NMBAQC

NE Atlantic Marine Biological Analytical Quality Control Scheme

Particle Size Analysis Component Annual Report
Scheme Operation 2015/2016 (Year 22)

Authors: Lydia Finbow (APEM), NMBAQCS Particle Size Analysis Administrator

Prof. Kenneth Pye (KPAL), NMBAQCS Particle Size Benchmark Analyst

Reviewer: David Hall (APEM), NMBAQCS Project Manager

Approved by: Claire Mason (Cefas), Contract Manager

Contact: nmbaqc@apemltd.co.uk

APEM Ltd.
Date of Issue: March 2016

PARTICLE SIZE COMPONENT ANNUAL REPORT FROM APEM Ltd

SCHEME OPERATION – 2015/16 (Year 22)

1. Introduction 3

1.1 Summary of Performance 4

1.1.1 Statement of Performance 5

2. Summary of PSA Component 5

2.1 Introduction 5

2.1.1 Logistics 6

2.1.2 Data returns 6

2.1.3 Confidentiality 6

*2.2 Particle Size Analysis (**PS**) Module* 6

2.2.1 Description 6

2.2.2 Results 8

2.2.3 Discussion 15

2.2.4 Application of NMBAQC Scheme Standards 17

*2.3 Particle Size Own Sample Analysis (**PS-OS**) module* 18

2.3.1 Description 18

2.3.2 Results 19

3. Conclusions and Recommendations 21

4. References 23

Linked Documents (hyperlinked in this report):

Particle Size Exercise Results – PS56

Particle Size Exercise Results – PS57

Particle Size Exercise Results – PS58

Particle Size Exercise Results – PS59

1. Introduction

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

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

APEM Ltd has been the administrative contractor for the Particle Size component since 2014 (Scheme year 21).

The Particle Size (PS) module followed the format of 2014/15. A series of exercises involved the distribution of test materials to participating laboratories and the centralised examination of returned data and samples.

The new Particle Size Own Sample (PS-OS) module, introduced in the 2014/15 Scheme year, followed the same logistical format as the previous year. There were changes to the reporting format. It was decided that for 2015/16 (Scheme year 22) a simpler, clearer report was required. A revised report was created to compare primary and AQC sieve and laser data separately along with data merging accuracy and assess whether a representative sample was supplied for reprocessing. The purpose of this module was to examine the accuracy of particle size analysis for participants' in-house samples. The Particle Size Own Sample module is a training / audit module. Participants' samples are re-analysed by the NMBAQC Scheme PSA contractor and the results are compared. PS-OS exercises receive a "Good" or "Review" flag for each element; a "Review" flag is provided with additional comments highlighting errors and areas for improvement. For 2015/16 the PS-OS results will not be used to assess the performance of a laboratory. The effectiveness of these flags is assessed in this annual report.

Sixteen laboratories signed up to participate in the 2015/16 PS Module exercises (PS56, PS57, PS58 and PS59); six were government laboratories and ten were private consultancies. Nine laboratories signed up to participate in the PS-OS Module exercises (PS-OS04, PS-OS05 and PS-OS06); seven were government laboratories and one was a private consultancy. One government laboratory had two Lab Codes to submit six PS-OS samples for AQC analysis.

To reduce potential errors and simplify administration, Lab Codes were assigned with a prefix to determine the Scheme component; all codes for the Particle Size component were prefixed with "PSA_".

As in previous years, some laboratories elected to be involved in limited aspects of the Scheme. Competent monitoring authorities (CMAs) completing PSA in support of biological analysis for monitoring programmes (including in assessment of MPA (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. The Scheme is aware of other PSA methodologies (*e.g.* those used in the Regional Seabed Monitoring Plan) and encourages those involved in any relevant PSA monitoring programmes to participate in this Scheme, especially where pass/fail criteria can be used to assess overlapping aspects of different methodologies.

1.1 Summary of Performance

In previous years the Particle Size (PS) module 'Pass/ Fail' criteria were based upon z-scores from the major derived statistics with an acceptable range of ±2 standard deviations (see Description of the Scheme Standards for the Particle Size Analysis Component). The annual report for 2009/10 (Scheme year 16) deemed the use of z-scores inappropriate for such a low number of data returns where two erroneous results can significantly alter the Pass / Fail criteria. The z-score method also assumes that the majority of respondents are correct and raised genuine concerns regarding technique and method bias. Following this, the 'Pass/ Fail' criteria are currently under review and alternative flagging criteria are being trialled. For the 2014/15 year, evaluation of the PS module results included z-score calculations for each half-phi interval, multi-variate analysis in the form of dendrograms and MDS (multi-dimensional scaling) plots, particle size ternary diagrams to determine sediment distribution, as well as assessment of sieve and laser metadata. Following a review of the 2014/15 data, a new method of Pass /Fail was developed for 2015/16 (Year 22) using z-scores with robust statistics. Z-scores were calculated on statistics from the merged data, the statistics used were the D_{10}, D_{50}, D_{90} and Mean particle size in microns. Participants received a Satisfactory, Questionable or Unsatisfactory result based on the z-score. Results between -2.0 and 2.0 were Satisfactory, between ±2.0 and ± 3.0 were Questionable and results greater than ± 3.0 were Unsatisfactory. Participants then received a score and a Pass or Fail based on their results for each statistic, as shown below (Table 1).

Table 1. Pass/Fail flag criteria based on z-scores using robust statistics.

Z- Score	Result	Score
-2.0 to 2.0	Satisfactory	5
±2.0 to ±3.0	Questionable	2
>3.0 or <-3.0	Unsatisfactory	0

Total Score	Pass/Fail	Level
20	PASS	Excellent
15 – 19	PASS	Good
12 – 14	PASS	Acceptable
6 – 11	FAIL	Poor
0 -5	FAIL	Bad

Comments are provided on the individual performance of the participating laboratories in each of the modules.

1.1.1 Statement of Performance

Each participating laboratory received a copy of the interim results for each exercise; these included a summary of results provided by each laboratory and a basic discussion of any major outliers. Further details and analysis can be found in this report.

