

₁ Supplemental Material: Data-Driven
₂ Discovery of Mechanistic Ecosystem Models
₃ with LLMs

₄

₅ October 17, 2025

6 S1 Curated Literature Collection

7 The local document collection used in this case study was carefully curated to
8 provide comprehensive coverage of marine ecosystem modeling approaches,
9 with particular focus on COTS-coral dynamics and management interven-
10 tions. The collection encompasses several key research areas:

- 11 • Ecosystem Modeling Frameworks: ¹² established foundational principles
12 for ecosystem approaches to fisheries, while ¹³ introduced Models of In-
13 termediate Complexity for Ecosystem assessments (MICE). ² explored
14 optimal model complexity levels.
- 15 • COTS Management and Ecology: ¹⁴ provided a comprehensive thirty-
16 year review of COTS research. ⁹ developed models for COTS outbreak
17 management, while ¹⁶ analyzed corallivore culling impacts under bleach-
18 ing scenarios.
- 19 • Ecological Regime Shifts: ¹ investigated predator-driven regime shifts
20 in marine ecosystems. ¹¹ provided insights into ecological tipping points
21 through ecosystem modeling.
- 22 • Management Interventions: ³ examined large-scale interventions on the
23 Great Barrier Reef. ¹⁵ explored harvest control implications using MICE
24 models.
- 25 • Model Application Guidelines: ⁵ provided critical guidelines for adapt-
26 ing ecosystem models to new applications. ⁷ demonstrated multispecies
27 production model applications for analyzing ecological and fishing ef-
28 fects.
- 29 • Integrated Systems: ⁸ and ¹⁰ explored integrated multi-trophic aquacul-
30 ture modeling, providing insights into coupled biological systems. ¹⁷ an-
31 alyzed trade-offs in seaweed farming between food production, liveli-
32 hoods, marine biodiversity, and carbon sequestration benefits.

33 These papers were selected based on their direct relevance to COTS pop-
34 ulation dynamics, coral reef ecology, and ecosystem modeling approaches.
35 The collection provided both specific parameter values and broader ecologi-
36 cal context for model development.

37 S2 RAG Architecture Implementation

38 The Retrieval-Augmented Generation (RAG) system facilitates parameter
39 search and extraction from scientific literature. The system employs two
40 primary search strategies: a local search of user-curated documents and a
41 comprehensive web search. For local search, the system uses ChromaDB as
42 a persistent vector store to maintain an indexed collection of scientific pa-
43 pers and technical documents specifically curated by research teams for their
44 ecological systems. These documents are processed into semantic chunks
45 of approximately 512 tokens with small overlaps to preserve context while
46 enabling precise retrieval of relevant information.

47 The parameter search process begins with the generation of enhanced se-
48 mantic descriptions for each parameter. These descriptions are crafted to im-
49 prove search relevance by capturing the ecological and mathematical context
50 in which the parameters are used. The system first searches the user-curated
51 local documents using embeddings generated through Azure OpenAI’s em-
52 bedding service. When necessary, it extends to web-based sources through
53 two channels: querying the Semantic Scholar database for highly-cited papers
54 in biology, mathematics, and environmental science, and conducting broader
55 literature searches through the Serper API to capture additional relevant
56 sources.

57 The search results from both local and web sources are processed through
58 an LLM to extract numerical values. The system applies consistent validation
59 across both search pathways, identifying minimum and maximum bounds,
60 ensuring unit consistency, and validating source reliability. When direct pa-
61 rameter values are not found in either the local collection or web sources,
62 the system defaults to the initial estimates from the coding LLM. All ex-
63 tracted information, including parameter values, valid ranges, and complete
64 citation details, is stored in a structured JSON database for reproducibility
65 and future reference.

66 The RAG system implements automatic retry mechanisms when initial
67 searches fail to yield usable results. Each retry attempt follows a structured
68 progression: first accessing the curated local collection through ChromaDB
69 queries, then expanding to Semantic Scholar for peer-reviewed literature, and
70 finally utilizing Serper API for broader scientific content. This progressive
71 broadening of scope, while maintaining focus on ecologically relevant sources,
72 ensures robust parameter estimation even in cases where direct measurements
73 are sparse in the literature.

74 S3 AI Prompts Used in Model Development

75 The development of the model relied on several carefully crafted prompts
76 to guide the artificial intelligence system. These prompts were designed to
77 ensure numerical stability, proper likelihood calculation, and clear model
78 structure. The following sections detail the exact prompts used at each stage
79 of model development.

80 S3.1 Initial Model Creation

81 The initial model creation utilized a comprehensive prompt that emphasized
82 three key aspects of model development. The prompt used for model initial-
83 ization was:

```
84 Please create a Template Model Builder model for the following topic
85 :[PROJECT_TOPIC]. Start by writing intention.txt, in which you
86 provide a concise summary of the ecological functioning of the
87 model. In model.cpp, write your TMB model with the following
88 important considerations:
89
90 1. NUMERICAL STABILITY:
91 - Always use small constants (e.g., Type(1e-8)) to prevent division
92 by zero
93 - Use smooth transitions instead of hard cutoffs in equations
94 - Bound parameters within biologically meaningful ranges using
95 smooth penalties rather than hard constraints
96
97 2. LIKELIHOOD CALCULATION:
98 - Always include observations in the likelihood calculation, don't
99 skip any based on conditions
100 - Use fixed minimum standard deviations to prevent numerical issues
101 when data values are small
102 - Consider log-transforming data if it spans multiple orders of
103 magnitude
104 - Use appropriate error distributions (e.g., lognormal for strictly
105 positive data)
106
107 3. MODEL STRUCTURE:
108 - Include comments after each line explaining the parameters (
109 including their units and how to determine their values)
110 - Provide a numbered list of descriptions for the equations
111 - Ensure all important variables are included in the reporting
112 section
113
```

```
114 | - Use ‘_pred’ suffix for model predictions corresponding to ‘_dat’
115 |     observations
116 |
```

117 S3.2 Parameter Enhancement

118 To enhance parameter descriptions for improved semantic search capabilities,
119 the following prompt was employed:

```
120 |
121 | Given a mathematical model about [PROJECT_TOPIC], enhance the
122 |     semantic descriptions of these parameters to be more detailed and
123 |     searchable. The model code shows these parameters are used in
124 |     the following way:
125 |
126 | [MODEL_CONTENT]
127 |
128 | For each parameter below, create an enhanced semantic search, no
129 |     longer than 10 words, that can be used for RAG search or semantic
130 |     scholar search.
131 |
```

