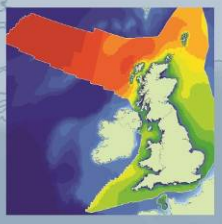
The Crown Estate

Marine Data Exchange: Lessons Learnt

Report R.2271

July 2014

Creating sustainable solutions for the marine environment


Page intentionally left blank

The Crown Estate

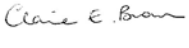
Marine Data Exchange: Lessons Learnt

Date: July 2014

Project Ref: R/4239/1

Report No: R.2271

Version	Details of Change	Date
1.0	Issue for Client Comment	16.05.2014
2.0	Issue for Client Comment - Re Additional Work	15.07.2014
3.0	Issue After Client Review	29.07.2014

Document Authorisation		Signature	Date
Project Manager:	T A Perks		29.07.2014
Quality Manager:	N M Warken		29.07.2014
Project Director:	C E Brown		29.07.2014

ABP Marine Environmental Research Ltd
Quayside Suite, Medina Chambers, Town Quay, Southampton, Hampshire SO14 2AQ

Tel: +44 (0) 23 8071 1840 Fax: +44 (0) 23 8071 1841
Web: www.abpmer.co.uk Email: enquiries@abpmer.co.uk

ABPmer is certified by:

Disclaimer:

- *Any 'Draft' issue of this report, and any information contained therein, may be subject to updates and clarifications on the basis of any review comments before 'Final' issue. All content should therefore be considered provisional, and should not be disclosed to third parties without seeking prior clarification from ABP Marine Environmental Research Ltd ("ABPmer") of the suitability of the information for the intended disclosure and should not be relied upon by the addressee or any other person.*

- *Unless previously agreed between the addressee and ABPmer, in writing, the 'Final' issue of this report can be relied on by the addressee only. ABPmer accepts no liability for the use by or reliance on this report or any of the results or methods presented in this report by any party that is not the addressee of the report. In the event the addressee discloses the report to any third party, the addressee shall make such third party aware that ABPmer shall not be liable to such third party in relation to the contents of the report and shall indemnify ABPmer in the event that ABPmer suffers any loss or damage as a result of the addressee failing to make such third party so aware.*

- *Sections of this report rely on data supplied by or drawn from third party sources. Unless previously agreed between the addressee and ABPmer, in writing, ABPmer accepts no liability for loss or damage suffered by the addressee or any third party as a result of any reliance on third party data contained in the report or on any conclusions drawn by ABPmer which are based on such third party data.*

Marine Data Exchange: Lessons Learnt

Contents

Tables

Figures

1. Introduction

The Marine Data Exchange (MDE) internet portal was launched by The Crown Estate in February 2013, to provide anyone with an interest in the marine environment access to a wealth of survey data collected from The Crown Estate's low carbon energy development partners (The Crown Estate, 2013). In July 2013, ABP Marine Environmental Research Ltd (ABPmer) were commissioned by The Crown Estate to undertake quality assurance (QA) checks upon the datasets and associated metadata present on the MDE. The datasets were QA'd using The Crown Estate's Quality Control (QC) Protocol, outlined within the 'Requirements for providing survey data to The Crown Estate via the Marine Data Exchange' document (Requirements document) (The Crown Estate, 2012). Datasets that fell within the scope of the Marine Environmental Data Information Network (MEDIN) data guidelines were checked against them, to ensure that all vital information was provided. The metadata was QA'd using the MEDIN discovery metadata standard (Seeley *et al*., 2014). At present, approximately 900 series, spanning 54 development sites, have been QA'd.

The purpose of this report is to provide details of the main issues that were flagged up during the QA checks, and to propose recommendations to reduce their occurrence. A quantitative review was performed on 286 series (32% of the total number of series QA'd), located within 18 development sites. During the review, the frequency of issues flagged during the QA checks was recorded in spreadsheets. Each reviewed series was assigned to a theme, to enable theme-specific issues to be identified. In addition, issues were tallied according to the series date, specifically whether the data was collected pre-2011 or post-2011. Table 1 shows the 20 devised themes, and the total number of series reviewed within each theme.

The results of the review are presented within this report, in which recurring issues across multiple themes are discussed, as well as recurring issues within specific themes. A number of recommendations have been proposed to reduce the future occurrence of the frequent issues.

Table 1. Number of series reviewed with the devised themes

Theme	No. of Pre-2011 Series Reviewed	No. of Post-2011 Series Reviewed	Total Number of Series Reviewed
Airborne Noise	5	1	6
Archaeology	11	1	12
Aviation	3	0	3
Benthic	13	5	18
Coastal Processes	15	0	15
Environmental Statements and Assessments	13	1	14
Fisheries	10	6	15
Geophysical	28	14	42
Geotechnical	18	6	24
Management and Monitoring Documents (e.g. licence documents)	10	2	12
Meteorology (including Meteorological Mast data)	6	2	8
Onshore Contaminated Land	2	0	2
Onshore Infrastructure (e.g. cable routing studies, substation design plans etc.)	14	2	16
Onshore Ecology	12	3	15
Ornithological and Marine Mammals	40	3	43
Physical Oceanography (e.g. buoy measurements, modelling etc.)	10	0	10
Shipping and Navigation	10	0	10
Socio-economics	4	1	5
Underwater Noise	5	1	6
Visuals, Landscape and Seascape	8	1	9
TOTAL	237	49	286

2. Recurring Issues Across Multiple Themes

The top seven recurrent issues are described in this section; these issues occurred in at least 5% of the total number series reviewed, and affected greater than three themes, figure 1.

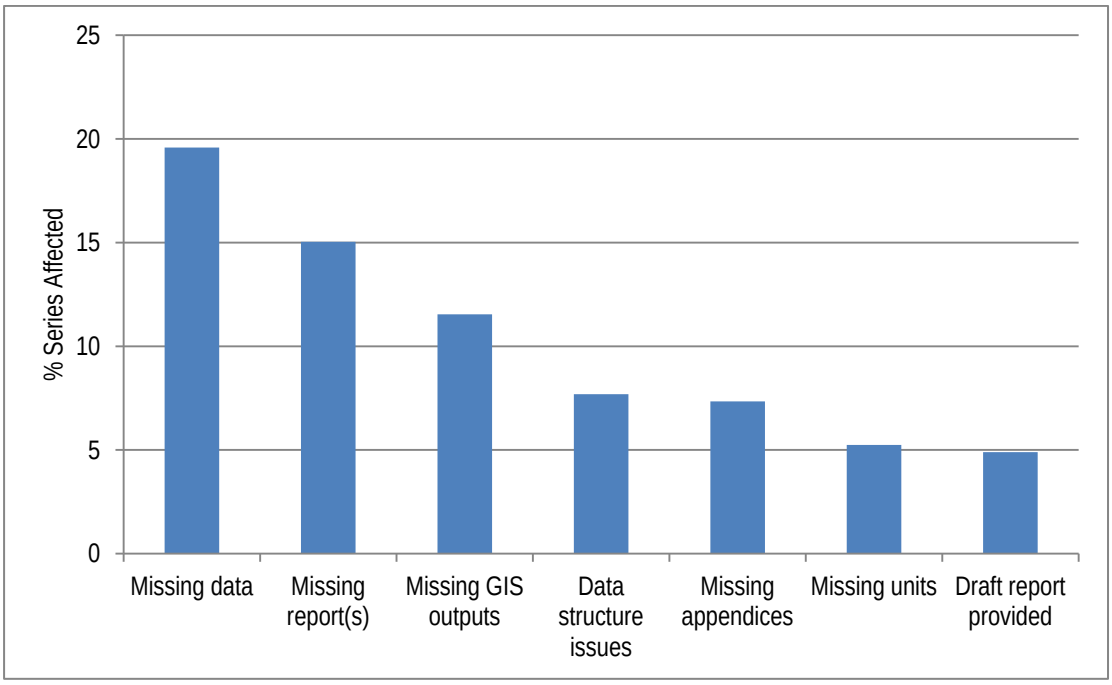


Figure 1. Top seven recurrent issues, occurring in greater than 5% of series reviewed, and affecting more than three themes

2.1 Missing Data

The issue of missing data was the most frequent issue identified during the review, with 20% of series across 13 themes, recorded as having missing data, figure 2. Missing data was usually identified by references within associated reports. Within the Geophysical theme, 40% of series were recorded as having data missing, the highest proportion of all themes. In many of these cases, data was provided for some, but not all of the techniques employed. The themes with the joint second-highest amount of missing data were Airborne Noise, Ornithology & Marine Mammals, and Underwater Noise, which all had 33% of data missing. In some cases within the Ornithology & Marine Mammals theme, the issue was gaps in the data provided, rather than entire datasets missing. In addition, summary results were often provided, rather than full observation datasets which contain details such as the date and time of the sighting, the distance from the boat, and weather conditions.

