NMBAQC

NE Atlantic Marine Biological Analytical Quality Control Scheme

Benthic Invertebrate Component Annual Report
Scheme Operation 2015/2016 (Year 22)

Author: Carol Milner, NMBAQCS Benthic Invertebrate Administrator

Reviewer: David Hall, NMBAQCS Project Manager

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

Contact: nmbaqc@apemltd.co.uk

APEM Ltd.

Date of Issue: July 2016

BENTHIC INVERTEBRATE COMPONENT ANNUAL REPORT FROM APEM Ltd

SCHEME OPERATION – 2015/16 (Year 22)

Linked Documents (hyperlinked in this report):

Ring Test Bulletin – RTB#49
Ring Test Bulletin – RTB#50
Own Sample Module Summary Report – OS59, 60 & 61
Description of the Scheme Standards for the Benthic Invertebrate Component
Guidelines for Processing Marine Macrobenthic Invertebrate Samples

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 macrobenthic invertebrate samples;
- The identification of macrofauna;
- The determination of physical parameters of sediments.

Scheme year 2015/2016 (year 22) followed the format of year 2014/15 (with the exception that the Macrobenthic Exercise was dropped through lack of participant interest). 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-one laboratories participated in the benthic invertebrate component of the NMBAQC Scheme in 2015/2016 (year 22). Sixteen participants were Competent Monitoring Authorities (CMAs) and twenty-five were private consultancies. One of the participants was a consortium of sole traders. Twelve of the CMA participants were responsible for the Clean Seas Environment Monitoring Programme (CSEMP) or Water Framework Directive (WFD) sample analysis. Laboratory Codes were assigned in a single series for all laboratories participating in the benthic invertebrate components of the NMBAQC Scheme. 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. Competent monitoring authorities (CMAs) completing benthic invertebrate biological analyses for monitoring programmes (including in assessment of MPAs (Marine Protected Areas), as evidence under MSFD (Marine Strategy Framework Directive) and WFD (Water Framework Directive) as well as the CSEMP (Clean Seas Environmental Monitoring Programme) must participate in this component of the Scheme. 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 OS 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 – OS59, 60 and 61 (2015/2016 (Year 22) 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 Invertebrates Component for year 2015/2016 (year 22) 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 own samples supplied by each of the participating laboratories;
- Invertebrate Ring Test module (RT) - identification of two sets of twenty-five invertebrate specimens; and
- LR, 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 2014/15 (year 21) of the Scheme. The results for each of the Scheme exercises are presented and discussed. Comments are provided on the performance for each of the participating laboratories in each of the exercises.

Two **Ring Tests (RT)** of 25 specimens were distributed (RT49 and RT50). Both sets contained 25 invertebrate specimens, the second (RT50) was targeted at amphipods and similar taxa.

For RT49 each participating laboratory (a total of 21 participants) recorded on average 3.5 generic differences and 5.7 specific differences. Seven taxa (two annelids, two sipunculans and three molluscs) were responsible for almost two thirds (61%) of the specific differences.

For RT50 each participating laboratory (a total of 22 participants) recorded on average 2.2 generic differences and 5.5 specific differences. Six taxa (five gammaridean amphipods and one caprellid) were responsible for almost half (49.6%) of the specific differences.

Laboratory Reference (LR): Nine laboratories signed up for the LR20 module and seven laboratories submitted their specimens for confirmation. Two of the laboratories submitted less than the allowed 25 specimens. Most misidentifications were found to be for Annelida, Crustacea and Mollusca belonging to genera which are either speciose, or for which the taxonomy has yet to be finalized. The majority of taxonomic errors could be attributed to the submitted polychaetes (50%) and crustacea (27%).

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 OS59-61, extraction efficiency (of individuals) was better than 90% in 91% of the comparisons and better than 95% in 77% of all comparisons. 100% of countable taxa were extracted from the sample residues in 55% of samples. The Bray-Curtis similarity index ranged from 0% to 100% with an average figure of 93%. The Bray-Curtis similarity index was greater than 95% in 72% of comparisons; in 84% of cases the value of the index was greater than 90% and, therefore, achieved 'Pass' flags. Fifteen samples (16%) 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 to be 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 and specific details can be found in the Scheme's annual reports for 1994/95 and 1995/96 (Unicomarine 1995 & 1996).

