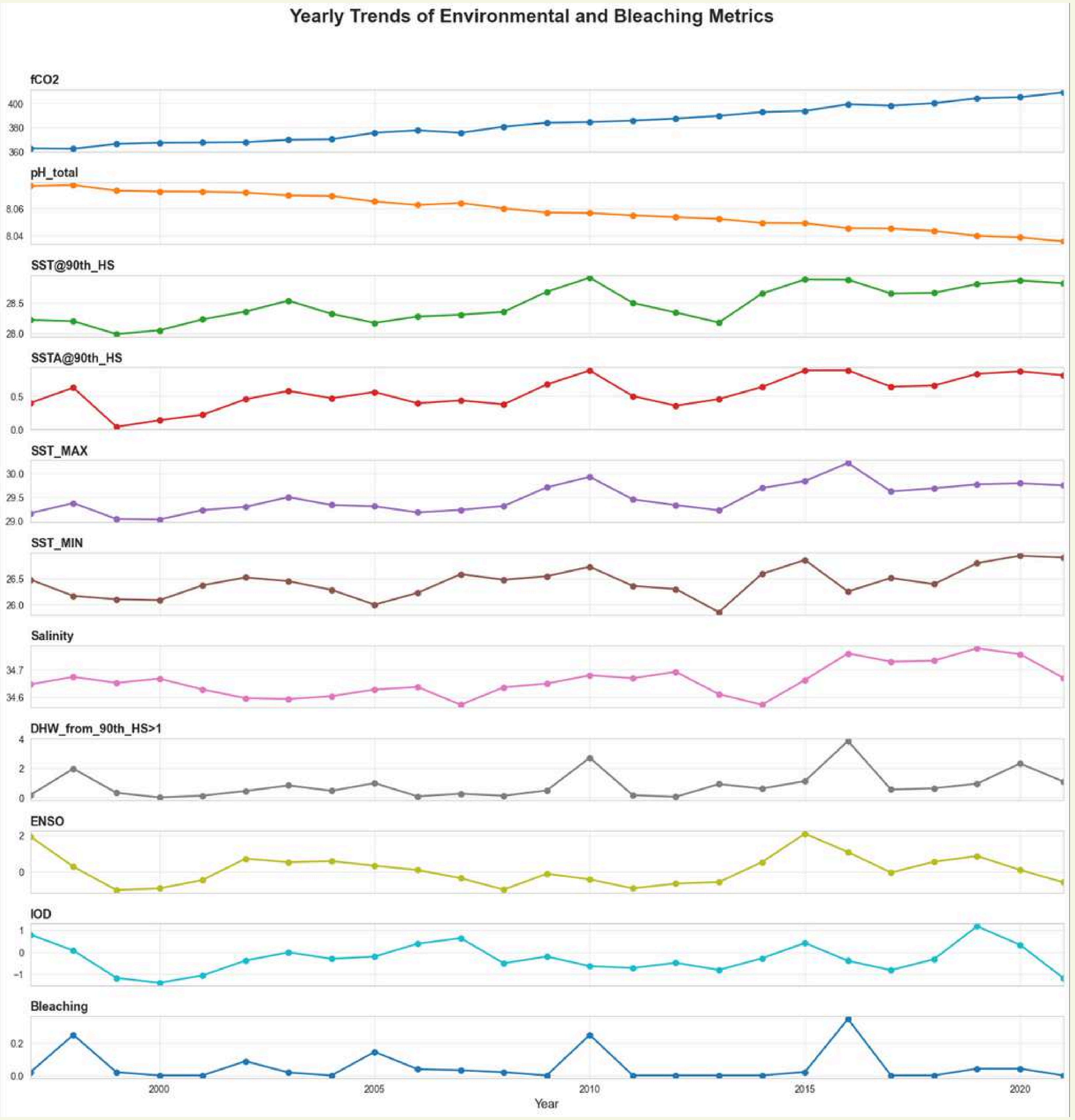
Data Integration & Aggregation:

1. Unified date formats and reef names across datasets.
2. Merged daily data into monthly (averaging salinity, CO_2, pH, IOD, ENSO; computed SST min, max, mean, and DHW).
3. Applied one-hot encoding on the genera; filtered data to 1995–2020 (final dataset: 1,201 rows).

