

1 Automated Diet Matrix Construction for Marine
2 Ecosystem Models Using Generative AI

3 Scott Spillias^{1,2*} Beth Fulton^{1,2} Fabio Boschetti^{2,3}
4 Cathy Bulman¹ Rodrigo Bustamante⁴ Javier Porobic Garate^{1,2}
5 Joanna Strzelecki³ Roshni Subramaniam^{1,2} Rowan Trebilco^{1,2}

6 February 5, 2025

7 ¹CSIRO Environment, Hobart, Australia

8 ²Centre for Marine Socio-Ecology, University of Tasmania, Hobart, Australia

9 ³CSIRO Environment, IOMRC Crawley, Australia

10 ⁴CSIRO Environment, St. Lucia, Australia

11 **Abstract**

12 This study introduces and validates a novel AI-driven framework for automated
13 species grouping in Ecopath with Ecosim (EwE) ecosystem models, addressing a critical
14 bottleneck in model development. We evaluate the framework across three contrasting
15 Australian marine regions, processing over 41,000 species through multiple validation iterations.
16 The framework successfully condensed 63 potential functional groups into 34-36 region-specific groups, achieving high classification stability (>99%
17 consistency) for most species. Notably, the framework demonstrated robust performance in the South East Offshore region with only 0.03% inconsistent classifications,
18 while showing greater variability in complex tropical systems (1% inconsistent classifications). Higher trophic level species maintained consistent classifications across
19 all runs, with the framework identifying 235-327 significant predator-prey interactions per region at >70% consistency. This systematic validation reveals that while
20 the framework can reliably automate species grouping, its performance varies predictably with ecosystem complexity and data availability. These findings provide
21 quantitative evidence for the framework's capability to accelerate ecosystem model
22 development while highlighting specific conditions where additional validation may
23 be necessary. Our results demonstrate the potential for AI to significantly reduce
24 model development time while maintaining ecological validity, offering a practical
25 pathway to expand the application of ecosystem-based management across diverse
26 marine environments.

32 **1 Introduction**

33 Ecosystem modeling is a critical tool for understanding and managing complex marine
34 environments, with Ecopath with Ecosim (EwE) emerging as one of the most widely
35 used frameworks (Christensen and Walters, 2004; Colléter et al., 2015). EwE models

*Corresponding author: scott.spillias@csiro.au

36 provide quantitative insights into ecosystem structure and function, enabling researchers
37 to assess cumulative impacts of multiple stressors and support ecosystem-based fisheries
38 management (EBFM) decisions (Coll et al., 2015; Villasante et al., 2016). However,
39 constructing these models presents significant challenges, particularly in parameterizing
40 diet matrices that capture the complex web of trophic interactions within an ecosystem.

41 Traditional approaches to EwE model development rely heavily on extensive field
42 data collection and expert knowledge, which are time-consuming and resource-intensive
43 (Holden et al., 2024a). The process of assembling diet matrices is particularly challenging,
44 requiring synthesis of diverse data sources including field studies, literature reviews, and
45 expert opinion. This creates a significant bottleneck in model development, especially
46 when applying models to new geographical contexts (Holden et al., 2024b).

47 Recent advances in artificial intelligence (AI) offer new opportunities to streamline the
48 model development process. AI tools have demonstrated success in tasks such as species
49 distribution modeling and remote sensing applications (Lapeyrolerie et al., 2022; Tuia
50 et al., 2022), but their application to process-based ecosystem modeling remains limited
51 (Karniadakis et al., 2021). Recent developments in marine science demonstrate significant
52 potential, with specialized AI systems achieving high accuracy in critical areas such as wa-
53 ter quality prediction and biomass estimation (Fernandes and D'Mello, 2024). A notable
54 breakthrough comes from the LITE framework, which achieved a 41.25% reduction in
55 prediction error for environmental ecosystem modeling through an innovative multimodal
56 approach combining semantic time-series and temporal trend images (Li et al., 2024).
57 This success demonstrates AI's ability to handle real-world challenges like incomplete
58 data and distribution shifts in environmental applications. However, implementation bar-
59 rriers including technical expertise requirements and system reliability concerns must be
60 carefully addressed (Fernandes and D'Mello, 2024). The key challenge lies in ensuring that
61 AI-driven approaches can effectively synthesize available information while maintaining
62 ecological validity.

63 This study presents a novel framework for assembling and synthesizing local and online
64 resources to parameterize EwE diet matrices using AI. Our approach integrates multiple
65 data sources, including global biodiversity databases, species interaction repositories, and
66 literature-derived information, to automate key steps in model development. Building on
67 recent advances in marine-specific AI applications (Zheng et al., 2023), the framework
68 employs natural language processing and machine learning techniques to understand re-
69 gional ecosystem characteristics, group species into functional units, and estimate trophic
70 interactions. This approach aligns with emerging work demonstrating how AI systems
71 can effectively analyze and classify marine conservation approaches (Chen and Xu, 2024).

72 We validate this framework through three case studies representing distinct Australian
73 marine ecosystems: the Northern Territory, South East Inshore, and South East Offshore
74 regions. These regions offer contrasting environmental conditions, species assemblages,
75 and ecological dynamics, providing a robust test of the framework's adaptability and
76 reliability.

77 The primary objectives of this study are to:

- 78 1. Present a systematic, AI-assisted framework for assembling and parameterizing EwE
79 diet matrices
- 80 2. Validate key steps in the AI decision-making process, including:
 - 81 • Species grouping decisions and their ecological validity

- 82 • Resulting diet matrix values and their reliability

83 3. Assess the framework's applicability across different marine ecosystems

84 By rigorously validating this AI-assisted approach across multiple regions, we aim to
85 demonstrate its potential for accelerating ecosystem model development while maintaining
86 scientific rigor. Recent work by Kuhn et al. (2024) emphasizes how machine learning
87 applications in fisheries must span multiple scales, from genomics to ecosystem-level anal-
88 yses, while maintaining interpretability and transparency in automated decision-making
89 systems. This work contributes to the growing need for rapid, data-driven methodologies
90 in ecology (Kelling et al., 2009; Michener and Jones, 2012), while ensuring their outputs
91 align with established ecological principles and can effectively support ecosystem-based
92 management decisions. Our validation approach is informed by emerging standardization
93 efforts in environmental AI applications (Guo et al., 2025), addressing the critical need
94 for rigorous evaluation frameworks in ecological modeling.

95 **2 Methods**

96 **2.1 AI-Assisted Framework Overview**

97 The development of ecosystem models requires substantial time organizing species into
98 functional groups and determining their interactions. This framework automates these
99 tasks through integration of artificial intelligence with ecological databases. The frame-
100 work executes four sequential steps: species identification within a region, biological data
101 collection, functional group organization, and interaction determination (Figure 1).

102 The framework utilizes the Claude-3.5 large language model (Anthropic, 2024) for data
103 synthesis and ecological analysis. Species identification begins with queries to the Ocean
104 Biodiversity Information System (Grassle and Stocks, 1999) and Global Biodiversity In-
105 formation Facility (GBIF, 2024). These queries extract occurrence records, temporal
106 distributions, and abundance data within defined geographical boundaries.

107 The biological data collection phase integrates information from multiple sources.
108 FishBase and SeaLifeBase (Froese et al., 2010) provide life history traits and ecologi-
109 cal parameters. The Ecobase repository (Colléter et al., 2015) supplies parameters from
110 existing ecosystem models. Additional data comes from systematic literature searches of
111 regional fisheries reports and peer-reviewed publications, using standardized search terms
112 (e.g., "[species name] AND diet OR feeding OR prey"). Natural language processing
113 extracts relevant ecological information from these documents.

114 Species grouping employs a vector database (Chroma (Chroma, 2024)) for charac-
115 teristic storage and retrieval. The database maintains embedding vectors derived from
116 ecological descriptions, life history traits, and habitat preferences. Analysis of these vec-
117 tors, combined with ecological rules regarding size classes, feeding guilds, and habitat
118 use, determines functional group assignments. Trophic level estimates from FishBase and
119 SeaLifeBase (Froese et al., 2010) validate the ecological coherence of these classifications.

120 Diet composition analysis merges quantitative records from databases with processed
121 literature data. Food items are extracted from FishBase and SeaLifeBase (Froese et al.,
122 2010) databases, supplemented with interaction data from the Global Biotic Interactions
123 (GLOBI) database (Poelen et al., 2014). Each predator-prey interaction includes source
124 documentation and confidence scores based on data quality metrics.

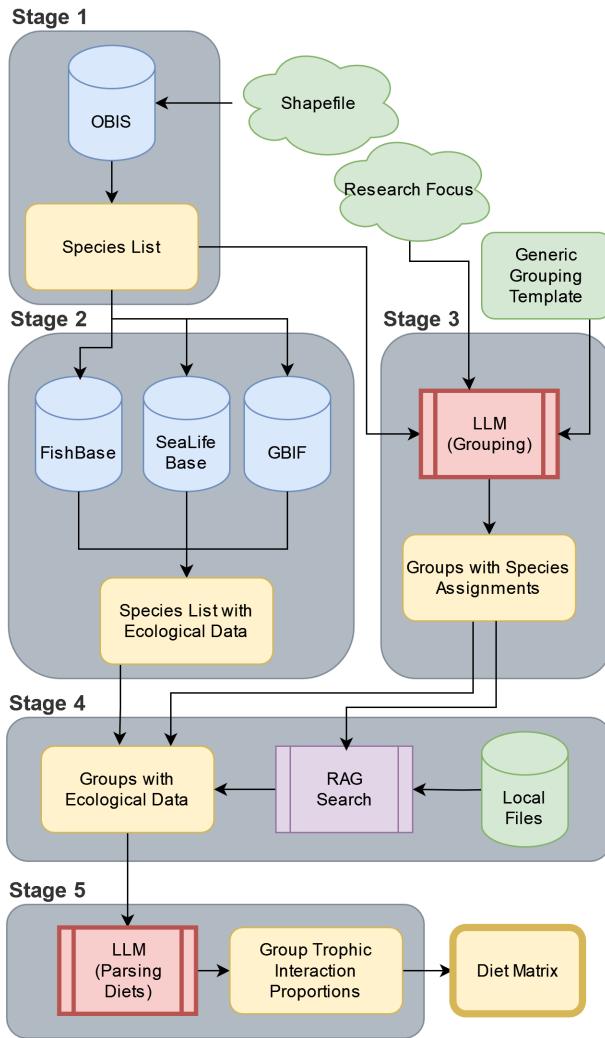


Figure 1: Overview of the AI-assisted framework for ecosystem model development. The process consists of four main steps: species identification, biological data collection, functional group organization, and interaction determination. Highlighted stages (functional group organization and interaction determination) undergo systematic validation across multiple iterations. Each step integrates multiple data sources and analytical approaches.

