Impact of Anthropogenic and Environmental Stressors on Bacterial Diversity and Abundance in Coral Reefs

Yijie Jiao

08/2024

A thesis submitted in partial fulfilment of the requirements for the degree of
Master of Science at Imperial College London
Submitted for the MSc in Computational Method in Ecology and Evolution

Declaration

The data used in this project was provided by the Emma Ransome laboratory, which had access to the necessary data sets. All subsequent tasks, including data processing, cleaning, analysis, and model development, were carried out solely by me. James Rosindell, my supervisor, did not directly contribute to the mathematical models or analyses presented in this thesis, but provided valuable guidance and feedback throughout the research process.

Gratefulness fills my heart for the support my girlfriend Yi has given me throughout my graduate studies.

Contents

1 Introduction **5**

2 Method **7**

2.1 Data Collection 7

2.2 Data Processing 9

2.3 Model Fitting and Visualisation 10

3 Results **12**

4 Discussion **16**

List of Tables

1 Description of Environmental Variables 8

List of Figures

1 Impact of Mean RESET Score on Hill numbers(N0,N1,N2) Across Different size fractionation 12

2 Influence of Environmental Stress Factors on Hill Number(N0, N1, N2) Predictive Error 13

3 Impact of Sea Surface Temperature(SST) on Hill Numbers(N0, N1, N2) under Different size fractionation 14

4 The Effects of Degree Heating Weeks (DHW) on Different Hill Numbers(N0, N1, N2) . 15

5 Impacts of Various Anthropogenic Stressors on Coral Reef Biodiversity 16

Abstract

Coral reefs are among the most biologically diverse ecosystems, crucial for both marine and human life. The health of these ecosystems heavily depends on bacterial communities, which play essential roles in nutrient cycling, coral symbiosis, and pathogen defense. Understanding the factors affecting bacterial diversity in coral reefs is vital for their conservation and management. This study examines the influence of various anthropogenic and environmental stressors on the diversity and abundance of bacterial communities within coral reefs. Using Hill numbers, which provide a comprehensive measure of biodiversity by considering species richness, evenness, and dominance, we assessed the diversity of these bacterial communities. The results indicate that comprehensive environmental stress, as measured by the RESET score, significantly decreases bacterial diversity, with Sea Surface Temperature (SST) and Degree Heating Weeks (DHW) being the most impactful factors, especially in reducing species richness. In contrast, factors such as SST variability and ocean currents showed minimal effects, suggesting a limited direct impact on microbial communities. In addition, anthropogenic pressures such as tourism and sedimentation emerged as key stressors. A key limitation of our study was the inability to observe temporal changes in bacterial diversity, which may have obscured short-term fluctuations and interactions. In addition, controlling for multiple factors that affect biodiversity proved challenging.

1 Introduction

Coral reefs are among the most biologically diverse ecosystems on Earth, playing a crucial role in maintaining global ecological balance and supporting human society [Graham and Nash, 2013]. Bacterial communities play multiple indispensable roles within the complex ecological network of coral reefs. Firstly, the symbiotic microbial community formed between corals and bacteria not only aids in nutrient acquisition and cycling—such as through nitrogen fixation, which converts atmospheric nitrogen into a form usable by corals—but also maintains the chemical balance of coral reef ecosystems through processes like organic matter decomposition and sulfide oxidation [Bourne and Webster, 2013]. Furthermore, the surface of corals is covered by a biofilm composed of various microorganisms, including a large number of symbiotic bacteria [Rohwer et al., 2002]. These bacteria significantly enhance the coral's defenses against external pathogens by competitively excluding harmful microbes, producing antimicrobial substances, and modulating the host's immune response [Bourne and Webster, 2013].

In the face of environmental change, a highly diverse bacterial community provides broad ecological support, enabling corals to adapt to various stressors such as temperature fluctuations, acidification, and increased pollution [Rohwer et al., 2002]. A highly diverse bacterial community provides broad ecological support, allowing corals to adapt to various stressors such as temperature fluctuations, acidification, and increased pollution [Zaneveld et al., 2016]. Different bacterial species may perform distinct functions under varying environmental conditions. When conditions change, some bacterial populations may decline, while others with adaptive traits may flourish, thus maintaining the functional stability of the entire coral reef system [Rohwer et al., 2001]. In essence, bacterial diversity provides corals with "functional redundancy", which means that even if certain bacterial functions fail, other bacteria can continue to perform similar functions, ensuring the resilience and sustainability of the ecosystem [Rosenberg et al., 2007]. The loss of bacterial diversity could lead to a decline in coral health, triggering a series of cascading effects that ultimately weaken the ecosystem services provided by coral reefs [Rohwer et al., 2001]. Therefore, understanding and preserving bacterial diversity within corals is not only vital for maintaining the health and functionality of coral reef ecosystems but also holds significant strategic importance for global biodiversity conservation and ecosystem management.

The bacterial diversity within coral reefs is increasingly being impacted by a range of environmental and anthropogenic pressures [Zaneveld et al., 2016]. Factors such as rising sea surface temperatures, ocean acidification, and the expansion of tourism and industrial development are driving significant changes in the composition and functionality of these critical microbial communities [Zaneveld et al., 2016]. These shifts not only undermine coral health, but also threaten the resilience and sustainability of the entire coral reef ecosystem [Rohwer et al., 2002]. To thoroughly investigate the effects of environmental and anthropogenic pressures on bacterial diversity within corals, we have utilized coral reef bacterial abundance data collected by various laboratories worldwide, along with corresponding environmental and anthropogenic pressure indicators. By analyzing these data, we aim to uncover the relationships between these pressures and changes in bacterial diversity, thereby providing crucial scientific insights for the conservation and management of coral reef ecosystems.

The bacterial abundance data for this study were collected using autonomous reef monitoring structures (ARMS). ARMS are standardised passive sampling devices designed to mimic the structural complexity of coral reefs, providing habitat for a wide range of marine organisms, including bacteria [Ransome et al., 2017]. These structures were deployed at various coral reef sites around the world to collect data from the microbial community over time [Ransome et al., 2017]. Upon retrieval, the ARMS units were processed to extract the bacterial communities.

To study the environmental impact, environmental stressor data was collected for each ARMS deployment site. The values of these variables were mapped to a range between 0 and 1 using a linear function [Williamson et al., 2022]. Finally, we obtain a comprehensive environmental stress exposure index (RESET score), which allows a unified assessment of the contributions of multiple variables [Williamson et al., 2022]. In addition to environmental stressors, this study also incorporates data on various anthropogenic pressures that affect coral reef ecosystems.

In this study, we used Hill numbers to assess the diversity of bacteria in coral reefs because Hill numbers provide a comprehensive and intuitive measure of biodiversity by integrating species richness, evenness and dominance into a single framework. By varying the parameter q, Hill numbers allow us to capture different aspects of diversity, allowing us to assess not only the number of species present (richness) but also how evenly these species are distributed and the degree to which a few species dominate the community [Ricotta and Feoli, 2024].Using Hill numbers, we can effectively compare diversity between samples with varying microbial loads and account for differences in community structure that might be overlooked by more simplistic diversity measures [Ricotta and Feoli, 2024]. This approach provides a more nuanced understanding of microbial community dynamics, especially when analyzing the impacts of environmental and anthropogenic pressures on the diversity of bacterial communities in coral reefs [Ricotta and Feoli, 2024].

Given the complexity of coral reef ecosystems and the multitude of factors that can influence their health, this study is based on several key hypotheses. First, it is hypothesised that changes in comprehensive environmental stress will have a significant impact on bacterial diversity within coral ecosystems. Second, we seek to identify which specific environmental factors exert the greatest influence on bacterial diversity and understand the mechanisms by which they operate. Finally, this study hypothesises that anthropogenic factors also play a crucial role in the formation of bacterial communities within corals, with some human-induced stressors potentially having more pronounced effects than others. These hypotheses will be tested through a comprehensive analysis of environmental and anthropogenic data, aiming to provide new insights into the interactions between coral health and external stressors.