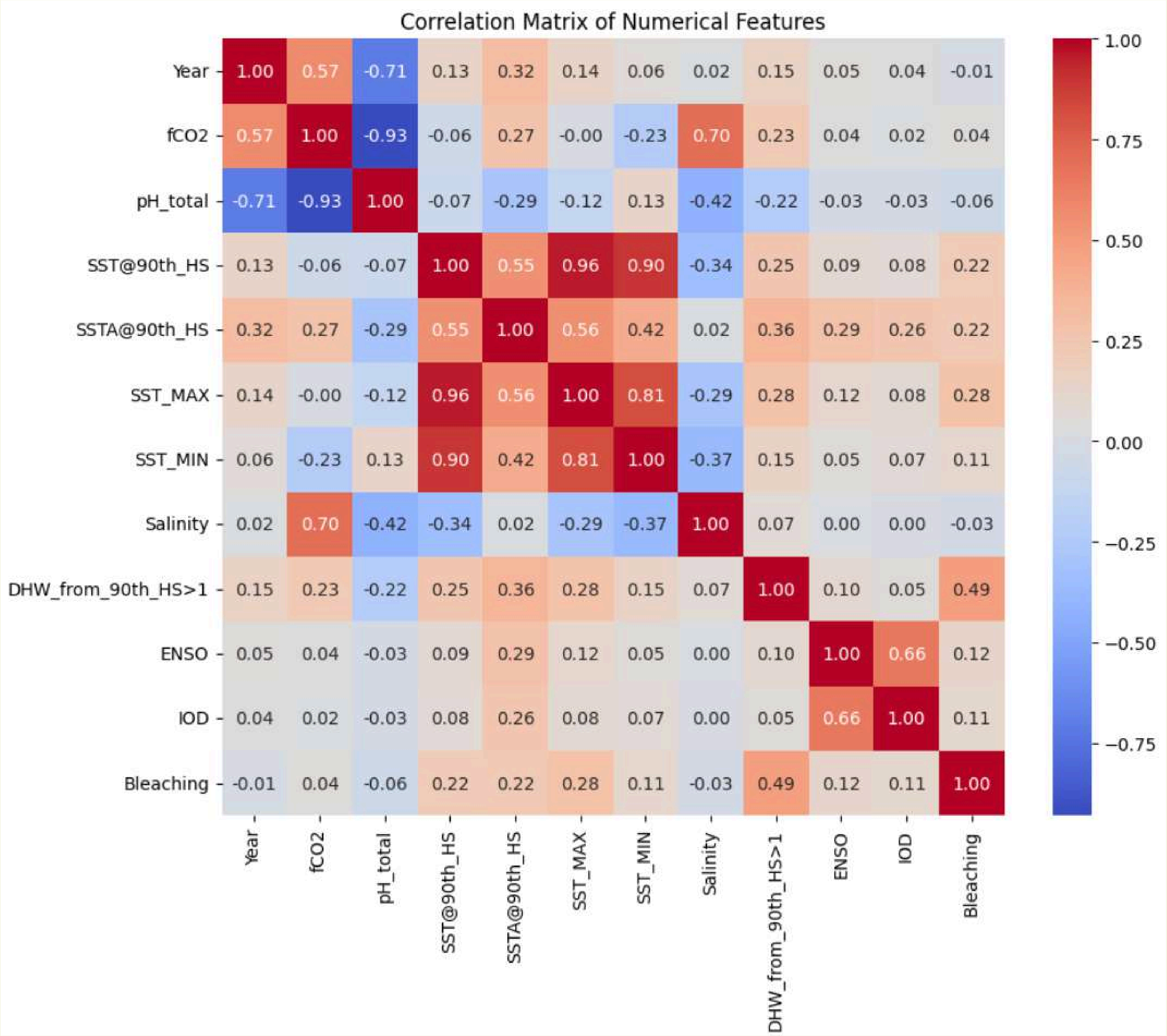
Final Work:

1. Distribution patterns for environmental variables.
2. Analyzed correlations and dropped unnecessary features
3. Visualisations to identify normal/skewed distributions

time series plot



correlation matrix



final data

All the four reefs are one-hot encoded and the dataset consists of monthly data (Month/Year) from 1997.

