The NE Atlantic Marine Biological Analytical Quality Control Scheme

www.nmbaqcs.org

Macroalgae Biomass

Component Report

Ring Test OMB RT13 2022

Søren Pears
Georgina Brackenreed-Johnston
D Hall

APEM Ltd.

22nd April, 2022

E-mail: nmbaqc@apemltd.co.uk

Contents

List of Tables

1. Introduction

1.1 Background

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 and seagrass. Good quality control ensures consistency of data being reported for management purposes, and for macroalgae and marine angiosperms this has been primarily driven by the requirements of the Water Framework Directive. This QC scheme aims to facilitate improvements in biological assessment whilst maintaining the standard of marine biological data. The scheme aims to ensure consistency between analysts with improved confidence in ecological quality status.

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

- The determination of algal biomass

This is the thirteenth 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 - RT12). Test material was distributed to participating laboratories along with data forms, which were completed with algal biomass results and returned for analysis.

Graphical representations of the performance of each participating laboratory were distributed in the OMB RT13 Bulletin Report. This bulletin included the z-score based ‘pass’ and ‘fail’ flags assigned to each result to highlight deviation from sample means and actual/expected weights. The current report describes the results in more detail and should be read in conjunction with the OMB RT13 Bulletin.

1.2 Participating Laboratories

Nine laboratories were issued test material, of which seven laboratories completed the macroalgae biomass component of the NMBAQC scheme. Of those participating, all seven laboratories were government organisations.

Due to the limited number of samples distributed, only a single set of results was permitted per laboratory unless more than one test was requested. It was possible for each sample to be completed by a different participant; however, this was not recorded within the results. Individual laboratories may look at such results internally.

Currently this scheme does not specify a definite qualifying performance level, and NMBAQC ring tests may be treated as training exercises. However, certain targets have been applied to the assessment of the results based on z-scores allowing ‘pass’ or ‘fail’ flags to be assigned; these may be used by competent monitoring authorities for internal monitoring of performance. These flags have no current bearing on the acceptability of data from such participating laboratories. 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 were 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 the Biomass exercise

1.3 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.

1.4 Description

This exercise examines the participants' ability to process macroalgae samples to extract values of biomass for wet and dry weight. The exercise assesses the 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 representative samples was distributed to each participating laboratory in January 2022. Participating laboratories were required to submit biomass results for both wet and dry weight. The sample material was consistent with that of OMB RT12 including j-cloths, wool and synthetic stuffing. Based on previous ring tests, these were chosen to be the most representative materials in terms of imitating the overall look and feel of various opportunistic macroalgae species. Cloths and wool were cut to different sizes and lengths to represent different foliose and filiform taxa (e.g. *Ulva* spp.). The synthetic stuffing represents finer algae such as *Chaetomorpha* spp. Each sample was mixed with sediment of a sandy-muddy nature consistent with the substrate type known to support opportunistic macroalgal blooms to simulate substrates that would be encountered in the field.

1.5 Logistics

Each sample was contained within a plastic sample bucket and distributed via a courier delivery service company. All instructions and additional test forms were distributed via e-mail attachments to each laboratory. The files contained a description of methodology and standardised forms for data submission. 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 test analysis.

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

1.6 Confidentiality

To preserve the confidentiality of participating laboratories, each participant was randomly allocated a four-digit laboratory code, which allowed them to identify their own results. The initial letters (MA) refer to the scheme, this is followed by two digits representing the current NMBAQC scheme year, and the final two digits representing the laboratory. However, it appears the macroalgae component is out of synchrony with the rest of the NMBAQC scheme components, since the current scheme year is twenty-eight, but the '28' prefix has already been used for macroalgae RT12. For the sake of continuity with the previous macroalgae ring tests, a '29' prefix has been used this year for RT13, but this will be corrected in the 2022-23 scheme year to ensure standardisation across all scheme components.

1.7 Preparation of the Samples

To assess the accuracy of determining biomass of opportunistic macroalgae, samples were prepared using j-cloth, wool and synthetic stuffing material that had been cut and shredded to mimic species of *Ulva*. These materials were deemed to be the most representative of actual opportunistic species and were based on suggestions from previous ring test feedback forms. Three representative samples were supplied for subsequent processing. Sediments commonly found within areas of opportunistic 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 samples were labelled A, B and C with samples of identical original dry weight provided to all participants.

Sample A – 11.4g

Sample B – 32.5g

Sample C – 75.1g

Due to the nature of the samples, they could be kept for several days retaining most of the moisture. However, only enough water was added to thoroughly soak the synthetic materials and liquify the sediments prior to distribution to reduce weight during transportation. It was therefore 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.

1.7.1 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 material 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 can be removed by hand (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.