2 Method

This study involved analyzing the impact of environmental and anthropogenic pressures on bacterial diversity within coral reefs, with all analyses conducted in R ver 3.6.3. We began by standardising the bacterial abundance data, calculating relative abundances, and computing Hill numbers to quantify diversity. To align temporal environmental data with bacterial abundance data, we applied data processing techniques. Following this, we constructed linear regression models and Generalised Additive Models (GAMs) to explore the relationships between environmental pressures and bacterial diversity. In addition, the Random Forest algorithm was employed to identify key environmental and anthropogenic factors that influence bacterial diversity. Finally, we visualized the results to better understand the complex interactions between these pressures and the microbial community structure in coral reefs.

2.1 Data Collection

We utilised bacterial abundance data derived from 228 samples. This data set comprises the detection results of different ESVs (Exact Sequence Variants) within each sample, where each ESV represents a specific bacterial species [Adrià Antich and Turon, 2021]. The abundance data is presented in a matrix format, with rows corresponding to different ESVs and columns representing individual samples. The values within the matrix indicate the quantity of each specific bacterium detected in the respective samples, while a value of zero indicates that the corresponding ESV was not detected in that sample.

The environmental variable data captures environmental stress-related data for each ARMS unit (Autonomous Reef Monitoring Structures) over different time periods. Each record in the dataset includes various variables such as the year, month, specific date, the name of the environmental variable(see Table:1), the Stress Exposure Score (SE score), the mean value of the variable (variable mean), the category of stress (whether the variable acts as a stressor or a reducer in the ecosystem), and the weighting category. The data is stored in a time-series format, showing the extent to which each ARMS unit is influenced by different environmental factors at various points in time.

Table 1: Description of Environmental Variables

Variable	**Description**
DHW (Degree Heating Weeks)	Measures accumulated heat stress over 12 weeks. Used to monitor coral bleaching events.
SST (Sea Surface Temperature)	The temperature of the surface layer of the ocean. SST is a key indicator of ocean-atmosphere interaction.
SST Anomaly	The difference between the observed SST and the long-term average. Positive values indicate above-average temperatures.
SST Variability	Variation in SST over time. High variability may indicate seasonal changes or extreme events.
Cloud (Cloud Cover)	The extent of sky covered by clouds, affecting solar radiation reaching the ocean surface.
Current (Ocean Currents)	Large-scale horizontal movement of water, important for the distribution of nutrients and heat.
Depth	The vertical distance from the surface to the bottom of the water body. Influences light penetration, temperature, and habitat distribution.
Salinity	The concentration of dissolved salts in seawater. Affects water density and the physiology of marine organisms.
Wind (Wind Speed and Direction)	The speed and direction of horizontal air movement, influencing heat exchange and ocean surface conditions.

The SE score is the primary metric used in this study to assess environmental stress on coral reefs. Calculating the SE score involves comparing the observed value of each environmental variable with predefined lower and upper threshold values. These thresholds are typically set based on known critical limits that impact coral health. If the observed value is below the lower threshold, the SE score is set to 0, indicating that there is no stress [Williamson et al., 2022]. If the value exceeds the upper threshold, the SE score is set to 1, indicating maximum stress. Values between the thresholds are interpolated linearly to assign an appropriate SE score. To minimise noise from data gaps or seasonal fluctuations, SE scores for each variable are often aggregated on a specific time scale. This metric is central to the analyses conducted in this study.

The comprehensive enviromental data records the RESET scores for each ARMS unit over different time periods. The RESET score is a comprehensive index that measures environmental stress exposure [Williamson et al., 2022]. The dataset includes information such as the event ID, year, month, RESET score, and specific dates. The RESET score is calculated based on the SE scores of multiple environmental variables using a specific algorithm, representing the level of environmental stress experienced at a particular location during a given time period [Williamson et al., 2022].

The anthropogenic stress dataset provides detailed information on various stressors impacting coral reef ecosystems. Primary variables include percentile rankings for fishing, coastal development, industrial development, tourism, sedimentation, and nitrogen pollution, all of which collectively influence the health and resilience of coral reefs. The data set also includes cumulative impact scores, categorisation of the top threats, and information on the region and climate-related scores where the samples

were collected. Additionally, it contains unprocessed raw data and specific climate stress indicators, such as historical thermal stress, projected future conditions, and recent thermal conditions. In this study, we investigate primarily the impact of percentile rankings for fishing activities, coastal development, industrial development, tourism, sediment accumulation, and nitrogen pollution on bacterial diversity.

2.2 Data Processing

First, we standardized the samples by calculating the relative abundance of each Exact Sequence Variant (ESV), which represents individual bacterial species within each sample. This calculation involves dividing the absolute abundance of each ESV by the total abundance within the sample.

To better assess the diversity within each sample, the hill numbers for each sample were calculated using the following steps [Ricotta and Feoli, 2024]:

- **N_0 (Species Richness)**: This is the count of unique species (or exact sequence variant, ESV) present in the sample. Mathematically, it is expressed as:

$$N_0 = \sum_{i=1}^{S} 1 = S$$

  where S is the total number of species observed in the sample.

- **N_1 (Shannon Entropy Exponential)**: This is a diversity measure that considers both species richness and the evenness of species distribution. It is defined as the exponential of the Shannon entropy:

$$N_1 = \exp\left(-\sum_{i=1}^{S} p_i \ln p_i\right)$$

  where p_i represents the relative abundance of species i.

- **N_2 (Reciprocal of Simpson's Diversity Index)**: This measure gives more weight to the most abundant species in the sample, reflecting the dominance structure of the community. It is calculated as:

$$N_2 = \left(\sum_{i=1}^{S} p_i^2\right)^{-1}$$

  where p_i represents the relative abundance of species i.

For each ARMS (Autonomous Reef Monitoring Structure), bacterial abundance data can only be obtained after the final collection of the device. However, environmental pressure data for each ARMS is continuously monitored over time. Consequently, for each sample, we only have a single set of bacterial abundance data, but a time-spanning set of environmental pressure data. To facilitate subsequent modelling and analysis, it is necessary to match these datasets accordingly.

We first applied double exponential smoothing to the time series data of RESET scores for each eventID. Visualisation revealed that the RESET score data exhibited significant seasonal variations, short-term fluctuations, and noise, leading to noticeable volatility. In such cases, directly using the raw RESET scores could introduce unnecessary noise into the analysis. Therefore, by applying double exponential smoothing, we were able to effectively smooth out these fluctuations, resulting in more stable and interpretable time series data. After the smoothing process, we calculated the mean of the

smoothed RESET scores for each eventID. This step was necessary because we needed a representative environmental pressure data point to correspond with the bacterial abundance data. Thus, by averaging the smoothed RESET scores, we simplified the complex time series data into a single, more manageable value.

For environmental variable SST, we applied a triple exponential smoothing model (Holt-Winters method) to analyze and smooth the SE scores for each eventID. The goal was to capture trends and seasonal components in the time series data while reducing short-term fluctuations and noise. After applying the smoothing process, we calculated the average of these smoothed values to generate a representative SE score for each eventID. Since a small portion of the ARMS devices were deployed for less than 12 months, the model failed to run successfully for those eventIDs. Consequently, we filtered the eventIDs where the smoothing process failed (resulting in NA values) to ensure that the final data set consists of a clean set of SE scores. For the environmental variable DHW, we simply calculated the mean values without applying any further smoothing or adjustments.

Finally, we consolidated the Hill number data, the corresponding sample information, the calculated mean RESET scores, and the mean SE scores of specific environmental variables into a single data set. This integration was performed to unify all relevant data into a cohesive framework, facilitating subsequent modeling analyses and data visualization.