At the end of the Scheme year each laboratory received a ‘Statement of Performance’, which included a summary of results for each of the Scheme’s modules and details the resulting flags where appropriate. These statements were first circulated with the 1998/1999 annual report, for the purpose of providing proof of Scheme participation and for ease of comparing year on year progress.

2. Summary of PSA Component

2.1 Introduction

The two 2015/16 year PSA modules, PS and PS-OS are described in more detail below. A brief outline of the information to be obtained from the 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). Email was the primary means of communication for all participating laboratories. This has considerably reduced the amount of paper required for the administration of the Scheme.

2.1.2 Data returns

Spread-sheet based workbooks were distributed to each participating laboratory via email for each circulation and data returned to APEM Ltd via the NMBAQC Scheme email address. 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

To preserve the confidentiality of participating laboratories, each was identified by a four-digit Laboratory Code prefixed with "PSA_", to identify the scheme component. In May 2015 each participant was given a confidential, randomly assigned 2015/16 (Scheme year twenty-two) Lab Code. Codes are prefixed with the Scheme year to reduce the possibility of obsolete codes being used inadvertently by laboratories, *e.g.* Laboratory number four in Scheme year twenty-two (2015/16) was recorded as PSA_2204.

2.2 Particle Size Analysis (**PS**) Module

2.2.1 Description

This component examined the percentage of sediment found in each half-phi interval from the particle size analysis of replicate sediment samples. Four samples of sediment, one fine (PS56), one coarser (PS57) and two diamictons (PS58 and PS59) were distributed in 2015/16. The samples were distributed in two stages; the first circulation (PS56 and PS57) was sent to participants on 13/05/2015 and the second circulation (PS58 and PS59) was sent on the 14/10/2015. For each circulation participants were given approximately 6 weeks to complete their analysis and send completed workbooks via email to APEM Ltd. PS56 replicate samples were derived from natural marine sediments; PS57 replicates were artificially prepared from commercial aggregate materials; PS58 and PS59 replicates were prepared from a combination of natural sediments and artificially prepared commercial aggregate; they were prepared at APEM's Letchworth laboratory as described below.

2.2.1.1 Preparation of the Samples

The first PS circulation, PS56, was a mud collected from natural marine environments from Barry Island. Approximately 30 litres of visually similar sediment was collected and returned to the laboratory where it was wet sieved at 0.5mm to remove any particles larger than 0.5mm. Sediment that passed through the 0.5mm sieve was retained in a large tray, mixed and left to settle before it was cored into replicate samples approximately 150 grams in weight. The second exercise, PS57, was artificially created from commercially acquired pea shingle that was split into half-phi intervals by dry sieving using a mechanical sieve shaker. The third exercise sample (PS58) was a diamicton sample consisting of a mixture of pre-sieved (1.0mm) sand from Cornwall and known quantities of maerl-based sediment from near Dean Quarry, dry sieved at half phi intervals. The final sediment (PS59) was a mixed sediment artificially created from combined natural sediments (sand from South-East coast and mud from Cornwall) and commercially acquired materials. For the mixed samples (PS58 and PS59) approximately 200g of water was added to help mix the sample together.

Five replicates were sent for particle size analysis to assess the degree of inter-sample variation and produce benchmark data. Where laser diffraction analysis was required, these replicates were analysed using a Coulter LS13320 laser diffraction instrument. The remaining replicates were randomly assigned to participating laboratories and distributed according to the Scheme timetable. Spare replicates were kept at the APEM Ltd. Letchworth laboratory in case of problems such as damaged samples during delivery or significant processing errors.

2.2.1.2 Analysis required

The participating laboratories were required to conduct particle size analysis on the samples following the NMBAQC's best practice guidance for particle size analysis to support biological data (Mason, 2011, (this version has since been updated)), either in-house or using a subcontractor. A written description of the sediment characteristics was to be recorded, with a visual estimate pre-processing and using the Folk (1954) textural classification Triangles post-processing as well as the percentage gravel, sand and silt/clay and an indication of any peroxide treatment or chemical dispersant used. Also requested was a breakdown of the particle size distribution of the sediment, to be expressed as a weight or percentage of sediment at half-phi (ϕ) intervals.

The 2015/16 workbooks had the same format as the previous year. As in the previous year, data provided in the "Participant Sieve Metadata" and "Participant Laser Metadata" spreadsheet tabs were for analytical purposes only and were not published in the Interim Results reports.

Approximately six weeks were allowed for the analysis of each pair of PS samples sent out (i.e. PS56 & PS57, PS58 & PS59).

2.2.2 Results

2.2.2.1 General comments

Sixteen laboratories subscribed to the exercises in 2015/16. For the first circulation (PS56 and PS57) all subscribing participants provided results; for the second circulation (PS58 and PS59) fifteen subscribing laboratories provided results, with one participant (PSA_2204) e-mailing details of their non-participation due to staffing changes.

Most participating laboratories now provide data in the requested format, although some variations remain. As reported previously, it should be remembered that the results presented may be from a more limited number of analytical laboratories than is immediately apparent since this component of the Scheme is often sub-contracted by participants to one of a limited number of specialist laboratories. One laboratory provided two sets of results run by two different analysts for each of the exercises. Detailed results for each exercise (PS56, PS57, PS58 and PS59) have been reported to the participating laboratories; additional comments are provided below.

2.2.2.2 Analysis of sample replicates (benchmark data)

Five replicate samples of the sediment used for the four PS distributions were analysed by Kenneth Pye Associates Ltd (KPAL) to examine variability and establish benchmark data. Replicate samples were analysed, where required, using Endecotts British Standard 300mm and 200mm test sieves, Endecotts EFL 2000/2 and Retsch AS2001 Control 'g' sieve shakers and a Beckman Coulter LS13320 laser size analyser. In previous Scheme years replicates were analysed by both laser diffraction and sieve / pipette methods; however, as the majority of laboratories are now conducting analyses by laser diffraction the testing of replicates for 2015/16 was undertaken only using a laser diffraction instrument.

An extra table was introduced into the 2015/16 PS reports to analyse the variability between the benchmark replicates. The Coefficient of Variation (CV) was calculated for the D_{10}, D_{50},

D_{90} and Mean particle size in microns. The CV is most commonly expressed as the standard deviation as a percentage of the mean and describes the dispersion of a variable in a way that does not depend on the variables' measurement units. A low CV indicates a smaller amount of dispersion in the variable. Good reproducibility was shown for replicates when the %CV was <3% for the D_{50} and <5% for the D_{10} and D_{90}, all limits were doubled when the D_{50} was less than 10µm, in line with recommendations in BS ISO 13320.