125 Diet matrix construction implements weighted averaging for prey proportions. Weights
126 derive from data quality metrics including sample size, study recency, and geographical
127 relevance. In cases of sparse direct diet data, the framework estimates trophic interactions
128 based on similar species' preferences and established ecological principles. Each
129 interaction maintains metadata documenting evidence sources and confidence levels.

130 Parameter estimation prioritizes values from comparable ecosystems and species groups
131 in Ecobase ([Colléter et al., 2015](#)). Empirical relationships from literature provide esti-
132 mates where direct parameters are unavailable. The framework validates all estimates
133 against biological constraints and ecological theory, identifying anomalies for expert re-
134 view.

135 Section [S3](#) of the supplementary material contains detailed documentation of all pro-
136 cessing steps, including database queries, literature search criteria, and ecological classifi-
137 cation rules. The complete codebase and configuration files reside at [GitHub repository
138 URL].

139 2.1.1 Species Identification

140 We begin species identification by accepting a GeoJSON file that defines the study re-
141 gion boundaries. We selected the Ocean Biodiversity Information System ([Grassle and](#)
142 [Stocks, 1999](#)) as our primary data source due to its extensive marine species coverage and
143 standardized taxonomic classifications. We access OBIS through the `robis` R package
144 ([Chamberlain, 2020](#)), which enables automated querying and data processing.

145 We process the GeoJSON file in two steps to ensure precise geographic filtering. First,
146 we extract a bounding box from the GeoJSON geometry. We then convert this box into
147 a polygon string for database querying. Through the OBIS checklist function, we retrieve
148 scientific names and complete taxonomic classifications from kingdom to species level for
149 all recorded occurrences within these boundaries.

150 We clean the raw occurrence data through three sequential steps. First, we filter the
151 dataset using OBIS's `is_marine` flag to eliminate terrestrial species that may occur in
152 coastal records. Second, we remove taxonomic redundancy using a rank-based approach
153 that retains only the most specific classification level available. Our algorithm processes
154 taxonomic ranks from most specific (scientific name) to most general (kingdom), keeping
155 only the entry with the highest taxonomic resolution for each organism. Third, we aggre-
156 gate occurrences by unique species while preserving their complete taxonomic hierarchies,
157 calculating occurrence frequencies for each species.

158 We store the final species list in a structured CSV file containing verified marine
159 species, their complete taxonomic hierarchies, and occurrence frequencies. This stan-
160 dardized format facilitates efficient data transfer to subsequent framework stages while
161 maintaining provenance information. The complete R implementation, including rank-
162 based filtering algorithms and geographic processing functions, is available in the project
163 repository.

164 2.1.2 Data Harvesting

165 Following species identification, we gather ecological and life history information for each
166 identified species through an incremental processing system. We access SeaLifeBase and
167 FishBase ([Froese et al., 2010](#)) through their publicly available PARQUET files using
168 DuckDB for efficient querying of large datasets. We construct temporary database tables
169 from our species list and execute structured queries to match species across the databases,

170 processing SeaLifeBase and FishBase data separately to maintain distinct source attribution.
171

172 Our species filtering protocol implements specific constraints to align with Ecopath
173 with Ecosim modeling requirements. When processing taxonomic data, we prioritize
174 species-level entries over genus-level classifications, only defaulting to genus-level when
175 species-specific data is unavailable. For diet composition data, we specifically filter for
176 adult life stages, as juvenile and larval stages are typically incorporated into planktonic
177 functional groups rather than treated as separate components of adult diets.

178 We supplement the base biological data with interaction information from the Global
179 Biotic Interactions (GLOBI) database ([Poelen et al., 2014](#)). For each species, we query
180 the GLOBI API using URL-encoded species names. The GLOBI data processing treats
181 directional relationships ('eats'/'preysOn' and 'eatenBy'/'preyedUponBy') as complemen-
182 tary evidence of trophic interactions, with interaction counts aggregated at the functional
183 group level to provide quantitative measures of interaction strength between groups.

184 Our data cleaning protocol consists of three stages: standardizing data types, handling
185 null values, and maintaining source attribution. We store the processed data incrementally
186 in a JSON document, implementing file locking mechanisms to handle concurrent access.
187 Each species entry contains complete taxonomic hierarchy, matched database records,
188 interaction data, and source attribution. This incremental approach enables processing
189 recovery after interruptions and allows parallel processing of different species batches.
190 Technical implementation details are provided in Section [S1](#) of the supplementary mate-
191 rial.

192 2.1.3 Species Grouping

193 We implement species grouping through a flexible system that adapts to regional ecological
194 contexts. The framework supports three approaches for defining functional groups: using
195 templates from existing EcoBase models, generating region-specific groups through AI
196 analysis of geographic characteristics, or applying a predefined template (provided in Sec-
197 tion ?? of the supplementary material). For geographic regions, we analyze oceanographic
198 conditions, habitat types, and ecological characteristics to inform group definitions.

199 We process taxa hierarchically from kingdom to species level, implementing an incre-
200 mental system with progress tracking to handle large datasets reliably. At each taxonomic
201 level, Claude evaluates taxa against the selected grouping template using the following
202 prompt:

203 You are classifying marine organisms into functional groups for an Ecopath with

Ecosim (EwE) model. Functional groups can be individual species or groups of species that perform a similar function in the ecosystem, i.e. have approximately the same growth rates, consumption rates, diets, habitats, and predators. They should be based on species that occupy similar niches, rather than of similar taxonomic groups.

Examine these taxa at the [rank] level and assign each to an ecological functional group.

Rules for assignment:

- If a taxon contains members with different feeding strategies or trophic levels, assign it to 'RESOLVE'
- Examples requiring 'RESOLVE':
 - A phylum containing both filter feeders and predators
 - An order with both herbivores and carnivores
 - A class with species across multiple trophic levels
- If all members of a taxon share similar ecological roles, assign to an appropriate group
- Only consider the adult phase of the organisms, larvae and juveniles will be organized separately
- Only assign a definite group if you are confident ALL members of that taxon belong to that group

Taxa to classify: [List of taxa]

Available ecological groups (name: description): [List of available groups and their descriptions]

Return only a JSON object with taxa as keys and assigned groups as values.

We implement several reliability mechanisms in the classification process:

- Exponential backoff retry system for handling temporary AI service interruptions
- Incremental progress tracking with file locking for concurrent processing
- Validation of group assignments against template definitions
- Separate handling of AI-suggested groups not in the original template

When the research focus indicates areas requiring higher resolution (e.g., commercial fisheries species), we modify the classification process with additional guidance:

Special consideration for research focus: The model's research focus is: [research

focus]

When classifying taxa that are related to this research focus:

- Consider creating more detailed, finer resolution groupings
- Keep species of particular interest as individual functional groups
- For taxa that interact significantly with the focal species/groups, maintain higher resolution groupings
- For other taxa, broader functional groups may be appropriate

213

214 We process each taxon through this classification framework, generating structured
215 JSON documents that map species to functional groups. Taxa marked as RESOLVE undergo
216 evaluation at finer taxonomic levels until reaching a definitive group assignment or the
217 species level. We maintain complete provenance information, including the source of
218 group definitions and any AI-suggested modifications.

219 2.1.4 Diet Matrix Construction

220 We construct diet matrices through a two-stage process designed for efficiency and re-
221 liability. The first stage gathers comprehensive diet data using parallel processing and
222 caching mechanisms. We query SeaLifeBase and FishBase ([Froese et al., 2010](#)) food items
223 databases through DuckDB, enabling efficient processing of PARQUET files without full
224 memory loading. For each species, we extract food items using specific codes that link to
225 standardized diet categories.

226 We supplement database records with interaction data from the Global Biotic Inter-
227 actions (GLOBI) database ([Poelen et al., 2014](#)). Our GLOBI processing differentiates
228 between direct observations (`eats`, `preysOn`) and inverse relationships (`eatenBy`,
229 `preyedUponBy`), maintaining separate interaction counts for each type. We further enrich
230 this data through retrieval-augmented generation (RAG) searches of regional literature
231 (detailed in Section [S3.2](#) of the supplementary material), focusing on specific feeding
232 relationships and dietary preferences.

233 For each functional group, we combine these data sources into a structured profile.
234 Claude ([Anthropic, 2024](#)) then analyzes this profile using the following prompt:

235

Based on the following information about the diet composition of [group], pro-

vide a summary of their diet. Include the prey items and their estimated proportions in the diet.

Available functional groups and their details: [List of groups with descriptions and top species]

Here is the diet data for [group]: [Combined data including RAG search results, compressed food categories, and GLOBI interactions]

Format your response as a list, with each item on a new line in the following format:

Prey Item: Percentage

For example:

Small fish: 40%

Zooplankton: 30%

Algae: 20%

Detritus: 10%

If exact percentages are not available, estimate percentages based on the information you have been provided. Ensure that all percentages add up to approximately 100%. Consider the RAG search results, compressed food categories, and GLOBI data when creating your summary. Pay special attention to the GLOBI interaction counts, which indicate frequency of observed feeding relationships. Note that some species may feed on juvenile or larval forms of other species, which are often classified in different functional groups than the adults.

236

237 We implement an incremental processing system with file locking mechanisms to han-
238 dle large datasets reliably. The system maintains caches for species-level diet data from
239 databases, combined diet information including literature results, GLOBI interaction net-
240 works, and intermediate AI analyses. This caching system enables recovery from inter-
241 ruptions and facilitates parallel processing of different functional groups.

242 The matrix assembly stage processes the AI's standardized diet descriptions through
243 automated parsing. We convert percentage strings to decimal values and implement a
244 validation system. Our validation ensures that all proportions sum to 1 for each predator,
245 prey items match defined functional groups, and mass-balance requirements are main-
246 tained.

247 When prey items do not exactly match functional group names, we employ a hierarchi-
248 cal matching system. The system first attempts exact matches, then falls back to partial
249 matching using taxonomic information, and logs unmatched items for expert review.

250 The final output consists of a CSV file containing the complete diet matrix, with
251 predators as rows and prey as columns. Each cell contains the proportion of predator diet
252 comprised by that prey item.