Ocean Temperature Metrics

- SST_MIN
- SST_MAX
- SST@90th_HS
- SSTA@90th_HS
- DHW_from_90th_HS>1

Ocean Chemistry Metrics

- Salinity
- pH_total
- fCO_2

Climate Indices

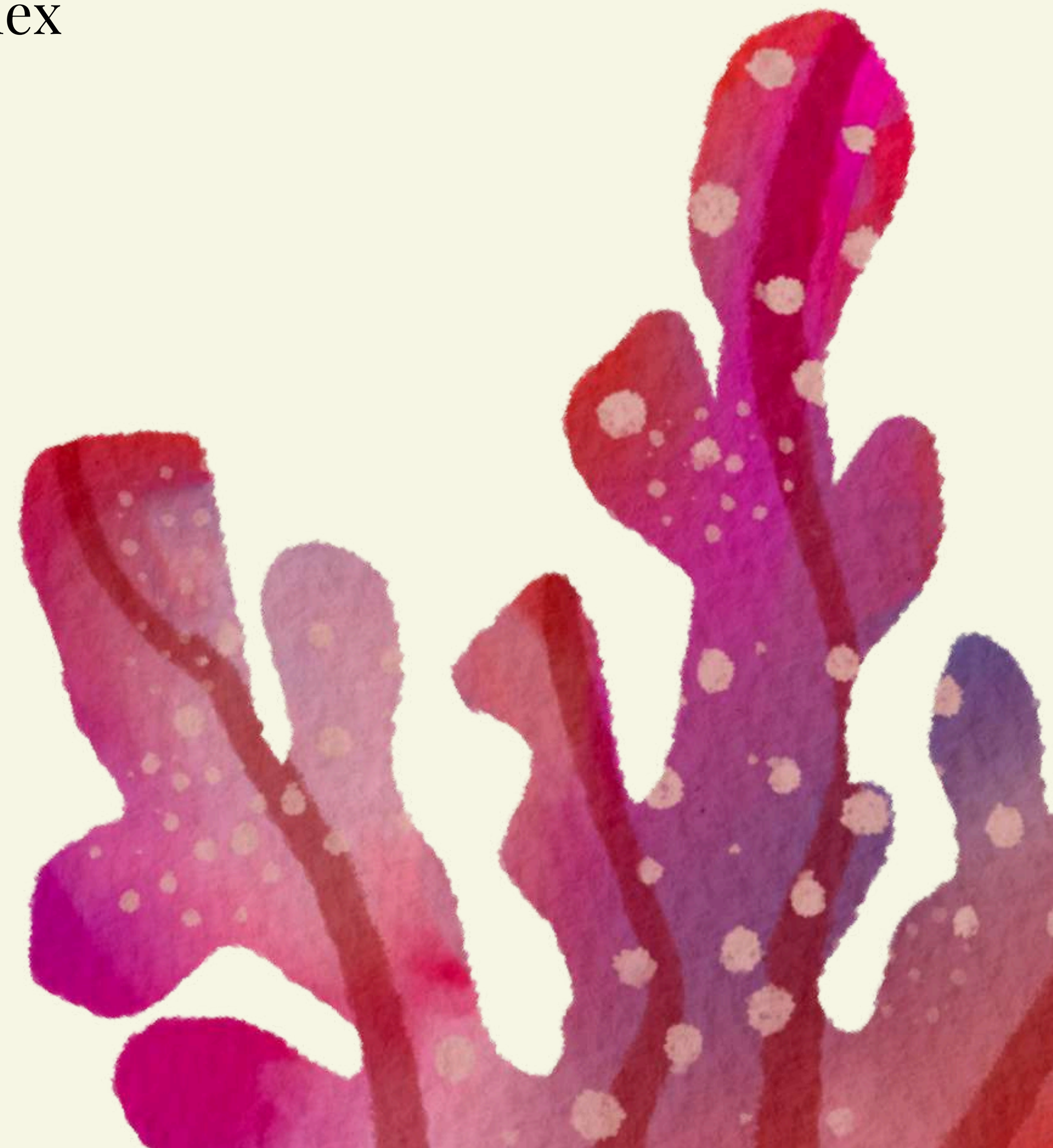
- IOD – Indian Ocean Dipole index
- ENSO – El Niño/Southern Oscillation index

Coral Genus (One-Hot)

- Acropora
- Porites
- etc. (55 columns)

