Climate Vulnerability and Protection Effectiveness on Rocky Reefs: Insights from 27 Years of Monitoring in the Gulf of California

Fabio Favoretto[1,2]*, Octavio Aburto-Oropeza[2]
[1]School of Biological and Marine Sciences, University of Plymouth, UK
[2]Marine Biology Research Division, Scripps Institution of Oceanography, UC San Diego, USA
*Corresponding author. Email: fabio.favoretto@plymouth.ac.uk

Main quantitative results in this report are derived by AutoDiscovery,
a discovery tool built by Allen Institute for AI

1 Introduction

The Gulf of California represents one of the most species-rich marine regions in the Eastern Pacific. Since 1998, a collaborative network of researchers from the Scripps Institution of Oceanography, the Centro para la Biodiversidad Marina y la Conservación, and the Universidad Autonoma de Baja California Sur has participated in standardized underwater visual censuses across more than 240 rocky reef sites, generating survey records spanning 27 years. Each survey follows a rigorous protocol: trained divers record species identity, size class, and abundance along replicate belt transects at two depth strata (5 m and 20 m), while satellite-derived environmental covariates (sea surface temperature, chlorophyll-a concentration) are matched to each survey location and date.

This dataset is uniquely suited for automated hypothesis discovery for several reasons. First, its temporal depth (nearly three decades) captures multiple climate cycles, including the severe 2014-2016 marine heatwave that affected the entire Eastern Pacific. Second, the spatial coverage spans strong environmental gradients—from the warm, oligotrophic waters of the southern Gulf to the cooler, nutrient-rich upwelling zones in the north. Third, the region contains a mosaic of marine protected areas with varying protection regimes, from the strictly enforced no-take zone at Cabo Pulmo National Park to weakly regulated reserves and fully open-access reefs. This natural experimental design allows simultaneous exploration of climate, biogeographic, and management hypotheses (*1*).

This combination of long time series, strong environmental gradients, and contrasting protection regimes creates an ideal system for testing ecological predictions about climate vulnerability and conservation effectiveness. Yet the sheer volume of data—hundreds of potential variable combinations and temporal patterns—has meant that many hypotheses have never been systematically explored.

2 How we are using AutoDiscovery

We applied AutoDiscovery (*2*) in two contrasting modes to evaluate the role of domain knowledge in automated hypothesis generation.

Agnostic Discovery (Iteration 1). AutoDiscovery analyzed the dataset with no prior ecological knowledge beyond that in the underlying language model's parameters, treating variables without sophisticated reference. The system identified correlations and temporal patterns without reference to ecological theory.

Knowledge-Informed Discovery (Iteration 2). We provided AutoDiscovery with a structured domain knowledge document compiled by local reef ecologists. This included conceptual definitions (alpha/beta/gamma diversity, trophic levels, production-to-biomass ratios, size spectra), theoretical frameworks (the insurance hypothesis, biogeographic homogenization theory, latitudinal diversity gradients), and system-specific context (Gulf of California oceanography, known thermal anomalies, marine protected areas' history).

3 Initial Analysis and Refinement

Our initial analysis revealed a striking qualitative difference between iterations. The agnostic run produced 47 hypotheses dominated by simple categorical comparisons ("fish biomass differs by protection status") and direct correlations ("Sea Surface Temperature (SST) correlates with productivity"). While statistically valid, these lacked theoretical depth.

The knowledge-informed run generated 52 hypotheses that engaged directly with ecological theory—testing predictions about resilience mechanisms, biogeographic dynamics, and functional diversity. Critically, these hypotheses came with explicit expectations derived from ecological frameworks, allowing AutoDiscovery to identify not just patterns, but departures from theory: if the analysis "surprises" the system, i.e., its initial theory-based confidence ("prior") changes significantly after the analysis ("posterior"), the finding is flagged as potentially interesting for the user and explored further.

We classified results into two categories based on Bayesian belief updates:

Surprisingly confirmed hypotheses (*posterior belief increased*):

- Species richness declines with warming (1.2 species per °C; prior 0.16 → posterior 0.71)
- Community mean trophic level has declined over time (prior 0.74 → posterior 0.90)
- Functional redundancy buffers year-to-year compositional turnover (prior 0.67 → posterior 0.85)

Refuted hypotheses (*posterior belief decreased substantially*):

- Rich reefs are more resistant to disturbance (prior 0.69 → posterior 0.23): Species-rich reefs actually lost more species during the 2014-2016 heatwave
- Protected reefs have more stable communities (prior 0.74 → posterior 0.35): Cabo Pulmo shows higher year-to-year turnover than unprotected sites
- High spatial heterogeneity buffers thermal stress (prior 0.73 → posterior 0.25): Opposite pattern found

An important refinement emerged regarding spatial scale. AutoDiscovery found that homogenization is occurring **within** regions (confirmed), but when examining **all** reefs together, communities are actually diverging. This scale-dependency was not apparent from either hypothesis in isolation—it emerged from comparing results across multiple related hypotheses. We are now refining our analysis to explicitly partition beta diversity dynamics by spatial scale.

