NMBAQC

NE Atlantic Marine Biological Analytical Quality Control Scheme

Benthic Invertebrate Component Annual Report
Scheme Operation 2017 / 2018 (Year 24)

Authors: Tim Worsfold, NMBAQCS Benthic Invertebrate AQC Analyst
David Hall, NMBAQCS Project Manager

Approved by: Myles O'Reilly, Contract Manager, SEPA

Contact: nmbaqc@apemltd.co.uk

APEM Ltd.
Date of Issue: July 2018

BENTHIC INVERTEBRATE COMPONENT ANNUAL REPORT FROM APEM Ltd

SCHEME OPERATION – 2017 / 2018 (Year 24)

Linked Documents (hyperlinked in this report):

Ring Test Bulletin – RTB#53
Ring Test Bulletin – RTB#54
Laboratory Reference Module Summary Report – LR22
Own Sample Module Summary Report – OS65, 66 & 67
Description of the Scheme Standards for the Benthic Invertebrate Component
Guidelines for Processing Marine Macrobenthic Invertebrate Samples
Ring Test Protocol
Laboratory Reference Protocol
Own Sample Exercise Protocol
Own Sample Interim Report Review and Remedial Action Processes

1. Introduction

The North East Atlantic Marine Biological Analytical Quality Control (NMBAQC) Scheme addresses three main areas relating to benthic biological data collection:

- The processing of macrobenthos samples;
- The identification of macrobiota;
- The determination of physical parameters of sediments.

Scheme year 2017 / 2018 (year 24) followed the format of year 2016 / 2017. A series of components, modules and exercises involved the distribution of test materials to participating laboratories and the centralised examination of returned data and samples. The labelling and distribution procedures employed previously have been maintained. Specific details can be found in previous Scheme annual reports.

Forty-nine laboratories (with multiple participants from some organizations counted separately) participated in the Benthic Invertebrate Component of the NMBAQC Scheme in 2017 / 2018 (year 24). Seventeen of the participants were UK Competent Monitoring Authorities (CMAs), responsible for the Clean Seas Environment Monitoring Programme (CSEMP) or Water Framework Directive (WFD) sample analysis; twenty-nine were private consultancies, one of which was a consortium of sole traders. Seven of the participants were non-UK laboratories (including three government organizations and four private consultancies). Laboratory Codes were assigned in a single series for all laboratories participating in the Benthic Invertebrate component. Separate Laboratory Codes were assigned for the other scheme components, such as the particle size component.

As in previous years, some laboratories elected to be involved in limited aspects of the scheme. UK Competent Monitoring Authorities (CMAs) completing benthic biological analyses for monitoring programmes, including the assessment of MPAs (Marine Protected Areas), as evidence under MSFD (Marine Strategy Framework Directive), WFD (Water Framework Directive) and the CSEMP (Clean Seas Environmental Monitoring Programme), must participate in the Benthic Invertebrate component. CSEMP / WFD laboratories are no longer required to participate in all components / modules of the scheme.

In this report, performance targets have been applied for the Own Sample module only (see Hall, 2010: Description of the Scheme Standards for the Benthic Invertebrate Component). These targets have been applied to the results from laboratories and 'Pass' or 'Fail' flags assigned accordingly. These flags are indicated in Table 1 of the Own Sample Module Summary Report – OS65, 66 and 67 (2017/2018 (Year 23) OS Module Summary) presenting the comparison of laboratory results with the standards.

1.1 Summary of Performance

This report presents the findings of the Benthic Invertebrate component for year 2017 / 2018 (year 24) of the North East Atlantic Marine Biological Analytical Quality Control (NMBAQC) Scheme.

This component comprised three modules (each with one or more exercises):

- Own Sample module (OS) - re-analysis by APEM Ltd. of three samples supplied by participating laboratories;
- Invertebrate Ring Test module (RT) - identification of two sets of twenty-five invertebrate specimens; and
- Laboratory Reference module (LR) - re-identification by APEM Ltd. of a set of up to twenty-five specimens supplied by participating laboratories.

The analytical procedures of the various modules were the same as for 2016 / 2017 (year 23) of the Scheme. The results for each of the Scheme exercises are presented and discussed. Comments are provided on the performance of participating laboratories in each of the exercises.

Two **Ring Tests (RT)**, each of 25 specimens, were distributed (RT53 and RT54). The second (RT54) was targeted on Spionidae, to follow the 2016 Scheme experts workshop, which included study of Spionidae and the development of an identification guide. A new Ring Test Protocol (Worsfold & Hall, 2017a) was produced to explain and standardise the methods and policies used in the module.

For RT53, the average numbers of differences per participating laboratory (for a total of 23 laboratories with 23 submissions) were 4.3 generic differences and 8.7 specific differences.

Six species (four annelids, one mollusc and one crustacean) were responsible for almost half (48%) of the specific differences.

For RT54, the average numbers of differences per participating laboratory (for a total of 21 participants) were 2.1 generic differences and 5.3 specific differences. Seven specimens (small, damaged *Malacoceros vulgaris, M. tetracerus, Dipolydora* 'species B'*, D. quadrilobata, Aurospio banyulensis, Pseudopolydora* 'species A'*,* and *Prionospio plumosa*), were responsible for three fifths (60%) of the specific differences.

Laboratory Reference (LR): Seven laboratories signed up for the LR22 module and four laboratories submitted specimens for confirmation, within the required deadline. A fifth laboratory submitted specimens for confirmation after the deadline; these were reported separately but not included in the statistics for this annual report. Most misidentifications were for Annelida (49%), followed by Mollusca (35%) and Crustacea (11%); many belonged to genera which are either speciose, or for which the taxonomy has yet to be finalized. A new Laboratory Reference Protocol (Hall & Worsfold, 2017) was produced to explain and standardise the methods and policies used in the module.