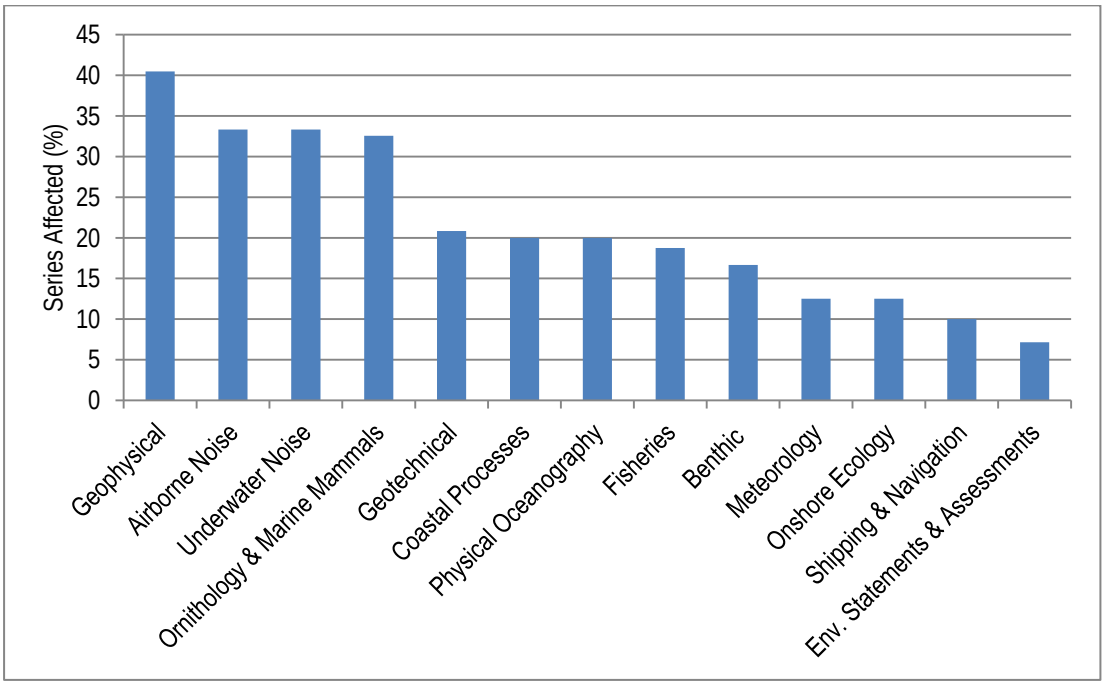


Figure 2. The percentage of reviewed series with missing datasets within each theme

2.2 Missing Reports

Missing reports was the second most-frequent issue; 15% of series reviewed had one or more reports missing, with 15 themes affected. Missing reports were identified by references in other reports, or by data being supplied without an accompanying report. The theme with the highest number of series with one or more reports missing was Onshore Ecology, with 25% of series reviewed missing at least one report, figure 3. The next highest was the Environmental Statements & Assessments theme, with 21% of series missing reports, followed by Physical Oceanography (20%) and Geophysical (19%). As multiple reports are usually produced by contractors for the Onshore Ecology and Geophysical surveys, this may account for the relatively high proportion of series within these themes with reports missing.

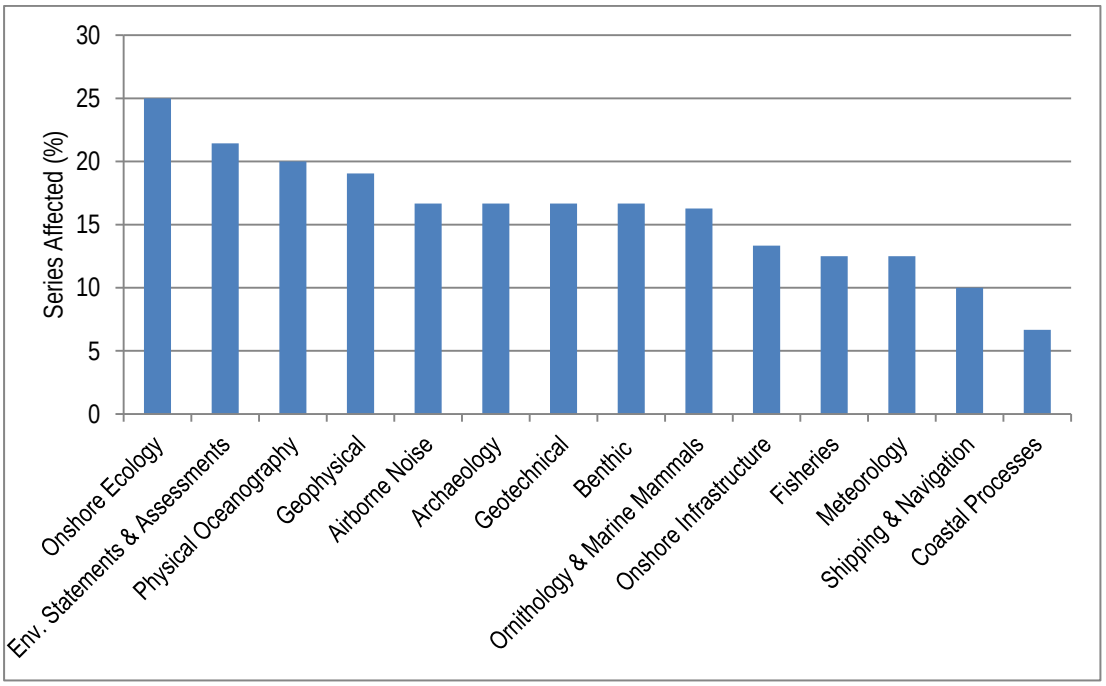


Figure 3. The percentage of reviewed series with missing reports within each theme

2.3 Missing GIS Outputs

As part of the QA process, GIS datasets were flagged as missing if they were presented in report figures, but not provided as shapefiles. In some cases, GIS datasets formed an integral component of the survey results, whilst in other instances, they were provided as supporting information. GIS outputs were missing from 12% of series reviewed, with 13 themes affected. They were found to be missing the most frequently in Shipping & Navigation series (30%), followed by Ornithology and Marine Mammals (21%), Onshore Infrastructure (20%), and Archaeology (17%), figure 4.

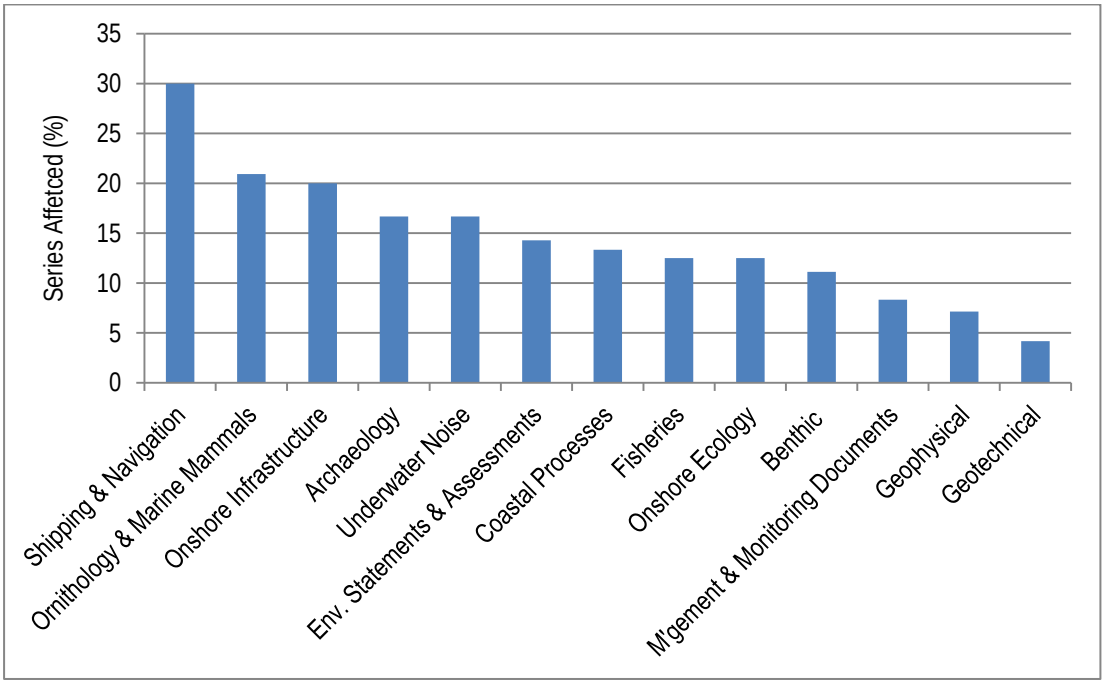


Figure 4. The percentage of reviewed series with missing GIS outputs within each theme

2.4 Data Structure Issues

Data structure issues were recorded when reports and datasets on the MDE were not structured according to the structure outlined within the 'Requirements Document' (The Crown Estate, 2012). Examples include reports being provided within 'dataset' packages (and vice versa), unrelated studies being uploaded within the same series (and vice versa), different types of data being bundled into one 'dataset' package, and series being uploaded to the wrong development site. Data structure issues reduce the discoverability of data for users of the MDE; they affected 8% of series reviewed, across 12 themes. Series assigned to the Environmental Statements and Assessments theme had the highest number of data structure issues (21%), followed by Airborne Noise and Geophysical (both 17%), figure 5. Data structure issues within the Environmental Statements & Assessments theme were often caused by documents relating to joint development sites being uploaded to individual sites, and vice versa.