253 2.1.5 Parameter Estimation (Likely to be removed - no validation)

254 The final stage estimates the required Ecopath ([Christensen and Walters, 2004](#)) param-
255 eters for each functional group. We query the EcoBase repository ([Colléter et al., 2015](#))
256 using a structured search strategy that combines group names with regional context (e.g.,
257 "Western Australian shelf species in [group]"). For each functional group, we retrieve
258 five key parameters from comparable models: habitat area fractions, biomass densities

(t/km^2), production/biomass ratios (P/B , $year^{-1}$), consumption/biomass ratios (Q/B , $year^{-1}$), and ecotrophic efficiency (EE).

We implement parameter-specific validation checks. For biomass densities, we verify values fall within expected ranges for each functional group type. For P/B and Q/B ratios, we check consistency with known allometric relationships and life history characteristics. EE values must fall between 0 and 1, with additional validation against typical ranges for similar species groups.

Claude ([Anthropic, 2024](#)) analyzes these parameter sets using standardized criteria. The analysis considers the relevance of each model to the group in question, the consistency of values across different models, trends or patterns in the data that might indicate suitable values, and the ecological context provided by model metadata.

For each parameter assignment, Claude provides a structured response containing the suggested parameter value, detailed reasoning for the selection, references to source models, and any assumptions or caveats.

When EcoBase data is insufficient or parameter values show high variability, we employ a hierarchical fallback system. First, we search for parameters from taxonomically similar groups. Next, we apply empirical relationships where available. Finally, if automated methods are unsuitable, we flag the parameter for manual parameterization.

The system maintains comprehensive logging of parameter assignments. This includes search queries and results, parameter validation checks, AI reasoning and decisions, source model references, and data quality assessments.

The final output consists of a structured JSON file containing the complete parameter set for each functional group. This format preserves the full provenance of each parameter value while enabling automated validation and mass-balance calculations in Ecopath.

2.2 Validation Framework

We executed model generation across three distinct phases. In phase one, we established baseline configurations for each study region by processing species occurrence data and research parameters. We terminated this phase prior to species grouping, creating standardized input states for subsequent validation iterations.

We executed five independent iterations per region in phase two. Each iteration began with species grouping and proceeded through the complete model construction sequence. We maintained fixed input parameters across iterations while allowing the AI's stochastic decision processes to generate natural variation in outputs.

2.2.1 Study Regions

We validate our framework using three Australian marine regions that present distinct ecological characteristics and modeling challenges (Figure 2). The Northern Territory region represents a tropical ecosystem characterized by complex reef systems, seasonal monsoon influences, and high biodiversity. This region tests the framework's ability to handle diverse species assemblages and complex trophic interactions in a dynamic environment.

The South East Inshore region represents a temperate coastal system with extensive ecological records. This region has comprehensive diet information in established databases, well-documented EwE models spanning multiple years, and active research programs. The rich availability of expert knowledge and historical data makes this region ideal for validating the framework's data integration capabilities.

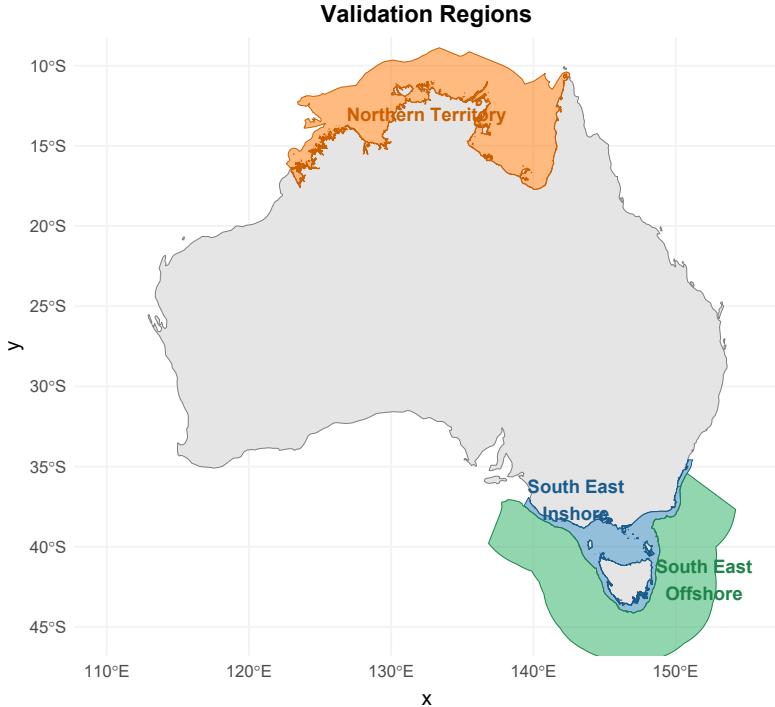


Figure 2: Map of the three validation regions used in this study: Northern Territory (orange), South East Inshore (blue), and South East Offshore (green).

304 The South East Offshore region presents a deep-water ecosystem that challenges the
 305 framework with data-limited conditions and unique ecological patterns. This region tests
 306 the framework's capacity to handle situations where direct observational data may be
 307 sparse and where species interactions may be less well understood. The contrasting char-
 308 acteristics of these three regions provide a robust test of the framework's adaptability
 309 across different ecological contexts.

310 2.2.2 Group Consistency Analysis

311 We tracked each species' group assignments across iterations and calculated a consistency
 312 score:

$$\text{Consistency Score} = \frac{\text{Number of occurrences in most common group}}{\text{Total number of iterations}}$$

313 We classified species with consistency scores below 0.95 as unstable, indicating variable
 314 group assignments across iterations.

315 We assessed group stability using the Jaccard similarity coefficient to measure consis-
 316 tency of group membership between consecutive iterations:

$$J(i, j) = \frac{|M_i \cap M_j|}{|M_i \cup M_j|}$$

317 where M_i and M_j represented species members in iterations i and j . We calculated
 318 the overall stability score by averaging Jaccard similarities across consecutive iteration
 319 pairs.

320 **2.2.3 Diet Matrix Analysis**

321 We analyzed diet matrix reliability by focusing on significant predator-prey interactions,
322 which we defined as those comprising more than 5% of a predator's diet. For each inter-
323 action, we calculated:

324 1. Presence ratio across iterations:

$$P_{ij} = \frac{\text{Number of iterations with interaction}}{n}$$

325 where n was the total number of iterations.

326 2. Mean diet proportion:

$$\mu_{ij} = \frac{1}{n} \sum_{k=1}^n x_{ijk}$$

327 where x_{ijk} represents diet proportion for predator i consuming prey j in iteration k .

328 3. Stability score:

$$S_{ij} = \frac{1}{n} \sum_{k=1}^n \frac{|x_{ijk} - \mu_{ij}|}{\max_k(x_{ijk})}$$

329 where x_{ijk} represents the diet proportion for predator i consuming prey j in iteration
330 k , μ_{ij} is the mean diet proportion across iterations, and $\max_k(x_{ijk})$ is the maximum
331 value across iterations. This score ranges from 0 (perfect stability) to 1 (maximum
332 instability), with lower values indicating more consistent diet proportions across
333 iterations.

334 We chose this stability metric over traditional variance measures for several reasons.

335 First, by normalizing deviations by the maximum value, the metric achieves scale inde-
336 pendence, allowing meaningful comparisons between interactions of different magnitudes.
337 For example, the sequences [0.2, 0.2, 0.2, 0.2, 0.1] and [0.02, 0.02, 0.02, 0.02, 0.01]
338 would yield the same stability score despite having different absolute variances. Second, the
339 bounded range between 0 and 1 provides an intuitive scale for assessing stability, unlike
340 the unbounded nature of variance. Third, this approach effectively handles small varia-
341 tions in small values, ensuring that minor fluctuations in trace diet components do not
342 disproportionately influence the stability assessment.

343 We classified interactions as unstable when their stability score exceeded 0.3, indicating
344 significant variation in diet proportions across iterations. To illustrate this metric:

- 345 • A stable interaction ($S = 0.08$) might show values [0.02, 0.02, 0.02, 0.02, 0.01], where
346 proportions remain very similar across iterations
- 347 • An unstable interaction ($S = 0.39$) might show values [0.027, 0.25, 0.25, 0.067, 0.25],
348 where proportions vary substantially between iterations

349 This metric provides a continuous measure of stability that handles both presence/absence
350 patterns and magnitude variations in a unified way, avoiding discontinuities at zero values.

351 **2.2.4 Expert Evaluation Protocol**

352 To evaluate the quality of diet matrices generated through artificial intelligence methods,
353 we implemented a single-blind validation framework comparing AI-generated and expert-
354 created Ecopath with Ecosim (EwE) diet matrices across three distinct regions. Two
355 independent domain experts were engaged to assess both matrix types without knowledge
356 of their origin. The evaluation protocol employed a five-point Likert scale assessment,
357 where a score of 1 indicated strong preference for the AI-generated matrix and 5 indicated
358 strong preference for the expert-generated matrix.

359 The validation framework encompassed two primary assessment criteria. First, ex-
360 perts evaluated the structural validity of functional group classifications within each ma-
361 trix. Second, they assessed the ecological plausibility of diet proportions assigned to each
362 functional grouping. To capture nuanced insights beyond numerical scores, experts pro-
363 vided qualitative feedback on the strengths and weaknesses of each matrix. This dual
364 quantitative-qualitative approach enabled comprehensive evaluation of both methodolog-
365 ical approaches to diet matrix construction.

366 **2.3 Statistical Analysis**

367 We conducted statistical analyses to evaluate framework performance across regions. We
368 assessed group consistency using chi-square tests and coefficients of variation, examining
369 the stability of species assignments across iterations. We analyzed regional differences in
370 group characteristics using one-way ANOVA with Cohen's f effect size calculations. We
371 evaluated trophic level patterns using Kruskal-Wallis tests to identify significant differ-
372 ences in trophic structure. We quantified diet matrix reliability through pairwise Spear-
373 man correlations between iterations, focusing on significant interactions (>0.05 propor-
374 tion).

375 **3 Results**

376 **3.1 Framework Validation**

377 **3.1.1 Processing Scale and Performance**

378 We evaluated our framework through five independent runs across three distinct Aus-
379 tralian regions, processing a total of 41,085 species. The framework handled 11,362 species
380 in the Northern Territory's tropical reef ecosystem, 13,901 in the South East Inshore's
381 coastal and pelagic environments, and 15,822 in the South East Offshore's deep-water
382 systems.

