The NE Atlantic Marine Biological Analytical Quality Control Scheme

www.nmbaqcs.org

Macroalgae/Angiosperms Percentage Cover

Component Report

Ring Test OMC RT14 2022

Søren Pears &
Georgina Brackenreed-Johnston

APEM Ltd.

28th April 2023

E-mail: nmbaqc@apemltd.co.uk

Contents

List of Figures

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 driven primarily 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 improve consistency between analysts and increase confidence in ecological quality status.

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

- Estimation of percentage cover
- Comparison of methodologies

This is the fourteenth year in which percentage cover estimations of macroalgae have been included as an element of the NMBAQC scheme and the twelfth year for which seagrass has been assessed as a separate entity. This included one exercise for macroalgae and one for seagrass, each of which were subdivided into three separate tests based on different methodologies. The format followed that of previous years (RT03 - RT13). Test material was distributed to participating laboratories along with standardised data forms, which were completed with macroalgae and seagrass percentage cover estimates and returned for analysis.

Graphical representations of the performance of each participating laboratory were distributed in the macroalgae and seagrass OMC RT14 Bulletin Report. This bulletin included the z-score based ‘Pass’ and ‘Fail’ flags assigned to each result to highlight deviation from sample means and image analysis values. The current report describes the results in more detail and should be read in conjunction with the OMC RT14 Bulletin.

1.2 Participating Laboratories

Eleven laboratories were issued test material. Eight laboratories completed the macroalgae/seagrass component percentage cover of the NMBAQC scheme with a total of 25 individual participants. Of those laboratories submitting results, all eight were government organisations. Due to the nature of the exercise, there was no limit on the number of participants per laboratory.

Judging percentage cover by eye is subjective and the preferred method of estimation varies between laboratories. Participants were given the option to complete the percentage cover test that best represented the methodology used within their laboratory. However, participants were also encouraged to complete all three variations of both the macroalgae and seagrass exercises so that the results obtained using different methodologies could be compared.

The NMBAQC scheme was originally set up for benthic invertebrate data submission to the NMMP (National Marine Monitoring Plan) to determine that data were fit for submission to the scheme. Macroalgal/angiosperms data are not submitted to any such scheme. However, they are used for classification, so it is important that they are correct. At present this scheme does not specify a definite qualifying performance level, and NMBAQC ring tests may be treated as training exercises. However, in previous years certain indicative targets have been applied to the assessment of the results based on calculated z-scores to allow "Pass" or "Fail" flags to be assigned, which provides competent monitoring authorities with an option for internal monitoring of performance. For consistency with previous years, these same criteria have been maintained for the current year. The ring tests offer a means to assess personal and laboratory performance and to identify training requirements or potential areas for improvement in existing field and laboratory procedures.

2. Summary of the Percentage Cover Exercises

2.1 Introduction

There was one exercise for the assessment of percentage cover of opportunistic macroalgae and one for seagrass. Each test included three methodology options. The exercise is described in full below to include details of distribution and logistics, procedures for estimation of percentage cover, completion of test result forms and full analysis and comparison of final submitted results.

2.2 Description

This exercise examined the participants' ability to accurately estimate various levels of opportunistic macroalgae and seagrass percentage cover. The exercise can determine the level of inter-laboratory variation and the degree of deviation from percentage cover estimations as calculated using image analysis software. It identifies areas of significant error, problematic coverage, or misuse of grid squares for aiding with estimations.

Three sets of 15 representative macroalgae and seagrass quadrat photos were distributed to each participating laboratory in January 2023. Participating laboratories were required to estimate the percentage cover of the opportunistic macroalgae and seagrass using one or more of the methodologies provided. The photographs were taken to be consistent with those provided for previous years, with two of the sets modified with overlaid grid systems. Opportunistic macroalgae consisted of species of *Ulva*, and seagrass was identified as *Zostera noltii*.

2.3 Logistics

The test material was distributed via an online file sharing link sent to each laboratory. The files contained the six tests, a description of methods and standardised forms for data submission. Participants were given six weeks to complete the test and return their results. There were no restrictions on the number of participants per laboratory.

Email has been the primary means of communication for all participating laboratories.

