Marine and Coastal

National Environmental Science Program

FINAL REPORT

Project 2.3

December 2024

Improving knowledge transfer to support Australian Marine Park decision making and management effectiveness evaluation

Keith R. Hayes, Peter Walsh, Neville Barrett, Carly Cook, Piers Dunstan, Emma Flukes, Scott Francis, Gretchen Grammer, Geoff Hosack, Steffan Howe, Daniel Ierodiaconou, Tim Langlois, Eric Lawrey, David Logan, Vanessa Lucieer, Jacquomo Monk, Joel Williams, Skipton Woolley

CSIRO
UNIVERSITY of TASMANIA
IMAS
Institute for Marine and Antarctic Studies
THE UNIVERSITY OF WESTERN AUSTRALIA
DEAKIN UNIVERSITY
MONASH University
Australian Government
AUSTRALIAN INSTITUTE OF MARINE SCIENCE
Government of South Australia
Department of Primary Industries and Regions
SARDI
SOUTH AUSTRALIAN RESEARCH AND DEVELOPMENT INSTITUTE
Australian Marine Parks

Milestone number: 11

Research Plan number: RP2022

Please address inquiries to: Keith Hayes; keith.hayes@csiro.au

Preferred citation

Hayes, K. R., Walsh, P., Barrett, N., Cook, C. N., Dunstan, P., Flukes, E., Francis, S., Grammer, G., Hosack, G. R., Howe, S., Ierodiaconou, D., Langlois, T., Lawrey, E., Logan, D., Lucieer, V., Monk, J., Williams, J. and Woolley, S. N. C. (2024). *Improving knowledge transfer to support Australian Marine Park decision making and management effectiveness evaluation.* Report to the National Environmental Science Program. CSIRO, Hobart, Australia.

Copyright

Acknowledgement

This work was undertaken for the Marine and Coastal Hub, a collaborative partnership supported through funding from the Australian Government's National Environmental Science Program (NESP).

NESP Marine and Coastal Hub partners

The Australian Institute of Marine Science, Bioplatforms Australia, Bureau of Meteorology, Charles Darwin University, Central Queensland University, CSIRO, Deakin University, Edith Cowan University, Flinders University, Geoscience Australia, Griffith University, Integrated Marine Observing System, James Cook University, Macquarie University, Murdoch University, Museums Victoria, NSW Department of Planning and Environment (Environment, Energy and Science Group), NSW Department of Primary Industries, South Australian Research and Development Institute, The University of Adelaide, University of Melbourne, The University of Queensland, University of New South Wales, University of Tasmania, University of Technology Sydney, The University of Sydney, University of Western Australia, The University of Wollongong

Disclaimer

This report is available on the NESP Marine and Coastal Hub website:
www.nespmarinecoastal.edu.au

Contents

List of figures

List of tables

Executive summary

Parks Australia requires data products that are “fit for purpose” in order to effectively manage Australia’s marine parks. The first objective of this project is to assist Parks Australia develop and implement a data product quality assurance framework by (i) identifying the use cases and data products it needs to answer key management questions; (ii) developing a suite of quality assessment criteria; (iii) assessing existing data products and their associated repositories against these criteria; and (iv) document recommendations that will help ensure data products become fit for their intended purpose.

A data product is a re-useable data asset designed to support one or more data use cases. A data use case is a description of a practical application of a data product. A use case typically describes how a user performs a function, for example how a user answers a key management question with a data product. Prior to this project Parks Australia identified six key questions (KQ’s) that define their most important data use cases.

1. What natural values are in the area of interest?
2. What state or condition are the natural values in?
3. How special are the natural values?
4. How might the natural values respond to a pressure, activity, or management action?
5. Where is the activity, incident, or pressure occurring in relation to where the values are?
6. What management actions can be taken to contribute to zone objectives and conservation goals for the natural values?

There are other management questions that park managers need to answer but these six questions were considered to be a priority and at an appropriate level of detail for the purposes of this project.

Prior to the project Parks Australia held two workshops to identify a list of data products, organised into a data product hierarchy, that the project might assess. By examining the workflows and decisions Parks Australia staff use when answering key questions 1 and 3, as well as 2 and 5, the project was able to identify 20 priority data products and associated repositories from among the 90 plus data products identified in these workshops.

The project was unable to describe the use cases for key questions 4 and 6 because Parks Australia has yet to establish workflows that identify the data products that they will use to answer these two questions. These questions therefore will require additional analysis that accounts for Parks Australia’s management approach.

In the workshops conducted prior to the project Parks staff also identified a range of quality concerns associated with data products. These concerns were categorised and organised into an initial data product quality assessment framework addressing 6 aspects of data quality: Requirements, Science, Production, Stewardship, Service and Use.

This framework provided the basis for a more detailed data product quality assessment undertaken in this project. For each quality aspect in the framework the project developed series of questions which were used to test if the product was “fit for purpose’ and elicit more detailed quality concerns.

The quality evaluation was performed using a questionnaire style survey addressing all six quality criteria. There was a total of 41 questions, 4 addressing the Requirements Quality criteria, 7 directed to the Science Quality criteria, 10 under Production Quality, 9 under Stewardship Quality, 7 addressing the Service Quality criteria, and 4 directed to the Use Quality criteria. These questions were answered by staff at Parks Australia and the project team. Answers provided for Production quality, Stewardship quality and Service quality were subsequently checked by conducting interviews with the data product’s owner, manager or representative.

Each of the data quality assessment questions were designed for three primary responses: “Yes”, “No”, or “Partially”. The overall quality of a data product was then based on how many of the quality assurance criteria a data product met – i.e. how many “Yes” answers, and how many almost/sometimes met – i.e. how many “Partially” answers.

The ranking of data products by counts of “Yes” and “Partially” response indicates that the more established data products with national or global governance – such as OBIS, the National Reef Monitoring Network and Seamap Australia’s National Benthic Habitat Layer – received the highest number of “Yes” responses. Newer data products that are still evolving and seeking ongoing support - such as Squidle+ deployments and the three GlobalArchive data products - had more “Partial” responses. The AMP Ecosystems data product had the highest number of “No” responses.

By implementing the quality assessment procedure on the 20 priority data products and associated repositories, the project was able to identify 111 recommendations that if implemented would help make these data product “fit for purpose”. These recommendations, together with additional recommended changes to the Natural Values Common Language, improvements to the way in which: (i) canyons and seamounts, (ii) shallow coral reef habitats, and (iii) oceanic vegetation in the Coral Sea are identified in the AMP ecosystems data product - and the use of higher resolution bathymetry and multinomial habitat distribution models to map (with associated level of uncertainty) habitat forming species within this data product – are colloquially termed as the “Roadmap” for improving knowledge transfer to support Australian Marine Park decision making and management effectiveness evaluation (Table 1).

In addition to individual data products and repositories, the project also examined the architecture and data product infrastructure that currently services Australia’s marine science and marine management community more generally. This analysis of the entire data supply chain, from requirements to end use cases, uncovered inconsistencies in the architecture used from minor issues to identifying data products and associated infrastructure that require significant work and support.

Table 1 Summary of the Roadmap issues identified during the course of NESP MaC Project 2.3

Issue	Summary
Status of the Natural Values Common Language	The Natural Values Common Language (NVCL) and its associated data products were designed to support a monitoring prioritisation process. Their status, and role within, the Authorisations, Incidence Response and State of Knowledge programmes has yet to be clarified or codified. At this juncture it is important for Parks Australia to confirm if it intends to proceed with its current classification of marine benthic habitats (via the NVCL ecosystems) or seek to adopt the more detailed (but data intensive) delineation of habitats provided by Seamap Australia via the National Benthic Habitat Layer.
Accuracy of the AMP ecosystems data product	The uncertainty associated with the ecosystems identified in the AMP ecosystems layer is unspecified. The NVCL implies that if an AMP ecosystem is shown as present within an area of interest, then all associated ecosystem components and sub-components are also present. A more rigorous approach is to use habitat distribution models to define the probability that habitat forming ecosystem components (such as seagrass, macroalgae and sessile invertebrates) are present with a specified level of certainty. The full use case for the AMP ecosystems needs to be articulated by Parks Australia to ensure that the utility of these models, and various representations of uncertainty that can be provided with them, will be adequate for future decision making.
Updating of the AMP ecosystems data product	A formal process to maintain and update the AMP ecosystems (including the addition of new ecosystems such as oceanic vegetation) need to be developed. This should include the process for inclusion of new data sources and the updating of any existing data sources. This process should be managed by Parks Australia.
Fit for purpose data products	The project identified 20 data products and repositories that Parks Australia rely on when addressing four of its six key questions. The quality assurance assessment performed in this project identified over 100 recommendations in order to make these products and repositories fit for Parks Australia's purpose. As a general rule the older, more established and well supported data products and repositories attracted fewer recommendations than the newer ones.
Data accessibility	During the course of the project a number of potentially useful, but not publicly accessible, data products were identified. Private data products (such as these examples) can be categorised into four types: (i) data from science activities that were commissioned and funded by Parks Australia, DCCEEW, or NESP/NERP and conducted under NESP/NERP; (ii) data from science activities that were commissioned and funded by Parks Australia, DCCEEW or other third parties, but not conducted under NESP/NERP; (iii) data collected in AMPs by industry, commercial consultants or research institutions following authorisations issued by Parks Australia; and (iv) data collected under other regulatory regimes (such as fisheries management) that is relevant to AMPs. Parks Australia should consider strategies that would help make these types of data sets publicly available into the future.

This broader review also identified eight very recent or currently in progress data infrastructure initiatives whose objectives overlap with those of Parks Australia, including the initiatives that are currently underway to reform Australia's national environmental information landscape in order to implement the new National Environmental Standards recommended following the recent review of the Environmental Protection and Biodiversity Act. This overlap suggests that there is an opportunity for future communication, collaboration and alignment of outcomes across marine and terrestrial environmental data systems.

The project's second major objective was to demonstrate an empirical and/or expert-based methodology for assessing the condition of marine park values, perhaps drawing upon the approaches, experience and methods developed in Victorian marine parks. The project provided demonstrations to Parks Australia of an expert facilitated approach to condition assessment that precipitated a series of internal discussions within Parks Australia from which it became clear that, despite the considerable progress that has been made to date, Parks Australia needs to resolve a number of policy questions to help inform ongoing development and implementation of a monitoring and management effectiveness system.

To assist in these discussions the project has provided a detailed outline of the steps involved in a condition assessment process, together with a summary of the important methodological and policy choices that each step entails, and a set of potential indicators for benthic natural values, that were identified during the course of the project.

1. Introduction

1.1 Project background

Data product hierarchy

Parks Australia, supporting the Director of National Parks, manages 60 Australian Marine Parks (AMPs) located within Commonwealth waters – those over 5.5 kms from the coast. These parks are vast, covering approximately 3.8 million km².

A key challenge associated with managing marine parks at this scale is the coordination of resources across the marine science community to specify, collect, manage, and collate the data needed to support evidence-based management of the AMP estate (Hayes *et al*., 2019). To meet the needs of Parks Australia and its management, partners' data must be synthesised into data products that are "fit-for-purpose".

Parks Australia recognises that for many data products achieving "fit-for-purpose" status will require a data product supply chain that involves adoption of standards and best practices by a range of organisations over time. These organisations and the shared infrastructure they rely on must work collaboratively, strategically, and respectfully to meet Australia's marine management challenges. The purpose of this project is to assist this process by *inter alia* helping to develop and implement a data product quality assessment framework.

Prior to the project Parks Australia held two workshops to identify a list of data products for the project to assess. The data products identified were considered important for answering six key questions that are common across three areas of Parks' management responsibilities namely: (i) authorisations advice, (ii) environmental incident response advice; and (iii) State of Knowledge reporting.

As part of the project preparations, Parks Australia also described a Data Product Hierarchy (DPH). The DPH attempts to organise Park's Australia's environmental information sources into several levels. The lower levels, starting at Level 0, identify the least processed inputs to data products. Data products become increasingly processed as they move up the DPH, culminating in the directly used, most value-added, product described at Level 5 (Figure 1).

Where data is of low quality or is otherwise unavailable, Parks Australia must also rely on publications and conversations with researchers and Traditional Owners to help it answer the six key questions (identified in DPH Levels -1, and -2). The placement of these information sources at these levels signifies their qualitatively different nature to data products rather than any implied notion of quality.

Whilst developing the DPH, Parks Australia conducted a review of current environmental information sources that it might access. This review identified 74 information sources, split across the 6 levels of the DPH, together with a range of supplementary information sources at Level 2, Level -1 and Level -2 (Appendix A).

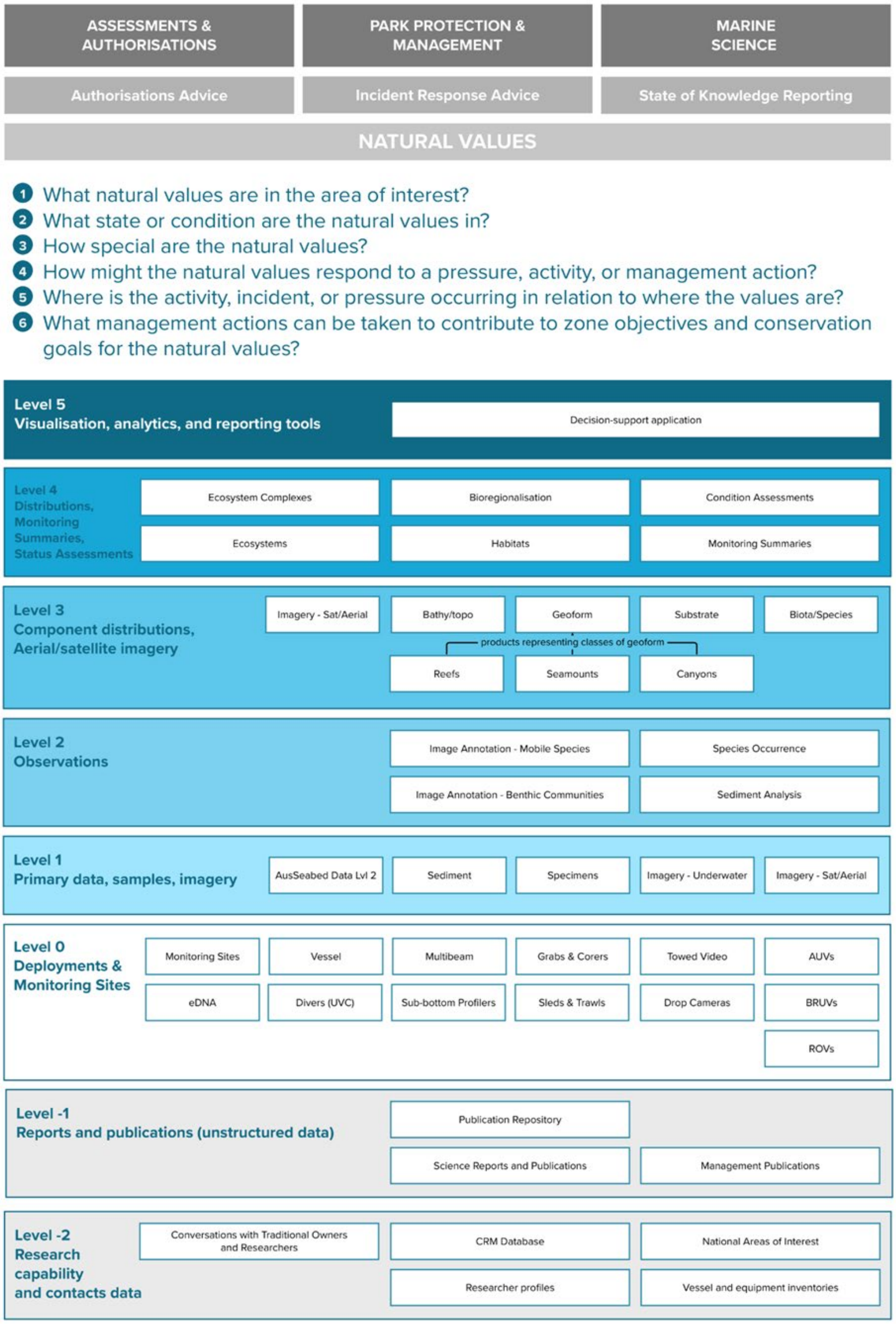


Figure 1 The Data Product Hierarchy organises the data products used by Parks Australia to answer six key management questions that are common to its three management programs (top of figure), from the least processed inputs to data products (Level 0) up to higher level, most processed, value-added, directly used products (Level 5). Parks also relies on publications and conversations with researchers and Traditional Owners to help it answer the six questions (Levels -1, and -2) where data is of low quality or is otherwise unavailable. A non-exhaustive list of inputs and data products is shown at each level.

SAFE 2.0 Framework

In May 2023, midway through this project, the SAFE 2.0 capability matrix was published (Figure 2). The SAFE 2.0 matrix describes key capabilities required at six stages of a data product lifecycle – collect, curate, integrate, analyse and use, supported by underlying quality culture – to achieve a new national supply chain of environmental information.

The SAFE matrix representation enables the elements of the environmental data and analytics landscape to be seen in relation to each other. It aims to be generally applicable regardless of the specific data types or institutions to which it may be applied, and was developed in response to the review of the Environmental Protection and Biodiversity Conservation (EPBC) Act to depict the key capabilities that underpin the recommended reforms to the current national environmental information landscape. These reforms are necessary to inform the implementation, evaluation, reporting and assurance of National Environmental Standards and other aspects of the EPBC Act reform, including State of the Environment reporting (Figure 3).

The data products and supply chains that Parks Australia use suffer from many of the issues that characterise the current national environmental information landscape. It is not surprising therefore that the data product quality assurance criteria identified by Parks Australia for this project are consistent with the, and address many of the same, key capabilities identified in the SAFE 2.0 framework.

The data product quality assessment framework developed by Parks Australia, with the assistance of this project, was also designed so that the quality assessment criteria could be logically grouped into the stages of a generalised data product lifecycle (Appendix C). The centrality of data products (data designed around use cases) and the inclusion of the Requirements and Use stages emphasise the active role of consumers (users) in the pathway to impact for data products from the marine domain.

The framework is circular with a directional arrow to emphasise that lessons learned during a data product's evaluation, ending in the evaluation of "Use" should eventually lead to subsequent evaluations, beginning with "Requirements". The components of the data product quality framework can also be readily aligned with the SAFE 2.0 data product lifecycle stages, and thereby emphasise the similarity with the SAFE 2.0's envisaged future supply chain for environmental data (Appendix C).

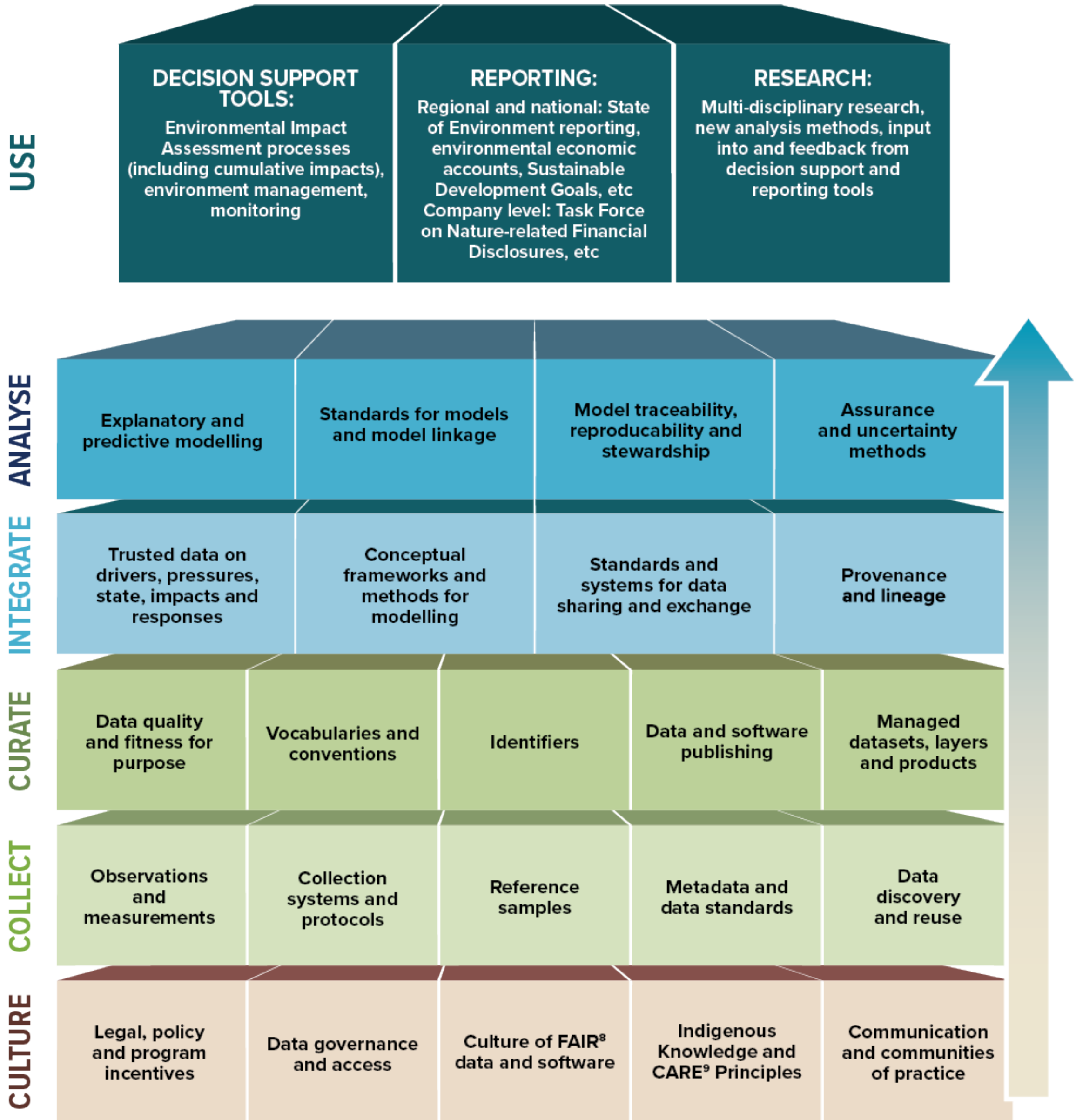
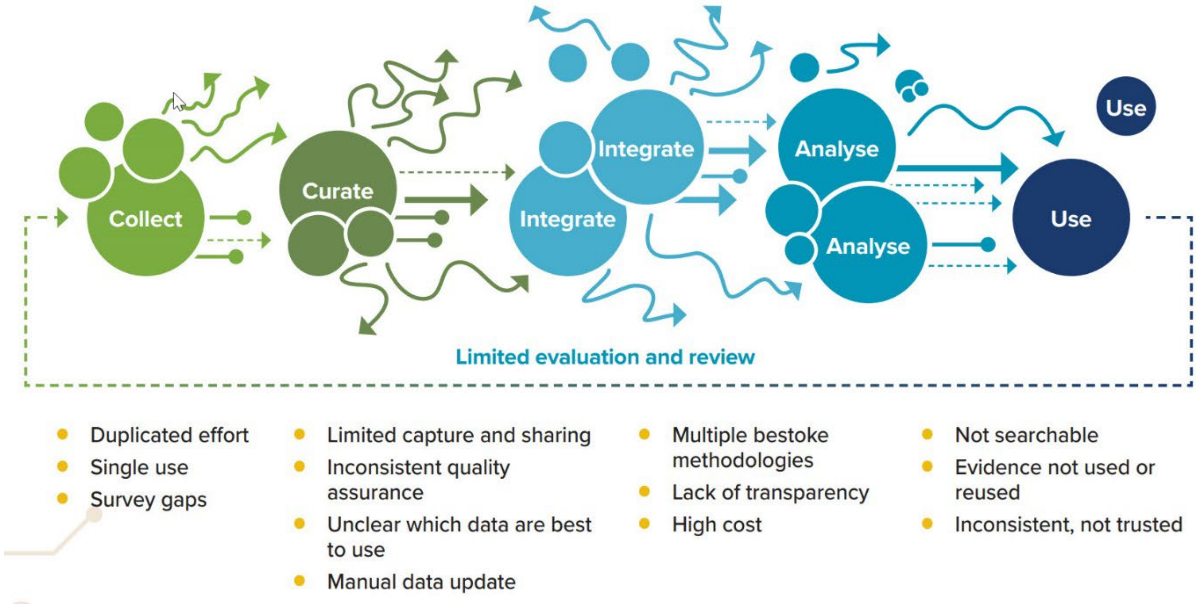


Figure 2 The SAFE 2.0 capability matrix identifies a set of key capabilities, grouped under six stages of a data product life cycle – collect, curate, integrate, analyse and use - supported by an underlying quality culture. These capabilities underpin the recommended reforms to the current national environmental information landscape, that were identified as part of the recent review of the Environmental Protection and Biodiversity Conservation Act (Source: WABSI and WAMSI, 2023).



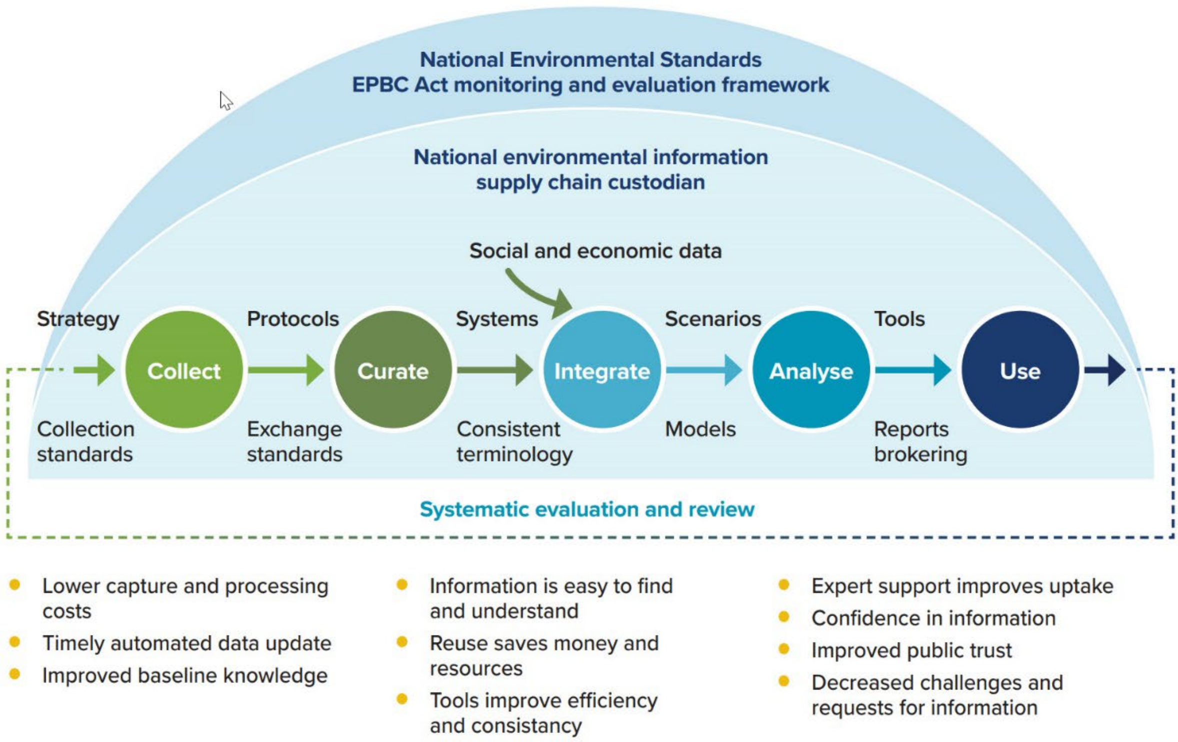


Figure 3 Characteristics of the current (top) national environmental information supply chain, versus an envisaged (bottom) future supply chain of national environmental data (Source: https://epbcactreview.environment.gov.au/resources/final-report/chapter-10-data-information-and-systems/103-recommended-reforms).

1.2 Project objectives

The project was originally conceived around two broad objectives. The first was to identify the data products that Parks Australia needs to protect and evaluate the condition of marine park values, and assess the availability of, and access to, these data products and whether they are currently fit for purpose. The second was to describe empirical methods for estimating the condition of AMP values and the effects of management, and if suitable, fit for purpose data products were available, demonstrate the use of these methods.

Availability and access to fit for purpose data products

The project's examination of the availability and access to fit for purpose data products was structured around five work items:

- describing and understanding the data use cases that Parks Australia currently employ to support knowledge transfer and decision making.
- identifying the data products needed to support these uses cases and examining the software and hardware infrastructure used by the data repositories that currently store and manage these data products.
- developing a suite of quality assessment criteria that meet Parks Australia's requirements and are consistent with other equivalent national or international criteria.
- assessing the data products and associated repositories against these criteria in order to determine the extent to which they meet Parks Australia's requirements for rapid and efficient knowledge transfer and decision support.
- document recommendations for advancing delivery of data and data products to meet Parks Australia's requirements.

Figure 4 provides a schematic summary of the project's first objective. The figure presents a stylised example of the workflow used to create an important data product – in this instance the AMP ecosystems product (a Level 4 product; Section 0) – showing the data products at the lower levels of the DPH that are used to create it, the data repositories from which Parks Australia may access each product, and examples of the issues addressed within the quality assurance process to ensure each product is fit for purpose.

At the outset the project recognised that very few data products or repositories would likely meet all of Parks Australia's quality assurance criteria because some criteria would require a high level of maturity, some are not applicable to particular products, and many products were not originally specified to meet these criteria. Importantly, by formally enunciating these criteria, and systematically assessing data products and repositories against them, the project aims to identify shortcomings - from Parks Australia's perspective - in the current marine data environment and thereby help identify mechanisms to improve this environment into the future.

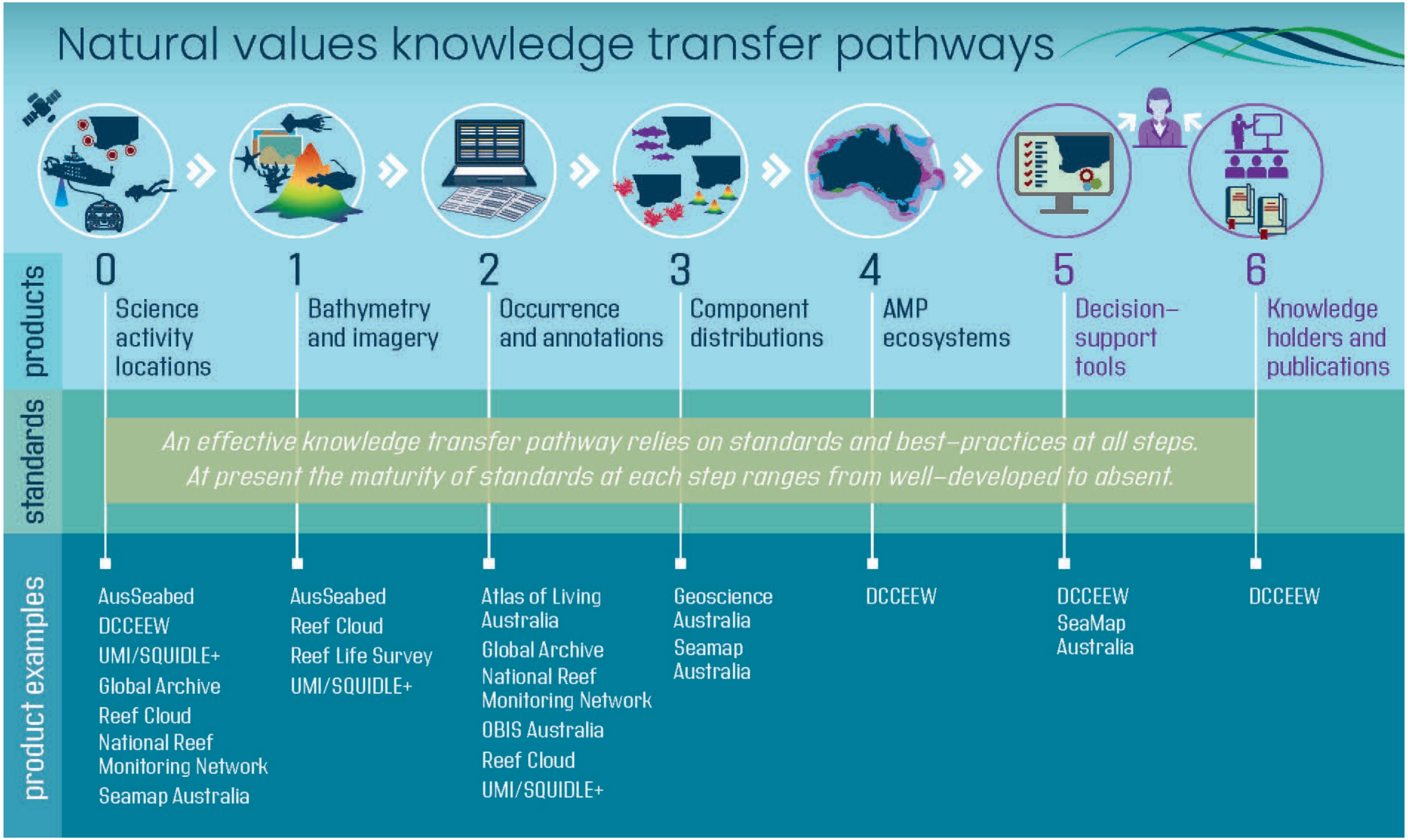


Figure 4 A stylised example of data products in a potential future supply chain for updating the AMP ecosystems data product.

Estimating condition and management effectiveness

At its inception the project envisaged that some existing data products might provide an opportunity to demonstrate an empirical (data driven) assessment of the condition of marine park values, perhaps drawing upon the approaches, experience and methods developed in State marine parks, such as the control charts approach implemented by Parks Victoria (Ierodiaconou *et al*., 2020). This aspiration was reflected in Project milestone 8.

Moreover, where existing approaches for establishing condition categories linked to management effectiveness such as control charts were not applicable because, for example, historical estimates of condition prior to management were not available, the project team agreed to work with Parks Australia to identify other appropriate methodologies for measuring management effectiveness into the future, including through the use of expert elicitation. This aspiration was reflected in Project milestone 9.

The project provided demonstrations to Parks Australia of an expert facilitated approach to condition assessment, however, it became apparent that the project would be unable to recommend a method for assigning condition categories to natural values within the park because internal discussions are still occurring with Parks Australia regarding options and preferred approaches to condition assessment for natural values within AMPs.

The project in effect precipitated a series of internal discussions within Parks Australia about important issues and options through which it became clear that despite the considerable progress that has been made to date, there are further design elements to be developed and decisions required in relation to Parks Australia's monitoring and management effectiveness system.

In light of these developments, the project amended its objectives at the request of Parks Australia to focus on identifying the critical steps and provide methodological options (including potential indices and indicators) to be considered when developing and applying a process for assessing the condition of natural values and evaluating the effectiveness of management interventions.

Roadmap

The fifth component of the project's first objective - document recommendations for advancing the delivery of fit for purpose data products – is viewed by Parks Australia to be particularly important. Over the course of the project many data product problems and associated recommendations were identified. These recommendations became colloquially known as the "roadmap" and are summarised within a dedicated section of this report under the same name. The project anticipates that some of these recommendations will be taken up by future NESP projects, particularly Project 4.20 and Project 4.21.

1.3 Report structure

The report is structured into four main sections. Section 2 examines the data products and information sources from the perspective of Parks Australia staff who are trying to answer six key questions and identifies a suite of critical issues that influence the decisions they make about the utility – i.e. fit for purpose – of the data or information sources.

Section 3 describes the quality assurance process developed by the project team to assess the extent to which the data products identified in Section 2 are, or are not, “fit for purpose”, while Section 4 describes the infrastructure and data product supply chain that Parks Australia and data providers should aspire to, and ultimately achieve, in order to meet management obligations.

Section 5 describes a number of generic national monitoring questions that align with one or more of the six key questions described in Section 2, identifies several types of monitoring and monitoring locations, and outlines the steps, choices and management options available to Parks Australia when assessing, and reporting on, the condition of natural values within the marine park estate, and the effect of management on these values.

Section 6 concludes the report with a roadmap – that is a list of issues identified by the project, together with a set of associated recommendations to improve the delivery of fit for purpose data products.

The main body of the report has been deliberately designed to be as concise as possible and therefore includes extensive Appendices that provide additional detailed information that some readers will find helpful.

2. Understanding data products and use cases

A data product is a re-useable data asset designed to support one or more data use cases. A data use case is a description of a practical application of a data product. A use case typically describes how a user performs a function, for example how a user answers a key management question with a data product. A use case can be about direct use, for example directly used to answer a key question, or an indirect use, for example the development of a (lower level) data product used in the creation or development of a (higher level) data product such as the AMP ecosystems (Figure 4) that is used to answer a key question.

Data product examples include geospatial layers, such as benthic habitat distribution, decision-support tools used to enable use of geospatial layers, and standardised data collections used to stage data for future updates to other products (Figure 1). In this project a data product is distinguished from unprocessed raw data and one-off data outputs. Data products are typically stored and/or viewable within a data product repository, defined here as a public or internal (to Parks Australia) database or portal.

Prior to this project Parks Australia identified six key questions (KQ's) that define their most important data use cases, specifically:

7. What natural values are in the area of interest?
8. What state or condition are the natural values in?
9. How special are the natural values?
10. How might the natural values respond to a pressure, activity, or management action?
11. Where is the activity, incident, or pressure occurring in relation to where the values are?
12. What management actions can be taken to contribute to zone objectives and conservation goals for the natural values?

These data use cases help identify the data products required by Parks Australia and guide the assessment of their fitness for purpose. There are other management questions that park managers need to answer but these six were considered to be a priority and at an appropriate level of detail. A finer scale analysis is possible but more feasibly conducted by individual data product producers and their specific end users within Parks.

The project focussed on data products to support the key questions in relation to benthic natural values. Characterisation of benthic natural values is an immediate priority for Parks, but the project recognises that additional analysis will be required into the future to address other data needs, for example monitoring and condition questions, pelagic values, Threatened Endangered and Protected Species (TEPs), and pressures.

2.1 Key questions

KQ1 and KQ3

Key question one (KQ1) – what natural values are in the area of interest - seeks to determine the type, location and extent of natural values in a park, network or area of interest. Key question three (KQ3) – how special the natural values are - seeks to identify natural values that are of a higher than usual importance for a variety of reasons. These questions are grouped together because when addressing KQ1 Parks Australia workflows often address KQ3 at the same time.

The area of interest in this context would typically be the park, a zone or ecosystem within the park or a particular feature – such as an island, reef, seamount, canyon or key natural value – within the park, that may traverse zone or park boundaries. The area of interest may also relate to other types of off-park areas. For incidence response and authorisations, for example, the area may be delineated by the (potentially changing) boundaries of an incident, such as the limits of an oil spill, or predicted footprint of a proposed development, such as a new offshore renewable energy facility, particularly where an incident or proposed activity outside the park boundaries is deemed to potentially impact the values within the park.

When addressing KQ1 and KQ3, Parks Australia typically refer to data products that are accessible through Seamap Australia, the AMP science atlas, an internal AMP features GIS application and the internal State of Knowledge report (Figure 5). The State of Knowledge report compiles information from data products held in a large number of repositories, including the AMP features application, Seamap Australia, Species Profiles and Threat (SPRAT) database, the National Introduced Marine Pest Information System (NIMPIS), eAtlas as well as a range of other sources including expert opinion (Figure 6).

The AMP Features application enables Parks Australia staff to identify what species are present in an area of interest by using the Microsoft Power BI application to visualise data products viewable in GlobalArchive, Squidle plus, the National Reef Monitoring network and the Ocean Biodiversity Information System. To check for completeness and currency of the data products, Parks staff may also consult the AusSeabed marine data portal and the Marine National Facility voyage catalogue to identify if additional data from any recent biological surveys may be available (Figure 6).

Many of the natural values that are likely to be priorities for Parks Australia are associated with hard bottom, hence Parks staff are often particularly interested in the presence/absence, extent and type of reef features both inside and outside AMP boundaries. Hard substrata and associated communities are best described for shallow waters (within state boundaries), and are viewable through Seamap National Benthic Habitat layer and Reef cloud. Information on hard substrata in deeper waters (outside of state boundaries) is not as readily available, so Parks staff will sometimes consult bathymetry, geomorphology and backscatter data products in an attempt to infer the presence of reef-like structures on the seabed (Figure 5).

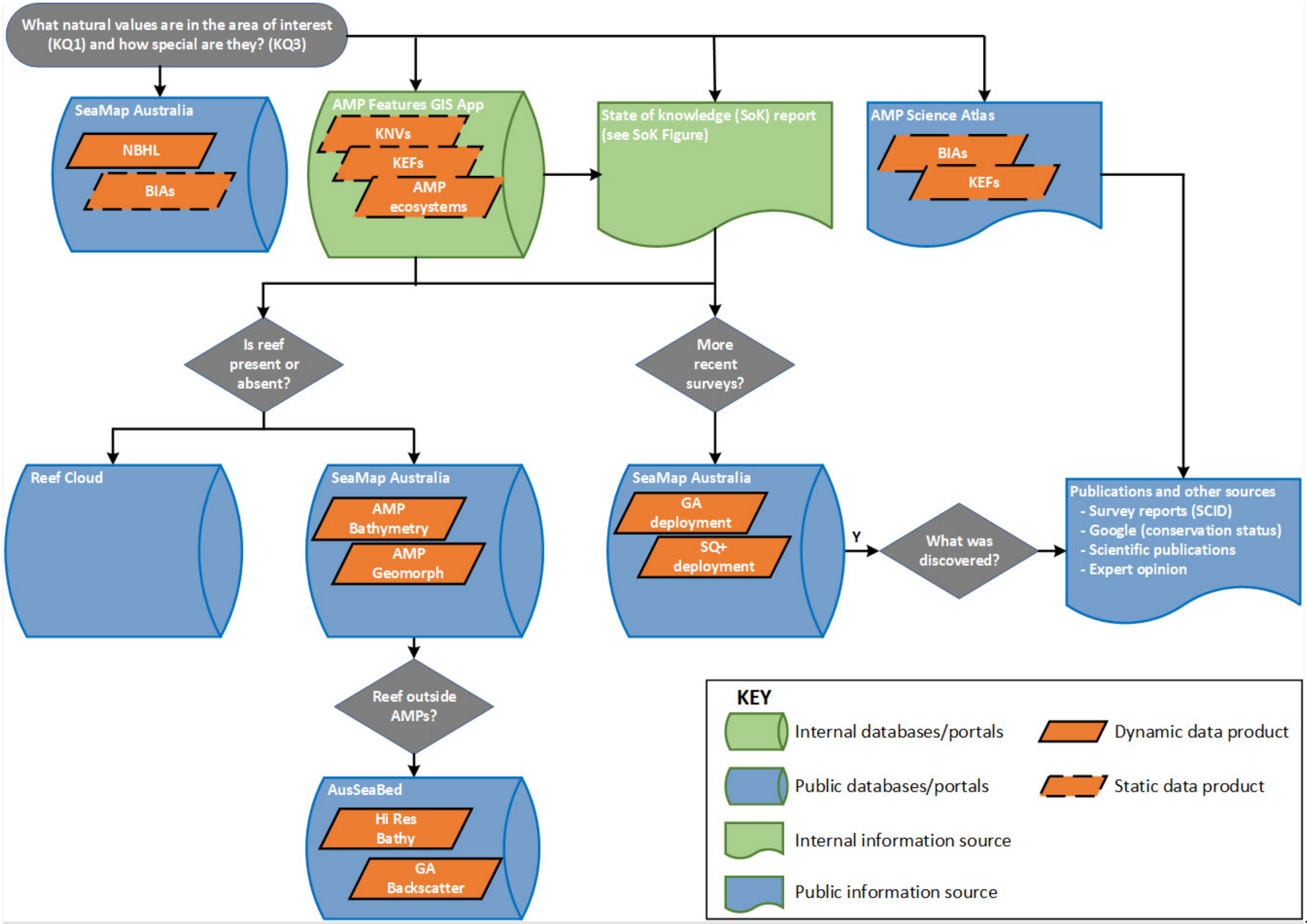


Figure 5 Data products and decision processes used by Parks Australia when answering key questions 1 and 3. The figure shows the regularly (dynamic) and rarely (static) updated data products used by Parks Australia when attempting to answering these key questions,

together with the public and internal (Parks Australia) data portals, databases or web-based applications from which these products are currently accessed.