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 have been 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 May 2015 each participant was given a confidential, randomly assigned 2015/2016 (Scheme year 22) LabCode. Codes are prefixed with the component initials, for example, BI for benthic invertebrates, the Scheme Year and a unique number (between 01 and 44) *e.g.* Laboratory number one in Scheme Year 2015/2016 (Year 22) was recorded as BI_2201. 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 inter-laboratory variation in the participants' ability to identify fauna and attempts to determine if errors were the result of inadequate keys, lack of reference material or the incorrect use of satisfactory keys.

Two sets of 25 benthic invertebrate specimens were distributed in 2015/16. The first circulation (RT49) was a general invertebrate ring test and included 9 (36%) annelids, 6 (24%) molluscs, 4 (16%) crustaceans, 2 (8%) echinoderms and 4 (16%) taxa belonging to other phyla. The second circulation (RT50) was targeted at amphipod and similar taxa. This test included 23 (92%) amphipods and 2 (9%) caprellids. Details of substratum, salinity, depth and geographical location 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 UK. Specimens were also donated by Scheme participants and other organizations. Every attempt was made to provide animals in good condition and of similar size for each laboratory. Each specimen was uniquely identifiable by means of a coded label and all material has been retained for subsequent checking. Where relevant, every effort was made to ensure all specimens of a given species were of the same sex. For both ring tests the specimens were taken from replicate trawls, grabs or cores within a single survey and in most cases they were replicates from a single sampling station.

2.2.1.2 Analysis Required

The participating laboratories were required to identify each of the RT specimens to species level. If a laboratory had not routinely identified the specimen to species level, they were asked to state this in the 'confidence level' field. Laboratories could also add brief notes and information detailing the keys, or other 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 data entries (*i.e.* different sets of results from different

analysts) for each exercise to enhance the training value of this module. One laboratory in each RT exercise chose to utilise this option in 2015/2016. 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 (RT49 and RT50).

2.2.2 Results

2.2.2.1 General Comments

A number of laboratories use the ring tests for training purposes and have selected them preferentially over other modules. The results are not used to assign 'Pass' or 'Fail' flags. In total 21 laboratories subscribed to RT49 and a total of 22 laboratories subscribed to RT50. For RT49, 19 laboratories returned data (20 individual data sets). For RT50, 20 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 number of differences. Where identification 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) 49 and 50 show identifications made by each of the participating laboratories for the twenty-five specimens, arranged by specimen and by laboratory respectively. For clarity, the name is given only in those instances where the generic or specific 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 was presented in brackets: "[name]". A dash, "-", in the tables indicates that the name of the genus (and / or species) given by the laboratory was considered to be 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 the tables (Tables 1 and 2 in RTB49 and RTB50). Two 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 particular difficulties lie within specific common taxa. Results for Scheme year 2015/2016 were presented in the Ring Test Bulletins (RTB) along with the reasons for each individual identification discrepancy. These bulletins contained images of the test material and the alternative, incorrectly recorded taxa, where these taxa where available. Participating laboratories were advised to retain their 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 potential future circulation.

2.2.2.3.1 Ring Test 49 (Type: General)

The results discussed below are given in Table 1 of RTB49 which displays the data arranged by species to enable quick reference to the range of answers received and in Table 2, which presents the results arranged by laboratory (see Ring Test Bulletin – RTB49).

Nine of the 25 specimens circulated were annelids, six were molluscs, four were crustaceans, four were other taxa and two were echinoderms. The agreement at generic level was generally good; 69 differences, or 14% of all genus identifications, were recorded in the 20 data sets received from 21 participating laboratories. There was less agreement at species level, with 114 differences recorded, equal to 23% of all species identifications.

Seven of the specimens circulated were incorrectly identified at species level by almost two-thirds (61%) of the participants. These were the annelids *Eulalia ornata* and *Aricidea wassi*; the molluscs *Pulsellum affine, Mya arenaria* and *Thracia phaesolina*; and the sipunculans *Golfingia elongata* and *Nephasoma minutum*.