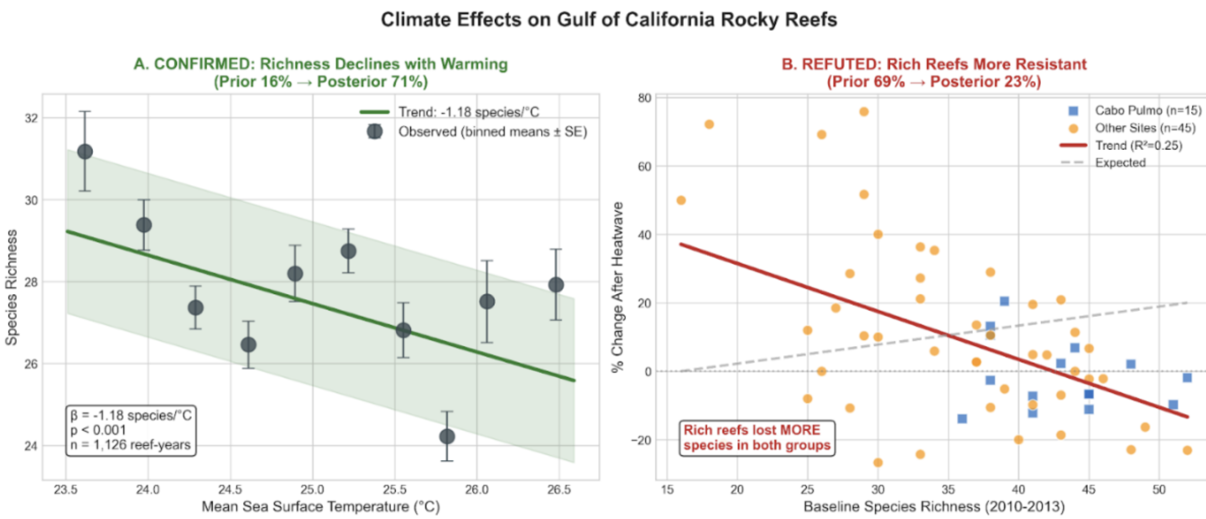

Figure 1: This finding provides a quantitative, empirically grounded estimate for projecting climate impacts on Gulf of California reef biodiversity. Under moderate warming scenarios (2–3 °C by the end of the century), reefs could lose an average of 1.2 species per transect, representing a substantial decline in local diversity (Figure 1A). The relationship is consistent across the full range of reefs in the dataset, suggesting it may generalize across the region. While the effects of warming on reef biodiversity are well documented, most previous work has focused on overall species loss, changes in dominant taxa, or recovery following discrete disturbance events (e.g. (*7*, *8*)). By contrast, the relationship between baseline species richness and the magnitude of biodiversity loss across continuous temperature gradients has received little explicit attention. Identifying this pattern requires both extensive, standardized biodiversity data spanning many reefs and an analytical approach that systematically evaluates a broad set of plausible relationships. AutoDiscovery enabled this by testing predictions without pre-selecting for expected outcomes, allowing a consistent but previously overlooked diversity–temperature relationship to emerge from the data.

4 In-Depth Finding: The richness vulnerability paradox

The most theoretically challenging result involves predictions about what confers resilience to climate disturbance. **The hypothesis:** Reefs with higher species richness should be more resistant to disturbance, losing a smaller proportion of their diversity during stress events. This prediction derives from the diversity-stability hypothesis—one of ecology's foundational ideas, supported by extensive theoretical and empirical work (see (*3*)).

AutoDiscovery assigned this hypothesis a moderate prior belief (69%), reflecting strong theoretical support but acknowledging mixed empirical results in marine systems. The insurance hypothesis (*4*) predicts that biodiversity buffers ecosystems against environmental fluctuations because species respond differently to perturbations. For example, a comprehensive study of 71,269 fish population time-series found that functional richness promotes community stability under marine heatwaves (*5*). However, empirical support is not universal. A synthesis of 41 field studies across 82 coral reefs found that species-rich regions were marginally less resistant to disturbance and did not recover more quickly (*6*). This context-dependence justified AutoDiscovery's moderate prior rather than high confidence.

During the 2014-2016 marine heatwave, the most severe thermal anomaly in our 27-year record, reefs with higher baseline richness (2010-2013) experienced larger proportional losses by 2017-2020 ($R^2 = 0.26$, $p < 0.01$). This pattern held for both protected (Cabo Pulmo) and unprotected sites, suggesting it is not an artifact of protection status.

This result does not contradict the diversity–stability hypothesis, but instead highlights its context dependence. Several non-exclusive mechanisms may explain the observed pattern. Species-rich reefs in this system may harbor a greater proportion of thermally sensitive, tropical-affinity species that live close to their upper thermal limits, with limited thermal safety margins (*9*) and low intraspecific variation in thermal tolerance (*10*), making them particularly vulnerable to extreme warming. Apparent losses may also reflect detection effects rather than true local extinctions, as species-rich assemblages include more rare or cryptic species that are prone to temporary nondetection during stress events; underwater visual census methods are known to underestimate species presence, especially for rare or behaviorally shy taxa (*11, 12*), and small populations are more likely to fall below detection thresholds during disturbance (*13*). In addition, highly diverse reefs may already operate closer to environmental or demographic limits, as theory predicts that diverse communities contain many rare species with small population sizes that are more susceptible to local extinction following disturbance (*14*), and may be closer to local carrying capacity, reducing their ability to buffer additional stress.