132 S3.3 Model Improvement

133 For iterative model improvements, the system utilized this prompt:

```
134 |
135 | Improve the fit of the following ecological model by modifying the
136 |     equations in this TMB script. Only make ONE discrete change most
137 |     likely to improve the fit. Do not add stochasticity, but you may
138 |     add other ecological relevant factors that may not be present
139 |     here already.
140 |
141 | You may add additional parameters if necessary, and if so, add them
142 |     to parameters.json. Please concisely describe your ecological
143 |     improvement in intention.txt and then provide the improved model.
144 |     cpp and parameters.json content.
145 |
```

146 S3.4 Error Handling Prompts

147 For compilation errors, the system used this prompt:

```
148 |
149 | model.cpp failed to compile. Here’s the error information:
150 |
151 | [ERROR_INFO]
152 |
```

153 Do not suggest how to compile the script
154

155 For data leakage issues, the system employed this detailed prompt:

156 Data leakage detected in model equations. The following response
157 variables cannot be used to predict themselves:
158
159 To fix this:
160 1. Response variables ([RESPONSE_VARS]) must be predicted using only
161 :
162 - External forcing variables ([FORCING_VARS])
163 - Other response variables' predictions (_pred variables)
164 - Parameters and constants
165 2. Each response variable must have a corresponding prediction
166 equation
167 3. Use ecological relationships to determine how variables affect
168 each other
169
170 For example, instead of:
171 slow_pred(i) = slow * growth_rate;
172 Use:
173 slow_pred(i) = slow_pred(i-1) * growth_rate * (1 - impact_rate *
174 cots_pred(i-1));
175
176 Please revise the model equations to avoid using response variables
177 to predict themselves.
178
179

180 For numerical instabilities, the system used an adaptive prompt that
181 became progressively more focused on simplification after multiple attempts:

182 The model compiled but numerical instabilities occurred. Here's the
183 error information:
184
185 [ERROR_INFO]
186
187 [After 2+ attempts: Consider making a much simpler model that we can
188 iteratively improve later.]
189 Do not suggest how to compile the script
190
191

192 S3.5 NPZ Case Study - Recovering Equations

193 The model implementation can be compared to the original NPZ equations
194 from⁴:

$$\begin{aligned}
\frac{dN}{dt} &= \underbrace{-\frac{V_m NP}{k_s + N}}_{\text{nutrient uptake}} + \underbrace{\gamma(1 - \alpha) \frac{gP^2 Z}{k_g + P^2} + \mu_P P + \mu_Z Z^2}_{\text{recycling}} + \underbrace{S(N_0 - N)}_{\text{mixing}} \\
\frac{dP}{dt} &= \underbrace{\frac{V_m NP}{k_s + N}}_{\text{growth}} - \underbrace{\frac{gP^2 Z}{k_g + P^2}}_{\text{grazing loss}} - \underbrace{\mu_P P}_{\text{mortality}} - \underbrace{SP}_{\text{mixing}} \\
\frac{dZ}{dt} &= \underbrace{\alpha \frac{gP^2 Z}{k_g + P^2}}_{\text{growth (assimilation)}} - \underbrace{\mu_Z Z^2}_{\text{mortality}} - \underbrace{SZ}_{\text{mixing}}
\end{aligned}$$

Our generated model captures several key ecological processes from the original system:

1. Nutrient uptake by phytoplankton following Michaelis-Menten kinetics
2. Quadratic zooplankton mortality
3. Nutrient recycling through zooplankton excretion
4. Environmental mixing effects

For evaluating the ecological characteristics of generated models against the NPZ reference model, the system employed a 4-level ordinal scoring system that compares each model component to both the ground truth equations and recognized alternate formulations from the ecological literature. The evaluation system assessed nine ecological characteristics organized by equation: nutrient uptake, recycling, and mixing (dN/dt); phytoplankton growth, grazing loss, mortality, and mixing (dP/dt); and zooplankton growth and mortality (dZ/dt).

The scoring rubric used for all evaluations was:

Scoring rubric per characteristic (choose exactly one category):
- 3 = TRUTH_MATCH
The mathematical structure is equivalent to the TRUTH model (modulo variable names, syntax, factor grouping, and coefficient naming). Quote the exact snippet that matches.
- 2 = ALTERNATE
The implementation matches one of the alternates enumerated in the literature catalog,

```

220     even if not identical to TRUTH. Name the family (e.g., "
221     Michaelis-Menten uptake",
222     "Ivlev grazing with threshold", "linear mortality", "Droop quota
223     ").
224 - 1 = SIMILAR_NOT_LISTED
225     The implementation plays the same ecological role and is
226     mathematically similar
227     (e.g., another saturating curve or plausible closure) but is not
228     represented in TRUTH
229     or alternates list.
230 - 0 = NOT_PRESENT_OR_INCORRECT
231     The ecological component is missing or cannot be identified.
232

```

233 The alternate formulations catalog was based on⁶ and included:

- 234 • Phytoplankton light response: linear, saturating (Michaelis-Menten,
235 exponential, tanh), and photo-inhibiting forms
- 236 • Nutrient uptake: Michaelis-Menten, Liebig minimum limitation, Droop
237 quota models
- 238 • Zooplankton grazing: linear, saturating with threshold, Holling/Ivlev
239 type, acclimating forms
- 240 • Mortality terms: linear and quadratic (density-dependent) for both
241 phytoplankton and zooplankton

242 Each characteristic was assigned a weight based on its contribution to its
243 parent equation: the three nutrient equation components each had weight
244 0.333, the four phytoplankton components each had weight 0.25, and the two
245 zooplankton components each had weight 0.5. The aggregate ecological score
246 was calculated as the weighted sum of individual scores, then normalized to
247 a 0-1 scale by dividing by the maximum possible score.

248 **S3.5.1 Validation of Scoring System**

249 To validate the ecological characteristics scoring system, we tested it on the
250 ground truth NPZ model itself (evaluating the model against its own equa-
251 tions). This test confirmed that the scoring system could correctly identify
252 and score all nine ecological characteristics when they were present in their
253 canonical forms.

254 The validation results demonstrated perfect performance:

- 255 • All nine characteristics received scores of 3 (TRUTH_MATCH)

- Raw total score: 8.997 (out of maximum 9.0, with small rounding due to floating point arithmetic)
- Normalized total score: 1.0000 (perfect score on 0-1 scale)
- Zero extra components identified (correctly recognized model contained only canonical NPZ processes)

The LLM evaluator correctly identified each ecological mechanism in the ground truth model, providing detailed explanations such as “algebraically identical to the TRUTH NPZ model” and specifically noting the presence of “Michaelis-Menten style nutrient limitation multiplied by a light/self-shading term for phytoplankton growth” and “a saturating $P^2/(P^2+P^2)$ (Hill/Type-III-like) grazing formulation.” This validation confirmed that the scoring system could reliably distinguish between different levels of ecological fidelity, from exact matches to the ground truth through recognized alternates to novel formulations, providing a robust framework for assessing LEMMA-generated models.