383 **3.1.2 Computational Efficiency**

384 The computational requirements of the AI framework varied across regions. Total pro-
385 cessing time ranged from 17.5 to 73.0 hours across regions. The most time-intensive stage
386 was the harvesting of biological data from online databases, accounting for approximately
387 65% of the total processing time. Species identification typically required 0.01 hours,
388 while the AI-driven species grouping process averaged 0.12 hours. Diet data collection
389 and matrix construction required 13.5 and 0.03 hours respectively, with final parameter
390 estimation taking 0.24 hours. On average, the framework required 6.8 seconds per species

391 for data harvesting and 3.5 seconds per species for diet data collection, though these rates
 392 varied considerably between regions due to differences in data availability and species complexity.

Table 1: Computational requirements by region and processing stage

Region	Species Count	Processing Time (hours)				
		Identification	Data Harvest	Grouping	Diet Collection	Matrix Construction
Northern Territory	11,362	0.01	9.6	0.1	7.6	0.03
South East Inshore	13,901	0.01	53.8	0.1	18.7	0.04
South East Offshore	15,821	0.01	15.3	0.1	14.1	0.03

*Includes periods of system inactivity between processing batches

393
 394 Processing times varied by region and stage (Table 1). Data harvesting required 9.6
 395 hours for the Northern Territory (3.0 seconds per species) and 53.8 hours for the South
 396 East Inshore (13.9 seconds per species). Diet data collection took 7.6 hours for the
 397 Northern Territory (2.4 seconds per species) and 18.7 hours for the South East Inshore
 398 (4.8 seconds per species). Species identification (0.01 hours), grouping (0.1 hours), and
 399 parameter estimation (0.1-0.3 hours) remained constant across regions.

400 3.1.3 Classification Consistency

401 The framework reduced 63 potential functional groups to 34-36 region-specific groups.
 402 Chi-square tests indicated consistency across regions ($p > 0.85$). Coefficient of variation
 403 measurements were: South East Inshore (mean = 0.002, SD = 0.004), Northern Territory
 404 (mean = 0.004, SD = 0.012), and South East Offshore (mean = 0.019, SD = 0.052).
 405 ANOVA revealed significant regional differences in group sizes ($F = 8279010.7$, $p < 0.001$,
 406 Cohen's $f = 2877.3$). Classification stability reached 99% in the Northern Territory (114
 407 inconsistent species), 99.6% in the South East Inshore (52 inconsistent species), and 98.8%
 408 in the South East Offshore (191 inconsistent species).

409 Figure 3 presents the quantitative analysis across regions. Panel A shows median
 410 group sizes of 150-200 species. Panel B displays consistency scores ranging from 0.4-1.0
 411 in the Northern Territory and 0.8-1.0 in South East regions. Panel C indicates Jaccard
 412 similarity indices of 0.995 in South East Offshore and 0.985 in Northern Territory. Panel
 413 D shows group size standard deviations of 5-10 species in South East regions and 15-35
 414 species in Northern Territory.

415 The stability heatmap (Figure 4) reveals near-perfect stability (Jaccard similarity $>$
 416 0.99) for most functional groups, with specific exceptions documented in Table 2. Clas-
 417 sification inconsistencies followed systematic patterns, primarily occurring between eco-
 418 logically similar groups. In the Northern Territory, anemones alternated between benthic
 419 infaunal carnivores and benthic filter feeder classifications, while flatfishes varied between
 420 benthivores and shallow demersal fish categories. The South East Inshore region showed
 421 similar patterns, with *Antigonia* species alternating between planktivore and benthivore
 422 classifications.

423 3.1.4 Diet Matrix Validation

424 The framework demonstrated consistent patterns in diet matrix construction across itera-
 425 tions. Using our stability score metric (where 0 indicates perfect stability and 1 indicates

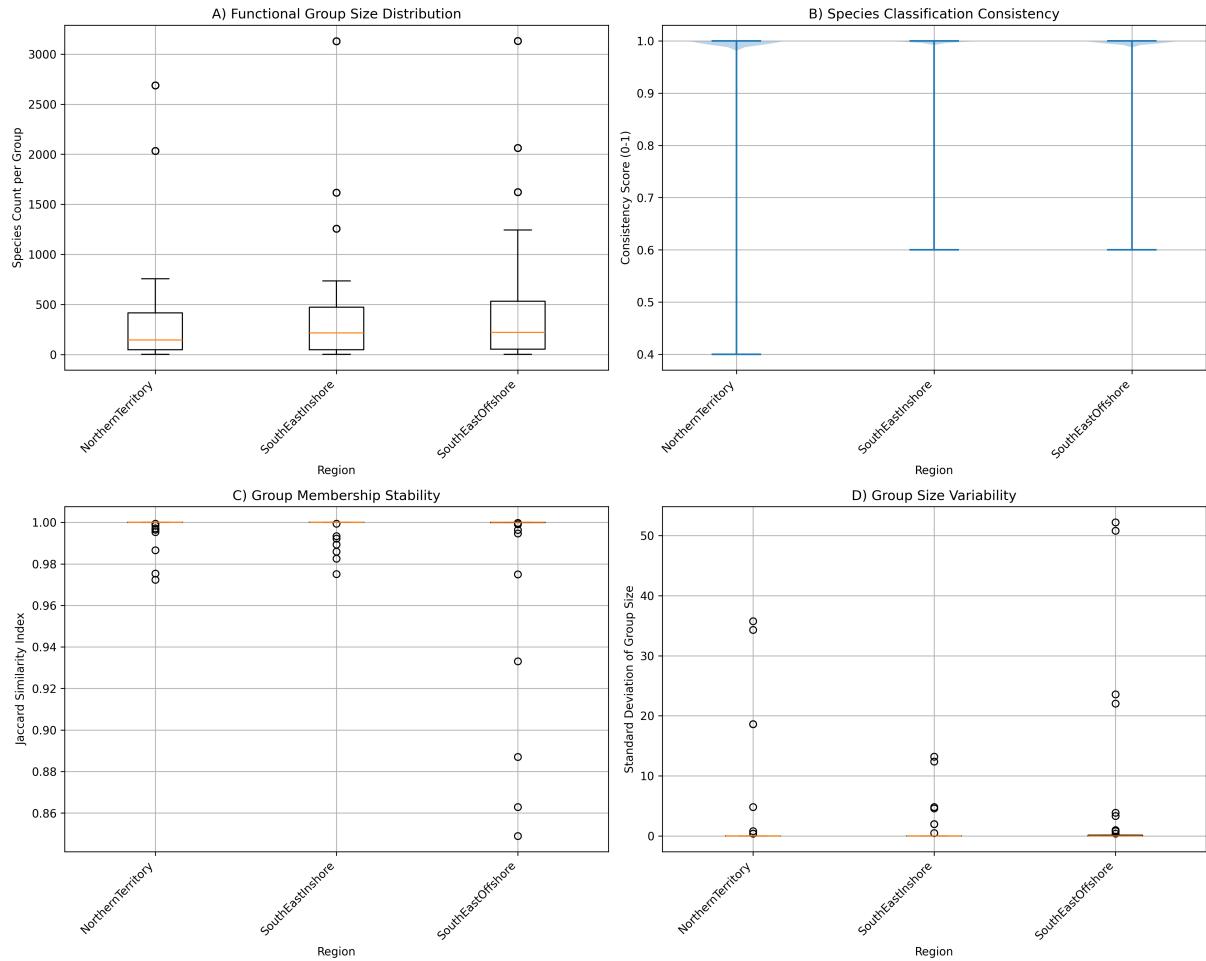


Figure 3: Multi-panel analysis of framework performance across regions. (A) Box plots of functional group sizes (0-3000 species), showing similar median sizes but varying distributions across regions, with outliers indicating some very large groups. (B) Violin plots of species classification consistency (0.4-1.0), where wider sections indicate more species with that consistency score; most species show high consistency (>0.9) with slightly more variation in the Northern Territory. (C) Box plots of group stability measured by Jaccard similarity (0.975-1.000), showing highest stability in South East Offshore and more variable stability in Northern Territory. (D) Box plots of group size variation (standard deviation 0-35), demonstrating larger fluctuations in group membership in the Northern Territory compared to South East regions.

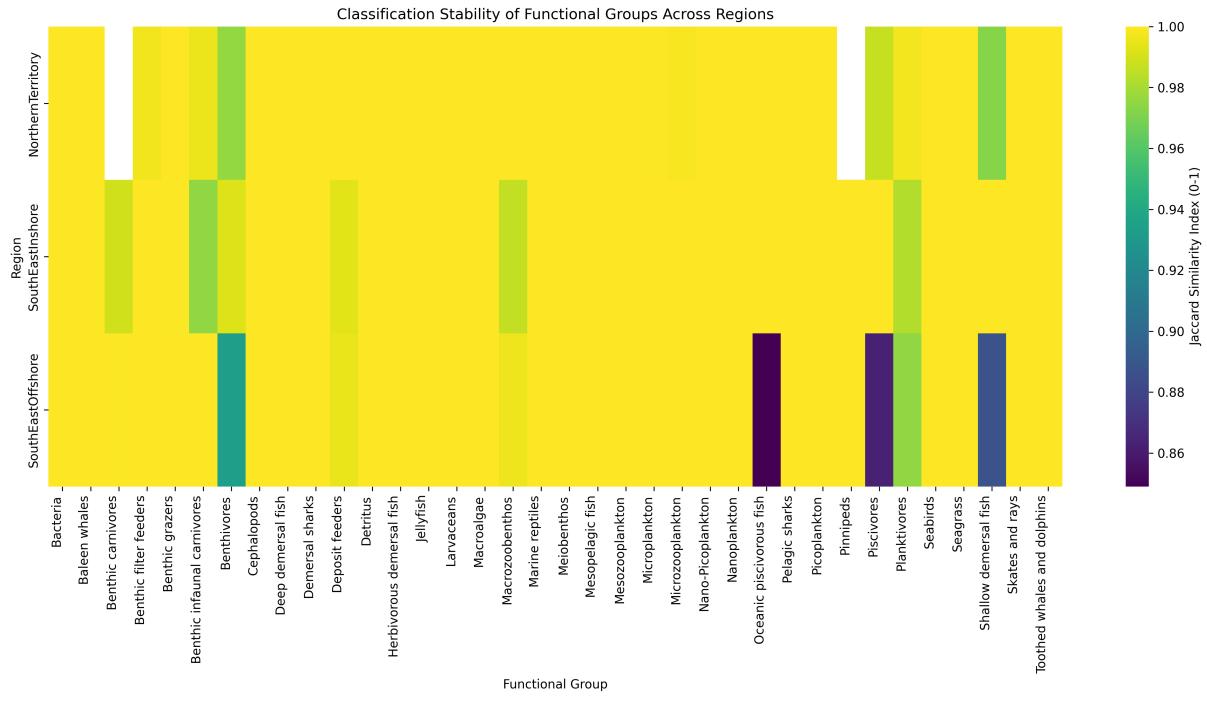


Figure 4: Heatmap showing the stability of functional group classifications across regions. Each cell displays the Jaccard similarity score (ranging from 0.975 to 1.000) between consecutive framework iterations, where 1.000 indicates perfect consistency in species assignments. Darker red colors represent higher stability (scores near 1.000), while lighter colors indicate more variable classifications (scores closer to 0.975). Most functional groups show high stability (>0.99) across all regions, with occasional variations in groups like benthic grazers and deposit feeders, particularly in the Northern Territory region.