2.4 Confidentiality

To preserve the confidentiality of participating laboratories, each participant was randomly allocated a four-digit laboratory code to allow them to identify their results. The two-letter prefix 'MA' refers to the scheme component and this is followed by two digits representing the current NMBAQC scheme year, and the final two digits representing the laboratory. It was noted during the previous scheme year that the macroalgae component was out of synchrony with the rest of the NMBAQC scheme components. Therefore the '29' year prefix has been repeated this year for RT14, to ensure that the macroalgal component is now consistent with the other NMBAQC scheme components. For those laboratories that provided multiple submissions, the laboratory code is followed by a letter suffix to distinguish each participant of that laboratory. For example, the third participant from laboratory twelve in scheme year twenty-nine would be recorded as MA2912c.

2.5 Preparation of the Samples

To assess the accuracy of opportunistic macroalgae and seagrass percentage cover determination, photographs were taken of quadrats placed to include varying amounts of macroalgae or seagrass cover. In total 15 representative photographs of macroalgae and 15 of seagrass were taken by APEM Ltd for the purpose of this exercise.

Each set of 15 photographs was modified with two different overlaid grids to produce the total of three tests for each component to facilitate different methods of percentage cover estimation.

2.5.1 Method A

Method A used an open quadrat, which allowed the participant to estimate the percent cover in the 0.25m² quadrat without visual obstruction or assistance from gridlines. A general estimation was conducted looking solely at the total area within the quadrat that is clearly covered by the opportunistic macroalgae or seagrass.

2.5.2 Method B

Method B used an overlaid grid to divide the divide the 0.25m² quadrat into 25 squares, with each square representing 4% of the total quadrat area. The percentage cover was estimated by counting the number of squares, to the nearest half square, that were covered by macroalgae/seagrass. Completely covered squares were counted as one each. Between 50% and 100% cover in individual squares was estimated to the nearest quarter and these portions were summed. For quadrats with sparse macroalgae cover (i.e., always < 50% cover per square) the participants accumulated the small portions of algal coverage (totalling to the nearest half square). The total number of covered squares was then divided by 25 and multiplied by 100 to provide the total percentage cover.

2.5.3 Method C

Method C used an overlaid 9 x 9 crosshair grid to divide the 0.25m² quadrat into 100 squares. The 'crosshair' refers to each point at which the gridlines cross and with a 9 x 9 grid there are a total of 81 crosshairs. The estimation of percentage cover was calculated by recording the presence or absence of macroalgae/seagrass under each of the crosshair points. Where seagrass or macroalgae were present a crosshair was given a score of 1 and where absent a score of 0. The total number of crosshairs with macroalgae/seagrass present was then divided by 81 and multiplied by 100 to provide a total percentage for the quadrat.

2.6 Quadrat Image Analysis

An image analysis programme called ImageJ was used to calculate a more objective measurement of percentage cover that could be compared with the traditional means of assessment following the methodology described by Xiong *et al*. (2019). Previous ring tests have sought a full, impartial image analysis comparison as part of the QC exercise. Initially this was attempted using GIS software, but this did not provide a fully independent analysis of percentage cover. ImageJ image analysis software was chosen to be less subjective by providing a more accurate analysis based on colour/tone contrast. Image analysis has been carried out to demonstrate how the comparisons would work but may still require further modification and discussion as to its applicability and accuracy, therefore cannot be taken as a definite measure of percentage cover.

Prior to analysis each quadrat photo was edited using Photoshop, cropping each image to the exact 0.25m² inside area of the quadrat and increasing the green colour saturation to ensure a substantial contrast between the seagrass and macroalgae against the background. The photograph was then processed using the ImageJ program. Firstly, the image measurements were calibrated according to the quadrat dimensions. Then each image was separated into two portions of green and non-green areas by adjusting the hue, saturation and brightness colour threshold settings to match the areas of macroalgae or seagrass. The resulting green area selection was used to calculate the area of coverage in cm² and this was converted to a percentage by dividing the result by 2500 (i.e. the total quadrat area) and multiplying by 100. The resulting percentages were used as a comparison against the skilled eye estimations as submitted by the participants.