S4 NPZ Validation

S4.1 Best Performing NPZ Model

S4.1.1 Model Description

The following model represents our framework’s attempt to recover the NPZ dynamics from⁴. The model aims to capture three key components:

- Nutrient uptake and recycling
- Phytoplankton growth and mortality
- Zooplankton predation and dynamics

S4.1.2 Model Intention

```
\section{Ecological Intention}

A key modification was made to incorporate direct nutrient recycling
from zooplankton grazing activity. In marine systems,
zooplankton feeding is often inefficient, with a significant
portion of consumed phytoplankton being released as dissolved
nutrients rather than being assimilated into biomass or entering
```

288 the detritus pool. This "sloppy feeding" process creates an
 289 important feedback loop where grazing can stimulate new primary
 290 production through rapid nutrient recycling.
 291
 292 The recycling efficiency is temperature-dependent, reflecting how
 293 metabolic rates and feeding mechanics vary with temperature. This
 294 creates an adaptive feedback where warmer conditions lead to
 295 both increased grazing pressure and faster nutrient recycling,
 296 better capturing the coupled nature of predator-prey interactions
 297 in planktonic systems.
 298
 299 The modification introduces a direct pathway from grazing to
 300 dissolved nutrients, complementing the slower recycling through
 301 the detritus pool. This better represents the multiple timescales
 302 of nutrient cycling in marine food webs and helps explain how
 303 high productivity can be maintained even under intense grazing
 304 pressure.
 305

306 S4.1.3 Model Implementation

³⁰⁷ **NPZ Model: Parameter and Equation Tables**

Symbol	Units	Meaning	Init. value	Bounds	Source	Literature (citekey)
log_mu_max	day ⁻¹ (log scale)	Log of maximum phytoplankton growth rate at reference conditions (day ⁻¹).	-0.02	[-0.22, 0.18]	literature	Yes (LitNPZ_log_mu_max)
log_K_N	g C m ⁻³ (log scale)	Log of half-saturation constant for nutrient uptake (g C m ⁻³).	-3.00	[-6.91, 0.00]	literature	Yes (LitNPZ_log_K_N)
I	W m ⁻²	Mean photosynthetically active irradiance proxy over the modeled period.	150.00	[0.00, 500.00]	initial estimate	No
log_K_I	W m ⁻² (log scale)	Log of light half-saturation constant for photosynthesis (W m ⁻²).	4.32	[0.00, 5.70]	literature	Yes (LitNPZ_log_K_I)
log_g_max	day ⁻¹ (log scale)	Log of maximum zooplankton grazing rate per unit Z biomass (day ⁻¹).	-0.69	[-3.00, 0.69]	literature	Yes (LitNPZ_log_g_max)
log_K_G	g C m ⁻³ (log scale)	Log of P half-saturation constant for grazing functional response (g C m ⁻³).	-2.30	[-6.91, 0.00]	literature	Yes (LitNPZ_log_K_G)
h_grazing	dimensionless	Holling type III shape exponent ($h \geq 1$).	2.00	[1.00, 3.00]	literature	Yes (LitNPZ_h_grazing)
logit_e_Z	dimensionless (logit scale)	Logit of zooplankton assimilation efficiency ($e_Z \in (0, 1)$); $e_Z = 0.5$ at value 0.	0.00	—	literature	Yes (LitNPZ_logit_e_Z)
log_m_P	day ⁻¹ (log scale)	Log of phytoplankton linear mortality rate (day ⁻¹).	-3.00	[-6.91, -1.20]	literature	Yes (LitNPZ_log_m_P)
log_m_Z	day ⁻¹ (log scale)	Log of zooplankton linear mortality rate (day ⁻¹).	-3.51	[-6.91, -1.20]	literature	Yes (LitNPZ_log_m_Z)
log_gamma_Z	(g C m ⁻³) ⁻¹ day ⁻¹ (log scale)	Log of zooplankton quadratic self-limitation coefficient ((g C m ⁻³) ⁻¹ day ⁻¹).	-4.61	[-9.21, -1.61]	initial estimate	No
logit_r_P	dimensionless (logit scale)	Logit of fraction of P mortality that is remineralized to N (0..1).	0.85	—	literature	Yes (LitNPZ_logit_r_P)
logit_r_Z	dimensionless (logit scale)	Logit of fraction of Z mortality that is remineralized to N (0..1).	0.85	—	literature	Yes (LitNPZ_logit_r_Z)
log_ex_Z	day ⁻¹ (log scale)	Log of zooplankton excretion rate to nutrients (day ⁻¹).	-4.61	[-13.82, -1.61]	initial estimate	No
log_k_mix	day ⁻¹ (log scale)	Log of vertical mixing rate driving nutrients toward N_* (day ⁻¹).	-3.91	[-13.82, -0.69]	initial estimate	No
N_*	g C m ⁻³	Deep/source nutrient concentration towards which mixing relaxes the system.	0.30	[0.00, 2.00]	initial estimate	No
log_q10	dimensionless (log scale)	Log of Q10 temperature scaling factor (dimensionless), typical $Q_{10} \approx 2$.	0.66	[0.61, 0.71]	literature	Yes (LitNPZ_log_q10)
T_C	deg C	Ambient temperature used for Q10 scaling (deg C).	15.00	[0.00, 35.00]	initial estimate	No
T_ref	deg C	Reference temperature for Q10 scaling (deg C).	15.00	[0.00, 35.00]	literature	Yes (LitNPZ_T_ref)
log_k_rem	day ⁻¹ (log scale)	Log of detritus remineralization rate to nutrients (day ⁻¹).	-2.30	[-4.61, 0.00]	conceptual addition	No
log_k_sink	day ⁻¹ (log scale)	Log of detritus sinking/export rate out of mixed layer (day ⁻¹).	-4.61	[-13.82, 0.00]	conceptual addition	No
log_sigma_N	log-scale SD	Log of observation SD for N on the log scale.	-2.30	[-5.00, 2.00]	initial estimate	No
log_sigma_P	log-scale SD	Log of observation SD for P on the log scale.	-2.30	[-5.00, 2.00]	initial estimate	No
log_sigma_Z	log-scale SD	Log of observation SD for Z on the log scale.	-2.30	[-5.00, 2.00]	initial estimate	No

309 S5 CoTS Model Convergence

310 S5.1 Model Evolution and Convergence

311 The evolutionary process exhibited consistent refinement across generations,
 312 with measurable improvements in model performance. On average, popu-
 313 lations reached their best-performing individual within 6.9 generations, and
 314 the mean improvement frequency across all populations was 38.0%. Figure
 315 1 shows the distribution of successful, culled, and broken models across gen-
 316 erations. Notably, two populations achieved convergence below the target
 317 threshold, representing 9.5% of all populations. Performance varied signif-
 318 icantly across populations. The fastest-converging population reached an
 319 optimal objective value of 0.0035 in just 3 generations, while others required
 320 up to 13 generations. This population also demonstrated a high improvement
 321 rate of -0.655 and a consistent improvement frequency of 50%. In contrast,
 322 several populations showed minimal or no improvement, with some failing to
 323 converge within the allotted iterations.

