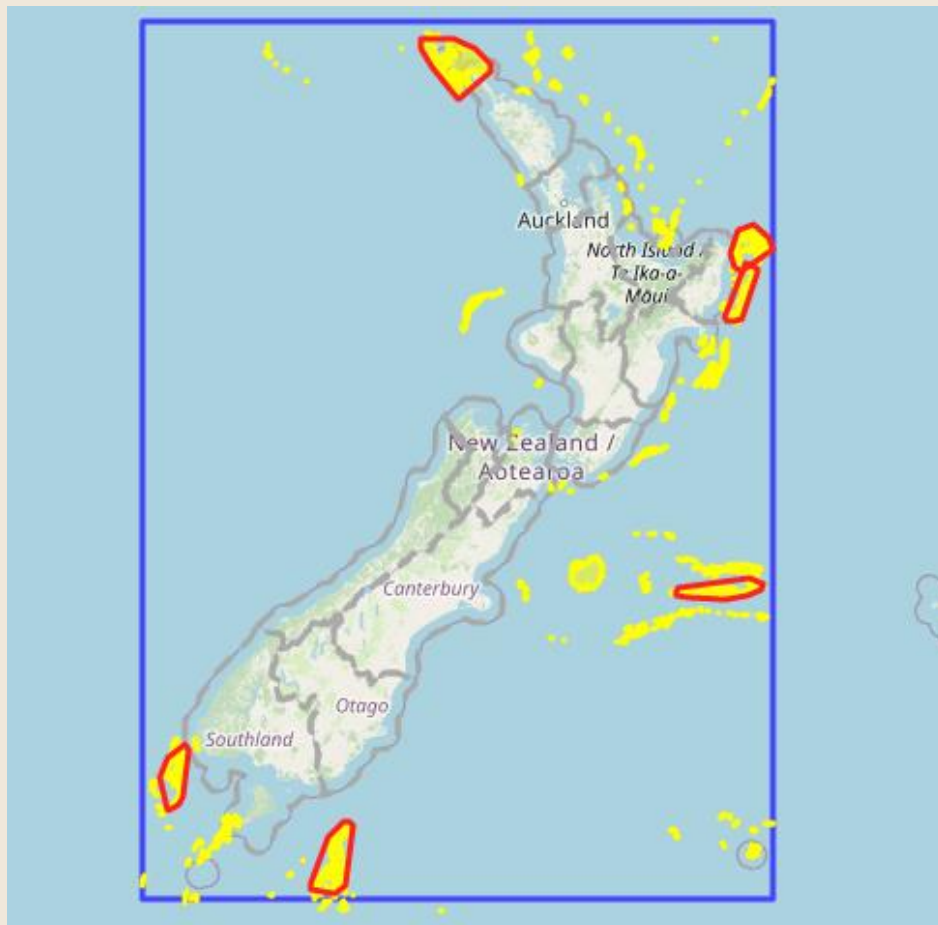# Enabling Integrated Marine Management

DOC use case: Mapping Marine Key Ecological Areas

WildCoast: Amélie Augé
DOC: Enrique Pardo, Shane Geange. Jodie Robertson

Department of
Conservation
Te Papa Atawhai

New Zealand Government

CONTENTS

1 Summary

Aotearoa New Zealand is nationally and internationally committed to manage its marine biodiversity; however, there is currently inconsistency on how different government bodies apply regulatory frameworks for managing marine and coastal biodiversity. One of the fundamental aspects to solve is how to effectively share existing data and allow potential users (e.g., regional councils, government agencies, etc) to analyse it in consistent ways. Te Papa Atawhai Department of Conservation has been developing a set of criteria for identifying Key Ecological Areas (KEAs) that can inform decisions related to the management of marine biodiversity. The DOC use case developed an online tool to delineate candidate KEAs within the marine environment using a set of criteria thresholds through the Datamesh. The use case explored how the Datamesh technology can support the development of an open-access analytical tool to delineate KEAs at national and regional scales. DOC, the National Institute of Water and Atmospheric Research (NIWA) and Oceanum collaborated during the use case to: (1) identify a series of test KEA criteria and sample of datasets to generate analyses; (2) link all required input datasets for the analyses to the Datamesh; (3) develop the analytical tool; and (4) conduct user testing internally and with regional councils.

The main output of the use case was a prototype online application, the 'KEA App'. The KEA App is Datamesh-enabled and fully performs the mapping of KEAs. The KEA App brings together data from multiple source servers (including the partner agencies' data portals) via the Datamesh without requiring downloading the data. The KEA App and Datamesh were extensively tested with key stakeholders such as regional councils, which showed strong interest and support for future development of these tools.

