NMBAQC

NE Atlantic Marine Biological Analytical Quality Control Scheme

Benthic Invertebrate Component Annual Report
Scheme Operation 2016 / 2017 (Year 23)

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

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

Contact: nmbaqc@apemltd.co.uk

APEM Ltd.
Date of Issue: August 2017

APEM

BENTHIC INVERTEBRATE COMPONENT ANNUAL REPORT FROM APEM Ltd

SCHEME OPERATION – 2016 / 17 (Year 23)

Linked Documents (hyperlinked in this report):

Ring Test Bulletin – RTB#51
Ring Test Bulletin – RTB#52
Laboratory Reference Module Summary Report – LR21
Own Sample Module Summary Report – OS62, 63 & 64
Description of the Scheme Standards for the Benthic Invertebrate Component
Guidelines for Processing Marine Macrobenthic Invertebrate Samples
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 2016 / 2017 (year 23) followed the format of year 2015 / 2016. 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-three laboratories participated in the Benthic Invertebrate Component of the NMBAQC Scheme in 2016 / 2017 (year 23). Sixteen of the participants were UK Competent Monitoring Authorities (CMAs) and twenty-seven were private consultancies, one of which was a consortium of sole traders. Thirteen of the UK CMA participants were responsible for the Clean Seas Environment Monitoring Programme (CSEMP) or Water Framework Directive (WFD) sample analysis. Seven of the participants were non-UK laboratories. 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 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 – OS62, 63 and 64 (2016/2017 (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 2016 / 2017 (year 23) of the North East Atlantic Marine Biological Analytical Quality Control (NMBAQC) Scheme.

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

- Own Sample module (OS) - re-analysis by APEM Ltd. of three samples supplied by each of the 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 twenty-five specimens supplied by each of the participating laboratories.

The analytical procedures of the various modules were the same as for 2015 / 16 (year 22) 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 (RT51 and RT52). The second (RT52) was targeted on bivalves.

For RT51, the average numbers of differences per participating laboratory (for a total of 20 laboratories with 22 submissions) were 4 generic differences and 8.4 specific differences. Eight taxa (three annelids, three molluscs, one crustacean and one cnidarian) were responsible for almost two thirds (65%) of the specific differences.

For RT52, the average numbers of differences per participating laboratory (for a total of 21 participants) were 4.0 generic differences and 5.9 specific differences. Four taxa

(*Scrobicularia plana*, *Cerastoderma edule* – 2 circulations at different sizes – and *Nucula nucleus*), all circulated as small sizes, were responsible for almost half (46%) of the specific differences.

Laboratory Reference (LR): Ten laboratories signed up for the LR21 module and six laboratories submitted specimens for confirmation. Most misidentifications were for Annelida (56%), followed by Mollusca (30%) and Crustacea (14%); many belonged to genera which are either speciose, or for which the taxonomy has yet to be finalized.

The revised protocols of Scheme Year 10 for 'blind' **Own Sample (OS)** audits were continued in this Scheme year. 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 Description of the Scheme Standards for the Benthic Invertebrate Component). In OS62-64, extraction efficiency (of individuals) was better than 90% in 88% of the comparisons and better than 95% in 76% of all comparisons. 100% of countable taxa were extracted from the sample residues in 45% of samples. The Bray-Curtis similarity index ranged from 43% to 100% with an average figure of 95.5%. The Bray-Curtis similarity index was greater than 95% in 74% of comparisons; in 86% of cases the value of the index was greater than 90% and, therefore, achieved 'Pass' flags. Nineteen samples (23%) achieved 'Pass-Excellent' flags with Bray-Curtis similarity scores of 100%.

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 protocols for LR21 (Milner & Hall, 2016a), RT51 (Milner & Hall, 2016b) and OS62-64 (Milner & Hall, 2016c).

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 June 2016, each participant was given a confidential, randomly assigned 2016 / 2017 (Scheme year 23) 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 2016 / 2017 (Year 23) was recorded as BI_2301. 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.