A new Own Sample Exercise Protocol (Worsfold & Hall, 2017b) was produced to explain and standardise the methods and policies used in the **Own Sample (OS)** module, including details of audit sample selection and determination of 'associated samples' for subsequent remedial actions. Laboratories were asked to submit full completed data matrices from their previous year's CSEMP / WFD, or similar alternative sampling programmes. The OS 'Pass / Fail' flagging system, introduced in Scheme Year 8, was continued (see Hall, 2010: Description of the Scheme Standards for the Benthic Invertebrate Component). In OS65-67, extraction efficiency (of individuals) was better than 90% in 96% of the comparisons and better than 95% in 86% of all comparisons. 100% of countable taxa were extracted from the sample residues in 65% of samples. The Bray-Curtis similarity index ranged from 42% to 100% with an average of 95.5%. The Bray-Curtis similarity index was greater than 95% in 73% of comparisons; in 89% of cases, the value of the index was greater than 90% and, therefore, achieved 'Pass' flags. Twelve samples (15%) achieved 'Pass-Excellent' flags with Bray-Curtis similarity scores of 100%.

An update to the Scheme's taxonomic literature database was produced as a text document Bibliography of taxonomic literature (Worsfold et al., 2018). This lists over 3,100 citations for identification literature for northeast Atlantic marine and brackish water biota by taxonomic group, with sections for benthic invertebrates, fish, benthic algae, zooplankton, phytoplankton and non-native species.

To a large extent as a result of work through the Scheme's Benthic Invertebrate Component, the contractor identified several anomalies in the World Register of Marine Species (WoRMS) through the Scheme year, some of which had caused problems with audits and ring tests. They were brought to the attention of WoRMS editors and, in most cases, resolved. This process had also been carried out in other years, including several (mainly cirratulids) that related to previous contract periods but were completed by the current contractor. The opportunity is taken to list those WoRMS edits initiated by the contractor over the current contract period:

- *Odostomia conspicua* to *Megastomia conspicua*; Serge Gofas, 07/07/2017;
- *Paraspio decorata* to *Spio decorata*; Geoff Read, 18/09/2017;
- *Parametaphoxus fultoni* to *Metaphoxus fultoni*; Tammy Horton, 05/10/2017;
- *Trichobranchus sikorskii* to *Octobranchus sikorskii*; Geoff Read, 15/12/2017;
- *Chrysallida sarsi* to *Parthenina sarsi*; Serge Gofas, 19/04/2018;
- *Palaemon yuna*, added; Sammy De Grave, 22/02/2018;
- *Palaemon leucurus*, authority corrected; Sammy De Grave, 22/02/2018.

1.1.1 Statement of Performance

Each participating laboratory was supplied with a 'Statement of Performance', which included a summary of results for each of the Scheme modules and details of the resulting flags, where appropriate. These statements were first circulated with the Year 5 annual report (1998 / 1999) for the purpose of providing evidence of Scheme participation and for ease of comparing year on year progress.

2. Summary of Benthic Invertebrate Component

2.1 Introduction

There are three modules within the Benthic Invertebrate component: Invertebrate Ring Test (RT), Invertebrate Laboratory Reference (LR) and Own Sample (OS) modules.

Each of these modules is described in more detail below. A summary of their performance with respect to standards determined for the CSEMP / WFD is presented. A brief outline of the information obtained from each module is given, together with a description of the preparation of the necessary materials and brief details of the processing instructions given to each of the participating laboratories.

2.1.1 Logistics

The labelling and distribution procedures employed previously have been maintained. Specific details can be found in the Scheme's new Benthic Invertebrate component protocols: Laboratory Reference Protocol (Hall & Worsfold, 2017), Ring Test Protocol (Worsfold & Hall, 2017a) and Own Sample Exercise Protocol (Worsfold & Hall, 2017b).

2.1.2 Data Returns

Return of data to APEM Ltd. followed the same process as in previous Scheme years. Spreadsheet-based forms (tailored to the receiving laboratory) were distributed to each laboratory via email. All returned data were converted to Excel 2010 format for storage and analysis. In this, and previous, Scheme years, slow or missing returns for exercises lead to delays in processing the data and resulted in difficulties with reporting and rapid feedback of results to laboratories. Reminders were distributed shortly before each exercise deadline.

2.1.3 Confidentiality

In July 2017, each participant was given a confidential, randomly assigned 2017 / 2018 (Scheme year 24) LabCode. Codes are prefixed with the component initials (*e.g.*, BI for Benthic Invertebrates), the Scheme Year and a unique number (between 01 and 46); *e.g.* Laboratory number one in Scheme Year 2017 / 2018 (Year 24) was recorded as BI_2401. Laboratory codes, with a PSA_ prefix, were assigned separately for the Particle Size component (also administered by APEM Ltd.).

2.2 Invertebrate Ring Test (**RT**) Module

2.2.1 Description

The Invertebrate Ring Test module is a training module which examines variation in participants' ability to identify different species and attempts to determine whether differences are the result of literature deficiencies, lack of reference material or

misinterpretation of identification resources. Details are explained in the new Ring Test Protocol (Worsfold & Hall, 2017a)

Two sets of 25 benthic invertebrate specimens were distributed in 2017 / 2018. The first circulation (RT53) was a general invertebrate ring test. It included 11 (44%) annelids, 5 (20%) molluscs, 8 (32%) crustaceans and 1 (4%) taxon belonging to other phyla. An effort was made to include a high proportion of species that had not previously been circulated through the module and that would highlight taxonomic problems. The second circulation (RT54) was targeted at spionid polychaetes, in order to progress the development of the identification guide under construction following the 2016 Scheme workshop. It included 25 (100%) spionid polychaetes. Basic notes on substratum, salinity, depth and geographical region were provided for all ring test specimens to assist identification.

2.2.1.1 Preparation of the Samples

The specimens distributed were obtained from a range of surveys from around the North-East Atlantic. Some specimens were donated by Scheme participants and other organizations. Care was taken to provide animals of similar size and condition and, where relevant, of the same sex, for each laboratory. Each specimen was uniquely identifiable by means of a coded label and all material has been retained for subsequent checking. For both ring tests, the specimens were taken from samples within a single survey and, in most cases, they were from a single sample, or replicates from a single sampling station. For RT54, the most recent update to the Scheme's spionid guide (Radashevsky, 2017) was circulated to participants for use with the RT and specimens from the same projects (usually the same samples) as the RT specimens were sent to Vasily Radashevsky for confirmation of identifications.

