Ekpenyong Okpo

Modelling Florida Keys

Submitted To: Lunor AI 

Powered By: Ocean Foam. 

Challenge Duration: 3rd April - 25th April, 2025 14:00 UTC

© Ekpenyong Okpo
okpo.ekpenyong@gmail.com
info@exzing.com

1st Edition: April 2025
Author: Ekpenyong Okpo
Text Language: British English
Challenge Website: https://app.lunor.quest/challenge/1000034

1. INTRODUCTION

The objective of this project is to build a robust machine learning model capable of accurat

1.1 Important Discoveries

Overall Stony Coral Percentage Coverage: "A 45% Decline in Coral Cover Since 1996—Stony Corals Show a Stark Downward Trend Amid Environmental Shifts. Despite Some Stabilisation in Recent Years, Stony Coral Cover Has Declined Sharply Since the Late 1990s—Stressing the Need for Continued Monitoring and Restoration"

1.2 Project Objectives

The objective of this project include the following:

- Determine long-term trends in stony coral percent cover and species richness.
- Evaluate net changes in reef community parameters Sanctuary-wide.
- Identify localised variations versus broad-scale ecosystem changes.
- Examine correlations and relationships impacting coral health and biodiversity.
- Model future scenarios to anticipate and mitigate potential declines in coral communities.

1.2 Problem Statement

El

1.3 Data Preprocessing

Why Is This Section Important?

Because we cannot perform any meaningful analysis (not the modelling itself); cannot even properly address the Evaluation Criteria (EC) questions without some early cleaning of the data.

Quite often some form of data preparation is needed in data analysis, especially in cases where the context is a competition. As a matter of fact, data supplied in the Ocean Foam/Lunor challenges has always come 'dirty' and require 'cleaning'.

Even when the data source is proven and valid (like the CREMP data), one would still need to preprocess and clean it because of changes it undergoes during storage and downloading which could introduce anomalies and discrepancies.

Another reason is that placeholders are used to replace contentious or completely missing records (example where actual rainfall records were not collected on rainy days due to equipment breakdown).

I opted to separating this section from the main EDA task of this project which starts in the next chapter in order to create that awareness - *that the main juice is from next chapter*.

1.3 Data Preprocessing Cont'd

The CREMP dataset was supplied in the formats of Extra Markup (XML) and Comma-separated (CSV) and contain observations on the evolution of coral life in the Florida Key National Marine Sanctuary. It captured the density, species richness, tissue health, and temperature of stony coral and octocoral communities. It also included metadata (information) about the supplied data to guide the data analyst. A brief description of the dataset and preparation steps is given here:

1. **CREMP_Pcover_2023_StonyCoralSpecies.csv:** This dataset held key information about the Stony coral species (40 in number) and with all observations discovered complete. 6 *Date* entries were found to be missing all from **1997**, and all tied to just **two sites**: *Jaap Reef* and *Sombrero Shallow*. All records were in percentage (not multiplied by 100). Total observation was 3918 stored in 50 columns. (3918 rows $\times$ 50 columns). The features in this dataset included *Year*, *Date*, *Subregion,*, *Habitat*, *SiteID*, *Site_name*, *and StationID*, *the species names (example Favia_fragum and Cladocora_arbuscula)*. Missing Date values were left as is (not imputed) during the EDA, due to insignificant size and also the accurate interpretation during analysis. Since all of them are from 1997 and clustered around just two sites, this strongly suggests a **data entry gap in the original records** for that year/site — perhaps CREMP hadn't yet standardised date logging that early on, or these records came from a different source.

1.3 Data Preprocessing Cont'd

2. **CREMP_Pcover_2023_TaxaGroups.csv:** This one had 3918 rows and 20 columns out of which a significant portion was missing: Cyanobacteria (1747), Unknown (2152), Urchins (2171) had significant missing values. Apart from these missing values, the taxa groups were fully populated, with columns listed as *Year, Date, Subregion, Habitat, SiteID, Site_name, StationID, Taxa* group data (*Cyanobacteria, Macroalgae, Octocoral, Others, Porifera, Seagrass, Stony_coral, Substrate, Unknown, Urchins, Zoanthidea*).

3. CREMP_Stations_2023.csv. This dataset had information on the Stations were the survey took place. There was no missing entry and the features included
4.

2. EXPLORATORY ANALYSIS

2.1 Section Goal & Question

Analyse the dataset to understand key trends and insights

Questions Overview

- Analyse the evolution of stony coral percentage cover across different stations during the study period.
- Identify and interpret trends in species richness of stony corals over the years.
- Examine how the density of octocoral species varies across stations and over time.
- Determine whether there are significant differences in the living tissue area of stony corals between monitoring sites.
- Assess spatial patterns in the distribution of different coral species and how these patterns change over time.