Four of the 25 specimens circulated, the annelids *Poecilochaetus serpens*, *Protodorvillea kefersteini* and *Spiophanes bombyx* and the crustacean *Eurydice pulchra* were correctly identified by all participants.

Further details and analysis of results can be found in the Ring Test Bulletin RTB49 which was circulated to each laboratory that supplied results for this exercise and was also posted on the Scheme's website (www.nmbaqcs.org/RTB49).

2.2.2.3.2 Ring Test 50 (Type: Targeted on Amphipoda)

The results discussed below are given in Table 1 of RTB50 which displays the data arranged by species to enable quick reference to the range of answers received and in Table 2 which presents the results arranged by laboratory (see Ring Test Bulletin – RTB50).

All 25 of the specimens circulated were crustaceans. The agreement at genus level was good; 47 differences, or 9% of all genus identifications, were recorded in the 21 data sets received from 20 participating laboratories. There was less agreement at species level, with 115 differences recorded, equal to 22% of all species identifications.

Six of the specimens circulated were incorrectly identified at species level by half of the participants. These specimens were *Ampelisca diadema, Socarnes erythrophthalmus, Caprella mutica*, *Dexamine thea* (x2 specimens) and *Gammarus tigrinus*.

None of the twenty-five specimens circulated were correctly identified by all participants.

Further details and analysis of results can be found in the Ring Test Bulletin RTB50 which was circulated to each laboratory that supplied results for this exercise and was also posted on the Scheme's website (www.nmbaqcs.org/RTB50).

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 graph related to Table 2 in RTB49 and RTB50 respectively. The laboratories are ordered by increasing number of differences at species level. The division of laboratories into three bands (Low, Medium and High) on the basis of the number of differences at the level of species is also shown.

2.2.2.5 Differences by Taxonomic Group

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

Taxon	No. species	Generic differences		Specific differences	
Polychaeta	9	9	7.8%	31	13.5%
Crustacea	29	53	45.7%	124	54.1%
Mollusca	6	30	25.9%	45	19.7%
Echinodermata	2	6	5.2%	7	3.1%
Others	4	18	15.5%	22	9.6%
Total	50	116	100%	229	100%

Most of the specific differences in the two ring test exercises can be attributed to crustacean species followed by molluscs.

2.2.3 Discussion

The results were in general comparable with those from previous exercises, with an average of 3.5 generic and 6 specific differences across the participating laboratories in RT49 and 2 generic and 5.5 specific differences across the participants in RT50. The RT component is considered a valuable training tool and can be an indicator of problem groups. It can highlight possible areas for further 'targeted' ring test exercises or for inclusion at taxonomic workshops. The ability of participants to submit multiple data entries and the inclusion of images in the Ring Test Bulletins have enhanced the training value of this component. No participants chose to submit multiple datasets for the Ring Test exercises in this Scheme year. All participating laboratories have been made aware of the variety of problems encountered during these ring tests via Ring Test Bulletins RTB49 and 50, which also include a list of useful literature which they can then source.

The best results were obtained by BI_2214, BI_2202, BI_2204, BI_2207 and BI_2211 for RT49 with between zero and one differences at genus level and zero to one difference at species level. In RT50 the best results were for laboratories BI_2202, BI_2204, BI_2209, BI_2207, BI_2217 and BI_2221 with zero differences at generic level and between one and three differences at specific level.

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

2.3.1 Description

The Laboratory Reference Module is a training module which encourages laboratories to build extensive, verified reference collections to improve identification consistency. The value of reference material in assisting the process of 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 area, or material with which they are familiar. Laboratories are also able to use this exercise to verify identifications of difficult or problematic taxa about which they are unsure. Specimens were, wherever possible, representatives from CSEMP/WFD reference collections. This was the twentieth Laboratory Reference exercise (LR20). 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 4 weeks. All specimens were re-identified and the identification made by APEM Ltd. compared with that made by the participating laboratories. All specimens were returned to the laboratories after analysis.