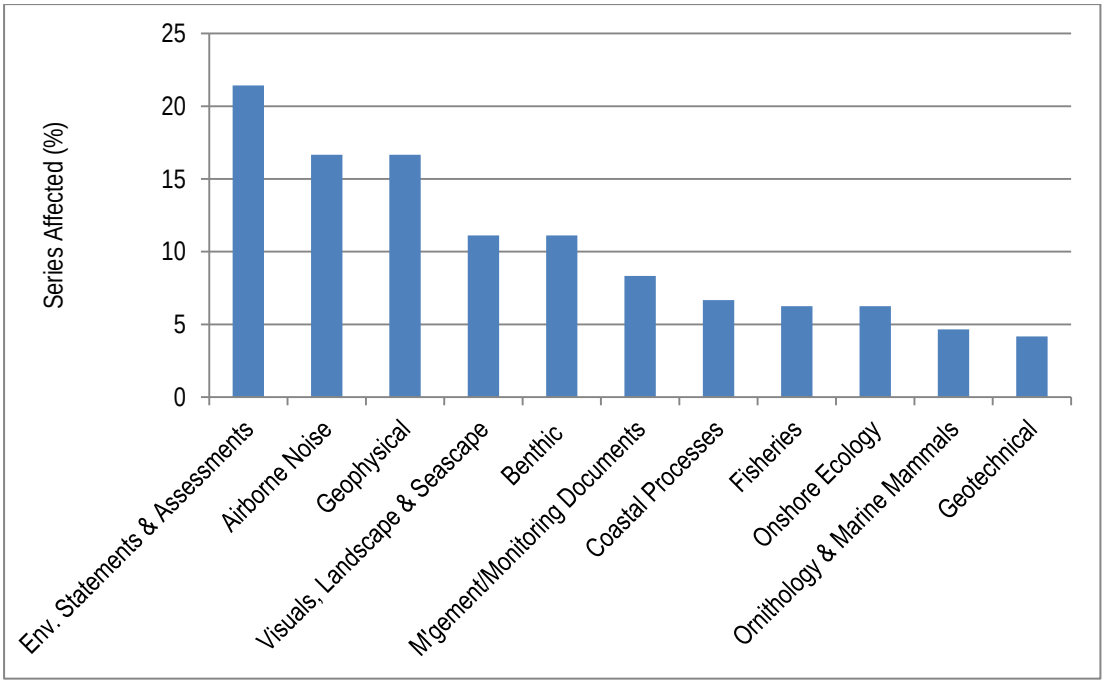


Figure 5. The percentage of reviewed series with data structure issues within each theme

2.5 Missing Report Appendices

Report appendices were missing from 7% of series reviewed, with 12 themes affected. The appendices usually contain important supporting information, such as survey logs, details of quality control (QC) checks and calibration information/certificates, hence it is important that they are provided. The proportion of Management and Monitoring Documents series affected was considerably higher than series within other themes, with 33% of series affected. In many cases, only parts of the appendices were missing, figure 6.

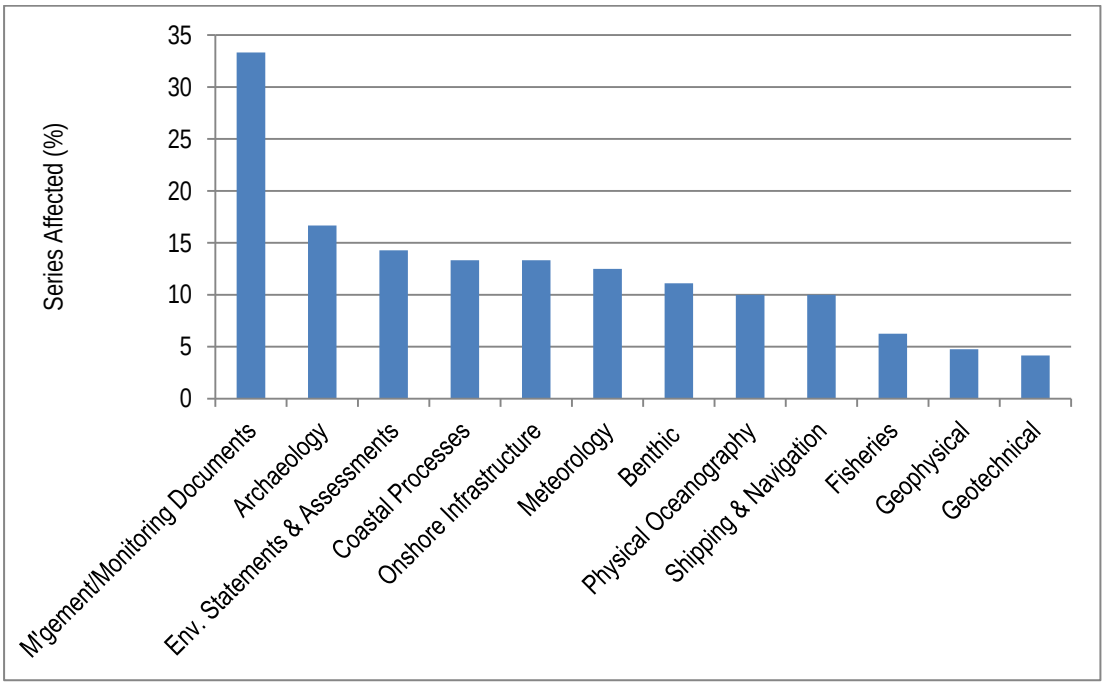


Figure 6. The percentage of reviewed series with missing report appendices within each theme

2.6 Missing Units

Data provided without the units being stated is a significant issue, since without them, the data cannot be used with confidence. 5% of reviewed series contained datasets with at least one set of units missing, across eight themes. This issue arose most frequently within the Meteorology (25%), Airborne Noise (17%) and Geophysical (14%) themes, figure 7. In many cases, particularly within the Meteorology and Geophysical themes, this issue was caused by the column headers not being provided, as this is where the units are usually stated. In many cases, the depth units were usually missing from the column headers of the results tables and survey logs.

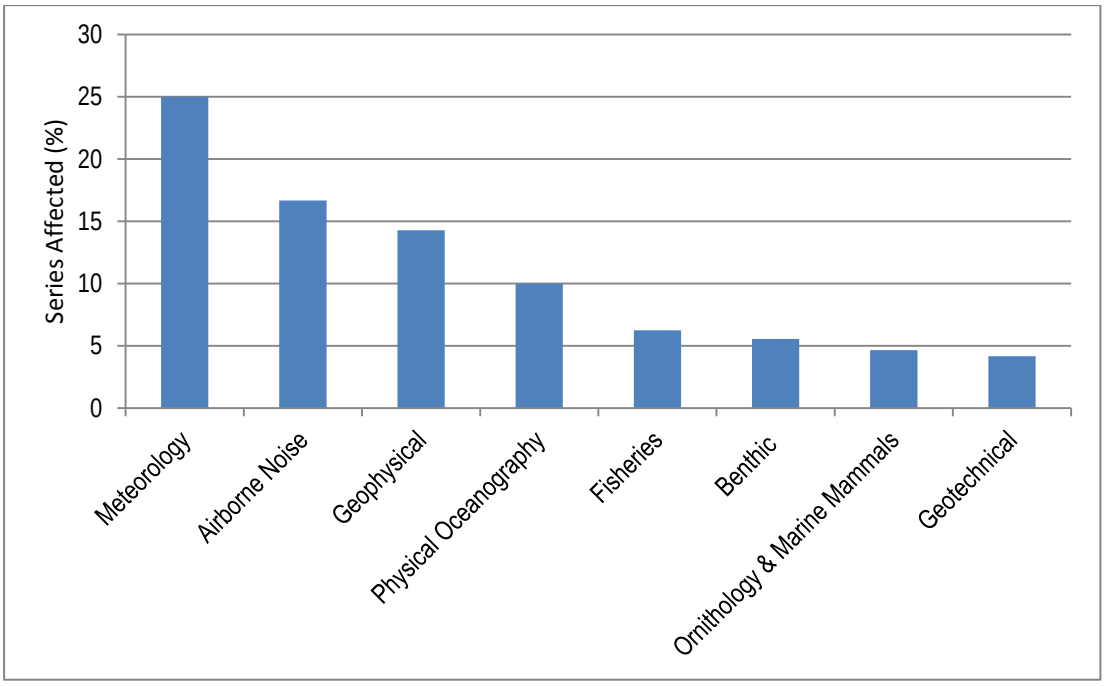


Figure 7. The percentage of reviewed series with missing units within each theme

2.7 Draft Report Provided

Draft versions of reports were provided within 5% of series reviewed, spanning 11 themes. Underwater Noise series were most affected (33%), followed by Coastal Processes and Onshore Infrastructure (both 13%), figure 8. In a couple of log files, the developer responded that the reports had outdated statuses, and were actually final versions.

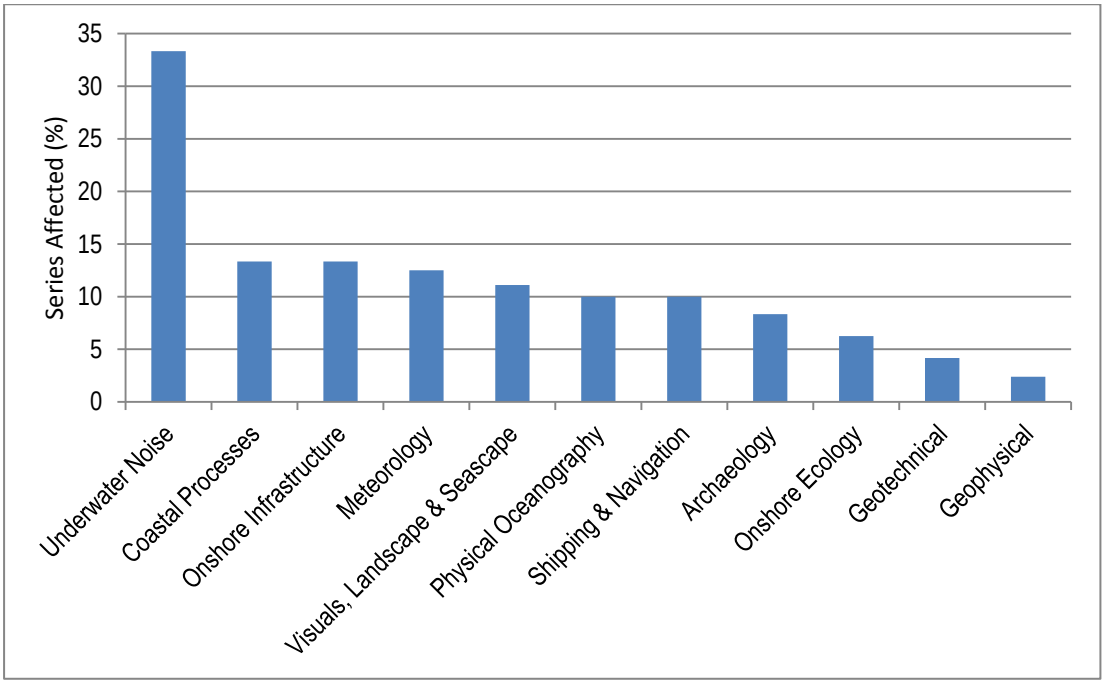


Figure 8. The percentage of reviewed series with draft reports within each theme

3. Comparison between Pre-2011 and Post-2011 Series

The majority of reviewed series were pre-2011 (83%), and as a result, there were only four data themes within which at least five post-2011 series were reviewed, and were therefore considered to be an adequate sample size for statistical analysis. Figure 9 shows the difference in the proportion of series affected by the most recurrent issues within pre and post-2011 series, within the four themes.