Two sets of 25 benthic invertebrate specimens were distributed in 2016 / 17. The first circulation (RT51) was a general invertebrate ring test and included 12 (48%) annelids, 5 (20%) molluscs, 4 (16%) crustaceans, 1 (4%) echinoderms and 3 (12%) taxa belonging to other phyla. The second circulation (RT52) was targeted at bivalves. This test included 25 (100%) bivalves. 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.

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. Two laboratories chose to utilise this option in RT51, none for RT52. The protocols followed for the two circulations, in particular the method of scoring results, were the same as for previous circulations. Approximately eight weeks were allowed for the analysis of both RT exercises (RT51 and RT52).

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, 23 laboratories subscribed to RT51 and 24 laboratories subscribed to RT52. For RT51, 20 laboratories returned data (22 individual data sets). For RT52, 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 the Ring Test Bulletins (RTB) 51 and 52 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 Scoring of RT Results

The laboratory's score was increased by one for 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 RTB51 and RTB52). Separate scores 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 2016 / 2017 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 51 (Type: General)

The results discussed below are given in Table 1 of RTB51, 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 RTB51).

Twelve of the 25 specimens circulated were annelids, five were molluscs, four were crustaceans, three were other taxa and one was an echinoderm. The agreement at generic level was generally good; 89 differences (16% of all genus identifications) were recorded in the 22 data sets received from 20 participating laboratories. There was less agreement at species level, with 184 differences recorded (33% of all species identifications).

Eight of the specimens circulated were responsible for almost two-thirds (65%) of participants' species level identification differences. These were the annelids *Melinna palmata, Terebellides shetlandica* and *Paramphitrite birulai*; the molluscs *Pharus legumen, Nucula nucleus* and *Vitreolina antiflexa*; the crustacean *Tanaissus danica*; and the cnidarian *Nematostella vectensis*.

Three of the 25 specimens circulated (the annelid *Sternaspis scutata*, the echinoderm *Echinocardium cordatum* and the pycnogonid *Nymphon brevirostre*) were correctly identified by all participants.

Further details and analysis of results can be found in the Ring Test Bulletin RTB51 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 52 (Type: Targeted on Bivalvia)

The results discussed below are given in Table 1 of RTB52 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 RTB52).

All 25 of the specimens circulated were molluscs. The agreement at genus level was moderate; 83 differences (16% of all genus identifications) were recorded in the 21 data sets received from 21 participating laboratories, of which 20 differences were for one species (1mm *Scrobicularia plana*). There was less agreement at species level, with 124 differences recorded (24% of all species identifications).

Four of the specimens circulated were responsible for almost half (46%) of participants' species level identification differences. These were 1mm *Scrobicularia plana,* 1mm and 2-3mm *Cerastoderma edule* and 2mm *Nucula nucleus*.

Two of the twenty-five specimens circulated (*Tellimya ferruginosa* and *Barnea parva*) were correctly identified by all participants.

Further details and analysis of results can be found in the Ring Test Bulletin RTB52 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 RTB51 and RTB52 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 for both exercises are shown below:

Major taxon	Species circulation	Generic differences		Specific differences	
Annelida	12	26	15.1%	80	26.0%
Crustacea	4	16	9.3%	26	8.4%
Mollusca	30	112	65.1%	184	59.7%
Echinodermata	1	0	0%	0	0%
Others	3	18	10.5%	18	5.8%
Total	**50**	**172**	**100%**	**308**	**100%**

Most of the specific differences in RT51 were for annelid species (RT52 was entirely molluscs).

2.2.3 Discussion

The results were in general comparable with those from previous exercises, with an average of 4% generic and 8.4% specific differences across the participating laboratories in RT51 and 4% generic and 5.9% specific differences across the participants in RT52.

In RT51, several of the most significant differences (e.g. for *Tanaissus danica*, *Terebellides shetlandica* and *Vitreolina antiflexa*) were the result of lack of knowledge of literature and recent taxonomic work (citations were provided in the bulletin). Others (e.g. for *Paramphitrite birulai*, *Ecrobia ventrosa* and *Nematostella vectensis*) were due to inherent difficulties in recognition of identification features for the species. The high error rate for *N. vectensis* is notable as this is a protected and non-native species. The majority of RT52 (and some RT51) differences were due to the inadequacy of identification keys for small bivalves (circulated sizes were included in RTB52). Growth series are often required for these and some were provided in the bulletins but it is ultimately the responsibility of participants to maintain their own reference collections.