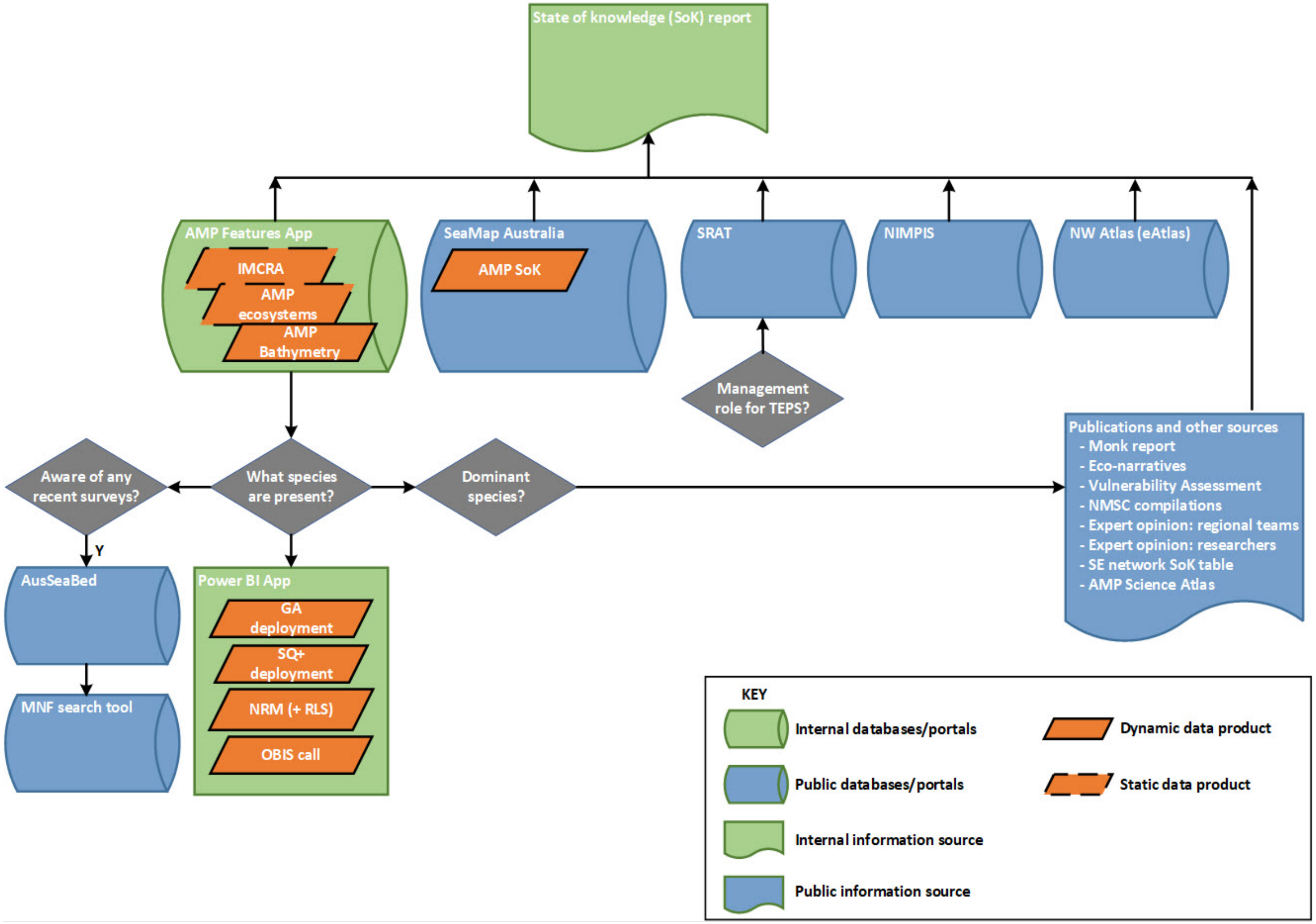


Figure 6 Data products and decision processes used by Parks Australia to create a “State of Knowledge” (SoK) report for an Australian Marine Park. The SoK reports are used by parks’ managers when addressing Key Questions 1 and 3 (refer to Figure 5).

KQ2 and KQ5

Key question two (KQ2) – what state or condition are the natural values in - seeks to determine the ecological condition of natural values. Key question five (KQ5) – where is the activity incident or pressure occurring in relation to the values - seeks to determine if incidents and pressures are strong enough and close enough to natural values to threaten their condition. Again, these questions are grouped together because when addressing KQ2 Parks staff will also often address KQ5.

The resources available to Parks Australia to help answer these questions are currently limited. For natural values associated with shallow reef ecosystem Parks staff can consult the Reef cloud interactive dashboard for summary information on current (and historical time series) macro-algae and coral cover at some sites around Australia, together with information on thermal stress and tropical cyclones. Similar summary statistics, but for a broader set of natural value condition indicators including large reef fish and habitat cover, is available on Reef Life Explorer with annotation data sourced from Squdle+/UMI (Figure 7).

For other natural values not associated with shallow reef ecosystems Parks staff may try to assess current condition by checking if a recent survey have been completed in the area of interest by reviewing for example the location of image collections on Seamap Australia (sourced from Squdle+/UMI and GlobalArchive), and if so consulting the grey literature (such as survey reports), the scientific literature and domain experts who they know were involved with the survey. These searches may be facilitated by the internal Scientific Information Database (SCID) (Figure 7) but this process could be time consuming, and any information retrieved may require additional interpretation, because Parks Australia has yet to finalise their condition assessment process for natural and cultural values (Section 5.5).

If Parks staff consider that the information retrieved by these checks and searches are inadequate for their particular query, they will likely initiate further investigations. They may, for example, infer possible impacts on natural values by using Seamap Australia to examine the benthic cumulative impact map and/or standardised pressure layers used to create it, and may again use bathymetry, geomorphology or backscatter data products to infer the presence of natural values associated with reefs, and hence the potential for overlap with, and consequently impacts arising from, anthropogenic pressures.

Parks staff may also retrieve information on potential impacts arising with offshore oil and gas infrastructure, or offshore renewable energy infrastructure, from the National Electronic Approvals Tracking System (NEATS). They may also refer to the Australian Maritime Information System (AMSIS) for information on the location of a variety of natural values and pressures that are by and large available in other data repositories. The substantial overlap in information provided by AMSIS and other data repositories, such as Seamap Australia, could be confusing to users who are unsure of where to access the most up-to-date information.

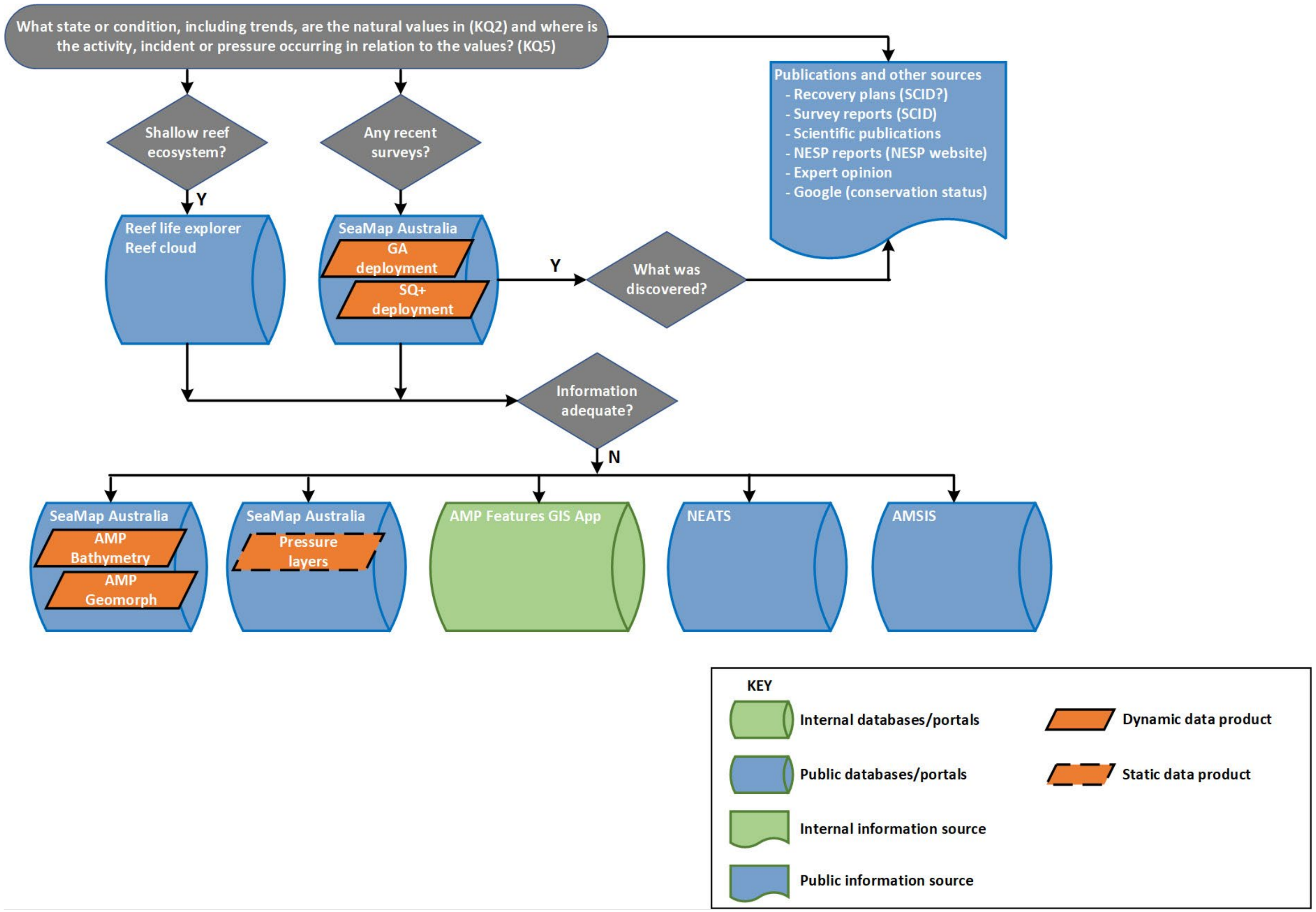


Figure 7 Data products and decision processes used by Parks Australia when answering key questions 2 and 5. The figure shows the regularly (dynamic) and rarely (static) updated data products used by Parks Australia when attempting to answering these key questions, together with the public and internal (Parks Australia) data portals, databases or web-based applications from which these products are currently accessed.

KQ4 and KQ6

Among the six KQs, key question four (KQ4) - how might the natural values respond to a pressure, activity, or management action - and key question six (KQ6) - what management actions can be taken to that contribute to zone objectives and anticipated effects of management (framed as monitoring questions) related to the natural values – are the least well developed in terms of how Parks staff approach the data use cases that are associated with them. Importantly, Parks Australia has yet to establish workflows that identify the data products that they will use to answer these two questions. These questions therefore will require additional analysis that accounts for Parks Australia's management approach.

2.2 Data products and repositories

The data products and repositories identified during the analysis of Parks Australia's data use cases and associated workflows (Section 2.1) are briefly summarised below.

Ocean Biodiversity Information System species records

The species records within Ocean Biodiversity Information System (OBIS) repository provide a global collection of occurrence data (with limited absence or sample effort data) that supports a broad range of Darwin Core fields. Parks requires all species observed during in AMP field studies to be published to OBIS. Records in OBIS can be retrieved by searching pre-assigned categories such as scientific name, institute and area. Data retrieval can be further delimited by depth, time and other specific characteristics such as Red List status. OBIS makes all species records conform to the World Register of Marine Species (WoRMS) which, however, is inconsistent with Codes for Australian Aquatic Biota (CAAB) for many endemic and range limited Australian fish and invertebrate species.

National Reef Monitoring Network fish and habitats

The National Reef Monitoring Network (NRMN) collates records from shallow reef surveys conducted around the world into a centralised database, including records from the Reef Life Survey and the Australian Temperate Reef Collaboration. Data products within the NRMN are organised into four streams covering the abundance and diversity of fish, invertebrates, cryptic fish and benthic coverage. This project distinguishes and analyses two types of data products stored in the NRMN repository: (i) the fish and invertebrate records; and (ii) the habitat records.

Validated and unvalidated National Benthic Habitat Layer

The Seamap Australia National Benthic Habitat Layer (NBHL) is a national-scale benthic habitat data product that amalgamates disparate benthic habitat data sets collected using different methodologies and classification schemes. These data sets are uniformly classified according to a national classification scheme which is implemented via a controlled vocabulary. The scheme is hierarchical in structure, enabling cross-compatibility with the common language used in the AMP Ecosystems data product (described below).

To be included in the NBHL, source habitat datasets must adhere to a set of Data Acceptance Guidelines. A key requirement for data acceptance is use of community-endorsed methodologies for ex-situ data collection, as outlined for example in the NESP

Field Sampling Manuals. Additionally, source datasets must have documented evidence of ground-truthing validation (e.g. via camera drops).

In addition to the formal Seamap Australia NBHL, an 'unvalidated' version also exists. This variant encompasses source habitat datasets that do not meet the validation requirements of the formal NBHL but have been synthesised and classified under the Seamap Australia scheme using identical methodology. This unvalidated dataset serves as a useful aggregation of benthic habitat information of potentially lower quality or certainty, but that may still be suitable for broader-scale modelling activities, such as development of the AMP Ecosystems product. The unvalidated dataset also highlights areas in which preliminary habitat mapping has occurred, that would benefit from field validation.

Bathymetry in Australian Marine Parks and high resolution bathymetry

The Bathymetry of Australian Marine Parks data product (visualisation here) is a multi-resolution compilation of the best available bathymetry in AMPs, created by Seamap Australia for the purposes of geomorphometry mapping (see Geomorphometry of Australian Marine Parks). It draws from the high-resolution surveyed bathymetry data managed by AusSeabed, who systematically publish Australian bathymetry data, typically on a per survey basis, but also as a compiled product such as the 50m Multibeam Dataset of Australia 2018 compilation (visualisation here). These bathymetry products are essential inputs into geomorphic and habitat distribution models.

Biologically Important Areas

Biologically Important Areas are areas where aggregations of individuals of a species are known to display biologically important behaviour such as breeding, foraging, resting or migration. The data product was used to assist with decision making in relation to meeting Principle 4 of the National Representative System of Marine Protected Areas (NRSMPA) Goals and Principles. BIA's are viewable in the Seamap Australia repository as a series of eight layers and are reportable in Parks Australia's internal AMP Features database.

Geoscience Australia backscatter

Geoscience Australia holds a large national collection of bathymetry datasets. The AusSeabed data access portal provides access to bathymetry data held and published by Geoscience Australia, and by CSIRO through the Marine National Facility (MNF). In most instances, bathymetry data collected by AusSeabed contributors will be acquired alongside backscatter data as per the NESP Marine Sampling Field Manuals for Monitoring Australia's Marine Waters (Chapter 3: AusSeabed Australian Multibeam Guidelines).

Seabed backscatter can provide an important measure of seabed hardness which can be used to differentiate between types of seafloor, such as hard rock or soft sediment. Backscatter data can also be used as a proxy to understand the characteristics of the seafloor, including surficial sediment characteristics, grain-size, and seafloor roughness. These measurements can be important inputs to seabed habitat mapping and can be used to explain spatial patterns of biology and biodiversity at continuous scales. Unlike bathymetry products, however, seabed backscatter is difficult to standardise between surveys and across depth ranges, and hence its use is typically restricted to relatively smaller spatial extents.

Seamap State of Knowledge Reports - Australian Marine Parks

Parks Australia has compiled State of Knowledge reports for the South-east Marine Parks Network and is considering similar products for other networks and the Coral Sea marine park. These products periodically document the current level of understanding and information available about the marine environments, ecosystems, species, and human activities in AMPs. These reports include a high-level summary of relevant scientific research, monitoring, and data collected in AMPs to support their management, conservation and sustainable use.

Parks Australia uses a variety of data products to compile State of Knowledge reports including Seamap Australia's State of Knowledge reporting functionality, relevant publications, and expert advice. Regularly assessing and updating the state of knowledge helps to identify knowledge gaps and research priorities to ensure that management strategies are evidence-based and effective.

Australian Marine Parks Geomorphology

The Geomorphometry of Australian Marine Parks data product (visualisation here) provides the geomorphic features of the AMP estate The data was generated by Seamap Australia using the Bathymetry of Australian Marine Parks multi-resolution bathymetry composites. All publicly available bathymetry data as of 30th June 2022 was included, with a specific ad-hoc update conducted for Macquarie Island in August 2023 to reflect the new Park boundaries. Morphological features were calculated and classified for each AMP using the Dove *et*. *al*. (2020) Seabed Morphology Features Glossary. Note Nanson et al. (2022) defines geomorphology as the study of the shape of the Earth's surface and the processes forming them, and notes that this may include, but is distinct from, geomorphometry, defined as the science of quantitative terrain characterization, which encompasses acquisition and processing of topographic data as well as analyses and applications related to geomorphology.

Squidle+ deployments and annotations

Squidle plus is a web-based framework that facilitates the exploration, management and annotation of marine imagery within a specified Terms of Reference (http://localhost:5000/wiki/documents/terms_of_reference) and Strategic Plan (http://localhost:5000/wiki/documents/strategic_plan). Squidle+ collates information on the deployment of marine image gathering platforms (such as AUVs). These deployments and annotations form the data product for the purposes of this project. Squidle+ does not store the imagery collected during these deployments. It encourages sharing and re-use of the imagery by providing a standard online interface hosted in a standard way by various other institutions.

Key Ecological Features

Key Ecological Features (KEFs) are elements of the Commonwealth marine environment that are considered to be of regional importance for either a region's biodiversity or its ecosystem function and integrity. KEFs were identified to assist with decision making in relation to meeting Goal 3 of the NRSMPA Goals and Principles, and are viewable in Seamap Australia under the Key Ecological Features layer.

AMP ecosystems

The AMP ecosystems data product was developed by the NESP SS2/D7 project (Hayes *et al*., 2021) and NESP Project 1.3 (Dunstan et al., 2023) to help articulate what ecosystems, habitats, communities and species occur within Commonwealth waters, and support Parks Australia with the design and implementation of a marine park management framework. The data product is viewable in Seamap Australia under the Natural Values Ecosystems (2022) layer.

GlobalArchive deployments, fish and habitats

GlobalArchive is an online centralised repository of fish and benthic image annotations from stereo-video, including body-size and height information, stereo calibration and associated information. This project distinguishes and analyses three types of data products stored in GlobalArchive: (i) the deployments; (ii) fish records; and (iii) habitat records. GlobalArchive supports schema validation against CAAB and WORMS for fish records but only supports CATAMI annotations for benthic assemblages. GlobalArchive provides the ability to upload data and serve out syntheses from a schema controlled database via Application Programming Interface (API) calls, for example fish abundance and body size in Geographe Marine Park and benthic habitat annotations currently being used in NESP Project 2.1.

National Pressures standardised sum

The national pressures standardised sum data product refers to the 36 categories of anthropogenic pressures identified in the pressures common language, and the standardised (and sometimes summed) data products that were collated under each of these categories in the NESP SS2/D7 project (Hayes et al., 2021). The data product can be viewed in the Seamap Australia repository as a series of data layers, such as demersal trawl fishing effort (under the “commercial fishing” pressure category) in the Pressures and Activities catalogue.

Key Natural Values

The Key Natural Values (KNVs) data product represents the distribution of important species, habitats and other natural features selected using criteria similar to those used to identify Ecologically or Biologically Significant Marine Areas, which also overlap with the criteria used to identify KEFs and BIAs. The KNV data product was developed in NESP SS2/D7 project to help prioritise monitoring locations across the AMP estate (Hayes et al., 2021; Dunstan et al., 2023), but it may also have other uses. This data product is currently under development and is not currently hosted in any public facing data repositories.

Reef cloud

Reef cloud is an open-source data product that allows users to manage their imagery in the field and upload it to the cloud. Using an automated pipeline and AI methods, imagery is automatically annotated with 80-90% confidence. Imagery can also be annotated manually by the user. Coupled with an interactive data dashboard that is automatically updated with ‘new’ data, this allow the user to quickly visualise trends at reef to regional scales. The platform is managed by the Australian Institute for Marine Science (AIMS). The platform is open all researchers collecting imagery for coral reefs.

3. Data Product Quality Assessment

3.1 Data product quality assessment approach

The quality of a data product is determined by its “fitness for use” in a particular application. Data quality is therefore determined by the requirements of the user and cannot be assessed independently of the user (Chapman, 2005; Kerney et. al., 2019). Hence the objective of a quality control procedure is to ensure that a data product meets the users’ needs. To achieve this it is important that the principles of data quality are applied to all stages of a data product’s life cycle. The various descriptions of this life cycle in the scientific literature (see for example Michener, 2003; Chapman, 2005; Peng et. al., 2022) are similar and can be readily mapped into the five stages - Collect, Curate, Integrate, Analyse and Use - defined by the SAFE 2.0 framework (Figure 3).

Prior to the project Parks Australia held two workshops to identify a list of data products that the project might assess. In those workshops Parks staff also identified a range of quality concerns associated with the data products. These concerns were categorised and organised into an initial data product quality assessment framework. The framework was inspired by the dataset quality aspects presented in Peng et. al. (2022) who identify four aspects of data quality - Science, Product, Stewardship, and Services - that arise throughout a data product’s life cycle. Two additional aspects - Requirements and Use - were added by Parks Australia to incorporate concerns associated with the quality of data product specification and the readiness of users to apply the product in their work. Each aspect is briefly described below:

- Requirements Quality: How well specified is the product. Are use cases documented and well specified?

- Science Quality: Were specific data collection and processing methods followed and did these align with best practise?

- Production Quality (including Data Quality): Does the data product content, structure, and format meet well-specified requirements?

- Stewardship Quality: How well is the data product governed, preserved, and managed?

- Service Quality: Can the product be used in functional platforms and applications? Is it findable? Are visualisation, analysis, and reporting requirements being met?

- Use Quality: How well are the products and services promoted? Are users informed and trained

This framework provided the basis for a more detailed data product quality assessment undertaken in this project. For each quality aspect in the framework a series of questions was developed and used to elicit more detailed quality concerns for each data product.

Requirement Quality

The questions on Requirements Quality (Table 2) were designed to be addressed by Parks Australia staff in order to understand how Park Australia uses the data product of interest to address the key questions outlined in Section 2.1. The questions also gauge how the use of the data products is documented internally at Parks Australia.

Science Quality

The questions on Science Quality (Table 3) were aimed at understanding if the data product is fit for its data use purpose from scientific point of view. This includes addressing how the data were collected, including if standard operating procedures were used, and if the survey design can be identified through the data product. The questions address aspects that are particularly important when data products are incorporated into higher level products by, for example, ensuring that the end users have statistically robust and reliable data for assessments and monitoring.

Production Quality

The production quality questions (Table 4) were first addressed by the project team to the best of their knowledge. The project subsequently conducted interviews with the data product's owner, manager or representative, that has sound knowledge of the data product and associated infrastructure to ensure that the questions were answered accurately.

Stewardship Quality

Questioning of stewardship quality (Table 5) naturally follow Production Quality and similarly the questions were first addressed by the project team before interviewing the data product's owner, managers, or representative. These questions are aimed at addressing the business side and continuity and future plans of the data product.

Service Quality

Service Quality questions (Table 6) are designed around understanding how the data is kept, managed and serviced. Again, the questions were initially addressed by the project team before interviewing the data product's owner, managers, or representative.

Use Quality

The final four questions under Use Quality (Table 7) were addressed by Parks Australia staff. These questions were designed to understand how the user identifies the data product and uses it in a business context, can it be interpreted and is the data product functional and ultimately fit for purpose. The last question (U4) is designed to check if feedback from user is provided to the owner or manager of the data product.

Table 2 Questions used to assess the requirement quality of data products and repositories used by Parks Australia to answer key monitoring and management questions

Ref	Question
R1	Which key questions does this product address?
R2	Are there published data gaps and use cases for this product?
R3	Are there documented specifications for this product?
R4	Is the data product regularly reviewed by end-users and the data provider?

Table 3 Questions used to assess the scientific quality of data products and repositories used by Parks Australia to answer key monitoring and management questions.

Ref	Question
Sc1	Are Standard Operating Procedures (SOPs) available and used for this data product?
Sc2	For data products derived from field studies: did the field study design follow the recommended procedures within the survey design SOP? If yes and the design was spatially balanced are the design covariates and site inclusion probabilities recorded in the meta-data?
Sc3	Where SOPs are not available, has best practice otherwise been applied?
Sc4	Are published and broadly adopted vocabularies used in this product?
Sc5	Has the product undergone expert/peer review?
Sc6	Is the data product fit for purpose for this intended scientific use cases?
Sc7	Is the confidence/uncertainty of within sample estimates or population estimates provided and are these meaningful in the context of the use case?

Table 4 Questions used to assess the production quality of data products and repositories used by Parks Australia to answer key monitoring and management questions.

Ref	Question
P1	Is the data accessible and the spatial and temporal extent of the data applicable?
P2	How much of the data is openly available? How much is embargoed?
P3	Are the data well-structured according to need?
P4	Does the product used published vocabularies and standard formatting conventions?
P5	Are spatial and temporal reference systems used documented?
P6	Does the product meet broad data quality requirements?
P7	Is the file or web service format relevant and consistent?
P8	Are uncertainty or confidence indicators provided?
P9	Can provenance and lineage of the product be tracked from creation to integration and use?
P10	Are the data real-time, near-real-time or provided in some other delayed mode?

Table 5 Questions used to assess the stewardship quality of data products and repositories used by Parks Australia to answer key monitoring and management questions

Ref	Question
St1	Does the provider have adequate funding, skilled staff and governance to effectively deliver the product?
St2	The provider provides standardised metadata?
St3	Are there formal and consistently applied data quality assurance processes in place?
St4	Does the provider have documented processes and procedures for managing archival storage of the data?
St5	Does the provider have a business continuity plan?
St6	Do the services have accreditation (e.g., Core Trust Seal or Accredited Data Service Providers)?
St7	The data provider has an explicit mission to provide access to and preserve data in its domain.
St8	Does the provider assume responsibility for long-term preservation and manage this function in a planned and documented way?
St9	Does the repository enable discovery of the data and provide consistent citation and/or persistent DOIs?.

Table 6 Questions used to assess the service quality of data products and repositories used by Parks Australia to answer key monitoring and management questions

Ref	Question
Se1	Does the provider have OGC compliant Web Mapping Services (WMS) and/or Web Feature Services (WFS) for the product?
Se2	Does the repository support other data delivery services (e g. Thematic Real-time Environmental Distributed Data Services)?
Se3	Does the provider support all of the FAIR data principles through platforms and applications?
Se4	Where applicable are there sensitive data assessment and treatment policy and procedures?
Se5	Are the data licensed for open access?
Se6	Where an embargo applies, is the product still released within a reasonable time period?
Se7	Are the providers platforms well maintained and updated?

Table 7 Questions used to assess the use quality of data products and repositories used by Parks Australia to answer key monitoring and management questions

Ref	Question
U1	Do users' (e.g. Parks Australia) procedures identify the product and characterises its uses in business context?
U2	Is the product accompanied by an appropriate level of interpretation?
U3	Does the product meet the users' functional requirements?
U4	Are User feedback captured and used in product reviews?

3.2 Evaluation method

Each of the twenty data products identified in Section 2.2 were evaluated under the data product quality assessment framework described in Section 3.1. The evaluation was done using a questionnaire style survey addressing all six “qualities”. There was a total of 41 questions, 4 addressing the Requirements Quality criteria, 7 directed to the Science Quality criteria, 10 under Production Quality, 9 under Stewardship Quality, 7 addressing the Service Quality criteria, and 4 directed to the Use Quality criteria.

Due to the vast differences in the data products, many questions were answered with either a simple yes or no. However, many questions needed further clarification and space was allowed for explanation and justification. Some questions were not relevant to the data product and were given an “NoAp”. If a question was deemed applicated but not answered a “NoAn” was recorded. It is important to note that each data product was assessed based on its current infrastructure and support. Many data products have future plans and aspirations, and this quality assessment will help identify gaps and weaknesses than can be addressed or included in future planning.

The responses from the questionnaire were used to identify the gaps in the “fitness for use” and these were used to develop a road map from data collection through to end user (Section 6.3). To further help identify these gaps and to provide a semi-quantitative assessment of each data product we ranked the responses through scoring. The quality assessment tool developed by Kerney *et. al.*, (2019) recommends a scoring system that allocates 0, 1, 2 or 3 points according to how well a data product meets a series of well-defined “best practice” criteria. This project opted for a similar but slightly simpler approach. Each of the data quality assessment questions were designed for three primary responses: “Yes”, “No”, or “Partially”. The overall quality of a data product was then based on how many of the quality assurance criteria a data product met – i.e. how many “Yes” answers, and how many almost/sometimes met – i.e. how many “Partially” answers.

3.3 Evaluation results

The semi-quantitative ranking of data products shows that the more establish data products with national or global governance were best placed with the highest “Yes” responses (Figure 8). These included OBIS and the National Reef Monitoring Network, and Seamap Australia’s National Benthic Habitat Layer (Figure 8). Each of these data products have dedicated staff maintaining and managing these data products.

Newer data products, such as Squidle+ deployments and GlobalArchive, that are still evolving and seeking ongoing support, had more “Partial” responses (Figure 8). The AMP Ecosystems data product was the one data product that clearly had the highest number of “No” responses (Figure 8). It is evident that production questions had the highest rate of “Yes” responses and requirement and use quality had the highest rates of “No” responses (Figure 8).

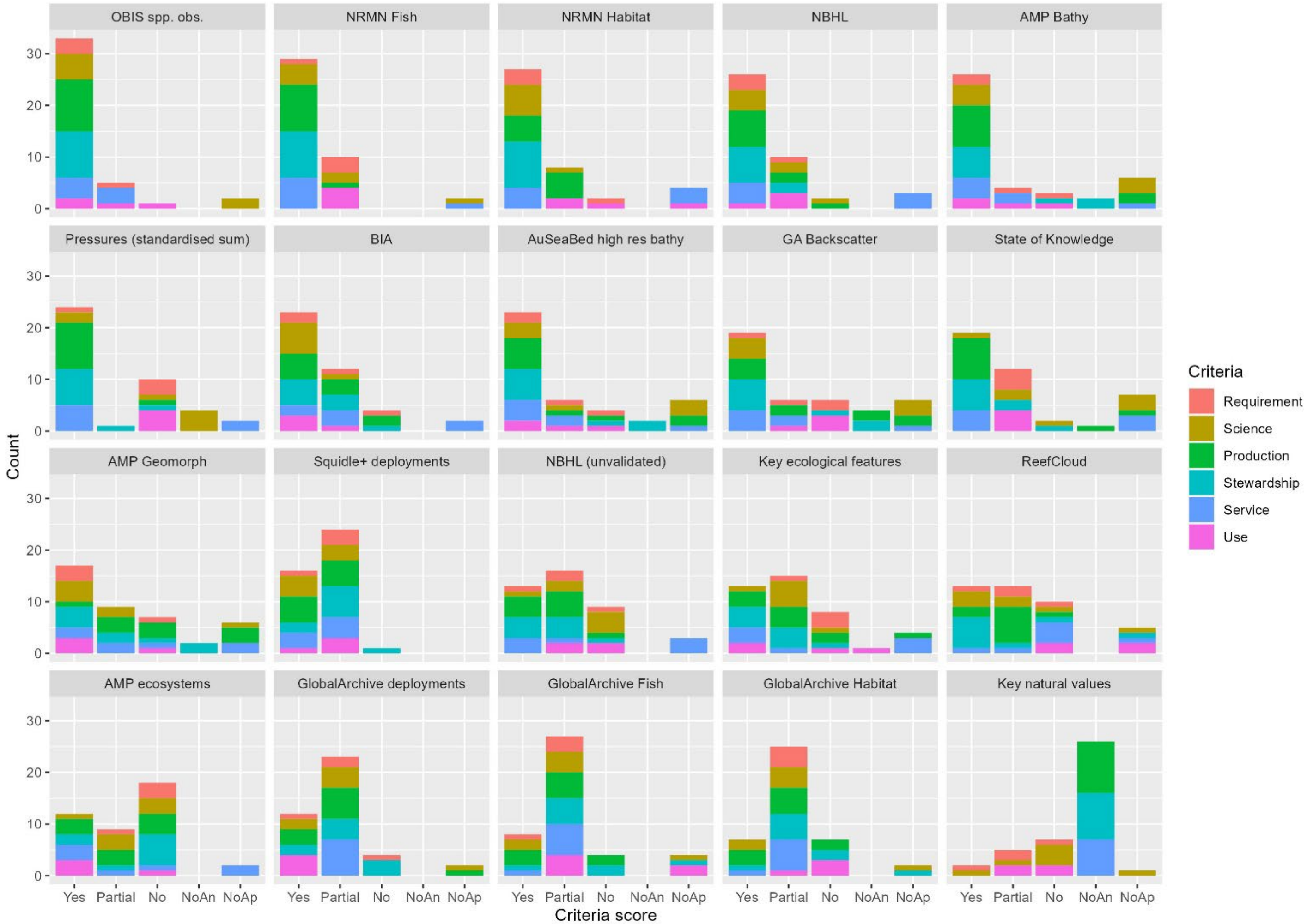


Figure 8 Summary of the outcomes of the data product quality assurance process designed to identify data products that are fit for Parks Australia's purposes. Data products used by Parks Australia when answering the Key Questions are organised from top left to bottom right according to how many of the quality assurance criteria they do meet (criteria score = Yes), partially meet (criteria score = Partial) or do not meet (criteria score = No)

4. Data and knowledge supply chain and architecture

Effective data and knowledge transfer relies on both well governed, funded and sustainable infrastructure as well as efficient and unblocked end-to-end data and knowledge supply chains. The SAFE 2.0 framework (Section 1.1) provides a useful overview of the data supply chain, the issues that can be encountered, and the benefits of resolving those issues.

Data infrastructure supporting the data supply chain must support standards-based interoperability between supply chain elements to ensure data from different platforms can be combined for modelling and reporting purposes.

4.1 Data and knowledge supply chain

Figure 9 expands on this supply chain and provides a conceptual view of data supply chains relevant to this report.

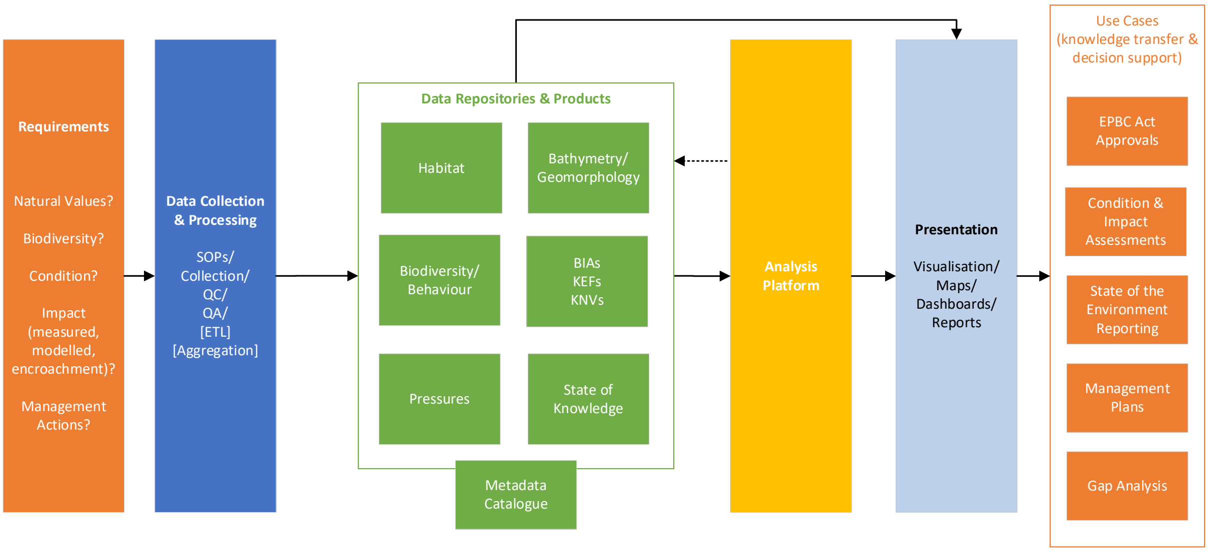


Figure 9 Conceptual data supply chain for monitoring and reporting on condition status.

Moving from left to right in the diagram:

- Requirements are identified by the end user. In this case the use cases associated with the key questions dictate what data and models are required to answer those questions.
- Data Collection and Processing follows Standard Operating Procedures (SOPs) to ensure data is collected using the same methodology, follows the same rules for quality control, provides applications for data import, quality assurance and aggregation with like data and may involve some form of Extract, Transform, Load (ETL) process.
- Data Repositories and Products represents the infrastructure used to store raw data and data products and uses standards based protocols and information about the data (e.g.

Open Geospatial Consortium standards like Web Mapping and Web Feature Services and metadata conforming to the ISO19115 standard).

- Analysis Platform is one of the various platforms and/or software used to process quality assured data products into (sometimes via modelling) higher-level data products to meet use cases and answer key questions (e.g. NeCTAR, NCI, R-Studio, Matlab, etc). Note the dotted line representing a feedback loop where an analytical process has created a higher level data product that is pushed back into the data repository for storage and access.

- Presentation represents (mostly web based) outputs for end users to interact with data or modelled outputs designed to answer key questions (Requirements) and may take the form of maps, dashboards, reports and other visualisations. Note that inputs to presentation can include raw data, derived data products and models.

- Use Cases are the predefined use cases required to answer key questions (Requirements) and drive how the Presentation elements are designed and built. Primary use cases are those with a direct relationship to key questions (Requirements). Secondary use cases allow for expanded use of the data beyond initial requirements which may include State of the Environment reporting, Environmental Economic Accounts or aggregation to Global Ocean Observing System (GOOS) Essential Ocean Variables (EOVs). Use cases may also include cross domain instances, for example economic analyses that include ecosystem services.

4.2 Data product architecture

The marine science sector has a long history of maintaining data infrastructure. The Australian Ocean Data Centre Joint Facility (AODC-JF), a partnership of six government agencies founded early this century, initiated the Bluenet Project in 2006 aimed at developing and socialising a system for recording project metadata and making it discoverable. The establishment of the Integrated Marine Observing System (IMOS) in 2007 leveraged the outcomes of the Bluenet project to lay the foundation for its data facility, the eMarine Information Infrastructure (eMII). In 2016, eMII was formally expanded to become the Australian Ocean Data Network (AODN) providing services and infrastructure for physical, biogeochemical and biological data. The AODC-JF was dissolved and resources and responsibilities transferred to the AODN, including being Australia’s representative to the International Ocean Data Exchange (IODE). The AODN Technical Advisory Group was established to guide marine research data management nationally. This early cooperative approach to data management has been the foundation for a strong data discovery and sharing culture and standardised approach to building infrastructure.

There is also well-established national governance in marine science represented by the National Marine Science Committee and its working groups including the Baselines and Monitoring working group and National Marine Science Data subcommittee (formed as a dual role of the AODN Technical Advisory Group co-chaired by the AODN Director as a standing role and a group member on a rotating basis). Additionally, IMOS have played a significant role nationally in leading the development of Standard Operating Procedures (SOPs) and other standards across the data collection and management lifecycle. The

NESP, through both the Marine Biodiversity Hub, and continuing through the Marine and Coastal Hub, have also developed a significant collection of SOPs for biological and ecosystem data.

Many of the data products included in this report are supplied by AODN partners and have followed the lead taken over the history of establishing the AODN, eventually leading to the AODN data contribution guidelines. These guidelines establish the requirements for interoperability within the AODN including metadata schema (ISO-19115), Open Geospatial Consortium (OGC) web mapping and web feature services for visualising and access to spatial and other data, developing and publishing vocabularies, and software recommendations (Geonetwork/Geoserver) to achieve compatibility and interoperability.

While this has led to a high degree of compatibility and interoperability between most AODN partners, the analysis of the entire data supply chain assessed in this project from requirements to end use cases has uncovered inconsistencies in the architecture used from minor issues to identifying data products and associated infrastructure requiring significant work and support.

There is also a significant amount of work recently completed or currently in process that will guide the further development of data infrastructure both generally and specific to the marine science domain. Of particular note are:

- The Global Open Research Commons (GORC) providing a high level specification for establishing and maintaining data infrastructure.
- The SAFE 2.0 framework completed in 2023 and outlined earlier in this report.
- Recommendation 10 of the National Marine Science Plan Midway Report identifying the need to “Establish national policy guidelines for open access to government-funded or regulatory data, provide historical dataset access, and expand the Australian Ocean Data Network (AODN)”.
- The Blue Economy CRC Data Infrastructure Design project due to completed mid-2024.
- The ARDC Planet Research Data Commons suite of programs developing standards, architecture and infrastructure across a broad range of research and management domains including the Trusted Environmental Data and Information Supply Chains program that will include a significant marine component for offshore renewables.
- The ARDC Food Security Data Challenges program project “Increasing food security through liberation of fishing and aquaculture data”.
- The Seafood Industry Australia and Blue Economy CRC project Futures of Seafood.
- DCCEEW Environment Information Australia establishing standards and infrastructure for environmental data and information management.

There will be considerable overlap between ongoing projects, and it will be critical for project teams to work closely on common goals and scope.

5. Monitoring and condition assessment

5.1 Management effectiveness context

Parks Australia has developed a park management model (Figure 10) that illustrates the core components of park management and the linkages between them. A draft set of 'management effectiveness' evaluation questions have been developed, based on this model, which are designed to help evaluate effectiveness of key aspects of AMP management, in achieving the objectives and desired outcomes for the Australian Marine Parks. These questions include:

a) Were enabling management services adequate for management?

b) Did management actions help achieve management outcomes and AMP objectives?

c) What was the level of pressures and drivers?

d) Were natural, cultural & heritage values conserved?

e) What was the level of use in the AMPs and what social, cultural and economic benefits were derived?

A major focus of the partnership between Parks Australia and the NESP Marine Biodiversity and Marine and Coastal Hubs has been on identifying monitoring needs for natural values and the pressures acting on these values (question C and D above). Parks Australia will endeavour to collect monitoring data for natural values and pressures to support broader management effectiveness evaluation.

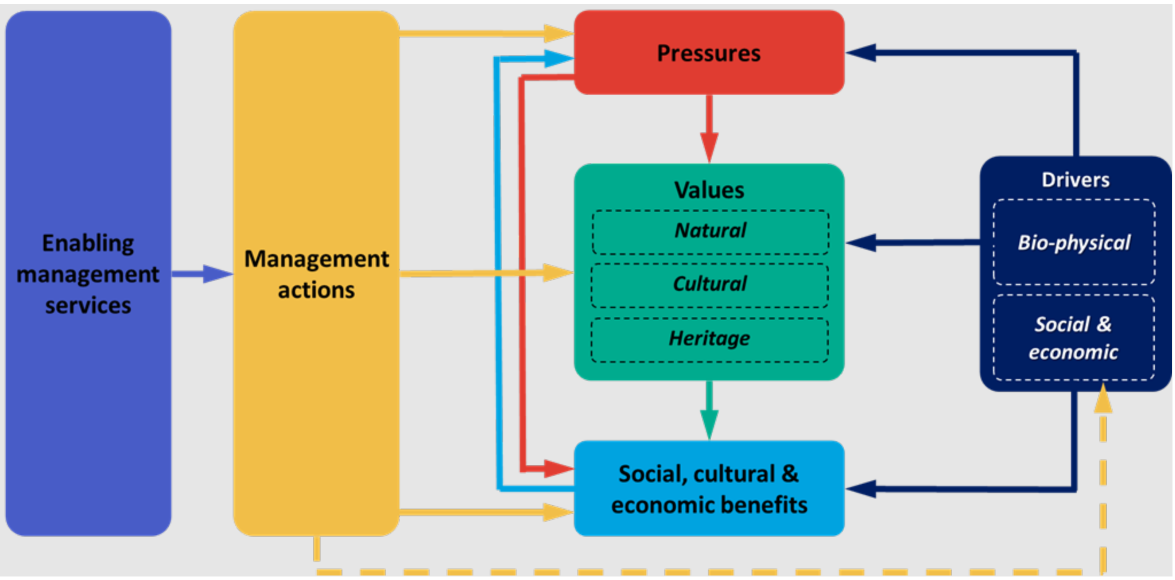


Figure 10 A reserve management model, with reference to Dambacher and Anthony 2019, Bryars et al., 2017a, Bryars et al., 2017b, Pocklington et al., 2012). The dotted line reflects limited management options for influencing social and economic drivers (with no real management options to influence for biophysical drivers).

5.2 Identifying monitoring needs for each AMP network

Over the last five years Parks Australia and the NESP have made significant progress towards establishing an AMP monitoring program to support evidence-based management of the marine parks. This has been based on a prioritization process that placed a strong emphasis on natural values and locations where Parks Australia anticipated an observable effect of management (Figure 12). This approach was piloted in the South-east marine parks network (Hayes et al., 2021) prior to being implemented nationally across the remaining four AMP networks and Coral Sea Marine Park (Dunstan et al., 2023).