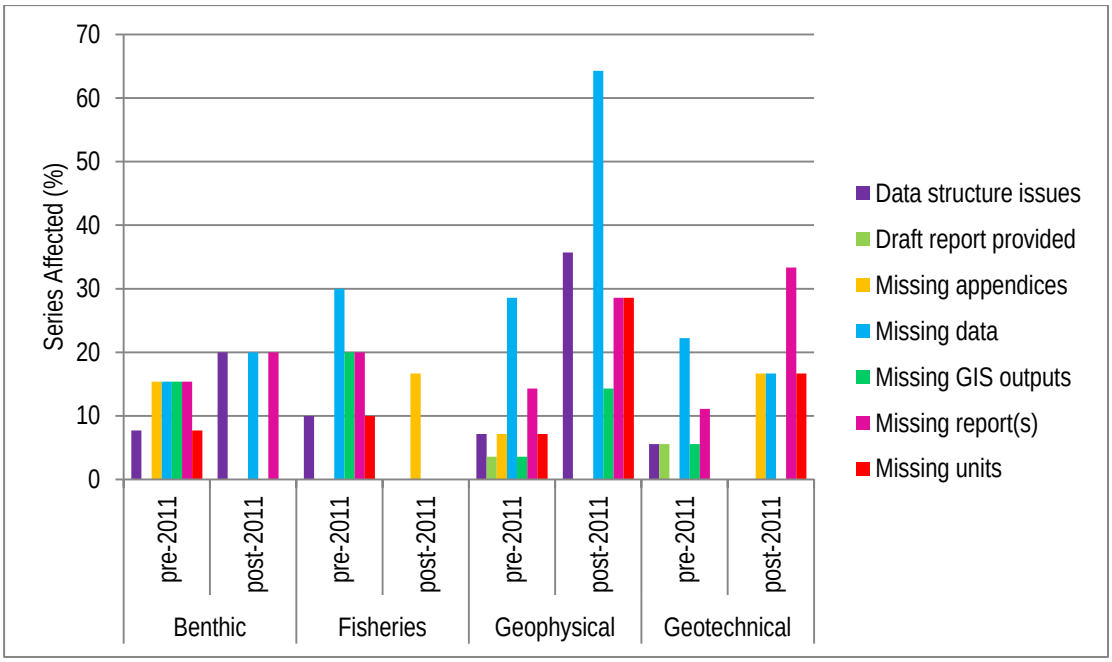


Figure 9. Comparisons between the proportion of pre-2011 and post-2011 series affected by the most recurrent issues within four themes

The results of the comparison show that the difference in the frequency of issues between pre-2011 series and post-2011 series is theme-dependent. No overarching patterns across all four themes and across the seven issues were identified. Fisheries series had the greatest decrease in issue frequency in post-2011 series relative to pre-2011 series, with five of the seven issues decreasing in frequency; the greatest decrease was the occurrence of missing data; 30% of pre-2011 series were missing data whilst none of the post-2011 series had data missing.

On the other hand, Geophysical series had the greatest increases in issue frequency in post-2011 series relative to pre-2011 series, with five issues increasing in frequency. Within this theme, the greatest increase was the occurrence of missing data; 29% of pre-2011 series were missing data, whilst 64% of post-2011 series had data missing, an increase of 35%.

The difference in the frequency of issues between pre-2011 series and post-2011 series was less clear-cut within the Benthic and Geotechnical themes. Within post-2011 benthic series, three issues increased in frequency relative to pre-2011 series, whilst four issues did not occur at all within post-2011 series. Within post-2011 geotechnical series, three issues increased in frequency relative to pre-2011 series, whilst four issues decreased in frequency.

4. Issues Within Specific Themes

To highlight the main issues within each theme, issues occurring more than once have been described. There were no issues recorded within series belonging to the Aviation and Socio-economics themes, however only three and five series were reviewed respectively. The issues associated with the Onshore Contaminated Land theme have not been analysed since only two series were reviewed. Within the Visuals, Landscape & Seascape theme, no issue occurred greater than once.

4.1 Airborne Noise

Since there was only one issue affecting greater than one series, a graph is not required for this theme.

Table 2. Details of Issues within Airborne Noise series

Issue	No. Series Affected	Comments
Missing data	2 (33%)	The soundings data was not provided for two of the six series reviewed

4.2 Archaeology

Since the same number of series were affected by each issue, a graph is not required for this theme.

Table 3. Details of Issues within Archaeology series

Issue	No. Series Affected	Comments
Missing GIS outputs	2 (17%)	GIS data was presented in the report figures, as supporting information, however the shapefiles were not provided
Missing survey/transect co-ordinates	2 (17%)	Survey co-ordinates were not stated in the survey report, and no track shapefiles were provided
Missing report(s)	2 (17%)	Not all reports were provided
Missing figures/drawings	2 (17%)	Report figures were missing from two series
Missing appendices	2 (17%)	The appendices contain important information, and therefore must be provided

4.3 Benthic

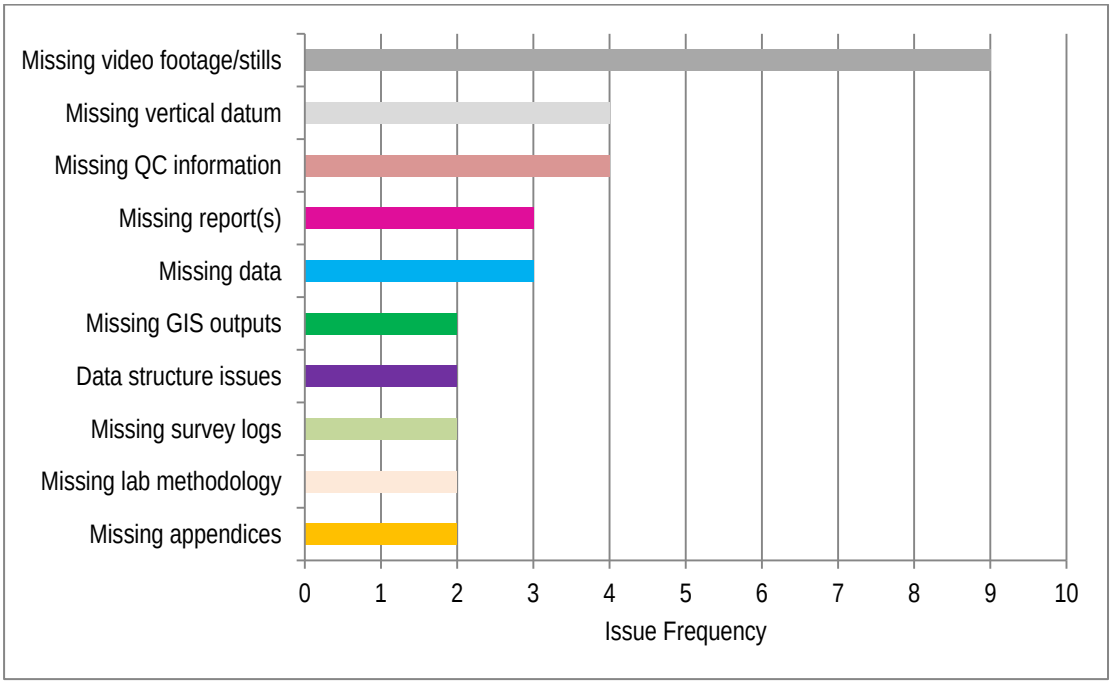


Figure 10. Issues within Benthic series

Table 4. Details of Issues within Benthic series

Issue	No. Series Affected	Comments
Missing video footage/stills	9 (50%)	This issue had a significantly higher frequency than others within this theme; the data may be being withheld due to large file sizes
Missing vertical datum	4 (22%)	The vertical datum should always be stated when the water depth data is stated
Missing QC information	4 (22%)	In most cases, there was no section on QC information specified in the report contents, therefore this information may not have been a deliverable from the contractor(s)
Missing report(s)	3 (17%)	There are usually multiple reports associated with benthic surveys
Missing data	3 (17%)	There are usually multiple datasets associated with benthic surveys, including infaunal abundance data and sediment data
Missing GIS Outputs	2 (11%)	GIS data was shown in the report figures but not provided as shapefiles
Missing survey logs	2 (11%)	Survey logs contain useful information, such as the survey co-ordinates and weather conditions
Missing lab methodology	2 (11%)	Details of the laboratory methodology were not provided
Missing appendices	2 (11%)	Appendices were missing from two reports