Target

Bleaching (0/1))



```
Year,Month,Reef Name,SST_MIN,SST_MAX,SST@90th_HS,SSTA@90th_HS,DHW_from_90th_HS>1,Salinity,pH_total,fCO2,IOD,ENSO,Bleaching,Acanthastrea,Acropora,Alveopora,Astreopora,Canthare
1997,1,Andaman & Nicobar Islands,25.84,27.97,27.506129032258062,-0.1143064516129032,0.0,32.69885775,8.090811,337.889160625,-0.839704,0.2013039999999999,0.0,0,1,1,1,1,1,1,1,0,
1997,1,Gulf of Kachchh,19.55,24.16,23.115806451612904,-0.4777838709677419,0.0,36.41402,8.062024,390.11514,-0.839704,0.2013039999999999,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1
1997,1,Gulf of Mannar,25.8,27.86,27.261612903225807,-0.1535838709677419,0.0,33.580544,8.077757,342.99323,-0.839704,0.2013039999999999,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,
1997,1,Lakshadweep,27.29,28.72,28.25032258064516,0.1485258064516129,0.0,34.97936766666667,8.082114895833334,359.6364529166667,-0.839704,0.2013039999999999,0.0,1,1,1,1,0,0,0,1
1997,2,Andaman & Nicobar Islands,25.54,27.73,27.32821428571429,-0.4619785714285714,0.0,32.778389375,8.093888312499999,337.500289375,-1.13003525,0.156985,0.0,0,1,1,1,1,1,1,1,0
1997,2,Gulf of Kachchh,19.67,23.74,22.892142857142854,-0.8360535714285715,0.0,36.33193,8.065048,386.8081,-1.13003525,0.156985,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,0,0,
1997,2,Gulf of Mannar,25.9,28.36,27.517142857142854,-0.2229571428571428,0.0,33.70213,8.084003,342.52118,-1.13003525,0.156985,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0,1,0
1997,2,Lakshadweep,27.17,28.52,28.088214285714287,-0.2263785714285714,0.0,34.681197250000004,8.083565333333334,356.4286545833333,-1.13003525,0.156985,0.0,1,1,1,1,0,0,0,1,0,1,
1997,3,Andaman & Nicobar Islands,25.92,29.57,28.483870967741936,-0.279541935483871,0.0,32.8279540625,8.084170625,346.547554375,-0.9050376,0.630737,0.0,0,1,1,1,1,1,1,1,0,1,1,1
1997,3,Gulf of Kachchh,21.47,25.45,24.483548387096775,-0.1406483870967741,0.0,36.283707,8.073111,377.17474,-0.9050376,0.630737,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,0,0
1997,3,Gulf of Mannar,27.4,29.92,28.626774193548385,-0.2004225806451612,0.0,33.918987,8.080267,352.78262,-0.9050376,0.630737,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0,1,0
1997,3,Lakshadweep,27.59,29.97,29.288709677419355,0.2782548387096774,0.0,34.596644625,8.083373979166668,356.46080375,-0.9050376,0.630737,0.0,1,1,1,1,0,0,0,1,0,1,1,0,0,1,0,1,1
1997,4,Andaman & Nicobar Islands,27.71,30.54,29.812666666666665,0.10064,0.0,32.8745801875,8.074581625,355.393995625,0.1358415,1.2039485,0.0,0,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,
1997,4,Gulf of Kachchh,24.27,27.86,25.999666666666663,-0.45997,0.0,36.35479,8.060422,391.9828,0.1358415,1.2039485,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,0,0,0,1,1,0,1,0,
1997,4,Gulf of Mannar,28.79,30.99,29.949333333333335,0.2842766666666667,0.0,34.041138,8.076875,356.52686,0.1358415,1.2039485,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0,1,0
1997,4,Lakshadweep,28.78,31.0,30.082666666666668,0.4344766666666667,0.0,34.709765999999995,8.079503729166666,361.26291583333335,0.1358415,1.2039485,0.0,1,1,1,1,0,0,0,1,0,1,1,
1997,5,Andaman & Nicobar Islands,27.9,31.8,30.011935483870968,0.3670483870967742,1.5571,32.9470068125,8.0781916875,352.854215,-0.4433465,1.53191625,0.0,0,1,1,1,1,1,1,1,0,1,1,
1997,5,Gulf of Kachchh,26.2,29.16,28.219032258064516,-0.1361677419354839,0.0,36.437794,8.059653,392.81116,-0.4433465,1.53191625,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,0,
1997,5,Gulf of Mannar,28.09,30.79,30.030967741935484,0.4410903225806452,0.0,34.19941,8.072525,358.62152,-0.4433465,1.53191625,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0,1,
1997,5,Lakshadweep,28.98,31.39,30.581935483870968,0.7715,0.3271,34.955354166666666,8.080037062499999,362.29366625,-0.4433465,1.53191625,0.0,1,1,1,1,0,0,0,1,0,1,1,0,0,1,0,1,1,
1997,6,Andaman & Nicobar Islands,27.63,30.13,29.253,0.3585366666666666,1.5571,33.0219770625,8.07992928125,350.54298125,-0.0187886,1.7849214,0.0,0,1,1,1,1,1,1,1,0,1,1,1,1,1,1,
1997,6,Gulf of Kachchh,27.34,30.18,29.141333333333332,0.3375833333333333,0.0,36.539436,8.060304,392.09216,-0.0187886,1.7849214,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,0,0
1997,6,Gulf of Mannar,27.77,30.86,29.35233333333333,1.1470166666666666,0.0,34.43527,8.060301,374.51636,-0.0187886,1.7849214,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0,1,0,
1997,6,Lakshadweep,28.34,31.0,29.959,1.149773333333333,0.3271,35.27767712500004,8.087001083333334,356.5923966666667,-0.0187886,1.7849214,0.0,1,1,1,1,0,0,0,1,0,1,1,0,0,1,0,1,
1997,7,Andaman & Nicobar Islands,27.54,29.07,28.504193548387097,0.2709290322580645,1.5571,33.041463625,8.084083,345.312556875,0.9782475,1.99414025,0.0,0,1,1,1,1,1,1,1,0,1,1,1
1997,7,Gulf of Kachchh,27.37,30.77,29.37032258064516,1.0931967741935484,0.5257,36.554104,8.056115,396.7945,0.9782475,1.99414025,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,0,
1997,7,Gulf of Mannar,26.48,30.34,29.32,1.0255838709677418,0.0,34.611534,8.047803,393.6829,0.9782475,1.99414025,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0,1,0,1,1,0,1,0,1,
1997,7,Lakshadweep,26.57,29.67,28.176451612903225,0.2935387096774193,0.3271,35.46550225,8.085127541666667,359.63739875000005,0.9782475,1.99414025,0.0,1,1,1,1,0,0,0,1,0,1,1,0,
1997,8,Andaman & Nicobar Islands,27.38,28.84,28.37548387096774,0.3632935483870967,1.2243,32.971632,8.08903140625,340.464855,0.5722552000000001,2.4541846,0.0,0,1,1,1,1,1,1,1,0
1997,8,Gulf of Kachchh,26.79,29.36,28.45903225806452,0.764941935483871,0.5257,36.52838,8.059383,394.45654,0.5722552000000001,2.4541846,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1
1997,8,Gulf of Mannar,26.0,29.93,28.91870967741935,0.640125806451613,0.0,34.71382,8.064054,375.1237,0.5722552000000001,2.4541846,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0
1997,8,Lakshadweep,26.59,28.96,28.278387096774193,0.5627548387096774,0.0,35.5926067916667,8.086068562500001,359.50168041666666,0.5722552000000001,2.4541846,0.0,1,1,1,1,0,0,0
1997,9,Andaman & Nicobar Islands,27.85,29.88,28.907,0.6533,0.0,32.8155781875,8.09068546875,338.58028375,1.03531775,2.9133675,0.0,0,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1
1997,9,Gulf of Kachchh,28.02,29.73,28.909333333333333,1.2583366666666669,0.5257,36.487938,8.052565,400.66675,1.03531775,2.9133675,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,
1997,9,Gulf of Mannar,26.17,29.9,28.950333333333333,0.5712733333333334,0.0,34.660736,8.073006,366.18753,1.03531775,2.9133675,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0,1,0
1997,9,Lakshadweep,27.61,29.44,28.743,0.7705766666666667,0.0,35.611874875,8.084623395833333,360.8121775,1.03531775,2.9133675,0.0,1,1,1,1,0,0,0,1,0,1,1,0,0,1,0,1,1,0,1,1,1,1,0
1997,10,Andaman & Nicobar Islands,27.62,30.58,29.60709677419355,1.1942000000000002,0.0,32.5996989375,8.0902760625,338.108778125,2.27434975,3.18873325,0.0,0,1,1,1,1,1,1,1,0,1,
1997,10,Gulf of Kachchh,27.2,29.4,28.52548387096774,0.8408645161290322,0.5257,36.443306,8.055789,396.9301,2.27434975,3.18873325,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,0,
1997,10,Gulf of Mannar,27.68,30.0,28.764516129032256,0.7457677419354839,0.0,34.575027,8.07583,358.81647,2.27434975,3.18873325,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0,0,1,
1997,10,Lakshadweep,28.6,30.22,29.72483870967742,1.4050709677419355,0.0,35.587941083333334,8.084227979166666,361.3246583333334,2.27434975,3.18873325,0.0,1,1,1,1,0,0,0,1,0,1,1
1997,11,Andaman & Nicobar Islands,28.06,30.32,29.28966666666667,1.00231,0.0,32.53549475,8.087462437500001,341.428559375,3.3124997499999997,3.374354,0.0,0,1,1,1,1,1,1,1,0,1,1,
1997,11,Gulf of Kachchh,24.07,28.24,27.27,0.5408033333333333,0.0,36.446175,8.071245,379.8312,3.3124997499999997,3.374354,0.0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,1,0,0,0,0,1,1
1997,11,Gulf of Mannar,27.63,29.59,28.783333333333335,0.7021666666666667,0.0,34.205025,8.073125,352.2769,3.3124997499999997,3.374354,0.0,0,1,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1
1997,11,Lakshadweep,28.68,30.39,29.744000000000003,1.1724066666666668,0.0,35.467304500000004,8.0817890625,362.76972125,3.3124997499999997,3.374354,0.0,1,1,1,1,0,0,0,1,0,1,1,0
1997,12,Andaman & Nicobar Islands,26.84,29.51,28.73064516129032,0.8051290322580645,0.0,32.60341675,8.08451075,344.425406875,3.934350999999999,3.1939956,0.0,0,1,1,1,1,1,1,1,0,
```