2.2.1.2 Analysis Required

The participating laboratories were asked to identify each of the RT specimens to species level and they were also asked to complete a 'confidence level' field to indicate whether they would ordinarily have left the specimen at a higher taxonomic level. Laboratories could also add brief notes and information detailing the literature used to determine their identifications. Specimens were to be returned to APEM Ltd. for verification, resolution of any disputed identifications and potential reuse in future Scheme exercises. The implementation of this part of the Scheme was the same as in previous years. Participating

laboratories were permitted to supply multiple returns (*i.e*. different sets of results from different analysts) for each exercise to enhance the training value of the module. No laboratories chose to utilise this option in RT53, or for RT54. The protocols followed for the two circulations, particularly the method of counting differences, were the same as for previous circulations. Approximately eleven weeks were allowed for the analysis of RT53 (including postage delays due to cyber-attack on the courier company); approximately twelve weeks were allowed for RT54, to allow time for circulation of the draft spionid guide.

2.2.2 Results

2.2.2.1 General Comments

Several laboratories use the ring tests for training purposes and select them preferentially over other modules. The results are not used to assign 'Pass' or 'Fail' flags. In total, 24 laboratories subscribed to RT53 and 23 laboratories subscribed to RT54. For RT53, 23 laboratories returned data (23 individual data sets). For RT54, 21 laboratories returned data (21 individual data sets).

2.2.2.2 Returns from Participating Laboratories

Identifications made by the participating laboratories were compared with those made by APEM Ltd. to determine the numbers of differences. Where identifications deviated from the APEM Ltd. identification due to the use of synonyms, or incorrect spellings of the name, the difference was ignored for the purpose of calculating the total number of differences.

Tables 1 and 2 of Ring Test Bulletins (RTB) 53 and 54 show identifications made by each of the participating laboratories for the twenty-five specimens, arranged with laboratories as rows and specimens in columns in Table 1; specimens as rows and laboratories as columns in Table 2. For clarity, the participant's identification is given only where the name given by the laboratory differed from the APEM Ltd. identification. Where it was considered that the name referred to the same species as the APEM Ltd. identification, but differed for one of the reasons indicated above, the name is presented in brackets: "[name]". A dash, "-", in the tables indicates that the name of the genus (and / or species) given by the laboratory was the same as the APEM Ltd. identification. A pair of zeros, "0 0", in the Tables indicates that the subscribing laboratory did not return data.

2.2.2.2.1 Counting RT Result differences

For each laboratory, a count was made of each difference between their identification and the APEM Ltd. identification (*i.e*. for each instance where text other than a dash or a bracketed name appears in the appropriate column in Tables 1 and 2 for RTB53 and RTB54). Separate counts were maintained for differences at genus and species level.

2.2.2.3 Ring Test Results

The intention of this training module is to discover where difficulties lie in the identification of certain taxa. Results for Scheme Year 2017 / 2018 were presented in the Ring Test Bulletins (RTB) along with the reasons for each identification discrepancy. These bulletins contain images of the test material and of all available taxa that were named as alternative identifications by participants. Participating laboratories were advised to retain ring test specimens for a few weeks after receiving their results, in order that they could review their identifications, if necessary. Participants are encouraged to question APEM Ltd. identifications if they still believe their original identifications to be correct. On completion of each exercise, specimens were required to be returned to APEM Ltd. for reference and / or potential future circulation.

2.2.2.3.1 Ring Test 53 (Type: General)

The results discussed below are given in Table 1 of RTB53, which displays the data arranged with columns for species to enable quick reference to the range of answers received and in Table 2, which presents the results arranged with columns for laboratories (see Ring Test Bulletin RTB53).

Eleven of the 25 specimens circulated were annelids, five were molluscs, eight were crustaceans and one was from other taxa (Sipuncula). The agreement at generic level was generally good; 98 differences (17% of all genus identifications received from participants) were recorded in the 23 data sets received from 23 participating laboratories. There was less agreement at species level, with 201 differences recorded (35% of all species identifications received from participants).

Six of the specimens circulated were responsible for almost half (48%) of participants' species level identification differences. These were the annelids *Amythasides macroglossus,*

Claviramus candelus, *Galathowenia fragilis* and *Paradoneis ilvana*; the mollusc *Littorina compressa*; and the crustacean *Nebalia kocatasi*.

Four of the 25 specimens circulated (the annelid *Levinsenia gracilis*, the crustaceans *Pasiphaea sivado* and *Praunus flexuosus* and the sipunculan *Onchnesoma steenstrupii*) were correctly identified by all participants.

Further details and analysis of results can be found in the Ring Test Bulletin RTB53, which was circulated to each laboratory that supplied results for this exercise and was also posted on the Scheme's website (www.nmbaqcs.org).

2.2.2.3.2 Ring Test 54 (Type: Targeted on Spionidae)

The results discussed below are given in Table 1 of RTB54, which displays the data arranged with species as columns to enable quick reference to the range of answers received and in Table 2 which presents the results with laboratories as columns (see Ring Test Bulletin RTB54).

All 25 of the specimens circulated were spionid polychaetes. The agreement at genus level was good; 45 differences (9% of all genus identifications received from participants) were recorded in the 21 data sets received from 21 participating laboratories, of which 10 differences were for one species (small, damaged *Malacoceros vulgaris*). There was less agreement at species level, with 112 differences recorded (21% of all species identifications received from participants).

Seven of the specimens circulated were responsible for three fifths (60%) of participants' species level identification differences. These were small, damaged *Malacoceros vulgaris, M. tetracerus, Dipolydora* 'species B'*, D. quadrilobata, Aurospio banyulensis, Pseudopolydora* 'species A'*,* and *Prionospio plumosa*.

Four of the twenty-five specimens circulated (*Aonides paucibranchiata, A. oxycephala, Pygospio elegans* and *Poecilochaetus serpens*) were correctly identified by all participants.

Further details and analysis of results can be found in the Ring Test Bulletin RTB54, which was circulated to each laboratory that supplied results for this exercise and was also posted on the Scheme's website (www.nmbaqcs.org).

2.2.2.4 Differences between Participating Laboratories

Differences recorded at genus and species level for each of the participating laboratories are summarised in the graphs related to Table 2 in RTB53 and RTB54 respectively. The laboratories are ordered by increasing number of differences at species level. The division of laboratories into three bands (Low, Mid and High) on the basis of the number of differences at species level is also shown.