The RT component is considered 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. Two multiple datasets were submitted for the Ring Test exercises RT51. All participating laboratories have been made aware of the problems encountered during these ring tests via Ring Test Bulletins RTB51 and 52, 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 / 97). This module can help assess the ability of participating laboratories to identify material from their own samples. Laboratories are also able to use this exercise to obtain second-opinion identifications of difficult or problematic taxa of which they are unsure. This was the twenty-first Laboratory Reference exercise (LR21). 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

Ten laboratories signed up for this exercise (LR21) but only six submitted specimens for examination. Detailed results have been separately reported to each participating laboratory. Taxonomic edits were made for submitted polychaetes (24; 56%), molluscs (13; 30%) and crustaceans (6; 14%). In addition, changes were made to taxonomic resolution,

recording notation and spelling for many specimens. A new report summarising the results from this module is presented in the Laboratory Reference Module Summary Report – LR21.

2.3.3 Discussion

As with all training exercises, detailed inter-laboratory comparisons are of limited value. Some of the differences resulted from policy changes and recent literature and workshop outcomes (e.g. *Syllis parapari*, *Dipolydora saintjosephi*, *Owenia borealis*, *Vitreolina antiflexa*). The submitted specimens also included several species that cannot yet be named and may be undescribed (e.g. *Sphaerosyllis cf. taylori*, *Scolelepis squamata* type 1, *Melinna* sp., Cochliopidae species A). The taxonomic resolution and recording policy differences were used to revise and standardize the notes made on such differences in future 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 and Worsfold, 2001), several changes to sample selection and scoring were implemented in Scheme Year 8 (2001 / 02). All participants must meet these Own Sample requirements. Own Sample participants must supply their previous year's CSEMP / WFD data matrices, where relevant, for Own Sample selection, *i.e.* 2014 / 2015 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 (an increase from the previous twelve in Year 22) 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 all animals 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, 84 selected Own Samples were received from 28 (of the 33 subscribing) laboratories, together with descriptions of their origin and the collection and analysis procedures employed. Samples were identified as OS62, OS63 and OS64 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 0 to 105, with the number of countable individuals from 3 to 2466. Of the 84 submitted Own Samples, 7 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 – OS62, 63 & 64.

Two laboratories (BI_2330 and BI_2334) transferred their Own Sample allocations (total six samples) to a project-level audit, which was reported outside of the Own Sample module. The six samples were audited following the NMBAQC OS protocols and reported along with additional audit samples to achieve a 5% audit of a sampling programme. All six samples passed the NMBAQC standards and data for the project were quality assured. It was not possible to include these samples in the standard OS module and reporting mechanism due to the specific demands of the project (reporting timescale).

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 analysis. In 45 samples out of the total 84, 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 39 cases, the difference was on average 2.9 with a maximum of 11 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 35.9%. The average difference between the samples with recorded differences was 6.6% (and 3.6% across all 84 samples), with 16 samples exceeding this average.

38 of the 84 samples reported showed 100% extraction of individuals from the residue (column 16), and in 46 samples between 1 and 206 individuals had been missed during processing. In just 14 samples only individuals attributed to taxa already recorded in the sample were found. In 39 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 1 to 11 with an average of 1.4 new taxa per sample. The poorest extraction records were a total of 11 missed taxa and 33 individuals, 10 missed taxa and 206 individuals, 9 missed taxa and 63 individuals, and 5 missed taxa and 86 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 84 samples) of missed individuals found

upon re-sorting the residue was approximately 8, and the average number of missed taxa was 1.4.