2.7 Analysis and Data Submissions

A results workbook was distributed to each participating laboratory along with the exercise instructions to standardise the format in which the results were submitted. These results will be retained and stored appropriately. Each participant had the option of completing the test which most represented their own procedures, but all participants were encouraged to complete all three tests of both macroalgae and seagrass to allow a comparison of methodologies and levels of accuracy achieved within each.

For each test the participant had to estimate the percentage cover of opportunistic macroalgae or seagrass species only, excluding any additional species that might be present within the quadrat and that were not considered to belong to either of these types of species. The assessment included a broad range of variation in percentage cover to represent the full range that could be experienced in the field.

2.8 Z-Scores

Z-scores were calculated to determine how many standard deviations each participant's percentage cover value was separated from the mean percentage cover value using the following formula:

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

Where:

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

$\boldsymbol{\mu}$ is the mean of the participants' percentage cover values for that test;

$\boldsymbol{\sigma}$ is the standard deviation of the participants' percentage cover values for that test.

Z-scores were calculated separately using the mean of the participants' percentage cover scores and then using the percentage cover score derived from the ImageJ analysis. For consistency with previous ring tests, a z-score value 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 a 'Fail' or 'Pass' flag on the submitted data.

3. Results

The results have been analysed using a variety of approaches to compare the results between participants, between the three different methods of estimation and to compare against ImageJ calculated percentage cover estimations for both macroalgae and seagrass.

3.1 Participant Data Received

Of the laboratories that submitted data, there were fourteen participants that completed method A, eleven that completed method B and fourteen that completed method C for the macroalgae exercise. For the seagrass exercise twenty completed method A, seventeen completed method B and fourteen completed method C. Four participants completed all three macroalgae tests and four completed all three seagrass tests. The results have been collated and represented in various formats to enable full comparisons between participants and against the percentage cover calculated using image analysis.

3.2 Macroalgae Results

3.2.1 Macroalgae Test A (Open Quadrat)

Test A was completed by fourteen participants and along with Test C was the joint most popular of the three methods. The ranges of percentage cover estimates per quadrat was generally higher than methods B and C, varying from 2% to 25% per quadrat. Six quadrats had ranges of less than 10% cover between participants and a further six quadrats had percentage cover ranges of between 10 and 20%. These were higher ranges than were recorded for the previous year, but this was most likely influenced by the much smaller sample size in the previous year. The largest ranges were recorded for quadrats 15, 9 and 2, with ranges of 22%, 23% and 25% cover, respectively. Six of the percentage cover values were deemed ‘fails’ when using the z-scores calculated from the mean percentage cover value, giving an overall pass rate of 97.14%.

In comparison, when using z-scores calculated from the ImageJ analysis percentage cover values there were a total of 30 ‘fails’. Eight of these were for quadrat 5, six were for quadrat 4 and five were for quadrat 10. Quadrat 14 had a further three ‘fails’, along with two each for quadrats 1 and 9 and one each for quadrats 2, 6, 12 and 15. The ‘fails’ resulted from a mixture of over and underestimation compared to the ImageJ values. The pass rate of z-scores against ImageJ results was 85.71%, which is higher than in the previous year.

Deviation from the mean varied between participants, ranging from 2.02% and 5.09% taken as an average across all quadrats. Deviation from ImageJ was slightly higher ranging between 2.78% and 7.51%. Both these results were slightly higher than the previous year, but as with the results above, this is likely due to the lower number of participants in the previous year. The average deviation of the mean participant percentage cover from ImageJ calculated values was -1.71%, indicating that there was an overall tendency towards underestimation.

3.2.2 Macroalgae Test B (5 x 5 Gridded Quadrat)

Test B proved slightly less popular than methods A and C, with a total of eleven participants choosing this method. The ranges of percentage cover estimates were slightly lower than Test A, varying from 1% to 21%. Nine of the quadrats had ranges of less than 10% cover between participants. Four quadrats had ranges between 10% and 20% cover and the remaining two quadrats each had a percentage cover range of 21%. These were very similar ranges to the previous year.