(1309, 122)

ml methods

Random Forest

Random Forest models combine the predictions of multiple decision trees to reach a single result.

- RF models are often used as a baseline in ecological modeling.
- Used in coral studies including McClanahan et al. (2007) and Udomchaipitak et al. (2022).

LSTM

LSTM is a type of recurrent neural network (RNN) that excels at learning and retaining long-term dependencies in sequential data.

- LSTM captures cumulative environmental stress over time, reflecting the biological reality that bleaching results from prolonged exposure to stressors.
- LSTM excels at high recall for rare event forecasting as it learns patterns from historical sequences.
- The model is successfully used in hydrological and climate modeling (Kratzert et al., 2019; Tian et al., 2020).

XGBoost

XGBoost is an ensemble learning method that builds decision trees sequentially based on previous trees.

- Known for handling nonlinearity, class imbalance, and tabular environmental data efficiently.
- Used by Udomchaipitak et al. (2022) for coral bleaching on a global scale.

LightGBM

- Similar to XGBoost, but optimized for faster convergence and larger datasets.
- Has not yet been applied to Indian reef datasets.

eval metrics

imbalanced classes

~~Accuracy~~, Precision, **Recall**, F1 Score, **F2 Score**

Recall

- Coral bleaching is a rare but high-impact event, hence false negatives (missed bleaching) are more ecologically damaging than false positives (false alarms).
- Recall quantifies how well the model identifies actual bleaching events.

F2 Score

- F_1 balances precision and recall equally, but F_2 puts more emphasis on recall, making it better suited for ecological early-warning systems.