2.3 Model Fitting and Visualisation

To investigate the impact of comprehensive environmental stress, measured by the RESET score, on the bacterial diversity of coral reefs (quantified by Hill numbers N0, N1 and N2), we constructed three linear regression models. Each model uses the mean RESET score and size fractionation as predictors, with an additional interaction term between these factors, to forecast the different Hill number indices (N0, N1, and N2):

```
lm(Hill numbers ~ mean RESET score * size fractionation)
```

Including size fractionation as a fixed effect is crucial, as bacterial diversity can vary significantly between different size fractionation. Taking into account this factor, we aim to reduce the standard error and enhance the accuracy of our model predictions. In addition, this approach allows us to explore whether the impact of environmental pressures on bacterial diversity differs between fractionation.

Similarly, we apply the same methodology to explore the relationship between the mean SE score and Hill numbers. For this analysis, we constructed three additional linear regression models:

```
lm(Hill numbers ~ Mean SE Score * size fractionation)
```

These models assess the influence of the mean SE score on the bacterial diversity indices (N0, N1, N2) while considering the interaction with size fractionation. By applying this approach, our aim is to determine how different environmental pressures, represented by the SE score, impact microbial diversity in various sizes of ecosystems in coral reefs.

To investigate the relationship between the mean SE score of Degree Heating Weeks (DHW) and bacterial diversity in coral reefs, we employed Generalised Additive Models (GAMs) for each Hill number (N0, N1, N2):

```
gam(Hill numbers ~ s(DHW) + s(size fractionation, bs=\re")
```

We chose to use Generalized Additive Models (GAMs) because linear models did not perform well in this context. GAMs offer greater flexibility in modeling complex, non-linear relationships that are often present in ecological data.By using GAMs, we can better account for these complexities and achieve more accurate and reliable results. Additionally, by treating size fractionation as a random effect within the model, we account for the inherent variability across different fractionations, thereby improving the overall robustness of the analysis.

We also used the Random Forest algorithm to analyse the impact of various environmental variables and anthropogenic stressors on coral reef bacterial diversity. In this study, we constructed separate Random Forest models for each Hill number (N0, N1, N2) and used these models to assess the importance of individual variables, with a particular focus on identifying which environmental factors and human-induced pressures significantly influence bacterial diversity.

By summarising the importance scores of the models, we were able to determine which variables contributed the most to the predictive performance of each Hill number index. The importance of variables in the Random Forest models is measured by the percentage increase in mean squared error (%IncMSE). This error quantifies the importance of each factor; a higher percentage increase in error suggests that the factor has stronger predictive power and explanatory capability for the corresponding Hill numbers and values less than or equal to 0 indicate no predictive power and explanatory capability for changes in Hill numbers.

In our study, we not only constructed and analysed various models, but also visualised the results to better understand the relationship between environmental pressures and bacterial diversity in coral reefs. For each model, we generated graphs that show the changes in the predicted Hill numbers (N0, N1, N2) compared to the observed values under different environmental pressure scores and size fractionations.

For the linear models and generalized additive models (GAMs), we created scatter plots along with regression lines (or smooth curves in the case of GAMs) to represent the trends predicted by the models. These visualizations help to clearly illustrate the relationship between environmental pressures and bacterial diversity.The results of the random forest models were visualized using bar charts, which ranked the importance of different environmental and anthropogenic factors in predicting the Hill numbers. This method of visualisation highlights the factors that have the most significant impact on bacterial diversity in coral reefs and also facilitates comparison of the influencing factors across different Hill number indices.

3 Results

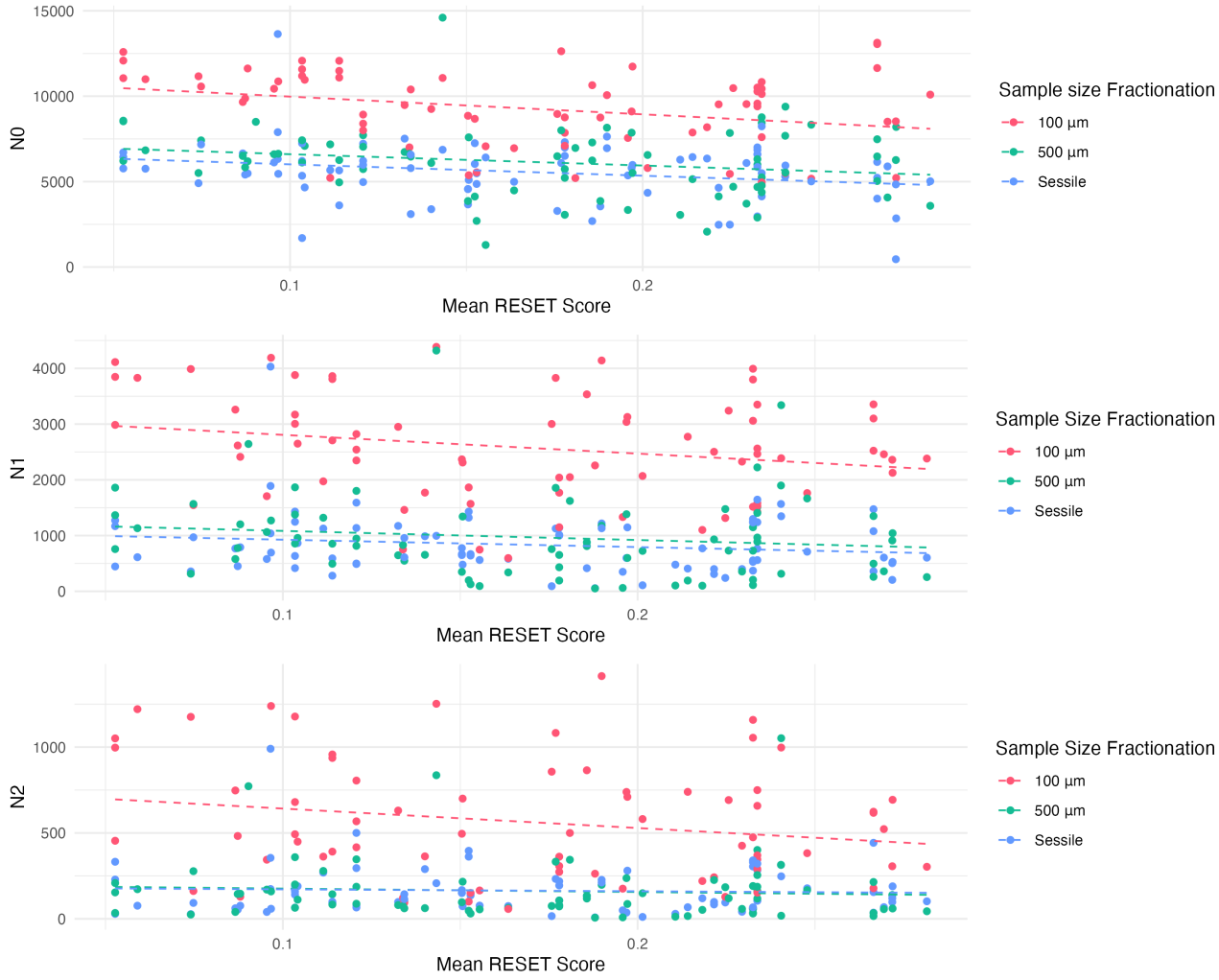

Figure 1: Impact of Mean RESET Score on Hill numbers(N0,N1,N2) Across Different size fractionation

The RESET score exhibits a significant negative influence on N0 (*coefficient = -10373.4,* $p = 0.0044$), shows that an increase in comprehensive environmental stress correlates strongly with a decrease in species richness. There were notable reductions in the species richness for size fractionation 500μm and sessile organisms compared to the baseline ($p < 0.0001$). The interaction between the RESET score and the sample size did not show a significant effect on N0, indicating that the response of bacterial communities to comprehensive environmental stress is consistent across different size fractionation ($p > 0.45$).

