The National Marine Biological Analytical Quality Control Scheme
www.nmbaqcs.org

Macroalgae Biomass Component Report – OMB RT06 2015

Emma Wells
Wells Marine Surveys
April 2015
Email: emma@wellsmarine.org

Contents

1 Introduction

To enable correct water quality classification and good management decision-making, quality control of biological data is a high priority. This extends through all biological elements including macroalgae. Good quality control ensures consistency of data being reported for management purposes, and for macroalgae this has been driven primarily by the requirements of the Water Framework Directive. The Healthy, Biologically Diverse Seas Evidence Group (HBDSEG), part of UK Marine Monitoring and Assessment Strategy, sets the key areas for UK agencies in which this external quality control is particularly needed. For 2015/2016 the components are: Benthic invertebrates, Fish, Particle Size Analysis, Macroalgae/Seagrass, Phytoplankton, Epibiota and Zooplankton. This QC scheme aims to facilitate improvements in biological assessment whilst maintaining the standard of marine biological data. The scheme should help to ensure consistency between analysts with improved confidence in ecological quality status.

The National Marine Biological Analytical Quality Control (NMBAQC) Scheme addresses several issues relating to macroalgae data collection, this report focuses on just one of these:

- The determination of algal biomass

This is the sixth year in which biomass of macroalgae has been included as an element of the NMBAQC scheme and was included as a single exercise. The format followed that of previous years of the test (OMB RT01 – RT05 - see NMBAQC website). Test material was distributed to participating laboratories from which data forms were completed with algal biomass results and returned for analysis.

Nine laboratories were issued with test material. All nine laboratories completed the macroalgae biomass component of the NMBAQC scheme with a single laboratory submitting two sets of results. All of the participating laboratories were government; no private consultancy took part in this component of the macroalgae exercises. To ensure consistency between scheme years, each participating laboratory was assigned the same laboratory code as in previous years except where a laboratory was new to the scheme.

Due to each sample being unsuitable for processing more than once, only a single set of results could be submitted per laboratory unless more than one test was requested. It was possible for each sample to be completed by a different analyst; however, this was not identified within the final results. Individual laboratories may look at such results internally.

Data for macroalgal blooming and seagrass are currently used in relation to WFD classification and assessments for other Directives/purposes. In the UK at present they are not reported through a national database such as Merman; consequently they do not have definite national qualifying performance levels. They may be treated as training exercises. However, certain indicative targets have been applied to the assessments of the results based on Z-scores allowing “Pass/Fail” flags to be assigned as appropriate. Ring tests offer a means of assessing personal and laboratory performance from which continued training requirements may be identified, or from which improvements in current field and laboratory procedures may be addressed.

Samples are synthetic, rather than composed of natural algal material. Natural samples would be subject to deterioration, and it is not feasible to ensure that each participant would receive a truly equivalent sample. This is in line with guidance on general requirements for proficiency testing (BS EN ISO/IEC 17043:2010).

2 Summary of Performance

A single test consisting of three biomass samples was distributed. Each sample consisted of a synthetic mix of dish cloths and wool, which are considered to imitate opportunist macroalgae species. Cloths were cut to different sizes to represent different taxa (e.g. laminar or tubular taxa). Each sample was contaminated with debris and sediment of a sandy-muddy nature consistent with the substrate type known to support opportunist macroalgal blooms.

Results for wet weight of biomass varied between laboratories with some laboratories producing high measures of biomass compared against the average biomass. The dry weights showed a similar level of variability. Two laboratories failed to remain within the Z-score of less than +/- 2.0 for the average sample wet weight and also 'Failed' at least one sample for the wet weight against expected weight. Two further laboratories showed significant deviation from the average sample dry weight and actual dry weight. A few of the dry weights were substantially higher than the actual dry weight and these dry weights skewed the results making the analysis unable to pick up smaller deviations. Most participating laboratory results were higher than the actual sample dry weight suggesting no loss of sample material during processing, but incomplete drying.