Analysis of the replicates for Sample PS56 indicated an average composition of 1.15% sand and 98.85% mud, classified as 'Mud' according to the Folk (1954) scheme and "Very Slightly Sandy Mud" according to the Blott & Pye (2012) scheme. Only laser analyses were required for this sample. The %CVs for each statistic were well within the limits showing that the replicates had good reproducibility. Results for the individual replicates are provided in Tables 1, 2, 3 and 4 and are displayed in Figures 1 and 2 (PS56 Report).

Sample PS57 was an artificial gravel sediment and contained an average of 99.50% gravel and 0.50% sand, classified as a 'Gravel' according to both the Folk (1954) and Blott & Pye (2012) schemes. The replicates were analysed by dry sieving only. The replicates showed extremely low variation, with %CV well below the limits for each statistic. Results for the individual replicates are provided in Tables 1, 2, 3 and 4 and are displayed in Figures 1 and 2 (PS57 Report).

Sample PS58 was a diamicton and both sieve and laser analyses were required. The sample contained an average of 10.30% gravel, 88.62% sand and 1.09% mud and was classified as a 'Gravelly Sand' according to the Folk (1954) and "Very Slightly Muddy Slightly Gravelly Sand" by the Blott & Pye (2012) scheme. The replicates showed extremely low variation, with %CV well below the limits for each statistic. Results for the individual replicates are provided in Tables 1, 2, 3 and 4 and are displayed in Figures 1 and 2 (PS58 Report).

Sample PS59 was also a diamicton and both sieve and laser analyses were required. The results showed an average of 48.06% gravel, 37.87% sand and 14.05% mud. The sediment is classified as 'Muddy Sandy Gravel' according to the Folk (1954) scheme and as 'Slightly Sandy, Muddy Gravel' according to the Blott & Pye (2012) scheme. Results for the individual replicates are provided in Tables 1, 2, 3 and 4 and are displayed in Figures 1 and 2 (PS59 report).

2.2.2.3 Results from participating laboratories

In each of the PS56, PS57, PS58 and PS59 reports, Table 5 shows summary data i.e. the percentage gravel, sand and silt/clay recorded as well as the participants' post analysis sediment descriptions. Where the summary statistics were not provided by participating laboratories they were calculated by APEM Ltd. The summary statistics were verified by APEM using the GRADISTAT program (Blott & Pye, 2001) based on the final half-phi frequency data provided by each laboratory. Table 6 provides a summary of the > 1mm and < 1mm wet separation weights determined by each participating laboratory and the benchmark data. The < 1mm weight should have been the sum of the oven-dried < 1mm fraction plus the weight of sediment in the sieved > 1mm fraction base-pan. Table 7 shows a summary of the final laser data submitted by the participants in one phi intervals, and the total column indicates whether or not the laser data has been re-proportioned; correctly re-proportioned laser data should equal exactly 100%. Table 8 shows the calculations for the robust mean and robust standard deviation or the standard deviation for proficiency assessment (SDPA). Robust statistics are a way of summarising results when it is suspected that there may be a small proportion of outliers, the robust mean is calculated as the median and the robust standard deviation is the normalised Median of Absolute Deviations (MADe) from the sample median. Table 9 shows a summary of the z-scores for each test statistic (D_{10}, D_{50}, D_{90} and Mean particle size); those greater or less than ±2.0 are highlighted in yellow to indicate the results that were "unsatisfactory" or "questionable". Table 10 summarises the results indicating whether the participant received a Pass or Fail flag and the level of that result e.g. "PASS – Excellent" or "Fail – Bad" etc.

Figure 3 shows the particle size distribution curves for each of the exercises. Included in each of these figures, for comparison, are the mean distribution curves for the replicate samples obtained by KPAL. Figure 4 displays comparative bar charts of the major sediment components (% sand, gravel and mud) for each laboratory and for each exercise. Figure 5 shows summary plots of z-scores achieved by each participating laboratory.

2.2.2.4 Fifty- sixth distribution – PS56

There was good agreement for PS56 between the results for the replicates and those supplied by some of the participating laboratories, although the latter showed considerable variation (see Figure 3). Table 5 shows the variation in data received from the participating laboratories; percentages of sand ranged from 1.2% (PSA_2203) to 38.6% (PSA_2201), and percentage mud ranged from 61.3% (PSA_2201) to 98.8% (PSA_2203). Three laboratories pre-treated their sample with a chemical dispersant; PSA_2203 used 3% Calgon solution and PSA_2214 (A and B) and PSA_2218 used Sodium Hexametaphosphate. Participant PSA_2209 stated they were using an in-house methodology but did not provide any details on how this differed from the NMBAQC methodology. Participant PSA_2218 used the British Standard Pipette method as they did not have a laser analyser. The majority of laboratories only used laser analysis, PSA_2201 used both sieves and laser and PSA_2218 only used sieves, as stated earlier they did not have a laser analyser. Participant PSA_2217 did not provide the summary data required for Table 5 so these were calculated by APEM based on the final merged data supplied. Two participants calculated the percentages of sand and silt/clay incorrectly; PSA_2208 had a 2.3% difference in the %sand and silt/clay and PSA_2210 had a 3.7% difference. Table 7 shows that most participants provided re-proportioned laser data. PSA_2209 did not provide re-proportioned laser data in the final laser data tab as it summed to 99.89% and PSA_2215 recorded final laser and final merged data at 100.11%, which is not possible.

This sample comprised fine sediment and most laboratories only used laser analysis therefore the final laser data should have been equal to the final merged data, however for a few laboratories this was not the case; comparisons of the final laser and final merged data can be seen in Appendix 2. PSA_2208 final merged data summed to 126.93, this was assumed to be in grams and converted to percentages for comparative analysis however this did not match up with their laser data. PSA_2210 final laser data summed to 100% however the final merged data only summed to 99.37%. PSA_2216 final merged data summed to 989.00, this was assumed to be in grams so converted to percentages for comparative analysis however, the sample was sent in a 200ml pot so it was not possible for it to weigh this much. It appears that these laser data are multiplied by 9.89 for some unknown reason.