Importantly, this pattern emerged from systematic testing across hundreds of hypotheses by AutoDiscovery. Because the diversity–stability relationship is so well established, it might not have been examined in a traditional hypothesis-driven framework. AutoDiscovery theory-agnostic approach revealed it precisely by evaluating predictions without pre-filtering for expected outcomes. Ongoing analyses aim to disentangle these mechanisms by assessing whether responses are concentrated within particular functional groups and whether observed declines represent true local extinctions or temporary absences.

5 Looking forward

This analysis shows how systematic hypothesis exploration can both confirm long standing ecological ideas and surface unexpected patterns that push theory forward. The confirmed results strengthen the empirical basis for projecting climate impacts on reef biodiversity, while the refuted predictions, most notably the apparent richness vulnerability paradox, highlight where established frameworks may need refinement rather than rejection. We are now focusing on understanding the mechanisms behind these patterns, including whether observed losses reflect true local extinctions or temporary changes in detectability, why communities at Cabo Pulmo appear especially dynamic despite, or possibly because of, intact predator populations, and how protection status shapes climate vulnerability across functional groups. More broadly, AutoDiscovery functioned like having a room full of PhD ecologists brainstorming together, systematically exploring ideas that no single researcher, or even a small team, could reasonably test by hand. It did not replace ecological insight, it amplified it. The system surfaced patterns at scale, while domain expertise provided the context, skepticism, and biological sense making needed to interpret them and decide what questions to ask next. What ultimately makes this approach powerful is not just the hypotheses it confirms, but the assumptions it forces us to question, opening space for new theory, new analyses, and a more exploratory way of doing ecology in a rapidly changing world.

Acknowledgments

FF and OAO gratefully acknowledge the generous support of the Alumbra Foundation, the Builders Initiative, and the National Geographic Pristine Seas. The Rocky Reef Long-Term Monitoring Program is led by the Gulf of California Marine Program. We recognize and appreciate the invaluable contributions of the many researchers, students, and colleagues who have participated in this effort over the years. Their dedication and commitment have made this work possible. Readers interested in learning more about the program and related initiatives are encouraged to visit the DataMares platform, which showcases this and other monitoring programs. Finally, we sincerely thank the AutoDiscovery team at Ai2 for making this work possible.

References

1. F. Favoretto, C. Sanchez, O. Aburto-Oropeza, Warming and marine heatwaves tropicalize rocky reefs communities in the Gulf of California. *Progress in Oceanography* **206**, 102838 (2022).

2. D. Agarwal, *et al.*, AutoDiscovery: Open-ended Scientific Discovery via Bayesian Surprise, in *The Thirty-ninth Annual Conference on Neural Information Processing Systems* (2025).

3. T. H. Oliver, *et al.*, Biodiversity and resilience of ecosystem functions. *Trends in ecology & evolution* **30** (11), 673–684 (2015).

4. S. Yachi, M. Loreau, Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proceedings of the National Academy of Sciences* **96** (4), 1463–1468 (1999).

5. L. Benedetti-Cecchi, *et al.*, Marine protected areas promote stability of reef fish communities under climate warming. *Nature Communications* **15** (1), 1822 (2024).

6. I. M. Côté, E. S. Darling, Rethinking ecosystem resilience in the face of climate change. *PLoS biology* **8** (7), e1000438 (2010).

7. T. P. Hughes, *et al.*, Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. *Science* **359** (6371), 80–83 (2018).

8. N. A. Graham, S. Jennings, M. A. MacNeil, D. Mouillot, S. K. Wilson, Predicting climate-driven regime shifts versus rebound potential in coral reefs. *Nature* **518** (7537), 94–97 (2015).

9. L. Comte, J. D. Olden, Climatic vulnerability of the world's freshwater and marine fishes. *Nature Climate Change* **7** (10), 718–722 (2017).

10. J. Nati, *et al.*, Intraspecific variation in thermal tolerance differs between tropical and temperate fishes. *Scientific Reports* **11** (1), 21272 (2021).

11. G. J. Edgar, N. S. Barrett, A. J. Morton, Biases associated with the use of underwater visual census techniques to quantify the density and size-structure of fish populations. *Journal of Experimental Marine Biology and Ecology* **308** (2), 269–290 (2004).

12. M. A. MacNeil, *et al.*, Detection heterogeneity in underwater visual-census data. *Journal of Fish Biology* **73** (7), 1748–1763 (2008).

13. K. F. Kellner, R. K. Swihart, Accounting for imperfect detection in ecology: a quantitative review. *PloS one* **9** (10), e111436 (2014).

14. K. S. McCann, The diversity–stability debate. *Nature* **405** (6783), 228–233 (2000).