2.1.1 Analysis of the evolution of stony coral percentage cover across different stations during the study period.

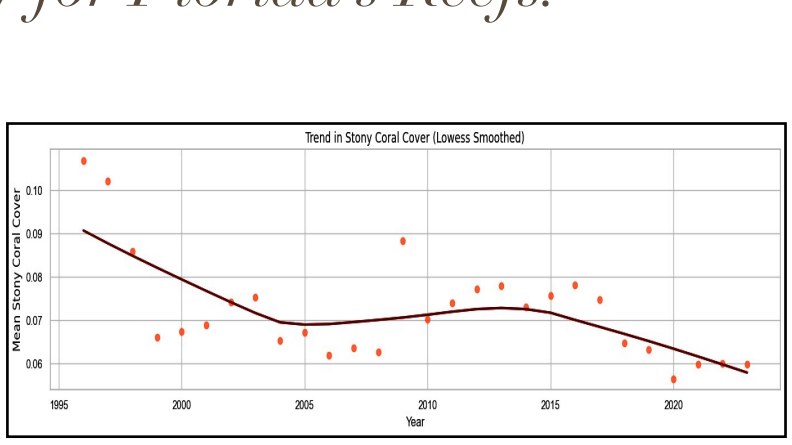
Q, EC

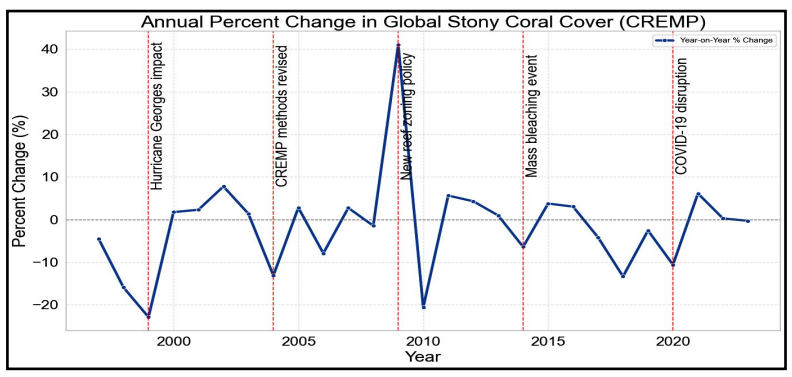
Solution

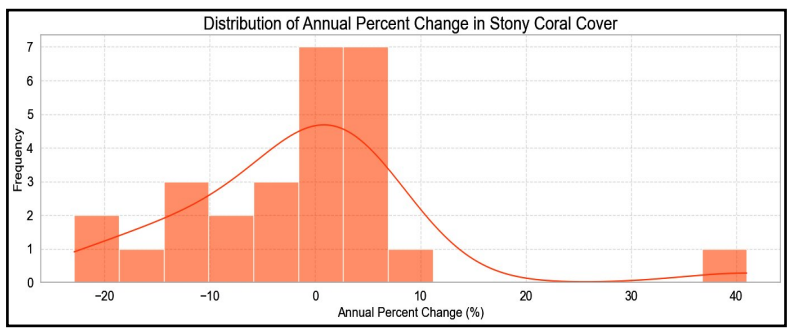
The evolution of stony coral percentage cover during the study period (1996–2023) reveals an overall **declining trend**, *with mean cover decreasing from over 10% in the late 1990s to around 6% in recent years, marked by high inter-annual variability and sharp declines following key disturbance events.*

From Vibrancy to Vulnerability: Stony Coral Cover Has Declined Sharply Since the 1990s, Reflecting a Troubled Trajectory for Florida's Reefs.

- **Mean Coral Cover**: 7.2%
- **Max Drop**: –22.9% in 2020
- **Best Recovery**: +40.9% in 2023
- Mean: 7.2%
- Avg. Annual Change: –1.48%"