5.3 Towards a national monitoring program

Since the completion of the process to identify what the monitoring needs for each AMP network, Parks Australia has built on this earlier work and is moving towards a holistic and integrated approach to a nation-wide AMP monitoring program, involving three different types or monitoring:

1. **Type 1**: monitor for long-term trends by regularly measuring the condition of natural and cultural values to detect the rate and direction of change and provide early warning of undesired outcomes.

2. **Type 2**: monitor to detect zone effects (where these are expected) by measuring and comparing the condition of natural values, and associated pressures, between zones in a given park.

3. **Type 3**: monitor to detect the effects of management on the condition of natural values and from this infer the causes of observed changes, including the effect of management interventions.

The monitoring needs and priorities identified to date (for each AMP network) have focussed mainly on types 2 and 3. These needs will be considered as Parks Australia seeks to establish a national monitoring program that also addresses the requirements of type 1. The national AMP monitoring program will be centred around a smaller number of priority parks or areas as ‘sentinels’, ideally where these three monitoring types intersect (Figure 13). This concept is based on similar approaches used nationally and internationally (Noble-James et. al., 2018; Howe et. al., in press).

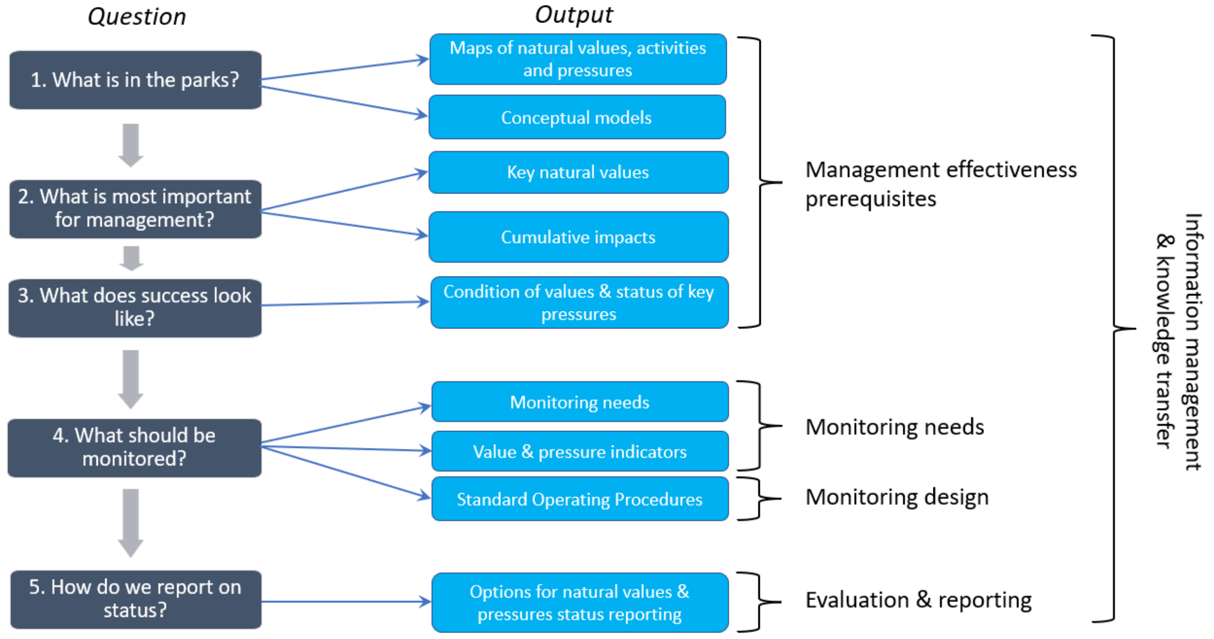


Figure 11 High level questions and outputs of Parks Australia's process for identifying and prioritising its monitoring, evaluation and reporting needs. This project (Project 2.3) focusses on *'Information management & knowledge transfer'*, as well as *'Monitoring needs'* and *'Evaluation & reporting'*.

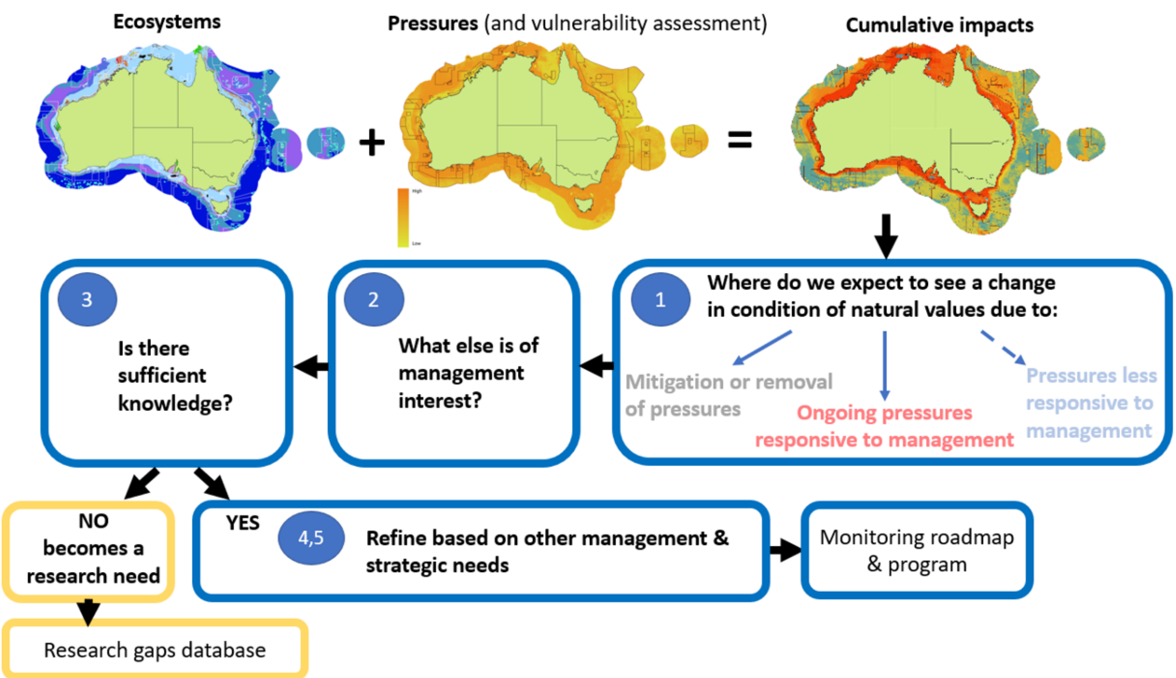


Figure 12 Summary of the process for identifying and prioritisation monitoring locations for the AMPs. This figure shows how the cumulative impacts assessment from the monitoring, evaluation and reporting process (Figure 11) was used as the first step (1) of the prioritisation process. Steps 2 and 3 were driven by Parks and NESP, with strategic / policy needs introduced at step 4, 5.

5.4 Monitoring questions

Parks Australia has developed a set of draft national monitoring questions that align with each of the monitoring types identified in Section 5.2, which in turn link to the broader management effectiveness questions C and D, and include:

1. Type 1 question: What is the rate and direction of long-term change in the AMPs?
2. Type 2 questions: What effect have zoning arrangements had on outcomes for values, benefits and pressures in Australian Marine Parks,
    - Have National Park Zones (NPZs) maintained condition of all natural values or allowed recovery from historical use?
    - Have Habitat Protection Zones (HPZs) maintained condition of habitat forming species or allowed recovery from historical use?
    - How do Multiple Use Zones (MUZs) compare to areas outside the AMPs?
3. Type 3 question: Have management approaches and interventions mitigated the effect of pressures on values and benefits in the AMPs?

These national monitoring questions align with ‘management effectiveness’ evaluation questions 3), 4) and 5), as well as KQ2, KQ4 and KQ6 and may address: (i) if NPZs or HPZs are maintaining the condition of natural values or allowing them to recover from historical use; and (ii) if the observed changes in biodiversity and ecological processes, or other features of natural values are ‘expected’ based on existing and historical management arrangements.

A series of more specific monitoring questions that sit under or align with the national questions have also been developed where Parks Australia anticipates an effect of management (Hayes et. al., 2021; Dunstan et. al., 2023). These specific questions align most closely with type 2 and type 3 monitoring and can be summarised as: (i) for natural values, has the condition of the ecosystem component(s) in the ecosystems in each park/zone improved, been maintained or declined?; and (ii) for pressures: has the status of pressure(s) in specific park/zone been mitigated (or reduced), remained stable or increased?

5.5 Monitoring and reporting hierarchy

Parks Australia has also developed a draft monitoring and reporting hierarchy based on the common language for natural values and pressures (Figure 14). The hierarchy is based on, and consistent with, the common language (Appendix B) for the AMPs and identifies the various scales and levels of Parks Australia possible monitoring, evaluation and reporting needs. It is anticipated that the draft national monitoring questions and more specific monitoring questions above will help inform ‘rolling up’ or aggregating the condition/status metrics associated with specific indicators (or indices) to different levels and scales. This approach will be used to support Parks Australia’s evaluation and reporting for AMPs, including evaluation of management plans for four AMP networks and the Coral Sea Marine Park scheduled for 2026-28.

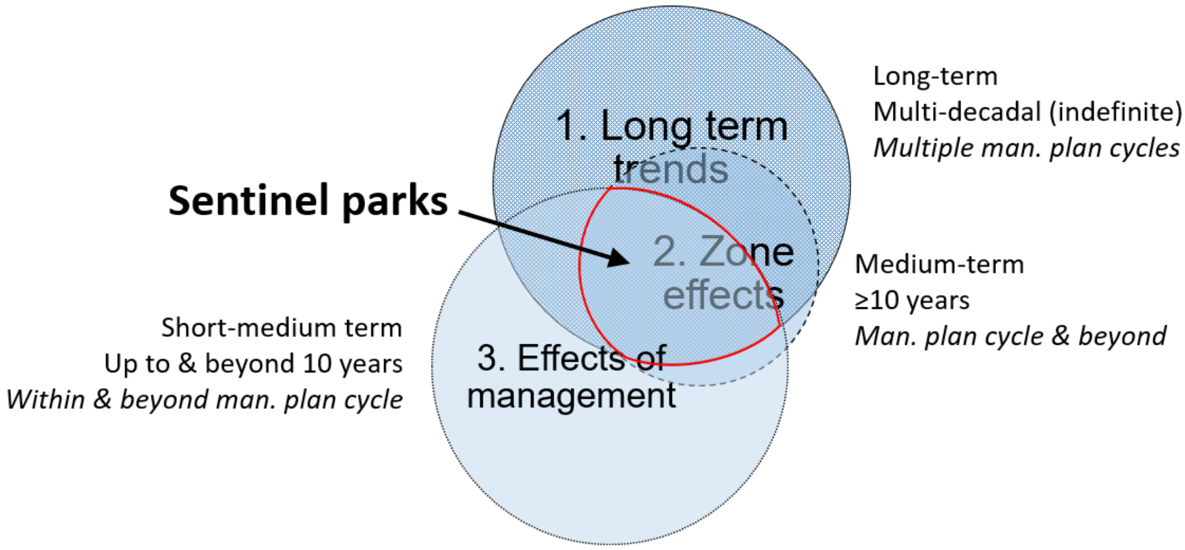


Figure 13 Parks Australia's three types of monitoring as part of the national AMP monitoring program. Sentinel Parks or areas will form the core of the national program and will be identified using a set of agreed criteria. They will be a subset of parks where the three types of monitoring intersect.

Reporting scales

1. National
2. Network / groups of parks
3. Park
4. Zone
5. *Feature?*

Finer scale reporting may occur based on footprint of values, benefits, pressures etc.

Natural values

1. Ecosystem
2. Ecosystem component
3. Ecosystem sub-component / index
4. *Indicator*
5. *Measure / metric*
6. *'Threshold'*

Pressures (& possibly activities)

1. Management plan pressure
2. Specific pressure
3. *Indicator*
4. *Measure / metric*
5. *'Threshold'*

Figure 14 Parks Australia's draft monitoring and reporting hierarchy for the AMPs, which serves as a framework for rolling up or aggregating natural values condition or pressure status to the required scale. The hierarchy is based on, and consistent with, the common language for the AMPs.

5.6 Condition assessment process

Condition assessments seek to understand the current state of a natural value and any change in its condition over time. Condition assessments are a central pillar of management effectiveness evaluation (TNC, 2007), and have the potential to provide a baseline, identify trends in condition over time and reveal the impacts of key pressures and management interventions on the condition of natural values.

Assessing the condition of complex marine ecosystems is challenging. The simplest approaches focus only on the biological components of the system, whereas more ambitious approaches may attempt to include aspects of ecosystem function or other socio-economic dimensions (Smith et al., 2021). In practise, however, many marine park monitoring programmes proceed without any validated mechanisms for assessing the condition of the natural values at the scale of individual parks (Noble-James *et al*., 2023).

Several different approaches to conducting condition assessments of natural values have been proposed in the literature (Hilton et al., 2022). These approaches attempt to address key challenges, such as data availability (data rich versus data poor scenarios), identifying baseline conditions, current condition, the limits of acceptable change (LACs) and associated thresholds that trigger management interventions (Hilton and Cook, 2022).

In data rich situations, many quantitative approaches to characterising the range of conditions (pristine through to collapse) and setting condition categories or thresholds are available. For example, Parks Victoria have used control charts designed around upper and lower LACs derived from the extreme values observed in control sites or in some of the earliest monitoring studies (Carey et al., 2015; Ierodiaconou et al., 2020).

In data poor environments, expert elicitation methods, designed to capture the inherent variability and expert's uncertainty in natural value indicators, under a range of different pressure-related scenarios (Hosack et al., 2017), provides an opportunity to quantify condition and avoid the ambiguity associated with terms such as “good” or “poor” condition.

Despite the diversity of available approaches, it is possible to identify a number of essential steps required to undertake a condition assessment (Figure 15). Note that the steps that occur within the blue square are the minimum information requirements necessary to progress a condition assessment (with expert elicitation) in the absence of initial or on-going observations of the natural values concerned.

Essential pre-requisites

The essential pre-requisites to the condition assessment process are illustrated in Figure 15 by dark green boxes. Within a management effectiveness framework, condition assessments should commence with a clear set of objectives, which flow from the management agency's monitoring questions and priorities. For AMP's these objectives are set forth by the National Monitoring Questions developed by Parks Australia (Section 5.3). Before commencing the condition assessment process, these monitoring questions would ideally be articulated using the SMART framework: Specific, Measurable, Achievable, Relevant and Timebound. This enables the condition assessment to be tailored to the context.

Natural values and pressures

The next step is to characterise the natural values within the park and identify those which are deemed to be important to the management agency and wider stakeholders, together with the relevant pressures that may impact those values. To support this step, Parks Australia has already developed a common language for: a) natural, cultural, and heritage values; (b) social, cultural, and economic benefits; (c) activities/uses and anthropogenic pressures; and (d) biophysical, and social and economic drivers (Hayes et al., 2021, Appendix B, Appendix C).

The Natural Values Common Language (NVCL) is a hierarchical, controlled vocabulary that was originally designed to assist in the identification and prioritisation of monitoring locations across the AMP estate. The NVCL distinguishes between 22 benthic ecosystems and 4 pelagic ecosystems in AMPs, together with their constituent components (as shown in the AMP ecosystems data product), each allocated to an ecosystem complex.

Ecosystems in the NVCL are delineated by habitat, depth, and other biological and/or spatial features, in a manner that ensures that their boundaries are identifiable. Defining ecosystems in this manner allows them to be mapped, thereby providing the basis for a whole-of-system, place-based approach to management. The NVCL could also support a future management evaluation process by enabling Parks Australia to aggregate or “roll-up” (Figure 14) condition assessments for individual ecosystem components into ecosystem, zone, park or regional level assessments.

For some natural values, the NVCL readily maps into other similar controlled vocabularies such as CATAMI and the controlled vocabulary that *inter alia* underpins the Seamap Australia NBHL (Section 2.2). The lowest levels of the CATAMI vocabulary include individual species identifications, and all levels of CATAMI are supported by unique CAAB codes. Individual species can therefore be unambiguously mapped into the NVCL via its relationship to CATAMI for the NVCL ecosystem components macroalgae, seagrass, sessile invertebrates and infauna (Appendix B).

The NVCL ecosystem component “mobile macro-invertebrates” can also be mapped into CATAMI but in this instance it may be useful to distinguish which NVCL ecosystems, for example coral reefs, shallow rocky reefs, shelf vegetated or unvegetated sediments, the mobile macro-invertebrates are associated with. The NVCL benthic and cryptic fish, and demersal fish, components can be mapped into the classification of fish developed by the Reef Life Survey, but unlike CATAMI this classification does not appear to have a similar CAAB code support (Appendix B).

Parks Australia has also developed a common language for pressures that may operate in marine parks and islands within those parks (Appendix C). This language facilitates the aggregation and analysis of pressures data and was used (for example) by Hayes et al. (2021) to develop the relative cumulative impact ratings for benthic habitats. Standardised estimates of pressures for the years 2013 to 2018, categorised according to the vocabulary, are available on Seamap Australia (Section 2.2).

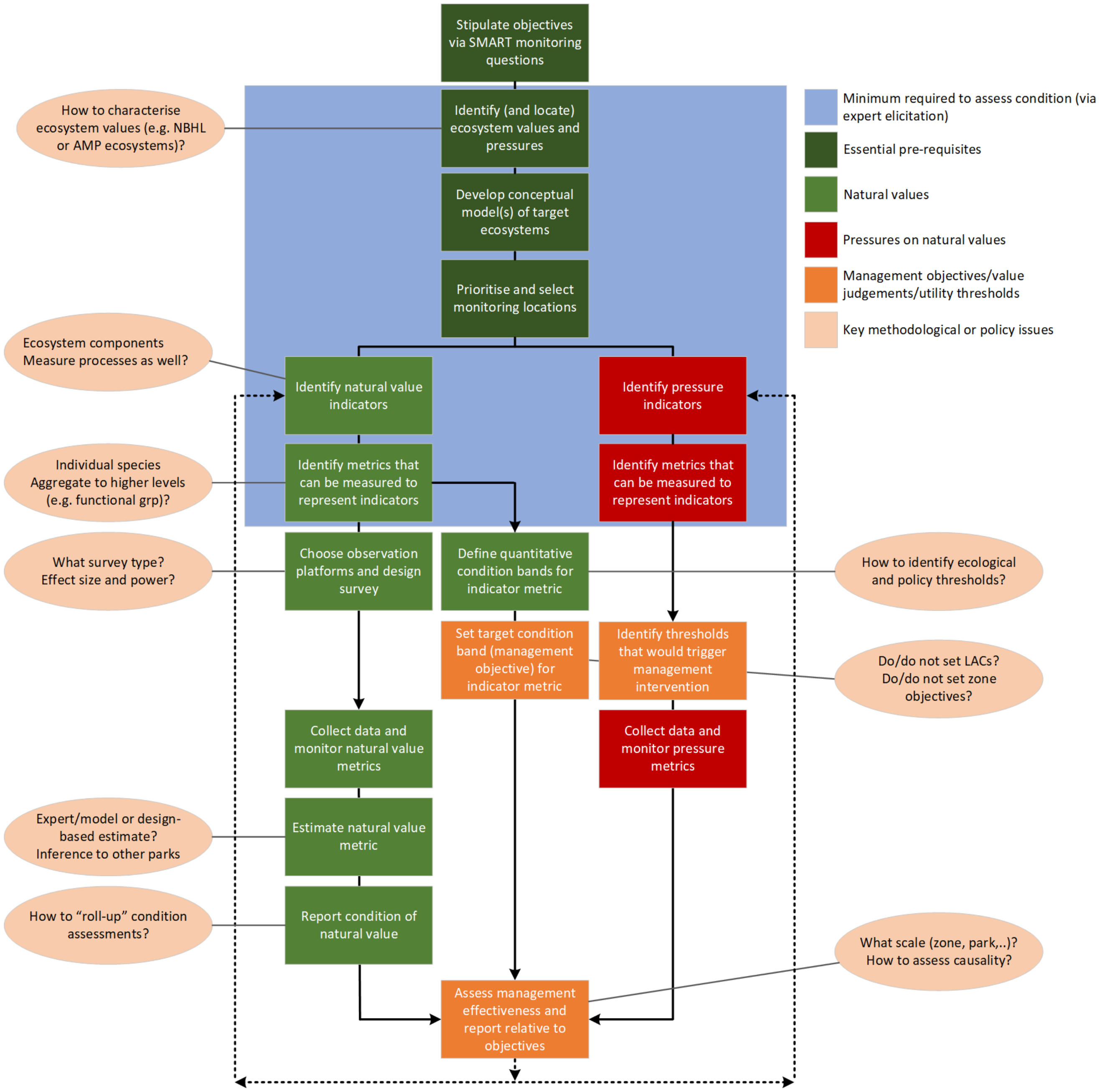


Figure 15 Summary of the essential steps required to complete an assessment of the condition of natural values within marine ecosystems, together with some of the policy and methodological issues that arise at some of these steps.

Conceptual models of the ecosystems, showing how their components interact, together with the pressures that are thought to influence these components, help to summarise current understanding of the system and identify the key values and ecological processes, as well as the drivers of relationships between them. Considering the relationships between natural values and pressures, both negative and positive, helps to identify the key components and pressures to monitor to understand changes in condition.

Conceptual models can be based on existing scientific data from research and monitoring, along with expert knowledge. There are several established approaches to developing conceptual models that represent the components and process of an ecosystem and the pressures that influence it (Gross 2003; Ogden et al., 2005; Pocklington et al., 2012). Block diagrams, cartoons and influence diagrams are often used, along with signed directed graphs (Hayes et al., 2015; Dambacher and Anthony, 2019) and other equivalent graphical structures (Mayfield et al., 2020).

Parks Australia has previously used conceptual diagrams to represent their current understanding of the linkages between different the ecosystem components of the benthic and pelagic ecosystems described in the NVCL, together with the pressures, drivers, social and economic benefits and management actions in Australian Marine Parks (Hayes *et al*., 2021).

Once the most ‘significant’ values and pressures have been identified, then monitoring locations can be identified. These locations could reflect sites that are already known to be in good condition, sites that might be vulnerable to pressures and/or sites that will help reveal zone effects, effects of management and/or long-term trends (Figure 12).

Identify indicators and metrics

Once the monitoring needs and locations have been identified, then indicators and associated metrics, i.e., what will actually be measured (see below) need to be identified. Indicators are an attribute that represents the natural value or pressure of interest, while metrics are what will actually be measured. The scientific literature identifies a variety of selection criteria that can be used to identify indicators whilst planning a monitoring strategy (Hayes et al., 2015), and many AMP relevant indicators have been identified based on experience and outcomes observed in existing marine parks (Barrett et al., 2007; Barrett et al., 2009; Babcock et al., 2010; Creswell et al., 2019; Allard et al., 2022).

Ultimately the choice of indicators and metrics should be informed both by the scientific understanding of the system, for example what best represents the natural value, as informed by conceptual models, by feasibility considerations, such as what data can be collected given realistic constraints, by relevant field experience and links to management objectives.

The project undertook a process to identify indicators and associated metrics for pressures and natural values associated with shelf reefs and unvegetated sediments in several AMPs, including Tasman Fracture and Dampier (Table 8). This initiative was conducted to support a condition assessment demonstration (in a data rich and data poor scenario) that was not concluded due to a late change direction for the project (Section 1.2).

The project team and Parks Australia recognises that condition assessments will likely require multiple indicators and metrics to be identified for natural values that can be described within a controlled vocabulary, for example the existing ecosystem components (or new sub-components) of the NVCL (Appendix B). In the first instance the indicators will also likely focus on the natural values themselves, rather than the processes or ecosystem services that they provide.

Moreover, the project recognises that by measuring the fundamental properties of individual species, Parks Australia retains the flexibility to combine and aggregate this information into a variety of indicators or indices that meet their management and reporting requirements, as and when these requirements are finalised (Table 8). Collecting data at the level of species offers Parks Australia the ability to be responsible to management and to create and report on indicator indices as needed.

More specifically, if Parks Australia monitors fundamental metrics such as abundance, density, percentage cover and size/height/length - of individual species, including habitat forming species, then it will retain the ability to aggregate this information to any desired level, such as a functional group, or to a defined entity within a controlled vocabulary, such as demersal fish, mobile and sessile invertebrates, seagrass and macroalgae within the NVCL, without jeopardising its ability to assess effect of management or influence of pressures on specific values.

The project also recognises that the process of selecting indicators for pressures can mirror that for natural values and that this is essential in order to understand drivers of change and interpret the effects of management. Future projects will therefore need to identify pressure indicators and one or more metrics that can be measured to represent that indicator.

Define quantitative condition bands and set target condition

Assessing condition requires an understanding of current condition (gained through monitoring) in relation to the spectrum of possible conditions that a natural value within a specific location might achieve. The condition of a natural value lies on a continuum from “pristine” to “collapse”, and so a condition assessment should consider all possible states, noting in particular that “pristine” condition will vary according to the value and its location.

After indicators and metrics have been identified, an optional next step (illustrated in Figure 15 by the solid arrow) is to define quantitative condition bands that represent the range of possible values but separate them into discrete categories. The purpose of condition bands is to identify the normal variation of natural values, separate this from an abnormal range, and thereby recognise that there is range of values that will likely represent fluctuations that are not of concern to managers. The width of the condition band can be tailored to recognise the range of natural variability and/or LACs. It also recognizes that there is uncertainty in our ability to define and measure condition (see Survey design below).

Where condition bands are viewed as representing the LACs, then the boundary between bands can create a threshold of management concern. The use of thresholds can be used to trigger reflection or management action and can be attached to either changes in the condition of the value itself, changes in the pressures, or both. Managers can use this information to decide whether or not a change in condition category is used to review management decisions. If LACs are not used, then condition bands can still be used to reflect natural variation in condition or uncertainty in estimates.

Table 8 Potential natural value indicators identified by NESP Project 2.3 for benthic ecosystem components and possible sub-components (Appendix B) in Australian Marine Parks

Ecosystem component/Sub-component	Potential indicators	Comments
Demersal, Benthic and cryptic fish **Hard bottom benthic fish** **Soft sediment benthic fish**	Abundance by size class (trophic and body-size group) Community temperature index (CTI) Assemblage abundance composition	Abundance by size class (e.g. “large” species) and/or trophic group (e.g. “large bodied generalist carnivores”) are well established indicators for fish that are known to show strong signals of MPA related management (Babcock *et* al., 2010; Claudet *et* al., 2010; Edmunds, 2017, Cresswell *et al*., 2019; Young *et al., 2023*). The data used to derive these indicators can also be used to estimate biomass (based on length weight relationships) or abundance by maturity status (based on length at maturity relationships) that can reflect differences across depths or bioregions. The community temperature index is used in terrestrial and marine ecosystems (Stuart-Smith *et al*., 2015) as an indicator of climate change. Demersal fish assemblage abundance composition has been demonstrated to be a sensitive indicator of ecological and environmental pressure and provides a basis for community (multivariate) control charts (Anderson and Thompson, 2004)
Mobile macro-invertebrates **Reef associated mobile macro-invertebrates** **Sediment associated mobile macro-invertebrates**	Abundance by sex ratio and size class (trophic and body-size group) Assemblage abundance composition	Similar to demersal, benthic and cryptic fish, abundance by size class by sex is a well-established indicator for macro mobile invertebrates such as *Jasus edwardsii* the southern rock lobster that are known to be sensitive to MPA related management activities (Young *et al*., 2022; Young *et al*., 2023). Similar to demersal fish assemblages, mobile macro-invertebrate assemblage composition has been demonstrated to be sensitive to environmental gradients and may provide a basis for an indicator of environmental pressures (Anderson 2008).
Seagrass and Macroalgae	Condition (presence of epiphyte/epifaunal fouling) Percent cover (species or morpho species)	Percent cover and condition are commonly used indicators of habitat forming species such as seagrass and macro-algae (Edmunds *et al*., 2017; Ierodiaconou et al., 2020; Young *et al*., 2023) that are sensitive to MPA related management activities and climate change.
Sessile invertebrates **Non-molluscan filter feeders**	Abundance by size class or shape (CATAMI/morpho species) Community diversity Index Condition (presence of epiphyte/epifaunal fouling) Percent cover (species or morpho species)	Abundance by size class of sessile invertebrates has been demonstrated to be a sensitive indicator of bottom disturbing activities, such as demersal trawling (Langlois *et al*., 2021b), and could be useful for assessing Habitat Protection Zones. Note CTI for sessile inverts is not possible due to difficulty of scoring sessile invertebrate imagery (i.e. sponges) to species level but may be possible for morpho-species.

Condition categories are different from management or zone objectives. Management objectives represent a desired or target condition that can be adjusted based on the expected condition given a range of factors or policy drivers. For example, setting management objectives may help interpret zone effects, where condition is expected to be different due to the effect of different activities or pressures. Management objectives may also reflect that it is unrealistic to expect natural values to achieve an ideal or pristine condition within a given time frame, but that the trajectory of change may be more important.

There are many approaches to setting condition bands (Hilton et al., 2022) but they can be broadly classified into expert-based or empirically based methods. As the name implies, empirically based methods rely on a series of observations that are deemed sufficiently extensive (in time and space) that the range of normal or natural variation can be discerned, and hence distinguished from abnormal variation. In practise this almost inevitably requires a practicality driven policy or expert-based determination of sufficiency, as observed for example in Parks Victoria's choice of LACs (Carey et al., 2012; Ierodiaconou et al., 2020).

Expert-based approaches can also vary from informal to formal approaches. The TNC (2007) for example encourages practitioners to adopt a pragmatic approach and simply identify an acceptable range of variation using any available information and necessary assumptions as soon as indicators and metrics are determined. More formalised approaches will likely require additional resources but provide the opportunity for a probabilistic representation of variability and the expert's uncertainty (Kuhnert et al., 2010). In both circumstances any initial expert-based determination can be amended over time. Formal (Bayesian) expert elicitation, however, has the additional advantage of providing a well-defined and mathematically coherent methodology for updating expert-based beliefs as observations are collected.

Survey design and power to detect change

Regardless of whether condition bands are set, once the natural values monitoring priorities have been identified, and indicators and metrics selected, then the process of designing the monitoring program commences. There are several decisions that need to be taken at this point, including identifying the appropriate observation platform and survey design.

Guidance on the design of surveys and the deployment of observation platforms is available through the NESP Standard Operating Procedures (SOPs) (https://marine-sampling-field-manual.github.io/). The SOPs are underpinned by a collaborative and iterative ocean best practice development process (Przeslawski *et al*., 2023) and currently provide detailed instructions for the deployment of multibeam sonar, autonomous underwater vehicles, baited remoted underwater stereo-video (stereo-BRUVs), Benthic Observation Survey System (BOSS), pelagic BRUVs, towed imagery, grabs and box corers, sleds and trawls, drop cameras, socioeconomic surveys, and microplastics, as well as statistical considerations associated with monitoring and survey design (Przeslawski and Foster, 2018; Przeslawski *et al*., 2019). These SOPs should be used to guide sampling design for Parks Australia's monitoring programme, while recognising that more detailed standard operating procedures may be required for particular monitoring programs or locations.

A critical element of survey design involves determining whether the design is likely to provide the ability to identify a change of interest if one is occurring (i.e., the power to detect an effect in relation to monitoring questions). If this is not considered at the design phase any subsequent observations may be unable to provide a manager with confidence that they can identify which condition band an indicator sits within or whether any observed change is within the bounds of natural variation, and hence within a level of acceptable change. These considerations can also affect the manager's ability to determine whether an observed change is a trend and where of interest, identify genuine differences in indicators between zones.

The ”classical” approaches to power analysis and sample size calculations are based on relatively simple formulas that rely on sometimes restrictive assumptions and require information such as the effect size and variance that is rarely available during the planning and design phase of a marine survey (Green, 1989; Fairweather, 1991, Gelman and Hill, 2007). This type of approach, however, has been used to re-evaluate the design of marine park surveys that had been operation for at least a few years (Keough et al., 2007).

More sophisticated approaches to power analysis are possible by simulating hypothetical outcomes from a model which an analyst expects to use, or has already used, to analyse data from a survey. This approach is more flexible, can be tailored so that the assumptions more closely match the (often multilevel) structure of the data gathering process (Gelman and Hill, 2007; Perkins et al., 2016), and can accommodate more realistic, but more complex, variance structures, such as spatial covariance in natural value metrics (Foster et al., 2014).

Estimating natural metric

Parks Australia’s draft monitoring and reporting hierarchy requires estimates of natural value metrics to be reported at the level of ecosystem by zone (and perhaps also for specific features). This implies that the population sample frame will be of the order of tens to hundreds of square kilometres.

The vast majority of the AMP estate lies in waters that are too deep to observe without remote sensing methods: 99.53% of the AMP estate lies in waters below 20m and 97.98% lies in waters below 40m (based on the 250m bathymetry and GDA 2020 Australian Albers projection, *pers comm* Emma Flukes, UTAS). Hence most observations of marine benthic natural values will be obtained by imagery from platforms such as drop-cameras, BRUVs or ROVs, taken at scales of square metres that are orders of magnitude smaller than the sample frame.

So whilst population-level metrics, such as percent cover, of habitat forming species, such as aquatic vegetation, in shallow waters, or island vegetation, may be observed from satellite imagery (Section 6.2), population-level estimates of natural value metrics for most of the AMP estate must be inferred from samples. Wherever possible these samples should be obtained in a manner that enables accurate and unbiased estimates of population-level metrics (Kermorvant et al., 2019; Nature United, 2023; Foster et al., 2024a; Aubry et al., 2024).

Population-level estimates of natural value metrics can be obtained in one of two ways: (i) via design-based; or (ii) via model-based estimators (Dumelle et al. 2022). Design-based estimators are relatively simple and entail less assumptions than model-based approaches. These estimators, however, can only be used with randomised sample-designs and where each sample’s inclusion probability is known. This can be problematic if randomised survey designs sometimes cannot be correctly implemented because of unpredictable weather, equipment failure, or other field constraints (Nature United, 2023), and because sample inclusion probabilities are often not recorded in meta-data records.

Model-based estimation assumes that the population metrics are a random realisation of a stochastic process represented by a statistical model. This approach is more flexible than design-based methods and can be used if the sample inclusion probabilities are unknown, either because the survey design was not randomised, or this information was subsequently lost. Expert elicitation (targeting prior distributions for the coefficients of explanatory variables) can also be readily incorporated into model-based methods, thereby enabling condition assessments to be conducted before surveys are conducted, and for these assessments to be coherently (in a mathematical sense) updated when survey data are made available.

Model-based procedures, however, make explicit assumptions about which explanatory variables (and the distribution of these variables) determine how natural value metrics vary in space (Conn et al., 2017). The accuracy of these approaches is therefore reliant on the validity of these assumptions. This approach is also more complicated and requires a relatively high-level of statistical training.

Report condition assessment and relative to policy thresholds

Once population-level estimates of the current condition of a natural value have been obtained - either through expert elicitation, or design or model-based estimators - then this estimate needs to be interpreted. When using a condition band approach managers can identify which band the estimate falls within, and how this relates to the target condition (management objective). In this scenario, reporting on condition can involve reporting on whether an indicator is within the target band (management objective has been achieved). Alternatively, the condition estimate can be reported alongside a trend in condition (e.g., improving, stable), which does not involve making a judgement about whether the value is in “good” or “poor” condition or whether targets have been met.

This step in the condition assessment process provides an opportunity to “roll-up” assessments conducted across different natural values, for example different ecosystem components or sub-components, into ecosystem by zone assessments, and ultimately up to higher reporting scales such as parks, networks and national (Figure 14). In the first instance the roll-up methodology could simply comprise a count (or proportion) of natural values metrics within a zone that are within target bounds or trending in a desired direction.

More sophisticated approaches, that introduce weighted summaries, with weights reflecting how close a metric is to a desired band, or how quickly it appears to be trending, are conceivable but this would require additional research, particularly around their interpretation and utility to Parks Australia.

Assessing management effectiveness

To assess the effectiveness of management, Parks Australia must be able to attribute observed changes in natural metrics to specific management interventions. This attribution addresses the causal type of questions (Leek and Peng, 2015) reflected in the national monitoring questions associated with Type 2 (what is the effect of zoning) and Type 3 (what is the effect of management approaches) monitoring (Section 5.3).

Causal type of questions are the hardest question types to address in a monitoring strategy. Addressing this type of question requires careful survey design, involving *inter alia*, careful selection of control sites, which may necessitate observations being collected outside of park boundaries, together with relatively advanced statistical analysis (Hayes *et al*., 2019; Nature United, 2023).

The gold-standard approach to this problem requires randomised, controlled, uninterrupted times series data that are almost never available in this context because this entails randomly assigning observation sites to “control” and “impact” groups before marine parks are established. Hayes et al. (2019) discuss the issues and available methods for addressing management effectiveness questions under the more realistic situation of monitoring commencing once the boundaries of a marine park have been established.

Table 9 provides a summary of the key steps in the condition assessment process together with the main methodological options at each step and a brief commentary on some the key issues, advantages and disadvantages associated with the step.

Table 9 Summary of condition assessment steps and key methodological issues

Step	Methodological issues	Comments
Identify and locate ecosystem values and pressures	Key issues are characterisation of natural values and mapping their location. Many statistical methods for inferring population characteristics from habitat and species (presence/absence or presence/only) data (e.g. Appendix D, Woolley and Foster, 2023, Foster et al 2024b)	Characterisation issues can be resolved through controlled vocabularies (such as the NVCL). Key challenge is high cost of data acquisition that can be approached by careful sampling design and habitat/species distribution modelling. Remote sensing techniques available for some environments (Appendix H)
Develop conceptual models of target ecosystems	Many methods for developing for conceptual models. Block diagrams based on Driver Pressure State Impact Response (DPSIR) framework primarily used by Parks Australia	This step is relatively low-cost, hence cost considerations are not a critical driver. Various methods are generally adequate for the role. Qualitative mathematical modelling has predictive utility over other methods
Prioritise and select monitoring locations	Methodology developed and implemented by NESP and Parks Australia (Figure 12)	Additional monitoring needs to be identified in relation to national monitoring question 1.
Identify natural values and pressure indicators and their associated metrics	Indicator and metric choice should be guided by monitoring questions. Again many possible choices.	Recommended indicators and metrics for benthic natural values provided in Table 8. Pressure indicators and metrics identified in Hayes *et al.*, (2021) should be re-confirmed. Need to identify indicators for both impact of pressures on values *and* for the pressures themselves.
Choose observation platforms and design survey	Several possible choices of observation platforms and survey designs. All should be ultimately guided by monitoring questions, and where-ever possible follow standard operating procedures.	Standard operating procedures for survey design considerations and deployment of observation platforms relevant to benthic natural values are available here: https://marine-sampling-field-manual.github.io/. Survey designs should minimally be randomised and stratified.
Define quantitative condition bands for indicator metric and management target	Method for defining condition bands broadly divided into empirical or expert-based. Management targets can be refined over time.	Parks Australia to confirm its preference regarding condition assessments
Collect data and monitor natural value and pressure metrics	Relevant collection platforms for benthic natural values include baited and unbaited, stereo or mono, underwater video, divers, drop cameras, towed-video, Remotely Operated Video (ROV), and Autonomous Underwater Video (AUV) and Benthic Observation Survey System (BOSS)	Observation methods will largely rely on underwater cameras with method contrasts around: (i) mono versus stereo camera, (ii) stationary versus mobile platforms; and (iii) baited versus unbaited. Mobile gears have good survey area coverage but relatively high initial and operational costs. Stationary platforms are generally simpler, with lower cost, but with limited spatial coverage. Stereo methods recommended in all cases.
Estimate and report condition of natural value	Estimators can be design-based or model-based.	Design-based estimators are relatively simple and unbiased for randomised samples. Model-based estimators are able to address complex (causal) monitoring questions and reflect the effects of spatial auto-correlation.

6. Roadmap

6.1 Natural values common language

Common controlled vocabularies are an important enabler of the FAIR data principles (WABSI and WAMSI, 2023) (Section 1.1). They also assist in the identification of indicators and the roll-up of condition assessments to management relevant scales (Section 5.6). The NVCL, however, was designed to support a monitoring prioritisation process. Its status, and role within, Parks Australia's Authorisations, Incidence Response and State of Knowledge programmes has yet to be clarified or codified.

The ecosystems (habitats) within the NVCL align closely with the marine components of the IUCN global ecosystem typology 2.0 (Keith et al. 2020) but they are not consistent with the controlled vocabulary used by Seamap Australia to describe marine benthic habitats.

The NVCL ecosystem components seagrass, macroalgae, sessile invertebrates and infauna can be readily mapped into the controlled vocabulary used by Seamap Australia to describe biota and the CATAMI classification (Appendix B). The NVCL ecosystem components mobile macroinvertebrates, benthic and cryptic fish, demersal fish, coastal epipelagic fish, oceanic epipelagic fish and mesopelagic fish, however, can be mapped to CATAMI but only at a very broad classification of “bony fishes”. The benthic and cryptic and demersal fish components, however, can be aligned with the Reef Life Survey classification is this were deemed to be useful.

At this juncture it is important for Parks Australia to confirm if it intends to proceed with its current classification of marine benthic habitats (via the NVCL ecosystems) or seek to adopt the more detailed (but data intensive) delineation of habitats provided by Seamap Australia via the NBHL. If Parks determines to continue with the NVCL we recommend that it: a) consider if there is utility in aligning the classification of Benthic and cryptic fish, and demersal fish, with the vocabulary developed by the Reef Life Survey; b) publish the NVCL on a controlled vocabulary service (e.g. Research Vocabularies Australia) to ensure reliable machine and human access; and c) establish a multi-stakeholder governance mechanism to ensure controlled changes to the classification system over time.

6.2 AMP ecosystems data product

Genesis and future role

The AMP ecosystems data product was developed by NESP and the AMP science unit to support a monitoring prioritisation process (Hayes *et al*. 2021, Dunstan et al. 2023). Unlike the NBHL, which lacks national coverage, the AMP ecosystem product maps the NVCL ecosystems at a national scale (Figure 16) and has therefore emerged a critical data product for answering Key Question 1 (Section 2.2). Like the NVCL, however, the AMP ecosystems data product was not designed to support the day-to-day operations of Parks Australia's other management programmes, and its role within the day-to-day operations of these other programmes has yet to be clearly articulated or codified. However, the AMP ecosystems produce is being used in the National Climate Risk Assessment and the National Ecosystem Accounts, and the NVCL has been aligned with the System of Environmental-Economic Accounting framework.

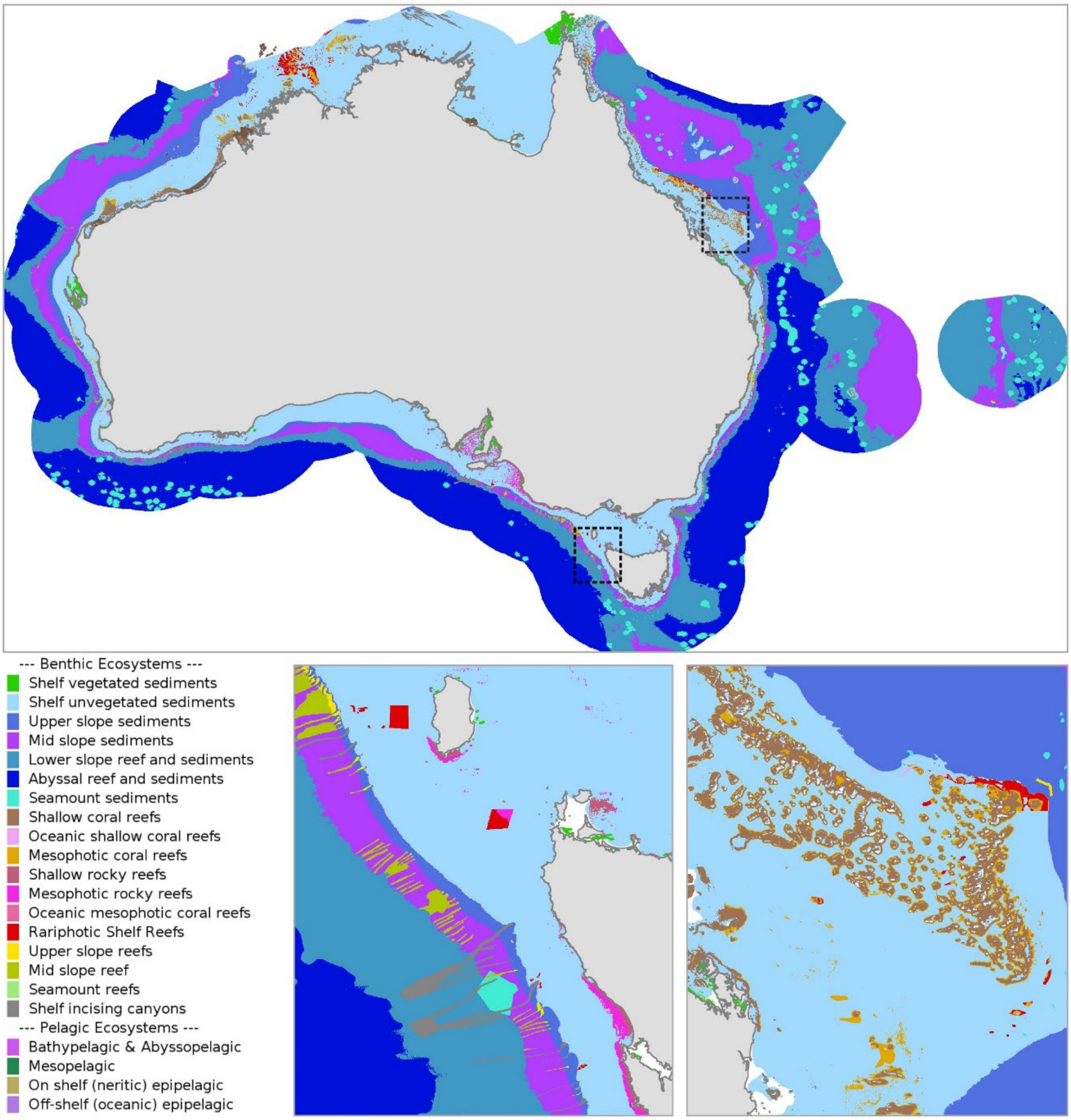


Figure 16 The [AMP Ecosystems]() data product (2022 version) available from Seamap Australia, including insert maps to show the detail of some of the Natural Value Common Language ecosystems depicted within the data product.