The use case was successful in showing how the Datamesh can increase connections between existing GIS web services, allowing users to query data from multiple providers while also providing access to user-friendly online analytical tools that can be run across multiple datasets. The use case illustrated how this technology can further improve the reuse of marine data and increase accessibility to standardised analyses to support marine management.

2 Problem

In 2020, Te Mana o te Taiao - the Aotearoa New Zealand Biodiversity Strategy was published, which shares a vision of ecosystems and species being protected, restored, resilient and thriving. A key to implementing several marine-related outcomes within Te Mana o te Taiao is the identification of areas that are important for marine biodiversity. To aid with this, a cross-agency Marine Protected Area Science Advisory Group (MSAG) developed a list of nine criteria for describing Key Ecological Areas (KEAs) that could be used as inputs into management planning (Figure 1). Subsequent contracts to NIWA (DOC investigation numbers 4735 and 4759) compiled 27 ecological datasets describing nine KEA criteria and evaluated the datasets for their spatial and taxonomic comprehensiveness, uncertainty (particularly for modelled layers), and gaps. However, while these investigations provided an information base that can be used to describe KEAs, the additional step of mapping where KEAs occur had not been taken. Such an exercise could be valuable for informing spatial planning processes and addressing biodiversity conservation at the site scale (i.e., at the level of individual marine protected areas, concessions and 'ocean' management units). Until this is achieved, it is likely that resources will continue to be spent on replicating similar data stocktakes and methodologies by various agencies, with data and analytical methodologies remaining restricted and not shared and reducing the benefits for Aotearoa marine management.

In recognition that there are a growing number of ways for delineating important sites for biodiversity, the IUCN have recently finalised and approved a harmonised approach that uses thresholds to map areas of biodiversity importance. These thresholds, with some modifications, can be used to map KEAs within the Aotearoa marine domain. However, a standing problem in mapping KEAs is how to access data to map KEAs and how to share agreed and standardised mapping methodologies among stakeholders, government, and partners so that they can inform management processes.
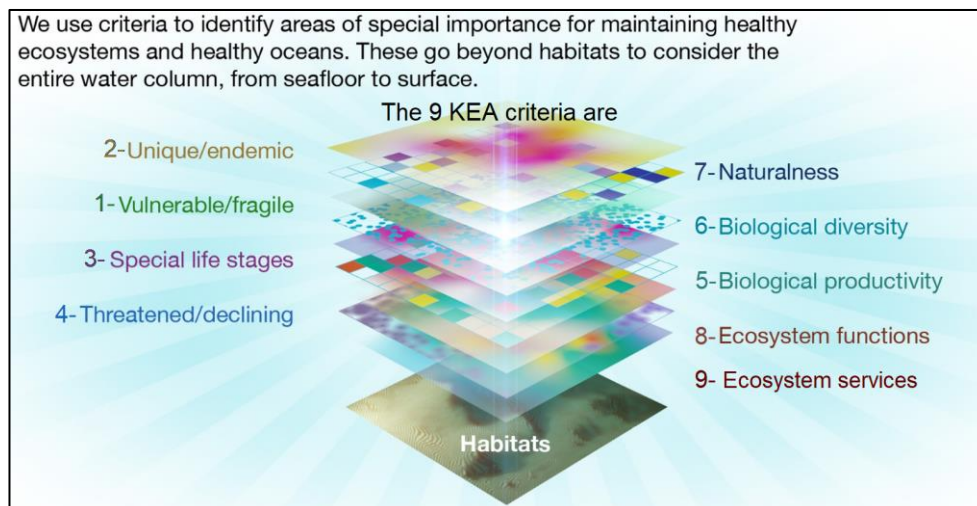

Figure 1: The nine Key Ecological Areas (KEAs) criteria as defined by Te Papa Atawhai - Department of Conservation

3   Purpose and objectives

The DOC use case aimed to integrate a set of habitats and species distribution layers into the Datamesh to enable the mapping of Key Ecological Areas (KEA) within the marine environment through the use of IUCN thresholds. This use case aimed to demonstrate how users can query data from multiple providers and stream it into analysis platforms to inform marine and coastal planning. The use case evolved into the development of a user-friendly online analytical tool that can be run across multiple datasets to map KEAs, the 'KEA App'.