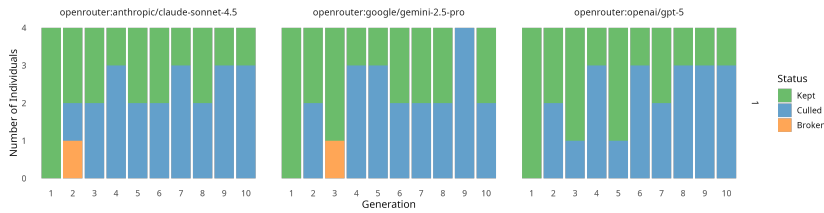


Figure 1: Evolution of model performance during the genetic algorithm optimization process. Each generation represents an iteration of model development, where models are evaluated and classified into three categories: the best performers according to the NMSE objective value (kept, green), those that are numerically stable but outcompeted (culled, blue), and those that failed due to numerical instability, data leakage, or syntax errors (broken, orange). The vertical axis shows the count of new models in each category per generation, while rows represent independent replicates using different LLM configurations. Gemini-2.5-Pro was included in the analysis but did not produce successful runs for some populations.

324 S5.2 Numerical Stability and Optimization

325 Numerical stability varied across LLM configurations, with runtime and gen-
 326 eration time metrics reflecting differences in optimization efficiency. The
 327 GPT-5 configuration showed moderate efficiency, with an average generation
 328 time of 12.0 minutes (SD = 13.0). The Claude Sonnet 4.5 configuration had

329 longer generation times, averaging 71.2 minutes ($SD = 155.2$), though this
 330 includes variability from a small number of outlier populations. In contrast,
 331 the Gemini-2.5-Pro configuration demonstrated the fastest generation cycles,
 332 averaging 4.1 minutes per generation ($SD = 0.54$), though it exhibited lower
 333 convergence rates and higher instability in some cases. Figure 2 illustrates
 334 the distribution of iteration counts required for successful model convergence
 335 across LLMs. Most models converged within 4 to 7 iterations, with some
 336 outliers requiring up to 11 iterations.

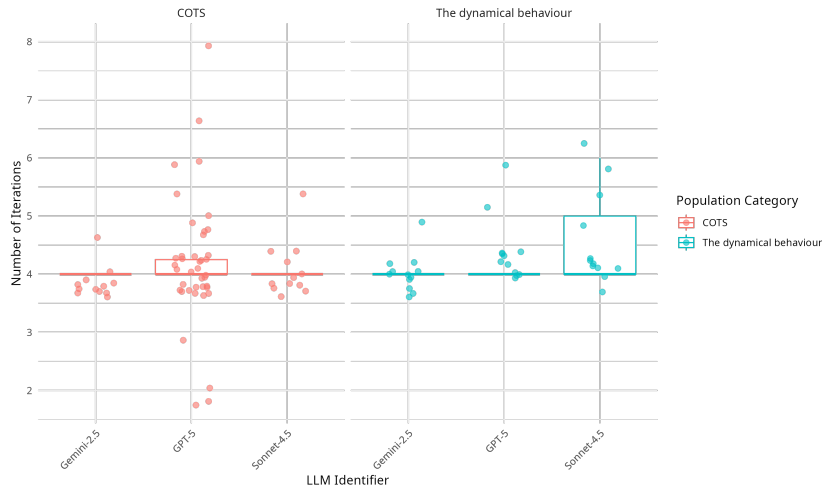


Figure 2: Distribution of iteration counts for successful model instances by LLM configuration. The boxplot excludes cases that reached maximum iterations or remained numerically unstable.

337 S6 Comparative Analysis of Best-Performing 338 Models

339 Before presenting the full code for each model, we analyze the key differ-
 340 ences between the best-performing models to understand their ecological ap-
 341 proaches and mathematical structures.

Table 1: Detailed Ecological Mechanisms

Mechanism	Human	Gemini 2.5 Pro	GPT-5	Sonnet 4.5
COTS state structure / life history	<i>Age-structured (3 classes):</i> $N_{t+1,1} = N_{t,0}e^{-M_0}$; $N_{t+1,2} = N_{t,1}e^{-fM_1} + N_{t,2}e^{-fM_2}$; recruits form $N_{t+1,0}$. Age-specific mortality $M_a = M_{\text{cots}} + \lambda/(1+a)$.	<i>Single compartment C_t</i> (no explicit ages). Growth from consumption + modifiers.	<i>Two-stage</i> (juveniles J and adults C): maturation $m_J J$; juvenile mortality $\mu_J J$; adult mortality $(\mu_C + \gamma_C C)C$.	<i>Single compartment</i> with logistic-like adult growth plus additive recruitment pulse; no age classes.
COTS stock–recruitment	<i>Beverton–Holt (BH)</i> from spawners: $R_{t+1} = \frac{\alpha(N_{t+1,2}/K_{\text{sp}})}{\beta + (N_{t+1,2}/K_{\text{sp}})}$, with α, β derived from slope steepness h and R_0 .	<i>Not BH</i> : recruitment implicit via $(e_F \cdot \text{Cons}_F + e_S \cdot \text{Cons}_S) \times \text{Allee} \times \text{temperature Gaussian}$ (no explicit SR function).	<i>BH-like taper on adults</i> : $\text{Rec} = \alpha_{\text{rec}} [C^\phi / (1 + C/C_{\text{sat}})] \cdot f_{\text{Allee}}(C) \cdot f_T(\text{SST}) \cdot f_{\text{food}} + \text{imm}_{t-1}$.	<i>Additive pulse recruitment</i> : $\text{recruit_pulse} = R_{\text{max}} \cdot \sigma(\text{fav} - \tau) \cdot \text{fav}$, added to adult change; adult growth also logistic-like.
COTS immigration & interannual pulses	<i>Background immigration</i> + year-specific deviations: $N_{t+1,0} = (R_{t+1} + \text{Imm}_{\text{COTS}}) \exp(\varepsilon_{t+1} + \sigma^2/2)$; ε_t applied in specified years.	Exogenous additive series $\text{cotsimm}(t)$ added directly to C_t dynamics.	Exogenous series cotsimm_{t-1} added to recruitment (lagged).	Used twice: (i) normalized within favorability index, and (ii) multiplicative $\text{immigration_boost} = 1 + \text{effect} \cdot \text{imm}$ on adult growth.
COTS mortality and its drivers	<i>Baseline + age-dependent</i> , modulated by coral availability: effective multiplier $f = (1 - \tilde{p}) + \tilde{p} \cdot \rho$ (see prey switching ρ); applied to age-1 and age-2+ mortality; mortality reduced when fast coral abundant.	<i>No age structure</i> ; mortality = survival Allee $\frac{m_{C,\text{max}} C}{1 + C/A_{\text{mort}}} + \text{quadratic density dependence } m_{C,\text{dd}} C^2$; not explicitly coral-modulated.	<i>No age structure</i> ; adult mortality $(\mu_C + \gamma_C C)C$; not explicitly coral-modulated.	<i>No age structure</i> ; baseline + density-dependent mortality, amplified by starvation: multiplier $1 + 2e^{-(F+S)/5}$ increases mortality when coral scarce.