We recommend that Parks Australia science unit confirm what role the AMP ecosystems data product plays in the authorisations, environmental incident response and State of Knowledge programmes, and whether or not the current classification of complexes, ecosystems and components serves their needs. The ecosystem components of the AMP ecosystem data product (and the NVCL) are also currently invariant across major bioregional boundaries. Parks Australia should confirm with these programmes if this is problematic and if so, explore ways to address these differences.

If it is determined that the AMP ecosystem data product serves a useful role beyond that for which it was initially designed, we recommend that: a) Parks Australia consider the changes discussed in the subsequent sub-sections of this report, and b) develop a formal process to maintain and update the AMP ecosystems. This should include the process for inclusion of new data sources and the updating of any existing data sources. This process should be managed by Parks Australia. The process must be able to accept data that is collected using standards and practices that are practically achievable by the Australian marine science community and be able to leverage both publicly and privately funded data.

Resolution

The 2022 version of the AMP Ecosystems dataset is largely derived from the Geoscience Australia 2009 250m Bathymetry Grid. Since the development of this data product the Australian national bathymetry grid has undergone substantial improvements with the release of an updated (2023) version at the same spatial resolution.

The updated version benefits from the integration of a considerably expanded national collection of high-resolution multibeam surveys in the construction of this DEM surface. Notable improvements have also been made in shallow waters, and projection inaccuracies in the Coral Sea that previously resulted in features being displaced by several kilometres have now been corrected. We recommend that this updated bathymetry grid be incorporated into the foundational layer of the AMP Ecosystems product.

Accuracy and uncertainty of ecosystem component maps

The structure of the NVCL and AMP ecosystems data product currently implies that all of the ecosystems components defined by the NVCL to be within an ecosystem are present whenever and wherever their associated ecosystem is identified as within an area of interest.

Accurately identifying the location of ecosystems such as reef in the Australian EEZ, and their associated ecosystem components, remains a significant challenge. The current AMP ecosystems data product contains initial predictions of the location of reef but does not include information on the certainty associated with these predictions. As a result, Parks Staff often turn to other data products, such as AMP bathymetry, geomorphometry and backscatter data products as a source of supplementary material (Section 2.1).

A more scientifically robust approach would be to describe the presence or absence of an ecosystem and its associated components in a probabilistic manner. This can be achieved through the use of habitat distribution models based on the multinomial family of generalised linear models (GLMs) (Figure 17). Importantly, this approach also provides for a coherent, probabilistic, representation of the uncertainty associated with the map.

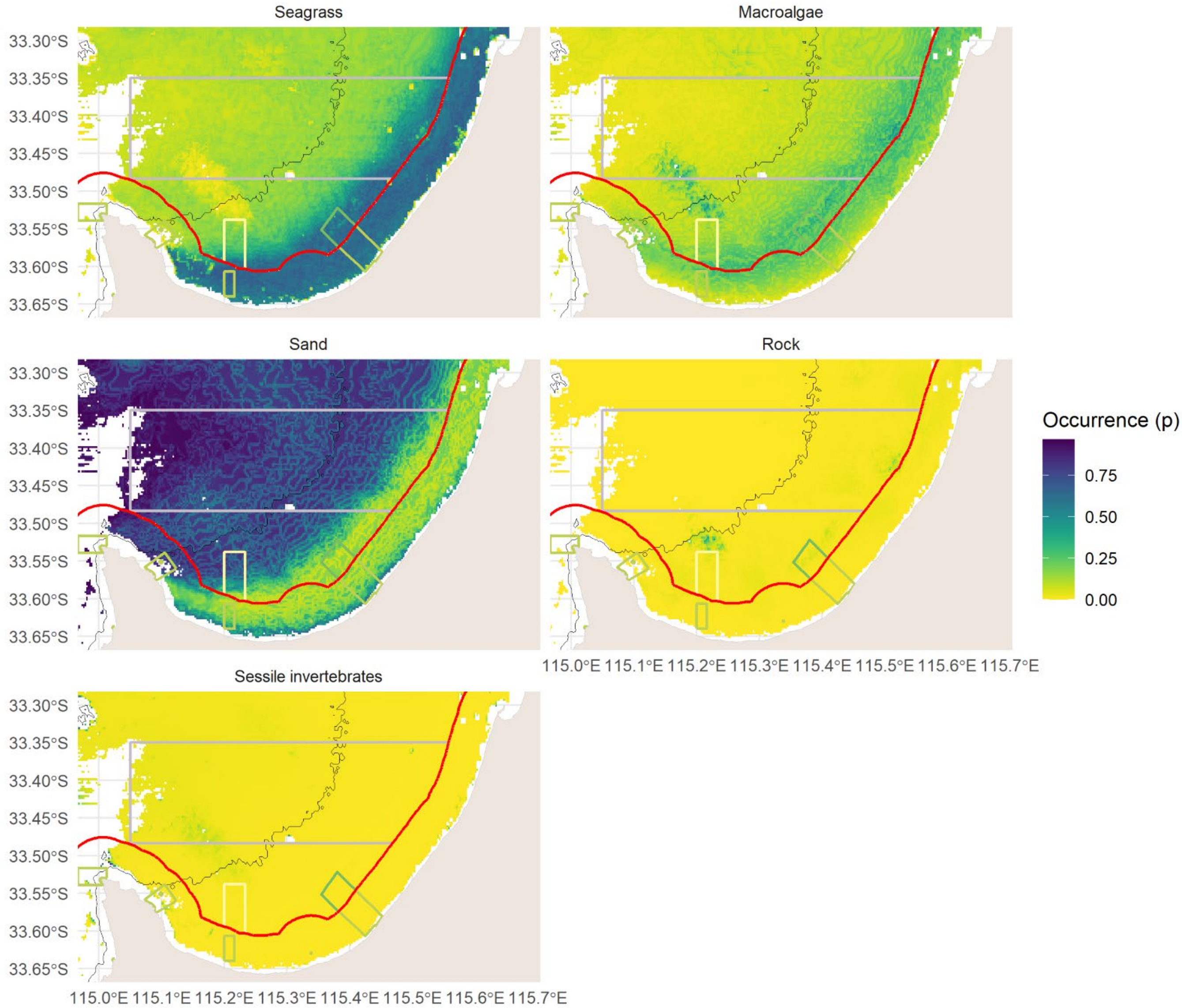


Figure 17 Multinomial GAM model predictions for the probability of occurrence of the AMP ecosystems shelf reefs (rock) and unvegetated sediment (sand), together with the ecosystem components seagrass, macroalgae and sessile invertebrate for the Geographe Bay region. State and Commonwealth marine park boundaries are shown. The red line delimits state and commonwealth waters. Bathymetric contour at 30 m is shown (Source: Spencer *et al.*, in review).

At Parks Australia's request, the project demonstrated the use of (Bayesian) multinomial GLMs to map habitat forming ecosystem components such sessile invertebrates, seagrass and macroalgae. The approach is described in detail in Appendix D, together with three options explored by the project for representing the uncertainty associated with the predicted distribution of these ecosystem components.

Canyons and Seamounts

Canyons and seamounts are not consistently located between the AMP Ecosystems and Key Ecological Features data products because each product uses a different source for the features. The AMP Ecosystems data product uses a combination of Yesson et al. (2011) and Heap et al. (2006) to derive seamount boundaries, with canyon features extracted from Huang and Nichol (2015), whereas the Key Ecological Features data product exclusively uses Heap and Harris (2006) to derive both feature types.

We recommend that common authoritative source(s) of information be used to consistently identify these important seabed features in both data products. Furthermore, the criterion for identifying seabed features as 'seamounts' versus other raised seabed features should be clearly defined. For example, Dove et al. (2020) define seamounts as *"prominent features rising more than 1000 m above the surrounding relief",* while Staudigel et al. (2010) offer a more inclusive definition: *"any geographically isolated topographic feature on the seafloor taller than 100m, but not including features that are located on continental shelves or that are part of other major landmasses"*

Although the morphological classification proposed by Dove et al. (2020) has gained traction for widespread adoption, many features identified as seamounts in the Yesson et al. (2011) and Heap and Harris (2006) datasets do not meet this elevation threshold. From an ecological standpoint, features rising prominently from the surrounding seabed may still function as seamounts, even if they fall below this arbitrary relief threshold. For instance, all seamounts within the Huon AMP fail to meet this elevation criterion and thus do not qualify as seamounts according to the Dove *et al*., (2020) definition.

The current data sources used in classifying seamount features in the AMP Ecosystems data product are coarse in resolution and now outdated. Given the availability of much higher resolution multibeam bathymetry, there is a compelling case for updating the footprints of these features in the AMP Ecosystems and KEFs data products. A more recent and detailed polygon dataset, *Seamounts and raised seabed features of the Australian Margin* by Flukes (2023), provides a refined alternative data source.

This dataset aggregates seamount and raised seabed feature footprints from multiple data sources around the Australian continental margin and validates them against the GEBCO 2023 global terrain model (15 arc-second interval grid) to remove any obviously erroneous features. Additionally, it re-digitises feature boundaries in areas where recent high-resolution MBES bathymetry is available. Incorporating this dataset into the AMP Ecosystems data product to replace existing seamount footprints would significantly enhance the spatial accuracy of the 'seamount sediments' classification (Figure 18; see also Appendix F).

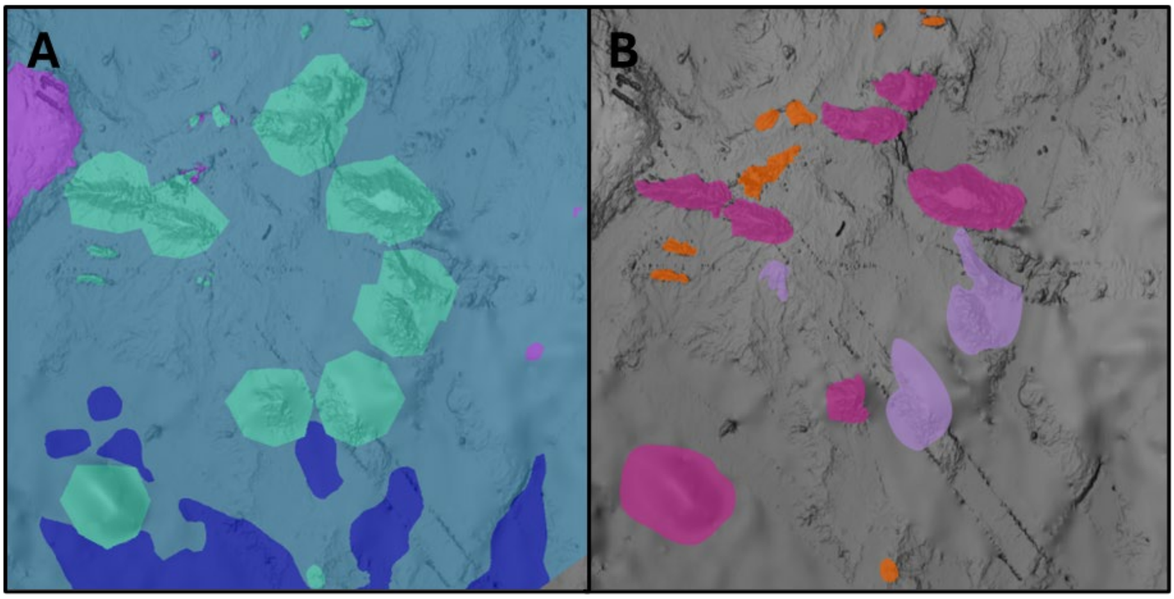


Figure 18 Sample region in the Norfolk Basin (https://seamapaustralia.org/map/#bc4deda4-9f84-4bc5-af49-266672384c66) for (A) the AMP Ecosystems data product; and (B) the refined footprints of raised seabed features contained in the Flukes (2023) dataset proposed for adoption in the AMP Ecosystems product. This figure demonstrates the inaccuracies with the current AMP Ecosystems classification of 'seamount sediments'. Teal features in (A) indicate regions currently classified as 'seamount sediments'. Pink features in (B) show refined feature boundaries that meet the 1,000m vertical relief threshold of Dove et al. (2020), while mauve boundaries delineate other large, raised seabed features that do not meet this relief threshold, with orange polygons showing smaller raised seabed features (cone/knoll/pinnacle, per the Dove et al. (2020) morphological features classification).

Shallow coral reef habitats

Shallow and mesophotic reef habitat classifications in the AMP Ecosystems data product were initially based on the NBHL supplemented by the Allen Coral Atlas and other sources of direction observation, that were subsequently supplemented by model predictions in waters outside of state jurisdiction (Hayes *et al*., 2021; Dunstan *et al*., 2023).

While this approach provided a consistent national approach it resulted in significant errors where the resolution of the available bathymetry data was poor. This resulted in most of the Pilbara, Kimberley, and parts of southern Great Barrier Reef having large false positive reef areas, overestimating the area of shallow reefs in these regions by an order of magnitude (Figure 19).

The reef predictions incorporated into the AMP Ecosystems data product would likely improve if the updated 2023 national bathymetry grid were incorporated as its foundation. This update, however, may not remove errors in reef predictions across much of the inshore regions of northern Australia. The bathymetry is sparse in the Kimberley below 15 m, and the inshore areas across the Northern Territory and the Gulf of Carpentaria are poor resolved in most places. Accurate estimates of reef locations in these locations will likely therefore need to incorporate additional sources of information.

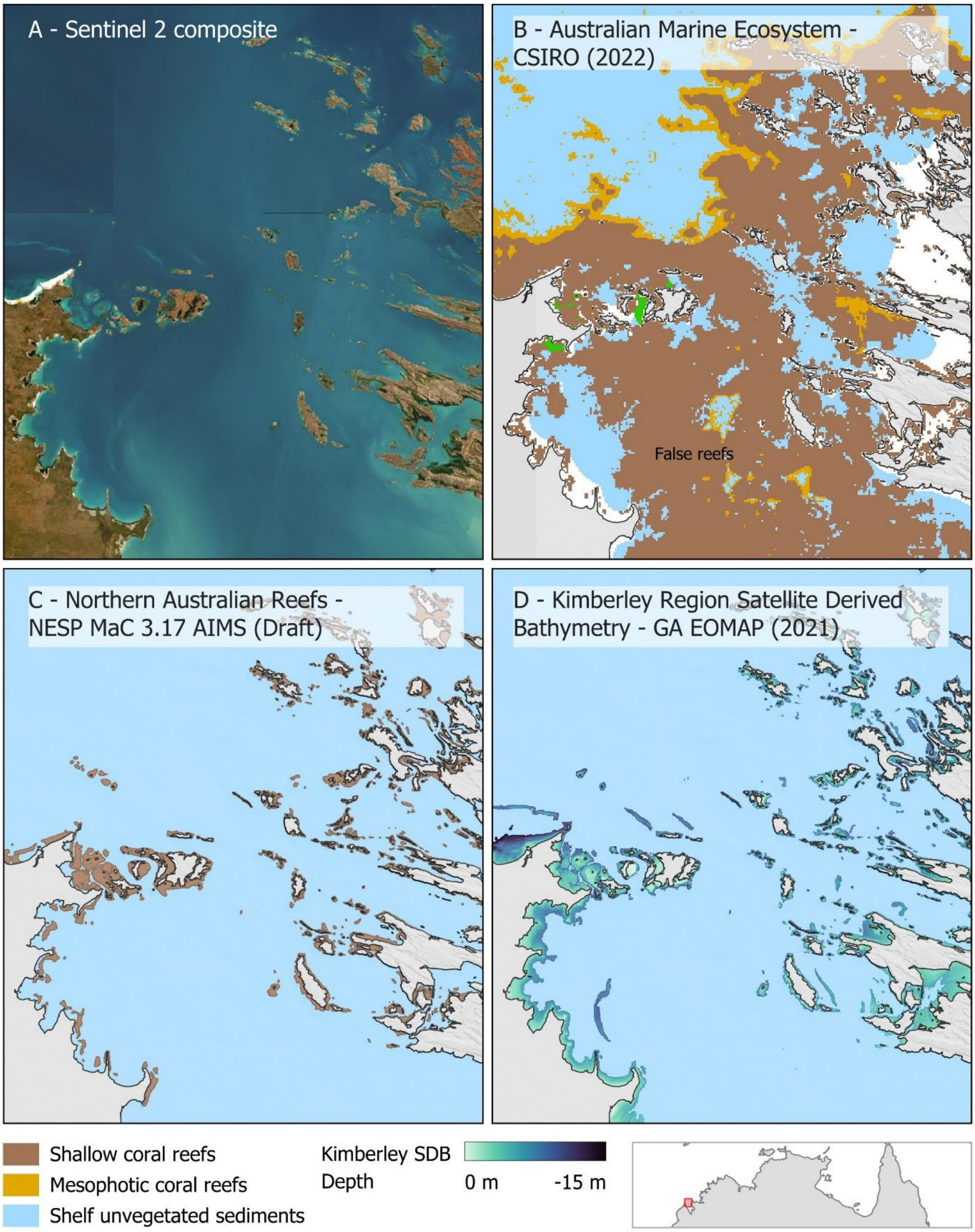


Figure 19 Map showing that the shallow and mesophotic reefs classification in the AMP Ecosystems layer in the Kimberley region (panel B) is substantially wrong and could be improved with new datasets. Panel A shows satellite imagery for the area. Panel C shows reef boundary mapping from manual review of satellite imagery. Panel D shows high resolution satellite derived bathymetry that aligns very closely with the coral reef mapping. These sources of data should be used to improve the reef mapping used in an updated version of the AMP Ecosystems layer.

The goal of the NESP MaC 3.17 project is to map the boundaries of the shallow and mesophotic reefs of northern Australian using satellite imagery and available bathymetry. All reefs will be mapped manually and carefully reviewed against multiple data sources and available existing field data. This project will map all shallow reefs (< 40 m deep) from the Houtman Abrolhos Islands in Western Australia through to the tip of Cape York in Queensland. The data from this project is expected to be published early 2025.

An example of the draft mapping from this project is shown in Figure 19 panel C. The data from this project could be used directly in the updated AMP Ecosystem habitat map, or if a modelling approach is used it could be used as model training and validation data.

The reef mapping undertaken by AIMS in the Coral Sea (Lawrey, 2024c) provides another dataset that could usefully be incorporated into an update of the AMP Ecosystem data product to improve the accuracy of the reef habitats in the Coral Sea. This dataset consists of mapped reef boundaries of shallow and mesophotic reefs across the Coral Sea, along with coral cays and the atoll platform boundaries.

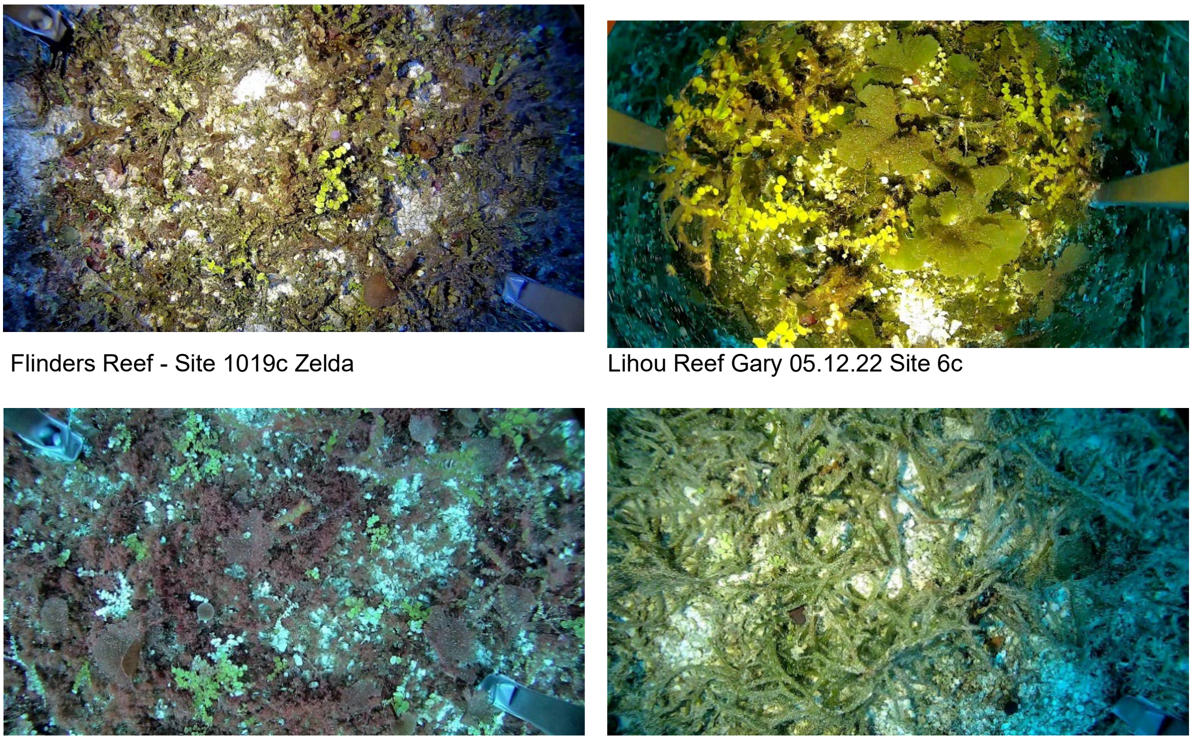
Other potentially useful sources of information in the Kimberley are the shallow reef features represented by the Kimberley Region and WA Reefs Satellite-Derived Bathymetry Acquisition (Twiggs, 2023) (Figure 19 panel D). While this data is incorporated into the GA 2023 250 m bathymetry it only corrects the shallow areas and so care needs to be taken not to incorrectly interpret the poor quality deeper bathymetry as reefs.

Coral Sea Oceanic Vegetation

The Coral Sea contains over ten very large coral atolls. Most have a gently sloping bowl-shaped lagoon filled with soft sediment, with an occasional patch reef every couple of kilometres. A significant portion of the deeper lagoonal sediments is covered by a diverse range of submerged aquatic vegetation, primarily erect macroalgae and erect calcareous algae such as *Halimeda* (Figure 20).

This lagoonal vegetation covers an area of almost 9000 km^2 but was previously unmapped due to the lack of surveys in the Coral Sea lagoons and its depth (30 - 60 m) making it very difficult to see from satellite imagery. The large extent of this vegetation habitat was first identified during the mapping of the reefs in the Coral Sea from satellite imagery (Lawrey, 2024c).

The existing Natural Values Common Language has a habitat classification of ‘Shelf vegetated sediments’ that refers to sediment habitats on the continental shelf that support marine macroalgae or seagrass. We propose that the NVCL should be expanded to include a corresponding ‘Oceanic vegetated sediments’ classification for the newly identified submerged aquatic vegetation habitats in the atoll lagoons in the Coral Sea’s atoll lagoons.



Flinders Reef - Site 1019c Zelda

Lihou Reef Gary 05.12.22 Site 6c

Lihou Reef Zelda 05.12.22 Site12d

Lihou Reef Zelda 06.12.22 Site 36b

Figure 20 Example drop camera images showing the vegetation on the bottom of atoll lagoons of Flinders and Lihou Reefs. The vegetation consists of a wide variety of erect red and green macroalgae and erect calcareous algae. Images: Tol *et al*., (2023)

The extent of the submerged aquatic vegetation was mapped as part of this project from satellite imagery and is available as the dataset [Coral Sea Oceanic Vegetation (NESP MaC 2.3, AIMS)] by Lawrey (2024). This dataset shows that a substantial portion of the lagoonal areas in the Coral Sea are covered in vegetation (Figure 21). This dataset should be considered for inclusion in a future version of the AMP Ecosystem map layer, or as a basis for further study and refinement. Further details of this mapping are provided in Appendix G.

This mapping offers a preliminary estimate of these habitats but comes with significant limitations. Most Coral Sea vegetation features are at depths of 30-60m, which is 2-4 times deeper than typical mapping from satellite imagery and was only possible due to the very high water clarity in the Coral Sea. Mapping these features at such depth required multiple techniques to reduce noise and anomalies, as well as manual mapping to leverage the visual structure of features to further minimise noise.

The final mapped vegetation features show good agreement against drop camera surveys conducted by JCU in 2023 (Tol et al., 2023; Appendix G), however, the key limitations of this validation were the limited coverage of the drop camera surveys, covering only 10% of the mapped area, and the significant scale difference between the drop camera survey samples ($1m^2$) and the satellite derived vegetation mapping (~ 400 m patch size).

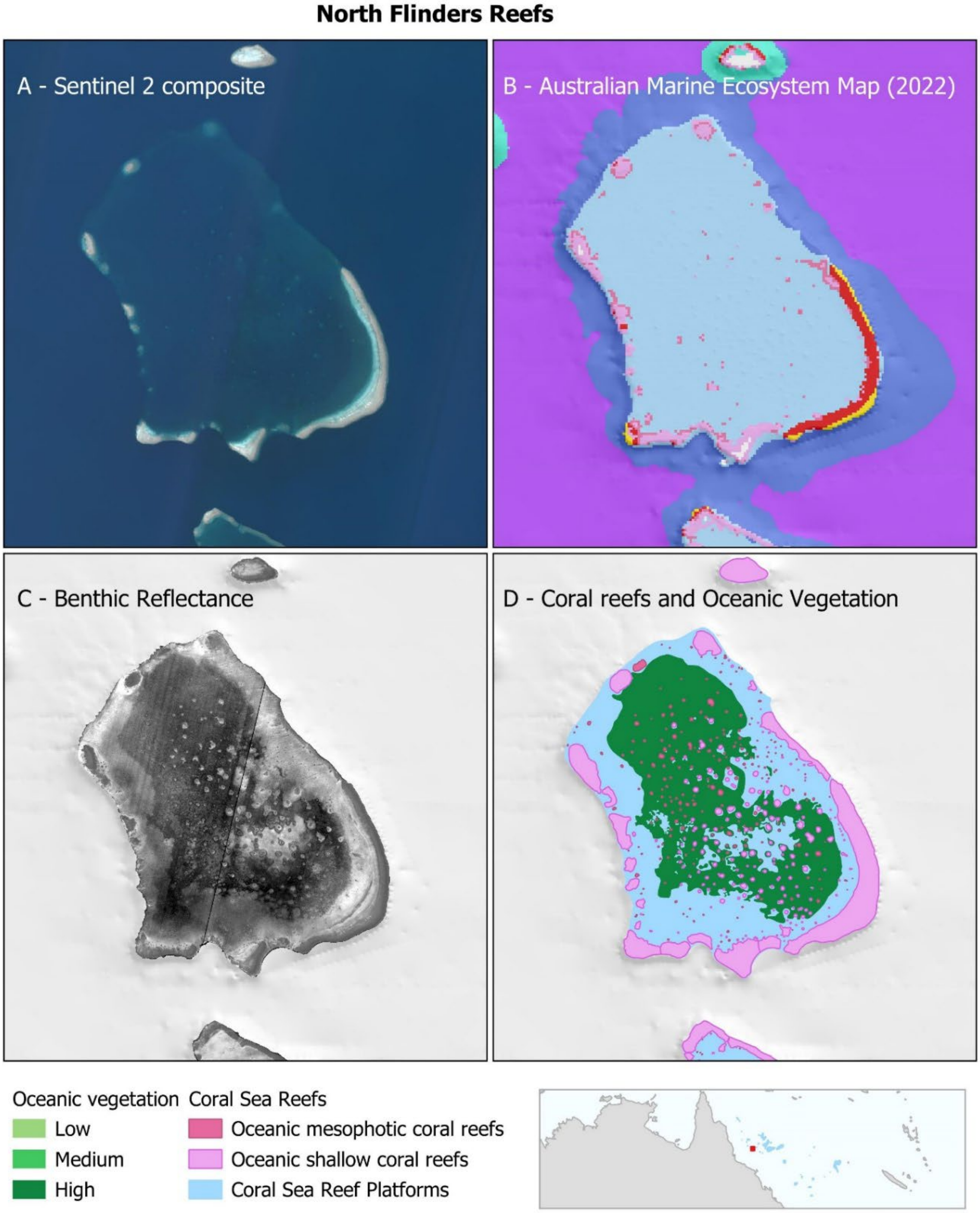


Figure 21 Mapped oceanic vegetation and reefs in the Coral Sea for North Flinders Reefs (Panel D). Panel A shows a composite satellite image of the region for reference. Panel B shows the existing AMP Ecosystems (2022) map highlighting the lack of oceanic vegetation, and so the lagoon would be considered as unvegetated sediment. Panel C shows the benthic reflectance estimated from the satellite imagery and high-resolution bathymetry (Lawrey, 2024a). Submerged aquatic vegetation and reefs appear as dark areas and sand as pale areas. Panel D shows the resulting map of Oceanic Vegetation (Lawrey, 2024b) and reef boundaries classified by the NVCL (Lawrey, 2024c). This mapping was completed across the entire Coral Sea (within the Australian EEZ).

6.3 Data products quality assessment

Data product assessments took place over the 2023/24 calendar years and it should be noted some recommendations may already have been met.

Through the process of the quality assessment and discussion with project participants and data custodians, some common requirements across the marine data landscape became apparent and are included here as observations or general recommendations for the marine science community and relevant stakeholders to consider.

General Recommendations

National Standards and Service Principles

It is critical our long-term monitoring data sources and services take a well governed, standards-based approach throughout the data lifecycle, utilising Standard Operating Procedures (SOPs) applying consistent field data collection methodologies, QA/QC, and applying FAIR data principles using national scale data infrastructure where needed. These well governed and implemented approaches help to create high quality, trusted datasets as the basis for modelling and decision support. The AODN has led the way establishing this approach and has had strong support and adoption by the Australian marine science community and by others including the NESP marine focused hubs and various Government departments (e.g. Geoscience Australia, AIMS). However, outside these organisations, it is evident from the assessment carried out by this project there remain gaps in the implementation of these standardised approaches by some other marine science data providers that affect the delivery of critical data products. It can be difficult to access these data easily, and with confidence in their quality.

There are some well documented frameworks available that could be used to establish a marine science specific data management and infrastructure framework. The Global Open Research Commons (GORC) developed by the GORC International Working Group under the Research Data Alliance (RDA) provides an overarching model that can be used to highlight important elements for such a framework. Additionally, the Shared Analytic Framework for the Environment (SAFE 2.0) developed by WABSI, WAMSI and the ARDC provides a more detailed set of requirements and a useful checklist.

Likewise, there is much to be gained through documenting the learnings and practices of facilities like the AODN, the NESP Marine and Coastal Hub, Geosocience Australia and others who have implemented well governed, standards-based approaches to managing and delivering their data.

Establishing this framework would provide the information and guidance required for individual groups to establish data infrastructure as part of an integrated federation of data portals for priority long term monitoring datasets. This is in line with the National Marine Science Plan 2015-2025 Midway Report recommendation 10 to “Establish national policy guidelines for open access to government-funded or regulatory data, provide historical-dataset access, and expand the Australian Ocean Data Network (AODN)” and consistent with broader goals of Environment Information Australia.

We recommend Parks Australia endorse establishing a marine science specific data management and infrastructure framework wherever possible, and where needed, request relevant data providers adopt and implement standards-based approaches as demonstrated by facilities like the AODN.

Access to sensitive and embargoed data

From several of the platforms assessed, user comments received, and a general understanding of these and other platforms, it's evident there is much data collected in the public good that is not always made openly accessible. This is often for sound reasons (e.g. defence, commercial or cultural sensitivities). But in many cases data is not published, released from embargo, or made otherwise accessible in a timely manner. Some generalised examples include:

- Data collected with public funding where the funder stipulates data must be made public but is neither published by the researcher nor verified by the funder as a follow up action.
- Data withheld or published under a limited time embargo but never released following expiry of the embargo as open data.
- Data collected for industry purposes, which are considerable in volume, but never released either publicly or for research purposes outside its original intention when commercial or other sensitivities have abated.

We recommend Parks Australia conduct an audit of their funded surveys and, where contractually or otherwise requested, ensure all data has been published in a data repository with open access and includes a well-developed metadata record conforming to the ISO19115 standard.

We also recommend Parks Australia contact operators of the Australian portals listed in the assessment and highlight the significant scientific benefits and associated outcomes that can be achieved by regularly asking their data contributors to maximise the amount of their data that is made openly available. In addition, where not already currently practiced, request portals have established policies and compliance reporting to ensure data is appropriately made openly accessible in a timely manner (e.g. time limits on embargoed data).

Sustainability and funding

Some of the portals included in the assessment cited ongoing funding as an issue. Funding for platform development is generally available under a range of competitive grant schemes. However, options aren't available to fund ongoing operational and maintenance costs for data platforms that operate outside Government departments or specific operational infrastructure funding schemes like NCRIS. Those platforms remain vulnerable to significant issues including security related risks, data quality issues, and ultimately platform availability. In short, the critical value of these datasets to sustainable management of the marine estate is not matched by funding to ensure their success.

This is a broader issue that is likely best addressed as part of a national conversation on research data infrastructure that falls into these funding gaps. However, it would be useful for Parks Australia and the NESP Marine and Coastal hub to introduce the issue into relevant forums including Environment Information Australia and the National Marine Science Committee.

Pressures and Activities

The standardised sum of national pressures and activities data product (Section 2.2) is derived from pressures and activities data sets that were (for the main part) collated over the five period 2011 to 2015 (Hayes *et al*., 2021). Pressures and activities often provide significant context to develop an understanding of causual agents effecting changes in ecosystem condition and are used for the development of cumulative impact layers. This product is produced by CSIRO and hosted on the IMAS Data Repository. Parks Australia has requested that these data sets and the derived data product be updated for the period 2018 to 2023.

Recommendations:

- Parks Australia develop a formal process to maintain and update the pressures and activities data sets and their derived products in the form of Standard Operating Procedures (SOPs). This should include functional requirements, the process for inclusion of new data sources and the updating of any existing data sources. This process should be managed by Parks Australia.
- Develop and publishing end-user documentation for the product (product specification, interpretation, etc).
- Consider data that might benefit from formally publishing vocabularies or adopting any existing pressures vocabularies (including those established by global research efforts).
- Update the data product for the years 2018 to 2023.
- Consider changes to the presentation of pressures and activities in Seamap Australia to improve use by Parks Australia staff.

GlobalArchive (General)

Information on the body-size of fish and height of sessile invertebrates have been recognised as potentially useful metrics to determine management effectiveness (Table 8). The only platform that currently supports the synthesising and sharing of quality controlled, stereo-image, body size and height measurements is GlobalArchive. GlobalArchive has built strong community support nationally and also includes many international contributions of data. Parks staff, however, are unable to interpret raw annotation data associated with the fish records and require derived products to effectively use the data for management purposes.

GlobalArchive is currently implementing tools based on the Open Geographical Consortium (OGC) (https://ogcapi.ogc.org/) web services standard to provide data supplemented with ISO19115 standard metadata to meet national and international guidelines for data accessibility and interoperability.

Availability of future funding, particularly ongoing operation funding, is unclear and development and maintenance of the platform is currently provided by one person. These represent significant risk factors for the platform moving forward.

Recommendations:

- The GlobalArchive steering committee should address issues highlighted in the existing business continuity plan and take steps to source long term funding for development and operational activities including adequate staffing to reduce risk.
- Coverage is biased towards shallower regions of MPAs. Parks Australia should consider producing reporting on data gaps for future survey planning, potentially differentiated for all included platforms.

- The deployments, fish and habitat data products derived from Global Archive should have a clear workflow to collect and respond to end-user feedback informing priorities and actions through the steering committee.

- The deployments, fish and habitat data products should include information on methods/SOPs used and include data to identify lineage (particularly in the synthesis product).

- The data that is open and accessible are predominantly from NESP or Parks Australia related projects. At the time of writing, 60% of data from deployments within AMPs are public (globally, 79% of deployments are private). Suggest making all data meeting quality criteria agreed by the steering committee public by default. If data is embargoed, consider an enforced embargo period. Consider introducing rules or guidelines for data embargoes (including length of embargo) endorsed by the steering committee. The steering committee should take steps to encourage users to make their data public.

- Implement clear licensing using a single Creative Commons license applied to all data contributions to avoid issues licensing aggregated and synthesised data products.

- Consider implementing a dashboard or similar tool on the website summarising available data (noting the Seamap Australia State of Knowledge tool already provides some information but may be too generic to meet all end user requirements).

- Implement metadata (and catalogue as needed) conforming to ISO19115 including a CSW endpoint for harvesting by upstream services such as the AODN and Research Data Australia. This needs to address lowest common elements at campaign and project level to ensure clear attribution of data.

- Implement a robust backup schedule and disaster recovery procedures (including advice from UWA IT services and/or NeCTAR as needed) and a maintenance schedule to ensure software upgrades and security patches are applied regularly.

- Complete the implementation of OGC compliant web services for observational data that meet end user needs. This should include both Web Mapping Services for deployments with information on underlying species data and Web Feature Services for deployment, observation, and synthesis data.

GlobalArchive (Deployments)

The quality of the deployments data in Global Archive is variable. Historically, deployment data formats have not always aligned with SOP requirements, although GlobalArchive now includes a tool (CheckEM) that requires consistent format and applies quality scores.

Recommendations:

- Include of one or more flags or scores to allow users to understand quality of the data in relation to a range of broadly identified use cases.

- Include information on methods used and inclusion probabilities. It is easy to identify deployments but difficult to assess which deployments have annotation data. Recommend adding data to the deployments product to understand what is available.

- Apply CheckEM to historic data in addition to new data and provide clearer guidance to data contributors on SOP requirements.

- A consistent application of spatial and temporal systems should be used across all deployments data enforced by quality assurance processes (noting WGA84 and UTC are most common in current data).

GlobalArchive (Fish and habitat)

Recommendations:

- In consultation with Parks Australia, develop additional resources (dashboard, maps, etc) providing interpretation of the data to address identified use cases.

- Consider products that strictly adheres to WoRMS taxonomy and remove non-compliant records from synthesis products and for OBIS extracts.

- Consider adopting and strictly enforcing relevant and broadly used vocabularies (e.g. AODN platform available in Research Vocabularies Australia).

- Include uncertainty/confidence indicators in both the fish and habitat data.

UMI/Squidle+ (General)

Understanding Marine Imagery (UMI) and Squidle+ are an IMOS sub-facility operated by Greybits Engineering through the Sydney Institute of Marine Science. The platform provides significant functionality to synthesise marine imagery data across groups and projects that might otherwise be disparate and is currently the largest repository of openly accessible survey, imagery and annotation data globally. It provides OGC compliant Web Mapping and Web Feature Services along with extensive API services for incorporation of data in other portals linking to source pages in the Squidle+ portal providing users with seamless access to comprehensive visualisations and data. Current ongoing operational funding for the UMI and Squidle+ platform is available to June 2025, but there is some uncertainty for sustained funding beyond this. Storage of the imagery data shown through UMI/Squidle+ is currently the responsibility of the data providers. While the science community benefit greatly from shared annotation data, Parks Australia don't have the knowledge required to interpret this level of data and would require higher level derived products to effectively use the data for management purposes.

Recommendations:

- The UMI steering committee consider addressing the funding issues as a high priority, including the formulation of a plan for long term preservation of the facility and its data. The plan should address future funding streams, particularly for operational activities, and more formalised and tracked development activities.

- Consider additional staff resources and/or greater integration with the AODN team.

- Complete the business continuity plan as a priority, including an IMOS commitment to long term preservation of the data should the platform be discontinued.

- Consider expanding metadata to project or campaign level for both deployment and annotation products to ensure attribution matches data contributor details. Consider creating DOIs for metadata records where appropriate.

- Establish a privacy policy for application users personal data.
- Consider automated release of embargoed data after a reasonable timeframe and in agreement with data contributors to ensure timely open access to data.

UMI/Squidle+ (Deployments)

Recommendations:

- Parks Australia should consider identifying data gaps in consultation with the science community to prioritise and forward plan new surveys and improve integration of this data product into their business processes (e.g. evaluation and licensing of proponent proposals, identification of benthic values).
- Deployment data should include information on methods and Standard Operating Procedures (SOPs) used, indicate if the study design was spatially balanced, stratified, or preferential sampling. Ideally this would also include inclusion probabilities. While this information is not available, metadata should include cautionary notes on use of the data.
- Consider locating and including data before 2007 and provide information on platform bias, such as towed video bias towards shallow waters verses AUV bias towards deep waters.
- Use relevant vocabulary services for primary copies of vocabularies, such as Research Vocabularies Australia for the Seamap Australia classification and for AODN vocabularies.
- Provide uniform spatial and temporal systems variables for all data and provide clarity on user and platform spatial and temporal systems used (both in data and metadata).
- Discuss with the NESP Marine and Coastal Hub Data Wrangler (southern node) potential improvements to Web Mapping and Web Feature Services delivering this product.

UMI/Squidle+ (Annotations)

Recommendations:

- Currently 60% of annotation data in AMPs are publicly available. Newly developed embargo guidelines will increase this amount. While this is a significant percentage, it will still be useful considering ways to engage with the UMI/Squidle+ community to increase the amount of data publicly available.
- Establish standards for imagery storage for integration with Squidle+ and include in relevant SOPs and consider revising relevant SOPs to ensure Squidle+ is adequately covered (relevant file formats, variables, access, etc).
- Consider possibilities to ensure long term storage of imagery, including a centralised repository or formal commitment from contributors to maintain storage facilities.
- Consider enhancing the platforms ability to include covariates and site inclusion probabilities as indicators of ensuring (or as a measure of) spatially balanced designs.

- Enforce Creative Commons Attribution licensing for all contributions to ensure uniform licensing of aggregated products.

- Integrate with Research Vocabularies Australia (RVA) and other vocabulary publishing services to maintain alignment with most recent versions (noting there may be technical issues with RVA affecting integration).

- Recommend establishing a Web Feature Service for annotation data.

- Consider whether filtering on data product outputs is required to protect Threatened, Endangered and Protected Species (TEPS).

- Recommend Parks Australia work with science users and Squidle+/UMI to establish derived products from annotation data to aid interpretation for their end use cases.

Seamap Australia (General)

Seamap Australia hosts the National Benthic Habitat Layer (NBHL) and all of the source benthic habitat data used in its development. It incorporates a State of Knowledge tool developed in consultation with Parks Australia and drawing on data provides by Geosiciences Australia, UMI/Squidle+ and GlobalArchive. It also integrates a large catalogue of marine data for synthesis with habitat data as a decision support tool. Seamap Australia is hosted by the Institute for Marine and Antarctic Studies (IMAS) at the University of Tasmania. Operationally, the Seamap Australia platform is funded and supported by IMAS and the University's Information Technology Services. Currently NBHL data product updates are underwritten by IMAS and the NESP Marine and Coastal Hub as and when resources allow. The current NESP hub finishes in 2027 and it will be critical to ensure updates continue to be funded.

Recommendations:

- The Seamap Australia Steering Committee should identify long term funding and establish a formal update process for the NBHL data product every 12 months and other data products listed below as needed. This should be carried out in consultation with major stakeholders, particularly those responsive to the value proposition the data products provided (e.g. DCCEEW) contributing significantly to programs such as State of the Environment and environmental accounting.

- Confirm or ensure that the IMAS/UTas business continuity plans include risk mitigation for critical on-line systems like Seamap Australia.