3 Summary of Exercise

3.1 Introduction

There was one exercise for the assessment of biomass of macroalgae which took the form of three representative artificial samples. This exercise is described in full below to include details of distribution and logistics, procedures for determination of biomass, completion of test result forms and full analysis and comparison of final submitted results. While it is recognised that synthetic samples are not true representations of natural algal samples, the point of the exercise is to test participants' ability to process samples using the same procedures as they would for algal samples.

3.2 Description

This exercise examined the participants' ability to process macroalgae samples to extract values for biomass for wet and dry weight. The exercise examines differences in sample processing efficiency and comparability of results using Z-scores. Comparison of participating laboratory results can highlight anomalies in processing at various stages of the methodology.

One set of three samples of different weights was distributed to each participating laboratory in January 2015. Participating laboratories were required to submit biomass results for both wet and dry weight. The sample material was consistent with that of OMB RT05 including a greater proportion of wool to further assist with the more accurate imitation of actual macroalgae samples, and added non-biological and non-algal biological material to simulate contaminating materials encountered in the field.

3.3 Logistics

Each sample was distributed within an airtight plastic container. Each sample within the container was separately sealed within a zip lock plastic bag to retain moisture. The samples were distributed either via first class mail or recorded delivery, depending upon the recipient's requirements. All instructions and relevant data sheets were distributed on CD, within the parcel, to each laboratory. Each disc contained a description of methods and data submission forms along with a feedback form. Participants were given six weeks to complete the test and return the results. Only one set of results

could be submitted per set of samples although it was possible to have up to three participants complete the sample analysis.

Email has been the primary means of communication for all participating laboratories subsequent to the initial postal distribution of test material.

3.4 Preparation of the Samples

In order to assess the accuracy of determining biomass of opportunistic macroalgae, samples were distributed consisting of both dish cloth and wool material that had been cut and finely shredded in order to mimic species of *Ulva* (previously known as *Enteromorpha*). The alternative materials were deemed to be the most representative of actual opportunist species and were based on suggestions from previous ring test feedback. Three samples representative of actual weights found in surveys were supplied for subsequent processing. Sediment and debris commonly found within areas of opportunist algal growth were mixed into the samples with small amounts of water. For each sample, wet weight and dry weight had to be ascertained.

The sample were labelled from A to C. Samples of identical original dry weight were provided for all participants.

Sample A – 74.2g

Sample B – 53.7g

Sample C – 18.0g

Due to the nature of the samples they could be kept for several days retaining most of the moisture. However, much of the water was removed prior to distribution to reduce weight during transportation; therefore it was necessary for participants to add additional water to each of the samples prior to commencement of the tests to enable rehydration of the material and aid with rinsing.

3.4.1 Method for Wet Weight

The laboratory instructions stipulated that each of the samples required rinsing free of all sediment. The samples should be fully washed in a bucket or sieve to ensure no loss of sample until the water runs clear and all debris is removed. Once the samples are adequately washed they are squeezed of excess water. This is achieved by hand, using samples no larger than the size of a tennis ball to ensure it fits in the palm of the hand and can still be squeezed properly. Where the sample was large, it should be divided into smaller clumps for squeezing. The samples are squeezed until no additional running water could be removed by hand, but the sample should not run green, as this indicates damage to cell membranes (over-enthusiastic squeezing of actual algal samples can damage cell membranes and lose 'genuine' weight). At this stage the whole sample is weighed on a calibrated balance to two decimal places. The exact method used for rinsing and squeezing should be consistent with that used in the field; this may vary between laboratories.

3.4.2 Method for Dry Weight

Once each of the samples has been wet weighed they are spread out on a sorting tray or similar container. By spreading the samples this aids with the drying process. The samples are left to air dry for at least 24 hours, but this may be longer depending on the size of the sample and the temperature of room. The samples should be checked regularly and the drying/weighing process is continued until

constant mass is achieved, recording weight to 2 decimal places. The unchanged dry weight is the final weight to be submitted.