Similarly, the RESET score significantly negatively impacts N1 (*coefficient = -3364.0,* $p = 0.0153$), indicating a decrease in species evenness due to increased comprehensive environmental stress. Significant decreases in N1 were also observed in the size fractionation of 500μm and sessile organisms ($p < 0.00001$). The p-value of the interaction term was relatively high ($p > 0.28$), showing that the effect of the comprehensive environmental stress score interacting with the size fractionation on N1 is not significant.

For N2, the persistent negative impact of the RESET score (*coefficient = -1132.04,* $p = 0.00998$) highlights a significant reduction in the composite biodiversity indicator of coral reefs associated with increased comprehensive environmental stress. Groups with size fractionation of 500μm and sessile

organisms also showed significant decreases in N2 ($p < 0.0001$).The interaction term's p-value was relatively high ($p > 0.088$), showing that the effect of the comprehensive environmental stress score that interacts with size fractionation on N2 is not significant.

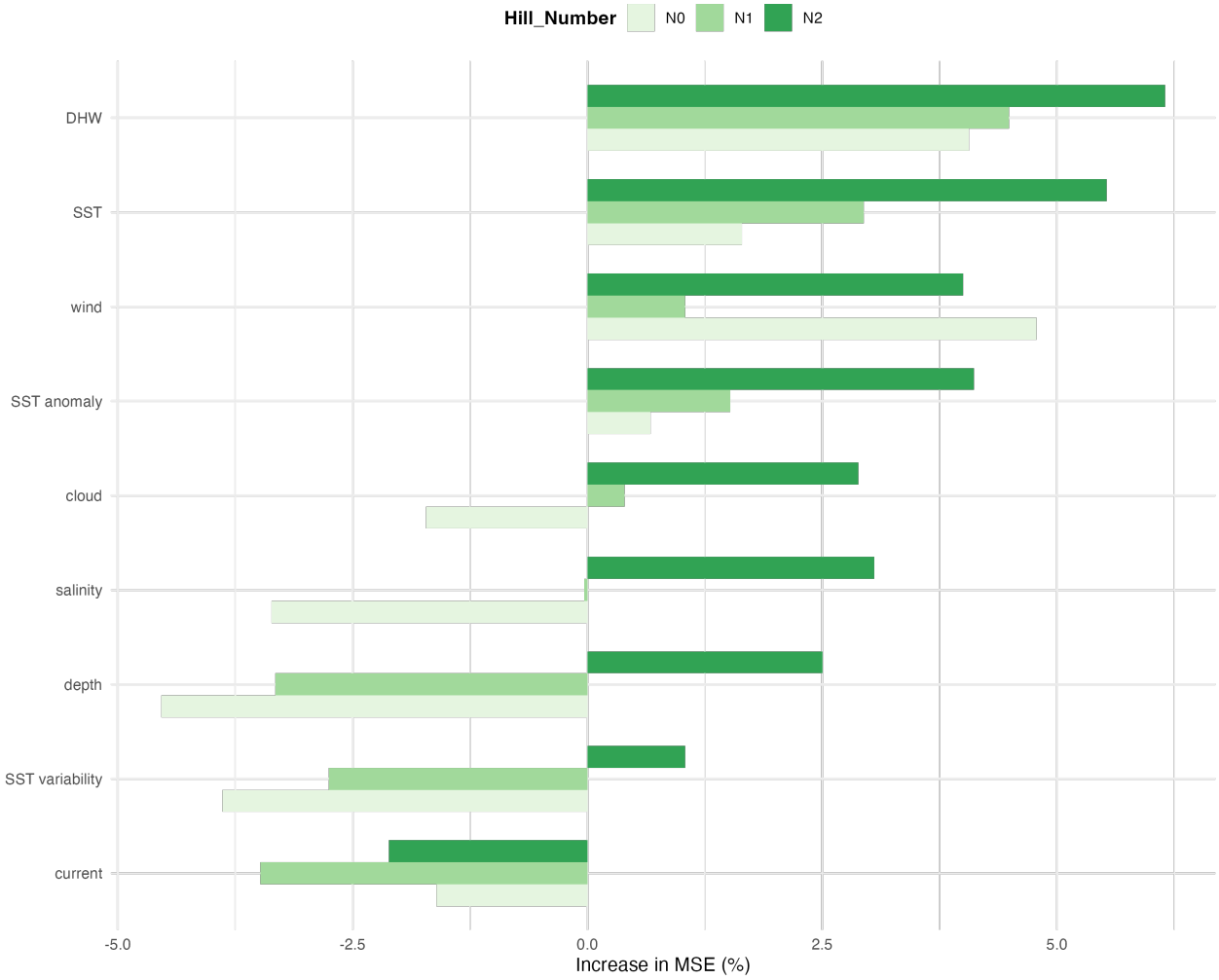

Figure 2: Influence of Environmental Stress Factors on Hill Number(N0, N1, N2) Predictive Error

The impact of various environmental stress factors on the predictive error associated with Hill numbers has been illustrated(see Fig.2). The importance of different environmental stressors on the diversity of bacterial communities in coral reefs varies significantly(between -0.06 and 0.06).

Degree Heating Weeks (DHW) and Sea Surface Temperature (SST) demonstrate the highest importance in the predictions across all three Hill numbers (N0, N1, N2), underscoring the profound impact of temperature factors on the diversity of coral reef bacterial communities. In contrast, Sea Surface Temperature Variability (SST Variability) and Ocean Currents (Current) have minimal impacts on bacterial diversity.Additionally, sea surface temperature anomaly (SST anomaly) and wind also show considerable importance in predicting some of the hill numbers. Moreover, apart from Current and SST Variability, other environmental factors exhibit significant relevance in the predictions for N2. However, for N0 and N1, it is primarily DHW, SST, and Wind that display robust explanatory power.