2.4.2.3 Uniformity of Identification

Taxonomic differences (columns 10 and 11) between the auditor and participating laboratories' results were found in 36 (43%) of the 84 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 less than one taxonomic error per laboratory was recorded; in the worst instance, 5 identification errors occurred. A large variety of samples (and biota) was received. Polychaetes accounted for 43%, Mollusca for 41%, Crustacea for 7%, Echinodermata for 7%, and others for 3% 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 2015 / 2016 (Year 22). The Bray-Curtis similarity index figures (Table 1, column 23) ranged from 43% to 100%, with an average figure of 95%. Twelve samples from eight laboratories achieved a similarity figure of less than 90%. Nineteen samples produced a similarity figure of 100%; these were submitted by fourteen different laboratories (BI_2304, BI_2309, BI_2311, BI_2315, BI_2317, BI_2328, BI_2333, BI_2335, BI_2336, BI_2337, BI_2338, BI_2340, BI_2342 and BI_2343). The best overall result was achieved by laboratories BI_2336 and BI_2337 with 100% similarity across all three Own Samples. The lowest overall result was achieved by BI_2207 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; 69 samples had not been supplied with species blotted wet weight biomass data. Consequently, only 15 of the 84 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 the participating laboratories varied greatly compared to those obtained by APEM Ltd. Differences in the recorded biomass ranged from -17% to +38%. The reason for these large differences is presumably 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 some laboratories 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 going 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 84 samples submitted for the Own Sample module, including the seven processed by the Scheme's external auditor. Of the 84 samples, 72 (86%) exceeded the 90% Bray-Curtis Pass mark and 62 (74%) of the samples exceeded 95% BCSI. Since the beginning of this module in Year 02 of the Scheme, 79% 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, 1493 admissible samples have been received (OS01-64). Of these, 257 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 been noticed in year 20 and 21 compared with the previous four years, there has been a marked improvement in 2015 / 2016 and this has continued in 2016 / 17. Some participating laboratories should be able to further improve their results by reviewing their extraction methods and their use of taxonomic literature and identification keys.

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 / 97). These standards were the subject of a review in 2001 (Hall and 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 this 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 of the 32 participating laboratories achieved a Bray Curtis of >90% ('pass' flag) for all three of their Own Samples. Overall, 73% of the comparisons were considered to have passed the enumeration of taxa standard, 83% exceeded the enumeration of individuals standard and 86% 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); 9 samples (11%) are flagged as 'Fail - Bad', 3 (4%) as 'Fail – Poor', 10 (12%) as 'Pass - Acceptable', 43 (51%) as 'Pass - Good'

and 19 (23%) 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 80% 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 – OS62, 63 & 64). The table shows the number of laboratories assigned 'Pass' and 'Fail' flags for the OS exercises over the past twenty two years based upon the current NMBAQC Scheme standards (see Description of the Scheme Standards for the Benthic Invertebrate Component). This year's 84 Own Samples resulted in a pass rate of 86% (see Table 5 in the Own Sample Module Summary Report), which is a 2% increase from the previous scheme year. Historically the highest pass rate achieved was 100% in exercise OS01 (1995 / 96; Year 2) that involved just fourteen samples; the lowest pass rate was 67% recorded in 2000 / 01 (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 replicates to improve the flagged data. For a CSEMP/WFD sample, the associated samples would normally be the samples (5-10 in number) collected from the same 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. The participating laboratories are expected to initiate remedial action according to the advice of APEM Ltd. 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 2016 / 2017 (Year 23). Eight laboratories were responsible for twelve '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_2304	**OS62**	Reprocess residues for associated samples; effectiveness to be assessed via a residue only audit	Remedial action completed 10/07/17
	OS63	Reprocess residues for associated samples; effectiveness to be assessed via a residue only audit	Remedial action completed 24/04/17
BI_2307	**OS62**	Review taxonomic errors, particularly Phaxus / Ensis individuals	Remedial action completed 27/04/17
BI_2317	**OS64**	Reprocess residues for associated samples; effectiveness to be assessed via a residue only audit	Remedial action completed 01/05/17
BI_2320	**OS62**	Review taxonomic error and reprocess associated residues; effectiveness to be assessed via a residue only audit	Remedial action completed 07/06/17
	OS63	Review taxonomic error and biota extraction methods	Remedial action completed 25/05/17
	OS64	Review taxonomic error and biota extraction methods	Remedial action completed 25/05/17
BI_2331	**OS63**	Review processing procedures for taxon extraction and reprocess associated residues for individuals	Remedial action not complete
	OS64	Reprocess associated sample residues	Remedial action not complete
BI_2340	**OS62**	Review extraction techniques for samples with few individuals	Remedial action completed 15/03/17
BI_2343	**OS64**	Reprocess taxonomic errors in associated samples	Remedial action completed 17/05/17
BI_2346	**OS62**	Review extraction procedures; Reprocess associated samples for taxonomic errors	Remedial action not complete