The same process is required for all 3 samples.

3.5 Data Submissions

A pre-prepared spreadsheet was distributed along with the exercise instructions to standardise the format in which the results were submitted. These results will be retained and stored appropriately. Each Laboratory was required to submit a dry weight and a wet weight for each of the 3 samples provided. Laboratories were permitted six weeks to complete the sample analysis and submit results.

3.6 Analysis

The macroalgae biomass exercise provides 'Pass' and 'Fail' flags to each data set to highlight deviation from sample mean and image results. Values of Z-scores were used to apply the 'Pass' and 'Fail' assessment.

Z-scores, calculated to indicate the level of deviation of biomass, used the following formula:

$$Z = \frac{X - \mu}{\delta}$$

X is a raw score to be standardized;

μ is the mean of the population;

σ is the standard deviation of the population.

Z-scores were calculated using the mean biomass and the actual biomass. A Z-score value of greater than +/- 2.0 was considered to be outside an acceptable limit of deviation from the mean. This value is considered standard practice and was used assign a 'Fail' or 'Pass' flag on the data.

3.7 Confidentiality

To preserve the confidentiality of participating laboratories, each participant is allocated a four digit laboratory code from which they can identify their results. These codes are randomly assigned. The initial letters (MA) refer to the scheme, this is followed by the scheme year which refers to the year in which the NMBAQC scheme originally commenced, and the final two digits represent the laboratory. For example, laboratory twelve in scheme year twenty two will be recorded as MA2212.

4 Results

4.1 General Comments

In total nine laboratories signed up for the biomass component of the macroalgae element for OMB RT06. Nine laboratories returned both wet and dry weight data with one lab submitting two sets of results, giving 10 in total. The results have been collated and presented in various formats to enable full comparisons both between laboratories and against actual sample weights.

Details of each participating laboratory performance were distributed in the macroalgae OMB RT06 Bulletin Report and the seagrass OMC RT06 Bulletin Report, which represent a summary of the results for RT06.

4.2 *Returns from Participating Laboratories*

The raw data (Table 1) indicates a wide range of both wet and dry weights. The range of results was greatest for the algae mass of the largest weight from both dry and wet weights. This is consistent with all previous OMB tests. For wet weight the range of results was 270.29 – 405.15 (Sample A), 177.81 – 311.16 (Sample B) and 56.97 – 81.11 (Sample C). This clearly indicates a degree of variation in data and lack of consistency between laboratories during the rinsing and squeezing of the samples particularly within the larger sample sizes (Samples A and B). The large degree of variation in wet weight results are primarily a result of the non-specific method of squeezing and rinsing as this is an element of the exercise that cannot be measured successfully and can vary significantly between participants. This is particularly evident with the larger sample sizes where there is a greater chance of error.

Table 1. Raw Data results from each laboratory including both dry and wet weights.

	Sample A		Sample B		Sample C	
Lab Code	Wet weight	Dry Weight 74.2g	Wet weight	Dry Weight 53.7g	Wet weight	Dry Weight 18g
MA2210	405.15	76.481	311.16	55.41	81.11	18.56
MA2203	365.8	74.9	243.7	54.2	77.5	18.6
MA2202	270.29	95.35	177.81	55.98	56.97	19.25
MA2209	309.2	156.2	215.5	65	76.7	21.9
MA2211	310.98	95.17	223.67	61.94	73.61	20.08
MA2233a	321.69	75.45	214.32	56.64	72.4	18.83
MA2233b	363.58	89.12	246.61	56.78	79.64	19.34
MA2234	355.9	78.8	239	54.1	69.4	18.4
MA2232	277.06	78.22	212.9	56.67	61.91	19.28
Max	405.15	156.2	311.16	65	81.11	21.9
Min	270.29	74.9	177.81	54.1	56.97	18.4
Range	134.86	81.3	133.35	10.9	24.14	3.5
Average	337.82	92.68	233.97	57.51	73.42	19.37