Specific objectives of the DOC use case were:

- *To test the connectivity of different geospatial datasets with the Datamesh, including some from the DOC Marine Data Portal*

- *To further develop a methodology for mapping Key Ecological Areas at regional and national scales*

- *To develop an app that uses Datamesh technology to implement methods for mapping Key Ecological Areas at regional and national scales*

4   Method

The DOC use case was based on the previous work completed by NIWA under contract to DOC, which included the identification of a set of datasets as inputs for mapping KEAs (Lundquist et al 2020). DOC worked with NIWA and Oceanum to develop the methodology for mapping KEAs and testing integration of the methodology with the Datamesh technology. The following steps were followed during the use case to explore how a Datamesh could be used to map KEAs at various spatial scales:

Step 1. Developing the methodology to map KEAs

Step 2. Linking input data to the Datamesh

Step 3. Developing the Datamesh-based KEA analytical tool (the 'KEA App')

Step 4. User testing of the KEA App and Oceanum Datamesh

4.1   Step 1. Developing methods for mapping Key Ecological Areas (KEAs)

Previously, nine KEA criteria were developed by the MSAG. Datasets relating to each of the nine criteria were identified and potential methodology for using these datasets to map KEAs using IUCN thresholds were developed by NIWA under contract to DOC (Lundquist et al 2020, Stepheson 2018, Stewart-Sinclair et al 2022).

As part of the DOC use case, NIWA was contracted to develop and test the methodology for mapping KEAs. A subset of existing geospatial layers for three of the nine KEA criteria were selected for this purpose. The selected KEA criteria for the use case were Criterion 1- Vulnerability, fragility, sensitivity, or slow recovery; Criterion 3- Special importance for life history stages; and Criterion 6- Biological diversity.

For KEA Criteria 1 and 6, the input datasets for the mapping were developed earlier by NIWA (Stepheson et al 2018). These data included Species Distribution Models (SDMs) which in this case specifically refer to predicted Habitat Suitability Index (HSI) models, for 583 species of marine fish, invertebrates and macroalgae. Each SDM was associated with a corresponding uncertainty spatial layer that estimates the accuracy of the HSI predictions within each cell.

For KEA Criterion 3, the input dataset was called 'Finfish Spawning Areas' and owned by Fisheries New Zealand. This dataset contains polygon shapefiles for the known spawning areas of 39 species of finfish, based on fishery observer data, catch data and expert knowledge.

The datasets included within the analysis for the three KEA criteria were selected because they were, or could easily be made, available in open-access data portals. NIWA developed R scripts to run the analyses to delineate KEAs and tested the analyses at different spatial scales. The details of the methodology to map KEAs, and the full descriptions of the datasets stated above, are available in the NIWA report (Bennion et al. 2022).

4.2   Step 2. Linking KEA input data to the Datamesh

Once the input data for the use case were identified, they were made available through existing APIs (e.g., on an open portal online). DOC prioritised the release of data for the use case that it held through the new DOC Marine Data Portal to allow integration with the Datamesh.

In total, 1168 datasets were linked from the DOC Marine Data Portal and 46 datasets were linked from other data portals (see Appendix 1). During the acquisition of the data, it was apparent that no or limited metadata was being incorporated from the source server into the Datamesh; consequently, some metadata fields (e.g., titles, descriptions, keywords) had to be completed manually. If the Datamesh progresses beyond a proof of concept, data suppliers (agencies, research bodies, universities) will need to be consistent and systematic in completing metadata fields in their datasets to meet the findability purpose. Once datasets were linked and metadata fields completed, the Datamesh provided flexible and innovative ways of managing the datasets (e.g., selecting multiple layers, grouping layers etc). For instance, Oceanum created bundles of 'virtual datasets' with the pre-defined groups of SDMs, based on the taxa selected for a KEA Criterion (example for the 'reef fish' group, Figure 2).
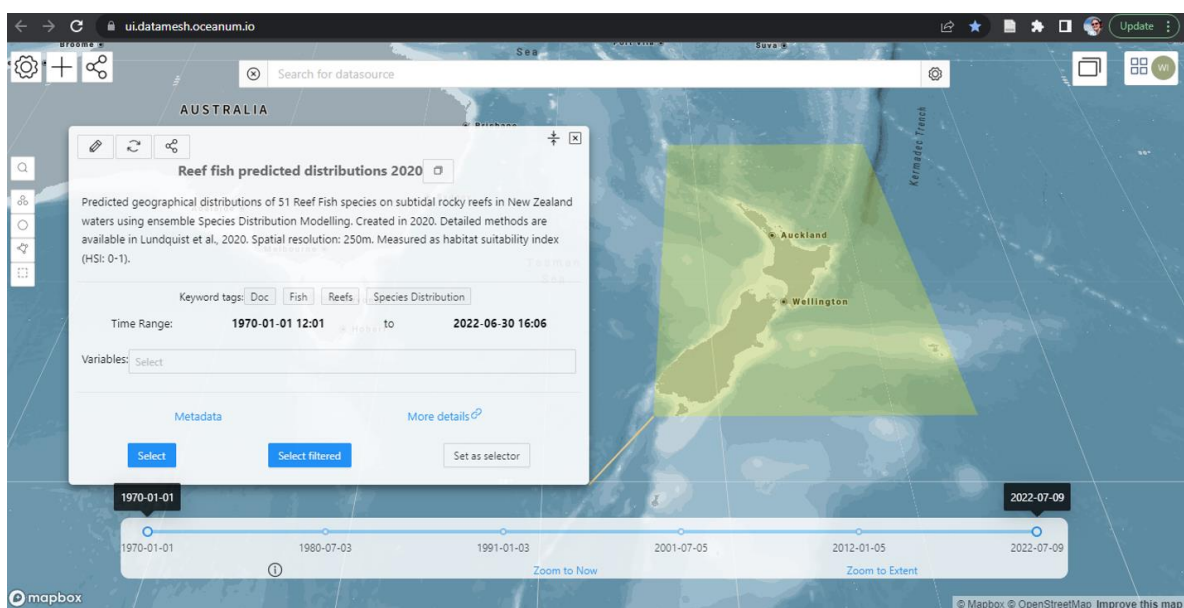
Figure 2: Example of a 'virtual dataset' of SDMs (HSI) for 51 reef fish linked into the Datamesh from the DOC Marine Data Portal.