Continued on next page

Detailed Ecological Mechanisms — continued

Mechanism	Human	Gemini 2.5 Pro	GPT-5	Sonnet 4.5
Prey switching / diet preference	<i>Abundance-driven switching:</i> $\rho = \exp(-\text{switchSlope} \cdot F/K)$. Predation share on fast coral $\propto (1 - \rho)$, on slow $\propto \rho$. As fast coral increases, switch toward fast prey.	Implicit via multi-prey Type II with separate a_F, a_S ; no explicit ρ rule.	Implicit via Type II/III blend (exponents η_F, η_S create low-prey refuge); no explicit ρ .	<i>Preference + availability:</i> weight combines fixed preference with fast-coral proportion (soft switching toward abundant prey).
Functional response (COTS \rightarrow coral consumption)	<i>Sigmoid-saturated vs COTS density:</i> $Q_F = F(1 - \frac{N_{1+2}}{N_{1+2} + p_{1F}})$ $Q_M = \frac{N_{1+2}}{1 + \exp(-N_{1+2}/p_{2F})}$ $S \rho p_{1M} \frac{N_{1+2}}{1 + \exp(-N_{1+2}/p_{2M})}$.	<i>Multi-prey Holling Type II:</i> per-capita $\frac{a_F F}{1 + a_F h F + a_S h S}$, $\frac{a_S S}{1 + a_F h F + a_S h S}$; totals scale with C .	<i>Type II/III blend:</i> $\text{Cons}_F = \frac{a_F F^{\eta_F} C}{1 + h(a_F F^{\eta_F} + a_S S^{\eta_S})}$, $\text{Cons}_S = \frac{a_S S^{\eta_S} C}{1 + h(a_F F^{\eta_F} + a_S S^{\eta_S})}$.	<i>Type II per prey</i> with separate handling; then <i>preference weighting</i> mixes fast vs slow consumption.
Coral intrinsic growth & space competition	<i>Logistic regrowth with shared space:</i> $F_{t+1} = F_t \left[1 + \rho_F(\text{SST}) r_f \left(1 - \frac{F_t + S_t}{K} \right) \right] - Q_F - M_{\text{ble}, F}$; analogous for S .	<i>Logistic regrowth</i> with shared K_{coral} ; losses: predation + bleaching mortality.	<i>Logistic regrowth</i> with shared K_{tot} ; losses: predation + heat-related terms.	<i>Logistic regrowth</i> with shared K_{coral} ; losses: predation + temperature-stress mortality.

Continued on next page

Detailed Ecological Mechanisms — continued

Mechanism	Human	Gemini 2.5 Pro	GPT-5	Sonnet 4.5
SST modulation of coral regrowth (performance curve)	<i>Gaussian performance multiplier:</i> $\rho_F(\text{SST}) = \exp\left[-\frac{(\text{SST} - \text{SST}0_f)^2}{2 \text{SST_sig}_f^2}\right]$, $\rho_M(\text{SST}) = \exp\left[-\frac{(\text{SST} - \text{SST}0_m)^2}{2 \text{SST_sig}_m^2}\right]$.	Not Gaussian on growth; SST enters via <i>logistic bleaching mortality</i> (see below).	<i>Heat-stress growth reduction:</i> multiplier $\exp[-\beta_{\text{bleach}} \max(0, \text{SST} - T_{\text{bleach}})]$ on growth (non-Gaussian).	<i>Threshold stress loss:</i> linear mortality above T_{stress} ; no Gaussian growth multiplier.
Coral bleaching mortality	<i>Logistic bleaching:</i> $M_{\text{ble},F} = F \cdot [1 + \exp\{-\eta_f(\text{SST} - M_{50,F})\}]^{-1}$; analogous for slow coral (impulse option commented).	<i>Logistic bleaching mortality:</i> $m_{F,\text{sst}}/[1 + \exp\{-k_{\text{bleach}}(\text{SST} - T_{\text{bleach},F})\}]$; applied proportionally to F ; analogous for S .	Two components: (i) multiplicative <i>growth reduction</i> under heat, (ii) <i>additional linear loss</i> $m_{\text{bleach}} \cdot \text{heat} \cdot \text{coral}$; no logistic 50% curve.	<i>Temperature-stress mortality</i> proportional to degrees above threshold; no explicit logistic bleaching curve.
Coral carrying capacity (space sharing)	<i>Shared K</i> via $1 - (F + S)/K$ in both coral equations.	<i>Shared K_{coral}</i> in coral logistic growth.	<i>Shared K_{tot}</i> in coral logistic growth.	<i>Shared K_{coral}</i> in coral logistic growth.