4.4 Coastal Processes

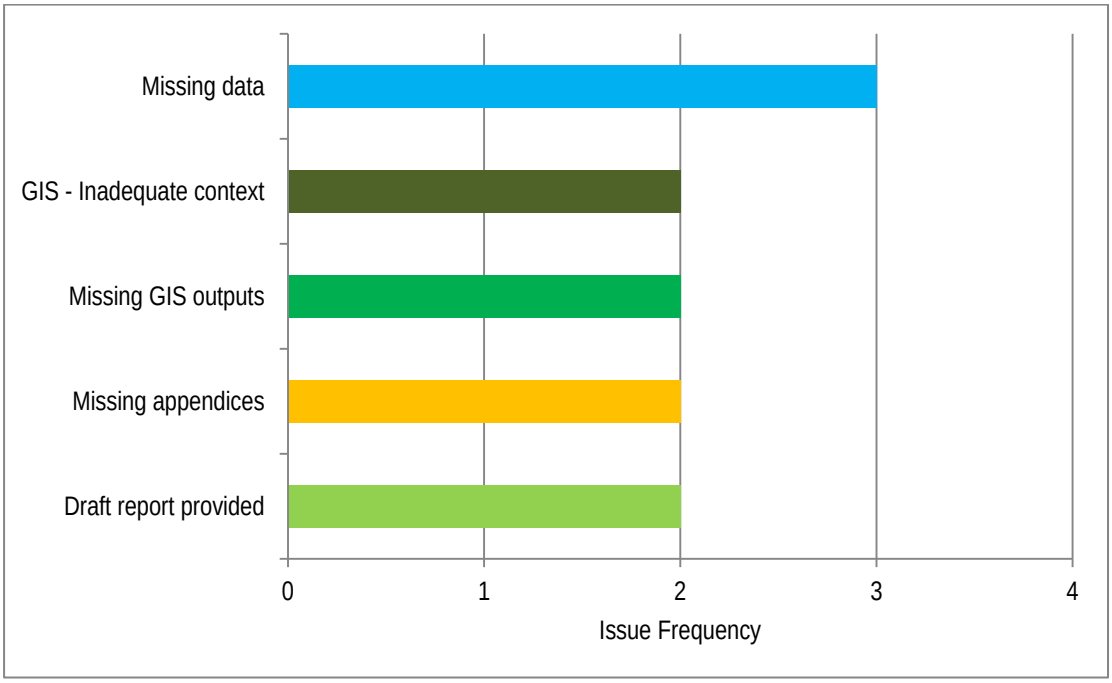


Figure 11. Issues within Coastal Processes series

Table 5. Details of Issues within Coastal Processes series

Issue	No. Series Affected	Comments
Missing data	3 (20%)	Data was missing from three series
GIS – Inadequate context	2 (13%)	Shapefiles were provided however they were not provided with enough context
Missing GIS outputs	2 (13%)	GIS data was shown in the report figures but not provided as shapefiles
Missing appendices	2 (13%)	Appendices were missing from two reports
Draft report provided	2 (13%)	A draft was provided in two of the fifteen series reviewed

4.5 Environmental Statements and Assessments

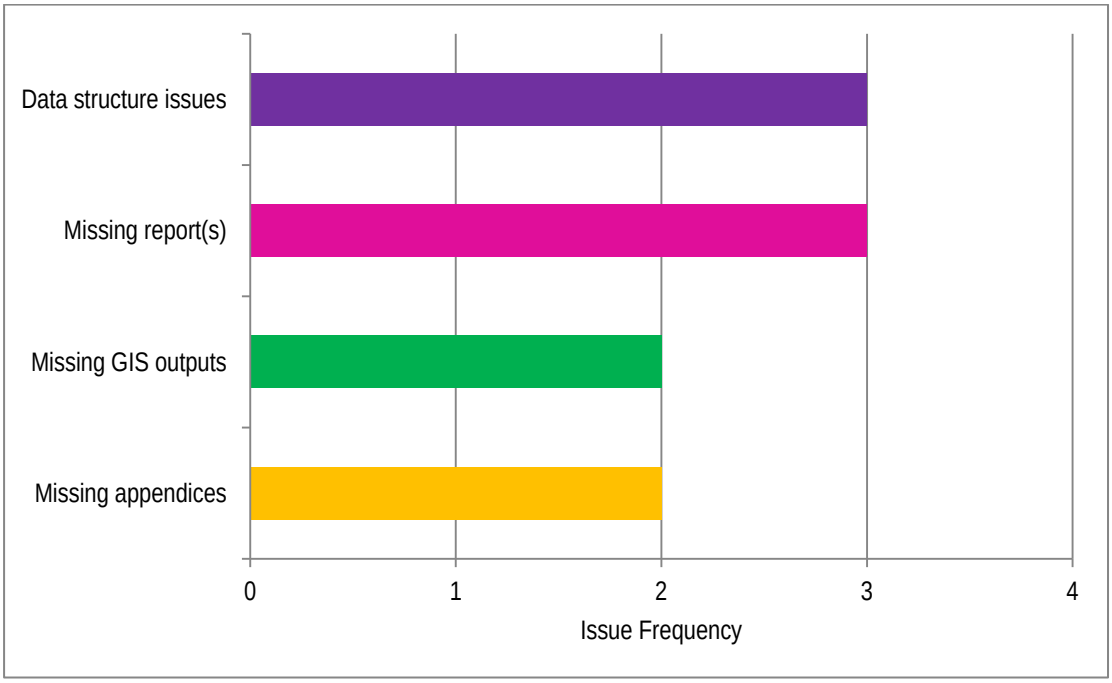


Figure 12. Issues within Environmental Statements and Assessments series

Table 6. Details of Issues within Environmental Statements and Assessments series

Issue	No. Series Affected	Comments
Data structure issues	3 (21%)	Environmental Statements were uploaded to the wrong site, or structured illogically
Missing report(s)	3 (21%)	Chapters of the Environmental Statements were missing
Missing GIS outputs	2 (14%)	GIS data was shown in the report figures but not provided as shapefiles
Missing appendices	2 (14%)	There are usually extensive appendices associated with Environmental Statements

4.6 Fisheries

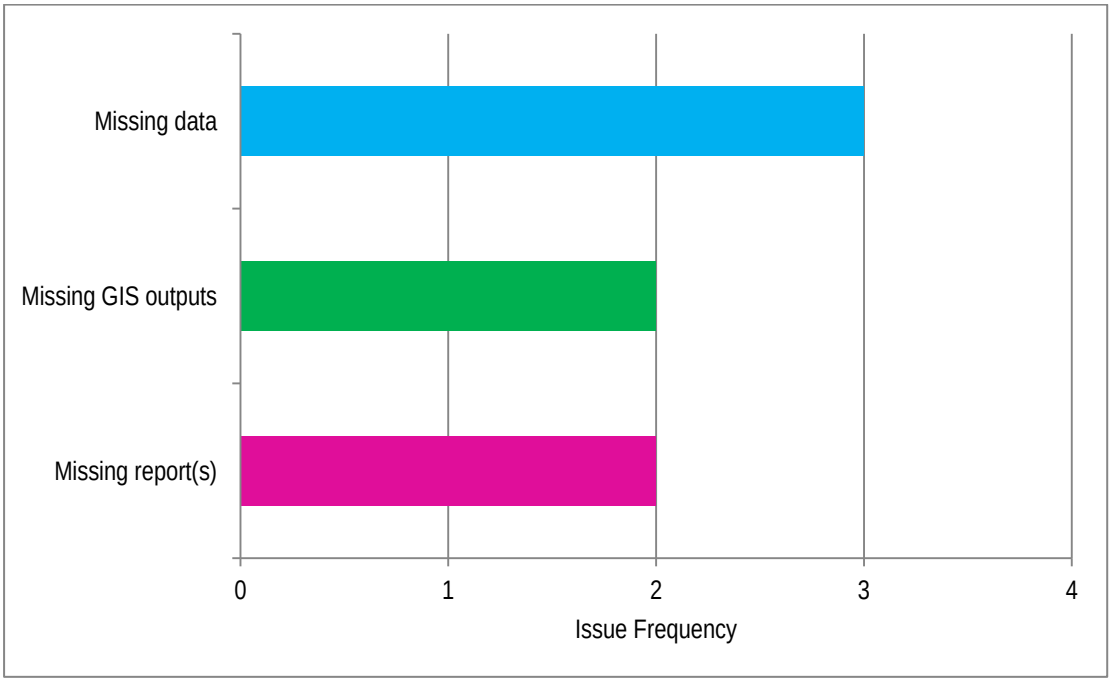


Figure 13. Issues within Fisheries series

Table 7. Details of Issues within Fisheries series

Issue	No. Series Affected	Comments
Missing data	3 (19%)	The full catch dataset, with fish sizes etc. was missing, and in another case there was temporal gaps in the data provided
GIS - Missing GIS outputs	2 (13%)	GIS shapefiles were presented in the report figures, as supporting information, but were not provided as usable shapefiles
Missing report(s)	2 (13%)	Numerous monthly reports are usually produced to accompany fisheries data , sometimes one or more is not provided

