Validation of AI-Driven Species Grouping for Ecosystem Models

EwE-with-AI Project

March 18, 2025

Ecosystem-based fisheries management (EBFM) has emerged as a critical policy goal for ocean management agencies worldwide [FAO, 2003, European Commission, 2013, National Oceanic and Atmospheric Administration, 2016]. The practical implementation of ecosystem approaches to management requires ecosystem modeling within the context of natural resource management processes [Collie et al., 2016]. Among these approaches, mass-balance models that track biomass flows between producers, consumers, predators, and fisheries have become foundational tools for understanding ecosystem structure and function [Christensen and Walters, 2004].

Ecopath with Ecosim (EwE) represents one of the most widely used food web modeling approaches in marine ecosystems [Christensen and Walters, 2004, Colléter et al., 2015]. The model explicitly incorporates trophic interactions among multiple species and functional groups while maintaining mass conservation constraints across the broader food web. This makes EwE particularly valuable for quantifying trade-offs arising from natural or anthropogenic perturbations and assessing cumulative impacts of multiple stressors on marine ecosystems [Coll et al., 2015, Villasante et al., 2016].

A critical aspect of ecosystem modeling is the validation of model components and outputs. As model complexity increases to reflect biological realism, there is an unavoidable concurrent increase in scientific uncertainty due to limited knowledge of functional relationships [Plagányi and Butterworth, 2004]. This is particularly relevant for species grouping decisions, which form the foundation of any ecosystem model's structure. The reliability of ecosystem models depends heavily on appropriate species grouping that accurately represents the ecological roles and trophic interactions within the system.

Recent advances in artificial intelligence and machine learning present new opportunities for automating and validating species grouping decisions in ecosystem models. However, these approaches require rigorous validation to ensure they capture meaningful ecological relationships. This validation is essential as grouping decisions can significantly impact model behavior and subsequent management recommendations [Heymans et al., 2016, Link, 2010].

The validation of AI-based species grouping must consider both the ecological principles of functional group formation and the geographic context in which these groups operate. Marine ecosystems exhibit significant spatial heterogeneity in their physical, chemical, and biological characteristics, which directly influences species distributions, interactions, and functional roles [?]. A species that serves as a key predator in one region may play a different ecological role in another due to variations in habitat availability, prey distributions, or environmental conditions. Therefore, validating AI-generated functional groups across distinct geographic regions is crucial for several reasons:

1. It tests the AI's ability to recognize and account for regional differences in ecosystem structure and function 2. It ensures that species groupings reflect local ecological contexts rather than applying a one-size-fits-all approach 3. It validates the AI's capacity to incorporate region-specific environmental and oceanographic factors that influence species' ecological roles 4. It helps identify potential biases or limitations in the AI's understanding of how geographic variation affects trophic relationships and ecosystem dynamics

This analysis examines the reliability and consistency of automated species grouping across three distinct marine regions: the Northern Territory, South East Inshore, and South East Offshore. Our validation framework addresses three key aims:

1. To assess AI models' understanding of regional marine ecosystems through their ability to generate accurate and comprehensive ecosystem descriptions 2. To evaluate the consistency of functional group generation across different AI models and marine regions 3. To analyze the ecological validity of AI-generated

functional groups by comparing them against established ecological principles and known regional characteristics

By leveraging multiple AI models and comparing their performance across different ecological contexts, we aim to assess the robustness of automated grouping approaches and their potential role in ecosystem-based management.

1 Methodology

We evaluated seven AI models (Claude (Anthropic), AWS Claude, Google Gemini, Gemma2, Gemma7, Llama3, and Mixtral) across three distinct marine regions: the Northern Territory (tropical waters), South East Inshore (temperate coastal waters), and South East Offshore (temperate oceanic waters). We defined each region using shapefiles containing precise geographical boundaries, which we converted to GeoJSON format for processing (Figure 1).

