



Quality Assurance procedures

The Marine Biological Association of the United Kingdom (MBA) is a centre of excellence in marine science, with over 125 years of experience of the delivery of quality science, advice and education to the United Kingdom, Europe and worldwide. The MBA is, therefore, committed to quality in all its activities.

The MBA has quality assurance procedures in place for its scientific research and knowledge exchange outputs and activities.

Scientific research

The scientific research activities of the MBA are governed by the MBA's Research Conduct Policy and, in particular, the section on Research Quality and Data Accessibility.

The responsibility for research quality rests with all staff. However, the relevant Principal Investigator (PI), Deputy Director (Research) and MBA Director are responsible ensuring the quality assurance of the research. It is their responsibly that:

- Research is conducted to the highest standard (in line with the Research Conduct Policy)
- Experimental design is statistically rigorous, state-of-the-art, recorded and described to enable repetition
- Statistical techniques and data interpretation are to the highest standard, rigorous, properly recorded and repeatable
- Scientific outputs (research proposals, scientific papers, research reports, posters, and presentations) are to a high standard.

To this end:

- PI approval is required for experimental design and statistical analysis.
- The PI 'signs off' all materials before they released or submitted to journals or other media

The MBA's scientific research is ultimately subject to peer review during publication in journals and by the Research Councils that fund the research.

Knowledge Exchange

The MBA's Knowledge Exchange activities cover a wide range of activities from contract research (Evidence Team), data management (Data Team) and education and community outreach (Education Team).

Responsibility for the quality of the outputs rests with all staff. However, the Team Leaders, Deputy Director (Knowledge Exchange) and MBA Director are responsible for the quality assurance. It is their responsibly that:

- All marine biodiversity data, its collation and management are subject to the Archive for Marine Species and Habitats Data (DASSH) (a MEDIN accredited marine Data Archive Centre) Quality Assurance procedure (appended).
- All research conducted under contract is subject to the MBA's Research Conduct Policy
- Experimental design is statistically rigorous, state-of-the-art, recorded and described to enable repetition
- Statistical techniques and data interpretation are to the highest standard, rigorous, properly recorded and repeatable
- Printed outputs (research reports, maps, posters, teaching materials) are to a high standard



- Electronic outputs (coding, software and databases) are to a high standard
- Online materials (web pages, digital publishing) are to a high standard.

To this end:

- The relevant Team Leader is required to sign-off all deliverables prior to release, and at each stage of delivery, that is draft and final stages.
- The Team Leader will also ensure that other relevant Senior Scientists at the MBA review the outputs where additional expertise is required.
- Sign-off procedures reviews the scientific content for accuracy, completeness and appropriate use of statistical and other analytical techniques.
- Sign-off procedures also ensure that the printed materials are written in plain English appropriate to the intended audience and that the printed materials are presented clearly, in line with the MBA's house style or clients preferred style.
- The Team Leader (or relevant web site section editor) is required to 'sign-off' relevant online materials in accordance with MBA house style and to ensure accuracy of content and writing style (materials include web page content, news items and press releases).
- The relevant Team Leader is required to test and sign-off data products (e.g. online and mobile device tools, and database solutions) before release.

HTW 2011-05-12



DASSH - Quality assurance of data and metadata.

Raw data

Raw data is initially checked against the metadata provided to ensure that the data provided matches the description of the data given. Data points are plotted in GIS to ensure that co-ordinates fit in the bounding box and match the reported extent of the dataset, i.e. does a Thames Estuary survey produce data points that map in the Thames Estuary. Records of data collection and other QA procedures applied by the data provider are recorded in the lineage element of the metadata.

Species records are checked against the NBN / NHM Species Dictionary on entry to Marine Recorder and using in-house macros to check Excel spreadsheet species information. Biotope data is checked against the relevant classification (i.e. Connor *et al.*, 1997; 2004). Any unusual or odd records are referred to the data provider and, if necessary, to local expert marine biologists for confirmation. Unusual records include new species sightings, species outside their recorded distribution range and/or in unexpected habitats.

Where data has been collected without quality assurance processes, for example volunteer records, then unusual records would need confirmation via either a specimen or photograph which can be referred to a taxonomic expert where required. For historical records where such rigorous confirmation may not be possible then the record will be marked as uncertain.

The raw data is archived in its original form. Any amendments to the datasets prior to dissemination, e.g. typographical errors, species name changes, etc are documented.

Metadata

Manual quality assurance

Initial QA of metadata is undertaken by checking the metadata information for each field against the source document(s) (e.g. data files and reports). Metadata cannot be entered and QA'd by the same person. If any updates are made the person reviewing the metadata then the date of metadata update is recorded. Final metadata is referred to the data provider for an additional check, to ensure accuracy where possible.

DASSH metadata QA procedure is based on the 'Metadata Guidelines for Geospatial Datasets in the UK' guidelines produced by the Department for Communities and Local Government and published by GI-Gateway.

Validation against the Marine Environmental Data and Information Network (MEDIN) Metadata XML Schema

Once manually checked, the metadata is converted into an eXtensible Markup Language (XML) document based on the MEDIN Discovery Metadata Standard. These XML documents must validate against the schema before they can be harvested by the NERC Data Grid software for display on the MEDIN Metadata Portal. XML validation is carried out automatically using the Schematron web services for the MEDIN Standard and ISO19139.

Processed survey data

DASSH routinely plots the survey bounding box and the survey points, using in-house GIS tools, to check the dataset.

Marine Recorder

Where data has been processed in Marine Recorder, the JNCC validation procedures are used to QA datasets (see <http://www.esdm.co.uk/MarineRecorder/index.html> for full documentation



relating to the validation procedure). Completed validation forms are stored with the archived data. Data is not entered and QA'd by the same person.

Guidelines for the number of survey events¹ to QA are outlined in Table 1 and some flexibility is allowed to allow for differences in QA of data inputted directly (e.g. from paper reports) and data imported electronically in which data copying errors are less likely. If errors are found in the QA sample the dataset the entire dataset is subject to further QA.

Table 1 – Guidelines for number of survey events per dataset to QA.

Number of survey events	Number of survey events to QA
1-30	50-100%
31-60	25-50%
61-100	20-25%
101+	10-20%

DASSH datasets sent to the National Biodiversity Network

Datasets that are sent to the National Biodiversity Network (NBN) are subject to additional validation. DASSH uses the NBN Validator software to QA records sent to the NBN from Marine Recorder and ensure that data meets NBN requirements.

Dealing with changes in taxonomy.

Taxonomy continues to change, as new research evidence comes to light. In some datasets, the taxonomic names provided are no longer recognised and these names are cross referenced with the NBN/NHM Species Directory and World Register of Marine Species (WoRMS) database and updated where necessary. In such cases, the original taxon and the taxon that it has been replaced with will be included in the lineage element of the metadata for reference purposes.

¹ Survey Event - A single geographic location within a survey e.g. a sampling station. Survey events are also temporally distinct so the same location surveyed at different times within a year would constitute different survey events.