In addition, administrative boundaries, coastline layers and similar nonspecific datasets were linked to the Datamesh. These were necessary for running the mapping methodology or interpretation of the mapping results and, at the same time, testing the compatibility of the Datamesh with a range of data and formats. DOC found the Datamesh intuitive and user-friendly with helpful pop-up notifications.

4.3 Step 3. Developing a Datamesh-based KEA analytical tool (the 'KEA App')

The output from Step 2 was an R script that uses thresholds to query inputs data layers and create maps for three KEA criteria. These R scripts and data layers were then available for use within the Datamesh. The purpose of Step 3 was to create an analytical tool that linked to the data in the Datamesh and operationalised the R scripts. Initial development of an analytical tool required the datasets to be downloaded, which the DOC use case team identified as a risk of duplicating datasets and losing traceability of the data. Oceanum suggested the development of a web-based application to run the R script analysis using the data directly from the Datamesh. Oceanum developed the prototype 'KEA App' for the DOC use case (Figure 3), which runs under a Python code and calls the data directly from the data source through the Datamesh.
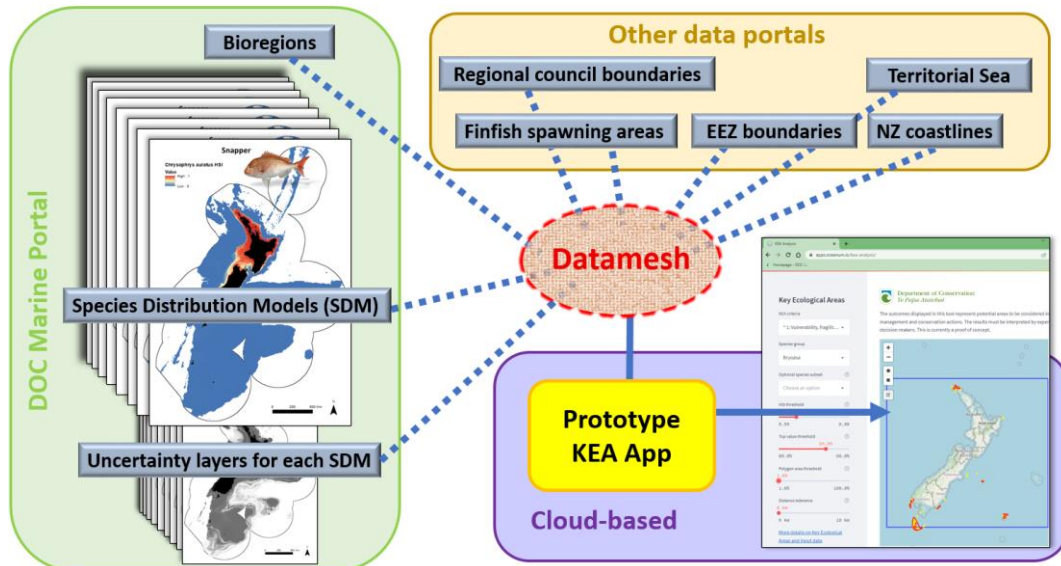
Figure 3: Conceptual model of the Datamesh-enabled KEA App