The level of variation in dry weight was also consistent with previous years. The dry weights results displayed a couple of large outliers, laboratory MA2209 submitted results considerably higher than the sample mean and actual dry weight for sample A and laboratories MA2202 and MA2211 also submitted results from sample A that could be considered significantly higher than the mean and actual dry weight, causing a slight skew in the overall results and a slightly higher mean and standard deviation than would be considered acceptable. The results from these laboratories indicate some problems during the processing of the samples. This may be due to procedures used, inadequate rinsing or incomplete drying. The average levels of wet weights suggests that possibly the samples were not dried fully prior to weighing. However, in contrast, sample C was much more comparable in terms of wet weight and dry weight, albeit still a little high, indicating the correct procedures were being used, if only in part. Sample C was the smallest sample size and as seen in previous tests had the most consistent set of results between laboratories.

The range of results for both the dry and wet weights (as seen in Bulletin OMB RT06) could generally be considered acceptable with only a couple of 'Fails' providing evidence of a good degree of consistency in practiced methods. However as with previous years it is evident that the level of error in the results submitted is related to the actual sample size provided. This is easier to see when comparing the % increase in sample size of dry weight against the actual weight. For sample A (74.2g) the average % increase is 25%, for sample B (53.7g) 7.1% increase and for sample C (18g) 7.6%. These

differences in % increase in sample size have also been seen in previous exercises providing evidence that the level of error is directly related to the size of the sample.

In total two results were flagged as 'Fail', when using Z-scores based on sample mean of wet weights. These were for Laboratories MA2210, with a Z-score of 2.02 for sample B and lab MA2202 had a Z-score of -2.1 for sample C, both falling just outside of the cut-off value. Two additional 'Fails' were flagged against Laboratories MA2209 for the comparison of dry weight against the sample mean with a Z-score of 2.347 (sample A) and 2.183 (sample C). In general a Z-score of <2 is considered satisfactory, one >2 and <3 indicates "questionable" performance and generates a warning signal and a Z-score of >3 indicates "unsatisfactory" performance (ISO/IEC 17043:2010). This means that no results were actually "unsatisfactory" but performance should be investigated.

A second Z-score was calculated based on deviation from the actual known dry weight using the same criteria to flag 'Pass' and 'Fail'. This resulted in a total of four 'Fails'. The largest anomalies were submitted by lab MA2209 with a z-score of 3.029 (sample A), 2.899 (sample B) and 3.365 (sample C). The actual dry weight of sample A from this lab was almost double the actual weight with a result of 156.2g dry weight compared with 74.2g actual weight. The high level of deviation from actual dry weight value as submitted by this Laboratory produced a higher standard deviation (27.07) than expected for the population mean of sample A and has prevented any smaller deviation from the actual weight of sample A becoming evident in this analysis. Lab MA2211 also 'Failed' sample B with a z-score of 2.114.

All results were higher than the original sample weight. This suggests no loss of material during the washing, squeezing and weighing process but may mean that the samples are not dried properly of there is residual debris within the sample.

The differences in sample processes have become evident over the years through the degree of variation in the results submitted. There are often a number of outliers which significantly skew the results and affect the average weight which is used to compare all other results. If this average is abnormally high or low it will affect the outcome of some laboratories results which might otherwise be considered acceptable. This can be problematic for those laboratories whose results were lower than average and based on the calculations have 'Failed' to fit within the z-score or +/- 2.0. With this in mind the biomass data sets from all 6 ring tests (including the present) have been collated and compared.

The following graph (Figure 1) illustrates the correlation between the average wet weight and the actual weight with a linear trend line indicting most points lie on or close to the line. The associated equation that refers to this trend line was used to calculate the expected wet weight from the known actual dry weight. Table 1 below details the expected wet weight for this years' ring test samples from which the submitted results could be compare and additional z-scores calculated. This is considered a more accurate comparison of the data and is not influenced by large outliers.