Table 2: Model Parameterisation Comparison

Human (name)	Role / Units	Gemini 2.5 Pro	OpenAI GPT-5	Claude Sonnet 4.5
Mcots	Baseline instantaneous COTS mortality (yr^{-1})	$m_{C,\text{max}}$ (low-density mortality scale); $m_{C,\text{dd}}$ (DD mortality)	μ_C (baseline adult mortality), γ_C (DD mortality)	<code>log_mort_base</code> (baseline), <code>log_mort_density</code> (DD)
lam	Age-dependence of mortality, $M_a = M_{\text{cots}} + \lambda/(1+a)$	— (no age structure)	— (two-stage but no age-specific M)	— (no age structure)
p _{til}	Fraction of M attributable to fast-coral availability (mortality modulation)	— (no coral-modulated M)	— (no coral-modulated M)	Starvation multiplier $1 + 2e^{-(F+S)/5}$ (different form)
h (BH steepness)	Shapes SR via $(R_0, h) \rightarrow \alpha, \beta$	— (no BH SR; growth from consumption)	$C_{\text{sat,rec}}$ (BH-like taper), $\alpha_{\text{rec}}, \phi$ (fecundity exponent)	Recruitment pulse parameters: <code>log_recruit_max</code> , <code>recruit_threshold</code>
R0	Recruitment at unexploited state (for SR derivation)	— (no explicit R_0)	α_{rec} (scales juvenile input; closest analogue)	<code>log_recruit_max</code> (caps pulse magnitude; different structure)
Imm_CoTS	Background immigration ($\text{ind m}^{-2} \text{yr}^{-1}$)	<code>cotsimm_dat</code> (t) (added to adults each step)	<code>cotsimm_dat</code> ($t-1$) (added into recruitment)	<code>cotsimm_dat</code> used in favorability and as growth boost
sigCoTS	SR process variability (lognormal on recruits)	— (no process noise term; observation SD only)	— (no explicit process noise on SR; observation SDs)	— (no explicit process noise on SR; observation SDs)
immigration	Year-specific recruitment deviations (vector by year)	Represent via time series in <code>cotsimm_dat</code>	Represent via time series in <code>cotsimm_dat</code>	Represent via time series in <code>cotsimm_dat</code>
COTS_init	Initial COTS abundance (age-2+)	Init from first data row (no param)	$C0$ (adults), $J0$ (juveniles)	Init from first data row (no param)
p _{lf} , p _{lm}	Per-COTS coral loss coefficients (fast/slow), units of $\% \text{cover} \cdot (\text{ind}^{-1} \text{yr}^{-1})$	a_F, a_S (attack), h (handling) with efficiencies e_F, e_S for COTS growth (coral loss \propto consumption)	a_F, a_S (attack), h (handling), q_F, q_S (loss efficiencies) (Type II/III via η_F, η_S)	<code>log_attack_fast/slow</code> , <code>log_handling_fast/slow</code> , plus preference weighting

Continued on next page

Model Parameterisation Comparison — continued				
Human (name)	Role / Units	Gemini 2.5 Pro	OpenAI GPT-5	Claude Sonnet 4.5
p2f, p2m	Logistic saturation vs COTS density (shape of Q_F, Q_M)	Denominator $1 + a_F h F + a_S h S$ (resource-based saturation; no explicit p2)	Denominator $1 + h(a_F F^{\eta_F} + a_S S^{\eta_S})$ (no explicit p2)	Separate Type-II denominators per prey (no explicit p2)
switchSlope	Controls prey switching: $\rho = \exp(-\text{switchSlope} \cdot F/K)$	— (no explicit ρ)	η_F, η_S (Type-III curvature; soft switching), no ρ	preference_fast blended with fast-coral share (soft switching)
K	Shared coral carrying capacity (% cover scale)	K_{coral} (log_K_coral parameterized)	K_{tot}	K_{coral} (log_K_coral)
rf, rm	Intrinsic regrowth (fast/slow), yr^{-1}	r_F, r_S (log_r_F, log_r_S)	rF, rS	$r_{\text{fast}}, r_{\text{slow}}$ (log_r_fast, log_r_slow)
Cf_init, Cm_init	Initial coral state (fraction of K)	Init from first data row (no param)	$F0, S0$	Init from first data row (no param)
Eta_f, Eta_m	Logistic bleaching slope (fast/slow)	k_{bleach} (common steepness)	β_{bleach} (growth reduction) and $m_{\text{bleachF/S}}$ (linear loss)	temp_stress_rate (linear loss; no logistic)
M_SST50_f, M_SST50_m	50% bleaching SST (fast/slow)	$T_{\text{bleach}, F/S}$	$T_{\text{opt}, \text{bleach}}$ (single threshold)	temp_stress_threshold (single threshold)
Ble_imp_f, Ble_imp_m	Optional impulse bleaching toggles	—	—	—
SST0_f, SST0_m	Coral Gaussian performance optima (fast/slow)	— (no Gaussian growth multiplier; bleaching only)	$T_{\text{opt}, \text{bleach}}$ used in heat reduction (not Gaussian peak)	— (no Gaussian growth multiplier)
SST_sig_f, SST_sig_m	Coral Gaussian performance widths (fast/slow)	—	—	—
—	—	$T_{\text{opt}, \text{rec}}, \beta_{\text{rec}}$ (Gaussian on recruitment)	log_temp_opt, log_temp_width $T_{\text{opt}, \text{cots}}, T_{\sigma, \text{cots}}$ (Gaussian in favorability & adult growth)	$T_{\text{opt}, \text{cots}}, T_{\sigma, \text{cots}}$ (Gaussian on COTS reproduction)

Continued on next page

Model Parameterisation Comparison — continued				
Human (name)	Role / Units	Gemini 2.5 Pro	OpenAI GPT-5	Claude Sonnet 4.5

342 COTS–Coral Model: Parameter and Equation 343 Tables

344 Implementation details (optimization stability)

- 345 • **Soft bounds:** Parameter ranges are enforced via a smooth quadratic
346 penalty with weight $w_{\text{pen}} = 10^{-3}$ (not hard constraints).
- 347 • **Non-negativity and % clamping:** States use a smooth positive-part
348 function for $x_+ \approx \max(0, x)$; coral % is smoothly clamped to $[0, 100]$.
- 349 • **Logit transform safety:** The logit of % cover uses a small ϵ to avoid
350 0/100 singularities.
- 351 • **Likelihood SD floors:** Observation SDs use a floor (≥ 0.05) for nu-
352 merical stability.
- 353 • **Food term default:** If no external driver is provided, $f_{\text{food}} = 1$ (neu-
354 tral), leaving recruitment unaffected by food.