- Consider adopting a product list and associated responsibilities as part of the Seamap Australia Steering Committee Terms of Reference.

Seamap Australia (Validated National Benthic Habitat Layer)

The validated National Benthic Habitat Layer (NBHL) currently provides information on methods/platforms used and is available in a spatial context (for each habitat polygon). Currently the classification used to describe the habitats is published in Research Vocabularies Australia (RVA). New data is identified by interested parties making contact through the website.

Recommendations:

- Extend the functions of the Seamap Australia Steering Committee to include identification of high priority data gaps in the NBHL product. This should be led by the Parks Australia member on the Steering Committee.
- Consider making data product specifications in line with ISO19131 available for the NBHL and extending this to include an indication of whether Standard Operating Procedures have been followed for data collection and processing wherever possible (this information might not be available in source data).
- Investigate methods to provide spatial resolution, accuracy, and uncertainty of source products or some other form of related, detailed provenance (provenance may be available in the downloaded product, but not the interface).
- Investigate adding a depth filter to the State of Knowledge analytics for the NBHL that will help align it with the Natural Values Common Language.
- Investigate pro-active options for capturing more data not identified through current channels.
- Parks Australia to consider improving integration of the NBHL data products into their business processes (e.g. evaluation and licensing of proponent proposals, identification of benthic values).

Seamap Australia (Unvalidated National Benthic Habitat Layer)

The primary difference between the validated and unvalidated NBHL layer is the removal of data acceptance guidelines for considering input data. Development of the unvalidated NBHL hasn't been formally addressed and ongoing updates are not funded (current work has been underwritten by the NESP Hub). The development of this layer should include identifying funding, data discovery (to capture more source data), identified interval for updates, end user feedback, metadata, and download capability.

Recommendations:

- Consider establishing a formal approach (recognising this product can contribute to the AMP ecosystems data product) for developing this layer. If the product is to be developed and updated more formally, expert/peer review will be an important element and should be included (through the Seamap Australia steering committee).
- If the product is to be developed and updated more formally, provide better provenance and spatial resolution information for users.

Seamap Australia (State of Knowledge)

Recommendations:

- Establish or adopt a standard approach, format, vocabularies (platform, method), etc for all observation deployment data. Consider where this may have been documented elsewhere by other organisations (AODN, OBIS, iOOS, etc).
- Work with data providers (UMI/Squidle+, GlobalArchive, etc) to ensure all source data products are regularly updated maintaining currency of information.

- Consider providing more interpretive information on analytics produced including quality ratings, available data (vs embargoed data), etc.

- Recommend adding State of Knowledge for named features and user drawn areas (on the map).

- Parks Australia to consider moving away from current, static State of Knowledge reports to an expanded Seamap Australia State of Knowledge online resource which is continuously updated (currently on a weekly basis). Consideration will also need to be given to ensuring all platforms contributing to the State of Knowledge product achieving more stable, long term operational funding where needed.

- Ensure State of Knowledge data sources are adequately acknowledged.

- Parks Australia to use the State of Knowledge aggregated resources to carry out gap analysis and achieve efficiencies in data collection, processing and production of mapped ecosystem products such as the NBHL.

AMP Bathymetry

The AMP Bathymetry compilation dataset is currently hosted by the Institute of Marine and Antarctic Studies (IMAS). This data product was produced as part of a Parks Australia funded project and was not intended to be regularly produced by IMAS.

Recommendations:

- Parks Australia approach Geoscience Australia (GA) with a request to take over production of this data compilation product as it more closely matches their remit. Irrespective of whether IMAS or GA continue to provide this product, it will be necessary to identify sustainable funding and resources to do so.

- Recommend generating this product at the finest possible scale (or an agreed scale, e.g. 10m) using point cloud data.

- Consider options for providing provenance and source data information. Consider if the GA bathymetry product may be adequate for this purpose.

- Investigate applying a DOI to this product and any associated implications such as the current policy on using DOIs with data that is updated regularly. Alternatively, consider publishing versions over time (however this could be misinterpreted as change over time).

- Consider enabling functionality for configurable colour ramps (noting that Seamap Australia currently provides two versions) and 3D visualisation.

- Provide review and feedback mechanisms for the product under existing governance (e.g. AusSeabed Steering Committee).

- Parks Australia to ensure public facing availability of co-located backscatter intensity products are set as a requirement of MBES guidelines and for delivery of bathymetry products.

AMP Geomorphology

This data product is also hosted by IMAS. While IMAS have expertise and mission alignment with production of this data product, it may align better with Geoscience Australia or AusSeabed. Furthermore, production of this data product is currently on an ad hoc basis. Cost estimates for ongoing production have been given to Parks Australia.

Recommendations:

- Continue the conversation with Parks Australia to establish ongoing production (with consideration to the next recommendation) and establish further discussion between parties to resolve longer term hosting issues (could be actioned by IMAS or the Seamap Australia Steering Committee).
- Consider applying a DOI to the product with each new release being a new version.
- Provide Parks Australia with interpretation materials to help them understand if/how the geomorphometric features it represents are important for them to consider for management actions (e.g. permit advice).

Biologically Important Areas (BIAs)

The BIA data product, hosted by Department of Climate Change, Energy, the Environment and Water (DCCEEW), was originally developed in 2015 and reviewed in 2023. There are multiple metadata records for this product online, which may be due to domain name changes made by the Department over time. This data product does not appear to meet all of the requirements of the FAIR data principles

Recommendations:

- Work with DCCEEW and community of experts to meet regularly to review and produce an update introducing new areas or removing areas as appropriate.
- Consider updating the protocol (https://www.dcceew.gov.au/sites/default/files/documents/bia-protocol.pdf) to include governance structure and terms of reference, and then publishing the protocol as a Standard Operating Procedure (SOP) on Oceans Best Practice.
- Where applicable, consider using established and widely used species vocabularies (e.g. WoRMS) to enhance interoperability with other platforms.
- Recommend rationalising discoverable metadata to a single location and publishing a metadata record with a DOI for each updated product version.
- OGC web services used may not be fully compliant. Suggest discussing with the NESP Marine and Coastal Hub (southern node) Data Wrangler to resolve issues found.
- Apply the FAIR Data assessment tool and consider recommendations (https://ardc.edu.au/resource/fair-data-self-assessment-tool/).
- Consider mechanisms to regularly capture feedback from end users.

Key Ecological Features (KEFs)

The KEFs data product is used for AMP establishment but doesn't have documented specifications or a formal review process. There are two metadata records online for this product, each record contains different metadata. There is limited information of data assurance procedures taken other than identifying subjective assessment (documented in the metadata records). This data product does not appear to meet all of the requirements of the FAIR data principles.

Recommendations:

- Provide documentation on the information sources that underly the data product to assist in its interpretation, publishing this information if identified use cases warrant, and review these information sources for quality assurance.
- Rationalise the metadata and only retain one record and publish the metadata with a DOI for this and future versions.
- OGC web services used may not be fully compliant and all elements are styled with the same colour. Suggest discussing with the NESP Marine and Coastal Hub (southern node) Data Wrangler to resolve issues found.
- Apply the FAIR Data assessment tool and consider recommendations (https://ardc.edu.au/resource/fair-data-self-assessment-tool/).
- Consider mechanisms to regularly capture feedback from end users.

Key Natural Values (KNVs)

Key question three requires further articulation of key natural values building on the work undertaken in the NESP SS2/D7 project and the NESP Marine and Coastal Hub 1.3 project. Parks has identified the need to develop a formal process to finalise, update and validate the existing KNV layer. This work will be undertaken through NESP Marine and Coastal Hub project 4.20.

AusSeabed (General)

AusSeabed is the Geoscience Australia (GA) led national seabed mapping collaborative initiative between federal and state agencies, research organisations, and industry and provides a well-managed and delivered range of survey, bathymetry and other seabed-related data products. These products are core requirements for a range of research and other activities in managing the marine estate including MPAs. AusSeabed has strong governance in place and broad community support. AusSeabed is working to include metadata about historic surveys from the Australian Hydrographic Office (AHO) and industry in their coverage products.

Recommendations:

- Geoscience Australia provide more information on data product coverage and associated tools (including coverage for data that is not public). Enabling the portal’s text searches to include more fields than just the layer title, such as the metadata fields, would assist in this.

- Parks Australia to consider whether greater use of the AusSeabed Survey Coordination tool can help identify data gaps and survey priorities. AusSeabed priorities for adding new data are currently unclear. Suggest prioritisation protocols be made public and the publication forecast (https://www.ausseabed.gov.au/data/publication-schedule) be updated more frequently and with more accurate timelines.

- Some Parks Australia and users have noted that the AusSeabed platform can sometimes be "glitchy/buggy". Suggest that Parks Australia staff and other users report any specific bugs to ausseabed@ga.gov.au. Additionally suggest that Geoscience Australia performs a dedicated bug sweep and additional testing of the code and platform. Consider stress testing or similar to ensure the system can withstand large or complex downloads.

- Parks Australia consider collaborating with Geoscience Australia and AusSeabed to develop a park manager user guide within a context of answering the key questions identified in this project.

AusSeabed (Bathymetry)

The majority of publicly accessible bathymetry data collected in Australian waters is made available through the AusSeabed Marine Data Portal. There are, however, considerable amounts of industry and AHO data that are not currently publicly available.

Publicly available data is typically provided at the resolution approved for publication by the data owner. However, there are exceptions (noting this may be caused as part of data delivery by data custodians rather than the AusSeabed publication process).

There are both individual surveys and compilation products available through the AusSeabed Marine Data Portal, but data and compilation surfaces do not yet exist for all Australian waters including all waters in Marine Parks. Parks Australia has conveyed to AusSeabed (via a needs analysis conducted by AusSeabed) its desire for a single, national, multi-resolution, highest-available resolution, bathymetry data product. AusSeabed have indicated an intention to produce 30m and 100m regional AusBathyTopo grids that will help to meet this need. A 250m AusBathyTopo grid covers all of Australia. The AusSeabed Marine Data Portal publishes approximately 60 multi-resolution best-available bathymetry grids (often produced by CSIRO) and will publish more in the future.

Recommendations:

- Recommend that Parks Australia work with the organisations they've funded to do surveys to ensure the highest resolution bathymetry and backscatter data is published through the AusSeabed marine portal.

- Recommend that GA does more outreach to find new data sources and ensure datasets are processed and published through the AusSeabed Marine Data Portal at highest available resolution in a timely manner to meet Parks Australia's needs.

- Recommend that GA continues to produce regional and national scale integrated bathymetric surfaces and publishes them through the AusSeabed Marine Data Portal in a timely manner to meet Parks Australia's needs.

- Recommend adding clipping regions (e.g. in a SHP file) where data is extracted repeatedly to increase usability.

- Recommend that all Australian institutions working in the seabed environment investigate if there is missing data where they are the custodian and consider publishing where possible (e.g. industry data).

- Parks Australia to investigate whether Australian Hydrographic Office (AHO) data that are sufficiently low sensitivity can be published at the finest resolution available (currently 30m resolution) and negotiate with AHO for more streamlined limited access to sensitive data at finest resolution available (< 30m).

- There are currently at least four providers serving high resolution bathymetry data, such as the NSW Govt and the WA Department of Transport, but their data are only available through AusSeabed as part of their coverage layers, not as individual surveys. Recommend that GA works with these agencies to establish them as contributing hubs to the AusSeabed Marine Data Portal to facilitate centralised data discovery and access.

- Recommend that AusSeabed engages with broader Government and research efforts to establish an approach that makes industry data available under conditions that protect industry commercial confidentiality while ensuring data is available to streamline and reduce uncertainty for regulatory and research processes. Consider setting timed embargoes for industry that release industry data into the public domain when there is no longer concern for commercial confidentiality. Note these changes will require considerable governance, planning, consultation and infrastructure for workflow and data management automation.

AusSeabed (Geomorphology)

This data product refers to a new geomorphology map with full AMP coverage that will: (1) support park managers to answer KQ1; and (2) be used as input to the AMP Ecosystems product (with particular focus on priority features: seamounts, canyons, and reefs). A limited number of geomorphology maps are available within several marine parks (e.g. https://dx.doi.org/10.26186/147998), and Geoscience Australia are developing a national-scale (250 m bathymetry) morphology map that will delineate key features. Ongoing maintenance and periodic version updates of this national product should be undertaken to ensure that higher-resolution maps of these features are incorporated as they become available.

Recommendations:

- Determine the scope and resource requirements to undertake this work, and give consideration to establishing this as a formal data product that can readily contribute to the AMP ecosystems data product. This may require building additional expertise and resources in GA.

- Geoscience Australia are currently publishing their geomorphology mapping methods (Part 1 Morphology: Dove et al., 2020; Part 2 Geomorphology: Nanson et al., 2023) vocabularies in the GA Vocabulary Register, to ensure standards-based Application Programming Interface (API) access. These vocabularies should be maintained to match versioning of the methods.

- Geoscience Australia have published a Part 1 (Morphology) Esri toolbox, and have drafted a new Part 2 (Geomorphology) Esri toolbox, which are designed to support researchers to apply the geomorphology mapping methods to their own datasets. These tools, and their supporting documentation, should be maintained to support versioning of methods. Consideration should also be given to developing alternative tools for implementing the methods in non-proprietary software (e.g. QGIS).

- Parks Australia should collaborate with Geoscience Australia to co-design a dashboard-style interface that will help users to readily interpret the management implications of the national geomorphology map product.

AusSeabed (Backscatter)

As of April 2024, only seven backscatter datasets are publicly accessible through AusSeabed, five of which are from MH370 search activities. There are currently more than 160 multibeam bathymetry datasets provided for public access through AusSeabed. Many of these MBES surveys will have accompanying backscatter data but this is not routinely made accessible. It should be noted that due the complexities of backscatter data collection, data collected between surveys or even across large depth gradients is very unlikely to be comparable and not suitable to combine into any larger scale product without integration of normalisation approaches for harmonisation (Misiuk et al., 2020).

Recommendations:

- The AusSeabed Steering Committee should establish a working group to review the AusSeabed/NESP multibeam community guideline to ensure the specifications for backscatter data product meet FAIR data requirements.

- AusSeabed steering committee and GA should review backscatter products at a regular interval with Parks Australia consultation (suggest every two years).

- The AusSeabed Steering Committee should establish a working group to write an Australian community guideline or endorse an existing guidelines for collecting, processing, and publishing backscatter data.

- The AusSeabed Marine Data Portal should, where practicable, supplement survey and coverage metadata with information to indicate whether backscatter data is available and community guidelines have been followed.

- GA should publish the backlog of backscatter data currently stored by GA, reducing future publication delays and identifying priorities for new data.

- Parks Australia should work with GA to confirm backscatter layers in AusSeabed are sufficient to meet Parks Australia staff needs (for example to delineate sediment types or marine habitats) and develop guidelines for Parks Australia staff who use the data product, to identify the presence of reef habitats.

National Reef Monitoring Network (NRMN)

The National Reef Monitoring Network brings together shallow reef surveys conducted around Australia by contributors including Reef Life Survey, the Australian Temperate Reef Collaboration, and others into a single data resource. The database is hosted and managed by IMOS with the AODN delivering a range of data products to meet end user needs. As with all data streams managed by IMOS, it is well funded, governed, managed and meets the criteria for FAIR data.

Recommendations:

- Parks Australia, working with the science community, consider verifying data products provided are fit for purpose in the context of MPA management (see https://imos.org.au/facility/autonomous-underwater-vehicles/national-reef-monitoring-network for a list of data products. If needed, consult with the IMOS office for required changes or additions.

Benthic cover

Recommendations:

- Parks Australia, working with the science community and the IMOS office, encourage users to annotate more imagery in Squidle+ and increase publicly available annotations for surveys in AMPs.

Fish and invertebrate abundance and biomass

There are no recommendations for this product.

ReefCloud

ReefCloud is a platform that has achieved significant success using Machine Learning and AI for detection of coral species and coral health. It provides a dashboard and mapped access to information and has the ability to share data and support collaboration within the ReefCloud environment. It is well managed and supported by AIMS. Data held in ReefCloud is only accessible by other users when a ReefCloud user specifically grants access to their data collections.

- Consider improving data discovery by generating ISO19115 metadata records for all data collections held within ReefCloud.

- Consider establishing public web services, following OGC standards, to make public data machine readable.

- Encourage researchers using ReefCloud to make their data openly available in a timely manner.

Ocean Biodiversity Information System (OBIS)

OBIS is the comprehensive, UNESCO led biodiversity and biogeographic data portal operating at global scale. CSIRO operate the OBIS Australia node for data contributions from many organisations nationwide.

Recommendations:

- Maintenance of taxonomy in Fishbase is inconsistently maintained which in turn affects WoRMS taxonomy for fish species. It is not a responsibility of OBIS to maintain taxa on Fishbase. However, it is used as a source for WoRMS fish taxonomy affecting support and use of its data and users have reported difficulty achieving timely additions or changes. Suggest OBIS or related parties consider and take steps to ensure Fishbase (or at least WoRMS) taxonomy is regularly maintained. For Australia, consider whether changes may be addressed in a more timely manner if channelled through OBIS Australia.

- While OBIS has support for absence and abundance data, the prevalence of that data is limited and source data products may more readily provide additional and important information for Parks Australia use cases. We recommend that Parks Australia consider this if sourcing data from OBIS.

We recognise, however, that OBIS is governed under UNESCO and is not an Australian managed service, hence this recommendation is provided as feedback through OBIS Australia (hosted by CSIRO).

6.4 Data Accessibility

Over the duration of the project a number of non-public, but potentially useful, data sets were identified and discussed. The project recognises, however, that there are many more private, potentially relevant, data sets that it is unaware of, and the individual data sets identified by the project are best characterised as examples of the following broader categories of historical and/or unpublished data (Table 10):

- data from science activities in AMPs that were commissioned and funded by Parks Australia, DCCEEW, or NESP/NERP and conducted under NESP/NERP
- data from science activities in AMPs that were commissioned and funded by Parks Australia, DCCEEW or other third parties, but not conducted under NESP/NERP
- data collected in AMPs by industry, commercial consultants and research institutions following authorisations (such as permits and licences) issued by Parks Australia
- data collected under other regulatory regimes (such as fisheries management) that is relevant to AMPs.

We recommend that Parks Australia consider ways to make these types of data sets publicly available (see also recommendations in Section 6.3). It should be noted, however, that many of the historical private data set may not be useful for establishing baselines and conducting management evaluations within AMPs, due to preferential sampling designs. In GlobalArchive, all metadata including spatial location of samples is publicly available, and this information can be used to assess the representativeness of the sampling design, and hence usefulness of the data sets.

Table 10 Examples of data products identified during the NESP 2.3 project that are not currently accessible to Parks Australia

Category	AMP specific examples
Data from science activities in AMPs that are commissioned and funded by Parks Australia, DCCEEW, or NESP/NERP and conducted under NESP/NERP	The initial characterisation of the physical, oceanographic and biological character of Arafura Marine Park undertaken by the NESP Marine Biodiversity Hub in 2020. Metadata record provides links to data (after submission of email and rationale). Image records from towed video available, but only BRUVs locations provided. Tasman Fracture rock lobster data from the 2021 NESP funded survey sits in Cray base data repository. NESP researchers have copies of this data, but it is not generally available.
Data from science activities in AMPs that was commissioned and funded by Parks Australia, DCCEEW or other third parties, but not conducted under NESP/NERP	A significant portion of benthic habitat and fish annotations, derived from imagery (Squidle+) and baited video (GlobalArchive), are not made publicly accessible by the original data providers. As of April 2024, 63% of benthic annotations from Squidle+ within AMPs are public, with 53% of annotations available across all waters (inside and outside AMPs). For fish assemblage and body size data delivered by GlobalArchive, 57% of scoring data from deployments within AMPs is publicly accessible, with just 19% available for all waters. Regular Holothurian surveys in the Ashmore/Cartier reef systems and mono-video BRUVs surveys conducted in the Coral Sea by AIMS.
Data collected in AMPs by industry, commercial consultants and research institutions following authorisations (such as permits and licences) issued by Parks Australia	Survey data in the Dampier Marine Park in 2020 and 2021 permitted by the Director of National Parks for the purposes of evaluating the impacts of port-related dredging. This survey was conducted prior to the Parks Australia requirements that benthic habitat information is provided to Seamap Australia in a manner that is consistent with Seamap Australia Data Acceptance guidelines.
Data collected under other regulatory regimes (such as fisheries management) that is relevant to AMPs.	Western Australia Fisheries have offered to provide Western Rock Lobster fishery survey data (three years) collected in the Abrolhos Marine Park to GlobalArchive but do not want this data to be made public. Longitudinal observations of rock lobster abundance from Tasmania, Victoria and South Australia are housed in the "Cray base" data repository which is not public facing. The project tried unsuccessfully to obtain Tasman Fracture Rock Lobster data from a 2015 rock lobster survey data into a publicly accessible form.

7. References

Aitchison, J. (1982). The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, **44**(2): 139–177. http://www.jstor.org/stable/2345821.

Albert, A. and Anderson, J. A. (1984). On the Existence of Maximum Likelihood Estimates in Logistic Regression Models. *Biometrika*, **71**(1):1-10. https://doi.org/10.2307/2336390.

Allard, H., Ayling, A. M. and Shears, N. T. (2022). Long-term changes in reef fish assemblages after 40 years of no-take marine reserve protection. *Biological Conservation*, **265**:109405. https://ui.adsabs.harvard.edu/abs/2022BCons.26509405A.

Althaus F., Hill N., Ferrari R., Edwards L., Przeslawski R., Schönberg C. H. L., et al. (2015). A Standardised Vocabulary for Identifying Benthic Biota and Substrata from Underwater Imagery: The CATAMI Classification Scheme. *PLoS ONE*, **10**(10): e0141039. https://doi.org/10.1371/journal.pone.0141039.

Anderson, M. J. (2008). Animal-sediment relationships re-visited: Characterising species' distributions along an environmental gradient using canonical analysis and quantile regression splines. *Journal of Experimental Marine Biology and Ecology*, **366(**1–2): 6-27. https://doi.org/10.1016/j.jembe.2008.07.006.

Anderson, M. J. and Thompson, A.A. (2004). Multivariate control charts for ecological and environmental monitoring. *Ecological Applications*, **14**:1921-1935. https://doi.org/10.1890/03-5379.

Aubry, P., Francesiaz, C. and Guillemain, M. (2024). On the impact of preferential sampling on ecological status and trend assessment. *Ecological Modelling*, **492.** https://doi.org/10.1016/j.ecolmodel.2024.110707.

Babcock, R. C., Shears, N. T., Alcala, A. C., Barrett, N. S., Edgar, G. J., Lafferty, K.D., McClanahan, T. R. and Russ, G. R. (2010) Decadal trends in marine reserves reveal differential rates of change in direct and indirect effects. *Proceedings of the National Academies of Science*, **107**(43):18256-61. DOI: 10.1073/pnas.0908012107.

Barrett, N. S., Edgar, G. J., Buxton, C. D. and Haddon, M. (2007). Changes in fish assemblages following 10 years of protection in Tasmanian marine protected areas. *Journal of Experimental Marine Biology and Ecology*, **345**: 141-157. https://doi.org/10.1016/j.jembe.2007.02.007.

Barrett, N. S., Buxton, C. D. and Edgar, G. J. (2009). Changes in invertebrate and macroalgal populations in Tasmanian marine reserves in the decade following protection. *Journal of Experimental Marine Biology and Ecology*, **370**: 104-119. https://doi.org/10.1016/j.jembe.2008.12.005.

Beaman, R. (2023). *AusBathyTopo 250m (Australia) 2023 Grid - A High-resolution Depth Model for Australia (20230004C)* [Dataset]. Geoscience Australia, Canberra. https://doi.org/10.26186/148758.

Beaman, R. J. (2017). *High-resolution depth model for the Great Barrier Reef - 30 m*. [Dataset]. Geoscience Australia, Canberra. http://dx.doi.org/10.4225/25/5a207b36022d2.

Buck, J. J. H., Bainbridge, S. J., Burger, E. F., Kraberg, A. C., Casari, M., Casey, K. S., Darroch, L., Rio, J. D., Metfies, K., Delory, E., Fischer, P. F., Gardner, T., Heffernan, R., Jirka, S., Kokkinaki, A., Loebl, M., Buttigieg, P. L., Pearlman, J. S. and Schewe, I. (2019). Ocean Data Product Integration Through Innovation-The Next Level of Data Interoperability. *Frontiers in Marine Science*, **6**:32. https://doi.org/10.3389/fmars.2019.00032.

Bryars, S., Page, B., Waycott, M., Brock, D. and Wright, A., (2017). *South Australian Marine Parks Monitoring, Evaluation and Reporting Plan*, DEWNR Technical report 2017/05, Government of South Australia, through Department of Environment, Water and Natural Resources, Adelaide. https://data.environment.sa.gov.au/Content/Publications/DEWNR-TR-2017-05.pdf

Chapman, A. D. (2005). *Principles of Data Quality, Version 1.0*. Report for the Global Biodiversity Information Facility, Copenhagen. https://www.gbif.org/document/80509/principles-of-data-quality.

Claudet, J., Osenberg, C. W., Domenici, P., Badalamenti, F., Milazzo, M., Falcón, J. M., Bertocci, I., Benedetti-Cecchi, L., García-Charton, J., Goñi, R., Borg, J. A., Forcada, A., de Lucia, G. A., Pérez-Ruzafa, Á., Afonso, P., Brito, A., Guala, I., Diréach, L. L., Sanchez-Jerez, P., Somerfield, P. J. and Planes, S. (2010). Marine reserves: Fish life history and ecological traits matter. *Ecological Applications*, **20**: 830-839. https://doi.org/10.1890/08-2131.1.

Conn, P. B., Thorson, J. T. and Johnson, D. S. (2017). Confronting preferential sampling when analysing population distributions: diagnosis and model-based triage. *Methods in Ecology and Evolution*, **8**:1535-1546. https://doi.org/10.1111/2041-210X.12803.

Cressie, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley and Sons Inc. https://onlinelibrary.wiley.com/doi/book/10.1002/9781119115151.

Cressie, N. and Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **70**: 209-226. https://doi.org/10.1111/j.1467-9868.2007.00633.x

Cresswell, A.K., Langlois, T. J., Wilson, S. K. Claudet, J., Thomson, D. P., Renton, M., Fulton, C. J., Fisher, R., Vanderklift, M. A., Babcock, R. C., Stuart-Smith, R. D., Haywood, M. D. E., Depczynski, M., Westera, M., Ayling, A. M., Fitzpatrick, B., Halford, A. R., McLean, D. L., Pillans, R. D., Cheal, A. J., Tinkler, P., Edgar, G. J., Graham, N. A. J., Harvey E. S. and Holmes T. H. (2019). Disentangling the response of fishes to recreational fishing over 30 years within a fringing coral reef reserve network. *Biological Conservation*, **237**: 514-524. https://doi.org/10.1016/j.biocon.2019.06.023.

Dambacher, J. M. and Anthony, K. R. N (2019). *Models and indicators for RIMReP DPSIR framework*, Great Barrier Reef Marine Park Authority, Townsville.

Datta, A., Banerjee, S., Finley, A. O. and Gelfand, A. E. (2016). Hierarchical Nearest-Neighbor Gaussian Process Models for Large Geostatistical Datasets. *Journal of the American Statistical Association*, **111**(514):800–812. https://doi.org/10.1080/01621459.2015.1044091.

DiFiore, B., Queenborough, S., Madin, E., Paul, V., Decker, M., and Stier, A. (2019). Grazing halos on coral reefs: predation risk, herbivore density, and habitat size influence grazing patterns that are visible from space. *Marine Ecology Progress Series*, *627*, 71–81. https://doi.org/10.3354/meps13074

Dove, D., Nanson, R., Bjarnadóttir, L. R., Guinan, J., Gafeira, J., Post, A., Dolan, M. F. J., Stewart, H., Arosio, R. and Scott, G. (2020). A two-part seabed geomorphology classification scheme (v.2); Part 1: morphology features glossary. Zenodo. https://doi.org/10.5281/zenodo.4075248.

Dumelle, M., Higham, M., Ver Hoef, J. M., Olsen, A. R., and Madsen, L. (2022). A comparison of design-based and model-based approaches for finite population spatial sampling and inference. *Methods in Ecology and Evolution*, **13**: 2018–2029. https://doi.org/10.1111/2041-210X.13919.

Dunstan, P. K., Woolley, S. N. C., Monk, J., Barrett, N., Hayes, K. R., Foster, S., Howe, S. A., Logan, D., Samson, C. R. and Francis, S. O. (2023). *Designing a targeted monitoring program to support evidence-based management of Australian Marine Parks: National Implementation*. Report to the National Environmental Science Program. CSIRO. https://www.nespmarinecoastal.edu.au/publication/designing-a-targeted-monitoring-program-to-support-evidence-based-management-of-australian-marine-parks/.

Edmunds, M. (2017). *Victorian Subtidal Reef Monitoring Program: Popes Eye, Port Phillip Heads Marine National Park*, January 2015. Parks Victoria Technical Report Series No. 103. Parks Victoria, Melbourne.

Fairweather, P. G. (1991). Statistical power and design requirements for environmental monitoring. *Marine and Freshwater Research*, **42**: 555-567. https://doi.org/10.1071/MF9910555.

Flukes, E. (2023). Seamounts and raised seabed features of the Australian margin. Institute for Marine and Antarctic Studies (IMAS), University of Tasmania (UTAS). Data accessed from https://metadata.imas.utas.edu.au/geonetwork/srv/eng/catalog.search#/metadata/dc4c761f-eb4d-4fe4-b071-e110bbe7c8f3 on 22/04/2024.

Foster, S. D. (2021). MBHdesign: An R-package for efficient spatial survey designs. Methods in Ecology and Evolution, **12**:415–420. https://doi.org/10.1111/2041-210X.13535.

Foster, S. D., Hosack, G. R., Hill, N. A., Barrett, N. S. and Lucieer, V. L. (2014). Choosing between strategies for designing surveys: autonomous underwater vehicles. *Methods in Ecology and Evolution*, **5**:287-297. https://doi.org/10.1111/2041-210X.12156.

Foster, S. D., Hosack, G. R., Lawrence, E., Przeslawski, R., Hedge, P., Caley, M. J., Barrett, N. S., Williams, A., Li, J., Lynch, T., Dambacher, J. M., Sweatman, H. P. A. and Hayes, K. R. (2017). Spatially balanced designs that incorporate legacy sites. Methods in Ecology and Evolution, **8**:1433-1442. https://doi.org/10.1111/2041-210X.12782.

Foster, S. D., Monk, J., Lawrence, E., Hayes, K. R., Hosack, G. R., Langlois, T., Hooper, G. and Przeslawski, R. (2024a). Statistical considerations for monitoring and sampling. In Przeslawski, R. and Foster, S. (Eds). *Field Manuals for Marine Sampling to Monitor Australian Waters, Version* 3. National Environmental Science Program (NESP). https://survey-design-field-manual.github.io/.

Foster, S.D., Peel, D., Hosack, G.R., Hoskins, A., Mitchell, D.J., Proft, K., Yang, W.-H., Uribe-Rivera, D.E. and Froese, J.G. (2024b), 'RISDM': species distribution modelling from multiple data sources in R. Ecography, 2024: e06964. https://doi.org/10.1111/ecog.06964.

Gefland, A. E. (1995). Model determination using sampling-based methods. Chapter in Gilks, W. R., Richardson, S. and Spiegelhalter, D. (Eds), Markov Chain Monte Carlo in Practice. Chapman and Hall/CRC. https://www.taylorfrancis.com/books/mono/10.1201/b14835/markov-chain-monte-carlo-practice?refId=0202ccdb-f547-40f1-b42a-4b3e8ce6438b&context=ubx.

Gelman, A. and Hill, J. (2007). Sample size and power calculations. Chapter 20 in Gelman, A. and Hill, J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge, UK.

Green, R. H. (1989). Power analysis and practical strategies for environmental monitoring, *Environmental Research*, **50**(1): 195-205. https://doi.org/10.1016/S0013-9351(89)80058-1.

Gross, J. E. (2003). *Developing conceptual models for monitoring programmes*. NPS Inventory and Monitoring Program, National Park Service, Washington DC. https://www.researchgate.net/publication/254513558_Developing_Conceptual_Models_for_Monitoring_Programs.

Hayes, K. R., Dambacher, J. M., Hosack, G. R., Bax, N. J., Dunstan, P., Fulton, E. A., Thompson, P. A., Hartog, J. R., Hobday, A. J., Bradford, R., Foster, S., Hedge, P., Smith, D. C. and Marshall, D. (2015). Identifying and evaluating biological indicators and essential variables for marine ecosystems. *Ecological Indicators*, **57**:409-419. https://doi.org/10.1016/j.ecolind.2015.05.006.

Hayes, K. R., Hosack, G. T., Lawrence, E., Hedge, P., Barret, N., Przeslawski, R., Caley, M. J. and Foster, S. D. (2019). Designing monitoring programmes for Marine Protected Areas within an Evidence Based Decision Making paradigm. *Frontiers in Marine Science*, **6**:746. https://doi.org/10.3389/fmars.2019.00746.

Hayes, K. R., P. Dunstan, S. Woolley, N. Barrett, S. A. Howe, C. R. Samson, R. Bowling, et al. (2021). *Designing a Targeted Monitoring Program to Support Evidence Based Management of Australian Marine Parks: A Pilot on the South-East Marine Parks Network*. Report to Parks Australia; the National Environmental Science Program, Marine Biodiversity Hub. Parks Australia, University of Tasmanian; CSIRO, Hobart, Australia. https://www.nespmarine.edu.au/document/designing-targeted-monitoring-program-support-evidence-based-management-australian-marine.

Heap, A., Harris, P., Sbaffi, L., Passlow, V. and Fellows, M. (2006). *Geomorphic Features 2006.* Geoscience Australia, Canberra. https://pid.geoscience.gov.au/dataset/ga/69797.

Hilton, M., Walsh, J. C., Liddell, E. and Cook C.N. (2022). Lessons from other disciplines for setting thresholds for conservation management. *Conservation Biology*, **36**: e13865. https://doi.org/10.1111/cobi.13865.

Hilton, M. and Cook, C.N. (2022). Defining performance thresholds for effective management of biodiversity within protected areas. *Conservation Biology*. **36**, e13963. https://doi.org/10.1111/cobi.13963.

Hoffman, M. D, Gelman, A. et al. (2014). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, **15** (1):1593–623.

Hosack, G. R., Hayes, K. R. and Barry, S. C. (2017). Prior elicitation for Bayesian Generalised Linear Models with application to Risk Control Option assessment. *Reliability Engineering and System Safety*, **167**:351-361. https://doi.org/10.1016/j.ress.2017.06.011.

Howe, S., Sams, M., Pocklington, J. and Rodrigue, M. (in press). *Parks Victoria's marine monitoring, evaluation and reporting program. Celebrating the 20th anniversary of the Marine National Parks and Sanctuaries*. Parks Victoria Technical Series. Parks Victoria, Melbourne, Australia.

Huang, Z. and Nichol, S. (2015). *National Submarine Canyons of Australia*. Geoscience Australia, Canberra. https://pid.geoscience.gov.au/dataset/ga/77010.

Ierodiaconou, D., Wines, S., Carnell, P., Tinkler, P., Allan, B., Carey, J., Young, M., Howe, S. and Pocklington, J. (2020). *An enhanced Signs of Healthy Parks monitoring program for Victoria's Marine National Parks and Marine Sanctuaries: Point Addis Marine National Park*. Parks Victoria Technical Series 114. https://www.parks.vic.gov.au/-/media/project/pv/main/parks/documents/get-into-nature/conservation-and-science/technical-series/pvts-114-an-enhanced-signs-of-healthy-parks-monitoring-program.pdf?rev=00350563b7b34e10874c0b50a4e8f953.

Kearney, T. D., Smith, L. M. and Rutherford, C. (2019). *Data Product Quality Best Practices: a white paper from the observatory best practices/lessons learned series*. Washington, DC, Consortium for Ocean Leadership, 62pp. DOI: http://dx.doi.org/10.25607/OBP-507.

Keough, M. J., Ross, D. J. and Knott, N. A. (2007). *Ecological performance measures for Victorian Marine Protected Areas: Review of existing biological sampling program*. Parks Victoria Technical Series No.51. Parks Victoria, Melbourne. https://www.parks.vic.gov.au/-/media/project/pv/main/parks/documents/get-into-nature/conservation-and-science/technical-series/pvts-051-ecological-performance-measures-for-victorian-marine-protected-areas.pdf?rev=d8afd783a9ee47aba1956085db3b9486.

Kermorvant, C., D'Amico, F., Bru, N., Caill-Milly, N. and Robertson, B. (2019). Spatially balanced sampling designs for environmental surveys. Environmental Monitoring and Assessment, **191**: 524. https://doi.org/10.1007/s10661-019-7666-y.

Kuhnert, P. M., Martin, T. G. and Griffiths, S. P. (2010). A guide to eliciting and using expert knowledge in Bayesian ecological models. *Ecology Letters*, **13**:900-914. https://doi.org/10.1111/j.1461-0248.2010.01477.x

Langlois T., Monk J., Giraldo A., Gibbons B., Adams K., Barrett N., Sims H., Leplastrier H., Boyd, M., Siwabessy J. and Nichol S. (2021a). *South-West Corner Marine Park Post Survey Report.* Report to the National Environmental Science Program, Marine Biodiversity Hub. The University of Western Australia. https://www.nespmarine.edu.au/document/south-west-corner-marine-park-post-survey-report.

Langlois, T.J., Wakefield, C. B., Harvey, E. S., Boddington, D. K. and Newman, S. J. (2021b). Does the benthic biota or fish assemblage within a large, targeted fisheries closure differ to surrounding areas after 12 years of protection in tropical northwestern Australia? *Marine Environmental Research*, **170**:105403. https://doi.org/10.1016/j.marenvres.2021.105403.

Lawrey, E., & Hammerton, M. (2022). *Coral Sea features satellite imagery and raw depth contours (Sentinel 2 and Landsat 8) 2015 – 2021 (AIMS)* [Data set]. eAtlas. https://doi.org/10.26274/NH77-ZW79.

Lawrey, E. (2024a). *Estimating benthic reflectance of deep coral atoll lagoons from satellite imagery and bathymetry - Analysis code and case studies (NESP MaC 2.3, AIMS)* [Data set]. eAtlas. https://doi.org/10.26274/s2a8-nw72

Lawrey, E. (2024b). *Coral Sea Oceanic Vegetation (NESP MaC 2.3, AIMS)* [Data set]. eAtlas. https://doi.org/10.26274/709g-aq12.

Lawrey, E. (2024c). *Coral Sea Features - Dataset collection - Reefs, Cays, Atoll reef platforms, Depth contours (AIMS)*. [Data set]. eAtlas. https://doi.org/10.26274/pgjp-8462.

Leek, J. T. and Peng, R. D. (2015). What is the question? *Science*, **347**:1314–1315. DOI: 10.1126/science.aaa6146.

Lindgren, F., Håvard R. and Lindström, J. (2011). An Explicit Link Between Gaussian Fields and Gaussian Markov Random Fields: The Stochastic Partial Differential Equation Approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **73**(4):423–98. https://doi.org/10.1111/j.1467-9868.2011.00777.x

Mayfield, H., Rhodes, J., Evans, M. and Maron, M. (2020). *Guidelines for estimating and evaluating species' response to management*. NSW Department of Planning, Industry and Environment, Australia. https://www.environment.nsw.gov.au/-/media/OEH/Corporate-Site/Documents/Animals-and-plants/Threatened-species/estimating-evaluating-species-response-management-guidelines-200415.pdf.

Michener, W. K. Quality Assurance and Quality Control (2003). Chapter 4 in Recknagel, F. and Michener, W. K. (Eds), *Ecological Informatics: Data Management and Knowledge Discovery*. Springer International Publishing, Switzerland.

Misiuk, B., Brown, C. J., Robert, K. and Lacharité, M. (2020). Harmonizing Multi-Source Sonar Backscatter Datasets for Seabed Mapping Using Bulk Shift Approaches. Remote Sensing, **12**(4):601. https://doi.org/10.3390/rs12040601.

Nanson, R., Arosio, R., Gafeira, J., McNeil, M., Dove, D., Bjarnadóttir, L., Dolan, M.F.J., Guinan, J., Post, A., Webb, J. and S. Nichol. (2022). A two-part seabed geomorphology classification scheme. Part 2: Geomorphology classification framework and glossary - Version 1.0. Published 14th April 2023. DOI: 10.5281/zenodo.7804019.

Nature United (2023). *Best Practices and Procedures for Operationalizing Marine Protected Area Network Monitoring: Synthesis of Global Insights and Recommendations for British Columbia's Northern Shelf Bioregion*. Report prepared by ReConnect Consulting and ESSA Technologies. 53 pp.

Neal, R. M. (2011). MCMC Using Hamiltonian Dynamics. Chapter in Brooks, S. Gelman, A., Jones, G. L. and Meng, X-L. (Eds) *Handbook of Markov Chain Monte Carlo*, Chapman Hall/CRC Press. https://www.taylorfrancis.com/chapters/edit/10.1201/b10905-6/mcmc-using-hamiltonian-dynamics-radford-neal.

Noble-James, T., Jesus, A. and McBreen, F. (2018). *Monitoring guidance for marine benthic habitats (Revised 2018)*. JNCC Report No. 598. JNCC, Peterborough, England. https://hub.jncc.gov.uk/assets/9ade4be8-63dd-4bbc-afd0-aefe71af0849.

Noble-James, T., Bullimore, R., McBreen, F., O'Connor, J., Highfield, J., McCabe, C., Archer-Rand, S., Downie, A. L., Hawes, J. and Mitchell, P (2023). Monitoring benthic habitats in English Marine Protected Areas: Lessons learned, challenges and future directions. *Marine Policy*, **157**. https://doi.org/10.1016/j.marpol.2023.105852.

Ogden, J.C., Davis, S.M., Jacobs, K.J., Barnes, T. and Fling, H. E. (2005). The use of conceptual ecological models to guide ecosystem restoration in South Florida. *Wetlands* **25**: 795–809. https://doi.org/10.1672/0277-5212(2005)025[0795:TUOCEM]2.0.CO;2.

Peng, G., Lacagnina, C., Downs, R. R., Ganske, A., Ramapriyan, H. K., Ivánová, I., Wyborn, L., Jones, D., Bastin, L., Shie, C-L and Moroni, D. F. (2022). Global Community Guidelines for Documenting, Sharing, and Reusing Quality Information of Individual Digital Datasets. *Data Science Journal*, **21**:8, pp. 1–20. DOI: https://doi.org/10.5334/dsj-2022-008.

Perkins, N. R., Foster, S. D., Hill, N. A. and Barrett, N. S. (2016). Image subsampling and point scoring approaches for large-scale marine benthic monitoring programs. *Estuarine, Coastal and Shelf Science*, **176**:36-46. https://doi.org/10.1016/j.ecss.2016.04.005.

Pocklington, J. B., Carey, J. M., Murshed, M. D. T and Howe, S. A. J (2012). *Conceptual Models for Victorian Ecosystems: Marine and Estuarine Ecosystems*, Parks Victoria Technical Series No. 66, Parks Victoria, Melbourne. https://www.parks.vic.gov.au/-/media/project/pv/main/parks/documents/get-into-nature/conservation-and-science/technical-series/pvts-066-ecosystem-conceptual-models-for-victorian-ecosystems.pdf?rev=d6cdd31c286b48b3981856147c1aec71.

Przeslawski, R. and Foster, S. (Eds.) (2018). *Field Manuals for Marine Sampling to Monitor Australian Waters, Version 1.* Report to the National Environmental Science Programme, Marine Biodiversity Hub. 212 pp. http://dx.doi.org/10.11636/9781925297669.

Przeslawski, R., Foster, S., Monk, J., Barrett, N., Bouchet, P., Carroll, A., Langlois T., Lucieer, V. Williams, J. and Bax, N. (2019). A Suite of Field Manuals for Marine Sampling to Monitor Australian Waters. *Frontiers in Marine Science*, **6**. https://www.frontiersin.org/articles/10.3389/fmars.2019.00177.

Przeslawski, R., Barrett, N., Carroll, A., Foster, S., Brooke, G., Jordan, A., Monk, J., Langlois T., Lara-Lopez, A., Pearlman, J., Picard, K., Pini-Fitzsimmons, J., van Ruth, P. and Williams J. (2023). Developing an ocean best practice: A case study of marine sampling practices from Australia. *Frontiers in Marine Science*, **10**. https://www.frontiersin.org/articles/10.3389/fmars.2023.1173075.

Ripley, B., Venables, W. and Ripley, M. B. (2016). *Package Nnet*. R Package Version 7 (3-12): 700.