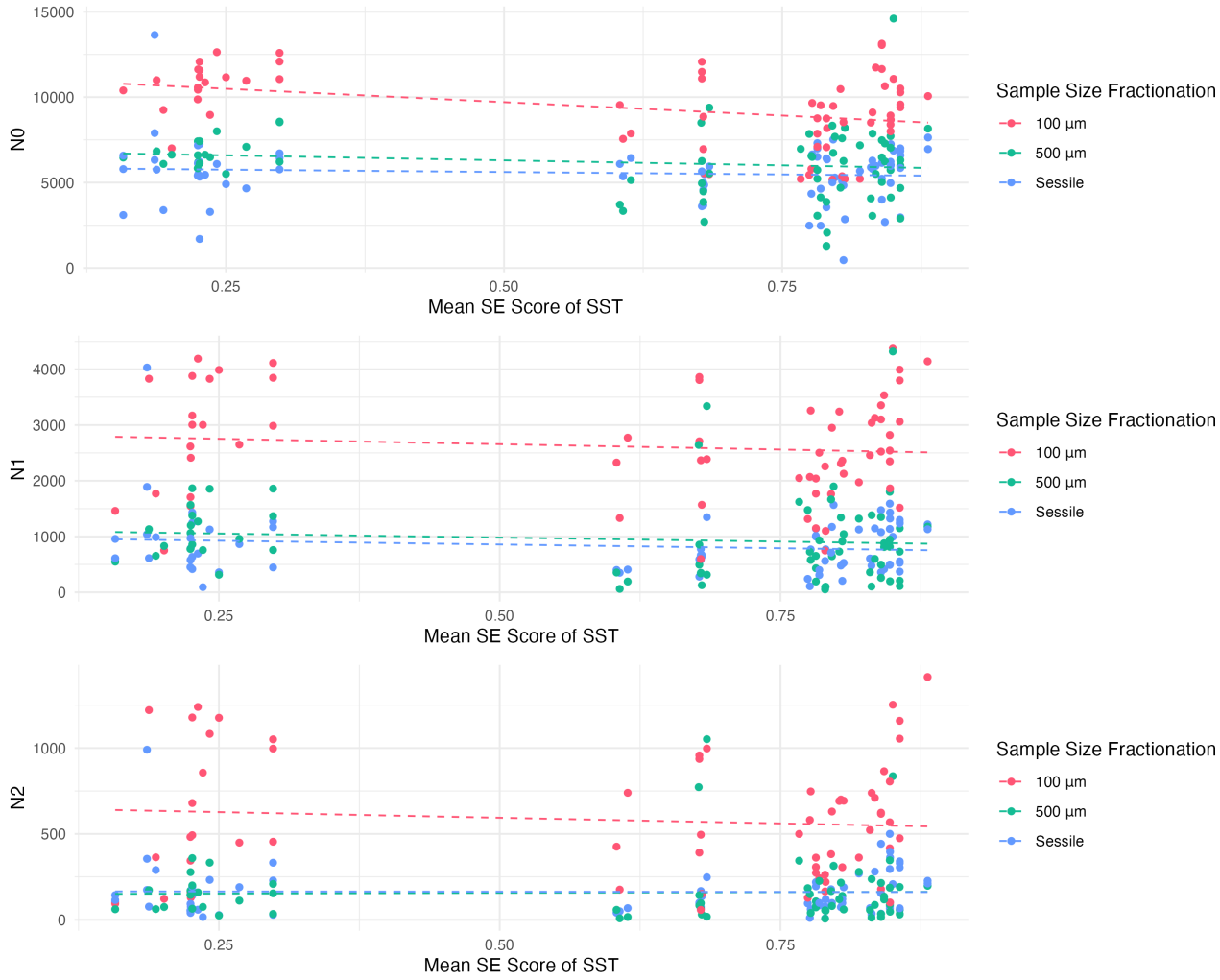

Figure 3: Impact of Sea Surface Temperature(SST) on Hill Numbers(N0, N1, N2) under Different size fractionation

Figure 3 provides a detailed examination of the impact of sea surface temperature (SST) scores on the diversity of bacterial communities in coral reefs. The findings indicate that an increase in the mean SE score of SST significantly negatively affects the species richness (N0), with a slope of *-3139.28* ($p = 0.0009$), showing that as the SST increases, the number of bacterial species in coral reefs decreases significantly. The coefficients for sample sizes fractionation 500μm and sessile organisms are *-4393.24* ($p < 0.0001$) and *-5379.76* ($p < 0.0001$), respectively, indicating that species richness in these groups is significantly lower than in the baseline size fractionation. The interaction terms between the mean SE score of SST and sizes fractionation 500μm and sessile organisms are *1970.58* ($p = 0.0134$) and *2578.51* ($p = 0.0339$), respectively, suggesting that the decline in species richness is somewhat mitigated under increased SST in these sample sizes.

For N1, the impact of the SST scores is *-385.64* ($p = 0.2950$), indicating that there is no statistically significant effect on species evenness. The effects for sample sizes fractionation 500μm and sessile organisms are significant, with coefficients of *-1723.60* ($p < 0.0001$) and *-1857.22* ($p < 0.0001$), indicating a significant reduction in species evenness for larger sample sizes.

Regarding N2, the impact of the SST scores is *-131.86* ($p = 0.2614$), which is also not statistically significant.

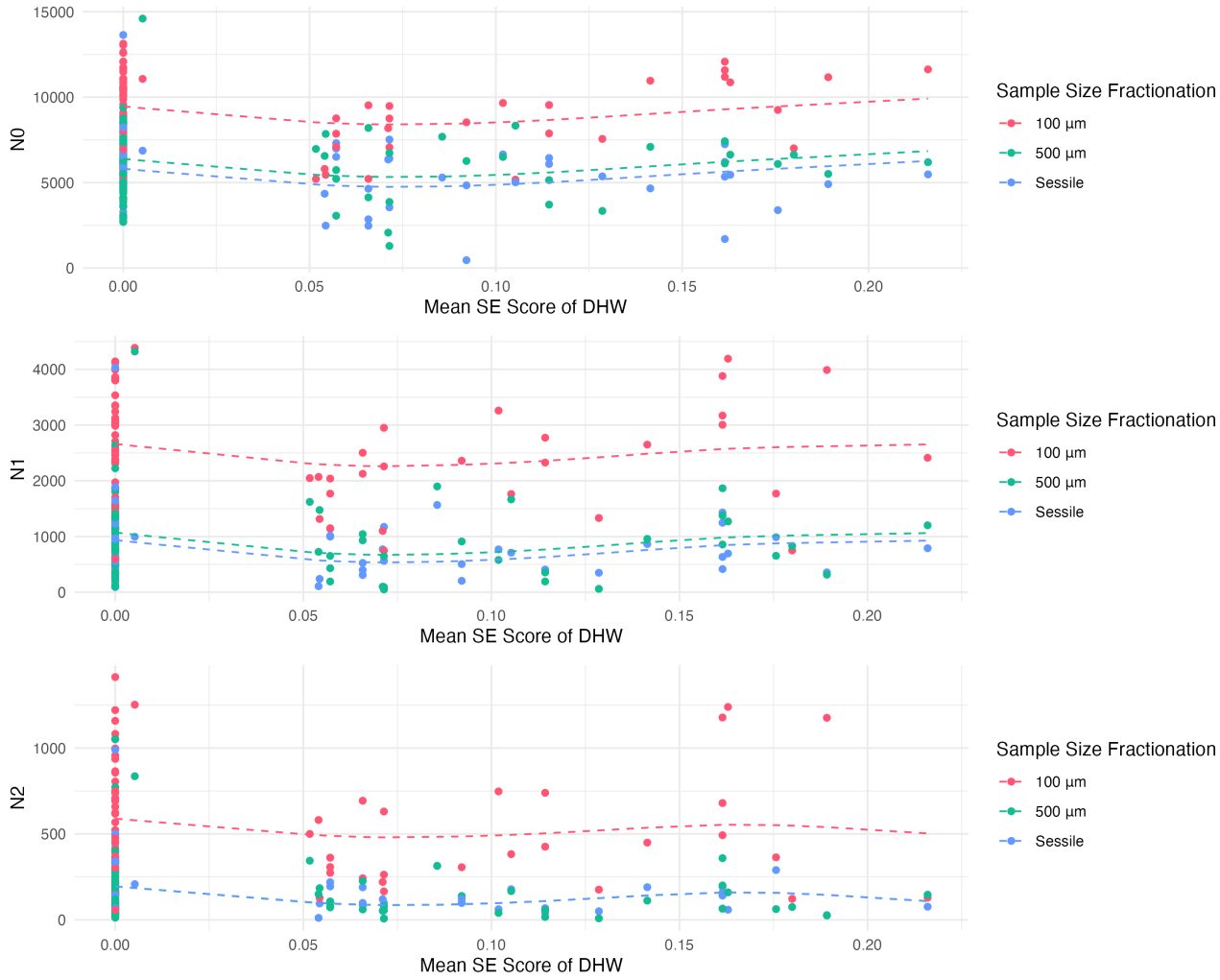

Figure 4: The Effects of Degree Heating Weeks (DHW) on Different Hill Numbers(N0, N1, N2)

For N0, DHW and size fractionation significantly affect the observations. As the mean SE score of DHW increase, N0 initially shows a significant decline, particularly in the range from 0.00 to approximately 0.05 ($p = 0.00659$, $F = 4.722$). Subsequently, as the mean SE score of DHW increases to higher values, N0 tends to stabilise or slightly rebound. Across different fractionations the trend in N0 remains largely consistent, but the rebound is more pronounced in the 500μm and sessile fractionation, possibly due to differences in sample processing or ecological response mechanisms ($p < 0.0001$, $F = 76.158$).

For N1, the impact of DHW shows a significant negative effect at midrange values, similar to N0, but with greater complexity in the response ($edf = 2.632$, $p = 0.0118$). N1 initially decreases as the mean SE score of DHW increases, although the decline is less pronounced than in N0. As the mean SE score of DHW reaches moderate levels, N1 stabilises and shows a slight recovery at certain points, similar to the adaptive or recovery response seen in N0. The pattern of change in N1 is quite consistent across different size fractionations.