4.7 Geophysical

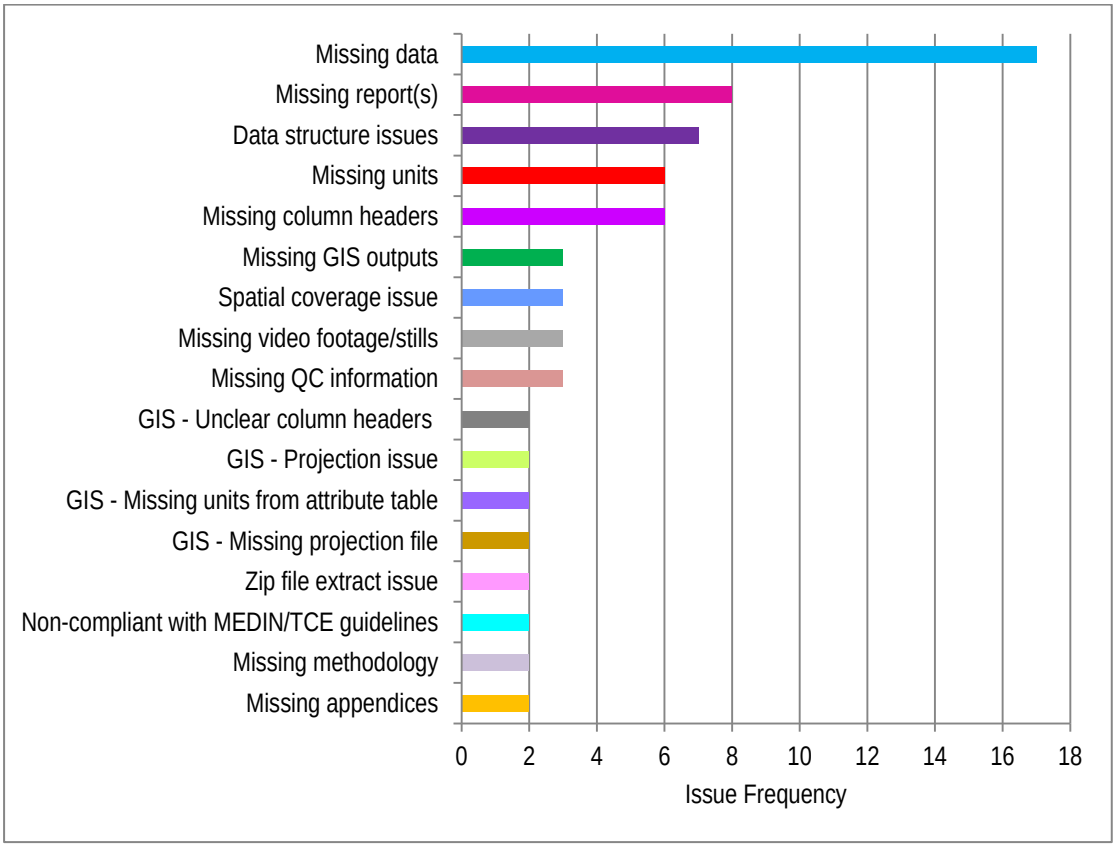


Figure 14. Issues within Geophysical series

Table 8. Details of Issues within Geophysical series

Issue	No. Series Affected	Comments
Missing data	17 (40%)	Entire datasets (e.g. side scan sonar - both raw and processed) were often missing
Missing report(s)	8 (19%)	There are usually multiple reports associated with geophysical surveys
Data structure issues	7 (17%)	There were issues with the data structure on the MDE e.g. all datasets bundled into one package
Missing units	6 (14%)	This issue was predominantly a result of missing column headers in the magnetometer data
Missing column headers	6 (14%)	Generally, this issue only occurred within magnetometer datasets.
Missing GIS outputs	3 (7%)	GIS shapefiles were presented in the report figures, as supporting information, but were not provided as usable shapefiles
Spatial coverage issue	3 (7%)	The datasets provided did not cover the entire survey area shown in the report
Missing video footage/stills	3 (7%)	This data may be being withheld due to large file sizes
Missing QC information	3 (7%)	Details of the quality control checks performed on the data were not provided
GIS – unclear column headers	2 (5%)	The column headers within the attribute tables were unclear
GIS – projection issue	2 (5%)	Shapefiles were not projecting to the correct location
GIS – missing units from attribute table	2 (5%)	The units are critical in order for the data to be usable
GIS – Missing projection file	2 (5%)	The shapefiles will not project in the correct location without the projection file
Zip file extract issue	2 (5%)	The zip files would not fully extract
Non-complaint with MEDIN/The Crown Estate guidelines	2 (5%)	Two series were missing mandatory MEDIN information that was considered to cause a material issue
Missing methodology	2 (5%)	The data collection methodology is valuable information that should always be provided
Missing appendices	2 (5%)	There are usually lengthy appendices accompanying Geophysical reports

4.8 Geotechnical

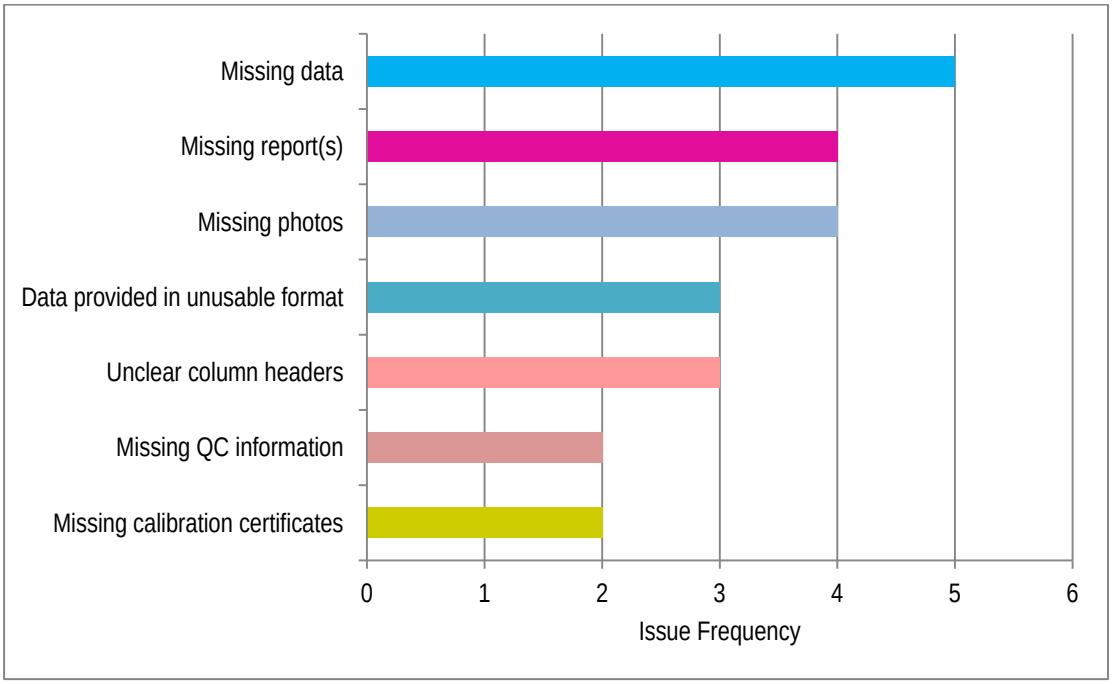


Figure 15. Issues within Geotechnical series

Table 9. Details of Issues within Geotechnical series

Issue	No. Series Affected	Comments
Missing data	5 (21%)	Data was not provided alongside the geotechnical reports
Missing report(s)	4 (17%)	There are usually operations and results reports associated with geotechnical surveys
Missing photos	4 (17%)	Photos of borehole samples were missing from four series
Data provided in unusable format	3 (13%)	The data was shown as graphs in the report, but not provided in a format that would enable further investigation and analysis
Unclear column headers	3 (13%)	The column headers did not adequately define the data
Missing QC information	2 (8%)	Details of the quality control checks performed on the data were not provided
Missing calibration certificates	2 (8%)	The cone penetrometer calibration certificates were missing from two series

4.9 Management & Monitoring Documents

Since there was only one issue affecting greater than one series, a graph is not required for this theme.

Table 10. Details of Issues within Management & Monitoring Documents series

Issue	No. Series Affected	Comments
Missing appendices	4 (33%)	Two of the twelve series reviewed were missing report appendices

4.10 Meteorology

Since the same number of series were affected by each issue, a graph is not required for this theme.

Table 11. Details of Issues within Meteorology series

Issue	No. Series Affected	Comments
Missing units	2 (25%)	This issue was predominantly a consequence of missing column headers
Missing column headers	2 (25%)	Column headers are vital for the data to be usable
Error within data	2 (25%)	Incorrect units were stated in one series, and there was a font issue with another

4.11 Onshore Ecology

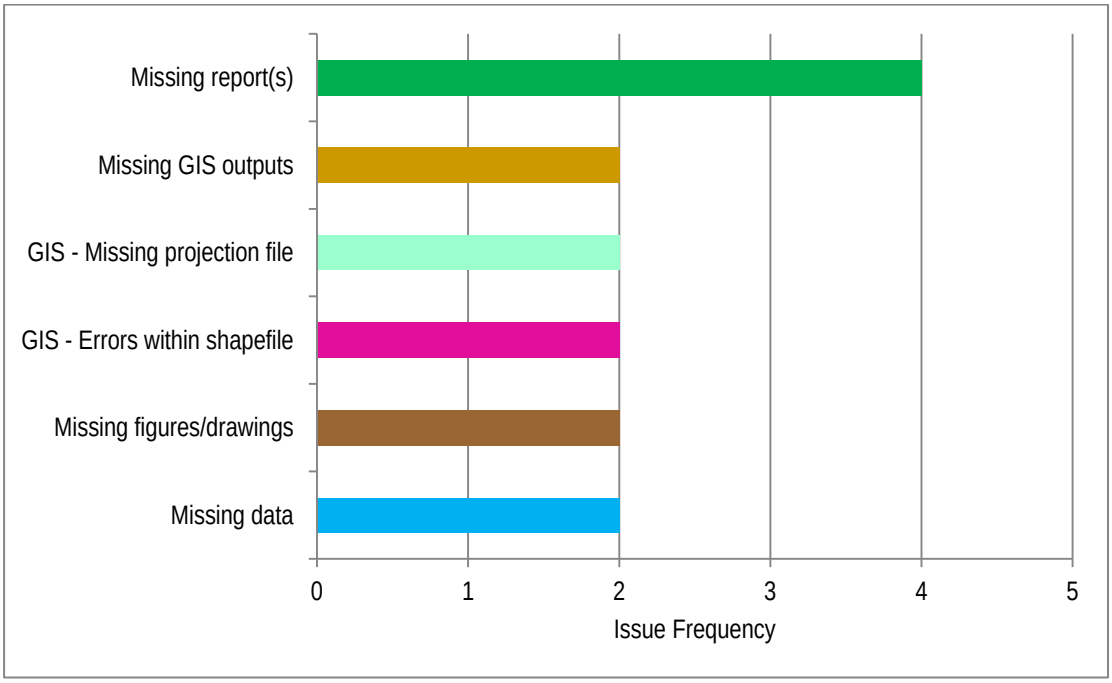


Figure 16. Issues within Onshore Ecology series

Table 12. Details of Issues within Onshore Ecology series

Issue	No. Series Affected	Comments
Missing report(s)	4 (25%)	Multiple reports are often bundled together to form an onshore ecology series, on occasion one or more can be missing
Missing GIS outputs	2 (13%)	GIS shapefiles were presented in the report figures, as supporting information, but were not provided as usable shapefiles
GIS – missing projection file	2 (13%)	Shapefiles will not project in the correct location without the projection file
GIS – errors within shapefile	2 (13%)	In both cases there were gaps in the attribute table entries
Missing figures/drawings	2 (13%)	Report figures were missing from two series
Missing data	2 (13%)	Data was not provided for two series