Table 2: Expected wet weights for samples A, B and C based on the linear trend line equation y=3.749x + 4.9964.

Dry Weight	Expected Wet Weight
74.2	283.1722
53.7	206.3177
18	72.4784

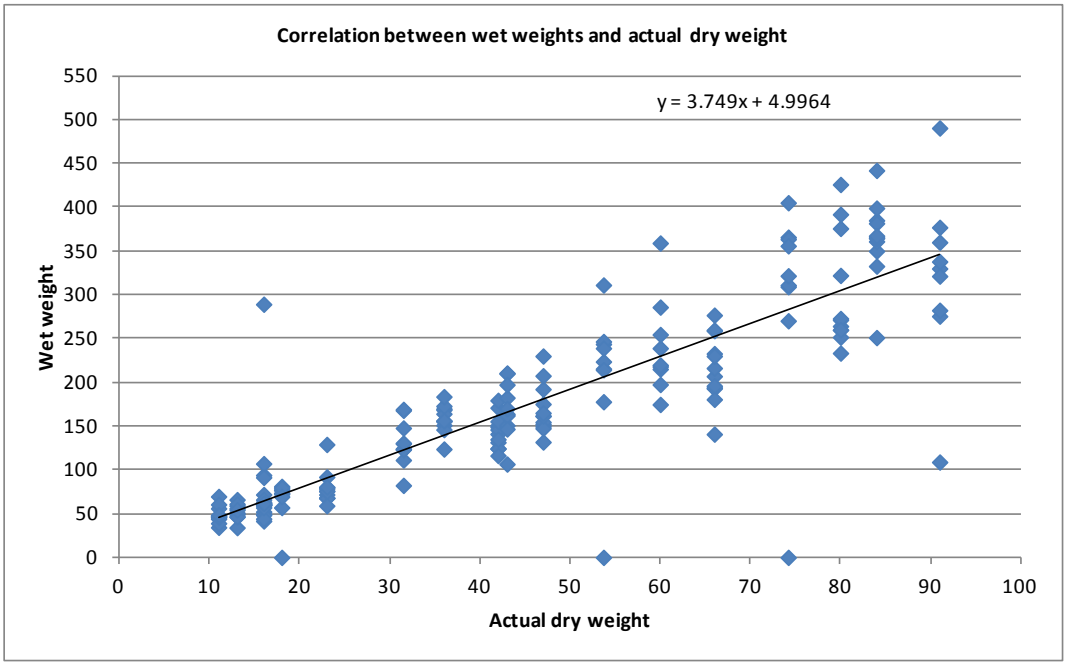

Figure 1: correlation between wet and dry weights including linear trend line and corresponding equation with R squared = 0.81873 and significance = 1.60438E-70

The results from this comparison were slightly different to those z-scores where wet weight is compared against the average. Where as previously there were only two 'Fails' against wet weight this method has resulted in three 'Fails'. Lab MA2210 'Failed' sample A (2.868) and sample B (2.746) and lab MA2202 marginally 'Failed' sample C (-2.019).

5 Discussion

Of the ten samples distributed to nine laboratories all submitted results. Although many of these laboratories do not routinely measure dry mass for macroalgae, this is still a necessary part of this exercise as it enables the procedure to be reviewed for inter-laboratory differences. If samples are dried to a level where the mass remains unchanged then a result that lies well above the actual dry weight is a clear indication that the sample has been insufficiently rinsed and it is the additional particles that are adding to this increased weight. This will contribute to both an overestimation of wet and dry weights. Seaweed is much harder to rinse especially in the field so may contribute to an overestimation of the levels of biomass present. Equally some laboratories do not measure wet weight only recording the final dry weight. Dry weight could be considered a much more accurate measure of biomass since this measure has fewer variables, i.e. it is only dependent upon the removal of debris and not the degree of pressure during squeezing. However, both measurements need to be incorporated into the test to cover all the different measurements and procedures utilised. Also the WFD methodology uses wet weight for the biomass metric.