Smit, K. P., Bernard, A. T. F., Lombard, A. T. and Sink, K. J. (2021). Assessing marine ecosystem condition: A review to support indicator choice and framework development. *Ecological Indicators,* **121**. https://doi.org/10.1016/j.ecolind.2020.107148.

Spencer, C., Langlois, T., Gibbons, B., Radford, B., Hickey, S., Mastrantonis, S., Navarro, M. and Griffin, K. (in review), *Establishing Standard Operating Procedures (including survey design) for the long-term monitoring of recognised and emerging natural and socio-economic values across the Geographe Marine Park, including a synthesis of available socio-economic natural value data*. Report to Marine Parks Branch, Parks Australia. University of Western Australia.

Staudigel, H., Koppers, A., Lavelle, J., Pitcher, T. and Shank, T. (2010). *Defining the Word "Seamount"*. *Oceanography*, **23**:20-21. http://dx.doi.org/10.5670/oceanog.2010.85.

Stuart-Smith, R., Edgar, G., Barrett, N. et al. (2015). Thermal biases and vulnerability to warming in the world's marine fauna. *Nature*, **528**: 88–92 https://doi.org/10.1038/nature16144.

The Nature Conservancy (TNC) (2007). *Conservation Action Planning Handbook: Developing Strategies, Taking Action and Measuring Success at Any Scale*. The Nature Conservancy, Arlington, VA, USA. http://conserveonline.org/workspaces/cbdgateway/cap/practices.

Tol, S. J., Coles, R. G., Shepherd, L., Scott, A., Hoffmann, L., Leeson, P., Ekins, M., Clarke, M., Grech, A., Rasheed, M. A. and York, P. (2023). *Benthic Habitat Mapping in the Central Coral Sea Marine Park: Preliminary voyage report for December 2022 survey*. JCU Publication, Centre for Tropical Water & Aquatic Ecosystem Research Publication 23/16, Cairns, 86 pp. https://www.tropwater.com/wp-content/uploads/2023/07/23-16-Benthic-habitat-mapping-in-the-cnetral-coral-sea-maine-park.pdf

Twiggs, E. (2023). *Kimberley Region and WA Reefs Satellite-Derived Bathymetry Acquisition (20210024S)* [Dataset]. Geoscience Australia, Canberra. https://dx.doi.org/10.26186/148669

The Western Australian Biodiversity Science Institute and Western Australian Marine Science Institution (WABSI and WAMSI) (2023). *Capabilities for a national supply chain of environmental information SAFE 2.0*. The Western Australian Biodiversity Science Institute, Perth, Australia. https://ardc.edu.au/resource/shared-analytic-framework-for-the-environment-safe-2-0/.

Whiteway, T. G. (2009). *Australian Bathymetry and Topography Grid*. Geoscience Australia, Canberra.

Woolley, S. N. C. and Foster, S. D. (2023). The ppmData R-package for setting up spatial point process models. Journal of Open Source Software, 8(83), 4771, https://doi.org/10.21105/joss.04771.

Yesson, C. Clark, M. R., Taylor, M. L. and Rogers, A. D. (2011). The global distribution of seamounts based on 30 arc seconds bathymetry data. *Deep Sea Research Part I: Oceanographic Research Papers*, **58**(4): 442-453. https://doi.org/10.1016/j.dsr.2011.02.004.

Young, M., Porskamp, P., Murfitt, S., Wines, S., Tinkler, P., Bursic, J., Allan, B., Howe, S., Whitmarsh, S., Pocklington, J. and Ierodiaconou, D. (2022). *Baseline habitat mapping and enhanced monitoring trials of subtidal and intertidal reef habitats in Victoria's marine national parks and sanctuaries*. Parks Victoria Technical Series 116.

Young, M., Ierodiaconou, D., Porskamp, P., Tinkler, P., Gray, S., Sams, M., Howe, S., Pocklington, J. and Whitmarsh, S. (2023). *An integrated monitoring program for Wilsons Promontory Marine National Park*, Parks Victoria Technical Series 120, Parks Victoria, Melbourne.

Appendix A Data Product Hierarchy

The data products identified in the tables below were collated by Parks Australia prior to the NESP project 2.3 during the initial development of the Data Product Hierarchy (Section 1.1). Many but not all of these data products were subsequently identified in the project as data products used by Parks Australia to answer Key Questions.

Table 11 Visualisation, analytics and reporting tools identified by Parks Australia as Level 5 data products in the Data Product Hierarchy. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key

QuestionsInformation sources	Description
Wylie (internal)	Wylie is a decision-support tool for analysis and reporting that allows users (internal to DCCEEW) to explore and discover matters of national environmental significance (MNES) in any location in Australia.
Protected Matters Search Tool	The Protected Matters Search Tool is used to generate a report that provides general guidance on Matters of National Environmental Significance (MNES) or other matters protected by the Environment Protection and Biodiversity Conservation Act (1999) that may occur in an area of interest. This tool is available to the public.
★ Seamap Australia	Seamap Australia decision-support tool enables use of a national collection of marine habitat data sets and related data sets.
★ AMP Science Atlas	The AMP Science Atlas, built on the Australian Institute of Marine Science's eAtlas platform, is used to communicate about activity and outputs from the Australian Marine Parks national marine science program.
★ AMSIS	Provides the Australian Government, Regulatory Agencies, and Management Bodies with an authoritative view of Australia's Marine Jurisdiction; providing clear and consistent geo-regulatory layers, and contextual information to support marine planning, and policy decisions related to the use and management of Australia's Marine Jurisdiction.
★ AMP Features GIS and analytics tools (internal)	AMP Features GIS application, analytical tool, and database enable Parks Australia staff to record and analyse information about geographic features (e.g. islands, reefs, seamounts, canyons) in Australian Marine Parks.
★ National Electronic Approvals Tracking System (NEATS)	The National Electronic Approvals Tracking System (NEATS) provides access to publicly available information concerning offshore petroleum titles and applications.

Table 11 Distributions, Monitoring Summaries, and Condition Assessments identified by Parks Australia as Level 4 data product in the Data Product Hierarchy. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key Questions.

Data product	Description
IMCRA Provincial and Mesoscale Bioregions	The bioregionalisation used in the design of Australian Marine Parks to help identify areas for protection. Provincial bioregions reflect broad patterns of biodiversity using fish populations and geomorphological features as proxies.
★AMP Ecosystems	AMP Ecosystems was developed to help articulate what ecosystems and habitats occur within IMCRA provincial bioregions.
★ Key Ecological Features (KEFs)	KEFs are elements of the Commonwealth marine environment that are considered to be of regional importance for either a region's biodiversity or its ecosystem function and integrity. Used to assist with decision making in relation to meeting Goal 3 of the NRSMPA Goals and Principles.
Biologically Informed Seascapes (internal)	Biologically Informed Seascapes represent a combination of physical and biological information that predicts where species are likely to occur. Used to assist with decision making in relation to meeting Goal 3 of the NRSMPA Goals and Principles.
★ Seamap Australia National Benthic Habitat Layer (NBHL)	The NBHL is the first Australian habitat dataset that seamlessly consolidates data from each of Australia's state and territory providers. The NHBL is available as a validated (in situ habitat observations) and unvalidated product.
★ Biologically Important Areas (BIAs)	BIAs are areas where species are known to display biologically important behaviour such as breeding, foraging, resting or migration. Used to assist with decision making in relation to meeting Principle 4 of the NRSMPA Goals and Principles.
Habitat critical to the survival of…	Distributions of habitat critical to the survival of threatened species listed under the EPBC Act.
Predicted Benthic Habitat	A benthic habitat distribution model for Oceanic Shoals Marine Park incorporated into the Seamap unvalidated NBHL.
Authorisation-holder benthic habitat	Surveys undertaken in Australian Marine Parks under authorisation (e.g. permit, license) from the Director of National Parks may produce benthic habitat maps derived from annotation of underwater imagery.
★ AMP Key Natural Values (KNVs)	The AMP KNVs product represent the distribution of species, habitats and other natural features selected using criteria based on Ecologically and Biologically Significant Area (EBSA) criteria, which overlap with the Key Ecological Feature (KEF) and Biologically Important Area (BIA) criteria.
Values Monitoring Summary Data Products	Undefined data products summarising consistently observed variables associated with a value in an AMP.
Values Condition Assessments	Various undefined data products representing the results of condition assessments of selected values.

Table 12 Supplementary data products identified by Parks Australia at Level 4 of the Data Product Hierarchy. These products do not represent natural values but were identified as sources of supplementary information needed to answer natural values questions. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key Questions.

Supplementary Data product	Description
★AMP common language activities and pressures layers	A collection of geospatial layers compiled under NESP projects that represent activities and pressures relevant to understanding the state or condition of a natural value.
AMP Cumulative Impacts (benthic)	An output of NESP SS2/D7 projects this represents the cumulative impact scores across the Natural Values Common Language (benthic and islands) ecosystems. Pixel values represent the cumulative impact score calculated as a weighted sum of the (standardised) activity/sub-activities pressure layer values in each raster cell, where the weighting reflects the vulnerability of the benthic ecosystem in each cell to the activity/sub-activities that exert pressure there. The map should be interpreted as showing the relative intensity of cumulative impacts. The absolute values of the scores have no ecologically meaningful interpretation.
AMP Cumulative Impacts (pelagic)	As above but for pelagic ecosystems
Emerging Pressures	Parks has identified the need to understand emerging pressures (e.g. renewable energy, offshore aquaculture). Specific data products for this category have not yet been identified.
Pressures Monitoring Summary Data Products	Undefined data products associated with consistently observing or measuring over time some variable associated with a pressure on a value in an AMP.
Pressure Status Assessments (various data products)	Undefined data products representing the results of status assessments of selected pressures on values in AMPs.
AMP Zoning	The AMP zoning data enables users to find networks, parks and zones on a map.
★ AMP Features (internal)	Used as a locator when areas of interest are expressed in terms of named features, this product represents named geographic features within or near Australian Marine Parks as polygons (unlike some locators that use points). Feature types include reefs, islands, cays/islets, and range of underwater feature types (e.g. seamounts, canyon).
AHO Charts	An AHO web service with up-to-date charts equivalent to their printed chart series.
Navionics Charts	Digital nautical charts by Navionics.

Table 13 Component distributions identified by Parks Australia at Level 3 of the Data Product Hierarchy. Items marked by a black star ★ were used to develop the AMP ecosystems layer. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key Questions.

Data product	Description
★ Bathymetry of Australian Marine Parks	This dataset is a compilation product of all publicly available surveyed bathymetry within the Australian Marine Parks (AMPs, 2023 boundaries), merged into a single multi-resolution composite per AMP.
★ Best Available Bathymetry, National (AusSeabed)	High-resolution bathymetry is managed by AusSeabed. AusSeabed generates Digital Elevation Models and publishes visualisations of bathymetry per survey and some multi-survey compilations through their data portal.
★ AMP Ecological Depth Ranges (Benthic)	AMP Ecological Depth Ranges are ecologically significant depth ranges derived primarily from the national 250 metre resolution bathymetry product from GA.
Bathomes (internal)	Bathomes within Australian waters are ecologically significant depth zones derived primarily from the distribution of demersal fishes and bathymetric data. The data product was used to assist with decision making in relation to meeting Goal 2 of the NRSMPA Goals and Principles.
★ GA Geomorphic Features (2006)	The data set provides outlines for the maximum extent of geomorphic units for Australia's Exclusive Economic Zone, including the offshore island territories, but not the Australian Antarctic Territory. These data were compiled as part of Geoscience Australia's integrated digital information system to provide improved accessibility and knowledge relating to the environmental management of Australia's oceans resources. The geomorphic units are to be used as surrogates for benthic habitats and can be best applied to the construction of bioregionalisations of the seabed.
★ Geomorphometry in AMPs	This dataset provides geomorphic features of the Australian Marine Parks (2022). The data was generated by Seamap Australia as part of an Our Marine Parks (Parks Australia) project with funding from the Australian Government to improve knowledge relating to classification of the Australian Marine Parks real estate. Bathymetry data was collated from existing AusSeabed data holdings and compiled into multi-resolution bathymetry mosaics for each Park.
★Allen Coral Atlas Reef Geomorphology (2021)	The Allen Coral Atlas uses Planet Dove image data, water depth derived from Sentinel 2, Landsat or Planet Dove satellite imagery, modelled waves and surrogates for texture and slope through machine learning random forest classification followed by an object-based cleaning approach using eco-geomorphological rules. Coral reefs are conceptually organised into four levels of physical and biological structure
★Knolls and seamounts in the world ocean	This dataset shows the global distribution of seamounts and knolls identified using global bathymetric data at 30 arc-sec resolution. Yesson, Chris; Clark, M R; Taylor, M; Rogers, A D (2011): Lists of seamounts and knolls in different formats [dataset publication series]. PANGAEA, https://doi.org/10.1594/PANGAEA.757564
★Reefs on the Australian Continental Shelf (2016)	This product collated all known mapping data from government and industry to provide an updated map of this key habitat around Australia.
★ CSIRO seamounts	An input to AMP Ecosystems. https://doi.org/10.3389/fmars.2020.00187

★National Submarine Canyons of Australia (2015)	This product contains 753 submarine canyons along the Australian continental margin and external territories, mapped from a range of bathymetry datasets.
Coral Sea Geomorphic Features (JCU, 2012)	Geomorphic features of the Coral Sea derived through manual digitisation of a 100 metre resolution bathymetry product.
Geomorphology of seven Coral Sea Seamounts (IMAS, 2019)	Seamounts extracted from Voyage IN2019_V04 data and classified to the Geoscience Australia Geomorphology Classification Scheme.
★Various survey-specific reef datasets	An input to AMP Ecosystems this data is described in a report that doesn't appear to be publicly available *Key Ecological Features of the East and South-east Marine Regions: "deep-reefs" within 150-700 m depths*, includes six NESP datasets and two CSIRO datasets.
Various survey- or park-specific geomorphology products	These can be from various methods including bathymetry surveys (e.g. multibeam) and remote sensing data analysis (e.g. satellite imagery).
Species Profiles and Threats Database	The database is designed to provide information about species and ecological communities listed under the Environment Protection and Biodiversity Conservation Act 1999. It includes species profiles with their conservation status, species lists, distribution maps, and other related data.
Australian National Expert Fish Distributions (CSIRO, 2012)	The CSIRO Marine and Atmospheric Research (CMAR) Bioregionalisation database contains expert assessed modelled distribution range information for fish species in Australian waters. Ranges for individual species are computed from values representing minimum and maximum depth, and linear start/end points around the 500-metre depth contour extrapolated towards and away from the coast. This product is used in Fishmap and Atlas of Living Australia (ALA)
★ Global Distribution of Seagrasses (UNEP-WCMC, Short FT, 2021)	This dataset shows the global distribution of seagrasses and is composed of two subsets of point and polygon occurrence data. The data were compiled by UN Environment World Conservation Monitoring Centre in collaboration with many collaborators (e.g. Frederick Short of the University of New Hampshire), organisations (e.g. OSPAR), and projects (e.g. the European project Mediterranean Sensitive Habitats "Mediseh"), across the globe.
★ Global Distribution of Coral Reefs (UNEP-WCMC, WorldFish Centre, WRI, TNC, 2021)	This dataset shows the global distribution of coral reefs in tropical and subtropical regions. It is the most comprehensive global dataset of warm-water coral reefs to date, acting as a foundation baseline map for future, more detailed, work. This dataset was compiled from a number of sources by UNEP World Conservation Monitoring Centre (UNEP-WCMC) and the WorldFish Centre, in collaboration with WRI (World Resources Institute) and TNC (The Nature Conservancy). Data sources include the Millennium Coral Reef Mapping Project (IMaRS-USF and IRD 2005, IMaRS-USF 2005) and the World Atlas of Coral Reefs (Spalding et al. 2001).

Coastal and Marine Resources Information System (CAMRIS) seagrass (1995)	This data contains information about the distribution of seagrass around the Australian coastline. It was prepared by Dr. Hugh Kirkman (CSIRO Division of Fisheries) from a review of published and unpublished sources, and updated by Dr. Ian Hahmdorf, (Bureau of Rural Sciences).
AMP Ecosystems Components distributions	Currently Parks has not adopted specific distribution products to represent the AMP Ecosystem Components. A finer categorisation of some components may be needed. Several of the ecosystem components are categories required/used for habitat mapping, (e.g. via Seamap), such as macroalgae, and sessile invertebrates.
Oceanic Shoals Marine Park Predicted Pelagic Biodiversity	The provides predictions of pelagic vertebrate species richness and relative abundance throughout sub-areas of the Oceanic Shoals Australian Marine Park (AMP). Predictive models were constructed from in situ observations made using mid-water stereo-BRUVS (baited remote underwater video systems) deployed at 116 sites within three sampling areas in the western half of the AMP.
CAMRIS Benthic Substrate (1995)	This database contains information about the distribution of 10 different types of sea floor sediment in the Australian region. It was derived from data collected and mapped by the Ocean Sciences Institute, University of Sydney.
Other national sediment or substrate distributions	For example, National Oceans Office benthic maps including gravel, sand, mud, carbonate, average grain size, Seabed Gravel Content Across the Australian Continental EEZ, 2011 which provides the spatially continuous data of the seabed gravel, sand, and mud content, presented in 0.01 decimal degree resolution raster format.
★ Backscatter per survey	Some backscatter data from multibeam surveys is submitted to AusSeabed.

Table 14 Observations identified by Parks Australia at Level 2 of the Data Product Hierarchy. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key Questions.

Data product	Description
★OBIS	OBIS provides a global collection of species occurrence data and supports a broad range of Darwin Core fields.
Atlas of Living Australia (ALA)	ALA provides a national collection of species occurrence (presence only). It aggregates from many sources of varying quality. The ALA Regions tool allows the user to select an AMP and get a report of species observations. Through ALA's spatial portal you can also overlay distributions.
Biodiversity Data Repository (BDR)	A new biodiversity data repository being developed by DCCEEW that uses a new data standard named ABIS (Australian Biodiversity Information Standard) and knowledge graph technology to support rich semantics
★ Squidle+/UMI	Image annotation data accessible via SQUIDLE+, the web-based framework that aims to facilitate the exploration, management and annotation of marine imagery.
★ GlobalArchive	Image annotation data accessible via GlobalArchive is an online centralised repository of fish and benthic image annotation, body-size and height information, stereo calibration and associated information.
★ National Reef Monitoring Network	The National Reef Monitoring Network brings together shallow reef surveys conducted around Australia (and globally) into a centralised database. This includes Reef Life Survey.
MARine Sediment (MARS) database	The MARine Sediment (MARS) database contains detailed information on seabed sediment characteristics for samples collected from Australia's marine jurisdiction using corers, dredges, and grabs. It includes survey and sample information such as locations, water depths and sample descriptions. Data are also provided from quantitative analyses of the sediments, such as grain size, mud, sand, gravel and carbonate concentrations, mineralogy, age determinations, geochemical properties, and physical attributes for down-core samples including bulk density, p-wave velocity, porosity and magnetic susceptibility. MARS holds >40,000 sample and sub-sample records, and approximately 200,000 records describing the characteristics of these samples.
Various per survey observations	Various per survey observations using varying standards, formats, structures.

Table 15 Primary data and samples identified at Level 1 of the Data Product Hierarchy. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key Questions.

Data and data portals	Description
Per acquisition satellite or aerial imagery	Individual acquisitions of satellite and aerial imagery occur as part of projects either through purchasing images or commissioning surveys (e.g. drone surveys).
National satellite or aerial imagery platforms or collections	For example, Digital Earth Australia (DEA) is a digital platform that catalogues large amounts of Earth observation data covering continental Australia.
★ Squidle+/UMI (imagery)	Imagery available via Squidle+/UMI. Squidle+/UMI does not store imagery but encourages sharing and re-use by providing a standard online interface hosted in a standard way by various institutions.
National Reef Monitoring Network (imagery)	Images are made available via an API that accompanies the National Reef Monitoring Network (NRMN).
Australian Marine Video and Imagery Data (GA)	This data package includes towed video and still images acquired on GA surveys from 2007 onwards. Between 2007 and 2013, this included 21 marine surveys (including Antarctic waters).
CSIRO Marine Invertebrates Image Collection (MIIC)	MIIC is a catalogue of geo-referenced images of invertebrates from surveys around Australia. Identifications have been done at CSIRO on the survey vessels or post voyage or the specimens have been passed to various Australian museums for identification. Images identifications use the Codes for Australian Aquatic Biota (CAAB) database and where possible CAAB records are cross linked to the World Register of Marine Species.
Per survey underwater imagery	Per survey underwater imagery exist in various repositories and institutional stores
National multibeam collection AusSeabed Data Level 2	Cleaned and/or derived variables (L1 data) undergoes cleaning and filtering to create the first ‘usable’ data. This is level 2 data under the AusSeabed Data Level Definitions.
Museums and institutional specimen archives	Museums and institutional specimen archives
GA Sediment samples	Sediment samples stored by Geoscience Australia
Index of Marine Surveys for Assessments (IMSA)	The IMSA is an online portal to information about marine-based environmental surveys in Western Australia.

Table 16 Deployments (including surveys and voyages) identified at Level 0 of the Data Product Hierarchy. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key Questions.

Data Portals	Description
★ GlobalArchive (deployment locations)	Deployment location data available via GlobalArchive, an online centralised repository of fish image annotation, stereo calibration and associated information.
★ Squidle+/UMI (deployment locations)	Deployment location data available via Squidle+, a web-based framework that aims to facilitate the exploration, management and annotation of marine imagery.
★ National Reef Monitoring Network (dive locations)	Deployment location data (dive locations) available via the National Reef Monitoring Network which brings together shallow reef surveys conducted around Australia (and globally) into a centralised database.
★ AusSeabed holdings dataset (multibeam)	This dataset represents the current extent of bathymetry surveys held by AusSeabed. This dataset is live and will continue to be augmented as coverage is supplied from AusSeabed collaborators.
MARine Sediment (MARS) database (sample locations)	Sample locations from the MARine Sediment (MARS) database which contains detailed information on seabed sediment characteristics for samples collected from Australia's marine jurisdiction using corers, dredges, and grabs.
Parks Standard Record of Deployments	Where not already provided through the use of platforms that provide access to this data Parks requires researchers to provide one or more spatial datasets representing all deployments of key platforms (e.g. AUV, BRUV, divers, towed video). Parks has developed a specification and template for a Standard Record of Deployments. This data does not include observations.
Institutional per survey locations	Various vessel operators **(**Marine National Facility, AIMS**)** provide voyage tracks
Index of Marine Surveys for Assessments (IMSA)	The IMSA is an online portal to information about marine-based environmental surveys in Western Australia.

Table 17 Reports and publications (unstructured data) identified by Parks Australia at Level -1 of the Data Product Hierarchy. The table shows example of tools park managers use to find relevant publications. It is not comprehensive. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key Questions.

Data Product	Description
★ AMP Science Atlas	The Atlas is designed to help communicate the historical science and research underpinning the design of Australian marine parks and share information about new and ongoing scientific research in these parks.
★ Parks Science Information Database Outputs List internal	An internal list of over 2000 science publications, curated by park managers.
★ Other websites (e.g. scientific publications)	For example, Parks Australia's website provide links to AMP-relevant science publications.

Table 18 Research capability and contacts data identified by Parks Australia at Level -2 of the Data Product Hierarchy. Items marked by a red star ★ were identified by this project as data products and repositories used by Parks Australia when addressing the Key Questions.

Information source	Description
Knowledge holder contacts	This includes online details for researchers and Traditional Owners
Conversations with Traditional Owners	Consultation with Traditional Owners
Researcher Profiles, ORCID	Research institutions publish profiles of researchers including their expertise and publications. ORCID is a unique identifier system used by researchers that can be used to help aggregate information about researchers, their expertise, affiliations, and outputs.
Institutional equipment inventories	Institutions publish information about their research capability including vessels and equipment
AODN Catalogue	AODN metadata catalogue brings together metadata from across Australia marine data community. Parks requires researchers to publish project-level metadata records in a repository harvested by AODN. The NESP Marine and Coastal Hub publishes project-level records to which it links related data outputs.
AusSeabed Survey Coordination Tool (including the National Areas of Interest layer)	The Survey Coordination Tool (SCT) is designed for the seabed mapping community to communicate their plans to survey. It includes the Areas of Interest layer which is compiled from submissions from private, academic, government and community groups across Australia's marine estate where they identify a need for seabed mapping or biodiversity characterisation. The layer provides metadata to help establish collaborations and understand the purpose, value and impact associated with data collection in the area of interest.
★ Expert advice	An example of expert advice being used to update a data product is - NESP Surveyed Reefs - Neville Barrett pers comm. 2021.

Appendix B Controlled vocabularies and the NVCL

The Natural Values Common Language (NVCL) is hierarchical, and the structure has been deliberately chosen to provide a balance between sufficient detail to allow unambiguous interpretation, whilst being sufficiently succinct so that its management role remains practical.

The use of the NVCL allows Parks Australia to use the same terminology across all AMPs. It could, however, also support condition evaluation in different habitats at different spatial scales, making it easier to scale or aggregate up from individual ecosystems or zones to a national scale as required, and thereby support management outcome reporting at park, regional or national scales.

The figures below summarise the relationship between the NVCL and other potentially relevant controlled vocabularies. In many instances these relationships could provide a mechanism that enables identification of individual species to be unambiguously allocated to NVCL ecosystem components, and for additional NVCL ecosystem sub-components to be identified. This mapping from individual species to NVCL would enable condition assessments based on metrics assigned to individual species to be “rolled-up” into ecosystem components, ecosystems, zone, park or regional level reports.

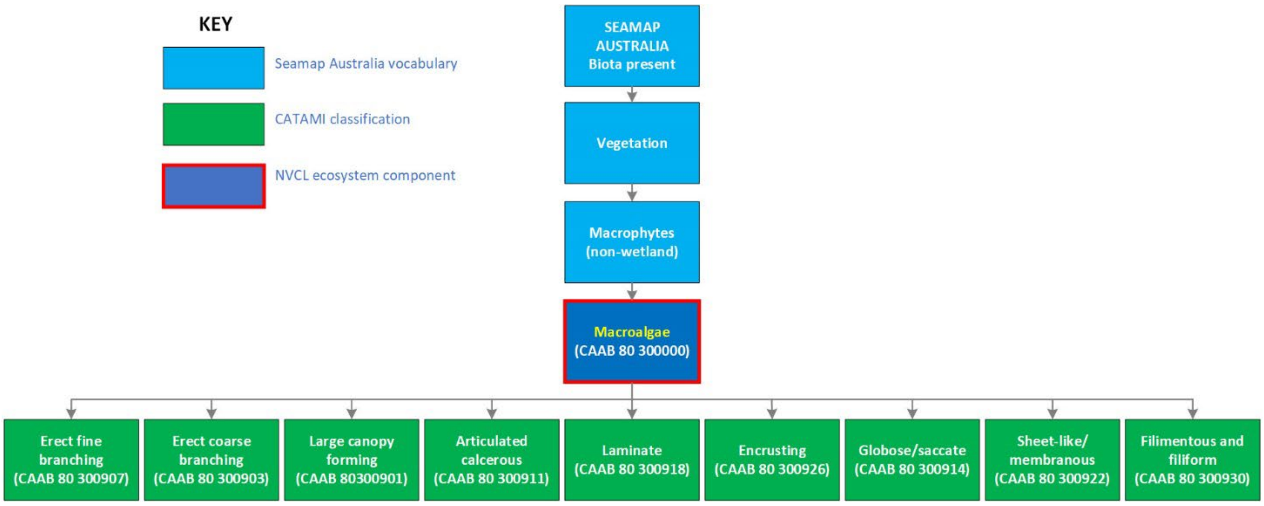


Figure 22 Illustration of the relationship between NVCL ecosystem component "macroalgae", the CATAMI controlled vocabulary and the controlled vocabulary established by Seamap Australia.

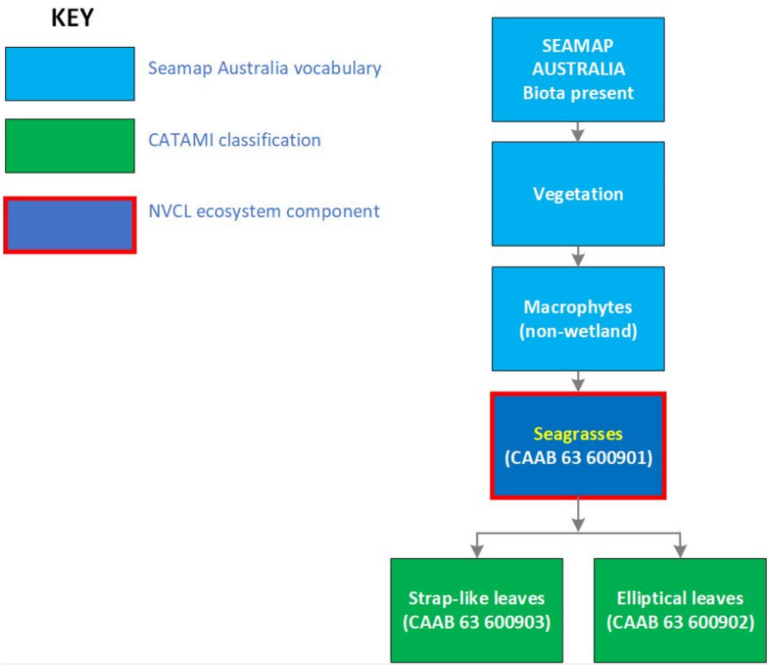


Figure 23 Illustration of the relationship between NVCL ecosystem component "seagrasses", the CATAMI controlled vocabulary and the controlled vocabulary established by Seamap Australia.

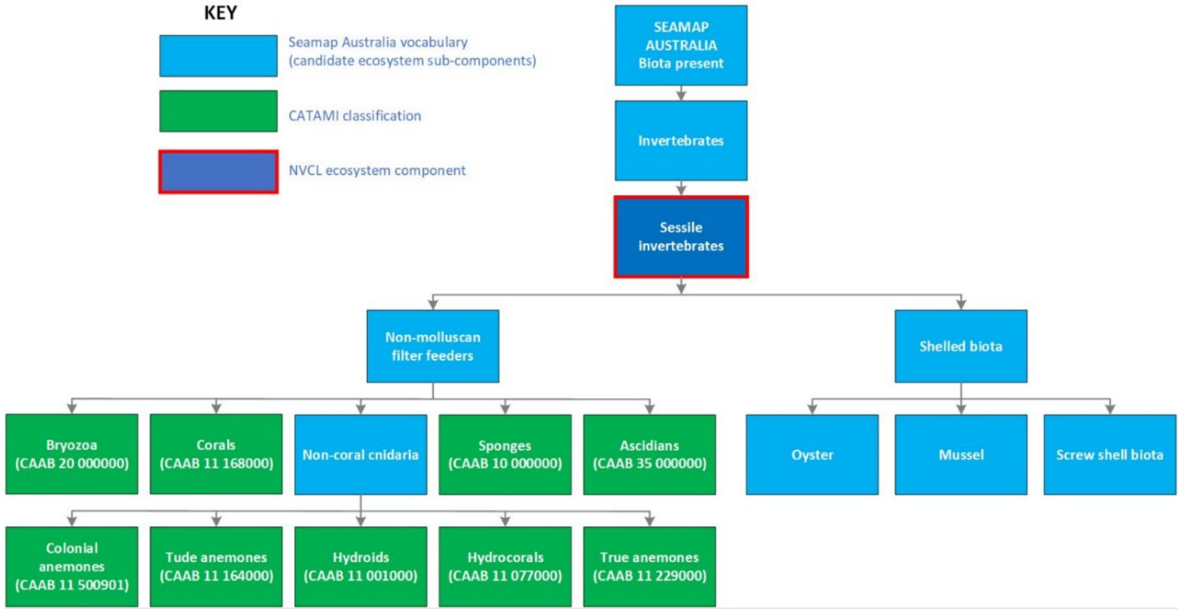


Figure 24 Illustration of the relationship between NVCL ecosystem component “sessile invertebrates”, the CATAMI controlled vocabulary and the controlled vocabulary established by Seamap Australia.

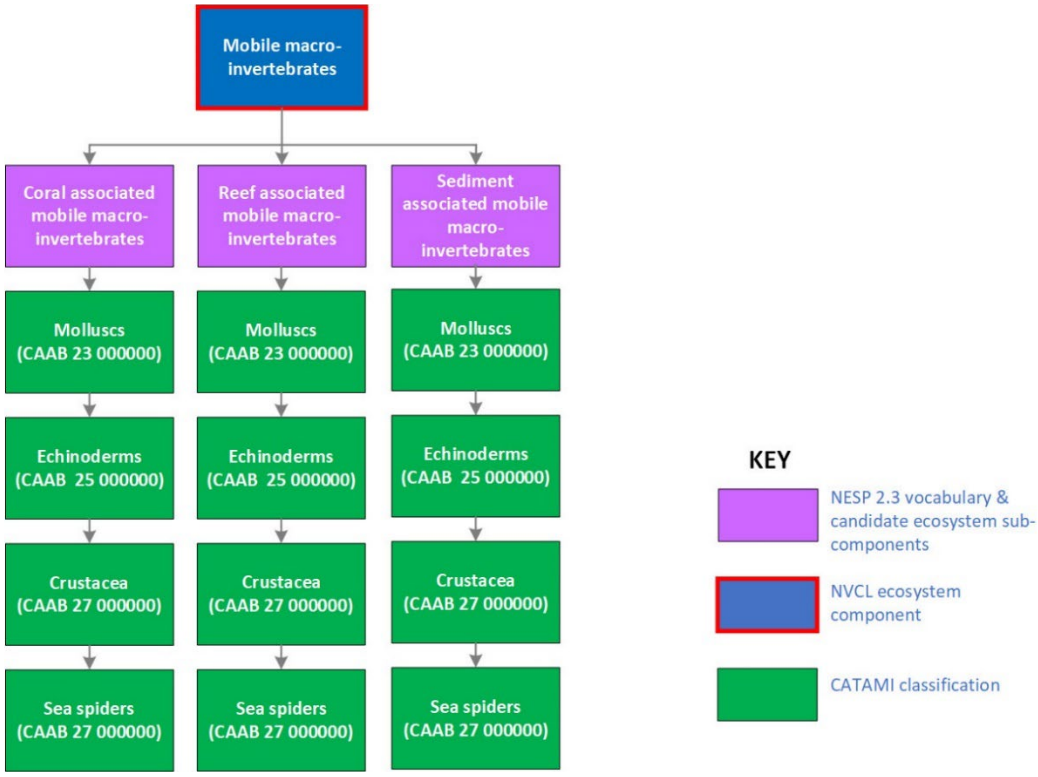


Figure 25 Illustration of the relationship between NVCL ecosystem component “mobile macro-invertebrates”, and the CATAMI controlled vocabulary. Potential additions to the NVCL vocabulary identified in this project are shown in purple.

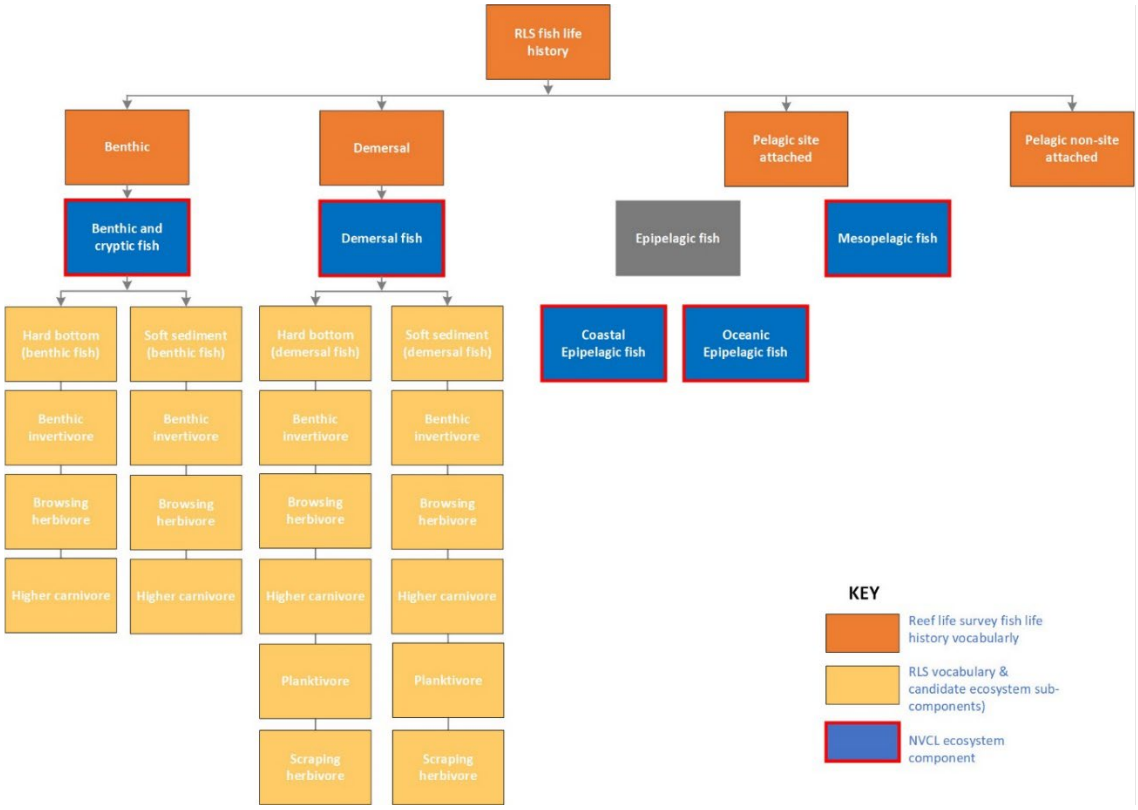


Figure 26 Illustration of the potential for a mapping between the NVCL “benthic and cryptic fish”, and “demersal fish”, ecosystem components, and the classification system developed by the Reef Life Survey.

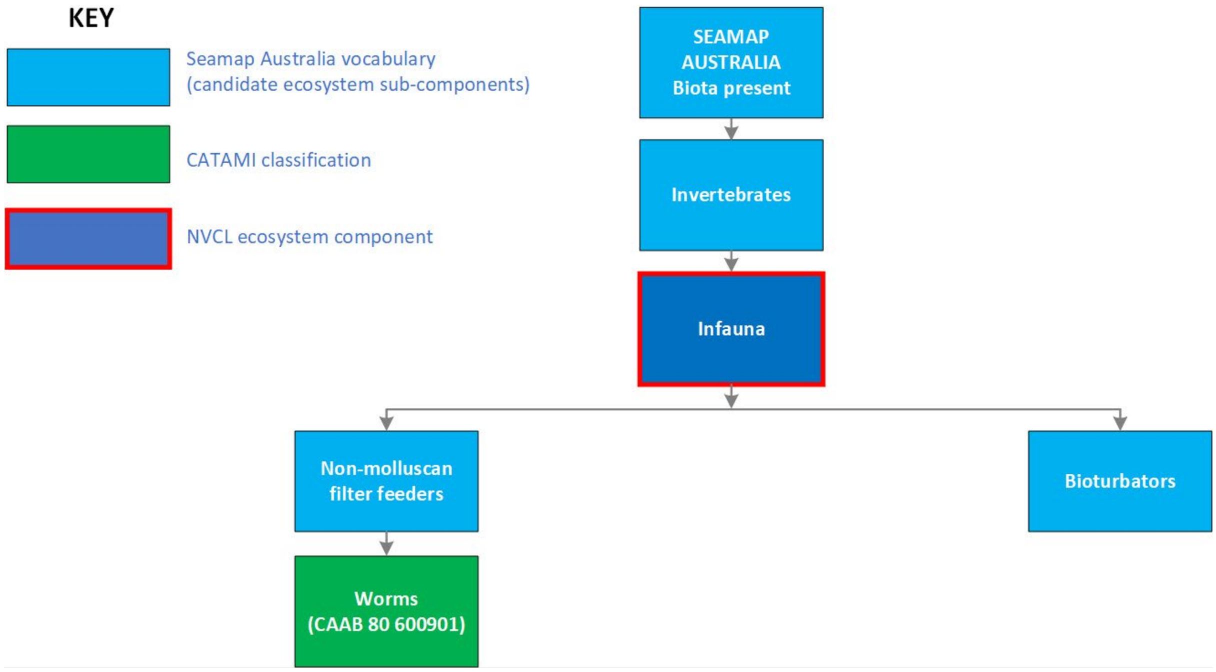


Figure 27 Illustration of the relationship between NVCL ecosystem component “Infauna”, the CATAMI controlled vocabulary and the controlled vocabulary established by Seamap Australia.

Appendix C Pressures common language

Pressure	Specific pressure	Notes
Climate change	Altered ocean currents	Strengthening of East Australian Current, changes in upwellings etc.
	Increased frequency and severity of weather events	
	Increased sea surface temperature	
	Ocean acidification	
	Sea level rise	
Extraction of living resources	Extraction of benthic mobile invertebrates	Lobster, crabs, abalone and scallops using pots, traps, hand removal etc.
	Extraction of fish and free-swimming invertebrates	Pelagic and demersal fish, squid etc. using lines, nets etc.
	Extraction of megafauna (excluding fish)	Marine mammals, marine reptiles, seabirds/shorebirds etc. includes bycatch, entanglement etc.
	Extraction of terrestrial flora and fauna	Terrestrial flora and fauna (birds, reptiles, mammals, harvesting of trees or woody shrubs for firewood etc.)
Physical disturbance	Habitat modification (physical disturbance and removal)	Physical damage to habitat caused by e.g. anchors, divers, fishing gear etc. to sessile invertebrates (including oysters and infauna), seagrass, macroalgae. Includes terrestrial flora
	Habitat modification (due to changes in nutrients and organic matter)	Marine sources only - Results in eutrophication and algal/epiphyte blooms - really only aquaculture (i.e. feeding finfish) and possibly fertiliser spills at sea
	Habitat modification (due to suspended sediments, including smothering)	Marine sources (results from dredging, excavation etc) and land based sources (run off from agriculture etc).
	Human presence (disturbance of mobile fauna communities or populations)	Vessel transit/vessel strike, disturbance of fish by divers etc. Human presence isn't used for habitat modification
Marine invasive species	Marine pathogens/disease	Abalone virus, coral disease, marine turtle fibropapillomatosis, etc. (may be exotic of native)
	Marine pests	Introduced species such as Northern Pacific Seastar, Japanese kelp, European fan worm, NZ screw shell, European shore crab, Pacific oyster, Codium etc.
	Overabundant marine native species	Black spined urchin, purple urchin, Sydney octopus, etc. and possibly plants/algae/sessile invertebrates as well
Terrestrial (islands) invasive species	Animal pests	Weeds, rats, cats, foxes, rabbits etc. (i.e. pest animals that that impact native fauna and don't fall into the category plant pests below)
	Weeds and plant pests	Includes weeds, as well as 'pest' animals that cause damage to plants, as well pathogens/disease (e.g. Xylella - a bacterial disease)

Marine pollution	Light pollution	
	Marine debris (including microplastics and litter on islands)	
	Noise pollution	
	Noxious substances (including chemicals & heavy metals)	
	Oil/fuel spill or leak	
	Sewage waste	
	Changes in nutrients and organic matter	Land based sources only (largely results in habitat modification through algal/epiphyte blooms)

Appendix D NESP 2.3 quality assurance criteria

Both the SAFE 2.0 Matrix and NESP 2.3 Data Product Quality Framework can be used to identify considerations for developing and maintaining trusted data assets. Both adopt a lifecycle view with key components representing stages across a lifecycle. Both include a stage representing collection or creation of data (Collect in SAFE and Science in NESP) and both incorporate a Use stage which feeds back to the beginning of the lifecycle through an evaluation and review process (Figure 28).

Neither model provides specific recommendations for specific data products or assets or their supply chains. For example, both point to the value of using vocabularies in data but deciding which specific vocabularies to use will be up to producers or consumers (users) in the supply chain for a specific product or asset.

For the purposes of NESP 2.3 neither model represents a supply chain at an adequate level of specificity. The SAFE matrix was applied in the EPBC Review to depict a very generalised national supply chain for environmental data. In NESP 2.3 supply chains are represented using a data product hierarchy inspired by a “data processing levels” depiction commonly used to represent the progression from raw data to data products. In NESP 2.3 the quality framework is used to help identify improvements in data products and associated supply chains.

The NESP 2.3 framework includes Requirements Quality to represent a stage in which a data product is specified or planned. This emphasises the importance of working with data product consumers from the outset to identify and clarify use cases for data, which will have implications for activities and capabilities in later stages.

Both models include a Use level or stage. The SAFE 2.0 Matrix identifies three high-level use cases: Decision Support Tools, Reporting, and Research. The importance of identifying use cases is emphasised in the Requirements stage of the NESP 2.3 Framework, however specific use cases are captured as the main drivers in the data product hierarchy. The Use stage in the NESP 2.3 framework emphasises the importance of adequately informed users to help ensure science data products have impact.