The model for N2 indicates that the impact of DHW is not significant ($p = 0.0514$, $F = 2.669$), with a complex pattern of change. Initially, N2 also shows a declining trend with increasing mean SE score of DHW, but significant fluctuations occur around a mean SE score of DHW = 0.1. Subsequently, N2 displays recovery at a certain mean SE score of DHW, particularly more pronounced at higher mean SE score of DHW.

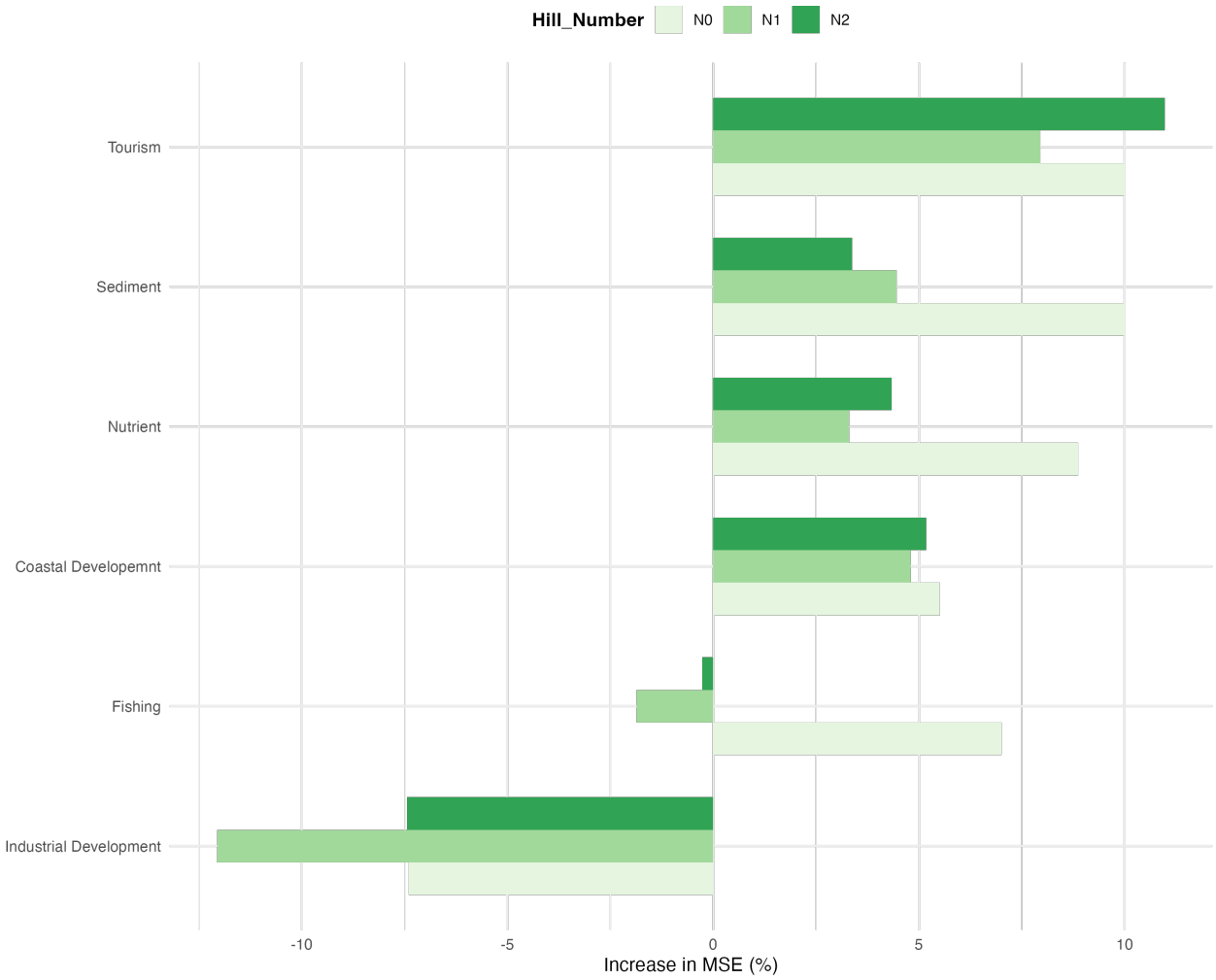

Figure 5: Impacts of Various Anthropogenic Stressors on Coral Reef Biodiversity

The importance of different anthropogenic stressors on the diversity of bacterial communities in coral reefs varies significantly (see fig.6). The results of the random forest model show that tourism has a high explanatory power for all three indices, N0, N1, and N2. Sediment has the most significant explanatory power for N0, but its explanatory power for N1 and N2 is not as strong. Nutrients also exhibit strong explanatory power, particularly for N0. Fishing demonstrates strong predictive power primarily for N0, but it does not significantly explain variations in N1 or N2. Conversely, industrial development does not show any predictive power for changes in N0, N1, or N2.

4 Discussion

In all of our models, linear or GAM, sample fractionation consistently showed a significant impact on bacterial diversity, measured by Hill numbers (N_0, N_1, N_2). Our findings indicate that the 100 micrometre fraction consistently exhibited the highest diversity across all metrics, followed by the 500 micrometre fraction, and the sessile fraction showed the lowest diversity. This can be attributed to several factors: 100 micrometre fractionation is more effective in capturing free-living bacteria and small microbial particles that are more abundant and diverse, leading to a higher species richness (N0) [Garren et al., 2014]. In addition, bacterial communities in 100 micrometre fractionation tend to be more evenly distributed, resulting in greater N1 and N2. This is because smaller bacteria and

microbial particles are less likely to dominate the community, leading to a more balanced distribution of species and function [Carvalho et al., 2013].In contrast, the 500 micrometer fractionation and sessile samples tend to capture bacteria attached to surfaces or larger microbial aggregates, where a few dominant species can out compete others [Schöttner et al., 2012]. This dominance results in lower N1 and N2, making these communities less diverse in terms of both species richness and functional capacity [Jizhong et al., 2002].

For N0, the significant negative impact of the RESET score ($p = 0.0044$) indicates that the species richness decreases markedly with increasing environmental stress. This suggests that comprehensive environmental stress exerts a direct and strong inhibitory effect on the number of species within coral reef bacterial communities [Bourne et al., 2008]. This may be because increased environmental stress makes it difficult for certain bacterial species to survive, thus reducing the number of species within the community [McDevitt-Irwin et al., 2017]. In addition, intensified competition for resources may result in some species being out-competed by more adaptable ones. Similarly, the negative impact of the RESET score on N1 ($p = 0.0153$) indicates that increasing environmental stress also leads to a decrease in species evenness. This is likely due to the suppression of growth in certain species, particularly those highly sensitive to specific environmental conditions, leading to an increase in the relative abundance of dominant species and a reduction in overall evenness [McDevitt-Irwin et al., 2017]. For N2, the significant negative impact of the RESET score ($p = 0.00998$) indicates a reduction in the abundance of dominant species within coral reefs, suggesting that even dominant species can be affected under extreme stress conditions, leading to a decrease in their numbers [McDevitt-Irwin et al., 2017].

Although size fractionation has some impact on these diversity indices, the negative influence of the RESET score on the diversity of the bacterial community remains relatively consistent between different fractionations. This suggests that comprehensive environmental stress leads to a decline in microbial diversity within coral reefs, regardless of size fractionation.

For individual environmental factors, sea surface temperature (SST) has a significant impact on the diversity of bacterial communities in coral reefs, particularly with regard to species richness (N0). The results show that as the mean SE score of SST increases, N0 significantly decreases ($p = 0.0009$), indicating that elevated temperatures have a strong negative effect on the number of bacterial species in coral reef ecosystems. Under high temperature conditions, bacterial proteins can denature and enzyme activity can decrease, directly affecting bacterial growth and metabolic efficiency [Ferreira et al., 2013a]. Additionally, high temperatures can disrupt bacterial metabolic pathways, leading to decreased energy metabolism efficiency, thereby inhibiting growth and reproduction. Metabolic disruptions may also hinder bacteria from effectively utilizing nutrients, impacting their survival in coral reefs. [Ferreira et al., 2013a] Moreover, rising sea temperatures are one of the main causes of coral bleaching, leading to the loss of symbiotic algae and bacteria, weakening the coral's ability to acquire nutrients and reducing the available habitat for bacteria [Deser et al., 2010].