1.7.2 Method for dry weight

Once each of the samples has been wet weighed, they are spread out on a sorting tray or similar container. Spreading the samples in this way 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 was required for all 3 samples.

1.8 Analysis and data submissions

A pre-prepared spreadsheet was distributed 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 both a dry weight and wet weight for each of the 3 samples provided within the allocated six-week time period.

1.9 Z-Scores

Values of z-scores were used to apply the ‘pass’ & ‘fail’ assessment.

Z-scores were calculated to determine how many standard deviations each participant’s weight results deviated from the mean, using the following formula:

$$Z = \frac{x - \mu}{\sigma}$$

Where:

$\boldsymbol{x}$ is the raw weight value to be standardised;

$\boldsymbol{\mu}$ is the mean of the participants’ weight values;

$\boldsymbol{\sigma}$ is the standard deviation of the participants’ weight values.

A z-score of greater than +/- 2.00 was considered to be outside an acceptable limit of deviation from the mean and this cut-off point was used determine ‘Fail’ or ‘Pass’ flag on the submitted data.

2. Results

2.1 Returns from participating laboratories

Of the seven laboratories that returned results for OMB RT13, six returned both wet and dry weight data and the seventh returned wet weight data only. 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's performance were distributed in the OMB RT13 Preliminary Bulletin Report. The Bulletin provided z-score derived 'pass' and 'fail' flags to each result set to highlight deviation from sample mean and actual/expected results.

Table 1 presents the range of wet and dry weights recorded by participating laboratories. Sample B had the largest range of results for wet weight (196.31g) and Sample C had the highest range of dry weights (3.89g).

Table 1 Raw data results from each laboratory including both dry and wet weights (where recorded)

Lab Code	Sample A		Sample B		Sample C	
	Wet Weight (45.07g)	Dry Weight (11.4g)	Wet Weight (127.26g)	Dry Weight (32.5g)	Wet Weight (293.17g)	Dry Weight (75.1g)
MA2901	72.27	11.64	289.78	34.07	219.09	77.87
MA2904	75	11	122	35	262.3	81
MA2905	42.43	11.53	93.47	33.68	238.32	79.8
MA2906	52.67	11.76	138.17	36.07	220.34	78.01
MA2907	52	-*	159	-*	291	-*
MA2908	42.36	11.46	145.39	32.77	210.45	77.11
MA2909	49.94	11.46	153.91	32.51	277.22	78.07
Max	**75**	**11.76**	**289.78**	**36.07**	**291**	**81**
Min	**42.36**	**11**	**93.47**	**32.51**	**210.45**	**77.11**
Range	**32.64**	**0.76**	**196.31**	**3.56**	**80.55**	**3.89**
Average	**55.24**	**11.48**	**157.39**	**34.02**	**245.53**	**78.64**

*Dry weight not attempted

Sample A consisted of synthetic stuffing and was the smallest of the three samples with an actual dry weight of 11.4g. The dry weight results were very consistent between participants, with a range of just 0.76g. Only one participant (MA2904) recorded a dry weight lower than the actual dry weight (11g), but since their results were provided as whole numbers this is likely to be due to the sensitivity of the balance used rather than loss of material. This sample also had the smallest range of wet weight results, between 42.36g and 75g. The average wet weight across all participants was 55.24g, which is 10.17g higher than the expected wet weight of 45.07g. None of the dry or wet weights were flagged as 'fails' using z-scores calculated using the mean participant value.

Sample B consisted of wool, with an actual dry weight of 32.5g. Participant dry weight results ranged from 32.1g to 36.07g, with an average of 34.02g. This sample had the widest range of wet weights, varying between 93.47g and 289.78g. Whilst the wool samples have had the largest ranges in previous years, the large range in the current results was primarily due to the exceptionally high weight recorded by laboratory MA2901, which was more than double the expected weight of 127.26g. The latter sample wet weight was the only 'fail' identified using the z-score of the mean participant value. This suggests that this sample was not rinsed and/or squeezed sufficiently to remove excess material and water.

Sample C was the largest of the three samples, consisting of shredded j-cloth with an actual dry weight of 75.1g. The participant dry weights were all higher than the actual weight, ranging between 77.11g and 81g. It may be that the larger sample size creates difficulties in sufficiently rinsing and/or drying the material, as has been concluded from the dry weight results for the largest samples in previous ring tests. The wet weights were all below the expected sample weight of 293.17g, ranging from 210.45g to 291g. The participant results were more consistent for the wet weights of Sample C than for the wool material comprising Sample B. There were no 'fails' using z-scores calculated using the mean participant value.