2.2.2.5 Differences by Taxonomic Group

The total differences by taxonomic group (combined for both exercises) are shown below:

Major taxon	Species circulation	Generic differences		Specific differences	
Annelida	36	92	64.3%	203	65.0%
Crustacea	8	27	18.9%	51	16.2%
Mollusca	5	24	16.8%	59	18.8%
Others	1	0	0%	0	0%
Total	**50**	**143**	**100%**	**312**	**100%**

Most of the specific differences in RT53 were for annelid species (RT54 was entirely spionid polychaetes).

2.2.3 Discussion

The results were in general comparable with those from previous exercises, with an average of 4.3% generic and 8.7% specific differences across the participating laboratories in RT53 and 2.1% generic and 5.3% specific differences across the participants in RT54.

In RT53, some significant differences (e.g. for *Nebalia kocatasi, Eugerda tenuimana* and *Ruditapes philippinarum*) were the result of lack of knowledge of literature and recent taxonomic work (citations were provided in the bulletin). Others (e.g. for *Littorina*

compressa, Potamopyrgus antipodarum and *Amythasides macroglossus*) were due to inherent difficulties in recognition of identification features for the species. For some (e.g. for *Claviramus candelus* and *Galathowenia fragilis*), there were acknowledged remaining taxonomic problems and nomenclature problems exacerbated some differences for *L. compressa* and *E. tenuimana*. The high error rates for *P. antipodarum* and *R. philippinarum* are notable as they are non-native species. Many RT54 differences (e.g. for *Dipolydora* 'species B'*, Aurospio banyulensis* and *Pseudopolydora* 'species A') were due to inherent difficulties in recognition of identification features for the species, particularly as they are rarely in perfect condition; the circulated specimens had varying degrees of damage (though as similar as possible within each circulated species), to reflect conditions encountered in samples and some high numbers of differences (e.g. for the smaller *Malacoceros vulgaris*) may have been mainly due to specimen condition. For some species (especially *Dipolydora* 'species B' and *Spio cf. symphyta*), the circulation highlighted taxonomic problems yet to be resolved. RT54 was circulated with a draft guide to UK spionids and significant progress has been made since, with help from this ring test.

The RT component is considered to be a valuable training tool and can be an indicator of problem groups. It can highlight possible taxa for further 'targeted' ring test exercises or for inclusion at taxonomic workshops. The allowance of multiple submissions per laboratory and the inclusion of images in the Ring Test Bulletins have enhanced the training value of this component. All participating laboratories have been made aware of the problems identified by these ring tests via Ring Test Bulletins RTB53 and RTB54, which also include literature citations that relate to the problem taxa.

2.3 Invertebrate Laboratory Reference (**LR**) Module

2.3.1 Description

The Laboratory Reference module is a training module which encourages laboratories to build reference collections to improve identification consistency and to seek additional opinions for difficult specimens. The value of reference material in assisting identification cannot be over-emphasized; the creation and use of reference collections is viewed as best practice. Accordingly, the Laboratory Reference (LR) module of the Scheme was introduced in Scheme Year 3 (1996 / 1997). This module can help participating laboratories to assess their ability to identify material from their own samples. Laboratories are also able to use this exercise to obtain second opinion identifications for difficult or problematic taxa of

which they are unsure. This was the twenty-second Laboratory Reference exercise (LR22). The participants were able to submit up to 25 specimens for re-examination by APEM Ltd.

2.3.1.1 Preparation of samples

A prepared results sheet was distributed with the exercise's instructions and attached labels for the laboratories to identify each of the specimens. Participating laboratories were asked to prepare and submit their reference specimens within 6 weeks. All specimens were re-identified by APEM Ltd., with comparisons to the original identifications. All specimens were returned to the laboratories after analysis.

2.3.2 Results

Seven laboratories signed up for this exercise (LR22) but only four submitted specimens for examination, within the required deadline. A fifth laboratory submitted specimens for confirmation after the deadline; these were reported separately but not included in the statistics for this annual report. Detailed results have been separately reported to each participating laboratory. Taxonomic edits were made for submitted polychaetes (18; 49%), molluscs (13; 35%) and crustaceans (4; 11%). In addition, differences were noted for taxonomic resolution, recording notation and spelling for many specimens. A report summarising the results from this module is presented in the Laboratory Reference Module Summary Report – LR22.

2.3.3 Discussion

As with all training exercises, detailed inter-laboratory comparisons are of limited value. Some of the differences related to recent literature updates (e.g. *Nebalia reboredae*) and to problems considered in recent workshops (e.g. *Spio symphyta*, *S. armata*, *Scolelepis bonnieri*). The submitted specimens also included some that could not be named due to condition or APEM experience (e.g. *Laonice, Chone*) and at least one probable undescribed species (*Polydora* 'species A'). The taxonomic resolution and recording policy differences were defined according to the current standardized format designed for these exercises (Worsfold, 2017), with a view to the later development of a taxonomic discrimination protocol.

2.4 Own Sample (**OS**) Module

2.4.1 Description

The Own Sample module examines analytical performance on material from each participating laboratory's annual CSEMP / WFD or other sample analysis batches. Following a review of the Own Sample module (Hall & Worsfold, 2001), several changes to sample selection and scoring were implemented in Scheme Year 8 (2001 / 2002). All participants must meet these Own Sample requirements. The Own Sample Exercise Protocol was updated in August 2017 and circulated to all OS participants ahead of the module for this scheme year. Own Sample participants must supply their previous year's CSEMP / WFD data matrices, where relevant, for Own Sample selection, *i.e.* 2016 CSEMP / WFD data. This is to ensure that all processing is completed (prior to selection of samples for audit), preventing reworking of the selected Own Samples and enabling samples to be audited earlier in the Scheme year. Each participating laboratory was requested to send data from which three samples were selected and the selection notified to the laboratories. Laboratories responsible for CSEMP / WFD samples were advised to use these samples if possible; otherwise, there was free choice, provided a minimum of twenty samples were included in the submitted data matrix.