Despite these issues the z-scores produced “Pass” flags for all participants except for PSA_2201 who received a “Fail – Bad”, with three unsatisfactory results (D_{50}, D_{90} and Mean) and one questionable result (D_{10}). Of the “Pass” flags, twelve received “Pass – excellent” (PSA_2202, PSA_2203, PSA_2204, PSA_2205, PSA_2209, PSA_2212, PSA_2213, PSA_2214 (A and B), PSA_2215, PSA_2217 and PSA_2218) and the remaining four received “Pass – Good” (PSA_2208, PSA_2210, PSA_2211 and PSA_2216).

2.2.2.5 Fifty-seventh distribution – PS57

There was generally good agreement for PS57 between the results from the analysis of the benchmark replicates and those from the participating laboratories (see Figure 3), although three laboratories stand out from the rest and received “Fail” flags. Participant PSA_2204 reported the highest percentage of sand (1.08%) see Table 5. As seen in the bar chart in Figure 4 they were the only laboratory to report a Coarse Sand fraction and produced the second highest percentage of Medium Gravel. PSA_2204 are also one of only two laboratories who attempted laser analysis on their sample, however their final laser data only sums to 0.81% and a closer investigation reveals this is the percentage of the final merged data rather than raw re-proportioned laser data. The other laboratory that attempted laser analysis was PSA_2209; their sieve data showed only 0.2grams of < 1mm sediment was recorded. Rather than attempting laser analysis, this small amount should have been added to the 0.00 to 050 phi (707µm), as this is not enough sediment for accurate laser analysis. Percentage gravel, sand and silt/clay for PSA_2209 did not add up to 100% (see the summary data in Table 5). PSA_2209 recorded gravel at 98.8%, sand at 1.00% and silt/clay at 0.0%, based on their final merged data these should have been recorded as gravel, 99.94%, sand, 0.06% and silt/clay, 0.01%. It appears that this laboratory measured the gravel fraction from 63mm to 4mm rather than 63mm to 2mm. Due to the summary data being reported at 1 decimal place the silt/clay fraction that was recorded was lost. Figure 4 shows that PSA_2209 also reported the highest proportion of Medium Gravel (74.80%) and lowest proportion of Fine Gravel (12.74%). The third participant to receive a “Fail” flag is PSA_2208, this participant did not record any sediment for intervals -4.5 to -4.0 and -4.0 to -3.50, and instead this sediment was incorporated into interval -3.5 to -3.0 causing this interval to have a higher weight compared to other participants.

Other than the three “Fail” flags mentioned previously, all other participants received “Pass” flags; two with “Pass - Good” (PSA_2210 and PSA_2212) and the rest receiving “Pass – Excellent” (PSA_2201, PSA_2202, PSA_2203, PSA_2205, PSA_2211, PSA_2213, PSA_2214, PSA_2215, PSA_2216, PSA_2217, PSA_2218).

2.2.2.6 Fifty-eighth distribution – PS58

There was generally good agreement for PS58 between the results from the analysis of replicates and those from the participating laboratories (see Figure 3). The percentage gravel recorded ranged from 7.60% (PSA_2213) to 11.66% (PSA_2217), sand ranged from 64.92% (PSA_2217) to 91.68% (PSA_2218) and the silt/clay recorded ranged from 0.00% (PSA_2215) to 23.42% (PSA_2217). Four participants received “Fail” flags based on the z-scores, three “Fail – Bad” (PSA_2208, PSA_2217 and PSA_2218) and one “Fail – Poor” (PSA_2215). From Figures 3 and 4 it is unclear why PSA_2208 received a “Fail” however,

Table 9 shows that this laboratory produced the third lowest D_{10} (188.94µm), the highest D_{50} (643.86 µm), the highest Mean (933.06µm) and the fourth highest D_{90} (1999.43µm). PSA_2215 stated that they used the NMBAQC method but this is not the case as, rather than using the laser to analyse the < 1mm fraction, they sieved down to 63µm; this means that they did not record any silt/clay fraction and recorded the second highest percentage sand (90.95%), resulting in the third highest D_{50} (411.34µm). PSA_2217 recorded the highest percentage of silt/cay (23.42%); however this participant contacted APEM Ltd after the interim reports were issued to re-submit their data as they had mixed up the < 1mm fractions of PS58 and PS59. Their re-submitted data reported Gravel at 11.66%, Sand at 87.88% and Silt/Clay at 0.41%; these data would have resulted in a "Pass – Good" overall. PSA_2218 recorded the highest proportion of Sand (91.68%) and the second lowest proportion of Silt/Clay (0.12%). It must be taken into consideration, as mentioned earlier (Section *2.2.2.4*), that this participant did not own a laser and therefore their data are not directly comparable with the other participants, particularly when dealing with fine sediments. PSA_2216 received a "Pass – Acceptable" based on the z-scores, however they stand out from the other participants in Figures 3 and 4 as they recorded the second highest proportion of Silt/Clay; aside from PSA_2217, who re-submitted their data due to a mix up with PS59.

Table 7 shows a summary of the final laser data provided by participants and shows that many laboratories did not provide re-proportioned laser data i.e. totalling 100%. PSA_2201 (99.94%), PSA_2202 (99.73%), PSA_2205 (98.53%), PSA_2208 (96.09%), PSA_2210 (100.12%), PSA_2211 (99.71%), PSA_2216 (98.51%) and PSA_2217 (65.62%) all provided non-re-proportioned laser data. Of these eight laboratories four (PSA_2205, PSA_2208, PSA_2211 and PSA_2216) have re-proportioned the laser data prior to merging as their final merged data equals 100%. The remaining four laboratories have discrepancies in their final merged data that could originate from either not re-proportioning laser data or from not merging data correctly. PSA_2201 has final merged data that sums to 99.96% this is directly caused by not re-proportioning their laser data from 99.94% to 100%. The same problem occurs for PSA_2202, their final merged data sums to 99.83%, again caused by not re-proportioning the laser data from 99.73% to 100%. PSA_2210 recorded final laser data of 100.12% which is not possible, however they have re-proportioned it to 100% before merging but data rounding has meant their final merged data only sums to 99.99%. PSA_2217 have reported final laser data as 65.62%, there is obviously an error here and their final merged data totals 100.02%, which is not possible, and clearly there are some issues.