2.3.2 Results

In total, nine laboratories signed up for this exercise (LR20) but only seven laboratories submitted specimens for examinations. Detailed results have been separately reported to each of the participating laboratories. Misidentifications were usually found for polychaete, amphipod and gastropod mollusc species and belonging to genera which are either speciose or for which keys are inadequate. The majority of taxonomic errors could be attributed to the submitted polychaetes (63 %) and crustaceans (27 %).

2.3.3 Discussion

In view of the different species that were sent by laboratories for identification it is inappropriate to make detailed inter-laboratory comparisons. Most laboratories elected to obtain a 'second opinion' on more difficult species.

2.4 Own Sample (**OS**) Module

2.4.1 Description

The Own Sample Module examines laboratory 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 new Own Sample requirements. Own Sample participants must supply their previous year's CSEMP/WFD data matrices, where relevant, for Own Sample selection, *i.e.* 2013/2014 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. The selection was, in turn, 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 twelve 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 their normal procedures. A summary of these in-house sample processing procedures was to be provided, on a standard form, with each Own Sample. Samples requiring sub-sampling were to be avoided where possible. All procedures were to be 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 fauna; and
- Other fractions - *e.g*. material containing fauna which had been counted *in situ*.

Identification was to be to the normal taxonomic level employed by the laboratory (presumed to be usually species), except for CSEMP/WFD samples where the NMBAQC guidelines for macrobenthic sample analysis were to be followed (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.

One month was allowed for the submission of data; and a further eight weeks was allowed for the preparation and submission of the Own Samples selected for re-analysis. The sorted residue was re-examined and any countable material extracted. Identified fauna was checked for the accuracy of enumeration and identification and, in cases where biomass was provided by the participant, 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, 96 selected Own Samples were received from 32 laboratories, together with descriptions of their origin and the collection and analysis procedures employed. Samples were identified as OS59, OS60 and OS61 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 2 to 141, with the number of countable individuals from 3 to 3425. Of the 96 submitted Own Samples, 7 had to be audited externally by Fugro EMU Ltd. due to the initial processing being carried out by APEM Ltd. Interim reports have been submitted to the participating laboratories. A summary of results from this module is presented in the Own Sample Module Summary Report – OS59, 60 & 61.

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 identified and enumerated by the participating laboratory were included in the analysis, except in instances where the fauna had been damaged and rendered

unidentifiable and uncountable. In 47 samples out of the total 96, the number of taxa recorded by the participating laboratories was identical to that obtained by the auditing laboratory (columns 2 and 3). In the remaining 49 cases, the difference was on average 2.5 with a maximum of 15 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 100%. The average difference between the samples with recorded differences was 10.3% (and 5.9% across all 96 samples), with 9 samples exceeding this average.

41 of the 96 samples reported showed 100% extraction of individuals from the residue (column 16), and in 55 samples between 1 and 671 individuals had been missed during processing. In just 15 samples only individuals attributed to taxa already recorded in the sample were found. In 43 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 2.7 new taxa per sample. The poorest extraction records were a total of 11 missed taxa and 34 individuals, 10 missed taxa and 119 individuals and five missed taxa and 671 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 96 samples) of missed individuals found upon re-sorting the residue was approximately 33, and the average number of missed taxa was 2.7.

2.4.2.3 Uniformity of Identification

Taxonomic differences (columns 10 and 11) between the auditor and participating laboratories' results were found in 44 (46%) of the 96 own samples. A summary of mis-identified taxa is presented in Table 3 of the OS Summary Report. In the samples with taxonomic errors an average of 3 taxonomic errors per laboratory was recorded; in the worst instance 19 identification errors occurred. A large variety of samples (and fauna) was received. Polychaetes accounted for 47%, Mollusca for 23%, Crustacea for 14%, others for 11%, Oligochaeta for 3% and Echinodermata for 1% of the taxonomic errors, 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 2014/2015 (Year 21). The Bray-Curtis similarity index figures (Table 1, column 23) ranged from 0% to 100%, with an average figure of 93%. Fifteen samples from

ten laboratories achieved a similarity figure of less than 90%. Fifteen samples produced a similarity figure of 100%; these were submitted by twelve different laboratories (BI_2201, BI_2202, BI_2205, BI_2208, BI_2209, BI_2211, BI_2220, BI_2226, BI_2228, BI_2232, BI_2234 and BI_2241). The best overall result was achieved by BI_2241 with 100% similarity across all three Own Samples. The lowest overall result was achieved by BI_2203 with an average similarity index of less than 2.3% over all three samples. The latter samples were exceptionally bad and represent the worst ever audited in the scheme.

