Simplification of complex ecological networks – species aggregation in Antarctic food web models

S.A. McCormack [a, b], J. Melbourne-Thomas [b, c], R. Trebilco [b], J. L. Blanchard [a] and A. Constable [b, c]

[a] *Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia*

[b] *Antarctic Climate and Ecosystems Cooperative Research Centre, University of Tasmania, Hobart, Tasmania, Australia*

[c] *Australian Antarctic Division, Department of the Environment and Energy, Kingston, Tasmania, Australia*

Email: stacey.mccormack@utas.edu.au

Abstract: Marine ecosystem models assist in understanding ecosystem changes resulting from fisheries and climate change, making them important tools for forecasting and guiding sustainable management decisions. One of the greatest challenges in formulating these models is determining their optimum scope that includes a suitable level of model complexity whilst still facilitating valid and robust predictions. For food web models this often involves reducing the system to a set of functional groups representative of the major components of the biological web. The process for choosing these functional groups varies considerably among modelling studies with many inferences made about the general structure of ecosystems.

In the Antarctic, food webs are taxonomically diverse, structurally complex and extremely variable in space and time, and our knowledge of the trophic linkages and energy flow is still fragmentary. This makes the development of ecosystem models for the Antarctic particularly challenging. Here we show that diet observations can provide a useful basis for justifying decisions regarding species aggregation in food web models. Using the Southern Ocean Dietary Database we investigated the trophic linkages present in the food web of the high latitude region of East Antarctica to create a simplified food web structure. Our aim was to determine an optimum level of aggregation whilst maintaining a high degree of realism in regards to the trophic interactions present in the web. Using results from these analyses we then made informed decisions, following best-practice guidelines, to assign species to functional groups for an Ecopath model of Prydz Bay and the southern Kerguelen Plateau region. We discuss our reasoning behind dealing with structural uncertainty in this manner, in particular our decisions regarding the formation of functional groups relevant to the system. We conclude by summarizing the implications of using dietary data to inform species aggregation in food web models and how transparency is an important concept during model formulation and documentation to ensure the development of robust ecosystem models.

Keywords: *Food webs, functional groups, structural uncertainty, Ecopath*

1.   INTRODUCTION

Food webs are the fundamental structures around which ecological communities are assembled and function. Food web structure governs fluxes of energy and underpins key processes such as productivity, ecosystem stability and resilience. As Antarctic marine ecosystems continue to undergo significant environmental and physical change driven by global climate change and harvesting (IPCC, 2013), the future integrated management of these marine resources will rely heavily on explorations of food web structure and the major pathways of energy flow that govern their functioning (Murphy et al., 2012). Antarctic marine food webs are taxonomically diverse, structurally complex and extremely variable in space and time. Our knowledge of the trophic linkages and energy flow patterns in Antarctic food webs is still fragmentary making the formulation of ecosystem models for these systems a challenging yet essential task.

Ecosystem models assist in understanding the structure and functioning of marine food webs by creating simplifications of reality. Natural ecological systems are inherently complex and it is therefore not possible nor desirable to model them exactly (Fulton, 2001). Simpler models can be more effective than more complex versions, as the sensitivity of the model to parameterisation is more easily assessed due to relatively fewer parameters and resulting parameter combinations. They may also be easier for non-experts to comprehend, they require less input information and computational efficiency and can be more economically and logistically attractive. Simplifying the underlying food web, often through means of aggregation from numerous species to fewer key representative "groups", is the most common approach for ecosystem simplification in food web model development. This has the advantage of the model not being tied to individual species and therefore being more generic. There are negative consequences of aggregation, the most common one being the model is more robust to change and is therefore less likely to predict small scale changes within an ecosystem (Fulton, 2001). This goal of constructing simpler models raises questions regarding the optimal level of complexity, in particular the appropriate degree of species aggregation (Gardner et al. 1982; Fulton, 2001; Fulton et al., 2003).