Although the SST impact coefficients for N1 and N2 are also negative, these relationships are not statistically significant (p *values of 0.2950 and 0.2614*, respectively). This suggests that while SST has a significant effect on species richness, its impact on species evenness and dominance is relatively weak or not significant. This may reflect that even with rising temperatures, although the abundance of certain species may decrease, elevated temperatures do not significantly alter the relative abundance relationships among the remaining species [Deser et al., 2010]. Furthermore, certain dominant species may have stronger adaptability, allowing them to survive under increased temperatures and thus maintain the overall structure of the community [Lamb et al., 2014].

When examining the interaction between size fractionation and SST, the results show that the decline in species richness in the 500μm and sessile organism fractionation is somewhat mitigated under increased SST (*p values of 0.0134 and 0.0339*, respectively). This may indicate that, while bacterial communities in 500 micrometre and sessile fractionation are sensitive to rising temperatures, certain populations may have the ability to adapt to or cope with elevated SST, thus partially alleviating the decline in species richness [Hoegh-Guldberg, 2011].

Degree Heating Weeks (DHW) have a complex impact on the diversity of bacterial communities in coral reefs, particularly on species richness (N0). As the mean SE score of DHW increases, N0 initially shows a significant decline, especially in the range of 0.00 to 0.05. This indicates that once temperature anomalies accumulate beyond a certain threshold, the adverse effects on bacterial diversity in coral reefs become apparent, leading to a reduction in the species richness [Hughes et al., 2018]. This decline reflects the direct impact of thermal stress on the most vulnerable species, which can cause some species to disappear. However, as the mean SE score of DHW continues to increase, N0 tends to stabilise or slightly rebound. This trend may suggest that bacterial communities are adapting to the ongoing thermal stress or that more heat-tolerant species are beginning to dominate, preventing further declines in species richness [Kayanne, 2017].

For N1, the effect of DHW is relatively gradual, showing a steady decrease as the mean SE score of DHW increases. This change is more moderate compared to N0, indicating that even though some species are affected, the community structure may still maintain a certain level of balance. Regarding N2, the relationship between the mean SE score of DHW and N2 is not significant. This could be because the complexity of the coral reef ecosystem provides a buffering capacity, which prevents any single species from becoming overly dominant [Bourne and Munn, 2005]. Even under environmental stress, the system can maintain balance through various biological responses [Bourne and Munn, 2005].

We evaluated the impact of various environmental variables on the prediction errors related to bacterial diversity and the results showed significant differences in the influence of different environmental stress factors. Firstly, Degree Heating Weeks (DHW) and Sea Surface Temperature (SST) consistently exhibited the highest importance in all three hill numbers (N0, N1, N2), highlighting the profound impact of temperature-related factors on the diversity of coral reef bacterial communities(see Fig:3 and Fig:4).

In contrast, Sea Surface Temperature Variability (SST Variability) and Ocean Currents (Current) had almost no impact on bacterial diversity. This might be because the changes in these factors are typically minor or periodic and do not directly lead to rapid changes in bacterial communities [van Hooidonk and Huber, 2012]. The SST variability reflects the fluctuations in the sea surface temperature, but these fluctuations might not reach the threshold necessary to cause large-scale changes in bacterial communities [van Hooidonk and Huber, 2012]. Furthermore, while ocean currents influence the distribution of nutrients and temperature in the ocean, the complex structure of coral surfaces and the secretion of mucus provide a stable microenvironment for bacteria, mitigating the direct impact of water flow [Munday et al., 2009]. As a result, unless there is a significant change in water flow speed, its impact on bacterial communities is minimal [Munday et al., 2009].

Sea Surface Temperature Anomaly (SST Anomaly) and Wind demonstrated substantial explanatory power for changes in Hill numbers. SST anomaly refers to deviations from long-term average sea surface temperature, often resulting from extreme climate events such as El Nio [Ferreira et al., 2013b]. Such anomalies can directly affect the health of the coral reefs, subsequently altering the composition of the bacterial community [Ferreira et al., 2013b]. Wind influences the mixing of seawater and gas

exchange, which can change the physical and chemical conditions of the surface waters of coral reefs, indirectly affecting the metabolism and distribution of bacteria [Munday et al., 2009].

The results of the random forest model reveal that different anthropogenic stressors also have significantly varying impacts on the diversity of bacterial communities in coral reefs. First, tourism stands out among the anthropogenic factors, showing a high explanatory power in the three Hill numbers (N0, N1, and N2). A key factor contributing to this impact is chemical pollution. Common pollutants, such as chemicals found in sunscreen used by tourists, have been shown to be toxic to corals and their symbiotic algae, leading to coral bleaching and inhibiting the growth of coral larvae, which subsequently alters bacterial diversity [Danovaro et al., 2008]. Activities such as diving and swimming, which often involve direct physical contact with coral reefs, can alter the biofilms on the coral surfaces and damage the structure of their microbial communities, thereby reducing their diversity [Lamb et al., 2014].

Sediment has the most significant explanatory power for N0 but relatively weaker explanatory power for N1 and N2. This suggests that sediment primarily impacts species richness (N0), meaning that high sedimentation rates may suffocate certain species or create unfavorable conditions for their survival, thereby reducing the total number of species. However, its smaller impact on N1 and N2 implies that, although sedimentation might reduce species numbers, it does not significantly alter the evenness of the remaining community or the structure of dominant species. Sediment buildup can physically block the coral surface and surrounding microenvironments, altering the original microhabitat conditions [Rogers, 1990]. These changes may make it impossible for some bacterial populations to survive, leading to a decrease in the total number of species [Rogers, 1990].

Fishing activities have primarily affected N0, indicating a significant influence on species richness. This likely reflects the disruptions caused by overfishing in coral reef ecosystems, where the depletion of key species disrupts the ecological balance, subsequently altering microbial communities. The lack of significant explanatory power for N1 and N2 suggests that fishing primarily affects species richness, with less impact on community evenness and dominant species structure. Fishing activities can alter the structure and dynamics of biological communities by directly reducing the number of certain species and indirectly affecting non-target species through food web interactions [Jessen et al., 2013]. When species that play crucial roles in maintaining the food web and the functional integrity of coral reef ecosystems are impacted, the richness of bacteria associated with corals may also be affected [Jessen et al., 2013].

In this study, one of the main limitations we encountered was the inability to observe temporal changes in the bacterial diversity. This constraint necessitated the use of smoothing and averaging techniques for the environmental data, which may have obscured important short-term fluctuations and interactions between bacterial communities and environmental variables, thereby significantly reducing the precision of our research. Additionally, since biodiversity is influenced by a multitude of factors [Hoegh-Guldberg, 2011], it was challenging to fully control for other variables in single-factor studies [Fabricius, 2005]. For example, even ARMS units placed at the same time and location could exhibit noticeable differences in bacterial diversity. This inevitable variability added complexity to the interpretation of results. In addition, the inconsistency in the duration of deployment of different ARMS units, some being deployed for one year and others for up to three years, probably also affected the results in bacterial diversity [Ransome et al., 2017].

Looking ahead, our next steps will focus on further refining the models to better understand the causal mechanisms behind the observed patterns. Future research should prioritise the incorporation

of longitudinal studies to more effectively capture the temporal dynamics of bacterial communities. Future efforts will focus on exploring a broader range of indices to measure biodiversity, moving beyond the exclusive use of Hill numbers. These approaches aim to provide a more comprehensive understanding of the factors driving diversity changes within coral reef ecosystems. In addition, the ultimate goal is to investigate potential mitigation strategies to alleviate the negative impacts of these stressors, contributing thus to the conservation and resilience of coral reef ecosystems.