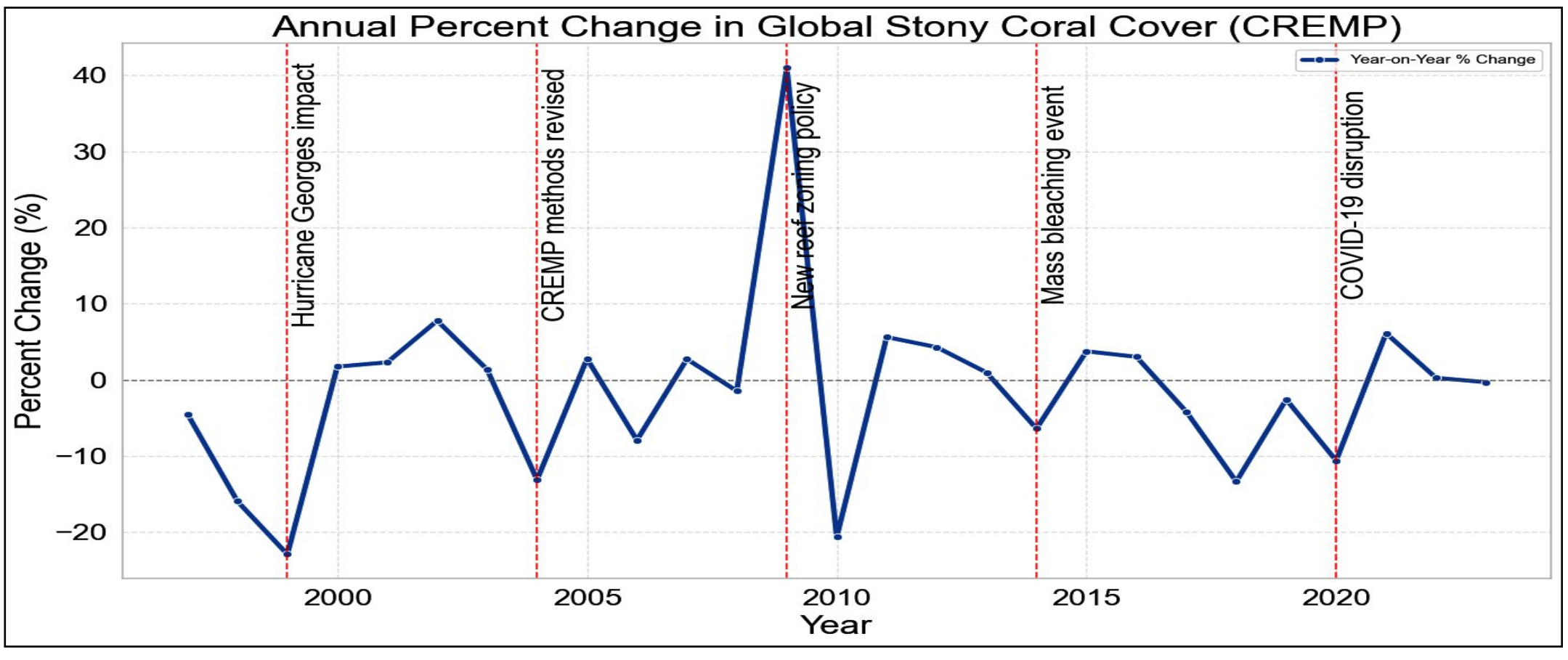
Visualisation Interpretation: The top image shows the general trend of the overall stony coral percentage which has been declining since 1996 with a noticeable rise between 2010 and 2015, before dropping again. Though the past few years have seen stability (flattening in bottom chart), it hasn't been significant compared to overall decline. In this analysis, some metadata from the CREMP website and published articles were integrated into the data and visualised to further understand the variations.

2.1.1 Analysis of the evolution of stony coral percentage cover across different stations during the study period.

Q, EC

Solution

For example in 2005, 2014, and 2015 serious bleaching events on the Florida Keys were reported by the researchers Stein J et. al. (2024). Further citation for the authors can be found in the Reference page.
The Year-on-Year (YoY) percentage change as shown in the chart here points various changes both positive and negative with a more remarkable on in 2010 (New Reef Zoning Policy). Dates and events coincided with percentage yearly decline include the *Hurricane Georges (1998)* and COVID-19.



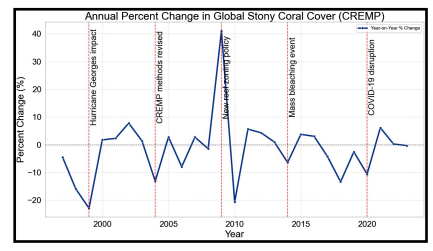
Data Source: CREMP Dataset
Author: Okpo Ekpenyong