However, as stated earlier in this section, participant PSA_2217 re-submitted their data due to a mix up between PS58 and PS59, and; their re-submitted merged data equalled 100%.

Other than the four participants that received “Fail” flags, two participants received “Pass – Acceptable” (PSA_2202 and PSA_2216), six received “Pass – Good” (PSA_2201, PSA_2203, PSA_2205, PSA_2211, PSA_2213, PSA_2214B) and the remaining four received “Pass – Excellent” (PSA_2209, PSA_2210, PSA_2212, PSA_2214A).

2.2.2.7 Fifty-ninth distribution – PS59

There was a lot of variation in results between laboratories and between the laboratories and the benchmark data (see Figure 3). All laboratories reportedly followed NMBAQC methods and used both sieve and laser analysis apart from three. PSA_2218 who used a combination of dry sieve analysis and pipette methods as they do not own a laser sizer. PSA_2209 stated “in-house” methodology was used but did not provide any details on how this differed from the NMBAQC method.PSA_2215 stated they used the NMBAQC method but wet split the sample at 2mm rather than 1mm. The weight of > 1mm sediment ranged from 256.45g (PSA_2210) to 301.4g (PSA_2209) and the weight of < 1mm sediment ranged from 237.57g (PSA_2216) to 323,23g (PSA_2212) (see Table 6). The majority of laboratories recorded a >1mm to <1mm ratio in line with the benchmark data (0.926), three laboratories recorded a lower ratio, 0.84 (PSA_2205), 0.86 (PSA_2210) and 0.85 (PSA_2213) and two laboratories recorded ratios greater than 1; PSA_2216 (1.26) and PSA_2218 (1.02). PSA_2215 appear to have filled in the sieve data tab incorrectly as they stated there was no material < 1mm even though they used laser analysis. Investigation into this reveals that they have entered their sieve data as a percentage of the final merged data rather than raw weights in grams. Table 5 shows that the percentage gravel recorded by laboratories ranged from 45.69% (PSA_2205) to 55.77% (PSA_2216) and that the percentage sand ranged from 26.76% (PSA_2205) to 50.54% (PSA_2217). There was a wide range of values reported from the percentage Silt/Clay with PSA_2205 reporting the highest (27.54%) and PSA_2216 reporting the lowest (0%). Z-scores in Table 10 show that only two laboratories (PSA_2215 and PSA_2216) received “Fail” flags using the robust statistics method. Table 5 shows that PSA_2215 had the second highest percentage gravel (50.99%), and fourth lowest percentage of Silt/Clay (6.47%), resulting in D_{10} and D_{50} statistics that are too high compared to other laboratories. Figure 4 shows that PSA_2216 recorded a significantly higher percentage of gravel compared to other laboratories and recorded no Silt/Clay fraction, resulting in D_{10}, D_{90} and Mean statistics that were too high in comparison to other laboratories (see Table 9).

PSA_2217 contacted APEM after the issue of the interim reports to re-submit data as there had been a mix up with their PS58 and PS59 samples. Their original data received a "PASS – Acceptable" flag whereas the re-submitted data would have received a "PASS – Excellent" flag. All other laboratories received "Pass" flags, with nine "PASS – Excellent" (PSA_2201, PSA_2202, PSA_2203, PSA_2205, PSA_2208, PSA_2209, PSA_2211, PSA_2212, PSA_2213), two "PASS – Good" (PSA_2210 and PSA_2218) and other than PSA_2217, one "PASS – Acceptable" (PSA_2214). Although the benchmark samples showed good homogeneity, the extreme variation between laboratories means that the calculation of z-scores is inappropriate since the calculated "robust" mean is unreliable. This causes laboratories to receive "Pass" flags that should perhaps be receiving Fails, for example PSA_2205 recorded much higher percentage of Silt/ Clay (27.54%) but still received a "PASS – Excellent" flag.

2.2.3 Discussion

The samples distributed as PS56 appeared from an analysis of replicates (Figure 2) to be good replicates with very little variance and the coefficient of variance for the D_{10}, D_{50}, D_{90} and Mean showed that the replicates had good reproducibility. Results from participating laboratories (Figure 3) showed a fair degree of variation in the distribution curves, this is not surprising given the nature of the sediment type and type of analysis required (laser diffraction). The mud from Barry Island contains aggregates, some containing coal and ash or slag as well as organics and shell fragments, clay, silt and a small amount of fine sand sized quartz; these need to be carefully dispersed before analysis. Procedures for dispersion are likely to differ between laboratories and can have a major affect on results, as can the optical model used, and different algorithms in different instruments are also likely to interpret the diffraction patters differently. PSA_2201 reported much too high a sand fraction and as a result was the only laboratory to receive a "FAIL – Bad" flag based on z-scores. Participant PSA_2211 and PSA_2216 received "PASS – Good" flags but their D_{50} (20.57µm and 13.91 µm, respectively) values are far outside what would be considered as an acceptable range. PSA_2208 reported no < 4 µm material at all and PSA_2205, PSA_2212 and PSA_2216 reported only small amounts, possibly due to sample dispersion and instrument measurement capabilities.

The samples distributed as PS57 appeared from an analysis of replicates (Figure 2) to be good replicates with very little variance and the coefficient of variance for the D_{10}, D_{50}, D_{90} and Mean showed that the replicates had good reproducibility. The majority of results from participating laboratories were similar (Figure 3). Two participants received "FAIL – Bad" flags (PSA_2204 and PSA_2209). Figure 4 shows that both of these laboratories reported too much medium gravel, and PSA_2204 were the only laboratory to report sediment classified as coarse sand. PSA_2208 received a "FAIL – Poor" flag, which, looking at Figures 3 and 4 may seem a bit harsh, however Table 9 shows they had an elevated D_{10} value compared to the average and lower than average D_{90} and Mean values.