Data and Code Availability

All the data and code used in this report are available on my GitHub repository [GitHub Repository](https://github.com/ekko857/ekk0/tree/master/PROJECT).

References

[Adrià Antich and Turon, 2021] Adrià Antich, Creu Palacin, O. S. W. and Turon, X. (2021). To denoise or to cluster, that is not the question: optimizing pipelines for coi metabarcoding and metaphylogeography. *BMC Bioinformatics*, 22:177.

[Bourne et al., 2008] Bourne, D., Iida, Y., Uthicke, S., and Smith-Keune, C. (2008). Changes in coral-associated microbial communities during a bleaching event. *The ISME Journal*, 2(4):350–363.

[Bourne and Munn, 2005] Bourne, D. G. and Munn, C. B. (2005). Diversity of bacteria associated with the coral pocillopora damicornis from the great barrier reef. *Environmental microbiology*, 7(8):1162–1174.

[Bourne and Webster, 2013] Bourne, D. G. and Webster, N. S. (2013). *Coral Reef Bacterial Communities*, pages 163–187. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Carvalho et al., 2013] Carvalho, J. C., Cardoso, P., Borges, P. A., Schmera, D., and Podani, J. (2013). Measuring fractions of beta diversity and their relationships to nestedness: a theoretical and empirical comparison of novel approaches. *Oikos*, 122(6):825–834.

[Danovaro et al., 2008] Danovaro, R., Bongiorni, L., Corinaldesi, C., Giovannelli, D., Damiani, E., Astolfi, P., Greci, L., and Pusceddu, A. (2008). Sunscreens cause coral bleaching by promoting viral infections. *Environmental Health Perspectives*, 116(4):441–447.

[Deser et al., 2010] Deser, C., Alexander, M. A., Xie, S.-P., and Phillips, A. S. (2010). Sea surface temperature variability: Patterns and mechanisms. *Annual review of marine science*, 2(1):115–143.

[Fabricius, 2005] Fabricius, K. E. (2005). Effects of terrestrial runoff on the ecology of corals and coral reefs: Review and synthesis. *Marine Pollution Bulletin*, 50(2):125–146.

[Ferreira et al., 2013a] Ferreira, B., Costa, M., Coxey, M., Gaspar, A., Veleda, D., and Araujo, M. (2013a). The effects of sea surface temperature anomalies on oceanic coral reef systems in the southwestern tropical atlantic. *Coral reefs*, 32:441–454.

[Ferreira et al., 2013b] Ferreira, B., Costa, M., Coxey, M., Gaspar, A., Veleda, D., and Araujo, M. (2013b). The effects of sea surface temperature anomalies on oceanic coral reef systems in the southwestern tropical atlantic. *Coral reefs*, 32:441–454.

[Garren et al., 2014] Garren, M., Smriga, S., Azam, F., and Thurber, R. V. (2014). A bacterial shed associated with coral bleaching. *PloS one*, 9(5):e92599.

[Graham and Nash, 2013] Graham, N. A. J. and Nash, K. L. (2013). The importance of structural complexity in coral reef ecosystems. *Coral Reefs*, 32(2):315–326.

[Hoegh-Guldberg, 2011] Hoegh-Guldberg, O. (2011). Coral reef ecosystems and anthropogenic climate change. *Regional Environmental Change*, 11(1):215–227.

[Hughes et al., 2018] Hughes, T. P., Kerry, J. T., Baird, A. H., Connolly, S. R., Dietzel, A., Eakin, C. M., Heron, S. F., Hoey, A. S., Hoogenboom, M. O., Liu, G., et al. (2018). Global warming transforms coral reef assemblages. *Nature*, 556(7702):492–496.

[Jessen et al., 2013] Jessen, C., Villa Lizcano, J. F., Bayer, T., Roder, C., Aranda, M., Wild, C., and Voolstra, C. R. (2013). In-situ effects of eutrophication and overfishing on physiology and bacterial diversity of the red sea coral acropora hemprichii. *PLoS One*, 8(4):e62091.

[Jizhong et al., 2002] Jizhong, Z., Beicheng, X., S., T. D., Y., W. L., L., M. T., V., O. R., V., P. A., and M., T. J. (2002). Spatial and resource factors influencing high microbial diversity in soil. *Applied and Environmental Microbiology*, 68(1):326–334.

[Kayanne, 2017] Kayanne, H. (2017). Validation of degree heating weeks as a coral bleaching index in the northwestern pacific. *Coral Reefs*, 36(1):63–70.

[Lamb et al., 2014] Lamb, J. B., True, J. D., Piromvaragorn, S., and Willis, B. L. (2014). Scuba diving damage and intensity of tourist activities increases coral disease prevalence. *Biological Conservation*, 178:88–96.

[McDevitt-Irwin et al., 2017] McDevitt-Irwin, J. M., Baum, J. K., Garren, M., and Vega Thurber, R. L. (2017). Responses of coral-associated bacterial communities to local and global stressors. *Frontiers in Marine Science*, 4:262.

[Munday et al., 2009] Munday, P., Leis, J., Lough, J., Paris, C., Kingsford, M., Berumen, M. L., and Lambrechts, J. (2009). Climate change and coral reef connectivity. *Coral reefs*, 28:379–395.

[Ransome et al., 2017] Ransome, E., Geller, J. B., Timmers, M., Leray, M., Mahardini, A., Sembiring, A., Collins, A. G., and Meyer, C. P. (2017). The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (arms) and metabarcoding to measure cryptic diversity on mo'orea coral reefs, french polynesia. *PLOS ONE*, 12(4):e0175066.

[Ricotta and Feoli, 2024] Ricotta, C. and Feoli, E. (2024). Hill numbers everywhere. does it make ecological sense? *Ecological Indicators*, 161:111971.

[Rogers, 1990] Rogers, C. S. (1990). Responses of coral reefs and reef organisms to sedimentation. *Marine Ecology Progress Series*, pages 185–202.

[Rohwer et al., 2001] Rohwer, F., Breitbart, M., Jara, J., Azam, F., and Knowlton, N. (2001). Diversity of bacteria associated with the caribbean coral montastraea franksi. *Coral reefs*, 20:85–91.

[Rohwer et al., 2002] Rohwer, F., Seguritan, V., Azam, F., and Knowlton, N. (2002). Diversity and distribution of coral-associated bacteria. *Marine Ecology Progress Series*, 243:1–10.

[Rosenberg et al., 2007] Rosenberg, E., Koren, O., Reshef, L., Efrony, R., and Zilber-Rosenberg, I. (2007). The role of microorganisms in coral health, disease and evolution. *Nature Reviews Microbiology*, 5(5):355–362.

[Schöttner et al., 2012] Schöttner, S., Wild, C., Hoffmann, F., Boetius, A., and Ramette, A. (2012). Spatial scales of bacterial diversity in cold-water coral reef ecosystems. *PLOS ONE*, 7(3):1–11.

[van Hooidonk and Huber, 2012] van Hooidonk, R. and Huber, M. (2012). Effects of modeled tropical sea surface temperature variability on coral reef bleaching predictions. *Coral Reefs*, 31:121–131.

[Williamson et al., 2022] Williamson, J. et al. (2022). The reset tool for coral reef environmental stress exposure: Development and application to coral reef health assessment. *Marine Environmental Research*, 145:101756.

[Zaneveld et al., 2016] Zaneveld, J. R., Burkepile, D. E., Shantz, A. A., Pritchard, C., McMinds, R., Payet, J. P., Welsh, R., Correa, A. M. S., Lemoine, N. P., Rosales, S., Fuchs, C., Maynard, J. A., and Thurber, R. L. V. (2016). Overfishing and nutrient pollution interact with temperature to disrupt coral reefs down to microbial scales. *Nature Communications*, 7:11833.