Powers, D. M. (2011). Evaluation: Precision, Recall, F-measure and ROC. *International Journal of Machine Learning Technology*

random forest

'n_estimators': 551, 'max_depth': 19, 'min_samples_leaf': 8, 'pos_weight': 17.25581183509492

	precision	recall	f1-score	support
0.0	0.990	0.891	0.938	229
1.0	0.265	0.818	0.400	11
accuracy			0.887	240
macro avg	0.627	0.855	0.669	240
weighted avg	0.957	0.887	0.913	240

F2 score = 0.909
F2 score = 0.577

- SMOTE was utilised to counter imbalanced classes.
- Hyperparameter tuning was done through Optuna and then Pareto-optimal trials were identified to pick the best model.
- The recall and precision plot also helped us to optimize the threshold.

```
Trial#5: Recall=0.651, Precision=0.528, Params={'n_estimators': 844, 'max_depth': 9, 'min_samples_leaf': 6, 'pos_weight': 2.43374
Trial#8: Recall=0.698, Precision=0.484, Params={'n_estimators': 147, 'max_depth': 35, 'min_samples_leaf': 6, 'pos_weight': 14.459
Trial#9: Recall=0.721, Precision=0.463, Params={'n_estimators': 551, 'max_depth': 19, 'min_samples_leaf': 8, 'pos_weight': 17.255
Trial#10: Recall=0.628, Precision=0.540, Params={'n_estimators': 978, 'max_depth': 50, 'min_samples_leaf': 4, 'pos_weight': 1.147
Trial#18: Recall=0.698, Precision=0.484, Params={'n_estimators': 854, 'max_depth': 9, 'min_samples_leaf': 7, 'pos_weight': 9.8903
Trial#21: Recall=0.581, Precision=0.595, Params={'n_estimators': 575, 'max_depth': 5, 'min_samples_leaf': 10, 'pos_weight': 1.213
Trial#28: Recall=0.698, Precision=0.484, Params={'n_estimators': 356, 'max_depth': 8, 'min_samples_leaf': 9, 'pos_weight': 8.0317
Trial#31: Recall=0.581, Precision=0.595, Params={'n_estimators': 592, 'max_depth': 5, 'min_samples_leaf': 10, 'pos_weight': 1.263
Trial#39: Recall=0.744, Precision=0.438, Params={'n_estimators': 227, 'max_depth': 16, 'min_samples_leaf': 2, 'pos_weight': 13.77
Trial#41: Recall=0.558, Precision=0.615, Params={'n_estimators': 987, 'max_depth': 12, 'min_samples_leaf': 2, 'pos_weight': 1.746
Trial#48: Recall=0.674, Precision=0.527, Params={'n_estimators': 406, 'max_depth': 6, 'min_samples_leaf': 1, 'pos_weight': 2.1614
Trial#49: Recall=0.605, Precision=0.591, Params={'n_estimators': 389, 'max_depth': 6, 'min_samples_leaf': 1, 'pos_weight': 1.0754
```

xgboost

‘subsample': 0.6, 'scale_pos_weight': 21.325581395348838, 'max_depth': 6, 'learning_rate': 0.01, 'colsample_bytree': 0.6, 'n_estimators': 300

	precision	recall	f1-score	support
0.0	0.990	0.878	0.931	229
1.0	0.243	0.818	0.375	11
accuracy			0.875	240
macro avg	0.617	0.848	0.653	240
weighted avg	0.956	0.875	0.905	240

F2 score = 0.898

F2 score = 0.555

- scale_pos_weight was utilised to counter imbalanced classes.
- Hyperparameter tuning was done using HalfingRandomizedSearchCV.
- Other methods were also utilised, such as the ROC plot.

lightgbm

	precision	recall	f1-score	support
0.0	0.98	0.98	0.98	230
1.0	0.60	0.60	0.60	10
accuracy			0.97	240
macro avg	0.79	0.79	0.79	240
weighted avg	0.97	0.97	0.97	240

F2 score = 0.98

F2 score = 0.60

- classweight = balanced and SMOTE was utilised to counter imbalanced classes.
- Hyperparameter tuning was done using RandomizedSearchCV, although the base case yielded superior results.

lstm

	precision	recall	f1-score	support
0	0.99	0.95	0.97	246
1	0.54	0.88	0.67	16
accuracy			0.95	262
macro avg	0.76	0.91	0.82	262
weighted avg	0.96	0.95	0.95	262