4.12 Onshore Infrastructure

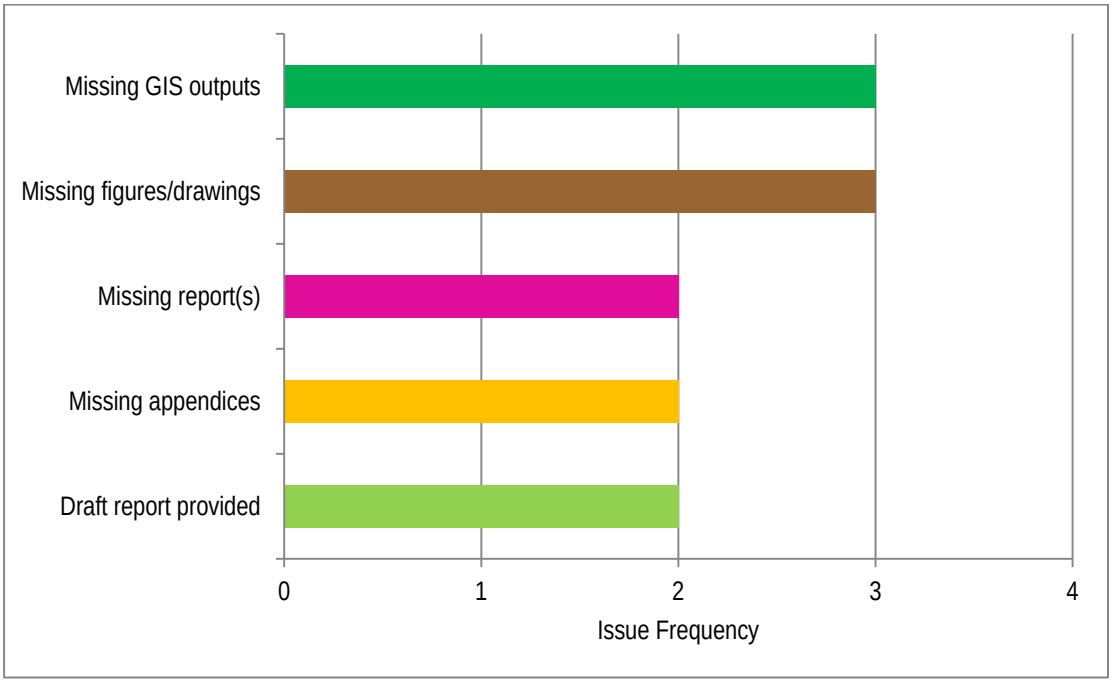


Figure 17. Issues within Onshore Infrastructure series

Table 13. Details of Issues within Onshore Infrastructure series

Issue	No. Series Affected	Comments
Missing GIS outputs	3 (20%)	GIS shapefiles were presented in the report figures, as supporting information, but were not provided as usable shapefiles
Missing figures/drawings	3 (20%)	Report figures were missing from two series
Missing report(s)	2 (13%)	Reports were missing from two of the 15 themes reviewed
Missing appendices	2 (13%)	Appendices were missing from two reports
Draft report provided	2 (13%)	A draft report was provided in two of the fifteen series reviewed

4.13 Ornithology and Marine Mammals

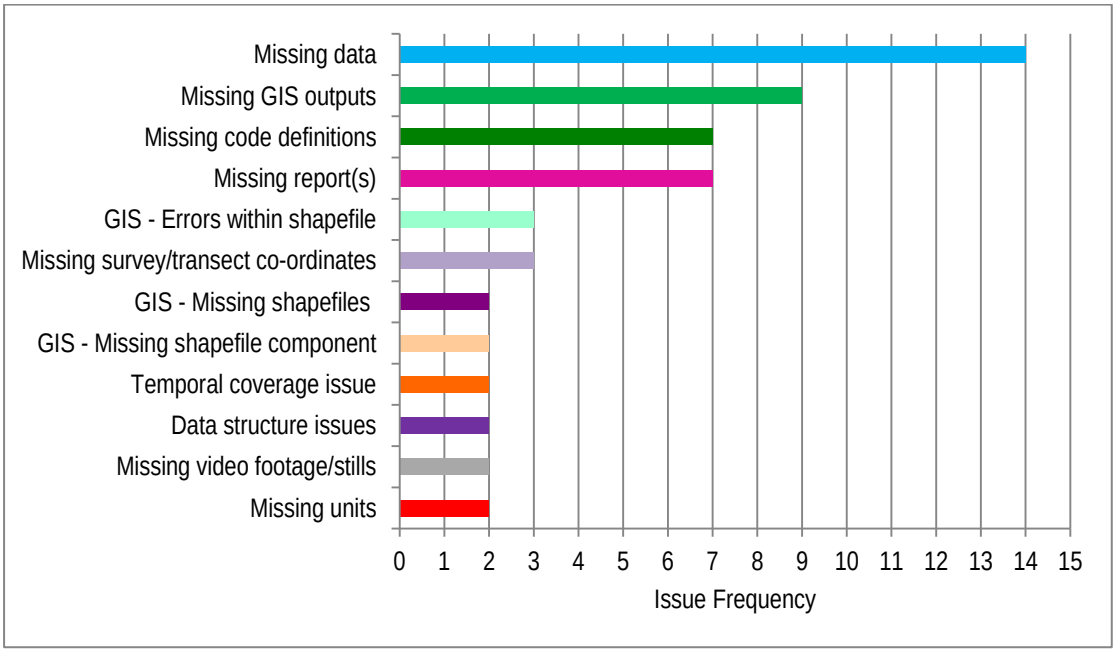


Figure 18. Issues within Ornithology and Marine Mammals series

Table 14. Details of Issues within Ornithology and Marine Mammals series

Issue	No. Series Affected	Comments
Missing data	14 (33%)	There were often gaps in the data provided e.g. missing one month
Missing GIS outputs	9 (21%)	GIS outputs are usually an integral component of the survey results (showing the location of sightings)
Missing code definitions	7 (16%)	Code definitions are vital for understanding the data, and were frequently missing for fields including species, distance, height and behaviour.
Missing report(s)	7 (16%)	Numerous monthly reports are usually produced to accompany ornithological data
GIS – Errors within shapefile	3 (7%)	In one series, data for the wrong month was provided within the shapefile
Missing survey/transect co-ordinates	3 (7%)	It is important that the survey locations are stated
GIS - Missing shapefiles	2 (5%)	Mxd files were provided, however the shapefiles were missing
Temporal coverage issue	2 (5%)	The data provided did not match the survey dates on two occasions
Data structure issues	2 (5%)	There were issues with the data structure on the MDE e.g. data provided in report packages
Missing video footage/stills	2 (5%)	Aerial video footage was missing from two series
Missing units	2 (5%)	The units are important to ensure that the data is usable

4.14 Physical Oceanography

Since the same number of series were affected by each issue, a graph was not required for this theme.

Table 15. Details of Issues within Physical Oceanography series

Issue	No. Series Affected	Comments
Missing report(s)	2 (20%)	Reports were missing from two of the ten reviewed series
Missing data	2 (20%)	Metocean data was missing from two series

4.15 Shipping and Navigation

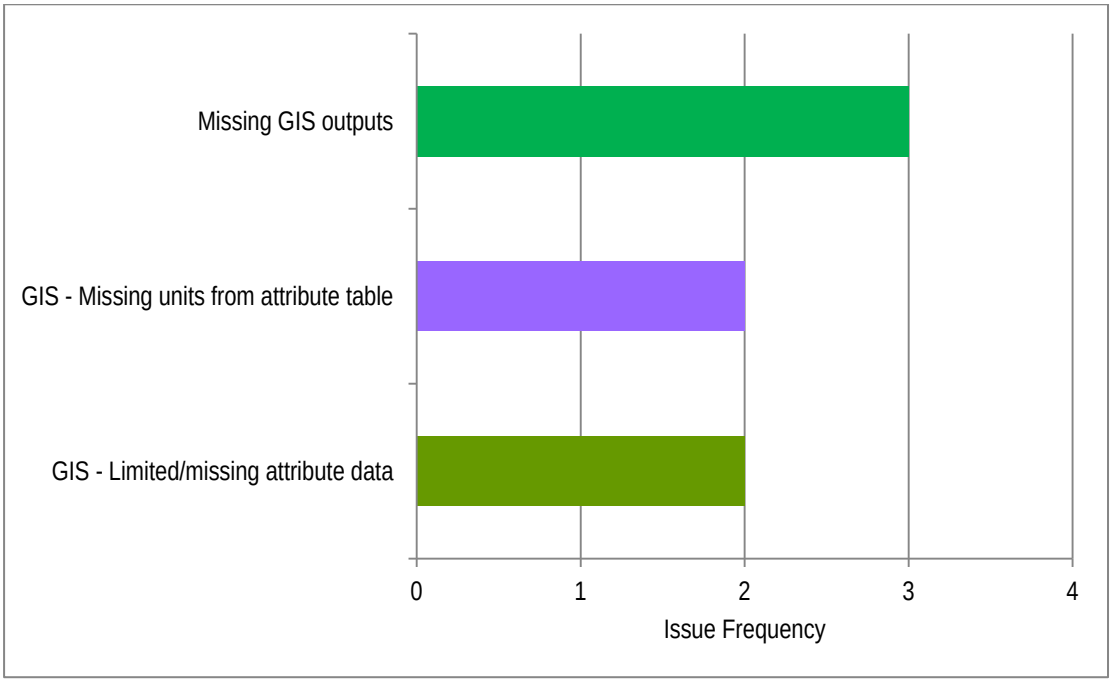


Figure 19. Issues within Shipping and Navigation series

Table 16. Details of Issues within Shipping and Navigation series

Issue	No. Series Affected	Comments
Missing GIS outputs	3 (30%)	Automatic Identification System (AIS) data is often not provided
GIS - Missing units from attribute table	2 (20%)	The usability of the data is limited without units
GIS - Limited/missing attribute data	2 (20%)	Useful information, such as date of ping, were missing from the attribute data of two series