Ecosystem models are constructed from conceptual models comprising hypotheses and assumptions regarding the ecosystem of interest. To ensure a model's structure is realistic, these assumptions are usually informed by available data from the real food web that indicate relationships between the objects within a model. In particular, observed trophic linkages can inform the most appropriate level of aggregation of species into groups. Species aggregation in food web models involves simplifying the underlying food web (i.e. the model) by combining organisms that share common predators, prey or other characteristics into functional groups. One of the most commonly used approaches for quantitative modelling of marine food webs is mass-balance network modelling of species and functional groups, in particular using the Ecopath with Ecosim tool (Polovina, 1984; Christensen and Pauly, 1992). Ecopath provides instantaneous snapshots of a food web, which can be used to depict the structure and function of aquatic ecosystems and to estimate changes in ecological indicators. Analysis of dietary information derived from an ecosystem of interest can provide useful information regarding species composition, similarities in species characteristics, dominant trophic linkages and general food web structure. This information provides a useful basis for decisions regarding species aggregation in food web models and highlights limitations in data prior to model parameterisation.

Model uncertainty often arises from incomplete information about how to represent real systems in models with model structure remaining one of the main sources of uncertainty. There will always be some degree of uncertainty regarding the most appropriate model structure for a particular research application with some uncertainty arising from biases and tendencies of the modelers themselves. For example, modelers that are more familiar with upper trophic levels may tend to aggregate the lower trophic levels to a higher degree compared with those familiar with bottom-up processes (see Pinnegar et al. 2005). Ecosystem modelers consider multiple informal model formulation criteria when developing ecosystem models (e.g. the number of elements to include in the model, the nature of functional relationships and the scaling of the model). These considerations are important components of the ecosystem modelling process although there are important tradeoffs among them (Fulton et al., 2003) with the optimal solution highly dependent on the research question being addressed. One method for dealing with such structural uncertainty is to maintain multiple alternative model structures and to then compare or combine their predictions in an ensemble type approach. Due to the subjective nature of the ecosystem modelling process, method documentation and justification are essential to ensure the limitations and assumptions within a model are transparent to external users. The process and justifications modelers use for species aggregation are rarely clearly documented or discussed despite the level of aggregation being one of the strongest determinants of model performance (Fulton et al., 2003).

This paper aims to investigate the trophic linkages, data availability and general food web structure of the East Antarctic oceanic region using the Southern Ocean dietary database (Raymond et al., 2011). We demonstrate how this information can then be used to inform the construction of functional groups through application to

an Ecopath model under development for Prydz Bay, Antarctica. We discuss our reasoning for dealing with structural complexity and uncertainty in this manner and present a decision framework that we used as a justification for the final functional group structure in Ecopath.

2. MODEL FORMULATION

2.1. The project

Both the Intergovernmental Panel on Climate Change (IPCC) and the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR) have highlighted the need for improved methods for assessing and predicting how Southern Ocean ecosystems are responding to climate change (IPCC, 2013). In order to address this challenge, the Antarctic Climate and Ecosystems Cooperative Research Centre (ACE CRC) proposed multiple strategies including the Ecopath model being developed for this project, which is a member of an ensemble of ecosystem models under development for assessing food web structure and energy flows for Australia's area of interest, the Indian sector of the Southern Ocean.
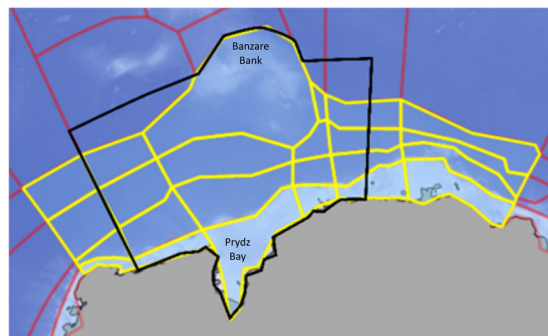

Figure 1. Ecopath model area (black) for Prydz Bay, Antarctica overlayed with the reduced domain for the implementation of an Atlantis ecosystem model (yellow).

Specifically, our Ecopath model of Prydz Bay aims to answer the following research questions:

- What are the contributions of the various mid-trophic energy pathways for transferring energy to higher trophic levels in Antarctic food webs? Particular focus is given to resolving the role of mesopelagic groups including fish, squid and zooplankton (copepods, gelatinous salps and krill).
- What trophic structures exist within the ecosystem and what eventual changes will result from environmental modifications (e.g. global climate change) and human activities (fisheries)?