The were four ‘failed’ quadrats when comparing z-scores against the mean, giving a 97.58% pass rate for this test component. In comparison the total number of ‘fails’ when compared with image analysis was 64, with a pass rate of only 61.21%. All eleven participants ‘failed’ at least three quadrats with 8 ‘fails’ attributed to one participant. The average deviation per participant compared to the mean was relatively low, ranging from 0.92 to 2.68. Average deviation from ImageJ analysis was slightly higher, ranging from 2.80 to 4.18. The average deviation of the mean participant percentage cover from ImageJ calculated values was 0.43%, indicating an approximately even mixture of overestimation and underestimation of percentage cover.

3.2.3 Macroalgae Test C (9 x 9 Crosshairs Quadrat)

Test C was the joint most popular of the three methods and was completed by fourteen participants. The range of results was the lowest of the three methods, varying from 1.23% to 12% per quadrat. Eight quadrats had ranges of less than 10% cover and the other seven varied by 10-20% cover between participants.

Using z-scores based on participants' mean percentages, two participants 'failed' one quadrat and one participant 'failed' two quadrats to give a total of four 'fails'. The overall pass rate was 98% which is higher than previous years. All four of the 'fails' were for different quadrats. Comparisons against image analysis resulted in 66 'fails' and a 68.57% pass rate, which is lower than the previous year. All but six of these 'fails' were overestimates compared to the ImageJ calculated results.

The average deviation from the mean across all quadrats was low, ranging from 1.12 to 5.65. Average deviation from the ImageJ results ranged from 1.88 to 4.30, which is lower than method A, similar to method B and lower than in previous years. The average deviation of the mean participant percentage cover from ImageJ calculated values was 1.71%, indicating more of a tendency towards overestimation of percentage cover than was seen with Tests A and B.

3.3 Seagrass Results

3.3.1 Seagrass Test A (Open Quadrat)

For the seagrass quadrats, Test A had the highest number of respondents, with twenty participants opting for this method. The results submitted showed more variation than the macroalgal test, with ranges between 1% and 50% per quadrat. The largest range of 50% was recorded for quadrat 8 with estimates varying between 30% and 80%. A further 9 quadrats had percentage cover ranges above 30% and only one quadrat (quadrat 7) had a range below 10%. The average range across all participants and quadrats was 27.6%, which is higher than the RT13 results.

Z-scores calculated using the population mean resulted in ten 'fails', giving a 96.67% pass rate for test A when using z-scores derived from the mean, which is slightly lower than the previous year. When comparing results against percentage cover calculated using ImageJ the number of 'fails' was 38, giving a pass rate of 87.33%, which was higher than the previous year.

The average deviation of results from the mean and image analysis percentage cover per laboratory ranged from 2.71 to 10.49 and 4.83 to 10.51 respectively. These deviations are higher than those recorded for macroalgae Test A but are consistent with seagrass test results from previous years. The average deviation of the mean participant percentage cover from ImageJ calculated values was 0.63%, indicating an approximately even mixture of overestimation and underestimation of percentage cover, but with a slight tendency towards overestimation.

3.3.2 Seagrass Test B (5 x 5 Gridded Quadrat)

Test B had three less respondents than Test A, with a total of 17 participants opting to complete the 5 x 5 square grid quadrat method, in contrast to most previous years where this has often been the least popular method. The range of results for this test was narrower than for Test A, varying between 0 and 41% with a mean range of 21.97%. Quadrat 13 had the highest range of results, with estimations varying between 49% and 90% cover. Three of the quadrats had ranges of less than 10% cover between participants, two had ranges between 10 and 20% cover, six ranged from 20 to 30% cover and a further three ranged between 30% and 40% cover between participants.

Comparing z-scores against mean percentage cover resulted in five 'fails' between three participants, with an overall pass rate of 98.04%. In comparison, the total number of 'fails' when comparing against ImageJ results was much higher at 55 and was distributed amongst all seventeen participants. The overall pass rates using image analysis % cover was 78.43%, which is lower than the results of Test A, but consistent with results from the previous year.

The deviation from mean percentage cover (2.72 – 8.55) was slightly lower than the previous year despite a higher number of participants in the current year. Deviation from ImageJ analysis (4.01 - 10.39) was slightly higher than deviation from the mean but was also lower than the previous year's results. The average deviation of mean participant results from ImageJ analysis values was lower than for test method A, at 0.39 and indicates an approximately even mixture of overestimation and underestimation of percentage cover values.