3. Conclusion 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 only the minimum 20 samples yet other labs offering up to 205 samples 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. Project / survey data have been supplied more than once for audit sample selection via the Own Sample module, from both CMA and processing laboratory subscriptions. Where data belong to CMAs the submitting participant should declare this so that audit results can be shared accordingly and CMA data audit can be tracked and co-ordinated. Revised data request and sample submission forms will be created for the 2017 / 18 OS module to capture **data / sample ownership**.

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 fauna**. The inclusion of growth series material is extremely useful for certain faunal groups, *e.g.* identifying certain 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 sample 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. Participants submitting data for the ring test exercises should attempt to identify the specimen / 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 still some 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 fauna 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 sample 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. It has been noted that some laboratories are producing data with an atypically high number of over-cautious identifications and multiple taxa recorded for a single species. For example, records of the bivalve *Thyasira flexuosa* may have been divided between false taxa (*Thyasira flexuosa*, *Thyasira flexuosa* juvenile, *Thyasira*, *Thyasira* juvenile, Thyasiridae, Thyasiridae juvenile, Bivalvia etc.), This will lead to exaggeratedly high numbers of total taxa and also to data comparison issues for spatial and temporal studies. For example, where additional taxa are recorded due to processing damage (e.g. loss of scales in scaleworms or amphipods breaking in half), data may appear to show artificial 'changes' with less or more damaging sample

processing. 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. A **newly introduced Laboratory Reference Module Summary Report** showing identification issues raised in all LR submissions should benefit all participants. In the RT module, after each RT exercise a bulletin was circulated, reviewing the literature used, detailing the correct 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 faunal groups in their Own Samples often receive additional support, as well as receiving their returned OS faunal material separated, according to the AQC identifications, for future reference. Nine of the 12 'failing' Own Samples in Scheme Year 2016 / 2017 (Year 23) 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*). The Scheme provides quality assurance for the UK's CSEMP/WFD programme. However the Scheme can provide project-level audits for any marine biological programme or development by applying the OS and LR protocols to examine project data. This service requires more extensive communication (Scheme website, information note etc.) to notify all potential users and maintain consistent quality assurance for European marine data.

15. Despite protocol documents being produced for this Scheme year, a number of misconceptions still exist regarding the purpose and methods for some of the Scheme's modules. **Protocol documents for all modules will be reviewed and re-issued ahead of 2017 / 18 circulations.**

16. 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., 2017. NE Atlantic Marine Biological Analytical Quality Control Scheme. Own Sample Module Summary Report OS62, 63 & 64. Report to the NMBAQC Scheme participants. 16pp, May 2017.

Milner, C. & Hall, D., 2016. Benthic Invertebrate component – Laboratory Reference Protocol. Report to the NMBAQC Scheme participants. 5pp, June 2016.

Milner, C. & Hall, D., 2016. *Benthic Invertebrate component – Ring Test Protocol.* Report to the NMBAQC Scheme participants. 5pp, June 2016.

Milner, C. & Hall, D., 2016. *Benthic Invertebrate component – Own Sample Protocol.* Report to the NMBAQC Scheme participants. 5pp, June 2016.

Milner, C., Worsfold, T., Hall, D. & Pears, S., 2016. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Ring Test Bulletin: RTB#51.* Report to the NMBAQC Scheme participants. APEM Report NMBAQC RTB#51, 39pp, Oct, 2016.

Worsfold, T., 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. & Hall, D., 2017. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Laboratory Reference Summary Report: LR#21.* Report to the NMBAQC Scheme participants. 12pp, May, 2017.

Worsfold, T., Hall, D. & Pears, S., 2017. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Ring Test Bulletin: RTB#52.* Report to the NMBAQC Scheme participants. APEM Report NMBAQC RTB#52. 33pp, Mar, 2017.

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.