Table 2: Species with unstable group assignments across validation iterations. Species are considered unstable if their consistency score is below 0.95, indicating they were assigned to different functional groups in different iterations. The consistency score represents the proportion of iterations where the species was assigned to its most common group.

Region	Species Pattern	Count	Primary Group	Alternative Groups
Northern Territory	Uranoscopidae	12	Shallow demersal fish (40%)	Benthivores (40%), Piscivores (20%)
	Actinodendron & allies	42	Benthic infaunal carnivores (60%)	Benthic filter feeders (40%)
	Platycephalidae	60	Benthivores (60%)	Shallow demersal fish (40%)
South East Inshore	Cirrhitabrus & <i>Antigonia</i> spp.	4	Planktivores (60%)	Benthivores (40%)
	Leptostrebla	13	Deposit feeders (60%)	Macrozoobenthos (40%)
	Polycladida	35	Benthic infaunal carnivores (80%)	Benthic carnivores (20%)
South East Offshore	Carangidae	34	Oceanic piscivorous fish (60%)	Piscivores (40%)
	Pleuronectiformes	42	Shallow demersal fish (80%)	Benthivores (20%)
	Clinidae	115	Shallow demersal fish (80%)	Benthivores (20%)

maximum variation), we found that most predator-prey interactions remained highly stable (Figure 5). The Northern Territory analysis identified 127 predator-prey interactions with a median stability score of 0.18, with 84% of interactions showing scores below our 0.3 threshold. The South East Inshore region produced 133 interactions with a median stability score of 0.16 (86% below threshold), while the South East Offshore achieved the highest stability with 147 interactions and a median score of 0.15 (89% below threshold).

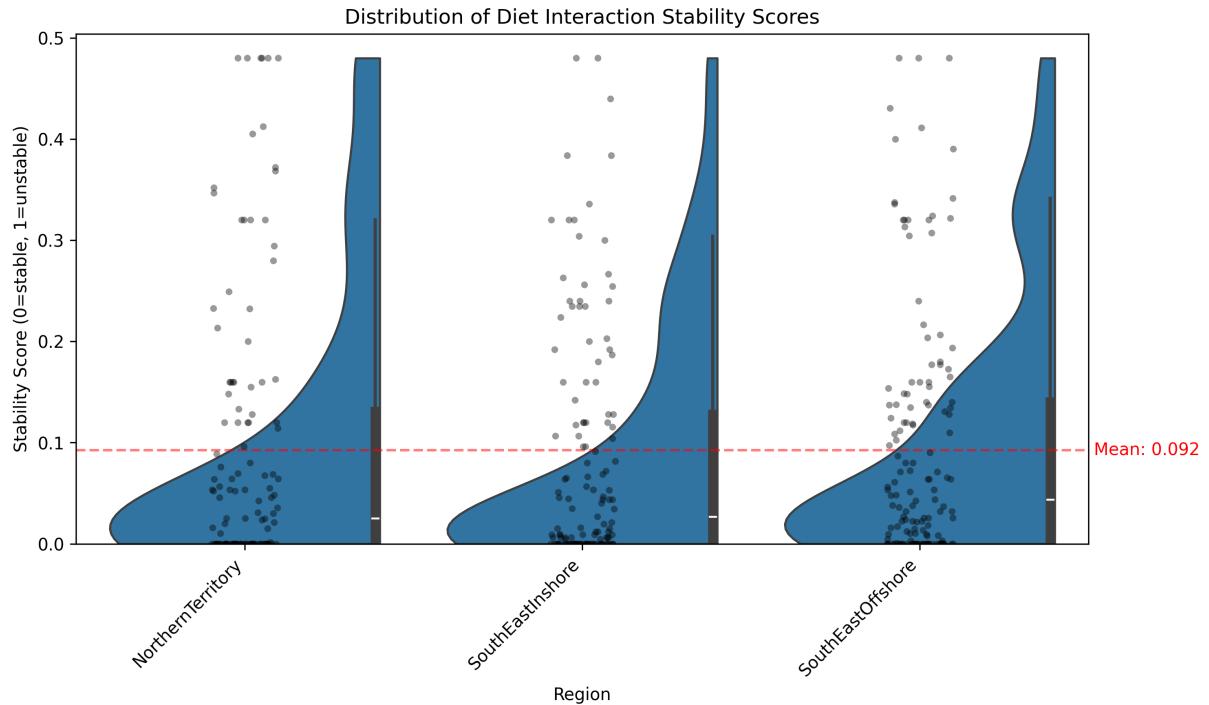


Figure 5: Distribution of diet interaction stability scores across regions. Half-violin plots show the density of stability scores (0=stable, 1=unstable), with embedded box plots indicating quartiles and median. Individual points represent specific predator-prey interactions, and the red dashed line shows the mean stability score across all regions. The distributions are bounded at zero, reflecting perfect stability, with most interactions showing scores below 0.3.

Analysis of stability scores by predator group revealed systematic patterns in diet consistency (Figure 6). Lower trophic level groups like nanoplankton and microplankton showed high stability (scores < 0.1) across all regions, reflecting their specialized diets. Mid-trophic level predators such as mesopelagic fish and cephalopods exhibited moderate variability (scores 0.1-0.2), while higher trophic level predators including demersal sharks and benthic carnivores showed the most variation in diet composition (scores > 0.2). This pattern suggests that diet variability increases with trophic level, possibly reflecting the more generalist feeding strategies of higher-level predators. Detailed diet matrices for each region are provided in the supplementary material (Figure S1).

3.2 Ecological Findings

3.2.1 Trophic Level Patterns

Analysis of trophic structure revealed significant differences in both the Northern Territory and South East Inshore regions (Kruskal-Wallis H-test, $H = 164.0$ and 172.0 respectively, p

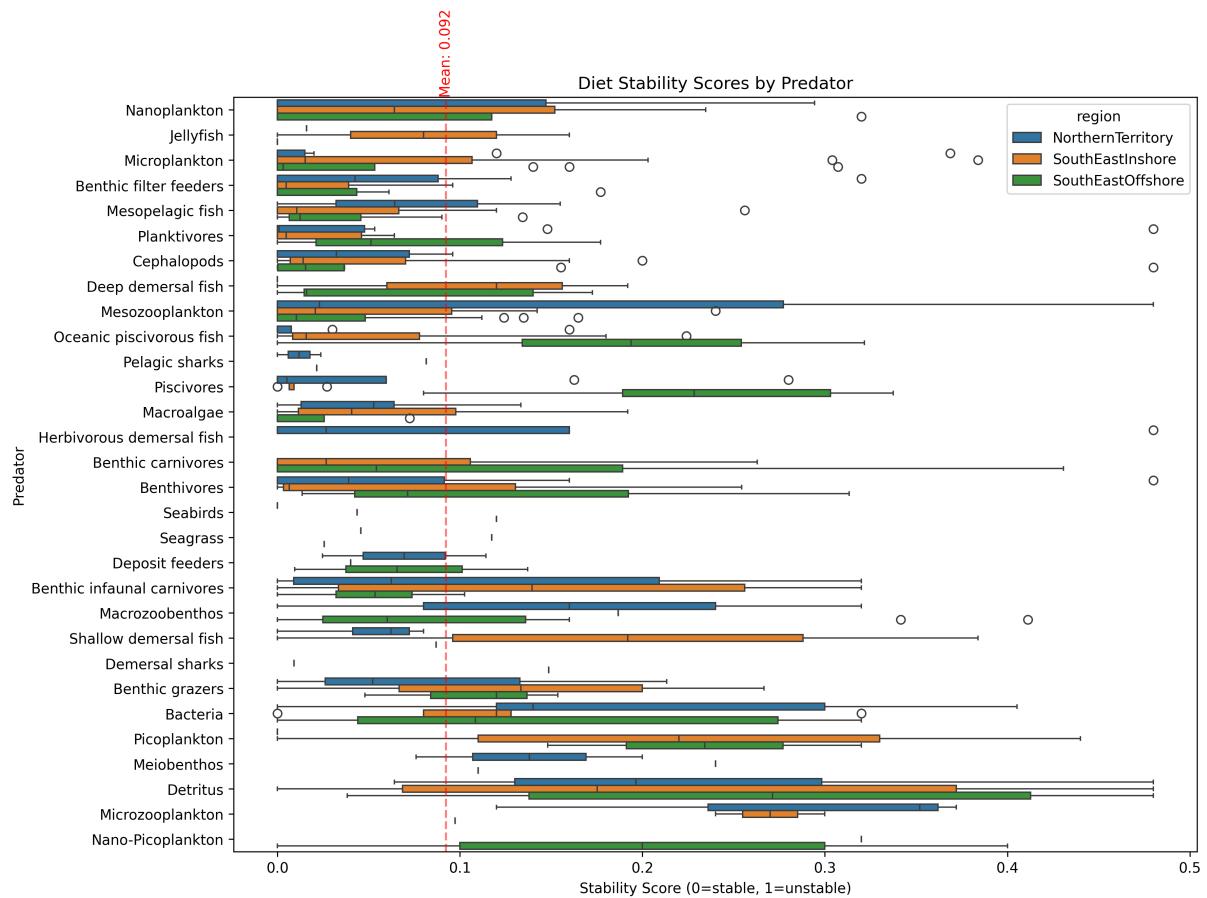


Figure 6: Diet stability scores grouped by predator, ordered by median stability. Box plots show the distribution of stability scores for each predator's diet across regions (colored by region). The red dashed line indicates the mean stability score across all predator-prey interactions. Lower scores indicate more consistent diet compositions across framework iterations.