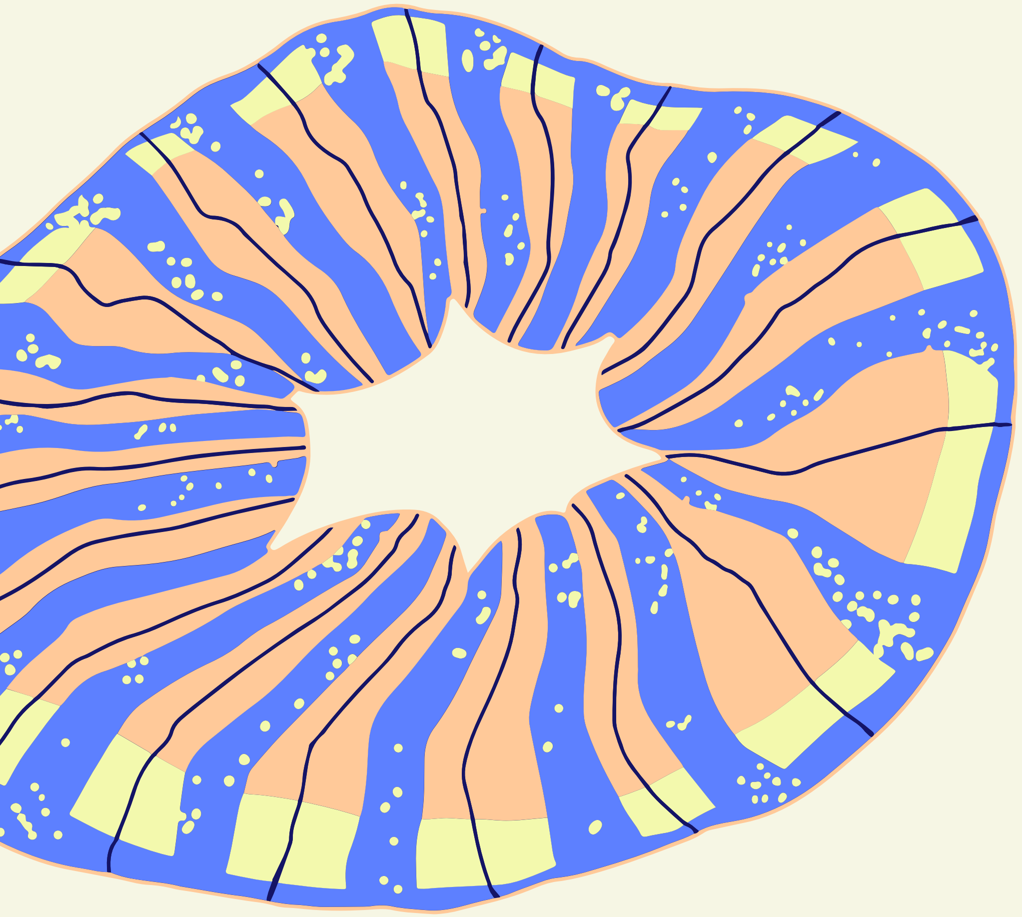
F2 score = 0.958

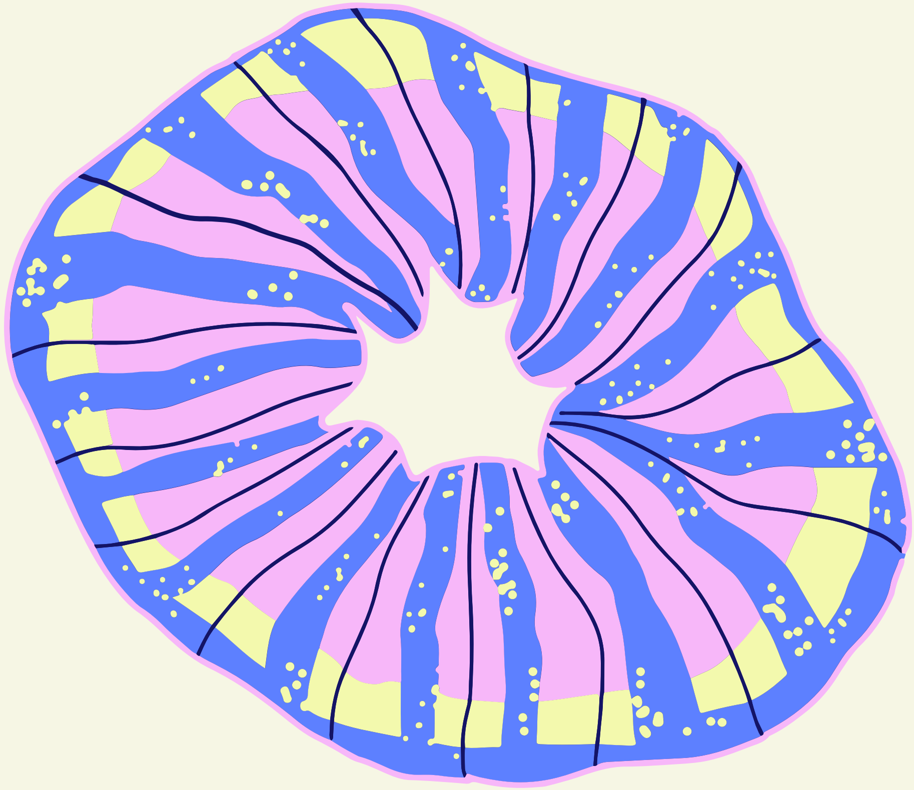
F2 score = 0.782

- The model was tuned by plotting the AUC curve and comparing the precision and recall based on difference thresholds.
- Early stopping was utilised to stop the training when then training and validation sets observed a gap.
- Another version of the model with class weights observed higher recall but at a lower F2 score, making the current model a better predictor.

```
name: lstm
trainable: True
dtype: {'module': 'keras', 'class_name': 'DTypePolicy', 'config': {'name': 'float32'}, 'registered_name': None}
return_sequences: False
return_state: False
go_backwards: False
stateful: False
unroll: False
zero_output_for_mask: False
units: 64
activation: tanh
recurrent_activation: sigmoid
use_bias: True
kernel_initializer: {'module': 'keras.initializers', 'class_name': 'GlorotUniform', 'config': {'seed': None}, 'registered_name': None}
recurrent_initializer: {'module': 'keras.initializers', 'class_name': 'Orthogonal', 'config': {'seed': None, 'gain': 1.0}, 'registered_name': None}
bias_initializer: {'module': 'keras.initializers', 'class_name': 'Zeros', 'config': {}, 'registered_name': None}
unit_forget_bias: True
kernel_regularizer: None
recurrent_regularizer: None
bias_regularizer: None
activity_regularizer: None
kernel_constraint: None
recurrent_constraint: None
bias_constraint: None
dropout: 0.0
recurrent_dropout: 0.0
```