2.2. The model domain: Prydz Bay, Antarctica

Prydz Bay is a high latitude embayment along the East Antarctic margin between 66°E and 79°E (Figure 1). For the purposes of this study, the boundaries of Prydz Bay and the surrounding oceanic region (i.e. the model domain) were defined as south of 60°S to the Antarctic continent and between 60°E and 90°E to encompass the Australian stations Mawson and Davis. The northern boundary of the model also encompasses Banzare Bank along the 3000m contour to capture important ecological processes relevant to the system and ensure congruence with the implementation of an Atlantis ecosystem model also currently under development (Figure 1). This region was chosen with data availability in mind; sampling that occurred in the same region as the model domain in early 2016 (*Aurora australis* voyage) provides updated biological information for model parameterisation.

3. SPECIES AGGREGATION

3.1. The Southern Ocean dietary database (SODD)

The Southern Ocean dietary database (SODD) is the only single consistent dietary data set for circum-Antarctic ecosystems (Raymond et al., 2011). The database consists of diet-related data from published and unpublished data sets and studies and is a combination of direct and indirect sampling methods of dietary assessment including gut, scat, and bolus content analysis, stomach flushing, and observed feeding. For the purposes of our study, we filtered the dataset to only include entries south of 60°S and within East Antarctica (between 30°E and 150°E). We included entries across all of East Antarctica to maximize the number of observations available for our analysis. Due to unavoidable inconsistencies in data reporting, the database contains various classifications for similar groups/species. To create a consistent list of species and groups we identified the lowest taxonomic resolution for each of the taxonomic groups within the refined list of observations. For example, many fish groups are entered into the database at the family or species level. Therefore to ensure the data entries were consistent, those reported at the species level were renamed to reflect the relevant family. This process resulted in 49 groups/trophospecies and 5507 observations for the East Antarctic region (Figure 2).

3.2. 'Best practice' guidelines for aggregation

Ecopath is a widely applied mass-balance food web modelling framework (Polovina, 1984; Christensen and Pauly, 1992) and one of many modelling approaches that requires the aggregation of system components into functional groups. There is no upper limit to the number or type of functional groups that can be included in Ecopath although many studies have suggested a range of useful best practice guidelines for creating functional groups in ecosystem models (e.g. Fulton, 2001; Heymans et al., 2016). The following guidelines were used as a basis for our aggregations:

1. Do not aggregate serially linked groups (i.e. predator and prey) (Gardner et al., 1982; Fulton, 2001)
2. Functional groups may be chosen with a specific policy question in mind, i.e. it may not be necessary to describe some groups in great detail when they do not apply to the specific research question being explored (Heymans et al., 2016)
3. Functionality and a lack of data may determine the level of aggregation for some groups (Heymans et al., 2016)

3.3. The decision framework

Using best practices as a basis, we constructed a decision framework consisting of seven attributes that we used during the aggregation of species into functional groups:

1. **Model scope**: This is one of the most important aspects of model formulation. Here is where we defined the research question to determine the biological and physical/spatial resolution of the model. The research question provides a foundation for determining which species/groups are important, which environments in the model domain require attention and if there are any aspects of the ecosystem that will impact the structuring of some of the functional groups (e.g. migrations).
2. **Data availability/limitations**: The second attribute we considered is the data available to parameterise the model. To include specific species in the model, it is ideal to have adequate data available for parameterisation. If there is no data available for a group/species, it may be more detrimental to include them. Here we identified data limitations through the presence and absence of species within the SODD, if there were species/groups that were not recorded within the database a literature search was conducted to make the final decision regarding their inclusion in the model.
3. **Taxonomy**: Taxonomic aggregation is not always an appropriate strategy to implement with regards to species aggregation although in this study it was an important step in forming plausible groups reflective of limitations in the data. We aimed to initially group to the lowest taxonomic resolution possible given the data.
4. **Spatial distributions**: The spatial extent of species/groups provides an easily identifiable divide between species within the model domain. Here we considered what environments each species/group inhabit and whether or not the species/group remain in the model domain for the whole year.
5. **Diet**: It is ideal to group species that share similar prey and predators (although this may not always be the case). For the remaining species we examined similarities in their diets using the dietary database to make justifiable and reproducible decisions regarding their representation in the model.
6. **Other characteristics**: There are other characteristics that can influence functional group structure such as life stage, size and behaviors. The main characteristic we considered for our decisions was size.
7. **Previous knowledge**: Before a final group structure is decided, previous findings from an ecosystem (or similar ecosystem) should be summarized to determine whether there are species or groups that may play unique roles within the ecosystem of interest and therefore require important consideration in regards to their representation.