During the development of the KEA App, the ability to easily customise the mapping of KEAs was identified as important for engaging with stakeholders and exploring different management options. A series of parameters were integrated in the App that allows the user to control the sensitivity of the mapping, which increases the flexibility of the App to inform different management objectives (e.g., application at different spatial scales, Figure 4). The parameters that the users can define include analysis area, the KEA criterion that is being mapped, and the

species or habitats to map against the criterion. An additional four parameters, which are constrained within 'acceptable' limits, allow user to customise the analysis by controlling the sensitivity of the mapping. The first two, "HSI values threshold" and "Top value threshold" allow the users to define how inclusive the mapping exercise is in defining KEAs (i.e. higher values map increasing important KEAs). The other two parameters "Polygon area threshold" and "Distance tolerance" allow the users to control the size and aggregation of KEAs to ensure they result in practical management units (see Bennion et al. 2022 for more details).
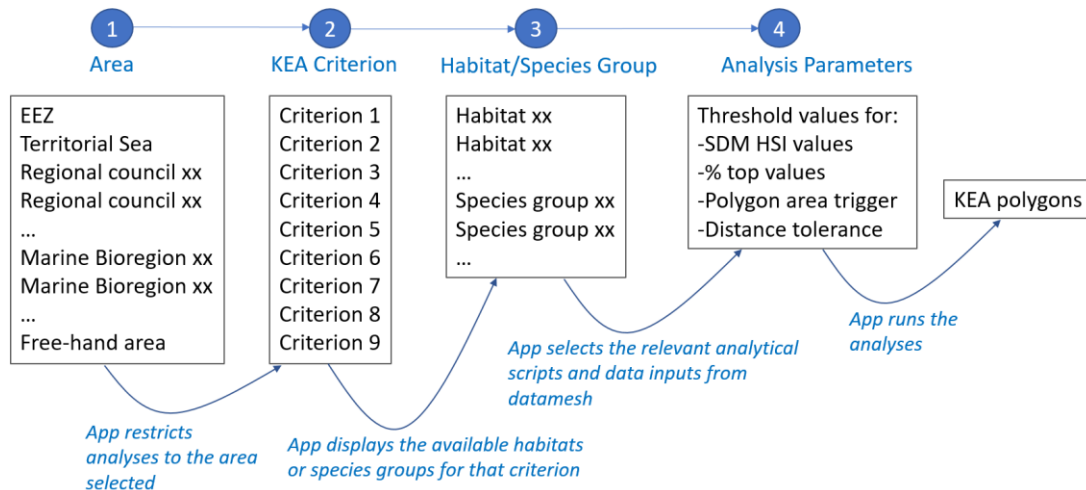
Figure 4: Sequential selection of parameters by the user in the KEA App to run KEA analyses. The user selects preferred parameters in sequence from steps 1 to 4 before running the analyses.

KEA App User Interface

The KEA App is accessed via a URL and is hosted on the Oceanum platform with a simple user interface (Figure 5) for the purpose of the DOC use case. The left-hand side panel displays the parameters for users to select and manipulate during the analyses. The mapping panel, on the right-hand side, allows users to select the area of analysis, zoom in and out, and visualise the mapping outputs (cells of potential KEAs in yellow, and candidate KEAs in red polygons). Underneath the mapping panel, three export buttons allow the users to download the outputs. The users can easily visualise the sensitivity of the analysis to the selection of different parameters.