2.4.1.1 Analysis Required

Participating laboratories were instructed to have conducted macrobenthic analysis of the samples using standard procedures. A summary of sample details, including codes, area and sample processing procedures was to be provided, on a standard form, for each Own Sample. Samples requiring sub-sampling were to be avoided where possible. All procedures were documented and details returned with the sample components. All material from the sample was to be sent to APEM Ltd., broken down as follows:

- Sorted residue - material from which biota had been removed and counted;
- Separated taxa - individually labelled vials containing the identified biota; and
- Other fractions - *e.g*. material containing biota that had been counted *in situ*.

Recording and identification were assumed to have followed NMBAQC guidelines for macrobenthic sample analysis (Worsfold, Hall & O'Reilly (Ed.) 2010). The names and counts of specimens were to be recorded on a matrix and linked to the vials through a specimen code number. In addition, measurements of the biomass of the recorded taxa were

submitted where required; measurements were to be blotted wet weights to 0.0001g for each of the enumerated taxa.

Two weeks were allowed for the submission of data; a further six weeks were allowed for the preparation and submission of the Own Samples selected for re-analysis. The sorted residue was re-examined and any countable material or new non-countable taxa extracted. Identified biota were checked for accuracy of enumeration and identification and, in cases where biomass was provided, all taxa were re-weighed using the procedure outlined in the NMBAQC Sample Processing Protocol (Worsfold, Hall & O'Reilly (Ed.) 2010).

2.4.2 Results

2.4.2.1 General Comments

Following the request to participating laboratories to submit data of suitable samples for re-analysis, 79 selected Own Samples were received from 27 (of the 32 subscribing) laboratories, together with descriptions of their origin and the collection and analysis procedures employed. Samples were identified as OS65, OS66 and OS67 and labelled with LabCodes. As would be expected, the nature of the samples varied considerably. Samples were received from estuarine and marine locations, both intertidal and subtidal. The sediment supplied for resorting varied from mud to gravel in various volumes of residue. The number of taxa per sample ranged from 1 to 186, with the number of countable individuals from 1 to 2,980. Of the 79 submitted Own Samples, 9 were audited externally by Marine Invertebrate Ecological Services (MIES), as the initial processing had been carried out by APEM Ltd. Interim reports were submitted to participating laboratories. A summary of results from this module is presented in the Own Sample Module Summary Report – OS65, 66 & 67.

2.4.2.2 Efficiency of Sample Sorting

Table 1 of the OS Summary Report displays a summary of the data obtained from the OS analysis. All taxa recorded by the participating laboratory were included in the AQC analysis (if required to be recorded by the NMBAQC PRP/TDP). In 42 samples out of the total 79, the number of taxa recorded by the participating laboratories was identical to that obtained by the auditing laboratory (columns 2 and 3). For the remaining 37 cases, the difference was on average 2.4 with a maximum of 27 taxa. Data for the numbers of individuals recorded (columns 16 and 17, Table 1) show a range of differences from re-analysis of 0% to 19%. The

average difference between the 47 samples with recorded differences was 3.4% (and 2% across all 79 samples), with 13 samples exceeding this average.

32 of the 79 samples reported showed 100% extraction of individuals from the residue (column 16) and, in 37 samples, between 1 and 71 individuals had been missed during processing. In 20 samples, only individuals attributed to taxa already recorded in the sample were found. In 28 samples, new taxa, as well as individuals attributed to already recorded taxa were recorded. Numbers of previously unrecorded taxa found in the residue ranged from 0 to 20, with an average of 0.8 new taxa per sample. The poorest extraction records were: a total of 20 missed taxa and 71 individuals, 6 missed taxa and 8 individuals, and 3 missed taxa and 29 individuals. A breakdown of the missed individuals by taxonomic group is presented in Table 2 of the OS Summary Report. The average number (across all 79 samples) of missed individuals found upon re-sorting the residue was approximately 4 and the average number of missed taxa was less than 1.

2.4.2.3 Uniformity of Identification

Taxonomic differences (columns 10 and 11) between the auditor and participating laboratories' results were found in 46 (58%) of the 79 Own Samples. A summary of misidentified taxa is presented in Table 3 of the OS Summary Report. For the samples with taxonomic errors, an average of just over one taxonomic error per laboratory was recorded; in the worst instance, 14 identification errors occurred. A large variety of samples (and biota) was received. Polychaeta accounted for 34%, Mollusca for 30%, Crustacea for 20%, other for 13%, Oligochaeta for 2%, and Echinodermata for 1% of the taxonomic errors (approximately), with a variety of species responsible for these errors.

2.4.2.4 Comparison of Similarity Indices (Bray-Curtis)

The procedure for the calculation of the similarity index was as used for the Own Sample exercise in Year 2016 / 2017 (Year 23). The Bray-Curtis similarity index figures (Table 1, column 23) ranged from 42% to 100%, with an average of 95%. Nine samples from six laboratories achieved a similarity figure of less than 90%. Twelve samples produced a similarity figure of 100%; these were submitted by six different laboratories (BI_2404, BI_2405, BI_2409, BI_2417, BI_2439 and BI_2442). The best overall result was achieved by laboratories BI_2417 and BI_2439, with 100% similarity across all three Own Samples. The

lowest overall result was achieved by BI_2441 with an average similarity index of less than 79% over all three samples (due to one 'bad' failing sample).

2.4.2.5 Biomass Determinations

It was not possible to make an accurate comparison of biomass determinations in all cases; 53 samples had not been supplied with species blotted wet weight biomass data. Consequently, only 26 of the 79 samples received were used for comparative analysis. Table 4 of the OS Summary Report shows the comparison of the participating laboratory and APEM Ltd. biomass figures by major taxonomic groups. The total biomass values obtained by some of the participating laboratories varied greatly compared to those obtained by APEM Ltd. Differences in the recorded biomass ranged from -48% to +35%. The reason for these large differences is likely to be a combination of variations in apparatus (*e.g.* calibration) and operator technique (*e.g.* period of and effort applied to drying). These figures are not comparable to those produced by the same module in each of the previous years due to the variability in the duration and method of drying and the consistency of results within each major taxonomic group. The APEM Ltd. biomass data were achieved using a non-pressure drying procedure as specified in the Green Book and the NMBAQC guidelines for macrobenthic sample analysis (Worsfold, Hall & O'Reilly (Ed.) 2010).

2.4.3 Discussion

It is evident that some laboratories use the Scheme as a complete audit check of their entire year's work, whereas others chose certain projects for submission, and may even do so prior to analysis. The latter approach would undermine the purpose of auditing, if the analyst(s) know beforehand which surveys, projects or samples were to be audited.