3.4. Functional groups for Prydz Bay

The aggregation process for the food web of Prydz Bay, Antarctica resulted in 26 functional groups. Our justifications for each group are summarized in Table 1 below.

Table 1. Functional groups for the Ecopath model of Prydz Bay, Antarctica along with their components and justifications. Numbers following justifications correspond to the relevant decision attribute used to justify the aggregation (1 - model scope, 2 - data availability, 3 – taxonomy, 4 – spatial distribution, 5 – diet, 6 – other characteristics, 7 – previous knowledge).

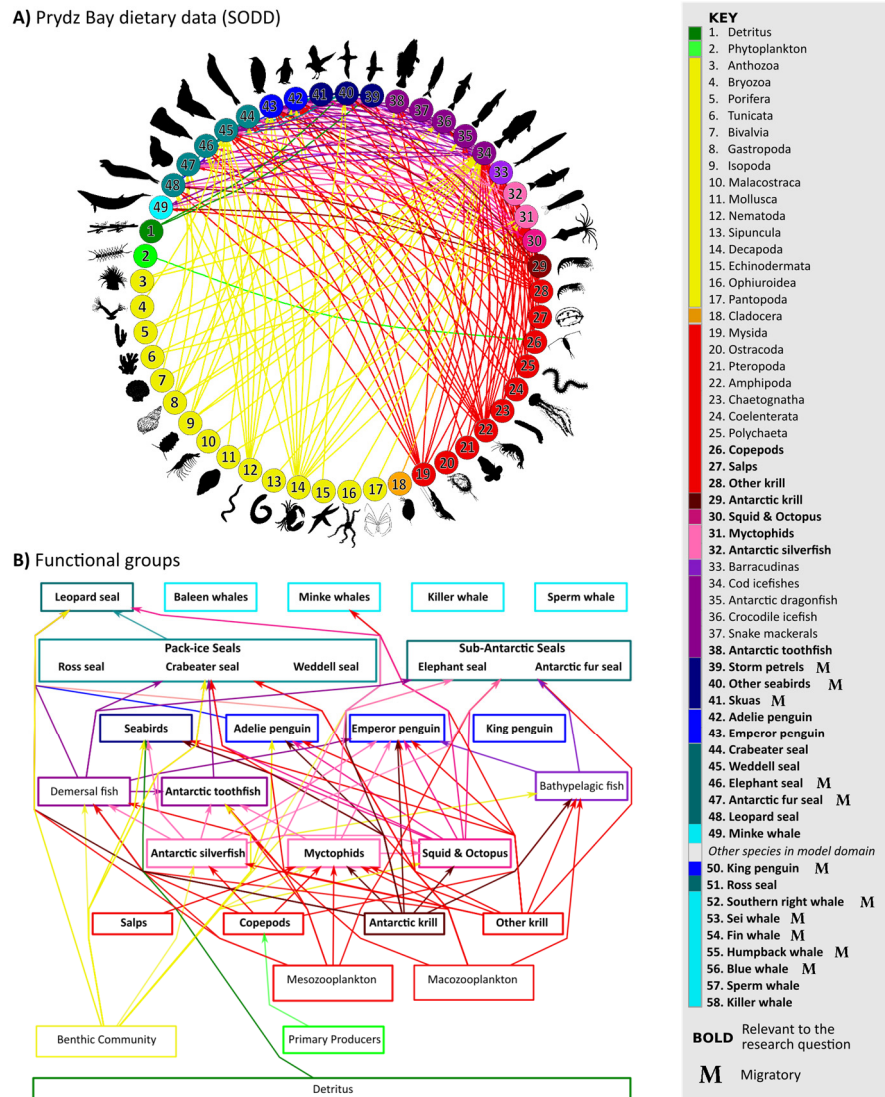| Functional group	Components	Justification
Detritus	Non-living material	It is a requirement to have a minimum of one detritus group in Ecopath.
Primary producers	Phytoplankton	Phytoplankton are generally highly aggregated in the diet database and therefore it would be difficult to parameterise this group at a higher resolution (2).
Benthic community	This group includes benthic invertebrates	Little to no relevance to the research question being addressed (1), occupy a similar habitat within the model domain (4)
Mesozooplankton	Order *Mysida*, Class *Ostracoda* and Superorder *Pteropoda*.	Defined according to size (0.2-20mm) (6) and diet (herbivorous) (5).
Macrozooplankton	Order *Amphipoda*, Phylum *Chaetognatha*, *Coelenterata* (phylas *Cnidaria* and *Ctenophora*) and Class *Polychaeta*	Defined according to size (greater than 20mm) (6) and diet (carnivorous) (5).
Salps	*Salpa thompsoni*	One of the primary organisms relevant to answering the research question (1).
Copepods	Subclass *Copepoda*	Important group for answering the research question (1).
Antarctic krill	*Euphausia superba*	Of interest to fisheries and relevant to answering the research question (1).
Other krill	Family *Euphausiidae*	Our analysis of circumpolar foodweb structure identified other krill as being important contributors to energy flow in the Indian sector of the Southern Ocean therefore this group was left at the family level to explore this (1)(3)(7).
Antarctic silverfish	*Pleuragramma antarcticum*	Antarctic silverfish play a "unique" role in Antarctic food webs (Pinkerton et al., 2010), therefore they have been separated from other notothenioids and included as one functional group (1)(6)(7).
Myctophids	Family *Myctophidae*	Important group for answering the research question (1)(3).