Figure 5: User interface of the prototype KEA App. Blue polygon is the analysis area selected by the user; yellow areas are KEA cells; red polygons are candidate KEAs.
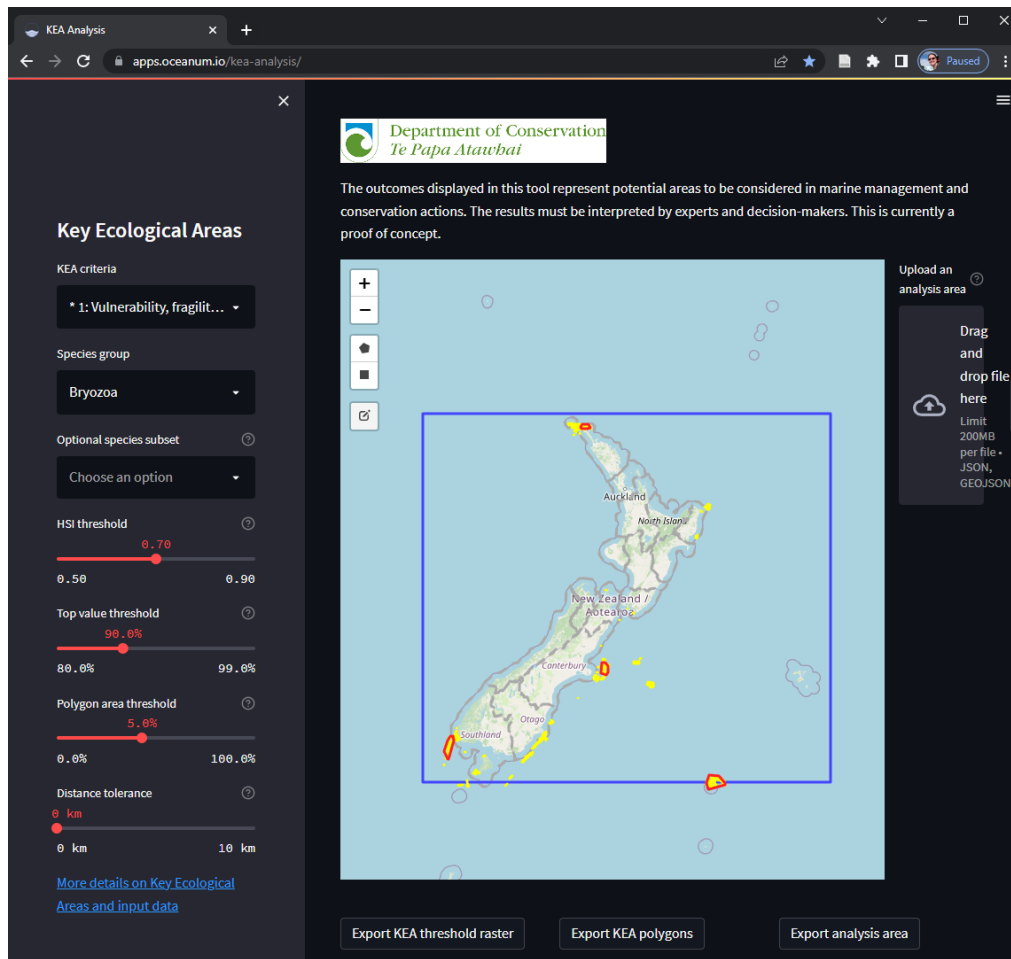
Benefits identified

While users of the KEA App do not interact directly with the Datamesh, it is important to note that the Oceanum platform enables the analytical tool by providing access to the data. The DOC use case demonstrated how a Datamesh could transform the way we access and use spatial data, not only within our own GIS systems, but also through the development of user-friendly online tools. A real-world application of this tool would improve the reuse and accessibility of marine data for a wide range of users, independent of level of expertise in spatial data.

One of the benefits of having an analytical tool performing analyses on data linked in a Datamesh is not having to download the data. The data remain on the original server, facilitating version control of data by the data custodian. Any changes to the data, by data custodians, will automatically be 'read' by the Datamesh into the KEA App.

Importantly, the Datamesh-based analytical tools are independent of the search function of a Datamesh. This means that if the Datamesh is not used as a central search place for marine data, it would still provide an effective platform for developing analytical tools such as the KEA App. The DOC use case demonstrated how the Datamesh environment has the potential of providing stand-alone analytical tools linked to the data in their source servers.

4.4   Step 4. KEA App and Datamesh user testing

With further development, the DOC use case can evolve in a useful set of tools for marine management. To help guide future development, potential end users were contacted to test the Datamesh and the KEA App. Four workshops attended by 21 people from ten organisations including eight regional councils, NIWA and DOC staff were run to introduce the DOC use case and test the prototype KEA App. Feedback on the Datamesh was also recorded during these workshops and is reflected in Table 1 and the overall recommendations.

Table 1. Key feedback from user testing

Oceanum Datamesh Interface comments	• Overall intuitive and user-friendly • Not enough details about the data in the descriptions (e.g., source of the data needing to be made clearer) • Filter should include 'by source repository' • Not enough metadata came with the downloaded datasets after export • Need for a function where the user selects the projection for the exports • Need for a way for the user to see which agency has linked their data
Issues and risks of a Datamesh	• Poor tagging and descriptions of datasets create a risk that a Datamesh does not provide excepted benefits • Need clear labels for Owner/Authority/Sharing license/Attributions. Significant risk of having the data downloadable directly from the Datamesh without people sufficiently checking this (data can be open access but attribution needs to remain for instance). • The concept of a Datamesh will only work if there are guidelines or mandates for agencies
Benefits and opportunities of a Datamesh	• The Datamesh would save time by being a single point for searching marine data and the ability to create and save data packages • A Datamesh could push agencies to share their data
KEA App-specific feedbacks	• There was a consensus that it could be a useful tool with good potential applications and clear interest in further development • The development of publicly available tools using the Datamesh could transform the ability of whanau/hapū/iwi, regional councils and other stakeholders to access data and analytical tools and undertake mapping exercises. • The regional councils all noted that it represents a great potential tool for engagement • The main limitations noted were the ability to view and understand the input data coming from the Datamesh for the analyses, and some improvements required in the formats of the exports (notably for the level of metadata included)

5 Outputs

The DOC's use case achieved all set objectives for the overall Datamesh Proof of Concept. Below, each objective is assessed against the DOC's use case findings. Key outputs achieved were listed, along with some key lessons learnt.