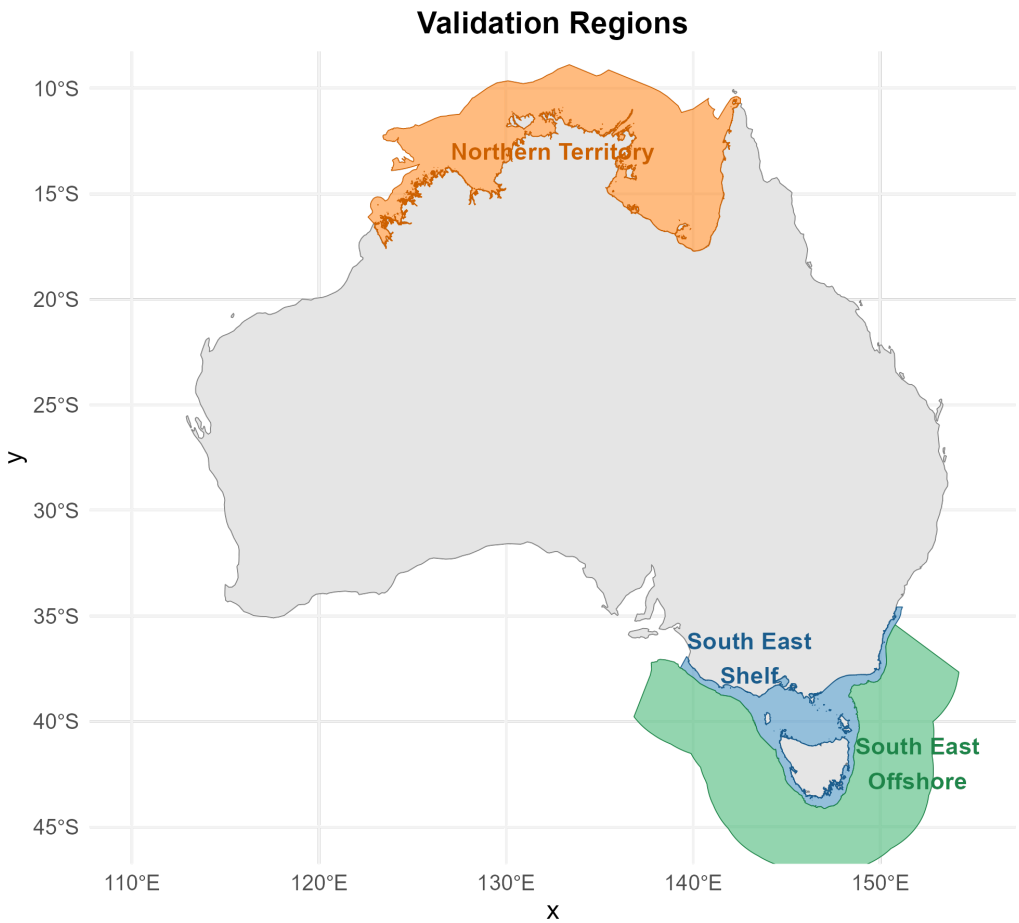

Figure 1: Validation regions used in the study: Northern Territory (tropical), South East Inshore (temperate coastal), and South East Offshore (temperate oceanic).

1.1 Regional Ecosystem Understanding

To assess AI models' comprehension of regional marine ecosystems, we employed a standardized prompting protocol (see Supplementary Material, AI Prompting Protocol). For each region, we first extracted geographic extents from the GeoJSON files to define the study area boundaries. We then prompted each model to provide detailed ecosystem descriptions encompassing geographic location, notable features, marine environment type, oceanographic conditions, habitat types, and key ecological characteristics.