The SAFE 2.0 Matrix includes a Culture level that highlights incentives, governance, ethical considerations, and the FAIR data principles. These considerations are scattered across stages in the NESP 2.3 Framework. The NESP 2.3 roadmap report should address culture in its recommendations particularly those related to the supply chain.

The tables below summarise the quality criteria used to guide the data product quality assessment grouped under the quality framework categories (e.g. Requirements quality, Science quality) and highlight the link SAFE 2.0 matrix level. Guidance on how to apply the criteria is provided for some criteria.

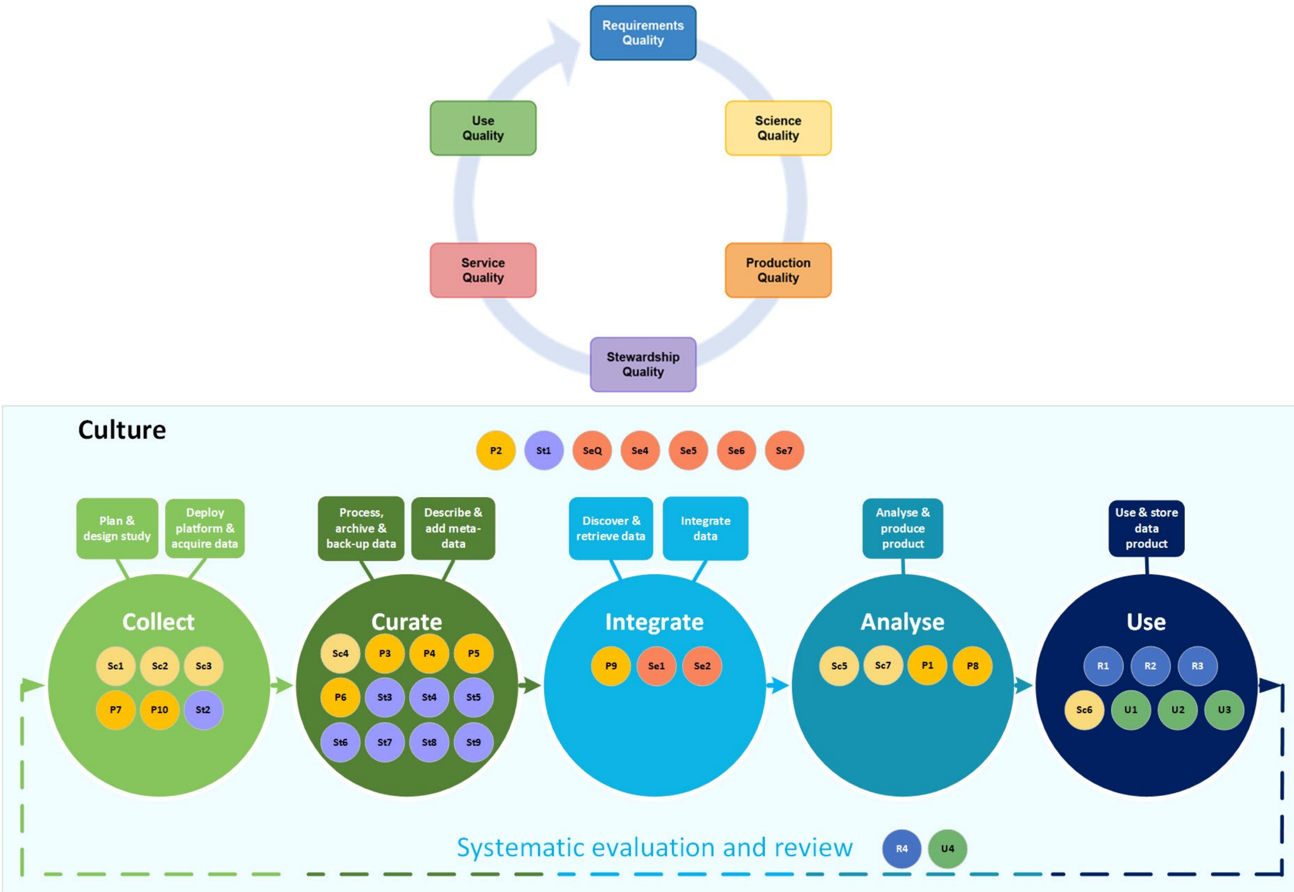


Figure 28 (Top) Parks Australia data product quality framework. (Bottom) The SAFE 2.0 framework together with equivalent steps identified within the scientific literature. The locations of the Parks Australia quality assessment criteria within the SAFE 2.0 framework are indicated by the small coloured-coded and annotated circles

Requirement quality

ID	Criteria	SAFE 2.0 level	Assessment Guidance
RQ1	Which key questions does this product address?	Use	
RQ2	Are there published data gaps and use cases for this product?	Use	This gives the science community visibility of high-level requirements
RQ3	Are there documented specifications for this product?	Use	Data products specifications are documented, preferably to a standard (e.g. ISO 19131) with elements including product scope, product identification (e.g. title), data content and structure, spatial and temporal reference systems, data quality, data production, data maintenance, portrayal, delivery information, metadata, and additional information.
RQ4	Is the data product regularly reviewed by end-users and the data provider?	Systematic evaluation and review	Product should be reviewed – consistent with Parks adaptive management approach product requirements should be reviewed over time based on user feedback to ensure the product continues to align with contemporary requirements.

Science quality

ID	Criteria	SAFE 2.0 level	Assessment Guidance
ScQ1	Are Standard Operating Procedures (SOPs) available and used for this data product?	Collect	
ScQ2	For data products derived from field studies: did the field study design follow the recommended procedures within the survey design SOP? If yes and the design was spatially balanced are the design covariates and site inclusion probabilities recorded in the meta-data?	Collect	
ScQ3	Where SOPs are not available, has best practice otherwise been applied?	Collect	Examples include classification systems like CATAMI and Semap Australia; authoritative species names directories like fishes of Australia, AFD, CAAB, WoRMS.
ScQ4	Are published and broadly adopted vocabularies used in this product?	Curate	
ScQ5	Has the product undergone rigorous expert/peer review?	Analyse	
ScQ6	Is the data product fit for purpose for this intended scientific use cases?	Use	
ScQ7	Is the confidence/uncertainty of within sample estimates or population estimates provided and are these meaningful in the context of the use case?	Analyse	

Production quality

ID	Criteria	SAFE 2.0 level	Assessment Guidance
PQ1	Is the data accessible and the spatial and temporal extent of the data applicable?	Analyse	
PQ2	How much of the data is openly available? How much is embargoed?	Culture	
PQ3	Are the data well-structured according to need?	Curate	Should include standardised and consistent naming conventions; tidy data principles
PQ4	Does the product used published vocabularies and standard formatting conventions?	Curate	Standardised and consistent format and units for a given parameter or field; preferred vocabularies are used for values
PQ5	Are spatial and temporal reference systems used documented?	Curate	
PQ6	Does the product meet broad data quality requirements?	Curate	e.g. currency, accuracy, precision or resolution, completeness, representation
PQ7	Is the file or web service format relevant and consistent?	Collect	e.g. species occurrence in Darwin Core, habitat in SHP files
PQ8	Are certainty or confidence indicators provided?	Analyse	The product includes a consistent and clear indication of uncertainty for each key feature, observation, measurement, or fact represented
PQ9	Can provenance and lineage of the product be tracked from creation to integration and use?	Integrate	
PQ10	Are the data real-time, near-real-time or provided in some other delayed mode?	Collect	Provide an estimate of currency.

Stewardship quality

ID	Criteria	SAFE 2.0 level	Assessment Guidance
StQ1	Does the provider have adequate funding, skilled staff and governance to effectively deliver the product?	Culture	
StQ2	The provider provides standardised metadata?	Collect	Based on standardised schema (ISO19115) and provides relevance and understandability for data users.
StQ3	Are there formal and consistently applied data quality assurance processes in place?	Curate	
StQ4	Does the provider have documented processes and procedures for managing archival storage of the data?	Curate	i.e. the data is backed up, should include copies to offsite locations.
StQ5	Does the provider have a business continuity plan?	Curate	To ensure ongoing access to and preservation of its holdings.
StQ6	Do the services have accreditation (e.g., Core Trust Seal or Accredited Data Service Providers)?	Culture	
StQ7	The data provider has an explicit mission to provide access to and preserve data in its domain.	Curate	
StQ8	Does the provider assumes responsibility for long-term preservation and manage this function in a planned and documented way?	Curate	
StQ9	Does the repository enable discovery of the data and provide consistent citation and/or persistent DOIs?	Curate	

Service quality

ID	Criteria	SAFE 2.0 level	Assessment Guidance
SeQ1	Does the provider have OGC compliant Web Mapping Services (WMS) and/or Web Feature Services (WFS) for the product?	Integrate	
SeQ2	Does the repository support other data delivery services (e.g. THReDDS)?	Integrate	
SeQ3	Does the provider support all of the FAIR data principles through platforms and applications?	Culture	
SeQ4	Where applicable are there sensitive data assessment and treatment policy and procedures?	Culture	e.g. personal data, culturally sensitive, ecologically sensitive, commercial-in-confidence, national security
SeQ5	Are the data licensed for open access?	Culture	e.g. Creative Commons and normally freely available
SeQ6	Where an embargo applies, is the product still released within a reasonable time period?	Culture	Generally 1-5 years case dependent
SeQ7	Are the providers platforms well maintained and updated?	Culture	

Use quality

ID	Criteria	SAFE 2.0 level	Assessment Guidance
U1	Do users' (e.g. Parks') procedures identify the product and characterises its uses in business context?	Use	
U2	Is the product accompanied by an appropriate level of interpretation?	Use	e.g. documented methodology, schema, how-to guide
U3	Does the product meet the users functional requirements?	Use	
U4	Are User feedback captured and used in product reviews?	Systematic evaluation and review	

Appendix E Bayesian habitat distribution models

Rational

Understanding the distribution of ecosystem complexes is a key part of Parks Australia's Management Evaluation (ME) framework (formally known as MERI - Management Evaluation, Reporting & Improvement). Parks have established an ecosystem common language that identifies ecosystems at three tiers. The first tier of the hierarchy are ecosystem complexes, where would be regions or areas that are made up of multiple ecosystem types. Tier two values are ecosystems (or ecosystem types), which essential represent key ecosystems and key habitats found within Australia's Commonwealth Marine Reserves. The third tier are ecosystem components, and these would be functional groups or species which are associated with specific ecosystems.

As part of Park Australia's Management Evaluation framework, a national layer of ecosystem types at a 250-meter grid cell resolution for all of Australia's EEZ was developed as part of NESP SS2 and D7 projects (Hayes *et al*., 2021; Dunstan *et al*., 2022). The layer integrated existing ecosystem type distributions, habitat distributions and rules about bathymetry and biogeographic provinces to identify a single ecosystem type for each 250-meter grid within Australia's EEZ. Despite the uptake of the approach in AMP reporting, there has been some desire to update the layer with more relevant information (where available) and understand explore how certain the ecosystem type classifications are. This case-study is designed to demonstrate how we could improve predictions of ecosystem types, while describing the uncertainty associated with those predictions. Here we use in-situ ecosystem type observations and other relevant spatial covariate information (such as multibeam data) to improve interpretation and reporting of ecosystem distributions.

Model description

Here we define a Bayesian implementation of multinomial family for a generalised linear model (GLM). The Bayesian implementation enables a few useful extensions compared to standard multinomial GLMs as available in statistical packages such as 'nnet' (Ripley *et al*., 2016). First, it enables us to sample the full posterior distribution for model parameters, and in turn generate uncertainty metrics for the predictive distribution of habitat classes. Secondly, it also enables the formal inclusion of expert opinion via prior elicitation. The prior elicitation step is beyond the scope of this report, but it should be noted that it provides a powerful option to inform model-based inference on the distribution of ecosystem classes, where empirical data are lacking, or extra gains can be made up consulting with experts from the field of marine ecosystems (or related disciplines).

The Multinomial GLM can be described as follows: if we have K independent multinomial vectors for each of the habitat classes. The observations of each class are defined as $(y_1, \dots, y_k)$, where each site observation (y_i) has an observation or count of each habitat type at each site (i). This framework enables the number of ecosystem classes scored per site to vary across sites. However, $(\sum_k^K \ y_{ik} = m_i)$ is fixed per site. This flexibility will prove itself to be useful when describe the data to be used in the ecosystem class modelling below. Thus, for each site we can describe the data with a multinomial distribution:

$$y_i \sim MN(m_i, \pi_{i1}, \dots, \pi_{ik})$$

To define the log-likelihood we consider (k) probability vectors $\pi_i \dots \pi_k$. We can express the probability vectors in terms of parameters in the model, typically we would do so using some linear predictor, where the regression coefficients are estimated against observed covariates for each site, the ecosystem class membership probabilities depend on covariates via a link-linear model:

$$\pi_{ik} = g(x_i; \beta_k)$$

$$\pi_{ik} = \frac{\exp\left(\beta_k x_i^\top\right)}{1 + \sum_j^{k-1} \exp\left(\beta_j x_i^\top\right)} \; if\; 1 \leq k \leq K-1$$

$$\pi_{ik} = 1 - \sum_j^{K-1} \pi_{ij} \text{ if } k = K$$

where β_k is a vector from the k^{th} row of (**B**) a $K-1 \times p$ matrix of parameters. x_i is a vector of covariates observed at site i, $\pi_{ik}$is the probability of each ecosystem class (k) at observed site. The link $g(.)$ is the additive logistic transform (Aitchison, 1982), which maintains the constraint that the elements of π_i sum to one and enables identifiability of the parameters in **B**. The reference class is set as the last class ($k = K$) in Y. The implication here is that these coefficients B represent the effects of X on the log-odds between categories $1 \leq k \leq K-1$ and K .

For each parameter β_{pk} in the β_k vectors, we assume is drawn from a Gaussian distribution, with a known mean (μ_{pk}) and variance (σ_{pk}):

$$\beta_{pk} \sim N\left(\mu_{pk}, \sigma_{pk}\right)$$

for the purposes of this report these priors are uninformative, and we set the mean for all $\mu_{pk} = 0$ and $\sigma_{pk} = 20$. Under a weakly informative or expert derived elicitation priors we could provide known values for μ_{pk} and σ_{pk}. It must be noted that the informative priors or the elicitation process would need to frame as the difference between the ecosystem class of interest and the ecosystem type which is set as the reference class for the model (typically the first or last column in **Y**. Where we are essentially estimating a series of log-odds ratios. For example, we might ask the question, given the reference class, sessile invertebrates, does the ecosystem class seagrasses tend to inhabit areas that have higher or lower current speeds?

Model code and software development

The code for the Bayesian Multinomial Generalized Linear Model is written in R, C++ and associated R and C++ libraries. The code is packaged up in the R package BAD (Bayesian Auto Differentiation) and uses CppAD to derive derivatives from a log-posterior (log-likelihood & log-prior) function that is written in C++. A gradient function is required to calculate the first derivatives of the log-posterior to use in gradient based Markov Chain Monte Carlo algorithms like Hamiltonian Monte Carlo (Neal *et al*., 2011) or NUTS (Hoffman *et al*., 2014). The code provided in this Rmarkdown document calls the appropriate function from BAD to fit the Bayesian Multinomial GLM. The BAD package is still in a prototype version, code to implement these functions can be requested from the first Author.

Ecosystem class observations

The data used to inform the distribution of ecosystem habitat composition comes from stereo-BRUV and BOSS deployments. The deployments are done using a spatial balanced design to distribute sites across each survey area in a spatially balanced pattern, while accounting for expected bathymetric gradients as part of the inclusion probability for survey sites. This approach follows the NESP Marine Biodiversity Hub Field Manuals and following methods given in Foster *et al*., (2017). This results in a spatial balanced sites for the South-West Corner region and ideally removes issues with preferential sampling for specific habitat types.

The data used to inform the distribution of ecosystem habitat composition comes from stereo-BRUV and BOSS deployments. Stereo-BRUVs and BOSS have a number of cameras that are pointed horizontally with the seafloor, as a consequence images only typically capture about 50% of the benthic habitat, with the other 50% pointing towards open water. Typically per image 20 randomly positioned points are scored per image (in the proportion of the image that intersects with the seafloor), annotation of the benthic habitat is done using a modified version of the CATAMI habitat classification scheme (Althaus *et al*., 2015). It must be noted that the total number of observations per-site (images and points scored per image) varies with across sites, and it defined as m_i, the sum of all scored benthic habitat points per site (see equations above).

The CATAMI habitat classification scheme tends to classify the habitat type at a finer resolution than that needed for the modelling in the NESP 2.3 project. We thus collapse some groups into more general groupings. For example, sessile invertebrates are a combination of corals, sponges, bryozoa, hydroids and other common sessile invertebrates. The other classes used in this models include, macroalgae, seagrasses, unconsolidated (sediment) and consolidated (bare reef/rock). For additional details on the methods used to survey, annotate and collate the ecosystem habitat data, see Langlois *et al*. (2021).

Physical data

Several physical datasets were used as covariates to describe the spatial distribution of ecosystem classes as part of model development. All oceanographic variables I smoothed up to 250-meter resolution Geosciences Australia bathymetric grid (Whiteway 2009), if near-shore physical variables were missing then these areas were spatial interpolated based on nearest known physical variables.

We ended up just using the bathymetry derived covariates, excluding depth. The other covariates were too spatially smoothed (from their coarser resolution) and depth resulted in an unidentifiable model due to full or partial separation of the ecosystem classes with depth as a covariate. Also, several of the environmental covariates where particularly coarse across this extent and didn't provide any meaningful contribution to the models, the development of high resolution (statistically downscaled) or regionalised climate variables would be useful for future modelling exercises.

Table 19 Physical covariates available for use for describing the distribution of ecosystem class distributions.

Variable	Data Source	Used
Depth	GA 250m bathymetry and topography grid	0
Roughness	Derived from GA 250m bathymetry and topography grid	1
Detrended bathymetry	Derived from GA 250m bathymetry and topography grid	1
Bathymetric Slope	Derived from GA 250m bathymetry and topography grid	1
Topographic position index (TPI)	Derived from GA 250m bathymetry and topography grid	1
Sea surface temperature	IMOS 1 monthly day and night average – made into a 10 year average	0
Sea level anomaly	IMOS DM02 dataset – made into a 10 year average	0
Southerly current velocity	IMOS DM02 dataset – made into a 10 year average	0
Easterly current velocity	IMOS DM02 dataset – made into a 10 year average	0
Primary productivity	IMOS Net Primary Productivity (GSM & Eppley-VGPM algorithm)	0
Openness	Euclidian distance from shore	0

Model fitting: Sampling the posterior distribution

Sampling the posterior distribution was done using Hamiltonian Monte Carlo (HMC) algorithm. Hamiltonian Monte Carlo is a relatively effective way to sample the posterior distribution when there is a known log-posterior and log-gradient for the posterior (Neal *et al.* 2011). We implemented our own sampler in R code and used the following steps to sample the posterior distribution.

For each $s\ = 1,\ \ldots,\ S\ in\ \beta_{ks}$ is sampled from the distribution from the log-posterior conditioning on the data and the current draws/values of $\beta_{kc},\ j\ =\ 1,\ \ldots,\ C$. Where C for represents an independent Markov Chain which uses the Hamiltonian Monte Carlo algorithm to sample the posterior. These chains can be run in parallel across each *s* to speed up inference if needed.

A total of 20,000 samples were taken per chain, chain convergence was assessed visual via trace plots. One of the challenges when sampling the posterior distribution using Hamiltonian Monte Carlo is finding the appropriate step-size (ϵ) to take for each new sample of the posterior, if ϵ is too small then the posterior distribution is not adequately sampled and you get chain wander, if ϵ is too large then the steps can jump into parts of the posterior where the values are too extreme, and the model will fail. To help with this problem we used a heuristic algorithm from Hoffman *et al*., (2014); algorithm No. 6 to find the initial step size. Additionally, the number of leapfrogs (steps) to take per sample also has an influence on model fitting and sampling of the posterior distribution. For initial model testing, we started with a relatively large number of leapfrogs (~100) but tuned this in to about 10 leapfrogs to balance accuracy and speed of model fitting. For parameter inference and prediction, we discard the first 10000 samples per chain and treat these as a burn-in.

Assessing posterior sampling

We assess the sampling of the posterior distribution by looking at the trace plots for each parameter. Trace plots where the chains are well mixed (looks like a fuzzy caterpillar) are a good indicator that the model is identifying the parameters well. We also look at the posterior distribution to make sure there are no problematic distributions that appear bimodal. We also assess fits via partial response plots, these plots enable up to assess the marginal effect of a covariate on the response of ecosystem type across a one-dimensional gradient (say Bathymetry).

Visualising uncertainty in ecosystem type predictions

One of the key objectives of this project was to demonstrate how to report uncertainty in the distribution of ecosystem types. We present three different ways to report uncertainty in the predicted ecosystem types

The first approach was to report a map of uncertainty for the predicted probability of occurrence for each ecosystem class as a separate map that can be viewed in parallel to a map of the expected (mean) probability of occurrence. The error is reported as the inter-quartile range (IQR), which is essentially the $P(y_{ik})_{0.75} - P(y_{ik})_{0.25}$ and provides a general summary of dispersion. From a Bayesian context this most analogous to standard errors, as commonly reported in maximum-likelihood models.

The second approach was to report a map of expected probability of occurrence combined with a map of uncertainty, we use the same mean probability of occurrence and the IQR of probability of occurrence as described above, to generate a bivariate map to represent these two axes in a single map. The colours scales are reported on two axes, one for the probability of occurrence and the second axis which represents the error in those predictions. It is the combinations of mean and error which gives the plot it'

The third approach represented uncertainty via an interactive map based on exceedance probabilities. Under this approach we provide a dynamic mapping tool (which is hosted on Seamap Australia: https://seamapaustralia-uat.imas.utas.edu.au/map/#30fe2ea4-1168-4399-83f0-97e7a472912c) where a slider bar can be used to view cells across the study area that are expect to have a probability greater than some probability of occurrence $E(\widehat{y_{ik}}) \geq p$, to report uncertainty in this framework we provide a map where each cell of the map (*i*) is reports the proportion of samples from the posterior predictive probabilities that satisfy $E(\widehat{y_{ik}}) \geq p$. Thus, for each cell we can report:

$$\widehat{v_{ik}}\left(E(\widehat{y_{ik}})\right) = \frac{\sum_{s=1}^{S} E(\widehat{y_{ik}}) \geq p}{S}$$

Where $E(\widehat{y_{ik}})$ is the expected probability of each ecosystem class (k) at each site (i) for each s from the posterior distribution. The variance layer for each class k has values that range zero, where for that specific cell and exceedance probability (p) has no samples from the posterior have values 1 where all samples from the posterior $\geq p$. The values $[0,1]$ are then used to adjust the transparency of the predicted probability of occurrence for each exceedance probability map. We set p as one of the following probabilities $\{0.000001, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.999999\}$, and thus result in 11 maps of exceedance probability uncertainty per ecosystem class.

Results Estimation of parameters

During model testing and fitting we ran a series of different model fits using all covariates in described above, but we found that some covariates like bathymetry were unidentifiable in the model. This is likely due to full or partial-separation of the multinomial log-likelihood, Albert and Anderson (1984) define this as, “there is a vector that correctly allocates all observations to their group”, we can see the distribution of unconsolidated sediment is largely in deeper habitats (Figure 32) and this could be a key reason for the issues associated with fitting this model with bathymetry as a covariate.

After some model testing, we settled on independent quadratic polynomials for roughness, detrended bathymetry, bathymetric slope and topographic position index (TPI), the coarser oceanographic variables like sea surface temperature are likely to be important for the ecology of some of the ecosystem classes but were to spatial coarse to provide meaningful model inference at the fine scale 250 meter we were working on with the bathymetry derived covariates. We present the trace plots for the intercepts (Figure 29) to demonstrate convergence of parameters across multiple chains. Trace plots for the remaining covariates are presented in Figure 30.

We also want to be able to assess if the posterior distributions for the parameters are unimodal and do not have any odd bimodal or other features (very long tails) that would suggest a poor fit. We can see that for an ecosystem type, the posterior distributions appear to be unimodal (Figure 31). We can also see that the ecosystem classes which contain few observations (seagrasses and consolidated) have the widest posterior distributions which reflects the higher uncertainty in those parameters, which is what we would expect under a joint Bayesian model.

Ecosystem type responses to covariates

Partial response plots demonstrated ecological reasonable responses to the environment covariates included in the models. The purpose of the partial response plots is to understand how each ecosystem type are predicted to respond to environmental covariates used in model fitting. Because we fit K-1 covariates, it is easier to interpret the responses on the probability scale (after it has been passed through the link function) and include the response of the reference class (sessile invertebrates). For the ecosystem types, sessile invertebrates and unconsolidated substrate to TPI seem ecological sensible (Figure 32). We can also see a mild response of sessile invertebrates and macroalgae to medium levels of roughness. We can see that seagrasses have a high response to roughness this could be a function of missing covariates that adequately describe seagrass distributions within the context of this region. We can also see that there is a large amount of uncertainty (credible interval) around the response for seagrass and roughness, which might reflect the lack of information available to accurately describe this response.

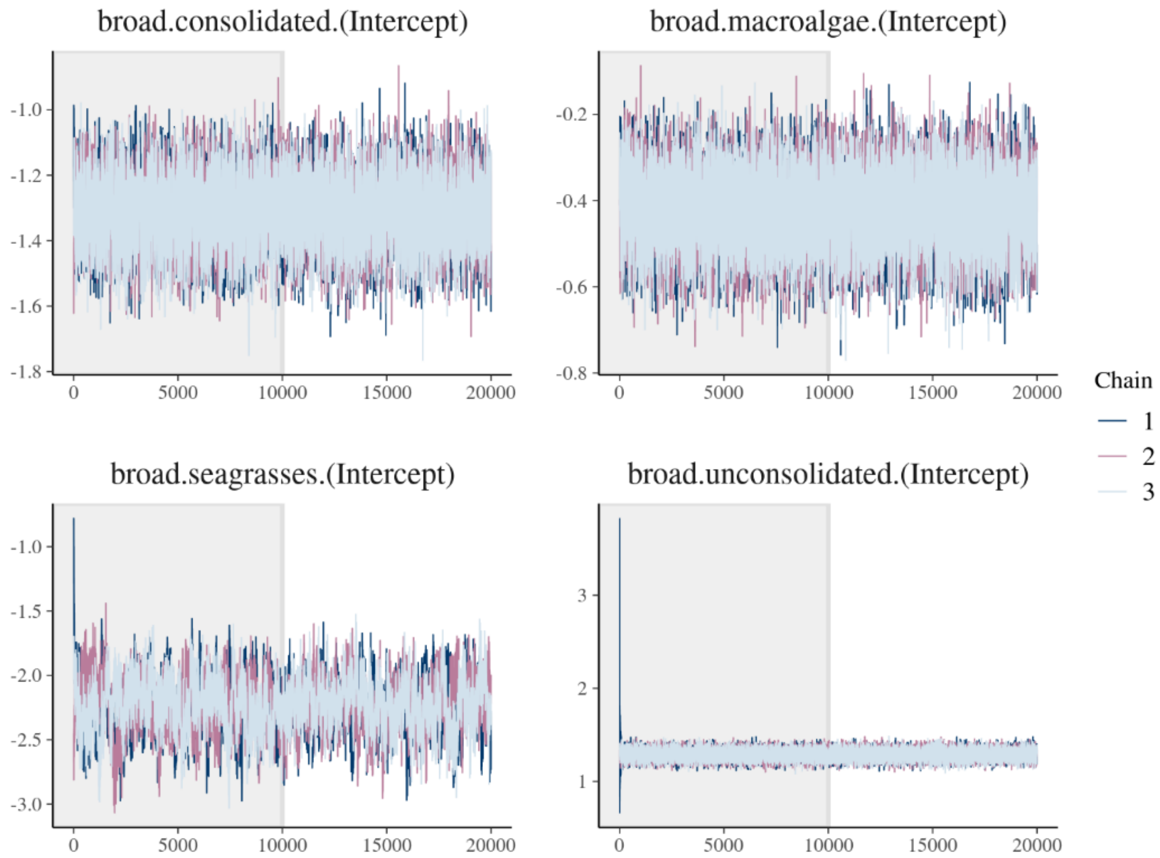


Figure 29 Trace plots for the intercepts for each ecosystem class. The grey areas represent the burn-in samples The trace plots are well mixed and demonstrate that the model is converging on a unimodal distribution for each intercept.

Figure 30 Trace plots for all coefficients in the model, each column represents a different ecosystem type.

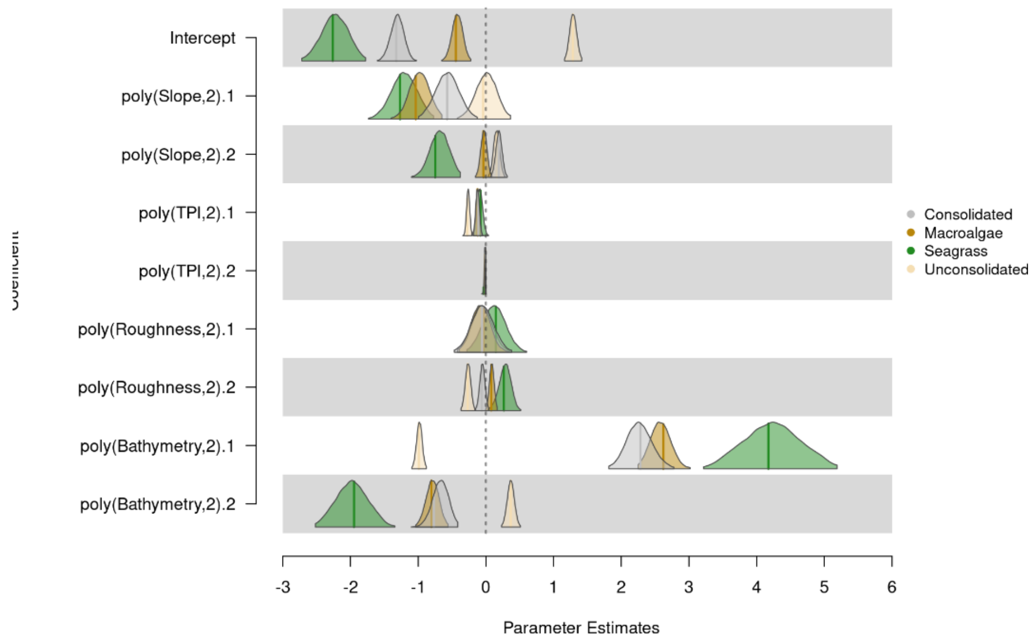


Figure 31 Posterior distributions for each model parameter. The shaded area represents the credible interval, and the line represents the median for each distribution. The colours represent each ecosystem type. Note we do not estimate parameters for the reference class (sessile invertebrates).

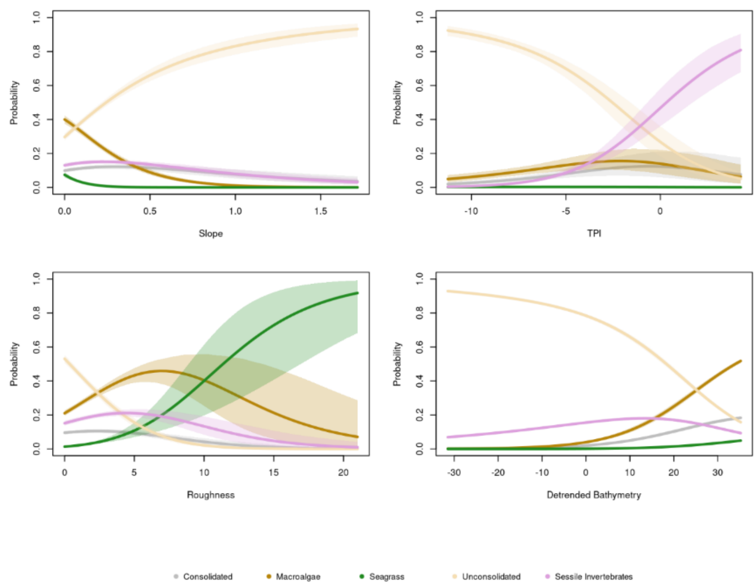


Figure 32 Partial response plots for each ecosystem type, excluding the reference class (Sessile invertebrates). We report the & credible intervals as the shaded regions, and the line represents the mean.

Reporting uncertainty

Here we present three main ways to report uncertainty (two a figure in this report and one online). The three approaches have different benefits, the first approach (Figure 33), is the classic way to report uncertainty. Here we report the inter-quartile range as a way to report uncertainty in a single figure (rather than presenting three maps one for the lower quartile, mean and upper quartile). The uncertainty maps in Figure 33 demonstrate for each class which sites are most variable from the posterior predictive distributions.

The second approach combines the mean probability of occurrence for each ecosystem class with the variance (IQR). This enables us to see areas that are highly likely and highly certain for the distribution of an ecosystem type (purple areas in Figure 34), areas that high likely by also highly uncertain (green regions in Figure 34), areas that are highly unlikely and highly uncertain (grey areas in Figure 34) and areas that are highly unlikely and highly uncertain for an ecosystem type to occur (yellow areas in Figure 34). One drawback of this approach is it can be difficult to work out the likelihood of occurrence and the certainty in prediction at mid-values for the colour key as the colours tend to blend into each other, in addition the colour palette is actually discretised so selecting the appropriate number of breaks and scale to represent the colours has a dramatic impact on the plot colours.

The third type of uncertainty plot was developed as a relatively simple way to communicate how uncertainty is intrinsically linked to the probability of occurrence, the plots developed for SeaMap Australia aimed to assist decision making, so managers could identify say areas of high probability for an ecosystem class and also see from the model simulations how confident the model was in those predictions (https://seamapaustralia.org/map/#f3c5062a-60d9-45af-9f15-3bda3016c355).

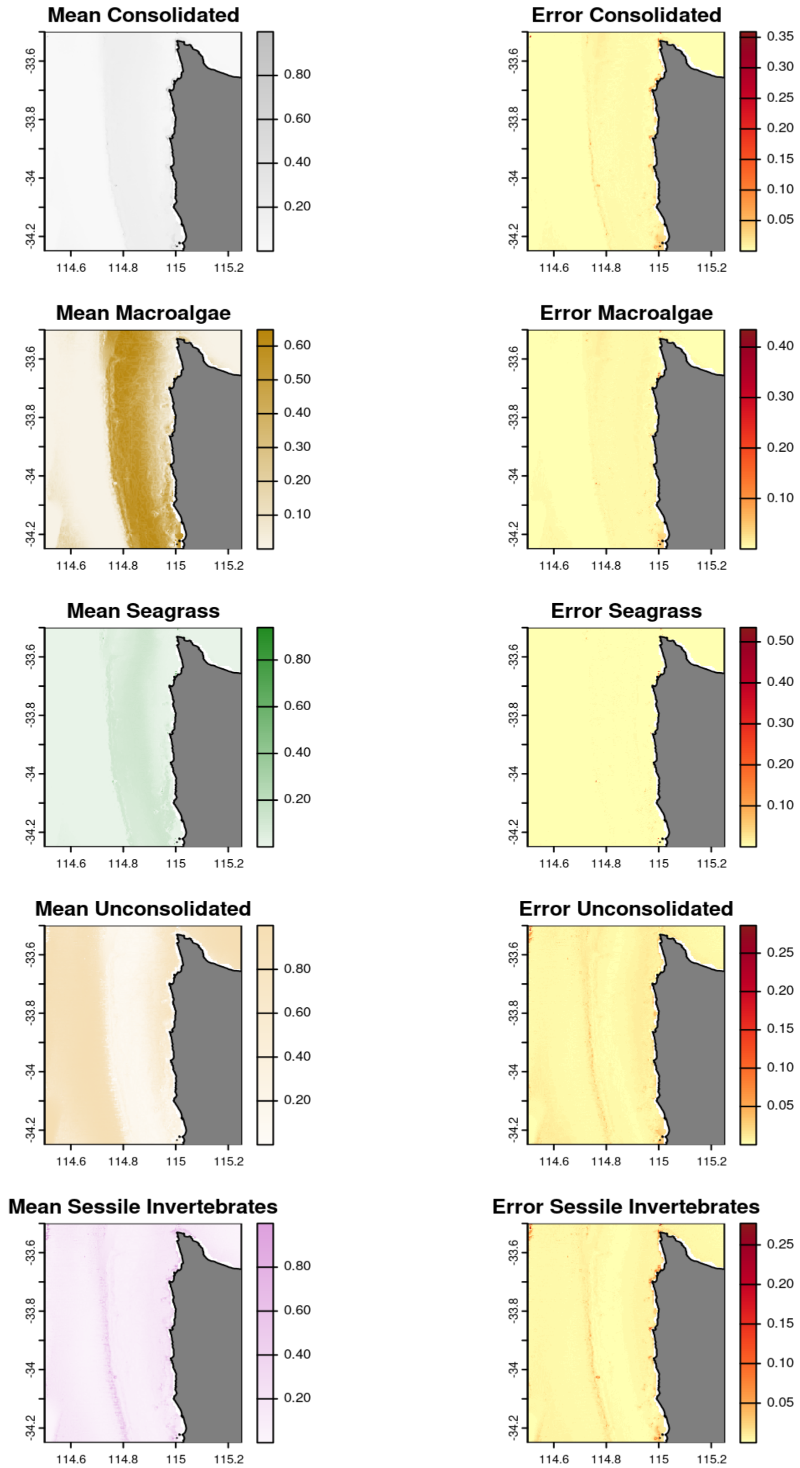


Figure 33 Mean prediction from the posterior predictive distribution for each ecosystem type, uncertainty is reported as the absolute difference between the 5 and 95 credible intervals from the posterior predictive distribution. The colours for each class highlight areas of higher probability where the colours for each class are bolder. The uncertainty in these predictions is represented by cells that are a stronger shade of red.

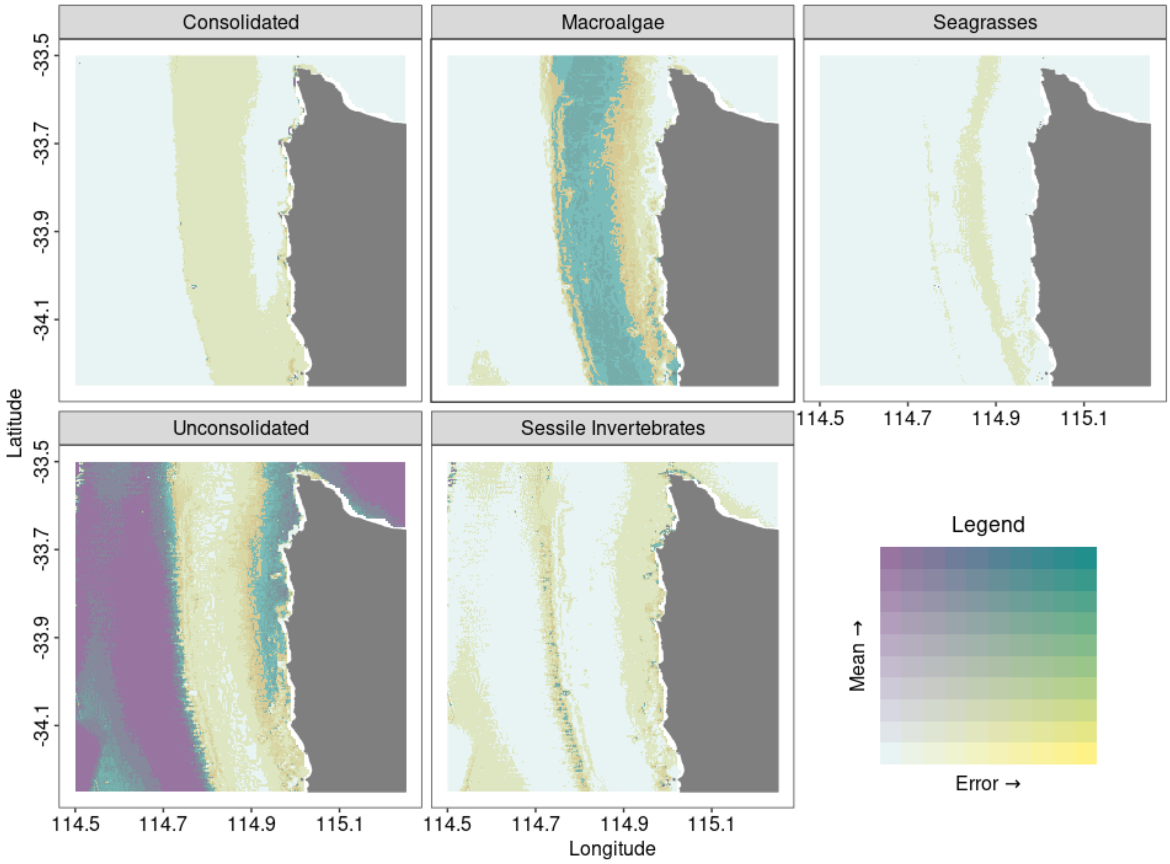


Figure 34 Bivariate uncertainty map for each ecosystem type. In this map, the colours represent two axes one for the mean probability of an ecosystem class and the other axis for the uncertainty in those predictions.

Appendix F High resolution bathymetry

Up until now we have been using spatial covariates based on the Geosciences Australia 250-meter bathymetric grid. To demonstrate the utility of including high resolution bathymetric products, we developed a model specificity for a subset of the Southwest Corner region. We used depth and depth-based derivatives, derived from the Geosciences Australia multi beam surveys of Southwest Corner (Langlois *et al*., 2021a). Due to fundamentally different scales that derivatives of bathymetry are calculated on a 250-meter grid, we could not simply predict into a finer 5-meter bathymetric product. We followed the above modelling steps but reduced the observed data to those observed within the multi-beam region (Figure 35).

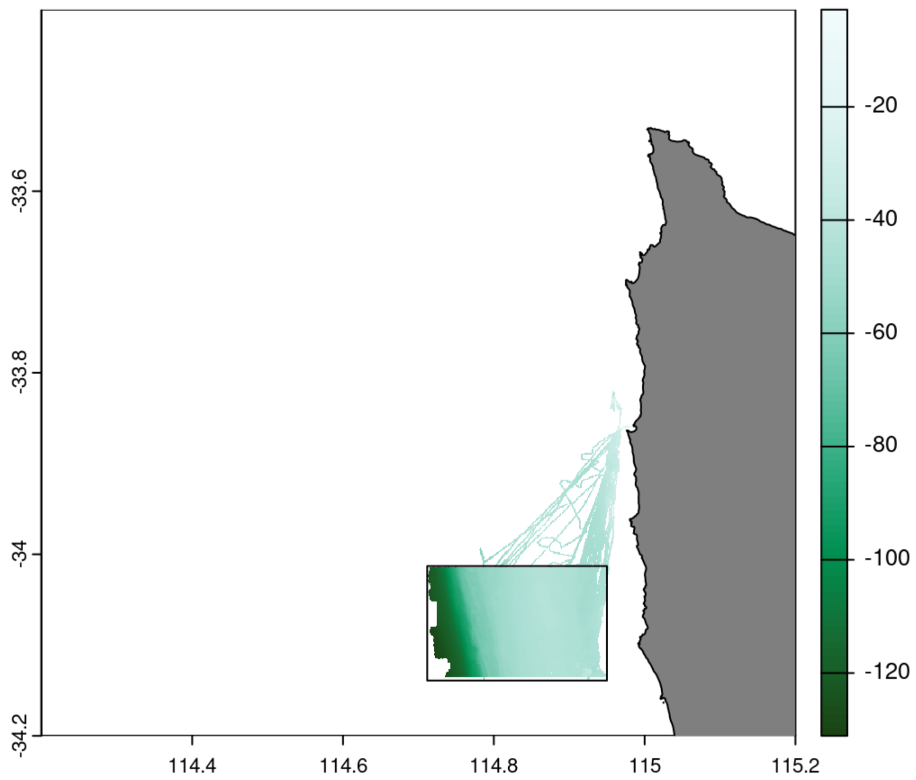


Figure 35 The distribution of high resolution bathymetry data (m) for the Southwest Corner region. The data come from: 2020. South-west Corner Marine Park survey – bathymetry (GA-4858). Geoscience Australia, Canberra. https://dx.doi.org/10.26186/145281. We use these data to generate a new high-resolution model for ecosystem types within the bounding box region in the above figure.

Updating the AMP ecosystem type layer with new quantitative predictions

The original Australia Marine Parks' ecosystem layer was developed using the best available ecosystem type distribution data, along with other geomorphic features including rules based around bathymetry to identify a specific ecosystem class per 250-meter grid cell for all out Australia's EEZ. For this study region, we compare the AMP layer as developed by Dunstan *et al*., (2022) and a hard-classification scheme derived from the model predictions above. The hard classification approach assigns an ecosystem type to each 250-meter grid cell, based on the rule: $\widehat{y_i} = \max\left(E(y_{i1}), \ldots, E(y_{ik})\right)$. We also generated a third map for the multibeam data region and use the same hard-classification rule to predict the distribution of each ecosystem type used in the Bayesian models.