2.2 Comparisons with expected wet weights and actual dry weights

The expected wet weight for each sample was calculated using historical NMBAQC biomass ring test participant data combined with the current year's results. Whilst the raw data for previous years was not made available, it was possible to extract most of the relevant data from the previous bulletins downloaded from the NMBAQC website. The only exception was RT04, for which a duplicate of that year's percentage cover bulletin appears to have been erroneously uploaded in place of the biomass bulletin. All other historical data were used to plot measured wet weights against known dry weights to generate a best fit trendline and corresponding linear equation. This linear equation was then applied to the known dry weights for the current year's samples to calculate an 'expected' wet weight for each sample. The linear equation applied to this year's data was y = 3.8951x +0.5844 (where x is the known dry weight and y is the 'expected' wet weight). The resulting expected wet weight for samples A, B and C were 45.07g, 127.26g and 293.17g respectively.

Comparing wet and dry weights using z-scores calculated from the expected wet weight and actual dry weight is less accommodating and more sensitive to slight deviations in results than comparisons against the mean. Consequently, six of the laboratories 'failed' at least one of the samples. The seventh laboratory only submitted data for wet weights. The wet weight for Sample A resulted in two 'fails', both of which recorded wet weights more than 50% larger than the expected value. There were no 'fails' for the dry weights of Sample A, which were all very consistent. The wet weights for Sample B resulted in only one 'fail', unsurprisingly for the same outlier that 'failed' against the mean value. The highest dry weight for sample B also resulted in a 'fail', most likely due to the lower standard deviation amongst the dry weights. Sample C had the highest number of 'fails', with three for the wet weight and four for the dry weight. The expected wet weight for this sample was higher than all the recorded wet weights, whereas the participant dry weights were all higher than the actual dry weight and therefore many results fell outside of the acceptable +/- 2.00 z-score range.

Overall, the range of results for both the dry and wet weights (as presented in Bulletin OMB RT13) when compared against the mean values could be considered acceptable with only

one ‘fail’. However, one of the limitations of using z-scores is that high standard deviation values reduce the chance of achieving a ‘fail’ based on the resulting +/- 2.00 cut-off value. The results for comparisons against expected and actual wet and dry weight, respectively, are also consistent with previous years with an increased number of ‘fails’ recorded using z-scores.

3. Discussion

Of the nine sets of samples distributed, seven laboratories submitted results. One of the laboratories did not attempt dry weight measurements. 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 an overestimation of both wet and dry weights. Macroalgae is much harder to rinse, especially in the field, which may contribute to an overestimation of the levels of biomass present. Conversely, some laboratories do not measure wet weight and instead only record 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 the different measurements and procedures utilised.

The level of accuracy remains greater for measurement of dry weight than of wet weight, for reasons given above. There is also a greater degree of consistency in results for smaller or mid-range sample weights, e.g. weight from 5g to 40g. The results overall suggest the techniques used to rinse and squeeze vary considerably between laboratories and may also vary 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 highly difficult to regulate between field workers. It is the wet weight that is most commonly used during routine monitoring of opportunistic macroalgae and therefore this lack of consistency in technique should be fully addressed within the standard operating procedures, especially in association with areas of high biomass. Each laboratory should have its own in-house training and competence assessment measures. In the method document distributed with the samples it is recommended that ‘where the sample is large it should be divided into smaller clumps for squeezing’ and ‘this should be 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 be properly squeezed’.

Most laboratories produced dry weights greater than the actual biomass of the sample; this indicates either insufficient drying or rinsing of the sample, some degree of which is to be expected during such a test. This is the sixth year in which the materials have been used separately, as opposed to mixed material samples as used in earlier years and it seems clear that there are differences in handling properties for the different materials.

In the report for RT12 a trend was identified potentially correlating increasing sample weight to increased deviation from actual/expected weight on the basis that the larger samples may

retain more debris and be more difficult to rinse free, squeeze or dry thoroughly. It was also noted that it appeared to be more difficult to obtain an accurate dry weight for the wool material than the other sample materials. However, for the tests in RT10, RT11 and RT12 the largest sample was also always comprised of wool, making it impossible to separate the effects of increased sample size from the effects of more difficult material handling on increased deviation in results. In the current year's test, the wool sample (Sample B) had a standard deviation more than double that of the largest sample (Sample C), comprised of j-cloth material. Whilst this was primarily due to one extreme participant result, the implication is that increased deviation from the actual/expected weight is co-dependent on both sample size and composition material. As such, it may be beneficial to try mixing materials again in future ring tests to reduce any 'material effect' and provide a sample more representative of the mixture of species that would be found in the field.

In general, the results for the current year were comparable with those from previous years. The ring test can provide evidence of problems in the measuring of biomass samples, such issues may need addressing through workshops and specifically aimed training. The results bulletin also provides those laboratories with outliers an opportunity to review the procedures used 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.