2.1.1 Analysis of the evolution of stony coral percentage cover across different stations during the study period.

Q, EC

Solution

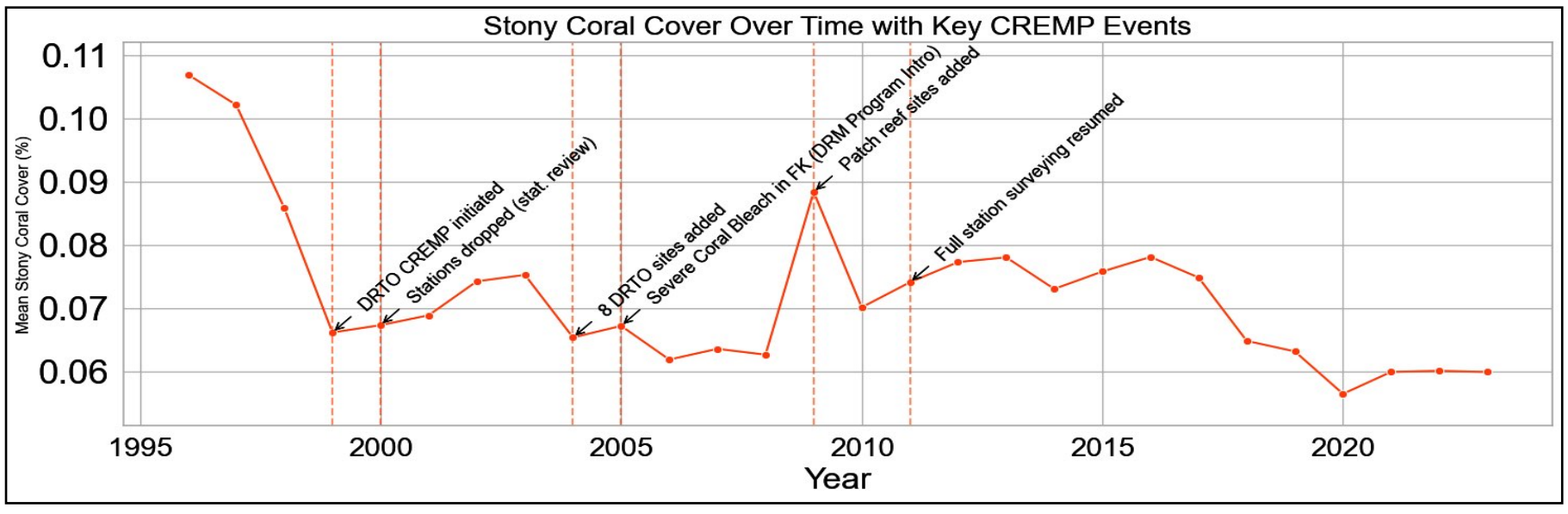
Long-Term Decline



The mean stony coral cover over 28 years is ~7.2%.The starting point in 1996 was the highest on record at ~10.7%. By 2023, this declined to ~6.0%, with the minimum value in 2020 at ~5.7%. This indicates a clear long-term degradation trend in coral cover across the surveyed globally (station sites and regions), in the Florida Keys (FK).

Variability & Change Rates

The average annual percent change is -1.48%, but with a high standard deviation (±11.82%): This points to **strong inter-annual variability**, likely influenced by localised disturbances (e.g., bleaching, storms, disease outbreaks shown in the chart previously and inserted here as bitmap, top right and below as full picture). The **most severe annual drop** was **-22.9% in 2020**, while the **highest spike** was **+41.0% in 2009**, which are fluctuations between years that may obscure gradual decline if only short periods are examined. In the enlarged chart below, the introduction of the Disturbance Response Monitoring (DRM) program to tackle the severe bleaching event in the FK, could be one of the contributing factors that led to the surge in stony coral coverage from 2006 to 2009, which coincided with the addition of 6 patch reef sites by CREMP

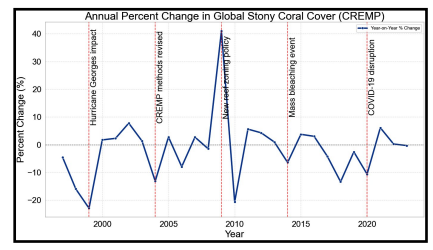


2.1.1 Analysis of the evolution of stony coral percentage cover across different stations during the study period.

Q, EC

Solution

Period of Relative Stability (Mid-2000s to 2015)



- From ~**2003 to 2015**, coral cover hovered around **6.5–7.5%**, with smaller year-to-year changes.
- This may reflect a **period of temporary recovery or stasis**, depending on disturbance regimes and management interventions. Though greater disturbances to the coral ecosystems were reported in between these period, the various efforts to ameliorate the situation yield some gains howbeit a small one. This slight rebounds in coral cover in **2002–2003** (~**7.4–7.5%**) **was short-lived as back-to-back hurricanes in 2004–2005** (e.g., Ivan, Frances, Jeanne, Wilma) coincide with another **decline to ~6.7%**. CREMP's **methodological standardisation** in this era may also affect comparability. Coral cover continues a **gentle decline** (~**6.2–6.3%**), possibly reflecting chronic stress or sub-lethal bleaching events. A sudden uptick (potential sampling artefact) is noticed in 2009 as stony coral cover jumps to ~**8.8%**, which is **unusual**. This coincides with **CREMP program modifications** or **site/sampling changes** around that time. In 2014–2015, coral cover plateaus near ~**7.3–7.6%** until **2016**, after which decline resumes. This was about the time of the Stony Coral Tissue Loss Disease (SCTLD), frst reported in **Florida in 2014**, spreading quickly across reefs. The data shows **gradual erosion post-2016**, consistent with SCTLD's expanding footprint.