Unlike in previous years the level of accuracy was similar between dry and wet weight. It is often the case that wet weight comparisons hold a greater degree of variability than dry weight with there being more variables involved. Therefore the similar number of 'Fails' across both dry and wet weight suggests either a greater level of proficiency and consistency when squeezing and weight wet samples or that the dried samples have been insufficiently dried leading to much higher results than expected. This was also across all three sample weights and did not seem to be as influenced by overall sample

size. However, there were still a number of outliers producing higher than expected results for both wet and dry weights. This suggests the techniques used between laboratories to rinse, squeeze and dry vary considerably and may also do so between participants within the same laboratory. The lack of consistency in wet weight indicates a high level of variation in pressure applied during squeezing of samples. However, this is difficult to regulate between analysts. It is the wet weight that is most commonly used during routine opportunist monitoring, therefore this lack of consistency in methodology should be fully addressed within the standard operating procedures especially in association with areas of high biomass. Each lab should have its own in-house training and competence assessment measures. Equally greater care should be taken to ensure the samples are properly dried prior to final weighing. It would be much more difficult to control quality during field processing rather than in the laboratory.

Most laboratories produced a dry weight greater than that of the actual biomass of the sample; this could be due to insufficient drying or rinsing of the sample a degree of which can be expected during such a test. The current ring test experiences no dry weights less than that of the actual biomass which suggests there is no loss of material during the rinsing process.

There was an obvious trend whereby the level of deviation from actual biomass increased as the sample biomass increased. There is no apparent reason for this, the larger biomass may be more difficult to rinse free of debris or possibly it is more difficult to squeeze or dry thoroughly. This is equally something that should be addressed within individual laboratories as well as across standard operating procedures to reduce this level of error. Laboratories may wish to check internal samples for this pattern.

In general the results were comparable with those from previous years. The ring test is able to provide evidence of problems in the measuring of biomass samples, such issues require addressing through workshops and specifically aimed training. Hopefully on receipt of the results bulletin those laboratories with outliers will also be able to review the procedures adopted during the processing of their samples.

It should be further highlighted that the 'Fails' do not necessarily signify poor quality data they merely flag those results which show significant deviation from either the actual sample weights or from the average, and should be investigated. These flags have no current bearing on the acceptability of data from such participating laboratories. These exercises cannot comment on what effect inconsistencies in assessing wet and dry weights may have on the calculation of the WFD biomass metric and classification of sites: this should be considered by the participating agencies.

6 Conclusions and Recommendations

A number of observations may be made from the results of the exercise and from participants' feedback which have been summarised below:

1. Most samples arrived in good condition, only one laboratory reported some leakage. Greater care shall be taken in the future to ensure any leakage is kept to a minimum by ensuring the individual bags are well sealed along with the plastic containers.

2. Apart from some extensive drying times the tests were considered quick and easy.

3. It seems there is now a general agreement that the use of artificial material to mimic algae is an acceptable surrogate for the test albeit less fragile and easier to rinse and squeeze than the real thing. There have been alternative material suggestions in previous years of using muslin, as well as the wool and dish-cloth, which was found to be insufficient in representing the algae and was

difficult to process. It was noted by some labs that there was a limited representation of small, finer, low biomass algae, such as *Cladophora*, which require more careful extraction. Further recommendations from this years' ring test includes the use of soft fillings such as those used in animal beds or soft toys which could be used to mimic such species and would represent the more tangled nature of such species.

It is appreciated that the use of wool and dish-cloths do not fully represent the conditions experienced within the field. However it may be possible in subsequent tests to incorporate alternative materials that may be more representative of the texture and general nature of opportunist algae. The use of soft filling to represent finer species of opportunist algae shall be considered for possible use in OMB RT07.