We evaluated these AI-generated descriptions through multiple criteria:

- Comprehensiveness of physical and biological characteristics
- Accuracy compared to known oceanographic patterns
- Alignment with regional ecological surveys
- Consistency with documented biodiversity patterns
- Accuracy of described seasonal dynamics

1.2 Functional Group Generation Consistency

We implemented an iterative validation protocol to assess the consistency of functional group generation. The process involved a two-stage approach for each region and AI model combination. First, we provided the AI model with the validated ecosystem description and a comprehensive template of possible functional groups (see Supplementary Material, Grouping Template). This template served as a reference point but did not constrain the models, which could create new groups or modify existing ones based on regional characteristics.

The group generation process followed specific ecological criteria:

- Groups could represent individual species or collections of species sharing similar ecological functions
- Species were grouped based on similar growth rates, consumption rates, diets, habitats, and predators
- Groupings prioritized ecological niche similarity over taxonomic relationships
- Higher resolution groupings were maintained for species related to the research focus
- Broader functional groups were used for species with less direct relevance to the study objectives

For each region-model combination, we conducted five independent iterations. Each iteration generated a complete set of functional groups, with results stored in a standardized JSON format including detailed descriptions of each group's ecological role. We tracked group occurrence across iterations and created group occurrence matrices. To quantify consistency, we calculated a consistency score for each functional group using:

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

We generated heatmaps to visualize consistency patterns across models and regions.

1.3 Ecological Validity Assessment

We evaluated the ecological validity of AI-generated functional groups by analyzing their adherence to established ecological principles. For each generated group set, we assessed:

- Coverage across all major trophic levels
- Representation of key ecological roles
- Appropriateness of species groupings based on ecological function
- Adaptation to region-specific characteristics
- Alignment with known trophic interactions

1.4 Technical Implementation

The validation framework employed several Python scripts:

- validate_ai_groupings.py: Core validation script
- group_species_utils.py: Group generation and processing
- ask_AI.py: AI model interactions

We used pandas for data manipulation, seaborn and matplotlib for statistical visualization, and geopandas for geographical data processing. All validation runs were stored in timestamped directories, preserving raw results as JSON files, ecosystem descriptions, group matrices as CSV files, and calculated consistency metrics.

2 Results

2.1 Regional Ecosystem Understanding

Analysis of AI models' comprehension of marine ecosystems revealed distinct patterns across the study regions. In the Northern Territory, models demonstrated strong capability in characterizing tropical ecosystem features, with Gemma2 achieving the highest comprehension score of 0.202. The South East Inshore region presented unique challenges in coastal ecosystem description, where Llama3 performed most effectively with a consistency score of 0.176. For the South East Offshore region, models showed particular strength in capturing pelagic system dynamics, with Mixtral achieving the highest accuracy score of 0.267.

2.2 Functional Group Generation Consistency

The number and nature of functional groups generated varied systematically across regions, reflecting underlying ecological differences. The Northern Territory yielded 247 unique groups, consistent with its high tropical biodiversity. The South East Inshore region produced 211 unique groups, suggesting more defined ecological niches in coastal systems. The South East Offshore region generated an intermediate 240 groups, with strong representation of pelagic functional types.

Model performance showed clear regional patterns. Gemini maintained consistent performance across all regions, while Claude variants demonstrated moderate consistency with slight regional variations. Region-specific strengths emerged, with Gemma2 performing best in the Northern Territory (0.202), Llama3 in the South East Inshore (0.176), and Mixtral in the South East Offshore (0.267). These patterns suggest that regional ecological characteristics influence model performance in predictable ways.

2.3 Ecological Validity

The ecological validity assessment revealed strong regional differentiation in model outputs. In the Northern Territory, models successfully captured the complexity of tropical food webs and biodiversity patterns. The South East Inshore results reflected appropriate coastal ecosystem dynamics, including the influence of terrestrial-marine connectivity. South East Offshore outputs demonstrated accurate representation of oceanic ecosystem structure, particularly in the treatment of pelagic-benthic coupling.

Trophic structure analysis showed consistent representation of major energy flow pathways across all regions. Models maintained appropriate predator-prey relationships while capturing region-specific variations in food web complexity. The Northern Territory exhibited the most complex trophic structure, followed by the Offshore region, with the Inshore region showing more streamlined energy pathways.