The average Bray-Curtis similarity index of 95% achieved for this Own Sample module shows that the agreement between the participating laboratories and APEM Ltd. was generally good.

There were 79 samples submitted for the Own Sample module, including the nine processed by the Scheme's external auditor. Of the 79 samples, 70 (89%) exceeded the 90% Bray-Curtis Pass mark and 58 (73%) exceeded 95% BCSI. Since the beginning of this module in Scheme Year 02, 83% of the samples received have exceeded the 90% Bray-Curtis Pass mark (see Table 5 of the OS Summary Report).

Since the beginning of the Own Sample module, 1,572 admissible samples have been received (OS01-67). Of these, 266 samples (17%) have fallen below the 90% Pass mark. Overall, these results are good and show the efficacy of the OS module, although a dip in quality was noticed in years 20 and 21 compared with the previous four years, there has been a marked improvement in 2015 / 2016 and this has continued in 2016 / 2017 and 2017 / 2018. Some participating laboratories should be able to further improve their results by reviewing their extraction methods and their use of taxonomic literature and identification aids.

2.4.4 Application of NMBAQC Scheme Standards

One of the original roles of the Benthic Invertebrate component of the NMBAQC Scheme was to assess the reliability of data collected as part of the CSEMP or WFD monitoring programmes; this has since been expanded to other data sets. With this aim, performance target standards were defined for certain Scheme exercises and applied in Scheme Year 3 (1996 / 1997). These standards were the subject of a review in 2001 (Hall & Worsfold, 2001) and were altered in Scheme Year 8; each performance standard is described in detail in the Description of the Scheme Standards for the Benthic Invertebrate Component document. Laboratories meeting or exceeding the required standard for a given exercise would be considered to have performed satisfactorily for that exercise. A flag indicating a 'Pass' or 'Fail' would be assigned to each laboratory for each of the exercises concerned. It should be noted that, as in previous years, 'flagging' has been applied only to the Own Sample module. A review of the formats used in recording identification differences was produced last year (Worsfold, 2017).

As the Scheme progresses, additional exercises may be included. In the meantime, the other exercises of the Scheme as presented above are considered of value primarily as training exercises or to inform policy and future developments.

2.4.4.1 Laboratory Performance

The target values for each Own Sample and the corresponding laboratory results, including the assigned flags are presented in Table 1 of the OS Summary Report. Although laboratories are requested to follow NMBAQC guidance, detailed comparisons of results between different laboratories are generally not applicable, due to the diversity of samples

analysed and some minor inter-laboratory variations in processing methodologies, especially in relation to identification. Development of more detailed taxonomic discrimination protocols may help resolve some of the latter discrepancies.

Table 1 (columns 5, 15 and 26) shows 'pass / fail' results for three of the OS targets: the enumeration of taxa, enumeration of individuals and the Bray-Curtis comparison. Twenty-one of the 27 participating laboratories achieved a Bray Curtis of >90% ('pass' flag) for all three of their Own Samples. Overall, 92% of the comparisons were considered to have passed the enumeration of taxa standard, 97% exceeded the enumeration of individuals standard and 89% passed the Bray-Curtis comparison standard (>90%). NMBAQC Scheme sample flags have been applied to each of the Own Samples, in accordance with the performance flagging criteria introduced in Scheme Year 08 (Table 1, column 26); 3 samples (4%) are flagged as 'Fail - Bad', 6 (8%) as 'Fail - Poor', 12 (15%) as 'Pass - Acceptable', 46 (58%) as 'Pass - Good' and 12 (25%) as 'Pass - Excellent' for their Bray-Curtis similarity indices. All the laboratories with 'Poor' or 'Bad' sample flags have been provided with specific recommendations of remedial actions to quality assure their Own Sample data sets (see 2.4.4.3 Remedial Action below).

Performance with respect to the biomass standard was generally good (Table 1, column 22), with 81% of the samples with submitted biomass values meeting the required standard.

2.4.4.2 Comparison with Results from Previous Years

A comparison of the overall results for recent years is presented in Table 5 of the OS Summary Report (Own Sample Module Summary Report – OS65, 66 & 67). The table shows the number of laboratories assigned 'Pass' and 'Fail' flags for the OS exercises over the past twenty three years based upon the current NMBAQC Scheme standards (see Description of the Scheme Standards for the Benthic Invertebrate Component). This year's 79 Own Samples resulted in a pass rate of 89% (see Table 5 in the Own Sample Module Summary Report), which is a 3% increase from the previous scheme year. Historically, the highest pass rate achieved was 100% in exercise OS01 (1995 / 1996; Year 2) that involved just fourteen samples; the lowest pass rate was 67% recorded in 2000 / 2001 (Year 7) from 45 samples.

2.4.4.3 Remedial Action

It is important that failing samples audited through the Own Sample module, are addressed (mandatory for CSEMP/WFD samples). Remedial action should be conducted upon the associated samples to improve the flagged data. The mechanism for identifying associated samples is described in the Own Sample Exercise Protocol. For a CSEMP/WFD sample, the associated samples would normally be those collected from the same station, stratum or water body. The revised NMBAQC Scheme OS standards, introduced in Scheme Year 08, give clear methods for discerning the level of remedial action required (see Description of the Scheme Standards for the Benthic Invertebrate Component). A failing Own Sample is categorised by a Bray-Curtis similarity index of <90%. The performance indicators used to determine the level of remedial action required are: % taxa in residue (missed taxa), % taxonomic errors, % individuals in residue (missed individuals) (see Table 1, columns 7, 10 and 17 in the OS Summary Report) and % count variance. Own Samples not achieving the required standards are monitored by the NMBAQC committee. Participating laboratories are expected to initiate remedial action according to the advice of the Scheme's contractor. APEM Ltd. or the NMBAQC Scheme Contract Manager should be notified when this has been completed. Any remedial action undertaken should be audited externally where required. The NMBAQC Contract Manager and Scheme's contractor, APEM Ltd., will provide clarification on specific details of remedial action or consider appeals relating to the remedial action process.