2.4.2.5 Biomass Determinations

It was not possible to make an accurate comparison of the biomass determination in all cases; 77 samples had not been supplied with species blotted wet weight biomass data. Consequently, only 19 of the 96 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 +20% to +40%. 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

The total numbers of samples for which the participating laboratories submitted data to APEM Ltd to chose audit Own Samples ranged from 6 and 8 (less than the requested minimum of 12) to 279. The average number of samples data for selection was 38. 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 or projects are going to be audited.

The average Bray-Curtis similarity index of 93% achieved for this Own Sample Module shows that the agreement between the participating laboratories and APEM Ltd. was generally acceptable, despite some exceptionally bad samples.

There were 96 samples submitted for the Own Sample Module, including the seven processed by the Scheme's external auditor. Of the 96 samples, 81 (84%) exceeded the 90% Bray-Curtis Pass mark and 69 (72%) of the samples exceeded 95% BCSI. Since the beginning of this module in Year 02 of the Scheme, results of Years 03, 04, 05, 06, 07, 08, 09, 10, 15, 16, 20, 2014/2015 (21) and now 2015/2016 (22) achieved 84% or less of the samples exceeding the 90% Bray-Curtis Pass mark (see Table 5 of the OS Summary Report).

Since the beginning of the Own Sample Module, 1409 admissible samples have been received (OS01-61). Of these, 245 samples (21%) have fallen below the 90% Pass mark. Overall, these results are acceptable 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. Some participating laboratories should be able to 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 key roles of the Benthic Invertebrate Component of the NMBAQC Scheme is to assess the reliability of data collected as part of the CSEMP or WFD monitoring programmes. 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 particular 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, only the Own Sample Module has been used in 'flagging' for the purposes of assessing data for the CSEMP/WFD programmes.

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 as more general indicators of laboratory performance, or as training exercises.

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 labs are requested to follow NMBAQC guidance, detailed comparisons of results between different labs are generally not applicable due to the diversity of samples analysed and some minor inter-lab variations in processing methodologies – especially in relation to identification. Development of more detailed taxonomic discrimination protocols may help resolve some of the latter discrepancies.

It can be seen from Table 1 (columns 5, 15 and 26) that 53% (17 of 32) of participating laboratories met or exceeded the required standard for three of the OS targets - the enumeration of taxa, enumeration of individuals and the Bray-Curtis comparison, for all three samples submitted as part of this exercise. Twenty-two laboratories achieved a Bray Curtis of >90% for all three of their Own Samples.

Overall, 88% of the comparisons were considered to have passed the enumeration of taxa standard, 85% exceeded the enumeration of individuals standard and 84% 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); 8 samples (8%) are flagged as 'Fail - Bad', 7 (7%) as 'Fail – Poor', 12 (13%) as 'Pass - Acceptable', 54 (56%) as 'Pass - Good' and 15 (16%) 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 only 79% 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 – OS59, 60 & 61). 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 96 Own Samples resulted in a pass rate of 84% (see Table 5 in the Own Sample Module Summary Report), the highest being 100% achieved in exercise OS01 that involved just fourteen samples; the lowest being 67% recorded in Year 7 from 45 samples.