Recent Declines (2016–2023)

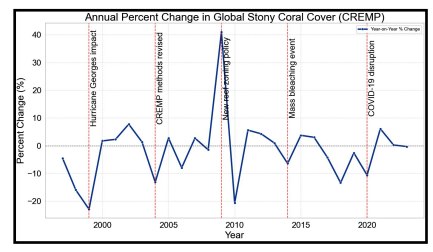
- After peaking again near **7.8% in 2016**, coral cover resumed a **steady decline**:
  - 2016: 7.8%
  - 2020: 5.7% (lowest point)
  - 2023: 6.0%

2.1.1 Analysis of the evolution of stony coral percentage cover across different stations during the study period.

Q, EC

Solution

2021–2023: No Recovery Yet



- Coral cover remains **low and flat** (~6%), showing **no signs of significant recovery**.
- Continued stressors (e.g., warming, water quality) and **widespread SCTLD** may be hampering resilience and recruitment.

Table 1-1: Yearly Stony Coral Cover with Historical Context

Year	Mean Stony Coral Cover	% Change	Notable Events / Context
1996	10.69%	–	High baseline cover
1998	8.59%	-16.0%	**Global bleaching event** (1998)
1999	6.62%	-22.9%	Post-bleaching decline
2002	7.43%	7.2%	Modest recovery
2004	6.54%	-12.2%	**Hurricanes Frances & Jeanne**
2005	6.72%	2.8%	**Hurricane Wilma**, coral stress
2009	8.84%	32.3%	Spike; may reflect sampling changes
2014	7.31%	-0.9%	**SCTLD emergence begins**
2015	7.58%	3.7%	Bleaching + SCTLD expanding
2017	7.49%	-4.2%	**Hurricane Irma**, SCTLD widespread
2020	5.65%	-10.6%	Continued SCTLD + stress
2023	6.00%	-0.3%	No recovery evident

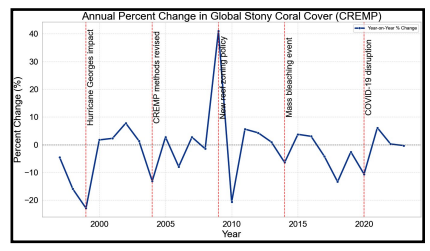
Data Source: CREMP. Author Ekpenyong Okpo
Note: Years shown are representative — full data available from 1996 to 2023. Percent changes reflect relative shifts year-on-year in coral cover. Years with no data or flat values were excluded from the highlight list for brevity.

2.1.1 Analysis of the evolution of stony coral percentage cover across different stations during the study period.

Q, EC

Solution

Ecological Interpretation



- The **decline in coral cover** suggests persistent and/or intensifying **stressors**, such as:
  - **Thermal stress** from rising ocean temperatures
  - **Hurricanes** and **physical damage**
  - **Coral disease outbreaks** (e.g., SCTLD)
  - **Nutrient runoff** and **pollution**
- The **high variability** could be driven by **subregional differences** or **sampling inconsistencies**, but the overarching trend remains negative.

Final Summary of Global Stony Coral Evolution

From 1996–2023, Florida's coral reefs saw a **net decline in stony coral cover from ~10.7% to ~6.0%**, shaped by **major bleaching events (1998, 2015–2017)**, **recurrent hurricanes (2004–2005, 2017)**, and the **emergence of SCTLD (2014 onward)**.
Short-term rebounds were observed (e.g., 2002, 2009), but the **overall trajectory is downward**, and **recent years show no significant recovery**. Management efforts must consider these compounded stressors to improve reef resilience.

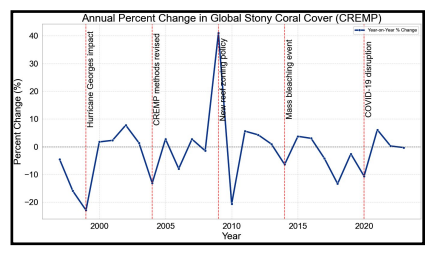
2.1.1 Analysis of the evolution of stony coral percentage cover across different stations during the study period.