We compare the three maps in (Figure 36) we can see for the region where we sample, that we gain more information on the distribution of macro algae and sessile invertebrates, which typically would be classified into shelf Vegetated Sediments (Macroalgae and Seagrass) and reef ecosystems (shallow rocky reefs, Mesophotic rocky reef, Rariphotic shelf reefs and upper slope reefs). For this area, we are underrepresenting reef and vegetated sediment habitats, and over representing the shelf unvegetated sediments.

We can also see the increase detail that is gained from developing a model with high resolution bathymetry data. Where there is a lot of finer detailed definition of sessile invertebrates, consolidated and macroalgae ecosystem types on the shelf break. Note , however, that if we extrapolate the model predictions to areas outside of the region that was surveyed, the model may under predict shelf vegetated sediments and other coastal reef ecosystems, that have directly been mapped using remote sensing or other monitoring approaches (see also Appendix G). We suggest that an appropriate rule for updating the AMP ecosystems layer might be to create a hierarchical product that identifies direct observations first, then model-based predictions.

Discussion

Here we have demonstrated a capacity to update the distributions of AMP ecosystem types using novel statistical models and *in-situ* observations. The model-based inference also enables us to explore uncertainty in the data and predictions. Although uncertainty can be a challenging concept to grasp on face value, it provides an important extra dimension in decision making, and enables decision makers to explore the confidence in the predicted distribution of an ecosystem type. For example, a decision maker might choose to prioritise areas for management and conservation where an ecosystem type is highly likely to occur, and the model is highly certain in those predictions (Figure 37; top right corner). Based on the data used here, our model demonstrates that there are far fewer cells (regions) where there is high confident that an ecosystem class is likely to occur. This is in stark contrast to AMP ecosystem layer, which treats all ecosystem classifications as present or absent, without considering uncertainty.

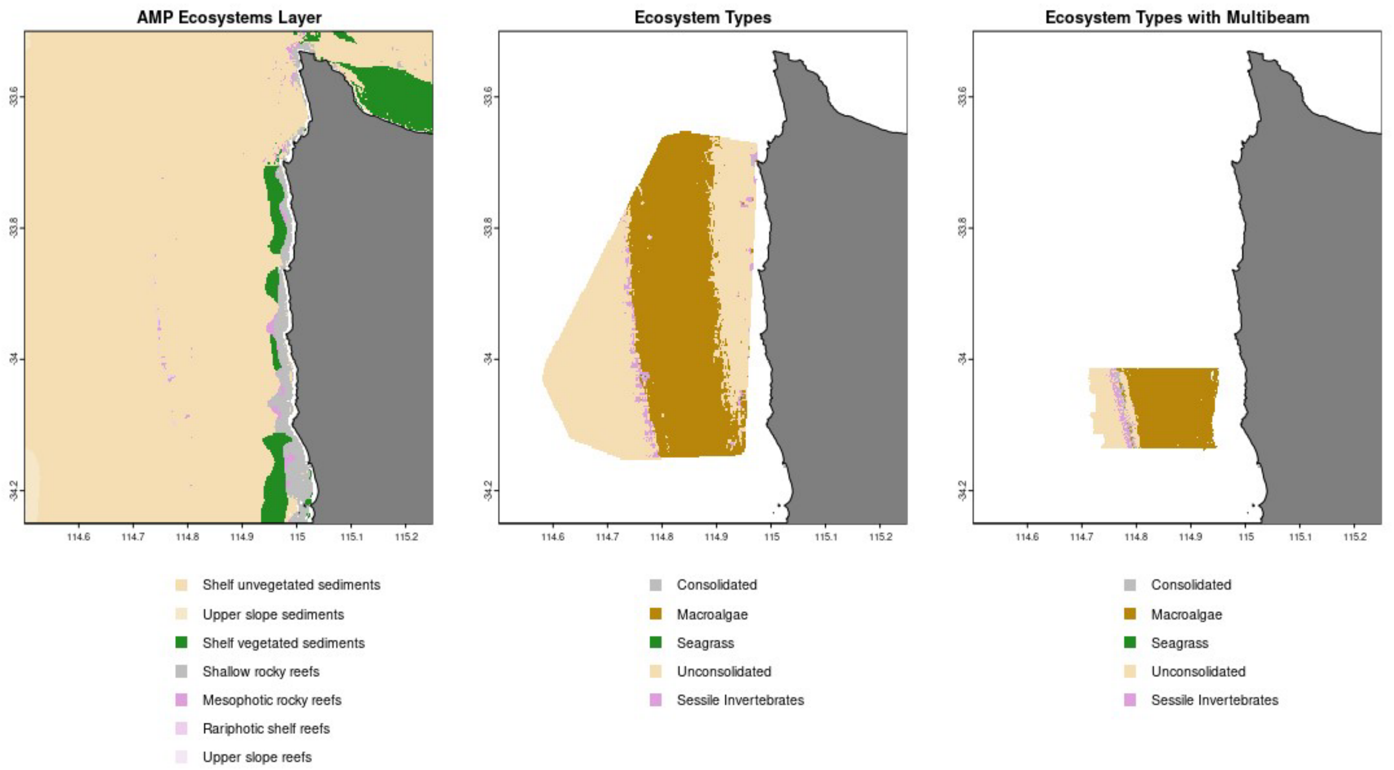


Figure 36 Comparison of hard classification from Bayesian model prediction versus the AMP Ecosystem layer for the South-west Corner study region. Points represent the sites surveyed using a BRUV or BOSS in-situ observations.

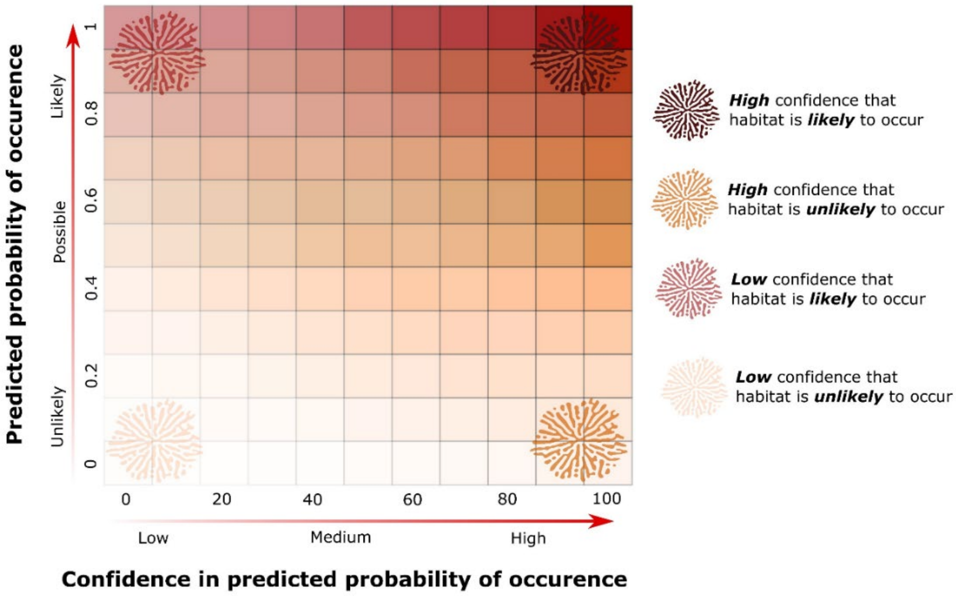


Figure 37 Mean probability of occurrence for five habitat classes in the South-west Corner marine park. Opacity for each habitat class indicates the likelihood of that habitat type occurring at that location (transparent = unlikely to occur; opaque = likely to occur)

Using high resolution bathymetry to improve spatial predictions

Superficially the inclusion of high-resolution bathymetry seems to produce a more accurate and better predictive map (Figure 38), but it needs to be made clear that the prediction is only as good as the model used to develop it. A future area of important research for understand ecosystem type classification would be to explore the precision and accuracy of predictions that build on these fine scale datasets. One way to assess this would be via some type of cross-validation procedure and then assess how accurately a fine resolution bathymetry dataset-based model predicted in-situ observations of ecosystem types compared to a courser resolution model.

Model extensions and improvements

There are a few natural extensions of this approach. The first would be to explicitly elicit priors on the distributions of ecosystem classes. This would be useful for understanding the distributions of ecosystem classes in low data situations or using expects to distinguish between ecological responses of ecosystem classes when the model struggles to do so, such as the problem we see with identifiable of certain covariate responses, or the uncertain responses of some ecosystem types against specific covariates, e.g. seagrasses and roughness. This is not a trivial process and requires a considered approach to engage with experts effectively.

The second logical extension would be to make the model ‘spatial’ by including a continuous spatial random effect, like a Gaussian random field (Cressie, 1993). The most basic version of this model would include a Gaussian random field just on the reference class. Where essentially, we could add the Gaussian random field as a spatial random effect to the linear predictor, the mathematics implies that it’d be the difference from the reference class (same as the intercept). We could also add a GRF for each class, this model might be harder to identify parameters with many of latent variables needing to be estimated.

One issue with the spatial random effect approach is it does not scale well; we would need to be careful about how we set-up the model so we can build an appropriate model which still remains useful and computational feasible at broad spatial scales. One approach is use approximate methods when estimating the spatial random effect, existing approaches which could be promising would be nearest neighbour Gaussian process (Datta *et al*., 2016) - which conditions the spatial process on a set neighbourhood, Fixed Rank Krigging (Cressie and Johannesson, 2008) or stochastic partial differential equations (Lindgren *et al*., 2011).

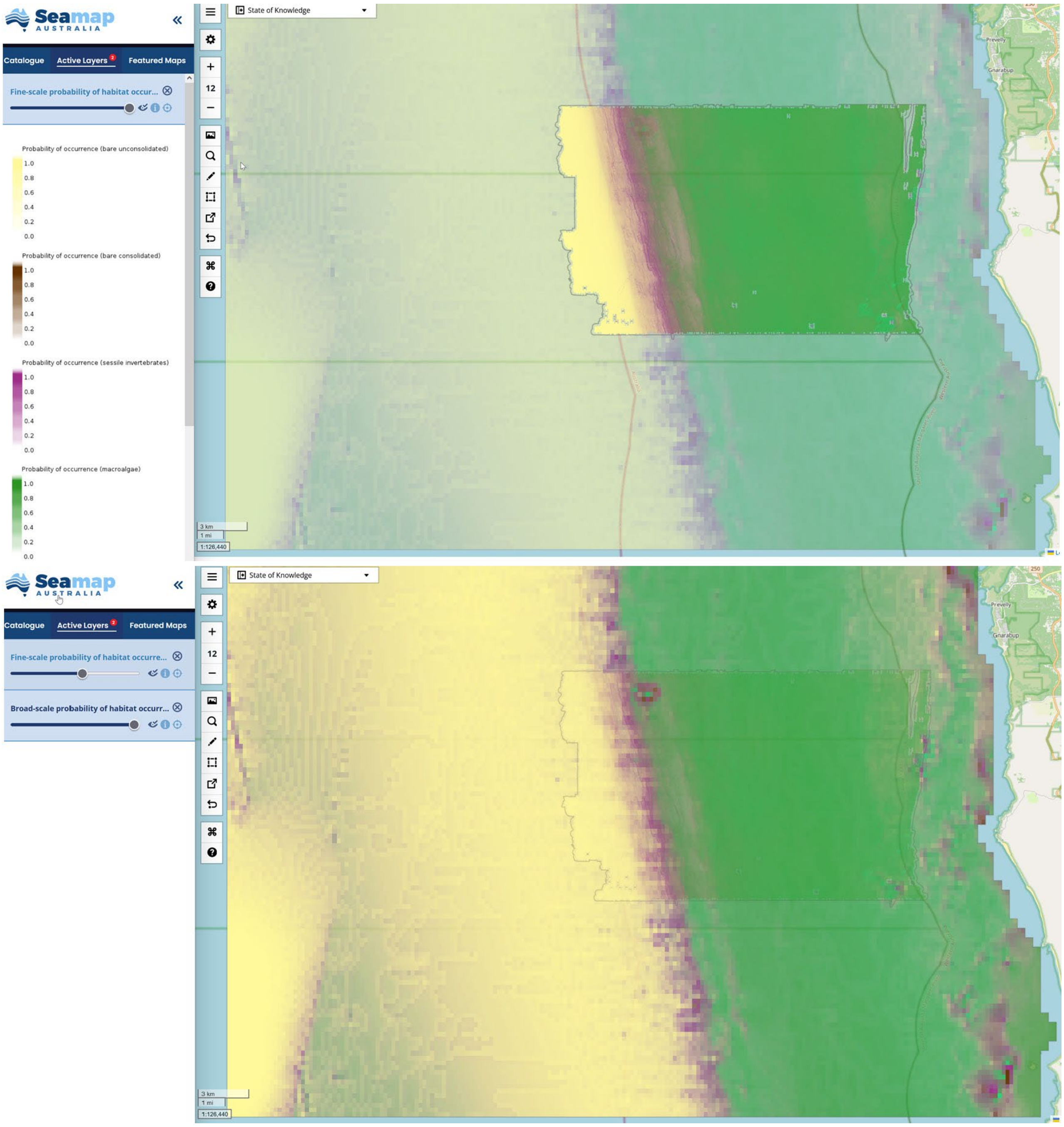


Figure 38 Illustration of the impact on the spatial extent and confidence of benthic assemblage prediction using (a) high resolution bathymetry data (~4 m) or (b) existing low resolution (250 m) national bathymetry products, at the scale of the National Park Zone within the western arm of the South-west Corner Marine Park

Appendix G Seamounts

The text below describes the creation of a new dataset, Seamounts and raised seabed features of the Australian margin (Flukes 2023), in support of the AMP Ecosystems data product.

Seamounts, elevated seabed features, and bathymetric highs constitute ecologically significant habitats for various marine species, often associated with unique ecosystems and high levels of biodiversity. In Australia, the absence of a comprehensive and spatially accurate inventory of these underwater structures has been a notable gap. Existing broad-scale modelling efforts, including the Parks Australia AMP Ecosystems map, have predominantly relied on coarse, global datasets. These datasets, characterized by low resolution, often fail to incorporate recent bathymetric data, leading to the omission or incorrect identification of features (Figure 39). This work aimed to address these limitations by developing an improved dataset of raised seabed features, specifically designed to enhance broad-scale modelling efforts like those used in the AMP Ecosystems mapping.

Seamount features were extracted from a range of data sources (listed below) for the area surrounding the Australian continental margin. These were cross-referenced with GEBCO's 2023 global terrain model (15 arc-second interval grid) and any obviously erroneous features removed. This dataset includes all features located inside the Australian Exclusive Economic Zone (EEZ). Rather than cropping exclusively to this boundary, those features falling outside the Aus EEZ but in the approximate vicinity were also retained for context

Existing feature boundaries were redigitised for areas in which more recent high-resolution bathymetry was available, utilising the 'Bathymetry of Australian Marine Parks (2024)' compilation dataset and individual survey datasets available through the AusSeabed data portal. Where available, fine-scale geomorphic mapping in which seamounts and pinnacles had been classified were extracted and merged with the larger-scale features. If fine-scale mapping disagreed with features classified in the broader-scale datasets, the finer-scale data was prioritised. Where multiple features occurred immediately adjacent to each other, the digitised area represents the "footprint" of the features and as such, a single polygon may encompass multiple peaks.

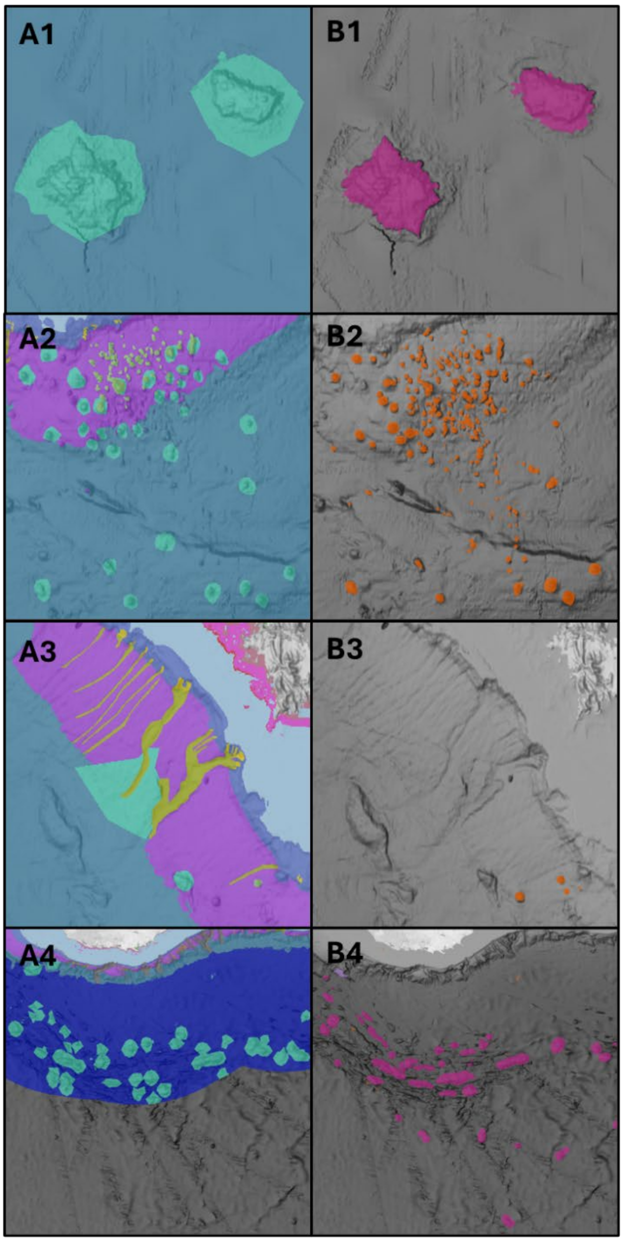


Figure 39. Sample region of AMP Ecosystems data product (left column: panels A1-A4); and corresponding refined footprints of raised seabed features contained in the Flukes (2023) dataset (right column: panels B1-B4). Teal features shown in A1-A4 indicate features currently classified as 'seamount sediments' in the AMP Ecosystems dataset. Pink features in B1 & B4 show refined footprints of seamount boundaries, with orange features in B2 & B3 showing other raised seabed features that do not meet the 1,000m relief threshold of Dove et al. (2020). A1/B1: Significantly improved accuracy of Sulu and Cassowary Seamounts boundaries derived from the geomorphic mapping of Lucieer 2020: https://seamapaustralia.org/map/#fa81b58a-1105-4f58-ba49-59628205a4dd. A2/B2: Improvements to Huon AMP seamount boundaries (although features do not meet the Dove et al. (2020) vertical relief threshold for seamount features) derived from high-resolution MBES and backscatter data of Lucieer 2018. https://seamapaustralia.org/map/#6ad8874f-60b5-470b-b999-478f507a6936. A3/B3: Spuriously classified seamounts (where no features exist) removed from Southwestern Tasmania: https://seamapaustralia.org/map/#f4e8b0b1-8ce3-4b72-be08-6fbb42da2683. A4/B4: Improved accuracy of seamount boundaries from higher-resolution MBES data off southern WA: https://seamapaustralia.org/map/#2fb49879-a2b9-4868-8107-d32626acdbd8

Input data sources:

- The GEBCO_2023 Grid - a continuous terrain model of the global oceans and land. https://www.gebco.net/data_and_products/gridded_bathymetry_data/gebco_2023

- Various bathymetry mapping datasets from the AusSeabed Data Portal (https://portal.ga.gov.au/persona/marine)

- Heap, A.D., Harris, P.T. 2008. Geomorphology of the Australian Margin and Adjacent Seafloor. Geoscience Australia, Canberra. http://dx.doi.org/10.1080/08120090801888669 [seamount/pinnacle features extracted]

- The Global Seamounts Database (KWSMTS v0.1) http://www.soest.hawaii.edu/PT/SMTS/main.html for data access

- Kim, S.-S., and P. Wessel (2011), New global seamount census from the altimetry-derived gravity data, Geophysical Journal International., 186, 615-631. https://doi.org/10.1111/j.1365-246X.2011.05076.x

- Lucieer, V. (2020). Geomorphology and Surface Classification of seven Coral Sea seamounts. Institute for Marine and Antarctic Studies (IMAS), University of Tasmania (UTAS). Data accessed from https://metadata.imas.utas.edu.au/geonetwork/srv/eng/catalog.search#/metadata/a56b2120-9753-4285-9767-f09060f8f3b8 on 10/05/2023.

- Lucieer, V (2018). Geomorphology map of the Huon Australian Marine Park. Institute for Marine and Antarctic Studies (IMAS), University of Tasmania (UTAS). Data accessed from http://metadata.imas.utas.edu.au/geonetwork/srv/en/metadata.show?uuid=6e9916d1-0e9c-4cf0-b08e-e5e631c1aa4d on 10/05/2023.

- Monk, J., Flukes, E. & Lucieer, V. (2023). Geomorphometry of Australian Marine Parks. Institute for Marine and Antarctic Studies (IMAS), University of Tasmania (UTAS). Data accessed from https://metadata.imas.utas.edu.au/geonetwork/srv/eng/catalog.search#/metadata/40e9283b-d4ed-4176-8fe6-112b8697003f on 10/05/2023.

- Beaman, R. (2012). Project 3DGBR: Great Barrier Reef and Coral Sea Geomorphic Features (JCU) [Dataset]. School of Earth and Environmental Sciences, James Cook University [producer]. eAtlas [distributor]. Available from https://eatlas.org.au/data/uuid/25685ba5-6583-494f-974d-cce2f3429b78

- Yesson, C et al. (2011): The global distribution of seamounts based on 30-second bathymetry data. Deep Sea Research Part I: Oceanographic Research Papers, 58(4), 442-453, https://doi.org/10.1016/j.dsr.2011.02.004
  see https://doi.pangaea.de/10.1594/PANGAEA.757563 for 2011 data and https://doi.pangaea.de/10.1594/PANGAEA.921688 for updated 2020 data

- Harris, P.T., Macmillan-Lawler, M., Rupp, J. and Baker, E.K. (2014) Geomorphology of the Oceans. Marine Geology, 352, 4-24. https://doi.org/10.1016/j.margeo.2014.01.011

- Guyots (via Arctic Biodiversity Data Service)
  https://geo.abds.is/geoserver/ebsa/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=guyots&outputFormat=SHAPE-ZIP
- Seamounts (via Arctic Biodiversity Data Service)
  https://geo.abds.is/geoserver/ebsa/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=seamounts&outputFormat=SHAPE-ZIP

Appendix H Coral Sea Oceanic Vegetation

In this section we detail mapping the vegetation in the atoll lagoons of the Coral Sea conducted as part of this project. Previously unmapped and poorly characterised, much of the soft sediment bottoms of these atolls are covered in a diverse collection of fleshy and calcareous macroalgae and small patches of seagrass (Tol, *et al.*, 2023). In the very clear waters of the Coral Sea, it is possible to see this vegetation in carefully processed Sentinel-2 imagery allowing it to be mapped to a scale of 1:400k (up to a 400 m boundary error).

Methods

To map the vegetation in the Coral Sea, the primary data sources used were Sentinel-2 image composites optimized for the marine environment (Lawrey and Hammerton, 2022), high-resolution bathymetry data covering part of the region (Beaman, 2017), and drop camera survey results for validation (Tol *et al.*, 2023).

Most of the vegetation in the lagoonal floors of the atolls in the Coral Sea occur at a depth of 30 - 60 m. At these depths only the blue channel of the satellite imagery provides any useful visual information. Additionally the contrast between bright sand and dark vegetation is very small in the imagery for area at such depths. Artefacts in the imagery due to clouds, sun glint, waves, and sensor noise can easily obscure these small differences.

To reduce the noise in the imagery a pixelwise statistical median composite was used, created from 4-10 of the clearest Sentinel-2 images of each scene manually selected from 2016-2021. Cloud masking and sun glint correction were applied before image composition (see Lawrey and Hammerton, 2022 for full details).

The atoll lagoonal areas were classified manually, and hand digitised as bare, vegetation or reef based on the estimated benthic reflectance. Lighter benthic regions were assumed to be bare sand, while darker regions assumed to be vegetation or reef features. Determining the benthic reflectance at such depth from satellite imagery is potentially ambiguous as areas might appear dark because they are deep, covered in vegetation or reef, affected by coloured dissolved organic matter in the water column absorbing light, or there is a tonal shift from different satellite sensors across the image swath. These factors make image interpretation challenging.

To resolve some of these confounding factors the mapping was done using visual cues to identify reference points across the scene to help compensate for tonal and contrast shifts due to depth, water clarity changes and the satellite sensor. These visual cues identify features where there is a high confidence in the benthic cover (sand or vegetation) and these act as local references for classifying the rest of the area between these reference locations. As most areas of the coral atoll lagoons are gently sloping, rapid changes in visual brightness are typically caused by changes in benthic reflectance, rather than changes in depth. We use this to find the edges of vegetation regions.

We employed the following multi-step process to map and verify the oceanic vegetation:

1. **Identifying Visual Cues:** We identified a set of potential visual cues to detect likely vegetated areas. These cues relied on distinguishing probable patches of sand to

estimate local depth and water conditions and observing transitions between light and dark regions to identify vegetation boundaries.

2. **Manually mapping:** The vegetation boundaries were manually hand digitised based on visual cues in the satellite imagery for Flinders and Holmes Reefs.

3. **Benthic Reflectance Estimation:** We developed benthic reflectance estimates for the North Flinders and Holmes Reefs regions using both high-resolution and accuracy bathymetry (Beaman, 2017) and satellite imagery (see Lawrey, 2024a for details).

4. **Compared Analysis:** The initial vegetation mapping from step 1 was compared against the benthic reflectance from step 3 to identify the most reliable visual cue techniques. This identified which were most robust against changes in depth and tonal shifts.

5. **Coral Sea Mapping:** Using the insights from the previous steps, we manually mapped the remaining Coral Sea region using only satellite imagery.

6. **Validation:** The final map was validated against the available drop camera survey data on Lihou and Tregrosse Reefs.

We previously, separately mapped reefs (Lawrey, 2024c). This mapping was used ensure that reef areas were not interpreted as vegetation. Reef areas were determined by their granular visual texture, their elevated central region, and by the grazing halos around their base.

Visual Cues for Benthic Cover Identification

The following is a summary of the key visual cues used to classify the areas as either vegetated or unvegetated.

1. **Grazing Halos Around Patch Reefs:** Grazing halos appear as pale rings of bare sand surrounding a textured dark, rounded feature (patch reef), see Figure 40 for examples. These occur because herbivorous fish forage and clear the surrounding sand of any algae. While grazing halos are well studied in shallow reef systems, (DiFiore *et al*., 2019) they are not well studied at the depths seen in the Coral Sea atoll lagoons. In this mapping we, however, assume the grazing halos in the Coral Sea are caused by a similar mechanism and thus where we see them, they indicate a central reef structure surrounded by a sandy area that is largely devoid of algae. Based on a review of bathymetry transects of patch reefs in North Flinders reefs, the depths of these grazing halos tend to be very close in depth to the surrounding lagoon. This allows them to act as an excellent reference for the brightness of sand at the depth of the lagoon in the area near the reef. Frequently, dark halos of dense vegetation surround these grazing halos, serving as a brightness reference for high-density vegetation.

2. **Atoll Plains:** On the atoll plains, particularly on the western side of Tregrosse Reefs platform there are large patches of dark substrate that have pale patches, unrelated to the presence of reefs. In this case, local tonal references were identified at locations where there was a clear step change in brightness and the shape and texture of the dark areas matched typical patterns of algae seen in other regions.

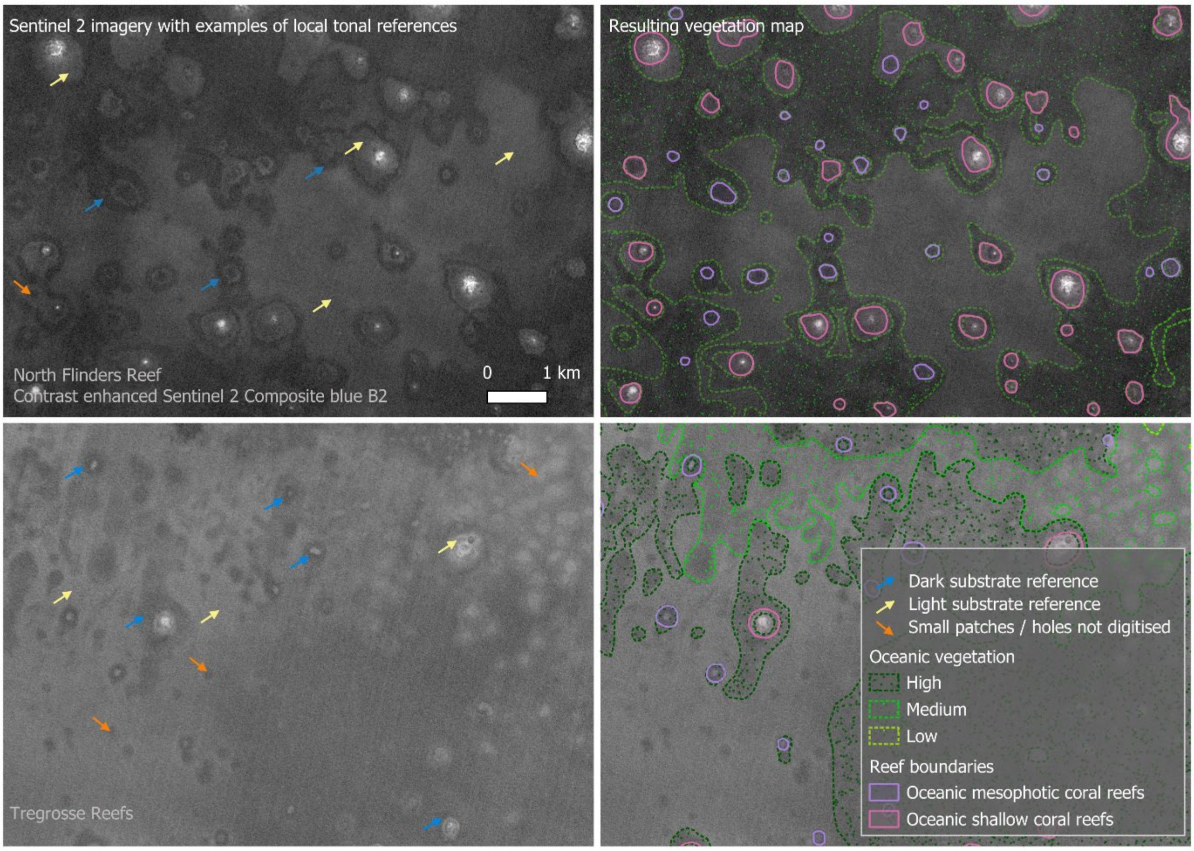


Figure 40 Using visual cues to classify the contrast-enhanced Sentinel 2 composite blue band (B2) images into vegetation and sand in the atoll lagoons of the Coral Sea. The top panels display a section of North Flinders Reef, while the bottom panels show part of Tregrosse Reefs. Local tonal references are used to classify the imagery into sand and vegetation. The left panels show examples of local tonal references with a light substrate (yellow arrows) and dark substrate (blue arrows). These correspond to locations where there is high confidence in the vegetation density. The right panels show the resulting digitised vegetation map, along with reef boundaries. The vegetation was digitised at 1:400k scale, so some small details are not captured (orange arrows).

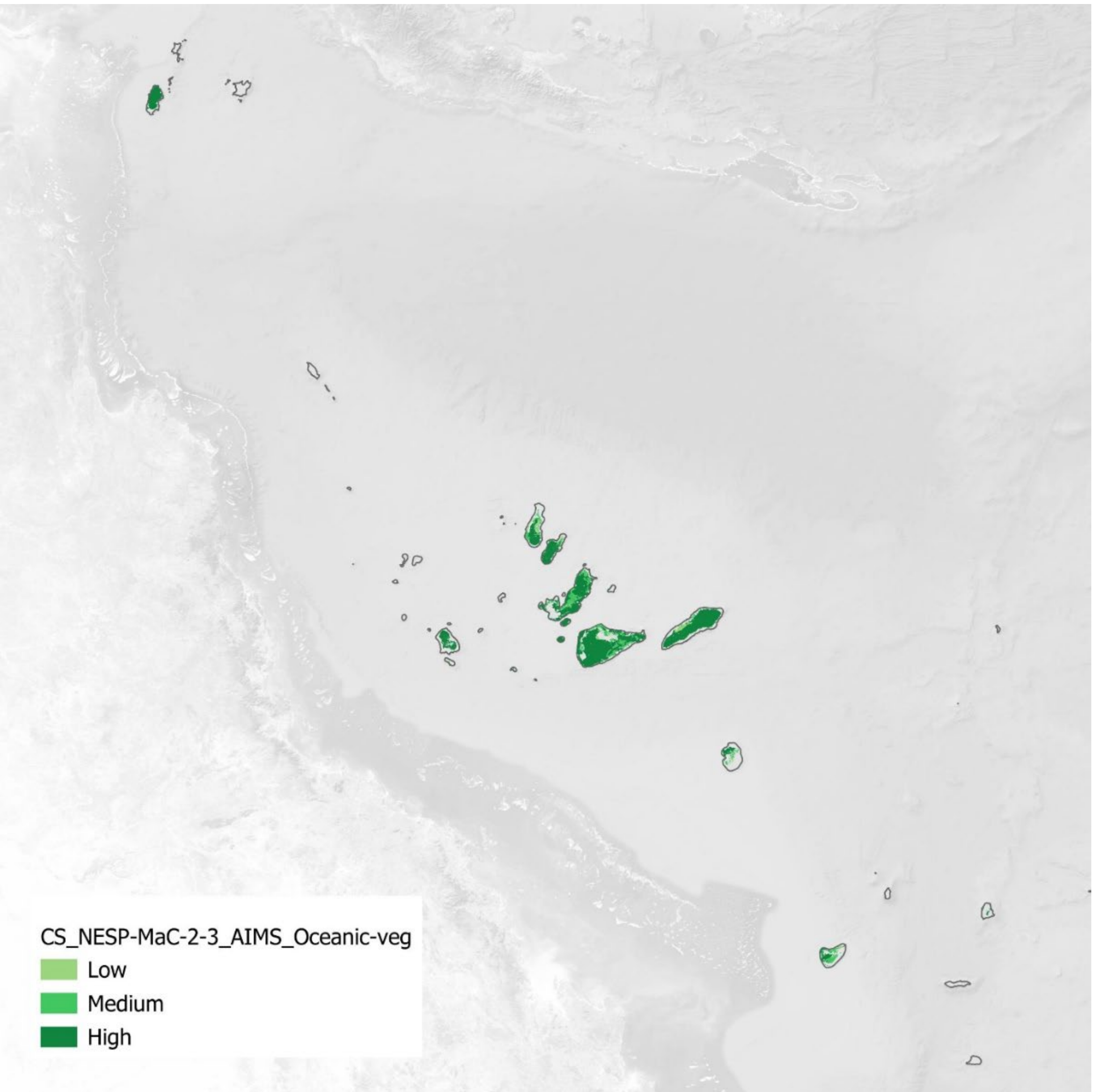


Figure 41 Vegetation mapped on the banks and coral atoll lagoons across the Coral Sea.

Table 20 Area of the mapped oceanic vegetation in the Coral Sea. It should be noted that the accuracy in distinguishing between the different density categories (low, medium and high) is low. The primary focus of the mapping was determining the extent of vegetated verses sandy areas, not in accurately estimating the absolute density of vegetation.

Vegetation Density	Area in Coral Sea (km^2)
Low	360
Medium	1409
High	7192
Total	8961

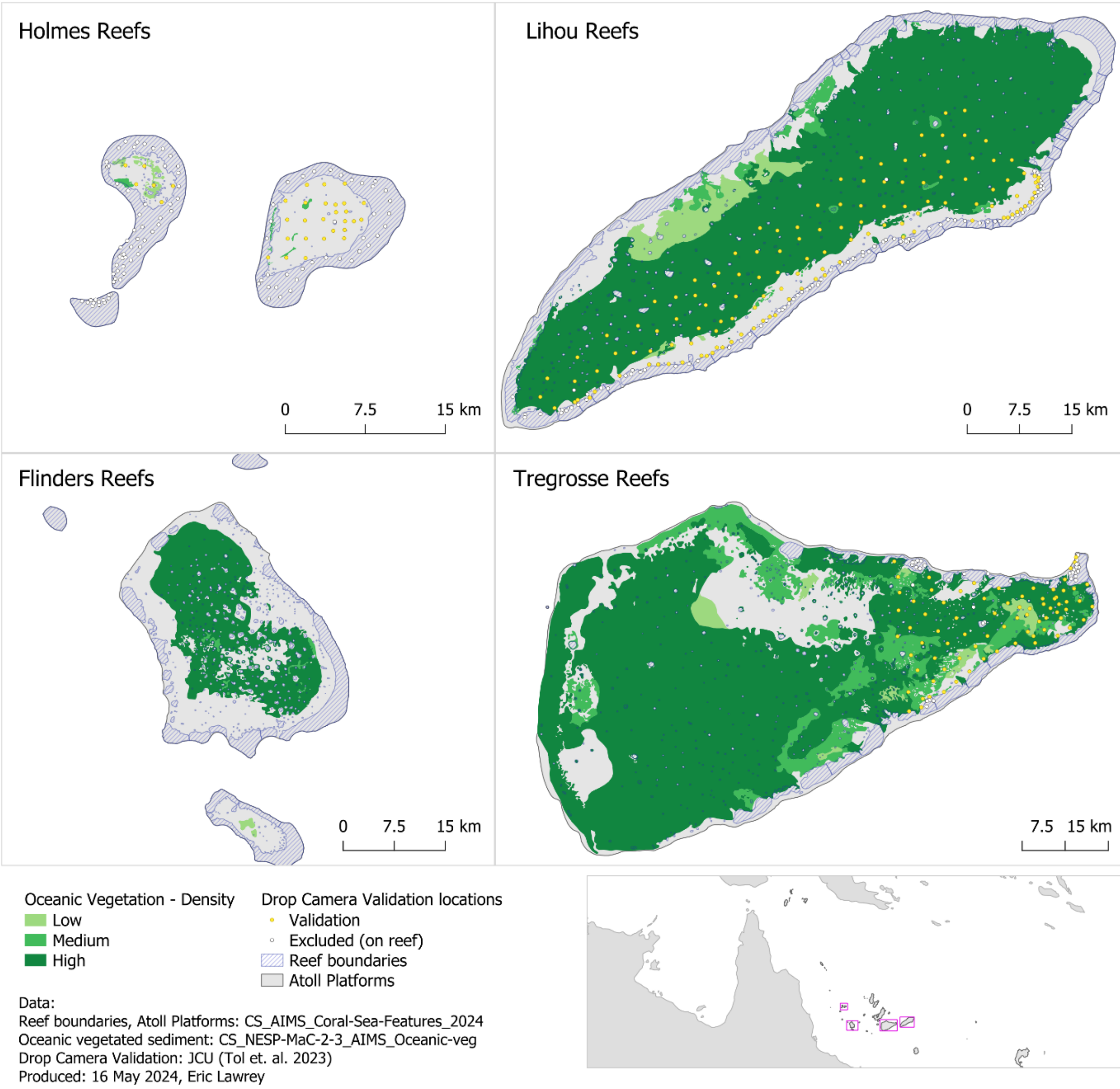


Figure 42 Oceanic vegetation on various coral atolls of the Coral Sea. Drop camera validation locations are shown as yellow dots.

Results

Most of the atolls with lagoons deeper than 30m were found to have vegetation covering much of the lagoonal floor.

Validation

Since this dataset was manually mapped from noisy and ambiguous imagery, validating this visual mapping approach was essential. The first form of validation was comparing the mapped vegetation boundaries with the benthic reflectance of Flinders reef. This comparison showed a very strong alignment between the manual visual mapping and benthic reflectance (Lawrey, 2024a), with the main deviations occurring around reef edges where the digitised vegetation did not capture all the details. It also deviated in areas where the benthic features were harder to see due to lower water clarity caused by coloured dissolved organic matter increasing the water column light absorption.

No significant adjustments were needed to the visual cue approach following this comparison. However, it highlighted the importance of identifying the local visual cues to compensate for varying depths, satellite sensor brightness shifts and changes in the water clarity.

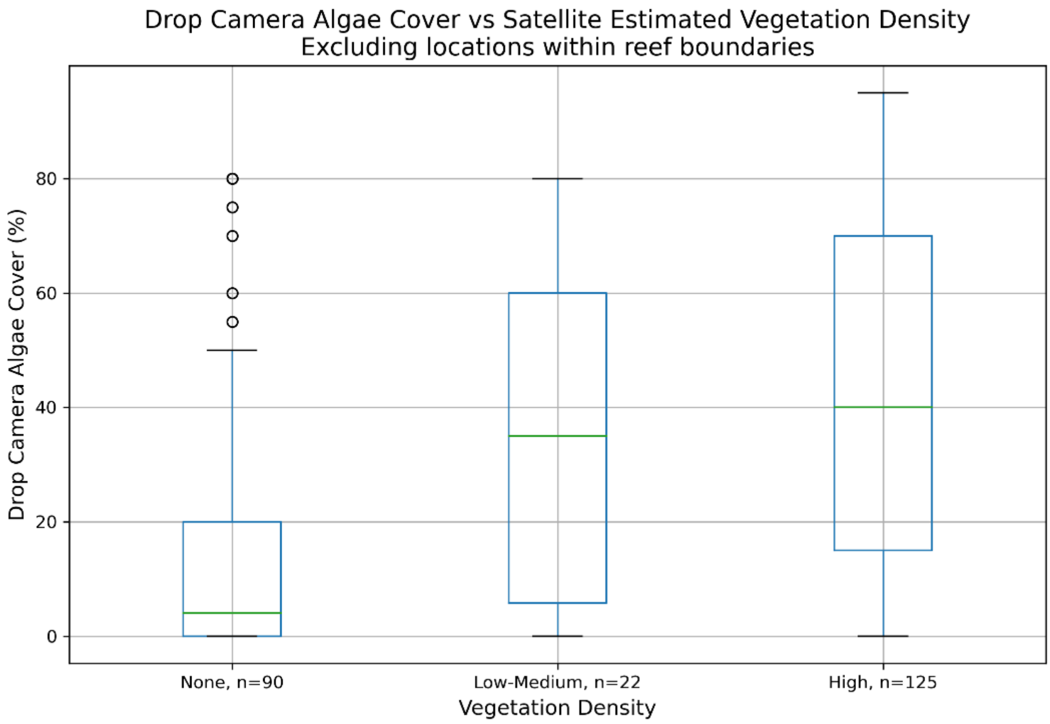


Figure 43 Comparison between drop camera estimated benthic cover and the vegetation density category mapped from satellite imagery. Low and Medium classifications were grouped due to the small number of samples in both categories.

The final validation involved comparing vegetation maps of Holmes, Tregrosse, and Lihou Reefs against the results of a drop camera survey conducted by JCU in December 2022 (Tol *et al*., 2023). The locations of the validation sites are shown in Figure 42. From this survey, 237 locations overlap the atoll lagoons. Figure 42 compares the vegetation density estimated from satellite mapping with the benthic cover assessed through the drop camera survey. This demonstrates a strong relationship between the mapped vegetation density and the benthic cover measured by drop cameras. The data show considerable variability, possibly due to fine-scale vegetation patchiness not captured by the satellite-based mapping. The drop camera results represent very small survey patches (less than 1 m across), while the satellite mapping represents patches around 400 m across.

Areas identified as having high benthic vegetation in satellite mapping showed 15-70% (average 40%) algal benthic cover. In contrast, lagoonal regions mapped as sand (outside the identified vegetation but not on reefs) had significantly lower algal benthic cover, ranging from 0-20% with an average of 4%.

Conclusion

This mapping process provided the first comprehensive estimate of the submerged aquatic vegetation habitat types in the atoll lagoons of the Coral Sea. Despite the inherent challenges due to the habitat's depth being at the very limit of what can be mapped from satellite imagery, the resulting vegetation map aligned well with the limited available validation data. This mapping could be further improved in the future with additional satellite image data, high resolution bathymetry in lagoonal areas, and well distributed field data. This new dataset should be incorporated into any update of the AMP Ecosystem layer to ensure that this large and significant habitat is represented.

National Environmental Science Program

CONTACT

Name: Keith Hayes, Peter Walsh

Email: keith.hayes@csiro.au; peter.walsh@utas.edu.au

This project is supported with funding
from the Australian Government under the
National Environmental Science Program.