2.4.4.3 Remedial Action

It is imperative that failing CSEMP/WFD samples, audited through the Own Sample Module, are addressed. Remedial action should be conducted upon the associated CSEMP/WFD replicates to improve the flagged data. For a CSEMP/WFD sample, the associated samples are the five sample replicates or the five dispersed samples in the same water body. For a WFD sample, the associated samples would normally be the samples (5-10 in number) collected from the same 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 indices 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 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 flagged with 'Fail' flags in Scheme Year 2015/2016 (Year 22). Ten labs and fifteen samples 'failed' (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_2203	OS59	Reprocess all associated residues; Ensure sieving is thorough; large proportion of residue <0.5mm	Remedial action not complete
	OS60	Reprocess all associated residues; Ensure sieving is thorough; most of the residue <0.5mm	Remedial action not complete
	OS61	Reprocess all associated residues; Ensure sieving is thorough; all residue <0.5mm	Remedial action not complete
BI_2205	OS59	Re-sort associated residues; Check taxonomic errors through batch and update data set; Submit revised data for random selection of another sample for full audit.	Remedial action not complete
BI_2213	OS60	Reprocess taxonomic errors in all associated samples; Review use of "dam.". Specimens described as "dam." were all identifiable. Review juvenile recording policy.	Remedial action not complete
	OS61	Review taxonomic errors in all associated samples; Review method of identifying live/dead molluscs.	Remedial action not complete
BI_2218	OS60	Re-sort associated residues	Remedial action in progress – to be evaluated
BI_2229	OS60	Reprocess associated residues and update dataset	Remedial action completed 5/5/16
BI_2232	OS60	Review taxonomic error for associated samples	Remedial action completed 22/4/16
BI_2234	OS59	Reprocess taxonomic errors and review extraction methods for all associated samples	Remedial action not complete
BI_2236	OS59	Reprocess taxonomic errors in associated samples	Remedial action completed 18/2/16

Lab Code	OS no.	Remedial action	Notes
BI_2239	OS59	Reprocess taxa and taxonomic errors and review counts throughout project; Review assignment of juveniles, damaged specimens and fragments.	Remedial action not complete
	OS60	Review estimation of taxa and reprocess taxonomic errors throughout project; Review assignment of juveniles, damaged specimens and fragments. Ampelisca (x78) should have been taken to species. *Dendrodoa grossularia* identifiable; not to be split into Ascidiacea juv.	Remedial action not complete
	OS61	Review estimation of taxa and reprocess taxonomic errors throughout project; Review assignment of juveniles, damaged specimens and fragments.	Remedial action not complete
BI_2240	OS61	Review taxonomic errors in all associated samples	Remedial action completed 19/1/16

3. Conclusion and Recommendations

A number of 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. Of the results submitted 20% of RT49, 9% of RT50, 57% of LR and 50% of all Own Samples were late. Late submission ranged from a day to seven months late (all of the Competent Monitoring Authority samples were delayed due to deadlines issues to their contractors for data delivery). 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 it within a week of report receipt; these considerations would greatly facilitate the analysis of results and effective feedback.

2. The range in the number of samples in data sets provided for selection of the Own Samples ranged from 6 to 279 and averaged 38 samples available for Own Sample selection. The number of project data sets submitted ranged from 1 to 7 with the average percentage audit being 15% of the submitted data. Best practice for commercial laboratories should be to use the Scheme as an external auditor and no 'cherry picking', pre-analysis selection, or pre-submission re-working of samples should be undertaken.

3. One set of Own Samples were submitted with residues of <0.5mm fractions only. Participants are reminded that Own Samples must include all sorted residues, including all extracted materials deemed 'unrecordable' during the initial processing; and all recorded taxa must be submitted to the auditing laboratory. Failure to supply all sample components according to the NMBAQC OS Protocol will result in the assignment of a 'Fail' audit flag.

4. There were continued problems associated with the measurement of biomass for individual species. 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 CSEMP/WFD sample should in no way compromise the effectiveness of an audit. Biomass procedures should not render the specimens unidentifiable; trials would help to derive the best protocol for the blotted weighing technique. Biomass must be reported to four decimal places with nominal weights recorded as 0.0001g. A standardised protocol is available and must be followed for CSEMP/WFD analysis.

5. There were continued problems with specimens being provided in containers which are 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.

6. One set of Own Samples specimens were submitted in formaldehyde solution and not Industrial Denatured Alcohol (IDA). A large number of Own Sample residues were also submitted in formaldehyde solution. Participants are reminded that Own Samples should be submitted to the APEM Ltd. in 70% IDA.

7. 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.

8. Participants submitting data for laboratory reference exercises should add a note on location of sample to aid identification. A similar 'Habitat Notes' section to that distributed with the ring test exercises will be distributed for completion in 2016/17 (Scheme year 23).