< 0.001 for both). Higher trophic level species, particularly apex predators and specialized feeders, maintained consistent classifications across all regions. In contrast, lower trophic levels showed greater variability, particularly among planktonic and benthic invertebrate groups. Benthivores exhibited the highest variation, ranging from 715 to 788 species between runs, while shallow demersal fish showed moderate variation between 586 and 659 species.

4 Discussion

4.1 AI Framework Consistency

Our framework demonstrates robust performance in automating the construction of complex ecosystem models across diverse marine environments. The successful processing of over 41,000 species across three distinct regions validates the framework's scalability and broad applicability. The framework's ability to maintain consistent species classifications while adapting to regional ecological differences suggests it effectively captures fundamental ecological relationships.

The computational efficiency analysis reveals important insights about framework scalability. Data harvesting and diet collection emerge as the primary computational bottlenecks, particularly evident in the South East Inshore region's extended processing times. These bottlenecks likely stem from API rate limitations and the complexity of extracting ecological information from diverse data sources. The relatively constant processing times for species identification, grouping, and parameter estimation across regions indicate these components scale efficiently with increasing species counts.

The framework's classification consistency merits particular attention. The high stability scores (98.8-99.6%) across regions demonstrate reliable species-to-group assignments. The systematic nature of classification inconsistencies provides valuable ecological insights. Species with ambiguous classifications often represent organisms that naturally span multiple ecological niches. For instance, the alternating classifications of anemones between benthic infaunal carnivores and filter feeders reflect their complex feeding strategies. Similarly, the variable classification of flatfishes between benthivore and demersal categories aligns with their known ecological plasticity. These classification patterns suggest the framework captures meaningful ecological uncertainty rather than arbitrary assignment errors.

The diet matrix validation reveals a nuanced picture of trophic relationship stability. The moderate negative correlations between iterations initially appear concerning. However, the high stability of significant feeding interactions (89.8-92.5%) suggests the framework maintains consistent broad-scale trophic structure while allowing flexibility in fine-scale interactions. This pattern aligns with ecological theory, where core trophic relationships remain stable while peripheral feeding interactions may vary with resource availability and environmental conditions.

The observed trophic level patterns provide compelling evidence for the framework's ecological validity. The consistent classification of higher trophic level species reflects the relatively constrained niches of specialized predators. Conversely, the greater variability in lower trophic level classifications mirrors the natural complexity and adaptability of these groups. The framework's ability to capture this fundamental ecological pattern suggests it successfully incorporates biological realism into its classification decisions.

489 The regional differences in group size variation and classification stability offer in-
490 sights into ecosystem complexity. The Northern Territory's higher variation in group
491 sizes and slightly lower classification stability likely reflect the increased ecological com-
492 plexity of tropical reef systems. The more stable classifications in the South East regions
493 may indicate more clearly defined ecological niches in temperate marine environments.
494 These regional patterns demonstrate the framework's sensitivity to underlying ecological
495 differences while maintaining consistent overall performance.

496 The reduction from 63 potential functional groups to 34-36 region-specific groups in-
497 dicates the framework's ability to identify ecologically relevant groupings while avoiding
498 artificial complexity. The statistical consistency across regions suggests these groupings
499 represent fundamental ecological units rather than arbitrary divisions. This optimiza-
500 tion of functional group complexity balances model detail with practical utility, a crucial
501 consideration for ecosystem modeling applications.

502 4.2 Limitations and Uncertainties

503 Our framework faces several AI-specific limitations identified in recent ecological mod-
504 eling research. The framework's performance depends on the quality and completeness
505 of available training data. The framework's classification patterns showed regional varia-
506 tions in stability, though further research is needed to determine the relationship between
507 data availability and classification performance. This consideration aligns with findings
508 from [Kuhn et al. \(2024\)](#) regarding machine learning applications in fisheries, where data
509 quality significantly impacts model reliability.

510 A significant limitation of our approach stems from its reliance on Claude 3.5 Sonnet,
511 a closed-source large language model. The proprietary nature of this model introduces
512 uncertainty regarding the training data used in its development and potential biases that
513 may affect ecological interpretations. While our validation demonstrates consistent per-
514 formance, the inability to examine the model's training data or internal decision-making
515 processes raises important considerations for scientific reproducibility. Future iterations
516 of the framework may benefit from exploring open-source alternatives or implementing
517 multiple model approaches as demonstrated by [Kommineni et al. \(2024\)](#) in their work
518 with various LLMs for biodiversity research.

519 The framework's interpretability presents another key challenge. While our validation
520 demonstrates robust performance metrics, the underlying AI decision-making processes,
521 particularly in parameter estimation, require careful scrutiny. This challenge mirrors
522 concerns raised by [Fernandes and D'Mello \(2024\)](#) regarding the "black box" nature of AI
523 systems in aquaculture applications. Our framework partially addresses these concerns
524 through explicit uncertainty quantification and validation metrics, but further work is
525 needed to enhance model transparency.

526 Technical limitations include computational resource requirements and processing time
527 constraints, particularly evident in data harvesting operations. These limitations align
528 with implementation barriers identified by [Fernandes and D'Mello \(2024\)](#), including ac-
529 quisition costs and technical expertise requirements. The framework's sensitivity to data
530 availability varies across ecological roles and regions, affecting both classification stability
531 and diet matrix reliability.

532 4.3 Applications for EBFM

533 The framework offers significant practical value for ecosystem-based fisheries management
534 through several key capabilities. Managers can now construct initial EwE models for new
535 regions in days rather than months, enabling faster response to management needs. This
536 addresses a key bottleneck identified by [Zheng et al. \(2023\)](#) in marine resource manage-
537 ment. The framework’s explicit quantification of uncertainty in species classifications and
538 trophic relationships enables managers to identify areas requiring additional data collec-
539 tion or careful monitoring, following approaches recommended by [Kuhn et al. \(2024\)](#).

540 The demonstrated adaptability across diverse ecosystems enables customization for
541 specific regions while maintaining methodological consistency, supporting standardized
542 approaches to ecosystem management across jurisdictions. Through systematic processing
543 of available data, the framework also reveals specific areas where additional research or
544 monitoring would most improve model reliability, aligning with recent work by [Chen and](#)
545 [Xu \(2024\)](#) on marine protected area management.

546 4.4 Future Directions

547 Future development should focus on several key areas to enhance the framework’s utility
548 and reliability. Integration with emerging marine AI systems, such as specialized models
549 like MarineGPT ([Zheng et al., 2023](#)) and multimodal frameworks like LITE ([Li et al.,](#)
550 [2024](#)), could enhance the framework’s capabilities in processing region-specific informa-
551 tion and handling incomplete data. The LITE framework’s success in combining semantic
552 time-series with temporal trend images suggests promising directions for improving our
553 framework’s robustness to data gaps and distribution shifts. Their sparse Mixture-of-
554 Experts approach to handling incomplete features could particularly benefit our diet ma-
555 trix construction process. Implementation of standardized evaluation frameworks, such
556 as those proposed in the ELLE dataset ([Guo et al., 2025](#)), would enable more rigorous
557 assessment of the framework’s performance across different ecological contexts. Develop-
558 ment of improved visualization and explanation tools for AI decision-making processes
559 would address current limitations in model transparency and support broader adoption
560 in management contexts.

561 4.5 Conclusion

562 Our validation analysis demonstrates both the capabilities and limitations of AI-assisted
563 ecosystem modeling. The framework shows remarkable stability in many aspects while
564 highlighting areas of ecological uncertainty that deserve attention. The clear regional
565 patterns in performance suggest that the approach can adapt to different ecological con-
566 texts while maintaining scientific rigor. These findings support the framework’s utility for
567 ecosystem-based management while providing clear directions for future improvement.

568 The observed trade-offs between consistency and complexity reflect fundamental chal-
569 lenges in ecosystem modeling rather than simple methodological limitations. By quan-
570 tifying these trade-offs and their regional variations, our analysis provides a foundation
571 for more informed application of ecosystem models in fisheries management. As marine
572 ecosystems face increasing pressures from climate change and human activities, this un-
573 derstanding of model behavior and limitations becomes increasingly crucial for effective
574 ecosystem-based management ([Geary et al., 2020](#)).

575 Acknowledgements

576 [Add acknowledgements here]

577 Data Availability

578 The complete codebase, including all scripts, configuration files, and analysis tools, is
579 available at [GitHub repository URL]. The validation framework, including reference
580 group definitions and classification rules, is documented in the project repository to ensure
581 reproducibility.

582 Author Contributions

583 SS: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation,
584 Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visual-
585 ization, Project administration. BF: Validation, Writing - Review & Editing, Supervision,
586 Funding acquisition. FB: Methodology, Software, Validation, Writing - Review & Editing,
587 Supervision. CB: Investigation, Validation, Writing - Review & Editing. RB: Investiga-
588 tion, Validation, Writing - Review & Editing. JPG: Investigation, Validation, Writing -
589 Review & Editing. JS: Conceptualization, Validation, Investigation, Writing - Review &
590 Editing. RS: Investigation, Validation, Writing - Review & Editing. RT: Methodology,
591 Software, Validation, Investigation, Writing - Review & Editing, Supervision, Funding
592 acquisition.

593 Statement on the Use of Generative AI

594 Generative AI tools, specifically Claude Sonnet 3.5, were utilized in the preparation of this
595 manuscript to assist with tasks such as language refinement, text structuring, and summa-
596 rization. All scientific content, data interpretation, and conclusions were independently
597 developed and verified by the authors to ensure accuracy and integrity.

598 References

- 599 Anthropic (2024). Claude 3 model card. Technical documentation, Anthropic.
- 600 Chamberlain, S. (2020). *robis: An R Client for the Ocean Biodiversity Information Sys-*
601 *tem*. R package version 0.2.0.
- 602 Chen, M. and Xu, Z. (2024). A deep learning classification framework for research methods
603 of marine protected area management. *Journal of Environmental Management*.
- 604 Christensen, V. and Walters, C. J. (2004). Ecopath with Ecosim: methods, capabilities
605 and limitations. *Ecological Modelling*, 172(2-4):109–139.
- 606 Chroma (2024). Chroma - the ai-native open-source embedding database. Accessed: 2024.