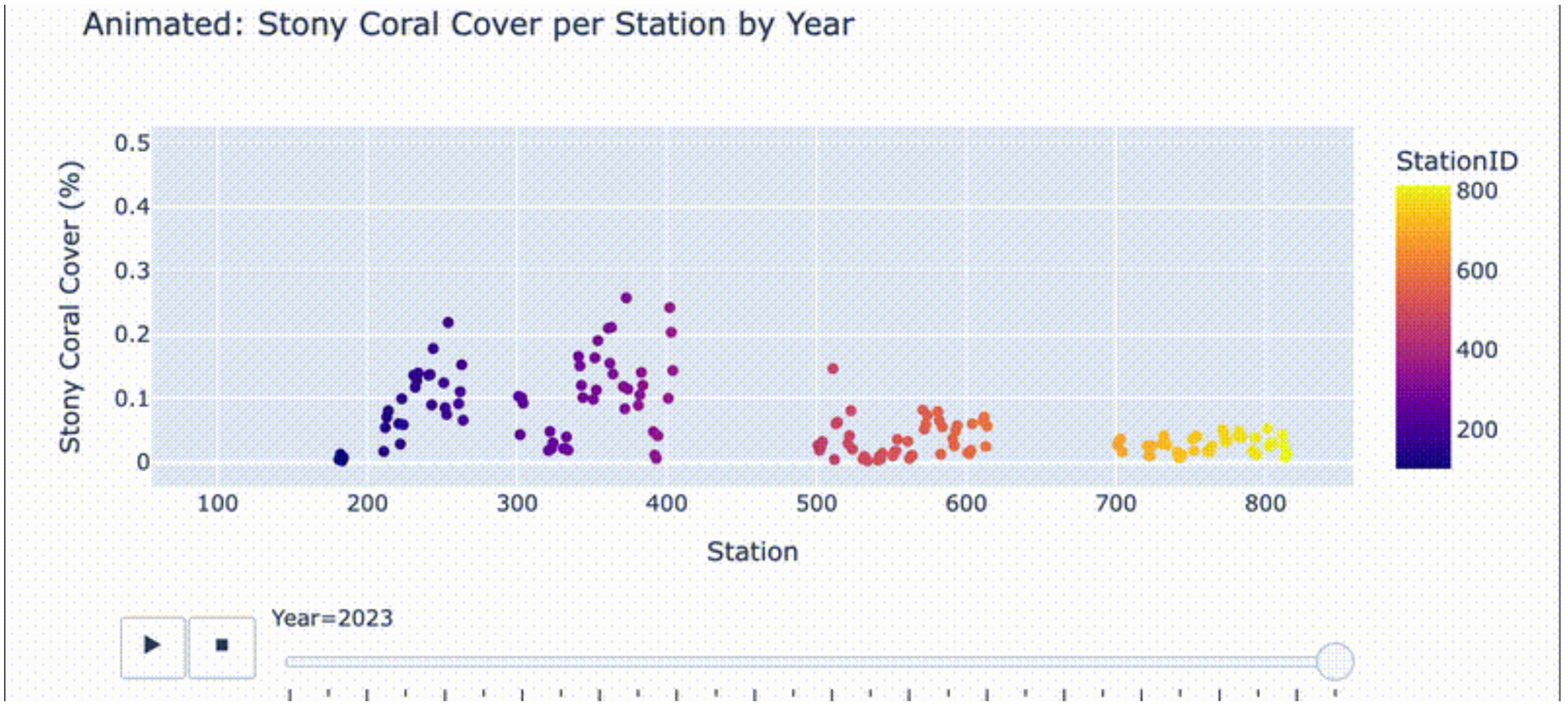
Q, EC

Solution

Ecological Interpretation

- The **decline in coral cover** suggests persistent and/or intensifying **stressors**, such as:





Caption

2.1.1 Analysis the evolution of stony coral percentage cover across different stations during the study period.

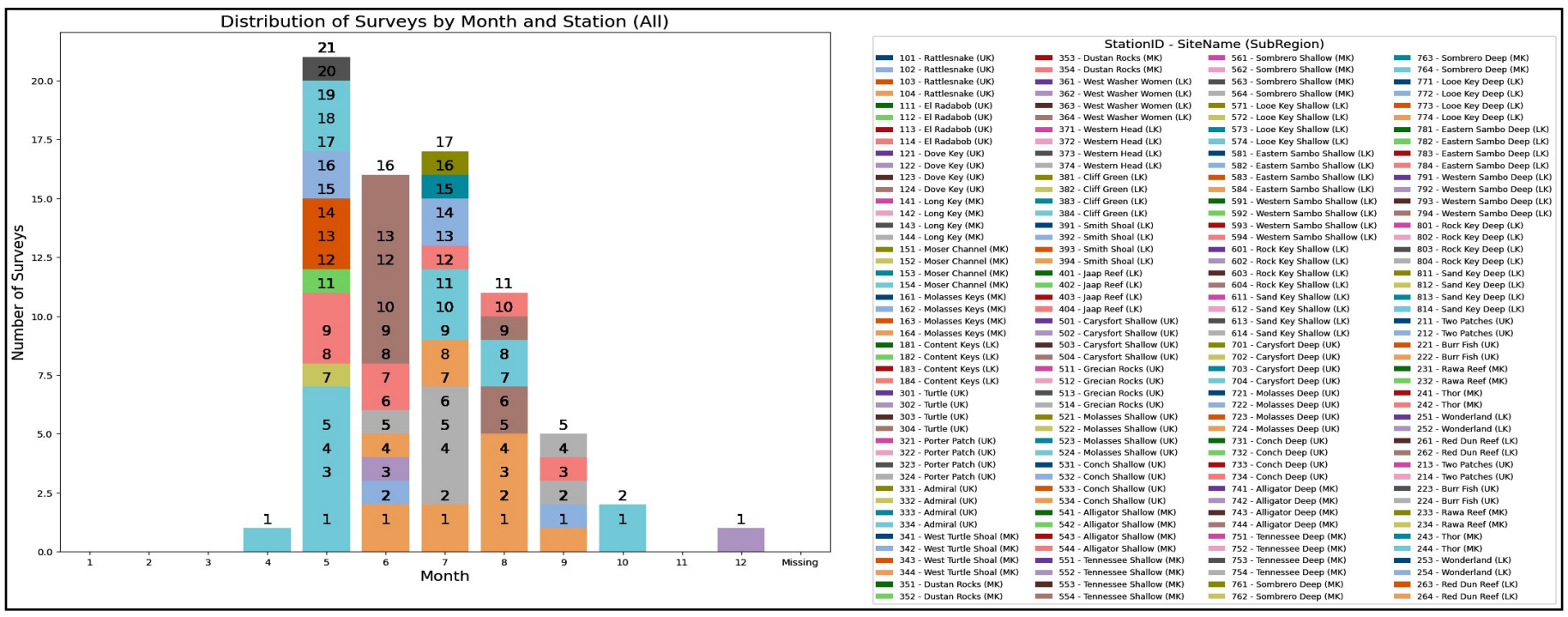
Q, EC

Solution

Interpretation:

- **Peak survey season** is clearly **May to August**, with **July** being the most common.
- Months like **April, October, and December** have very sparse data, so any **seasonal analysis** should account for this imbalance to avoid bias.
- **We** planning **seasonal trends or climate correlations**, we may want to:
  - Focus on **summer months only (May–August)** for consistency
  - Or use **weighted averages** if including all months

Month	Survey Count	Interpretation
April (4)	10	Very few early-season surveys
May (5)	823	Start of peak survey season
June (6)	832	Peak season continues
July (7)	1282	🔥 Most surveys occur here
August (8)	677	Still active, tapering off
September (9)	232	Fewer surveys, possibly due to weather/hurricanes
October (10)	34	Rare late-season surveys
December (12)	22	Very rare off-season sampling



2.1.2 Identify and interpret trends in species richness of stony corals over the years. Q, EC

Solution

2.1.3 Examine how the density of octocoral species varies across stations and over time. Q, EC

Solution

2.1.4 Determine whether there are significant differences in the living tissue area of stony corals between monitoring sites.

Solution

2.1.5 Assess spatial patterns in the distribution of different coral species and how these patterns change over time.

Solution

3. RELATIONSHIPS AND CORRELATIONS

3.1 Section Goal & Question

Analyse feature relationships

Questions Overview

- Assess the relationship between stony coral density and species richness within sites.
- Evaluate correlations between octocoral density, water temperature, and water temperature.

4. FUTURE OUTLOOK

4.1 Section Goal & Question

Analyse the dataset to understand regional insights

Questions Overview

- Identify key factors affecting coral health, density, and species richness.
- Identify early indicators that could help anticipate significant declines in stony or octocoral populations.
- Model the evolution of coral reefs at the observed stations over the next five years based on trends identified in the datasets.