challenges





Lack of an existing region specific dataset

Dataset curation using intensive research and a wide variety of sources, along with extensive preprocessing.

Imbalanced Data

- SMOTE
- scale_pos_weight
- Incorporation of Australian training data

Low Recall/F2 Scores

- Hyperparameter tuning
- Handling imbalanced data

Computational Time for Hyperparameter Tuning

- Regularization
- Virtual GPUs on Google Colab

deployment

Plaksha can deploy this coral bleaching prediction solution as a research project in the fields of marine biology and environmental monitoring.

This initiative can serve as a foundation for interdisciplinary projects and over time, Plaksha can collaborate with Indian **government agencies** to scale the model and integrate it into national reef monitoring programs, and contribute to data-driven marine conservation efforts in India.

scaling challenges

- Data Restrictions - Access to environmental data is limited.
- Deployment Platform - Needs shift from prototype (e.g., Colab) to scalable, secure systems.
- Funding & Maintenance - Continuous support needed for updates and upkeep.
- Policy Integration – Adoption by government depends on trust and validation.

references

[1] Arora, M., et al. (2019). Bleaching stress on Indian coral reef regions. Current Science. https://doi.org/10.18520/cs/v117/i2/242-250
[2] Brown, B. E. (1997). Coral bleaching: Causes and consequences. Coral Reefs, 16(1), S129–S138. https://doi.org/10.1007/s003380050249
[3] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953
[4] Glynn, P. W. (1993). Coral reef bleaching: Ecological perspectives. Coral Reefs, 12(1), 1–17. https://doi.org/10.1007/BF00303779
[5] Hoegh-Guldberg, O., et al. (2007). Coral reefs under rapid climate change and ocean acidification. Science, 318(5857), 1737–1742. https://doi.org/10.1126/science.1152509
[6] Hughes, T. P., Kerry, J. T., & Álvarez-Noriega, M., et al. (2017). Global warming and recurrent mass bleaching of corals. Nature, 543(7645), 373–377.
[7] Kratzert, F., Klotz, D., Herrnegger, M., Hochreiter, S., & Nearing, G. S. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. Hydrology and Earth System Sciences, 23(12), 5089–5110. https://doi.org/10.5194/hess-23-5089-2019
[8] Liu, G., Strong, A. E., & Skirving, W. (2006). Remote sensing of sea surface temperatures during coral bleaching. Eos, Transactions American Geophysical Union, 87(15), 137–144. https://doi.org/10.1029/2006EO150001
[9] McCalla, S., & Williams, B. (2023). Developing a machine learning model for predicting coral bleaching. Journal of Oceanographic Research, 8(2), 112–128.
[10] McClanahan, T. R., et al. (2007). Predictability of coral bleaching using temperature and coral stress history. Ecological Modelling, 212(2–3), 180–199. https://doi.org/10.1016/j.ecolmodel.2007.10.033
[11] McPhaden, M. J., Zebiak, S. E., & Glantz, M. H. (2006). ENSO as an integrating concept in earth science. Science, 314(5806), 1740–1745. https://doi.org/10.1126/science.1132588
[12] Ministry of Environment, Forest and Climate Change (MoEFCC). (2014). India's Fifth National Report to the Convention on Biological Diversity.
[13] Powers, D. M. (2011). Evaluation: Precision, Recall, F-measure and ROC. International Journal of Machine Learning Technology
[14] NOAA Coral Reef Conservation Program. (2023). What are Coral Reefs? https://coralreef.noaa.gov/
[15] NOAA Coral Reef Watch. (2022). Coral Bleaching Data and Products. https://coralreefwatch.noaa.gov/
[16] Patel, J., et al. (2024). Automated detection of coral bleaching using YOLOv8. The Academic, 2(12), 938–953.
[17] Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLOS ONE, 10(3), e0118432. https://doi.org/10.1371/journal.pone.0118432
[18] Thangadurai, T., et al. (2024). History of recurrent coral bleaching events in Indian coral reefs. bioRxiv. https://doi.org/10.1101/2024.05.18.594244
[19] Tian, Y., et al. (2020). Deep learning for predicting air quality in China using LSTM and 1D CNN. Scientific Reports, 10, 20988. https://doi.org/10.1038/s41598-020-77860-4
[20] Udomchaipitak, T., et al. (2022). Forecast coral bleaching by machine learning of remotely sensed geospatial data. International Journal of Design & Nature and Ecodynamics, 17(3), 423–431.

(inclusive of literature review)

thank you