Symbol	Units	Meaning	Init. value	Bounds	Source	Citation
C0	ind m ⁻²	Initial adult COTS density at first time step	0.1	[0.0, 50.0]	initial estimate	No
J0	ind m ⁻²	Initial juvenile COTS pool at first time step	0.1	[0.0, 50.0]	initial estimate	No
F0	% cover	Initial fast coral (Acropora) cover at first time step	30.0	[0.0, 100.0]	initial estimate	No
S0	% cover	Initial slow coral (Faviidae/Porites) cover at first time step	30.0	[0.0, 100.0]	initial estimate	No
alpha_rec	ind m ⁻² yr ⁻¹	Recruitment productivity scaling controlling outbreak potential	1.0	[0.0, 10.0]	initial estimate	No
phi	dimensionless	Fecundity density exponent shaping recruitment curvature	1.5	[1.0, 3.0]	initial estimate	No
k_allee	m ² ind ⁻¹	Steepness of smooth Allee effect on recruitment	2.0	[0.01, 20.0]	initial estimate	No
C_allee	ind m ⁻²	Allee density where mating success increases rapidly	0.2	[0.0, 5.0]	initial estimate	No
C_sat_rec	ind m ⁻²	Adult density scale for stock-recruitment taper (Beverton-Holt)	2.0	[0.01, 50.0]	improvement	No
muC	yr ⁻¹	Baseline adult COTS mortality rate	0.6	[0.0, 3.0]	initial estimate	No
gammaC	m ² ind ⁻¹ yr ⁻¹	Density-dependent mortality coefficient	0.5	[0.0, 10.0]	initial estimate	No
mJ	yr ⁻¹	Annual maturation fraction from juvenile to adult	0.5	[0.0, 1.0]	initial estimate	No
muJ	yr ⁻¹	Annual proportional mortality of juvenile COTS	0.5	[0.0, 1.0]	initial estimate	No
T_opt_rec	degC	Optimal SST for COTS recruitment success	26.5	[20.0, 34.0]	literature	Yes ^{3,9,14}
beta_rec	degC ⁻²	Curvature of Gaussian temperature effect on recruitment	0.2	[0.0, 2.0]	initial estimate	No
K_food	units of food_dat	Half-saturation constant for larval food limitation on recruitment	1.0	[0.001, 1000.0]	improvement	No
T_opt_bleach	degC	SST threshold where bleaching stress starts impacting coral	32.65	[20.0, 34.0]	literature	Yes ^{3,16}
beta_bleach	dimensionless	Multiplier controlling growth reduction under heat stress	0.5	[0.0, 5.0]	initial estimate	No
m_bleachF	yr ⁻¹ degC ⁻¹	Additional proportional loss of fast coral per °C above threshold	0.2	[0.0, 2.0]	initial estimate	No
m_bleachS	yr ⁻¹ degC ⁻¹	Additional proportional loss of slow coral per °C above threshold	0.1	[0.0, 2.0]	initial estimate	No
rF	yr ⁻¹	Intrinsic regrowth rate of fast coral	0.5	[0.0, 2.0]	literature	Yes ^{3,14}
rS	yr ⁻¹	Intrinsic regrowth rate of slow coral	0.2	[0.0, 2.0]	literature	Yes ^{3,14}
K_tot	% cover	Total carrying capacity for combined coral cover	70.0	[10.0, 100.0]	literature	Yes ^{3,16}
aF	yr ⁻¹ % ^{-η_F} m ² ind ⁻¹	Encounter/attack parameter on fast coral	0.02	[0.0, 1.0]	initial estimate	No
aS	yr ⁻¹ % ^{-η_S} m ² ind ⁻¹	Encounter/attack parameter on slow coral	0.01	[0.0, 1.0]	initial estimate	No
etaF	dimensionless	Shape exponent for fast coral	1.5	[1.0, 3.0]	initial estimate	No
etaS	dimensionless	Shape exponent for slow coral	1.2	[1.0, 3.0]	initial estimate	No
h	yr % ⁻¹	Handling/satiation scaler controlling saturation	0.02	[0.0, 1.0]	initial estimate	No
qF	dimensionless	Efficiency converting fast coral feeding into % cover loss	0.8	[0.0, 1.0]	literature	Yes ¹⁶
qS	dimensionless	Efficiency converting slow coral feeding into % cover loss	0.5	[0.0, 1.0]	literature	Yes ¹⁶
sigma_cots	log-space SD	Observation/process error SD for COTS	0.5	[0.01, 2.0]	initial estimate	No
sigma_fast	logit-space SD	Observation/process error SD for fast coral	0.3	[0.01, 2.0]	initial estimate	No
sigma_slow	logit-space SD	Observation/process error SD for slow coral	0.3	[0.01, 2.0]	initial estimate	No

Process and observation equations

25

Name	Equation	Description
Allee effect	$f_{\text{Allee}}(C) = \frac{1}{1 + \exp\{-k_{\text{Allee}}(C - C_{\text{Allee}})\}}$	Smooth logistic Allee effect on recruitment
Recruitment temperature	$f_{T,\text{rec}}(\text{SST}) = \exp\{-\beta_{\text{rec}}(\text{SST} - T_{\text{opt},\text{rec}})^2\}$	Gaussian peak temperature modifier
Food modifier	$f_{\text{food}} = \frac{\text{food}}{K_{\text{food}} + \text{food}} \quad (= 1 \text{ if no driver})$	Monod saturation; neutral if driver missing
Stock term	$s(C) = \frac{C^\phi}{1 + C/C_{\text{sat},\text{rec}}}$	Beverton–Holt taper at high adult density
Juvenile recruitment	$\text{Rec}_t = \alpha_{\text{rec}} s(C_{t-1}) f_{\text{Allee}}(C_{t-1}) f_{T,\text{rec}}(\text{SST}_{t-1}) f_{\text{food}} + \text{IMM}_{t-1}$	Recruitment with environmental modifiers and immigration
Adult mortality	$\text{Mort}_{\text{adult}} = (\mu_C + \gamma_C C) C$	Baseline + density-dependent adult mortality
Juvenile flows	$\text{Mat} = m_J J, \quad \text{Mort}_{\text{juv}} = \mu_J J$	Maturation and juvenile mortality
Adult update	$C_t = \max\{0, C_{t-1} + \text{Mat} - \text{Mort}_{\text{adult}}\}$	Nonnegative adult state update
Juvenile update	$J_t = \max\{0, J_{t-1} + \text{Rec}_t - \text{Mat} - \text{Mort}_{\text{juv}}\}$	Nonnegative juvenile state update
Heat stress	$\text{heat} = \max\{0, \text{SST} - T_{\text{opt},\text{bleach}}\}$	Bleaching temperature exceedance
Bleach growth mod	$g_{\text{bleach}} = \exp\{-\beta_{\text{bleach}} \text{heat}\}$	Growth reduction under heat stress
Space limitation	$\text{space} = 1 - \frac{F+S}{K_{\text{tot}}}$	Shared coral space carrying capacity
Coral growth	$G_F = r_F F \text{space} g_{\text{bleach}}, \quad G_S = r_S S \text{space} g_{\text{bleach}}$	Fast/slow coral intrinsic regrowth
Bleaching losses	$B_F = m_{\text{bleach}F} \text{heat} F, \quad B_S = m_{\text{bleach}S} \text{heat} S$	Additional proportional bleaching losses
Functional response denom	$\text{denom} = 1 + h(a_F F^{\eta_F} + a_S S^{\eta_S})$	Type II/III multi-prey denominator
Feeding losses	$\text{Cons}_F = q_F \frac{a_F F^{\eta_F} C}{\text{denom}}, \quad \text{Cons}_S = q_S \frac{a_S S^{\eta_S} C}{\text{denom}}$	COTS consumption of fast/slow coral
Coral updates	$F_t = F + G_F - \text{Cons}_F - B_F, \quad S_t = S + G_S - \text{Cons}_S - B_S$	Fast/slow coral cover updates
Observation: COTS	$\log Y^{(C)} \sim \mathcal{N}(\log C, \sigma_{\text{cots}}^2)$	Lognormal observation with Jacobian in NLL
Observation: fast	$\text{logit}_{\%}(Y^{(F)}) \sim \mathcal{N}(\text{logit}_{\%}(F), \sigma_{\text{fast}}^2)$	Normal on logit-% cover
Observation: slow	$\text{logit}_{\%}(Y^{(S)}) \sim \mathcal{N}(\text{logit}_{\%}(S), \sigma_{\text{slow}}^2)$	Normal on logit-% cover