Table 2. Summary of DOC case finding against Proof of Concept Objectives

Objective	Findings	Result
Connections between existing GIS web services and/or datasets (without replicating the data across to this system) from the four partner organisations supporting faster and lower-cost computation of bulk data than existing systems.	**All data needed for KEA analyses linked successfully to the Datamesh.** **The Datamesh allowed KEA analyses to run directly from the data source via the Datamesh, providing lower-cost computation of large datasets.**	**Achieved**
Interoperability of various data formats, standards, and scales.	**All types of data successfully integrated (from various platforms, in several formats)**	**Achieved**
A web service (conforming to OGC API standards) enabling users to query and stream data into their own GIS systems and analysis platforms.	**The KEA App developed in the use case is Datamesh-based and queries, streams, and uses datasets through the Datamesh.**	**Achieved**
Safe and secure access to sensitive data repositories.	**This was not applicable directly to the DOC's use case as the fundamental principle was to create a fully open-access analytical tool. The function was, however, successfully tested during development.**	**Achieved**
A spatial catalogue viewer with query functionality such as searching and filtering.	**The Datamesh interface provided a useful and user-friendly spatial catalogue with good search, filter and download functions**	**Achieved**
Access to analytical tools (source code) using open formats.	**Successful development of a prototype KEA analytical tool (the KEA App). Open-access and web-based, using data through the Datamesh.**	**Achieved**
Insights/models from analytical queries run across multiple datasets.	**KEA analyses use multiple large datasets from different sources, that are run and visualised using the KEA App.**	**Achieved**
Performance of the system meets partners' and users' needs.	**The Datamesh provides an excellent platform to improve marine and coastal management, using data from different sources through a single portal. The Datamesh enables the use of data from their original source, maintaining version control by data custodians.**	**Achieved**

Note: see DOC use case Project Outcomes in Proof of Concept report for more details

6   Key findings

The Datamesh provides an intuitive and innovative solution to access standardised data from different organisations and enable the DOC use case purpose of mapping KEAs. The DOC use case showed how a data mesh can be transformational in how marine spatial data are accessed in Aotearoa New Zealand providing a single access point to a range of data held on different data servers. The set of flexible and functionality data architecture provided by the Datamesh could save time and resources if fully operationalised, which ultimately could help create efficiencies in marine management. The Datamesh allows the user to create packages of data, download data from different organisations at once, and develop new datasets, among other things.

The KEA App also represents an innovative way of running analyses on spatial data, while maintaining the data on source servers The KEA App can help to improve coastal planning initiatives with a user-friendly interface to map KEAs.

The two significant limitations identified during the use case were: (1) The Datamesh partially incorporated metadata from the linked datasets, and (2) Data provider organisations need adequate data management systems and completed metadata. The DOC use case noted that these can be overcome with proper planning and governance in place to drive the implementation of the Datamesh.

Overall, the DOC use case provided an excellent test of the functionality of the Datamesh. Use case' goals were achieved, as all data were successfully linked and the mapping of KEAs was successful, with the additional achievement of building the prototype KEA App. The use case also demonstrated that the Datamesh or a similar solution can provide a range of benefits to government departments, including regional councils, and improve marine management in Aotearoa New Zealand.

7   Recommendations

In developing the DOC use case and KEA App several improvements for the Datamesh and the KEA App and potential next steps were identified.

Table 3: Summary recommendations

Recommendations for improving the Datamesh
Better acquisition of metadata from source servers
Better description of the functionality in the Datamesh
Additional filtering options, e.g., 'by source repository', 'by owner' etc.
Better metadata associated with exported datasets from the Datamesh
Additional option for the user to select the projection for the export function

Additional clear descriptions of Owner/Authority/Sharing license/Attributions

Regular automatic refreshing of data links

Automated notifications/alert when source datasets used in analyses or selected in a package are updated

Recommendations for improving the KEA App

Include an 'execution/enter button'

Include an option to incorporate new datasets in the analysis

Integrate additional analyses such as consideration of existing protected areas and uses and activities (in progress)

Display of statistic data from each analysis (in progress)