4. Conclusions and Recommendations

Observations made from the results of this year's exercise and from participants' feedback are summarised below:

1. Despite the artificial nature of the sample material, the test has been generally well accepted by all laboratories with constructive comments on points of possible improvements. All laboratories found the test useful despite the differences between the processing of artificial material compared with actual macroalgae samples.
2. All samples arrived with participants in good condition. This year an objection was raised to the use of anoxic mud as part of the distributed samples. This sediment was selected to provide a realistic representation of the type of habitat in which opportunistic macroalgae may occur in the field, but it is acknowledged that this can make the samples unpleasant to process if they are left for any length of time. This will be taken into account when preparing the samples for future ring tests.
3. There is general agreement that whilst synthetic materials cannot fully replicate real algae samples, the use of artificial material is an acceptable surrogate for the test. This is the fifth year in which synthetic stuffing has been used to mimic much finer opportunistic algae such as *Pilayella* and *Chaetomorpha*. This was considered the most difficult to process due to the difficulty in rinsing free of sediment, but this was somewhat offset by the smaller sample size. Wool and j-cloths were considered the most representative materials, with j-cloths also being easiest to process whereas the wool was the most time-consuming. It may be possible in the future to source alternative materials that may be more representative of the texture and general nature of opportunistic algae but at this stage alternative materials have not been

tested with the same success rate. Throughout the thirteen years of the OMB ring test there has so far been no consensus on the preferred material of use and can depend on the current opportunistic blooms being experienced in the field. The only suggestion for a potential alternative material was for fine plastic bag strips or bubble-wrap to represent some fine *Ulva* species, although it was noted that these would not be easy to process. Further investigation of the viability of alternative materials will be carried out before the next round of samples is prepared.

4. This has been the sixth year in which each sample has consisted of a different artificial material. Due to the mixed opinions on which material is the most representative all three materials will continue to be used for future tests or until a more realistic alternative can be found. However, it has been suggested again that at least one of the samples should be a combination of all three materials to provide a more realistic representation of mixed algal stands in the field. It should also help to reduce differences in results caused by the handling properties of the different materials.
5. During this thirteenth cycle of the macroalgae biomass exercise seven 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 one week prior to the completion of the exercise to aid with this process. If laboratories suspect that they will not be able to submit results within the designated timescale prior notice is required (preferably two weeks prior to the deadline) to allow for this to be factored into the reporting time scales.
6. This year six out of the seven participating laboratories submitted data for both wet and dry weights for all samples and the seventh laboratory did not attempt dry weights. Although many in-house field procedures do not incorporate dry weight of algal samples these values are included in the NMBAQC scheme to allow comparison of laboratory procedures. The values provide evidence of insufficient rinsing of samples, whereby the dry weight is 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 accurate. However, the dry weight can be compared directly with the original weight of the samples which was measured very accurately prior to addition of debris. In addition, some laboratories only measure dry weights and therefore, for such an exercise to be appropriate for these laboratories this measure of biomass needs to remain within the test. It is in the interest of all participating laboratories to complete both aspects of the test as submission of partial results may hinder any explanation of outliers and skew statistics due to the relatively small datasets. During future ring tests, it is recommended that all laboratories should continue to complete the full exercise even if it is not part of their routine monitoring to maximise the usefulness of the ring tests.
7. It is evident that the larger samples can increase the margin of error and reduce consistency between laboratories. However, it has been suggested that these samples are more appropriate in terms of representing natural conditions. Larger samples are more difficult to handle and process with higher risk of outliers or loss of material during the rinsing phase, but they are a necessary component of the test to ensure a full range of sample sizes are represented.
8. There were an increased number of participating laboratories in the current exercise than in RT12, providing a sample size more comparable to pre-pandemic levels.

Larger participant sample sizes provide more accurate mean values and help to identify outliers that could easily skew the data in a smaller sample of participants.

9. This year all participants entered their results into the spreadsheets provided. This has made the analysis process smoother and reduced the risk of errors during subsequent calculations. It is requested that participants continue to submit only final dry and wet weight results using the workbooks provided to reduce the risk of transcription errors.
10. There is some question as to whether the methodology for both wet weight and dry weight is being read and followed consistently across all laboratories. This applies to the appropriate squeezing of samples and the removal of debris. It is clear in the methods that when working with a large biomass this should be split into smaller sizes such as the size of a tennis ball, to ensure they can be squeezed properly. Any attempts to squeeze the sample as a whole will result in too much residual water being retained within the sample and increase the wet weight. This can affect the whole sample weight and increase the average. The length of time required to dry the samples may also vary from sample to sample and from lab to lab and if the samples are not completely dried or thoroughly checked prior to weighing this can result in a dry weight significantly greater than the actual dry weight.

If anyone has further thoughts on this, or disagrees with any of the interpretation, please pass forward your comments to nmbaqc@apemltd.co.uk. The biomass ring test is now in its thirteenth year and although proving successful it is still open to continual refinement.