The samples distributed as PS58 appeared from an analysis of replicates (Figure 2) to be good replicates with very little variance and the coefficient of variance for the D_{10}, D_{50}, D_{90} and Mean showed that the replicates had good reproducibility. Results from participating laboratories were generally in accordance (Figure 3). Participants PSA_2215 and PSA-2218, did not use laser diffraction in accordance with the NMBAQC methodology, however PSA_2218 do not own a laser. These two laboratories, as well as PSA_2202, PSA_2208, PSA_2209, and PSA_2210, reported very little or no Silt/Clay fraction. PSA_2216 and PSA_2217 reported too much Silt/Clay (10.4% and 23.4%, respectively); however PSA_2217 re-submitted their results after the interim report after a mix up with PS59, their updated results reported a Silt/Clay fraction of 0.41%. The laser data in Table 7 showed that PSA_2208 and PSA_2216 recorded too much coarse sand (0.00 to 1.00 phi) and that seven laboratories did not provide re-proportioned laser data.

The samples distributed as PS59 appeared from an analysis of replicates (Figure 2) to be good replicates with very little variance and the coefficient of variance for the D_{10}, D_{50}, D_{90} and Mean showed that the replicates had good reproducibility. Results from participating laboratories (Figure 3) showed a fair degree of variation in the distribution curves. PSA_2216 and PSA_2215 were the only two laboratories to receive "Fail" flags based on the z-scores. PSA_2216 reported the highest amount of gravel and no silt/clay; both laboratories had exceedingly high D_{50} results. Although they received a "PASS – Excellent" flag, PSA_2205 and PSA_2209 reported unacceptably high silt/clay fractions (27.54% and 21.82%, respectively).

Participating laboratories were asked to provide a visual description of the PS56, PS57, PS58 and PS59 samples prior to analysis and instructed to describe the sediment using the Folk triangle post analysis, as well as to report the percentages of gravel, sand and silt/clay in each exercise. Data were provided by all but one (PSA_2217) participating laboratories for PS56 and by all laboratories for PS57, although PSA_2217 did not provide a post analysis sediment description. Two participating laboratories (PSA_2205 and PSA_2208) did not provide any summary data for PS58 and PSA_2217 did not provide the post analysis sediment description. Three participating laboratories (PSA_2205, PSA_2208 and PSA_2216) did not provide any summary data for PS59 and PSA_2217 did not provide the post analysis sediment description. APEM Ltd checked participants' calculations using GRADISTAT based on the participants' final merged data. Of the data provided for PS56, the majority were correct apart from PSA_2208 and PSA_2210, who provided data a few per cent out for sand and silt/clay. For PS57 the majority of laboratories were correct; PSA_2209 data only summed to 99.8% and PSA_2210 only summed to 99.9%. PSA_2214 reported a small silt/clay fraction (0.06%) that was not recorded in the APEM verification. All data provided for PS58 and PS59 was correct except for PSA_2217 and PSA_2218 for PS59 and PSA_2215 for PSA_2215. The summary data in each report showed variability in how it was reported in terms of decimal places, with some participants reporting data at two decimal places, some at one decimal place and a few as whole numbers.

2.2.4 Application of NMBAQC Scheme Standards

One of the key roles of the Particle Size Analysis component of the NMBAQC Scheme is to assess the reliability of data collected as part of the Clean Seas Environment Monitoring Programme (CSEMP; formerly UK NMMP) and Water Framework Directive (WFD) monitoring programmes. With this aim, performance target standards were defined for certain Scheme modules and applied in 1996/97 (Scheme year three). These standards were the subject of a review in 2001 (Unicomarine, 2001) and were altered in Scheme year eight; each performance standard is described in detail in the Description of the Scheme Standards for the Particle Size Analysis Component document. In previous years 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. As the Pass/Fail criteria are still under review for the PS exercises, in 2015/16 (Scheme year 22) a "Pass" or "Fail" flag will be assigned to each laboratory for these particular exercises but will not be used to assess the performance of a laboratory.

2.2.4.1 Laboratory Performance

An overall summary of the data reported by each participant are presented in each of the PS exercise reports along with z-scores and a "Pass" or "Fail" flag. However these "Pass/Fail" criteria are still under review and are not to be used to assess the performance of a laboratory. Each laboratory was issued with a Statement of Performance outlining their results and participation in the Scheme.

2.3 Particle Size Own Sample Analysis (**PS-OS**) module

2.3.1 Description

The Particle Size Own Sample (PS-OS) module is a new module introduced in Scheme year 21 (2014/15) and is a training/ audit module. Participants' "own" samples are re-analysed by the NMBAQC Scheme PSA contractor and the results are compared. The purpose of this exercise was to examine the accuracy of particle size analysis for participants' in-house samples. In its first year (2014/15) the PS-OS exercises carried a trial pass/fail criteria based on the correlation between the participant data and the AQC data. After discussions between KPAL, APEM and Scheme manager Claire Mason, it was decided that a more simplistic approach to analysing the results would be more appropriate in identifying errors in participants' results. The results were split into sieve processing, laser processing, data merging and whether a representative sample was supplied. Participants received a "Good" or "Review" flag based on their results. Where a "Review" flag was issued comments were supplied detailing problems that had arisen and where to find information to help address them.

2.3.1.1 Analysis required

Laboratories were requested to submit details of a survey with at least 12 samples from their previous year's Clean Seas Environment Monitoring Programme (formerly NMMP) samples, or similar alternative sampling programmes (if not responsible for CSEMP samples), along with the associated PSA data. Once these data were provided, three samples were randomly chosen by APEM Ltd to be re-analysed by the NMBAQC Scheme's PSA contractor.

Spread-sheet based workbooks were distributed to each participating laboratory via email for each PS-OS exercise. These were to be returned to APEM Ltd via the NMBAQC Scheme email address (nmbaqc@apemltd.co.uk). 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.

In each workbook a written description of the sediment classification was to be recorded, a visual estimate for pre-processing and using the Folk Triangles post-processing along with the percentages of gravel, sand and silt/clay and an indication of any peroxide treatment or chemical dispersant used. Also requested was a breakdown of the particle size distribution of the sediment, expressed as a weight or percentage of sediment in half-phi (ϕ) intervals, as well as sieve and laser metadata to provide insight into laboratory procedures, especially for the laser analysis.