Below is a summary of the samples that have been assigned 'Fail' flags in Scheme Year 2017 / 2018 (Year 24). Six laboratories were responsible for nine 'failed' samples (some of these may include data that is reported to the CMA's, e.g. WFD samples). Remedial action, outlined below, was required for associated replicates of the following Own Samples:

Lab Code	OS no.	Remedial action	Notes
BI_2409	**OS65**	Review taxonomic errors	Remedial action completed 16/01/18
BI_2418	**OS66**	Review taxonomic errors	Remedial action completed 27/09/17

Lab Code	OS no.	Remedial action	Notes
BI_2425	**OS65**	Review methods/protocols for the extraction of individuals from sample residues, particularly bivalves	Remedial action completed 10/07/18
	OS67	Review methods/protocols for the extraction of taxa from sample residues	Remedial action completed 10/07/18
BI_2435	**OS65**	Review taxonomic errors, especially *Nucula* and *Bathyporeia*	Subcontracted analysis; remedial action completed 11/07/18
	OS66	Reprocess associated residues and review taxonomic errors	Subcontracted analysis; remedial action not complete
	OS67	Review extraction of taxa and taxonomic errors (particularly *Nucula*)	Subcontracted analysis; remedial action completed 11/07/18
BI_2441	**OS67**	Review processing/taxonomic error for dominant taxon	Remedial action completed 12/06/18
BI_2449	**OS66**	Review taxonomic errors and extraction of taxa	Subcontracted analysis; remedial action completed 11/07/18

Data captured 11th July 2018

3. Conclusions and Recommendations

Several observations may be made from the results of the exercises described above. The following is a summary of the major points of importance:

1. The majority of participating laboratories submit data / samples in accordance with the Scheme's timetable. **Late submissions**, however, are still the major contributing factor for delaying the production of exercise bulletins / reports. Laboratories should endeavour to report their results within the requested time, according to the deadlines circulated at the beginning of each Scheme year. It would be helpful if laboratories wishing to query Ring Test specimen identifications did so within a week of report receipt. These considerations would greatly facilitate the analysis of results and effective feedback.

2. The number of samples in **data sets provided for selection of Own Samples** varied considerably, with several laboratories offering less than the minimum 20 samples (due

to low volumes of sample processing) and other laboratories offering up to 556 samples across 18 projects for audit selection. Best practice for commercial laboratories should be to use the Scheme as an external auditor for most or all of their samples and no 'cherry picking', pre-analysis selection, or pre-submission re-working of samples should be undertaken. **Retention of sample residues** will be required to facilitate this and to ensure that any subsequent remedial actions can be adequately completed.

3. Revised data request and sample submission forms were introduced for the 2017 / 2018 OS module to capture **data / sample ownership**. Where data belong to CMAs, the submitting participant was required to declare this so that audit results could be shared accordingly and CMA data auditing could be tracked and co-ordinated.

4. There were continued **problems associated with the measurement of biomass** for individual species in the Own Sample module. In this and previous Scheme years, several laboratories, despite using blotted wet weight biomass techniques, rendered some of their specimens too damaged to be re-identified. Additionally, some laboratories had erroneous results where it appeared that biomass had been estimated or mis-transcribed. The initial processing of a sample should in no way compromise the effectiveness of an audit. Biomass procedures should not render the specimens unidentifiable. Biomass must be reported to four decimal places with nominal weights recorded as 0.0001g. A standardised protocol is available in the NMBAQC guidance document (Worsfold, Hall & O'Reilly (Ed.) 2010) and must be followed for CSEMP / WFD analysis.

5. There were some instances (OS & LR modules) of **specimens being provided in vials / containers that were not airtight** and, as a consequence, specimens were dry and in some case identification was impossible. Participants are reminded that specimens should be stored in suitable air-tight containers so that viability is maintained for the audit process. Participants should also ensure that OS & LR samples are transported to APEM in accordance with the H&S regulations. Participants should use rigid crates when submitting heavy sample residues to **prevent damage in transit**.

6. The maintenance of a comprehensive reference collection has numerous benefits for improving identification ability, maintaining consistency of identification between

surveys and access to growth series material. The LR exercise can be used as a means of verifying reference specimens. Laboratories are strongly recommended to **implement and expand in-house reference collections of biota**. The inclusion of growth series material is extremely useful for certain groups, *e.g.* molluscs. All surveys should have an associated reference collection to enable ease of cross-checking or adopting future taxonomic developments.

7. Participants submitting data for **laboratory reference exercises should add a note on habitat / location** of samples, to aid identification. A similar 'Habitat Notes' section to that distributed with the ring test exercises was distributed for completion in this year's exercise and will continue into the next exercise to support AQC identifications.

8. Laboratories participating in the ring test exercises should attempt to identify all specimens to species and **complete the 'confidence level' section of their ring test datasheets** to enable additional information to be gathered regarding the difficulty of ring test specimens.

9. The Own Sample module has shown **repeated taxonomic errors** for some laboratories over several years. Participating laboratories are encouraged to redress or resolve disagreements for taxonomic errors reported in their Own Samples even if their samples achieve an overall 'Pass' flag.

10. There are problems of **individuals and taxa missed at the sorting stage** of Own Sample analysis. This is an area that is often the major contributing factor in samples with 'Fail' flags or low Bray-Curtis similarity indices. When taxa and individuals are missed during the extraction of biota from the sediment, laboratories should determine why certain taxa have not been extracted. This could be due to the taxon not being recognised as countable, or due to problems with the effect of stains upon the specimens. There may also be a problem within certain taxonomic groups (*e.g.* crustaceans floating within samples or molluscs settled within the coarser sediment fractions). Additional training may be required and a review of existing extraction techniques and internal quality control measures may be beneficial. Remedial action should concentrate on the specific causes of the failure and should be targeted accordingly *e.g*. analyst or method related discrepancies.

11. It is apparent that some laboratories are **not utilizing the NMBAQC guidelines** for processing macrobenthic samples (Worsfold, Hall & O'Reilly (Ed.), 2010) issued with MB18 in Scheme Year 17 to improve the consistency of analysis*, i.e.* all analysts extracting and recording all biota. A detailed **taxonomic discrimination policy (TDP) needs to be developed** and added to the processing requirement protocol (PRP) to ensure that macrobenthic data from multiple analysts are as consistent and inter-comparable as possible. The Own Sample pass / fail criteria will be reviewed to ensure that they are fit for purpose and uphold data consistency between the Scheme participants.