9. Participants submitting data for the ring test exercises should complete the 'literature used' section to enable additional information to be gathered regarding incorrect identification. In some cases this information could result in a laboratory being marked correctly for what could be perceived as a mis-identification without that information.

10. 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 datasheets to enable additional information to be gathered regarding the difficulty of ring test specimens.

11. There was a problem with mixtures of specimens being sent out under both RT exercises this year which led to a significant amount of extra time required in sorting out these issues and delayed the production the ring test bulletins. APEM appreciate

the discussions that these errors created and believe they led to a greater understanding in the taxonomy of these groups (for example, specimen 21 in RT50, the talitrid) but also appreciate that all specimens sent out under the ring tests should be the same for all participants. Extra vigilance will be employed when preparing difficult specimens for future ring tests.

12. All Own Sample submissions must be accompanied with a 'processing details sheet' to ensure that the re-analysis (audit) matches that of the initial processing. Laboratories should also ensure that these sheets are completed accurately. Own Samples processed for CSEMP/WFD must be processed according to the NMBAQC guidelines (Worsfold, Hall & O'Reilly (Ed.) 2010).

13. 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.

14. 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.

15. 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 atypical number of over-cautious identifications and multiple taxa recorded for a single species, which will lead to data comparison issues for spatial and temporal studies. 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.

16. An improved learning structure to the Scheme through detailed individual exercise reports has been successfully implemented and was continued in this Scheme year. For the LR and OS Modules, detailed results have been forwarded to each participating laboratory as soon after the exercise deadlines as practicable. After each RT exercise a bulletin was circulated, reviewing the literature used and detailing the correct identification of the taxa circulated. Participants are encouraged to review their exercise reports and provide feedback concerning content and format wherever appropriate. Valuable feedback was received from participants for RT49 and RT50. Hopefully the feedback and additional discussions will result in further revision of the literature. APEM Ltd. wish to thank all participants that submitted feedback, photos and were involved in further discussions regarding both RT49 and RT50.

17. Positive, constructive feedback has been received from participants during Scheme Year 2015/2016 (Year 22). As in previous years, participants have expressed the benefits of the modules, especially RT and OS. 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. Three of the fifteen 'failing' Own Samples in Scheme Year 2015/2016 (Year 22) have already been rectified via the recommended remedial action

18. Additional guidance for Own Sample 'next steps' following audit results will be created to ensure that all participants and other stakeholders are aware of the route to quality

assured data. The Scheme provides quality assurance for the UK's 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.

19. If participants have queries, or wish to raise issues regarding Own Sample or Ring Test specimen identifications this must be done in a timely manner. Issues have been raised up to two months after the interim reports which led to delays with exercise, module and Annual Report.

20. APEM Ltd. strives to ensure smooth running and transparency of the Scheme at all times. Consideration should be given by participants as to the tone of correspondence with APEM Ltd. Participants should remember that APEM Ltd. must log and make available all correspondence to the Benthic Invertebrate Contract Manager (Myles O'Reilly, SEPA). As such participants should not communicate anything regarding the Scheme or Scheme Contractor that they would not wish to be shared with the Contract Manager. 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. and 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.

Milner, C., Hall, D. and O'Reilly, M. (Ed.) 2016. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Own Sample Module Summary Report OS59, 60 & 61. Report to the NMBAQC Scheme participants.* 23pp, March 2016.

Milner, C., Worsfold, T., Hall, D. & Pears, S., 2016. *National Marine Biological Analytical Quality Control Scheme. Ring Test Bulletin: RTB#49. Report to the NMBAQC Scheme participants. APEM Report NMBAQC RTB#49.* 32pp, Feb 2016.

Milner, C., Worsfold, T., Hall, D., Ashelby, C. and Pears, S., 2016. *NE Atlantic Marine Biological Analytical Quality Control Scheme. Ring Test Bulletin: RTB#50. Report to the NMBAQC Scheme participants. APEM Report NMBAQC RTB#50*. 23pp, March, 2016.

Worsfold, T.M., Hall, D.J. and 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.