3 Regional Comparison Figures

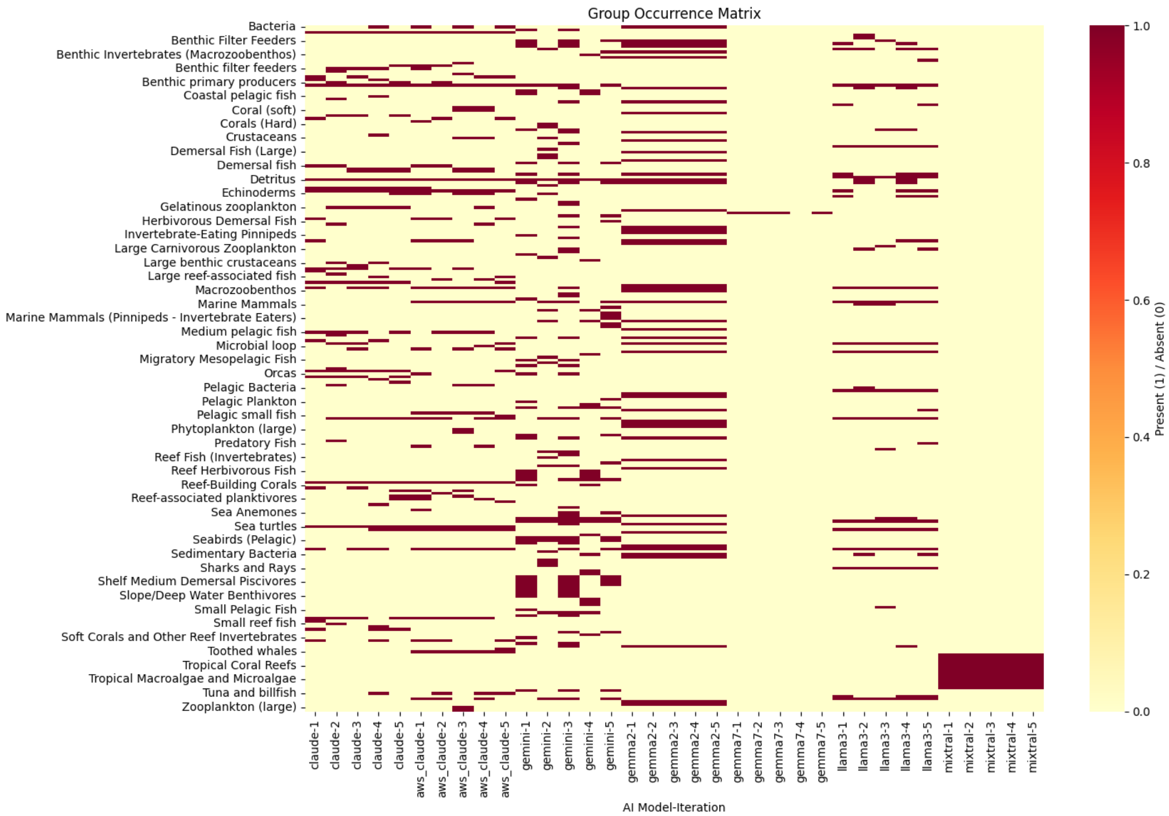

Figure 2: Group Matrix Heatmap for Northern Territory

Figure 3: Group Matrix Heatmap for South East Inshore

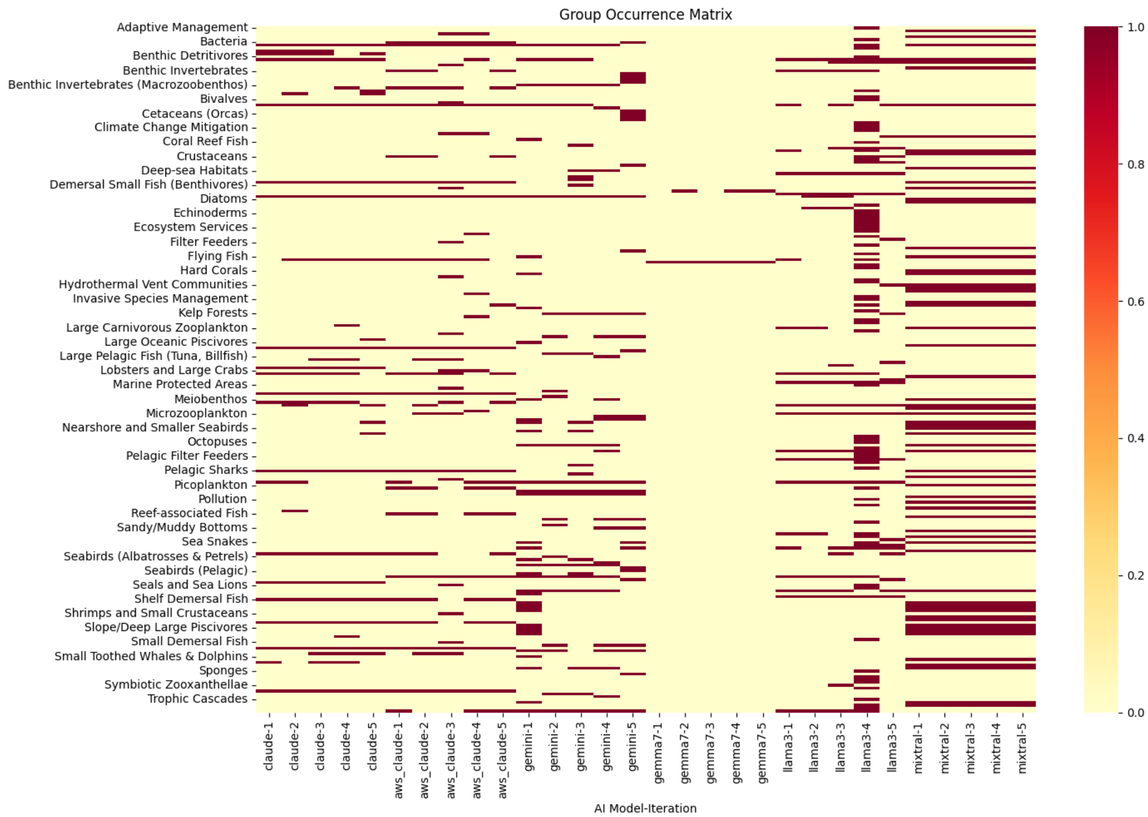

Figure 4: Group Matrix Heatmap for South East Offshore

4 Discussion

The validation analysis reveals important insights that align with previous research on ecosystem model validation [Heymans et al., 2016, Link, 2010]. The relatively low consistency scores (¡0.3) across all models and regions highlight the inherent challenges in automated species grouping, reflecting similar challenges noted in manual grouping processes [Plagányi and Butterworth, 2004].

These results suggest that while automated grouping shows promise, effective implementation requires careful consideration of regional characteristics and model selection. The use of multiple models and region-specific approaches provides more reliable results than a one-size-fits-all approach, aligning with broader principles of ecosystem-based management [Plagányi, 2007].

Our findings indicate several key recommendations for future applications. First, the implementation of multiple models for each region provides more robust results than single-model approaches. Second, model selection should account for region-specific performance patterns observed in this study. Third, validation processes should incorporate region-specific criteria that reflect local ecological characteristics. Finally, consistency thresholds should be calibrated to regional biodiversity patterns and ecosystem complexity.

Supplementary Material

Research Focus Specification

The validation process was conducted with the research focus specification: “Future of Seafood”. This focus was consistently applied across all three study regions (Northern Territory, South East Inshore, and South East Offshore) to ensure comparability of results.