Recommendation for the Proof of Concept

Overall, DOC recommends progressing the use of a Datamesh amongst government agencies beyond this Proof of Concept. The Marine Geospatial and Information Steering Group would be best placed to develop the roadmap on how this can be achieved. A multiagency business case could provide the necessary resources to undertake this process

Recommendations for marine management

If the Datamesh is progressed beyond a Proof of Concept its success in informing marine management will be reliant upon its uptake. This includes ensuring that both data custodians make data available to the Datamesh and end users are aware of the Datamesh and its functionality. We therefore recommend investment in developing a management plan for the Datamesh and promoting it amongst data providers and end users to maximise its ability to inform marine management. A key component of this will be highlighting its dual functionality – that of providing centralised access to spatial datasets from multiple providers, and as a platform for developing apps to run standardised spatial analyses.

Develop a standardised process to map areas of ecologically significant

Increase the number of marine datasets publicly available

8   Acknowledgments

Oceanum, Dave Johnson and Peter McComb provided excellent support with the Datamesh and the KEA analyses; NIWA, Matt Bennion, Carolyn Lindquist and Phoebe Stewart-Sinclair developed the KEA analytical methodology and scripts; All the user testing participants from the regional councils. LINZ, MPI and Te Arawhiti, as partner agencies.

9   References

Bennion, M., Stewart-Sinclair, P., Lundquist, C. (2022). Application of KEA Criteria at National and Regional level: a Proof of Concept. NIWA report prepared for DOC

Lundquist, C., Stephenson, F., McCartain, L., Watson, S., Brough, T., Nelson, W., Neill, K., Anderson, T., Anderson, O., Bulmer, R., Gee, E., Pinkerton, M., Rowden, A., Thompson, D. (2020) Evaluating Key Ecological Areas datasets for the New Zealand Marine Environment. Report prepared by NIWA for DOC. Available at https://www.doc.govt.nz/globalassets/documents/conservation/marine-and-coastal/marine-protected-areas/mpa-publications/evaluating-kea-datasets-2020.pdf

Stephenson, F., Rowden, A., Anderson, T., Hewitt, J., Costello, M., Pinkerton, M., Morrison, M., Clark, M., Wadhwa, S., Mouton, T., Lundquist, C. (2018) Mapping key ecological areas in the New Zealand marine environment: Data collection. Report prepared by NIWA for DOC. Available at https://www.doc.govt.nz/globalassets/documents/conservation/marine-and-coastal/marine-protected-areas/mpa-publications/key-ecological-areas-report-2018.pdf

Stewart-Sinclair, P., Lundquist, C. (2022) Application of the Key Biological Area (KBA) process to identify Key Ecological Areas. Report prepared by NIWA for DOC.

10   Appendices

Appendix 1: Spatial data linked to the Datamesh for the DOC's use case and used for the KEA analyses with sources, formats, sizes and uses for the DOC's use case.

Dataset	Source	Format	Size	Used for
Macroalgae predicted distributions 2020 and uncertainty layers	DOC Marine Portal	172 individual rasters	0.74 GB	Analyses
Reef Fish predicted distributions 2020 and uncertainty layers	DOC Marine Portal	102 individual rasters	0.46 GB	Analyses
Demersal Fish predicted distributions 2020 and uncertainty layers	DOC Marine Portal	478 individual rasters	3.90 GB	Analyses

Subtidal Invertebrates predicted distributions 2020 and uncertainty layers	DOC Marine Portal	414 individual rasters	3.6 GB	Analyses
Finfish Spawning Distributions	MPI Open Geospatial Data Portal	39 polygon shapefiles	0.85 GB	Analyses
NZ Exclusive Economic Zone	LINZ Data Service	1 polygon shapefile	<0.1 GB	Clip
New Zealand Territorial Sea	LINZ Data Service	1 line shapefile	0<0.1 GB	Mapping
NZ MHW shoreline	LINZ Data Service	1 line shapefile	<0.1 GB	Mapping
Regional Councils boundaries	Stats NZ Geographic Data Service	1 polygon shapefile	0.1 GB	Clip
Marine Bioregions	DOC Marine Portal	1 polygon shapefile	<0.1 GB	Clip
Sanctuaries to Protect Marine Mammals	DOC Marine Portal	1 polygon shapefile	0.22 GB	Mapping
NZ Street Address	LINZ Data Service	1 point shapefile	1.18 GB	Analyses (trial only)
Northland Wetland	NRC Open Data Site	1 polygon shapefile	0.2 GB	Analyses (trial only)
Ramsar Sites boundaries	Ramsar Sites Information Services	1 polygon shapefile	<0.1 GB	Mapping