4.16 Underwater Noise

Since the same number of series were affected by each issue, a graph was not required for this theme.

Table 17. Details of Issues within Underwater Noise series

Issue	No. Series Affected	Comments
Missing data	2 (33%)	Soundings data was not provided for two series
Draft report provided	2 (33%)	A draft report was provided in two of the six series reviewed

5. Metadata Issues

Issues with metadata on the MDE are fairly common; it is rare that the metadata supplied by the developer complies with the MEDIN discovery metadata standard. Recurrent issues are listed below.

5.1 Series, Report and Dataset Packages

- Vague titles – e.g. Bathymetry Data – missing contractor name, the year the survey was undertaken, type of data (raw or processed).
- Vague abstracts – often consisted of a broad statement copied and pasted from the survey report, rather than stating the content of the resource, which is required by the MEDIN standard.
- Limited and uncompliant keywords – The Crown Estate keywords were usually missing and the range of keywords was often limited (only one or two are stated). Keywords that are not listed in the MEDIN vocabulary were sometimes used e.g. 'Round 3'.

5.2 Dataset Packages Only

- Incorrect spatial reference system – incorrect European Petroleum Survey Group (EPSG) code stated.
- Incorrect or missing temporal extent – did not match the survey dates, or is not stated at all.
- Incorrect creation and publication date – did not match the creation or publication date of the associated report(s).
- Incorrect spatial resolution – spatial resolution stated was incorrect, or stated when it was not required.
- Incorrect originator – developer contact details were stated rather than the survey contractor.
- Incorrect geographic bounding box co-ordinates – co-ordinates provided were usually for the relevant wind farm lease area, regardless of whether the data covered this area.
- Lineage information missing - this is mandatory under the MEDIN standard.

5.3 Report Packages Only

- Incorrect creation and revision dates – often the creation date stated was the date the report was added to the MDE, rather than the date the report was created. A publication date was often stated even when the document had not been made publically available. Also, a revision date was usually stated, even when the document had not been revised. Perhaps developers assume all three dates have to be stated.

6. Case Studies

The following case studies provide examples of series that passed and failed the quality checks, to demonstrate the key differences and the impact upon the usability of the data.

6.1 QC Fail - Geophysical Series, Zone 2 Firth of Forth Offshore Wind Farm

There were numerous issues with this series that led to it failing the quality checks. Firstly, the operations report was missing; this report would provide important information including instrument details and methodology. Secondly, there were various data structure issues. Automatic Identification System (AIS) data had been provided in a package, however it was not associated with the geophysical survey in any way, and hence should have been provided in a separate series. Another data structure issue concerned the 'bathymetry xyz' package; items unrelated to the xyz bathymetry data were provided within this package, such as the Acoustic Ground Discrimination System (AGDS) data, and survey charts. This data structure does not conform to the structure outlined in the 'Requirements Document' (The Crown Estate, 2012), and is likely to hinder the ability of the data user in finding the information needed. In addition, there were issues with a couple of datasets provided; for example, definitions of codes present within the AGDS data were not provided. The column headers and units were also missing from the magnetometer and sound velocity data. All of these issues combine to reduce the usability of the data provided within this series.

6.2 QC Pass – Physical Oceanography Study, Lincs Offshore Wind Farm

This series passed the quality checks. The reports and data had been organised logically into three separate packages, one for the survey report and two for the datasets, which were divided by location. A large proportion of the data was presented in the report appendices, all of which were provided. The column headers within the spreadsheet data were clear, as the acronyms had been defined, and the units and vertical datum had also been provided. Hence the reports and datasets provided were comprehensible, and therefore usable.

7. Recommendations

Based on the issues identified during this review, a number of recommendations have been proposed to reduce their occurrence in future:

- Request that developers ensure all reports, report appendices, datasets and other deliverables are present before uploading a series to the MDE - particularly thorough checks need to be made on intricate series, for example, if a wide range of data has been collected, or if the data collected covers a long time period.
- Request that developers check that the statuses of reports are 'final' before uploading them to the MDE
- Remind developers to refer to the Requirements document when submitting data – guidance within this document, such as the mandatory data structure, is not always being followed.

7.1 Amendments to the Requirements Document

Additional details could be added to the Requirements document, which may reduce the occurrence of certain issues, these include:

- Provision of additional context on the MDE - developers seem to be unaware that some users of the MDE are unfamiliar with the subject field of the data they are downloading. With additional contextual details, for example, details on the various types of MDE users and likely uses for the data, the need to supply coherent data may be further appreciated by developers.
- Addition of 'comprehensible data requirement' to the Data Clause Obligation – reference could be made to vital information that needs to be provided, such as column headers, code definitions, units, and vertical datums. In addition, developers could be instructed to check that the data is comprehensible before delivery to The Crown Estate.
- Addition of 'GIS data requirement' to the Data Clause Obligation – GIS shapefiles should be provided by developers if they are available, as they are a valuable resource for data users. This includes supporting information (e.g. survey area extent polygon).
- Widen the range of data structure examples provided - at present, an example of a Geophysical series with one dataset is provided. Most geophysical series have multiple datasets, hence this example could be expanded upon to show how separate datasets should be structured within large series. In addition, examples could be added for the other common themes, such as Ornithology and Marine Mammals, Geotechnical and Benthic. Presenting examples visually, for example, in the format of a data tree, may aid interpretation.

7.2 Metadata Recommendations

To improve metadata compliance with the MEDIN discovery metadata standard, enhancements could be made to Appendix 5 (Metadata Guidance Note) of the Requirements' document, such as:

- Improve the summary table of MEDIN metadata requirements – greater detail could be provided within this table, for example, example titles and abstracts, explanation of the various date types, a recommended number of keywords and reference to The Crown Estate keywords.
- Request that developers always follow the MEDIN recommended title format – MEDIN's recommended title format is <'Date' 'Originating organisation/programme' 'Location' 'Type of survey'>. This should improve consistency and data discovery for MDE users.
- Add the list of The Crown Estate keywords – introduce a mandatory requirement for one development area and one development phase keyword to be specified.
- Emphasize the importance of accurate metadata for data discovery – this may reduce the occurrence of incorrect temporal extents, spatial resolutions etc.
- Remind developers to include all mandatory details - this could reduce instances of metadata being provided without lineage information.

8. Conclusion

Numerous issues have been highlighted within this report, some of these issues occur within a wide range of themes, whilst others are confined to a small number of themes. A large proportion of the issues flagged involved missing components, such as reports, datasets and appendices. Developer-specific statistics have been provided, which will enable developers to be targeted based on the most prevalent issues associated with their development sites. A range of recommendations have been proposed, which include communicating with developers, and enhancing the Requirements document. These recommendations should result in improvements to the discoverability and comprehensibility of the supplied data, and ultimately improve the experience for users of the MDE.

9. References

Seeley, B., Rapaport, J., Merritt, O. and Charlesworth. M (2014) Guidance notes for the production of discovery metadata for the Marine Environmental Data and Information Network (MEDIN), Marine Environmental Data Information Network Standards Working Group, Version 2.3.8 [Online]: http://www.oceannet.org/marine_data_standards/medin_disc_stnd.html

The Crown Estate (2013) Marine Data Exchange to boost 'Blue Economy', [Online] Available from: http://www.marinedataexchange.co.uk/announcements/2012/marine-data-exchange-to-boost-“blue-economy”.aspx, Accessed 14th May 2014.

The Crown Estate (2012) Requirements for providing survey data to The Crown Estate via the Marine Data Exchange, The Crown Estate, Requirements Note 001, Version 1.2.

10. Abbreviations

ABPmer	ABP Marine Environmental Research Ltd
AIS	Automatic Identification System
AGDS	Acoustic Ground Discrimination System
EPSG	European Petroleum Survey Group
GIS	Geographic Information System
MDE	Marine Data Exchange
MEDIN	Marine Environmental Data and Information Network
QA	Quality Assurance
QC	Quality Control

Cardinal points/directions are used unless otherwise stated.

SI units are used unless otherwise stated.

ABP Marine Environmental Research Ltd (ABPmer)
Quayside Suite, Medina Chambers, Town Quay, Southampton SO14 2AQ

T +44 (0)23 80 711840
F +44 (0)23 80 711841
E enquiries@abpmer.co.uk

www.abpmer.co.uk

Creating sustainable solutions for the marine environment