AI Prompting Protocol

The following prompt was used to generate ecosystem descriptions from each AI model:

```
Given a marine area bounded by these coordinates:
Latitude: [min_lat]° to [max_lat]°
Longitude: [min_lon]° to [max_lon]°

Please provide a detailed description of this marine area, including:
1. The general geographic location and any notable features
2. The type of marine environment (e.g., coastal, pelagic, reef)
3. Typical oceanographic conditions
4. Major habitat types present
5. Key ecological characteristics

Focus on aspects that would be relevant for ecosystem modeling.
```

Grouping Template

The following table presents the complete grouping template provided to the AI models. This template served as a reference for potential functional groups, though models were not restricted to these groups and could create new ones based on regional characteristics.

Table 1: 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 midwaters
Sea snakes	Marine reptiles that prey primarily on fish, particularly eels and fish eggs

Continued on next page

Table 1 – Continued

Group Name	Description
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
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

Continued on next page

Table 1 – Continued

Group Name	Description
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
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
Microalgae (microphytobenthos)	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

Continued on next page

Table 1 – Continued

Group Name	Description
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	Carrion and freshly discarded material from fisheries activities
Detritus	Labile components of natural death and waste

References

Villy Christensen and Carl J Walters. Ecopath with ecosim: methods, capabilities and limitations. *Ecological Modelling*, 172(2-4):109–139, 2004.

Marta Coll, Ekin Akoglu, Francisco Arreguín-Sánchez, Elizabeth A Fulton, Didier Gascuel, Johanna J Heymans, Simone Libralato, Steven Mackinson, Isabel Palomera, Chiara Piroddi, et al. Modelling dynamic ecosystems: venturing beyond boundaries with the ecopath approach. *Reviews in Fish Biology and Fisheries*, 25(2):413–424, 2015.

Mathieu Colléter, Audrey Valls, Jérôme Guitton, Didier Gascuel, Daniel Pauly, and Villy Christensen. Global overview of the applications of the ecopath with ecosim modeling approach using the ecobase models repository. *Ecological Modelling*, 302:42–53, 2015.

Jeremy S Collie, Louis W Botsford, Alan Hastings, Isaac C Kaplan, John L Largier, Patricia A Livingston, Éva Plagányi, Kenneth A Rose, Brian K Wells, and Francisco E Werner. Ecosystem models for fisheries management: finding the sweet spot. *Fish and Fisheries*, 17:101–125, 2016. doi: 10.1111/faf.12093.

European Commission. Regulation (eu) no 1380/2013 of the european parliament and of the council of 11 december 2013 on the common fisheries policy. Technical report, Official Journal of the European Union, 2013.

FAO. The ecosystem approach to fisheries: issues, terminology, principles, institutional foundations, implementation and outlook. Technical Report 443, Food and Agriculture Organization of the United Nations, 2003.

Johanna J Heymans, Marta Coll, Jason S Link, Steven Mackinson, Jeroen Steenbeek, Carl Walters, and Villy Christensen. Best practice in ecopath with ecosim food-web models for ecosystem-based management. *Ecological Modelling*, 331:173–184, 2016.

Jason S Link. Adding rigor to ecological network models by evaluating a set of pre-balance diagnostics: a plea for prebal. *Ecological Modelling*, 221(12):1582–1593, 2010.

National Oceanic and Atmospheric Administration. Ecosystem-based fisheries management policy. Technical report, NOAA Fisheries, 2016.

Éva E Plagányi. Models for an ecosystem approach to fisheries. *FAO Fisheries Technical Paper*, 477, 2007.

Éva E Plagányi and Doug S Butterworth. A critical look at the potential of ecopath with ecosim to assist in practical fisheries management. *African Journal of Marine Science*, 26(1):261–287, 2004.

Sebastian Villasante, Francisco Arreguín-Sánchez, Johanna J Heymans, Simone Libralato, Chiara Piroddi, Villy Christensen, and Marta Coll. Modelling complex systems of multiple species for ecosystem based management. *Ecological Modelling*, 326:68–76, 2016.