The different components of each PS-OS sample (< 1mm, > 1mm and laser sub-sample) were to be sent to APEM's Letchworth laboratory to be passed on to the NMBAQC Scheme PSA contractors. The two sets of results were then compared by APEM Ltd.

2.3.2 Results

2.3.2.1 General comments

Nine laboratories subscribed to the PS-OS module in 2015/16. Two of the nine lab codes belonged to the same participant to facilitate multiple PS-OS submissions due to the sub contraction of samples. One participant did not participate but sent an email confirmation of their non-participation.

Each laboratory received detailed comparisons of their data with the re-analysis results obtained by the NMBAQC Scheme's contractor. Data was split into sieve processing, laser processing, data merging and whether a representative sample was supplied. At the end of each report participants received a "Good" or "Review" flag based on their results; where "Review" flags were issued, comments were made on errors that had arisen and links were provided to information to help resolve problems.

Laboratories generally provided workbooks with all the correct information. Three laboratories (PSA_2208, PSA_2210 and PSA_2213) provided all necessary fractions of their sample for re-analysis. Three laboratories (PSA_2211, PSA_2206 and PSA_2207) did not provide any laser sub-sample, therefore the dried < 1mm fractions were used for laser analysis but this required soaking the sample for 72 hours and disaggregation in 3% sodium hexametaphosphate solution to achieve adequate dispersion. PSA_2206 and PSA_2207 did report that laser sub-samples were not available as they are only stored for 30 days after results have been reported. PSA_2212 only provided a very small volume of sample, participant reporting that samples were very small especially after organic content analysis had been completed; however the AQC lab reported that these samples were too small to be considered representative of sediment in the field. PSA_2219 provided freeze dried bulk sample which required soaking and disaggregation in 3% sodium hexametaphosphate solution to achieve adequate dispersion. Participant PSA_2219 reported that they were only interested in the < 1mm fraction, therefore although there was > 1mm sediment present in the samples it had not been analysed.

There was generally good agreement between the participants and the AQC results, particularly in terms of basic sediment textural classification. There were a few discrepancies in the sieve data but these are to be expected due to factors such as breakage of particles during repeat analysis and variations in sieving time and vibration amplitude. In the laser results the AQC laboratory detected a higher clay fraction due to the higher resolution and sensitivity of the Coulter 13320 instrument used, and this was taken into consideration when comparing data. The main issue in the PS-OS module related to data merging, with a few of the participants not re-proportioning laser data to 100%; this had a knock-on affect on the final merged data.

2.3.2.2 Discussion

As in the previous year, the PS-OS module raised issues over the interpretation of the methodology set out in the NMBAQC Best Practice Guidelines (Mason, 2015), in particular how the laser analysis is undertaken. These guidelines, originally written in 2011, were based on the widespread use at that time amongst participants of Malvern Instruments laser diffraction instruments that have 15 – 25 second standard run times and generally are restricted to the analysis of material < 1mm in size. The original methodology suggested that:

1. A homogenised sub-sample of approximately 100ml is taken from the bulk sample for laser analysis (Laser Pot).

2. A small representative sub-sample is taken from the Laser Pot and passed over a 1mm sieve using as little water as possible (Replicate 1).

3. Replicate 1 is then run through the laser at the desired obscuration, producing three run results.

Steps 2 and 3 are then repeated to create Replicates 2 and 3, giving a final result of 9 runs to create the final laser data, the average of these 9 runs. The completion of nine analyses, and subsequent merging of results is necessarily a time consuming process, especially if standard run times longer than 15 to 25 seconds are used (e.g. 60 seconds is standard with Beckman Coulter instruments, which are used by some NMBAQC Scheme participants).

It has been demonstrated by KPAL that, for the vast majority of samples, there is little practical benefit in routinely carrying out analysis of three replicate sub -samples if instruments are calibrated properly and accuracy is checked in the normal way using standards and laboratory reference materials, and if samples are homogenised properly both before the sub-sample is taken from the bulk sample and when the representative sample is taken from the laser pot. In relatively rare instances where samples consist very largely of > 1mm size material and it is impractical to obtain a representative test sub-sample for laser analysis from the bulk sample, more consistent laser results can be obtained by taking a test sample from the wet separated < 1mm fraction of the sediment, rather than from the bulk.

Where samples display, or are suspected of, unstable behaviour, such as time-dependent agglomeration, repeat runs of the same laser test sample should be carried out. Sometimes this may require repeat runs of more than three replicates to fully characterise agglomerative behaviour, and to establish the best dispersal procedures required to obtain repeatable results (e.g. ultrasonic treatment before as well as during the analysis run, and/ or use of chemical dispersants). The guidance has now been updated to incorporate most of these findings, with some further follow up expected at future NMBAQC PSA workshops. The guidance can be viewed in Mason (2016).

The returns for the 2015/16 PS-OS module showed that some laboratories, particularly those using Coulter instruments, in routine case work usually only run one replicate through the laser, with replicates run every 20th or 50th sample. The results obtained by KPAL, for the NMBAQC replicates samples prepared by APEM since 2014/15, demonstrate that the high degree of repeatability which can be obtained when strict analysis protocols are followed, and that a high degree of confidence can be placed in the results obtained for any individual analysis.

3. Conclusions and Recommendations

A number of observations may be made based on the results of the exercises described above. The following is a summary of the major points of importance.

1. Laboratories should ensure that their PS results are reported in the requested format. Data should be provided at half-phi intervals to enable the direct comparison of data from all participants and simplify the creation of cumulative curve figures. The workbook was modified for use in 2014/15 to assess whether laboratories are merging data correctly in their in-house methods. It is therefore even more important that that data are reported correctly. Raw sieve data should be reported in grams, with the > 1mm and < 1mm weights provided. Raw laser data should be

provided re-proportioned to 100% and reported as volume percentages. Final merged data should ideally be reported in percentage of final weight.

2. Participants should review their data prior to submission. Errors in datasets can often be spotted in the summary statistics, e.g. percentage gravel, sand and silt/clay, before the data are submitted. All parts of the workbook should be double checked before submission to ensure that it is all filled in correctly. This will help eradicate typing and transcription errors. The 2015/16 Scheme year revealed inconsistencies on how results were reported in terms of decimal places; it is requested that for 2016/17 all results be reported to at least 2 decimal places.