- 607 Coll, M., Akoglu, E., Arreguín-Sánchez, F., Fulton, E. A., Gascuel, D., Heymans, J. J.,
608 Libralato, S., Mackinson, S., Palomera, I., Piroddi, C., et al. (2015). Modelling dynamic
609 ecosystems: venturing beyond boundaries with the Ecopath approach. *Reviews in Fish
610 Biology and Fisheries*, 25(2):413–424.
- 611 Colléter, M., Valls, A., Guittot, J., Gascuel, D., Pauly, D., and Christensen, V. (2015).
612 Global overview of the applications of the Ecopath with Ecosim modeling approach
613 using the EcoBase models repository. *Ecological Modelling*, 302:42–53.
- 614 Fernandes, S. and D’Mello, A. (2024). Artificial intelligence in the aquaculture industry:
615 Current state, challenges and future directions. *Aquaculture*, page 742048.
- 616 Froese, R., Pauly, D., et al. (2010). FishBase.
- 617 GBIF (2024). GBIF: The global biodiversity information facility. *Global Biodiversity
618 Information Facility*. What is GBIF?
- 619 Geary, W. L., Bode, M., Doherty, T. S., Fulton, E. A., Nimmo, D. G., Tulloch, A. I.,
620 Tulloch, V. J., and Ritchie, E. G. (2020). A guide to ecosystem models and their
621 environmental applications. *Nature Ecology & Evolution*, 4:1459–1471.
- 622 Grassle, J. F. and Stocks, K. (1999). A global ocean biogeographic information system
623 (OBIS) for the census of marine life. *Oceanography*, 12(3):12–14.
- 624 Guo, J., Li, N., and Xu, M. (2025). Environmental large language model evalua-
625 tion (ELLE) dataset: A benchmark for evaluating generative AI applications in eco-
626 environment domain. *arXiv preprint arXiv:2501.06277*.
- 627 Holden, M. H., Akinlotan, M. D., Binley, A. D., Cho, F. H., Helmstedt, K. J., and
628 Chadès, I. (2024a). Why shouldn’t I collect more data? Reconciling disagreements
629 between intuition and value of information analyses. *Methods in Ecology and Evolution*,
630 15:1580–1592.
- 631 Holden, M. H., Plagányi, É. E., Fulton, E. A., Campbell, A. B., Janes, R., Lovett, R. A.,
632 Wickens, M., Adams, M. P., Botelho, L. L., Dichmont, C. M., et al. (2024b). Cost–
633 benefit analysis of ecosystem modeling to support fisheries management. *Journal of
634 Fish Biology*, 104:1667–1674.
- 635 Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L.
636 (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440.
- 637 Kelling, S., Hochachka, W. M., Fink, D., Riedewald, M., Caruana, R., Ballard, G., and
638 Hooker, G. (2009). Data-intensive science: a new paradigm for biodiversity studies.
639 *BioScience*, 59(7):613–620.
- 640 Kommineni, V. K., König-Ries, B., and Samuel, S. (2024). Harnessing multiple LLMs
641 for information retrieval: A case study on deep learning methodologies in biodiversity
642 publications. *arXiv preprint arXiv:2411.09269*.
- 643 Kuhn, B., Cayetano, A., and Fincham, J. (2024). Machine learning applications for
644 fisheries—at scales from genomics to ecosystems. *Reviews in Fisheries Science*.

- 645 Lapeyrolerie, M., Chapman, M. S., Norman, K. E., and Boettiger, C. (2022). Deep
646 reinforcement learning for conservation decisions. *Methods in Ecology and Evolution*,
647 13:2649–2662.
- 648 Li, H., Liu, J., Wang, Z., Luo, S., Jia, X., and Yao, H. (2024). Lite: Modeling environ-
649 mental ecosystems with multimodal large language models. *Conference on Language*
650 *Models (COLM)*.
- 651 Michener, W. K. and Jones, M. B. (2012). Ecoinformatics: supporting ecology as a
652 data-intensive science. *Trends in Ecology & Evolution*, 27(2):85–93.
- 653 Poelen, J. H., Simons, J. D., and Mungall, C. J. (2014). Global biotic interactions:
654 An open infrastructure to share and analyze species-interaction datasets. *Ecological*
655 *Informatics*, 24:148–159.
- 656 Tuia, D., Kellenberger, B., Beery, S., Costelloe, B. R., Zuffi, S., Risso, B., Mathis, A.,
657 Mathis, M. W., van Langevelde, F., Burghardt, T., et al. (2022). Perspectives in
658 machine learning for wildlife conservation. *Nature Communications*, 13:1–15.
- 659 Villasante, S., Arreguín-Sánchez, F., Heymans, J. J., Libralato, S., Piroddi, C., Chris-
660 tensen, V., and Coll, M. (2016). Modelling complex systems of multiple species for
661 ecosystem based management. *Ecological Modelling*, 326:68–76.
- 662 Zheng, Z., Zhang, J., Vu, T. A., Diao, S., Tim, Y. H. W., and Yeung, S. K. (2023).
663 MarineGPT: Unlocking secrets of ocean to the public. *arXiv preprint arXiv:2310.13596*.

664 **Supplementary Material**

665 **S1 Data Harvesting Implementation**

666 Our data harvesting system employs DuckDB for efficient querying of PARQUET files,
667 enabling complex joins and aggregations without full memory loading. For species matching
668 across databases, we use structured SQL queries that join on concatenated genus and
669 species names:

670 `SELECT * FROM sealifebase_df`
671 `WHERE SpecCode = {spec_code}`
672 `AND PreyStage LIKE '%adult%'`
673 `AND PredatorStage LIKE '%adult%'`

674 When combining interaction data from GLOBI with diet information, we implement
675 a comprehensive interaction mapping system that creates bidirectional records:

676 `interaction_data[source_group]['preys_on'][target_group] = count`
677 `interaction_data[target_group]['is_preyed_on_by'][source_group] = count`

678 Our data cleaning protocol standardizes types by converting numerical values to con-
679 sistent formats and timestamps to ISO format. We handle null values by removing empty
680 values, ‘NA’ strings, and null entries while preserving data structure. Source tracking
681 maintains database origin information for all data points.

682 The system implements file locking mechanisms for concurrent access, with separate
683 locks for species data and interaction networks. We use exponential backoff retry logic
684 for API interactions, with configurable parameters including maximum retries (5), initial
685 delay (1 second), and maximum delay (60 seconds).

686 The completion check system verifies the presence of required fields including:

- 687 • Complete taxonomic hierarchy
- 688 • Species-specific database records (when available)
- 689 • Interaction data
- 690 • Source attribution
- 691 • Data quality indicators

692 The final JSON output maintains a consistent structure across all species entries,
693 facilitating automated processing in subsequent framework stages.

694 **S2 Diet Matrix Analysis**

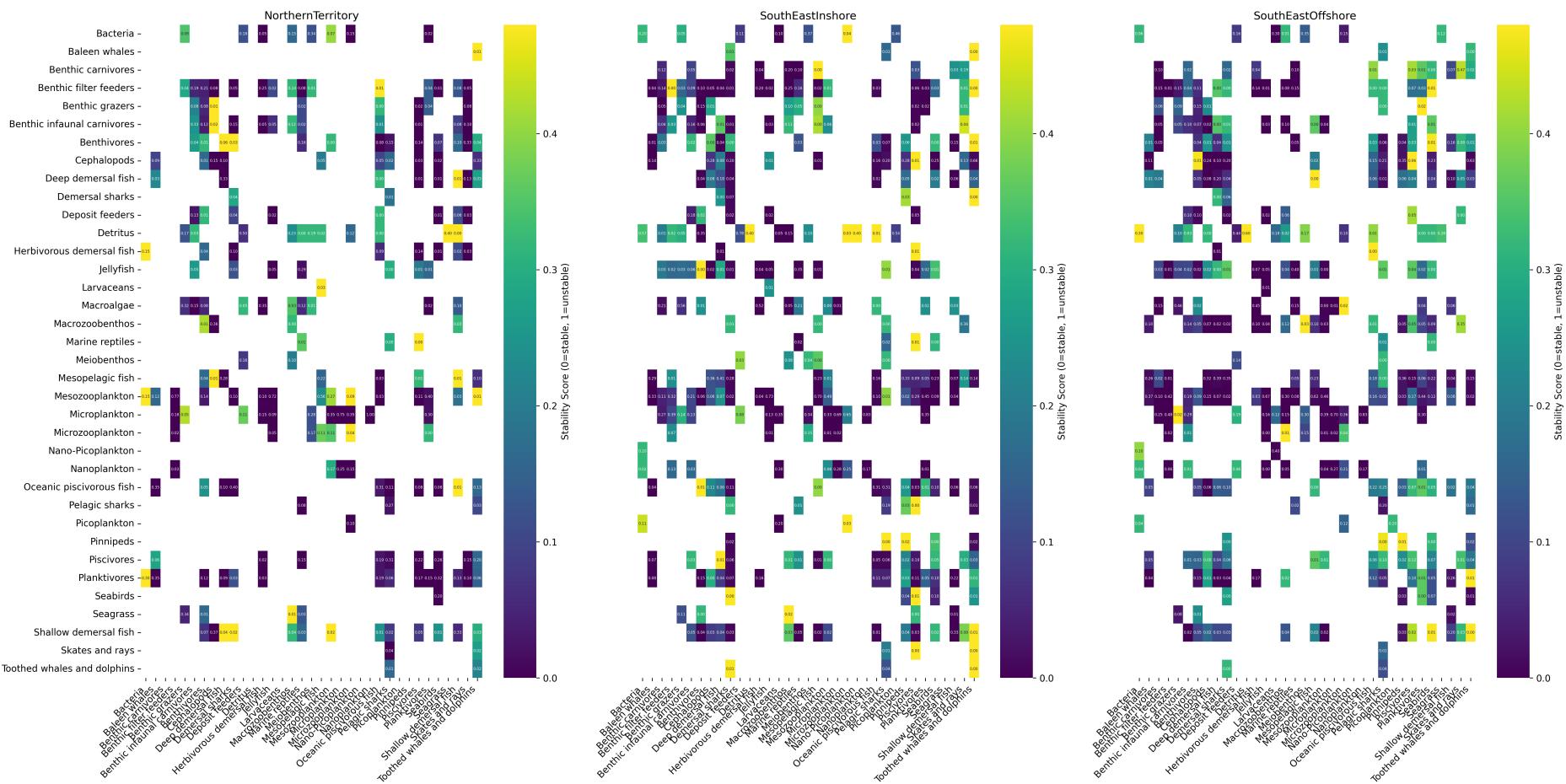


Figure S1: Detailed diet matrix consistency across five iterations for each geographic region. Column names represent predator groups and row names represent their prey groups. Numbers in each cell indicate the mean diet proportions across five iterations, while cell colors indicate the stability score (0-1, where 0 represents perfect stability and 1 represents maximum variation). White cells represent absent feeding relationships. This comprehensive visualization complements the stability score distributions and predator-specific analyses presented in the main text (Figures 4 and 5).