5. CONCLUSION & RECOMMENDATIONS

Analyse the dataset to under

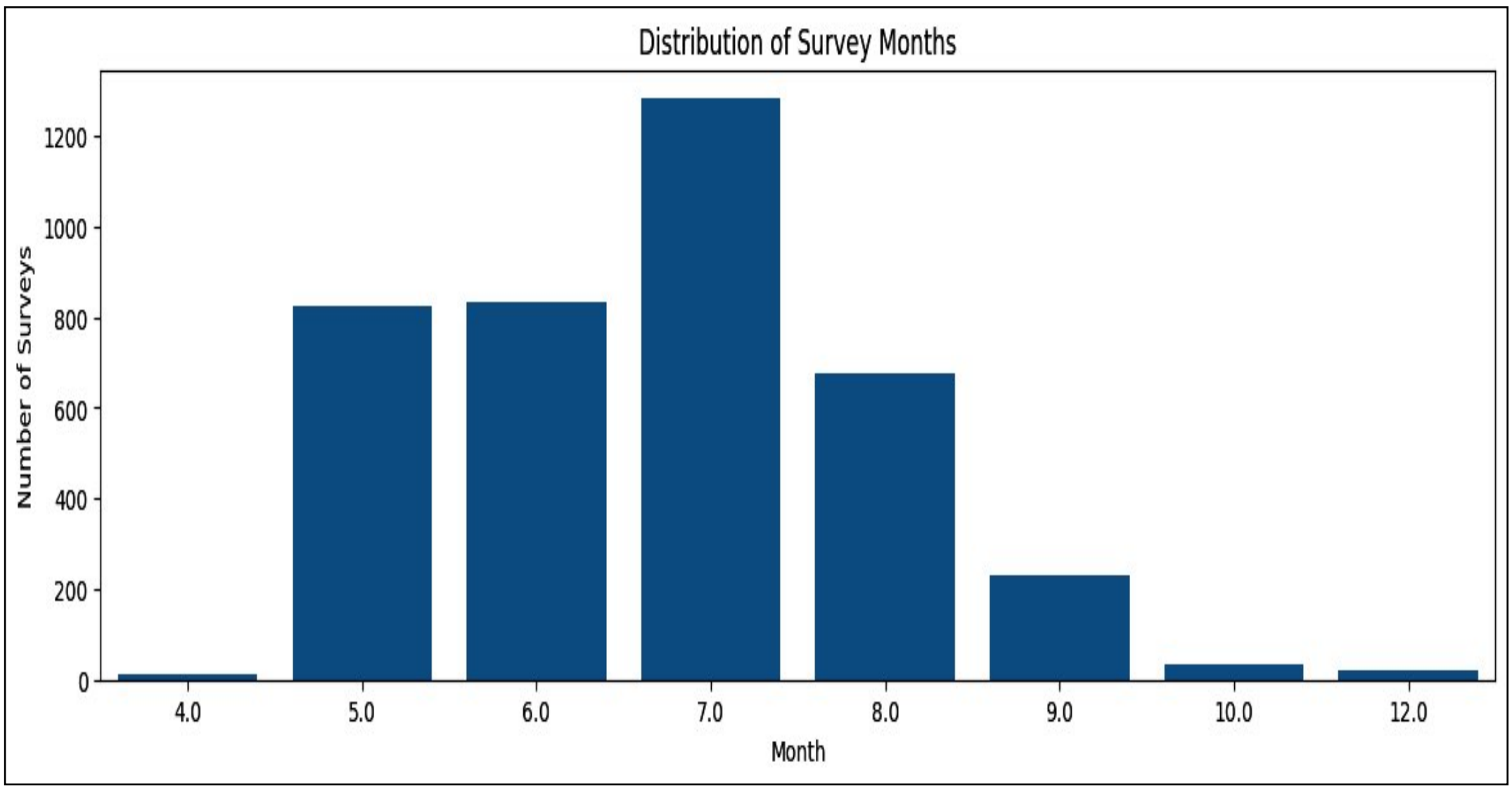
6. REFERENCES & ACKNOWLEDGEMENT

1. Cover Page Credit: Rapture Reef
2. Data Source Credit: CREMP Dataset
3. Data Source Metadata: CREMP Metadata
4. My JUPYTER Notebook: GITHUB Source Code
5. J Stein J, Huebner LK, Colella M, Harrell C, and Ruzicka R. 2024. Florida's Coral Reef 2023-2024 Post-Bleaching Assessment Quick Look Report.Florida Fish and WildlifeConservation Commission, Fish and Wildlife Research Institute, St. Petersburg, FL.23pp.https://ocean.floridamarine.org/FRRP/Home/Abouthttps://myfwc.com/research/habitat/coral/
6. Hurricane Georges On Florida

7. APPENDIX A: DISTRIBUTION OF SURVEY DAYS

1. The Monthly coral survey table observed shows the highest records on July (7th month) of the year and illustrated in the table here

Table A1



Missing rows were reported in the StonyCorel dataset as shown here in Table B1.

Table B1: Missing Date Rows In the StonyCoral Dataset

Index	Year	Subregion	SiteID	StationID	Site Name
224	1997	LK	40	401	Jaap Reef
225	1997	LK	40	402	Jaap Reef
226	1997	LK	40	403	Jaap Reef
227	1997	LK	40	404	Jaap Reef
252	1997	MK	56	561	Sombrero Shallow
253	1997	MK	56	562	Sombrero Shallow

8. APPENDIX B: MISSING VALUES IN STONYCORAL DATASET