12. Since the beginning of the scheme, continual improvement to the learning structure of the Scheme reports has been maintained. For the LR and OS modules, detailed results have been forwarded as **individual exercise reports** to each participating laboratory as soon after the exercise deadlines as practicable. The **Laboratory Reference Module Summary Reports introduced last year** show identification problems found in all LR submissions and should benefit all participants. In the RT module, after each RT exercise a bulletin was circulated, reviewing the literature used, detailing the accepted identification of the taxa circulated, and including images of relevant specimens. Participants are encouraged to review their exercise reports and **provide feedback concerning content and format** wherever appropriate.

13. The primary aim of the Benthic Invertebrate Component of the Scheme is to improve the quality of biological data via training and audit modules. An informal constructive reporting system exists to assist in the overall improvement of data quality. For example, laboratories struggling with particular taxonomic groups in their Own Samples often receive additional support, as well as receiving their returned OS material separated, according to the AQC identifications, for future reference. Eight of the 9 'failing' Own Samples in Scheme Year 2017 / 2018 (Year 24) have already been rectified via the recommended remedial action. **Participants are encouraged to provide feedback and request further information for any of the scheme exercises to improve the quality and consistency of their data.**

14. **Additional guidance for Own Sample 'next steps' following audit results** has been created to ensure that all participants and other stakeholders are aware of the route to quality assured data (*Own Sample Interim Report Review and Remedial Action Processes*).

15. There remain some misconceptions about the nature of the Scheme and the services it provides. It is not an accreditation scheme but provides quality assurance for the UK's CSEMP/WFD programme. In addition, the Scheme can provide **audits of samples** for any marine biological programme or development. It also provides **project-level audits** by applying the OS and LR protocols to examine project data. These services require more extensive communication (Scheme website, information note etc.) to notify all potential users and maintain consistent quality assurance for European marine data. A best practice guidance protocol for NMBAQC project-level audits needs to be produced and published on the scheme website. Meanwhile, it should be understood that a project level audit includes a review of data and check of reference collection specimens for the whole project, as well as for selected samples. Audits of samples from a project without more extensive reviews of data and other material do not constitute quality control of the whole project through the Scheme.

16. Despite protocol documents being produced for a recent Scheme year (Year 21, 2015-2016), misconceptions still exist regarding the purpose and methods for some of the Scheme's modules. **Protocol documents for all modules were reviewed and re-issued ahead of the exercises for this scheme year (Ring Test Protocol, Laboratory Reference Protocol, Own Sample Exercise Protocol).**

17. APEM Ltd. strives to ensure smooth running and **transparency of the Scheme** at all times. APEM Ltd. log and make available all correspondence to the Benthic Invertebrate Contract Manager (Myles O'Reilly, SEPA). Participants can be assured that their anonymity will be protected if this correspondence is required to be shared with the Committee.

4. References

CEFAS (ed.), 2012. *Clean Seas Environment Monitoring Programme.* *Green Book.* July 2012.

Hall, D.J. & Worsfold, T.M., 2001. *National Marine Biological Analytical Quality Control Scheme.* *Own Sample Format and Standards Review: Current Problems and Proposed Solutions.* Report to the NMBAQC Committee. April 2001.

Hall, D.J., 2010. *National Marine Biological Analytical Quality Control Scheme.* *Description of Scheme Standards for the Benthic Invertebrate Component from Scheme Year 8 (2001/02).* Report to the NMBAQC Scheme participants. February 2010.

Hall, D.J., 2016. *NE Atlantic Marine Biological Analytical Quality Control Scheme.* *Benthic Invertebrate component – Own Sample Interim Report Review and Remedial Action Processes.* Report to the NMBAQC Scheme committee and participants. 5pp, June 2016.

Hall, D.J., 2018. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Own Sample Module Summary Report OS65, 66 & 67.* Report to the NMBAQC Scheme participants. 14pp, July 2018.

Hall, D.J. & Worsfold, T.M., 2017. *Benthic Invertebrate component – Laboratory Reference Protocol.* Report to the NMBAQC Scheme participants. 5pp, August 2017.

Radashevsky, V.I., 2017. Identification keys and comments on the taxonomy of spionid polychaetes (Annelida: Spionidae) from the continental shelf of northern Europe. (unpublished draft circulated to ring test participants, 3rd November 2017).

Worsfold, T.M., 2017. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Review of recording and identification policy differences in Benthic Invertebrate Component exercises (OS, LR, MB) for Scheme Operation 2014 - 2016 (Years 21, 22, 23).* Report to the NMBAQC Scheme committee and participants. 18pp, July 2017.

Worsfold, T.M. & Hall, D.J, 2017a. *Benthic Invertebrate component – Ring Test Protocol.* Report to the NMBAQC Scheme participants. 6pp, August 2017.

Worsfold, T.M. & Hall, D.J, 2017b. *Benthic Invertebrate component – Own Sample Exercise Protocol.* Report to the NMBAQC Scheme participants. 16pp, August 2017.

Worsfold, T.M & Hall, D.J, 2018. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Laboratory Reference Summary Report: LR22.* Report to the NMBAQC Scheme participants. 9pp, March, 2018.

Worsfold, T.M, Hall, D.J. & Pears, S., 2017. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Ring Test Bulletin: RTB#53.* Report to the NMBAQC Scheme participants. APEM Report NMBAQC RTB#53. 41pp, Nov, 2017.

Worsfold, T.M, Hall, D.J. & Pears, S., 2018. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Ring Test Bulletin: RTB#54.* Report to the NMBAQC Scheme participants. APEM Report NMBAQC RTB#54. 32pp, Mar, 2018.

Worsfold, T.M., Hall, D.J. & O'Reilly, M. (Ed.), 2010. *Guidelines for processing marine macrobenthic invertebrate samples: a Processing Requirements Protocol: Version 1.0, June 2010.* Report to the NMBAQC Committee. June 2010.

Worsfold, T.M, Hall, D.J. & O'Reilly, M., 2018. Bibliography of taxonomic literature for marine and brackish water Fauna and Flora of the North East Atlantic. NMBAQC Scheme, 198 pp., February 2018.