3. Particle size (PS) exercises over the past twenty years have shown differences in the results obtained by different techniques (laser and sieve / pipette), in-house methods (*e.g.* pre-treatment) and also differences between equipment (*e.g.* Malvern Mastersizer 2000, Mastersizer X and Coulter LS230 / LS13320 lasers). The PS data also indicate that the variance between laser and sieve results is further emphasised by certain sediment characteristics, notably particle shape and density (Blott and Pye, 2006; Blott *et al*., 2004). The overall range of these variances needs to be determined if combining data sets derived from different methods. The NMBAQC's Best Practice Guide was developed for use in 2010/11 (Scheme year 17); this has helped to reduce the amount of variation between methods. Sieve and laser metadata information sheets were added to workbook for 2014/15 to give more detailed information on methods used, particularly for laser analysis. It is essential that particle size data are presented with a clear description of the method of analysis and equipment used, including nature of any ultrasonic or other dispersion process, and the optical model values which have been assumed.

4. The current NMBAQC Scheme Pass/Fail criteria for the PS module are under review. For 2015/16 alternative flagging criteria using z-scores on descriptive statistics combined with robust statistics were used following a review of this method on data from 2014/15. However, this year's results have shown that even with robust statistics z-scores are not appropriate for creating "Pass" or "Fail" flags as variability in results can lead to participants receiving false "Pass" results. For 2016/17 (Scheme year 23) reports will follow a similar format to that of PS-OS reports with each section broken down for review, for example sieve processing, laser processing, data merging and summary statistics. Laboratories will then receive a "Good" or "Review" flag

based on their results; "Review" flags will come with accompanying comments as to where mistakes have been made and how to correct them.

5. <u>The 2015/16 PS-OS module highlighted differences between the sensitivity of laser instruments and affects of dispersants.</u> Comparison of laser data in the PS-OS results showed that the Beckman-Coulter LS13320 instrument used by the AQC lab, which includes a Polarization Intensity Differential Scattering (PIDS) which gives enhanced measurement capability in the size range 0.4 and 0.04 microns, indicates a higher clay content compared to other lasers models used by many of the NMBAQC scheme participants. It is therefore even more important that participants provide metadata regarding the laser model and optical model used, and about the dispersion methods, whether or not ultrasonics were used before or after the run in addition to the possible use of chemical dispersant . Although laser models will not be directly linked to participants, in order to keep participant confidentiality, the range of laser models used will be specified in future reports. As well as this, the possibility of developing conversion factors between laser sizers will be explored when enough data have been collected.

6. <u>The 2015/16 PS-OS module highlighted that participants do not always supply the samples in the requested format</u>, i.e. dried > 1mm fraction, dried < 1mm fraction and a laser subsample taken from the bulk sample. The need has been identified to update the Guidance with more detailed advice on how to store samples; these amendments are included in the guidance and can be viewed in Mason (2016).

4. References

Blott, S.J. and Pye, K., 2001 GRADISTAT: a grain size distribution and statistics package for the analysis of unconsolidated sediments. *Earth Surface Processes and Landforms* 26, 1237-1248.

Blott, S.J. & Pye, K. 2006 Particle size distribution analysis of sand-sized particles by laser diffraction: an experimental investigation of instrument sensitivity and the effects of particle shape. *Sedimentology* 53, 671-685.

Blott, S.J. & Pye, K. 2012 Particle size scales and classification of sediment types based on size distributions: review and recommended procedures. *Sedimentology* 59, 2071-2096.

Blott, S.J., Croft, D.J., Pye, K., Saye, S.E. & Wilson, H.E. 2004 Particle size analysis by laser diffraction. In Pye, K. & Croft, D.J. (eds.) *Forensic Geoscience - Principles, Techniques and Applications*. Geological Society, London, Special Publications 232, 63-73.

Blott, S.J., Croft, D.J., Pye, K., Saye, S.E. & Wilson, H.E. 2004 Particle size analysis by laser diffraction. In Pye, K. & Croft, D.J. (eds.) *Forensic Geoscience - Principles, Techniques and Applications*. Geological Society, London, Special Publications 232, 63-73.

Finbow, L. & Hall, D., 2015. National Marine Biological Analytical Quality Control Scheme. Particle Size Results: PS56. Report to the NMBAQC Scheme participants. Apem Report NMBAQCps57, 41pp, August 2015

Finbow, L. & Hall, D., 2015. National Marine Biological Analytical Quality Control Scheme. Particle Size Results: PS57. Report to the NMBAQC Scheme participants. Apem Report NMBAQCps57, 41pp, August 2015

Finbow, L. & Hall, D., 2016. National Marine Biological Analytical Quality Control Scheme. Particle Size Results: PS58. Report to the NMBAQC Scheme participants. Apem Report NMBAQCps58, 38pp, January 2016

Finbow, L. & Hall, D., 2016. National Marine Biological Analytical Quality Control Scheme. Particle Size Results: PS59. Report to the NMBAQC Scheme participants. Apem Report NMBAQCps59, 38pp, January 2016

Folk, R.L., 1954. The distinction between grain size and mineral composition in sedimentary-rock nomenclature. Journal of Geology 62, 344-359.

Hall, D.J. 2010 *National Marine Biological Analytical Quality Control Scheme. Description of Scheme Standards for the Particle Size Analysis Component from Scheme Year 8 (2001/02) to Year 16 (2009/10).* Report to the NMBAQC Scheme participants. Unicomarine report NMBAQCpsa_stds, February 2010.

Mason, C. 2015. *NMBAQC's Best Practice Guidance. Particle Size Analysis (PSA) for Supporting Biological Analysis.* National Marine Biological AQC Coordinating Committee, 72pp, January 2015.

Unicomarine. 1995 *National Marine Biological Quality Control Scheme. Annual Report (Year one).* Report to the NMBAQC Committee and Scheme participants. September 1995.

Unicomarine. 1996 *National Marine Biological Quality Control Scheme. Annual Report (Year two).* Report to the NMBAQC Committee and Scheme participants. September 1996.

Unicomarine. 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.