695 S3 Technical Implementation

696 S3.1 Default Grouping with Descriptions

697 Table S1 presents the complete template of potential functional groups used by the system.
698 This template serves as a reference for group classification, though the system can create
699 new groups or modify existing ones based on specific ecosystem characteristics.

Table S1: Complete Functional Group Template

Group Name	Description
Skates and rays	Bottom-dwelling cartilaginous fish that play a role in controlling benthic prey populations
Nearshore and smaller seabirds	Small gulls, terns etc that feed near shore (possibly include penguins here too) - avian predators that link marine and terrestrial ecosystems
Albatrosses	Large seabirds that forage exclusively at sea, feeding on marine prey (fishes, squids, gelatinous organisms)
Skuas and giant petrels	Large predatory seabirds that feed both at sea and on land, including predation on other birds
Fish-eating pinnipeds	Marine mammals (seals, sea lions) that primarily prey on fish in coastal and pelagic ecosystems
Invertebrate-eating pinnipeds	Marine mammals (particularly Antarctic seals) that primarily feed on krill and other invertebrates
Baleen whales	Large filter-feeding marine mammals that regulate zooplankton populations and contribute to nutrient cycling
Orcas	Apex predators that uniquely prey upon other top predators including marine mammals, sharks, and large fish
Sperm whales	Deep-diving cetaceans that primarily feed on deep-water squid and fish
Small toothed whales and dolphins	Smaller cetaceans that primarily feed on fish and squid in surface and mid-waters
Sea snakes	Marine reptiles that prey primarily on fish, particularly eels and fish eggs
Crocodiles	Large predatory reptiles in coastal and estuarine waters that prey on fish, birds, and mammals
Turtles	Herbivores and omnivores that breed on land
Planktivores	Small fishes that feed on plankton, crucial in transferring energy from plankton to larger predators
Flying fish	Epipelagic fish capable of gliding above the water surface, important prey for many predators
Remoras	Fish that form commensal relationships with larger marine animals, feeding on parasites and food scraps
Large oceanic piscivorous fish	Fish-eating predators in open ocean environments, mid-sized non-migratory species (e.g. barracuda)
Tuna and Billfish	Large oceanic predatory fish, highly mobile, often dive to feed deeper into the water column

Continued on next page

Table S1 – Continued

Group Name	Description
Shelf small benthivores	Small bodied fish that feed on benthic organisms, playing a key role in benthic-pelagic coupling, live in shelf waters
Shelf demersal omnivorous fish	Medium sized demersal fish that feed on invertebrates as well as smaller fish, live in shelf waters
Shelf medium demersal piscivores	Medium sized demersal fish living near the bottom in shallow waters, often important in benthic food webs, feed on other fish primarily, live in shelf waters
Shelf large piscivores	Fish-eating predatory fishes found in various marine habitats, important in controlling prey fish populations
Herbivorous demersal fish	Bottom-associated fish that primarily feed on plants, important in controlling algal growth
Slope/deep water benthivores	Small to mid sized fish that feed on benthic organisms and live on the shelf or seamounts
Slope/deep demersal omnivorous fish	Medium sized demersal fish that feed on invertebrates as well as smaller fish, live in slope or seamount waters
Slope/deep medium demersal piscivores	Medium sized demersal fish that feed on other fish primarily, live in slope or seamount waters
Slope/deep large piscivores	Fish-eating predatory fishes found in various marine habitats in deeper water, live in slope or seamount waters
Migratory mesopelagic fish	Fish living in the mesopelagic zone, undertake diel vertical migration, important in energy transfer between depths
Non-migratory mesopelagic fish	Fish living in the mesopelagic zone, non-migratory species, important in energy transfer between depths
Reef sharks	Top predators in coral reef ecosystems, controlling fish populations and maintaining reef health
Pelagic sharks	Open-ocean predators that help regulate populations of fishes and squids
Demersal sharks	Bottom-dwelling sharks, including dogfishes, that control populations of fishes and invertebrates on and near the seafloor
Cephalopods	Intelligent mollusks like squid and octopus, important predators in many marine ecosystems
Hard corals	Reef-building colonial animals that create complex habitat structure through calcium carbonate deposition
Soft corals	Colonial animals that contribute to reef habitat complexity without building calcium carbonate structures
Sea anemones	Predatory anthozoans that can form symbiotic relationships with fish and crustaceans
Hydrothermal vent communities	Specialized organisms living around deep-sea vents, including chemosynthetic bacteria and associated fauna
Cold seep communities	Organisms adapted to methane and sulfide-rich environments on the seafloor
Deep-sea glass sponges	Filter-feeding animals that create complex deep-water habitats and are important in silicon cycling
Sea cucumbers	Deposit-feeding echinoderms important in sediment processing and bioturbation

Continued on next page

Table S1 – Continued

Group Name	Description
Sea urchins	Herbivorous echinoderms that can control macroalgal abundance and affect reef structure
Crown-of-thorns starfish	Coral-eating sea stars that can significantly impact reef health during population outbreaks
Benthic filter feeders	Bottom-dwelling organisms that filter water for food, important in nutrient cycling and regulating water quality in various depths - bivalves, crinoids, sponges
Macrozoobenthos	Mobile large bottom-dwelling invertebrates in both shallow and deep waters, important in benthic food webs and bioturbation (predatory or omnivorous)
Benthic grazers	Bottom-dwelling organisms that graze on algae and detritus, influencing benthic community structure
Prawns	Small crustaceans that are important in benthic and pelagic food webs
Meiobenthos	Tiny bottom-dwelling organisms, important in sediment processes and as food for larger animals
Deposit feeders	Animals that feed on organic matter in sediments, important in nutrient cycling
Benthic infaunal carnivores	Predatory animals living within the seafloor sediments
Sedimentary Bacteria	Microscopic organisms crucial in nutrient cycling and the microbial loop in marine ecosystems
Large carnivorous zooplankton	Fish larvae, arrow worms and other large predatory zooplankton
Antarctic krill	Key species in Antarctic food webs, particularly important as prey for whales, seals, and seabirds
Ice-associated algae	Microalgae living within and on the underside of sea ice, important primary producers in polar regions
Ice-associated fauna	Specialized invertebrates living in association with sea ice, important in polar food webs
Mesozooplankton	Medium-sized zooplankton (200 µm to 2 cm) that feed on smaller plankton and serve as food for larger animals
Microzooplankton	Tiny zooplankton (20 µm to 200 µm) that graze on phytoplankton and bacteria, forming a crucial link in the microbial food web
Pelagic tunicates	Including larvaceans, salps, and pyrosomes, important in marine snow formation and carbon cycling
Jellyfish	Predatory gelatinous species
Diatoms	Larger phytoplankton (20 µm to 200 µm), silica dependent important primary producers in marine ecosystems
Dinoflagellates	Mixotrophic species (20 µm to 200 µm) that can switch between primary production and consumption as needed
Nanoplankton	Plankton ranging from 2 µm to 20 µm in size, including small algae and protozoans
Picoplankton	Plankton ranging from 0.2 µm to 2 µm in size, including both photosynthetic and heterotrophic organisms

Continued on next page

Table S1 – Continued

Group Name	Description
Microalgae (microphyto-benthos)	Microscopic algae that live on the seafloor or attached to other organisms
Pelagic bacteria	Watercolumn dwelling bacteria, consume marine snow amongst other things
Seagrass	Marine flowering plants that form important coastal habitats and nursery areas
Mangroves	Salt-tolerant trees forming critical coastal nursery habitats and protecting shorelines
Salt marsh plants	Coastal vegetation adapted to periodic flooding, important in nutrient cycling and shoreline protection
Macroalgae	Seaweeds of various sizes that provide habitat and food for many species, including both canopy and understory forms
Symbiotic zooxanthellae	Photosynthetic dinoflagellates living within coral and other marine invertebrates
Cleaner fish and shrimp	Species that remove parasites from other marine animals, important in reef health
Discards	Carriion and freshly discarded material from fisheries activities
Detritus	Labile components of natural death and waste

700 S3.2 Retrieval-Augmented Generation Implementation

701 We implement a retrieval-augmented generation system using ChromaDB for vector stor-
 702 age and document management. Document processing begins with LlamaParse conversion
 703 of source materials to markdown format, preserving structural elements while enabling
 704 consistent text extraction across document types. We segment documents using a token-
 705 aware chunking strategy with a 2000-token maximum size, determined through empirical
 706 testing to balance context preservation with model limitations.

707 Document processing follows a two-phase approach. The initial phase generates em-
 708 beddings for each document chunk using Azure OpenAI’s text-embedding-3-small model,
 709 storing them in ChromaDB’s PersistentClient. The system maintains an indexed_files.json
 710 registry to track processed documents. The second phase handles incremental updates,
 711 identifying and processing only new content when documents are added to the source
 712 directory.

713 For diet composition analysis, we implement a two-stage query process. The first stage
 714 employs a simple query to retrieve relevant document chunks:

What do [group] eat?

715 The system embeds this query using the same Azure OpenAI model and performs
 716 vector similarity search to identify relevant document chunks. These results combine
 717 with structured data sources including species occurrence frequencies, food category clas-
 718 sifications, and GLOBI interaction data to form a comprehensive input for the second
 719 stage.

720 We implement comprehensive error handling throughout the pipeline. The system em-
 721 ploys exponential backoff retry logic for API interactions, with configurable parameters in-

722 cluding maximum retries (10), initial delay (1 second), and maximum delay (300 seconds).
723 For model interactions, we utilize LlamaIndex’s query engine with zero-temperature sam-
724 pling to ensure deterministic responses. The system supports multiple language model
725 backends including Claude-3 Sonnet (200k token context), GPT-4, and AWS Claude,
726 enabling flexible deployment based on availability and performance requirements.

727 The system maintains separate storage contexts for different document collections
728 through ChromaDB’s collection management. This separation prevents cross-contamination
729 between knowledge bases while enabling efficient parallel processing. We track document
730 citations throughout the retrieval process, maintaining provenance information for all re-
731 trieval content. The complete implementation, including embedding generation, chunking
732 algorithms, and query processing functions, is available in the project repository.