Squid & octopus	Class *Cephalopoda*	Important group for answering the research question (1). This group was left highly aggregated due to data limitations for other parameters (2)(3).
Demersal fish	Families *Nototheniidae*, *Bathydraconidae* and *Channichthyidae*	Little to no relevance to the research question being addressed (1), occupy a similar habitat within the model domain (4).
Antarctic Toothfish	*Dissostichus mawsoni*	Of interest to fisheries and are relevant to answering the research question (1).
Bathypelagic fish	Families *Gempylus* and *Paralepididae*	Grouped according to habitat (4) and general diet (5)
Seabirds	Flying birds that forage within the model domain	Known to migrate out of the system for a portion of the year (4), do not require high levels of disaggregation to answer the research question (1)
Adelie penguin	*Pygoscelis adeliae*	Left at species level as their diet is unique compared to other penguin species (5)
Emperor penguin	*Aptenodytes forsteri*	Despite similarities to the diet of king penguins, they remain in the model domain all year round (4)
King penguin	*Aptenodytes patagonicus*	Despite similarities to the diet of emperor penguins, they migrate out of the model domain for a portion of the year (4)
Pack-ice seals	*Ommatophoca rossii*, *Lobodon carcinophagus*, *Leptonychotes weddellii*	Despite small variations in diet compositions, these seal species remain in the system all year round (4)
Sub-Antarctic seals	*Mirounga leonina, Arctocephalus gazella,*	Despite small variations in diet compositions, these seal species migrate out of the system for a portion of the year (4)
Leopard seal	*Hydrurga leptonyx*	Has a unique diet compared to other seal species (5) and no consistent predators (7)
Baleen whales	*Eubalaena australis, Balaenoptera borealis, Balaenoptera physalus, Megaptera novaeangliae, Balaenoptera musculus*	Baleen whales (3) have extremely similar diet compositions (5) and all migrate out of the model domain for a portion of the year (4). The data available for these species is limited and therefore better combined into one group (2).
Minke whales	*Balaenoptera bonaerensis*	Remain in the model domain for the whole year (4) and have a unique diet compared to other whale species (5).
Sperm whale	*Physeter macrocephalus*	Have a unique diet compared to other whale species (5) and are important for the squid pathway being explored by the model (1).
Killer whale	*Orcinus orca*	Has a unique diet compared to other whale species (5) and no predators (7)