4. During this sixth cycle of the macroalgae biomass exercise all participating laboratories submitted results within the designated timescale. All laboratories should continue to submit results within the requested deadlines as detailed at the beginning of the exercise. Reminders will continue to be distributed two weeks prior to the completion of the exercise to aid with this process.

5. This year all laboratories submitting results managed to complete both wet and dry weights for all samples, however some participants still question the necessity to incorporate both dry and weights within the ring test. Although many in-house field procedures do not incorporate dry weight of algal samples these values are included within the NMBAQC scheme to enable comparison of laboratory procedures. The values provide evidence of insufficient rinsing of samples, whereby the dry weight would be considerably higher than the actual dry weight. Also there is no definitive wet weight from which to compare the individual laboratories submissions so it is difficult to conclude which results are the most representative. The dry weight however can be compared directly with the original weight of the samples which was measured very accurately prior to addition of debris. The dry weights are also now being used to calculate an expected wet weight from which to compare results.

   Most laboratories submitted dry weight values that were considered well within an acceptable limit of the actual biomass; however wet weight still remains highly variable. Therefore the level of squeezing still remains an issue within the overall procedure and should be addressed. In addition, some laboratories only measure the dry weight therefore, for such an exercise to be appropriate for such laboratories; this measure of biomass needs to remain within the test. It is in all laboratories' own interest to complete all aspects of the test. Submission of partial results may hinder any explanation of outliers and skew statistics due to the relatively small data sets. During subsequent ring tests, all laboratories should continue to complete the full exercise even if it is not part of their routine monitoring in order to maximise the usefulness of the ring tests.

6. It was suggested that the mud added to the sample, to enable a more realistic comparison with field procedures, should include a variety of debris. There was a suggestion that shells and *Hydrobia* could be added to the sample as well as thicker and more gloopy mud to reduce the ease with which the samples can currently be rinsed and since this type of substrate is more consistent with that found in the field. It has also been commented that the artificial material is easier to rinse free of sand than mud, silt or clay. This will be taken into account for subsequent tests.

7. Many labs not only weigh the complete biomass sample but also separate species for both identification and separate weighing to enable relative percentages to be ascertained. This process of separating species has not been incorporated into the biomass test to date but given this request it will be discussed for subsequent years whereby different materials could represent

different species. This will require additional work on part of the participating labs and may only be a requirement of some labs therefore this will be discussed in full prior to future tests.

8. It is evident that the larger samples create a greater margin of error with far less consistency between laboratories. However, it has been suggested that these samples are more appropriate in terms of representing natural conditions. This will be taken on board when compiling future tests whereby they will be aimed at including a good range of weights but focusing on some much larger biomass weights.

9. There may be future requirements to include biomass analysis within a workshop to further discuss processing procedures and levels of intensity for manual removal of debris and water. This has been suggested by some participating laboratories and may be considered a more realistic measure of quality assurance. This is something that requires further discussion as to the nature of the approach.

10. A number of laboratories submitted results to a lesser degree of accuracy than others. It is stipulated that both wet and dry weights be provided to 2 decimal places where possible. This will highlight smaller variations in weight as the samples are relatively small compared with some field samples. However if this is not feasible for some laboratories then measurements to the nearest gram are also acceptable but it needs to be recognised by participating laboratories that such measurements will be less accurate particularly with smaller sample sizes. In the instance where the dry weight recorded is less than the actual weight this may be an indication of loss of material but may also be linked to the accuracy of the scales. It is recommended that all laboratories use calibrated scales so as to reduce such minor discrepancies.

11. All laboratories must use the result spreadsheets provided and include *all* the required information. Data presented in Word files or within emails is very inconvenient when collating and storing the results and will not be accepted. If this does occur a request will be sent for the data to be completed in the correct format. .

If anyone has further thoughts on this, or disagrees with any of the interpretation, please pass forward your comments to Dr Emma Wells (emma@wellsmarine.org) and copy to Dr Clare Scanlan (clare.scanlan@sepa.org.uk). This ring test is now in its fifth year and although proving successful it is still open to continual refinement.