References

- [1] Blamey, L. K., Plagányi, E. E., and Branch, G. M. (2014). Was overfishing of predatory fish responsible for a lobster-induced regime shift in the Benguela? *Ecological Modelling*, 273:140–150.
- [2] Collie, J. S., Botsford, L. W., Hastings, A., Kaplan, I. C., Largier, J. L., Livingston, P. A., Plagányi, E., Rose, K. A., Wells, B. K., and Werner, F. E. (2016). Ecosystem models for fisheries management: finding the sweet spot. *Fish and Fisheries*, 17:101–125.
- [3] Condie, S. A., Anthony, K. R. N., Babcock, R. C., Baird, M. E., Beeden, R., Fletcher, C. S., Gorton, R., Harrison, D., Hobday, A. J., Plagányi, E. E., and Westcott, D. A. (2021). Large-scale interventions may delay decline of the Great Barrier Reef. *Royal Society Open Science*, 8:201296.
- [4] Edwards, A. M. and Brindley, J. (1999). Zooplankton mortality and the dynamical behaviour of plankton population models. *Bulletin of mathematical biology*, 61(2):303–339.
- [5] Essington, T. E. and Plagányi, E. E. (2014). Pitfalls and guidelines for “recycling” models for ecosystem-based fisheries management: evaluating model suitability for forage fish fisheries. *ICES Journal of Marine Science*, 71:118–127.
- [6] Franks, P. J. (2002). Npz models of plankton dynamics: their construction, coupling to physics, and application. *Journal of Oceanography*, 58(2):379–387.
- [7] Gamble, R. J. and Link, J. S. (2009). Analyzing the tradeoffs among ecological and fishing effects on an example fish community: A multispecies (fisheries) production model. *Ecological Modelling*, 220:2570–2582.
- [8] Hadley, S., Wild-Allen, K., Johnson, C., and Macleod, C. (2015). Modeling macroalgae growth and nutrient dynamics for integrated multi-trophic aquaculture. *Journal of Applied Phycology*, 27:901–916.
- [9] Morello, E. B., Plagányi, E. E., Babcock, R. C., Sweatman, H., Hillary, R., and Punt, A. E. (2014). Model to manage and reduce crown-of-thorns starfish outbreaks. *Marine Ecology Progress Series*, 512:167–183.
- [10] Oca, J., Cremades, J., Jiménez, P., Pintado, J., and Masaló, I. (2019). Culture of the seaweed *Ulva ohnoi* integrated in a *Solea senegalensis* recirculating system: influence of light and biomass stocking density on macroalgae productivity. *Journal of Applied Phycology*, 31(4):2461–2467.

- 392 [11] Plagányi, E., Ellis, N., Blamey, L., Morello, E., Norman-Lopez, A.,
 393 Robinson, W., Sporcic, M., and Sweatman, H. (2014a). Ecosystem mod-
 394 elling provides clues to understanding ecological tipping points. *Marine*
 395 *Ecology Progress Series*, 512:99–113.
- 396 [12] Plagányi, E. E. (2007). Models for an ecosystem approach to fisheries.
- 397 [13] Plagányi, E. E., Punt, A. E., Hillary, R., Morello, E. B., Thébaud, O.,
 398 Hutton, T., Pillans, R. D., Thorson, J. T., Fulton, E. A., Smith, A. D. M.,
 399 Smith, F., Bayliss, P., Haywood, M., Lyne, V., and Rothlisberg, P. C.
 400 (2014b). Multispecies fisheries management and conservation: tactical
 401 applications using models of intermediate complexity. *Fish and Fisheries*,
 402 15:1–22.
- 403 [14] Pratchett, M., Caballes, C., Wilmes, J., Matthews, S., Mellin, C., Sweat-
 404 man, H., Nadler, L., Brodie, J., Thompson, C., Hoey, J., Bos, A., Byrne,
 405 M., Messmer, V., Fortunato, S., Chen, C., Buck, A., Babcock, R., and
 406 Uthicke, S. (2017). Thirty Years of Research on Crown-of-Thorns Starfish
 407 (1986–2016): Scientific Advances and Emerging Opportunities. *Diversity*,
 408 9:41.
- 409 [15] Punt, A. E., MacCall, A. D., Essington, T. E., Francis, T. B., Hurtado-
 410 Ferro, F., Johnson, K. F., Kaplan, I. C., Koehn, L. E., Levin, P. S., and
 411 Sydeman, W. J. (2016). Exploring the implications of the harvest control
 412 rule for Pacific sardine, accounting for predator dynamics: A MICE model.
 413 *Ecological Modelling*, 337:79–95.
- 414 [16] Rogers, J. G. D. and Plagányi, E. E. (2022). Culling corallivores im-
 415 proves short-term coral recovery under bleaching scenarios. *Nature Com-*
 416 *munications*, 13:2520.
- 417 [17] Spillias, S., Cottrell, R. S., Layton, C., O’Brien, K. R., and McDonald-
 418 Madden, E. (2024). Having our kelp and eating it too: Minimizing trade-
 419 offs from seaweed farming. *Journal of Cleaner Production*, 448:141150.