Figure 2. Two food web diagrams for Prydz Bay, Antarctica constructed from dietary data from the Southern Ocean Dietary Database (SODD). Colours correspond to general group classifications (e.g. red for zooplankton, light blue for whales) and edges are coloured according to prey species/group with arrows pointing towards the predator species/group. A) Disaggregated food web representing the lowest taxonomic resolution for species/groups and their interactions present in the database, numbers correspond to those in the Key B) Aggregated food web/functional group structure in Ecopath constructed from species/groups represented in (A). Edges represent the aggregated trophic linkages of those present in (A).

4. DISCUSSION AND CONCLUSIONS

Guidelines for best practices in ecosystem model development (e.g. Heymans et al. 2016) provide a useful basis for developing appropriate conceptual models of a system but can be challenging for applications in Antarctic environments as data quality and quantity is often insufficient. The Southern Ocean dietary database has yet to be utilised in Ecopath models for the Southern Ocean despite 10 Ecopath models already developed for various regional food webs. The number of functional groups varies considerably among these models, ranging from 23 (Pruvost et al., 2005) to 63 groups in an Ecopath model of the Western Antarctic Peninsula (Suprenand and Ainsworth, 2017). Here we argue that the addition of information from the Southern Ocean dietary database to the model formulation process provides a stronger basis for decisions regarding species aggregation in mass balance modelling approaches. These decisions remain influenced by the subjective nature of model development therefore we would recommend that alternative model structures be maintained and compared to generate various possible outcomes for the research question and evaluate structural uncertainty. Although the decisions presented here may not be final, the process of determining them is invaluable for exploring limitations in the available data as well as improving the modelers' understanding of the system.

We have only considered aggregation at the level of species classification although aggregation also exists in the underlying biological processes of each group, for example the different life stages of species/groups are rarely defined (larvae, juveniles, adult etc.). An important question therefore remains regarding how the representation of groups will affect representation of the underlying biological processes such as recruitment, growth and reproduction.

The decision framework presented is an important step in ensuring transparency in the development and documentation of ecosystem models. Despite the suitability of its application to food web models with functional groups it may not be appropriate for application to other ecosystem modelling approaches (e.g. size-based models or models of intermediate complexity – 'MICE' models). With the ever-growing demand of end-to-end models for understanding the structure and function of oceanic ecosystems (Murphy et al., 2012), the development of a general framework appropriate for application to these other modelling approaches remains a challenge.

ACKNOWLEDGEMENTS

REFERENCES

Christensen, V., and D. Pauly. (1992). ECOPATH II – a software for balancing steady-state ecosystem models and calculating network characteristics. *Ecological Modelling*, 61(3-4), 169-185.

Fulton, E.A. (2001). The effects of model structure and complexity on the behaviour and performance of marine ecosystem models. Ph.D. thesis, University of Tasmania, Hobart.

Fulton, E.A., Smith, A.D., and Johnson, C.R. (2003). Effect of complexity on marine ecosystem models. *Marine Ecology Progress Series*, 253, 1-16.

Gardner, R., Cale, W., and O'Neill, R. (1982). Robust analysis of aggregation error. *Ecology*, 63(6), 1771-1779.

Heymans, J.J., Coll, M., Link, J.S., Mackinson, S., Steenbeek, J., Walters, C., and Christensen, V. (2016). Best practice in Ecopath with Ecosim food-web models for ecosystem-based management. *Ecological Modelling*, 331, 173-184.

IPCC. (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Murphy, E.J., R.D. Cavanagh, E.E. Hofmann, S.L. Hill, A.J. Constable, D.P. Costa, M.H. Pinkerton, N.M. Johnston, P.N. Trathan, J.M. Kinck, D.A. Wolf-Gladrow, K.L. Daly, O. Maury, and S.C. Doney. (2012). Developing integrated models of Southern Ocean food webs: Including ecological complexity, accounting for uncertainty and the importance of scale. *Progress in Oceanography,* 102, 74-92.

Pinkerton, M.H., Bradford-Grieve, J.M., and Hanchet, S.M. (2010). A balanced model of the food web of the Ross Sea, Antarctica. *CCAMLR Science*, 17, 1-31.

Pinnegar, J.K., Blanchard, J.L., Mackinson, S., Scott, R.D., and Duplisea, D.E. (2005). Aggregation and removal of weak-links in food-web models: system stability and recovery from disturbance. *Ecological Modelling*, 184 (2-4), 229-248.

Polovina, J.J. (1984). Model of a coral reef ecosystem. *Coral Reefs*, 3, 1-11.

Pruvost, P., Duhamel, G., and Palomares, M.L.D. (2005). An ecosystem model of the Kerguelen Islands' EEZ. *Fisheries Centre Research Reports*, 13(7), 40-64.

Raymond, B., Marshall, M., Nevitt, G., Gillies, C.L., Van Den Hoff, J., Stark, J.S., Losekoot, M., Woehler, E.J., and Constable, A.J. (2011). A Southern Ocean dietary database. *Ecology*, 92(5), 1188-1188.

Suprenand, P.M., and Ainsworth, C.H. (2017). Trophodynamic effects of climate change-induced alterations to primary production along the western Antarctic Peninsula. *Marine Ecology Progress Series*, 569, 37-54.