3.3.3 Seagrass Test C (9 x 9 Crosshairs Quadrat)

Test C was the least popular method, having three less participants than Test B. The percentage cover ranges varied from 0% to 43.95%. Three quadrats had ranges of less than 10% cover between participants, one had a range between 10 and 20% cover, four ranged from 20-30% cover between participants. Six quadrats had ranges of 30% or above, with the highest range for quadrat 15, which had results varying between 40% and 83.95% cover.

Comparison of z-scores calculated from the mean resulted in nine 'fails' distributed between four participants, giving a total pass rate of 95.71%. Comparing results against the ImageJ calculated values gave 46 'fails' with a pass rate of 78.10%. Eight of the 'fails' were underestimates compared to the ImageJ analysis calculated values and all of the rest were overestimates.

Deviation from mean percentage cover varied between 1.55 and 10.57, which was comparable to Tests A and B and data from the previous year. The deviation from the ImageJ analysis values (5.14 – 17.68) is higher than the deviation from the mean but is a slight decrease from the previous year. The average level of deviation between percentage cover estimates and image analysis across all quadrats and participants was 7.26, indicating a tendency towards overestimation of % cover compared to ImageJ results as noted above.

4. Discussion

The percentage cover of opportunistic macroalgae or seagrass in 0.25 m² quadrats is usually estimated in the field based on a skilled eye observation using either an open quadrat or gridded quadrat with the aim of achieving a variance of less than 5% between surveyors. It is highly unlikely that this method of percentage cover estimation is 100% accurate due to the subjectivity of individuals, although over time people can become highly skilled. It is difficult to establish an unambiguous 100% reliable method for determining percentage cover of opportunistic macroalgae or seagrass. Based on the methodology established in previous ring tests, OMC RT14 used the population mean and an image analysis software program (ImageJ) to calculate a more objective percentage cover for comparison with individual participants' results. The use of image analysis software is considered to provide less subjectivity than skilled eye estimations.

The exact methodology used to prepare and analyse images in ImageJ in ring tests RT2-12 was not specifically defined, and the program offers multiple possible techniques for the calculation of percentage cover. These include manual definition of areas of macroalgae/seaweed, defining areas based on colour thresholds or converting the image to binary (i.e. black and white) based on a defined contrast threshold. The resulting calculation of percentage cover therefore still has scope for variability depending on the settings used and in particular the way the selection threshold is defined. In keeping with RT13, the methodology described by Xiong *et al*. (2019) for determination of vegetation cover was used for the current year. This involved defining the areas of macroalgae or seagrass cover by adjusting hue, saturation and brightness threshold settings to match the observed areas of macroalgae or seagrass.

Z-scores were used to establish a level of acceptance for results submitted by participants following the same methodology used in previous ring tests. Separate z-scores were calculated using both the mean percentage cover per quadrat recorded by participants and the percentage cover as calculated using ImageJ analysis. The results could then be compared between participants and between methodologies of cover estimation for both macroalgae and seagrass. As in previous years the number of 'fails' was higher when comparing results against ImageJ analysis values rather than against the population mean. This is unsurprising given that the mean is calculated directly from the participant data whereas the ImageJ value is derived from independent analysis. The benefit of comparing participants' results against the mean is that it fully represents the range of results submitted and this is not the case for the ImageJ results. This does not negate the value of using ImageJ analysis for comparison as this still the most objective method determined so far and using the mean is naturally going to allow more estimates to sit within the z-score +/- 2.00 range.

The range of results provided was higher this year than the previous year, which is likely to be due to an increased number of participants this year but was consistent with results seen in earlier years of the percentage cover exercises. The ranges are much greater than the recommended +/- 5% between surveyors, with some quadrats having estimated ranges up to 50%. The average range per test varied between 6.37% and 14.27% for macroalgae and between 21.97% and 27.60% for seagrass. As noted in the previous year, one of the limitations of using z-scores is that when standard deviation values are high, the chance of achieving a 'fail' are reduced based on the resulting +/- 2.00 cut-off value. Conversely, low standard deviation values give much less tolerance for outliers, which can lead to higher rates of ‘fails’ when comparing participant results to ImageJ results, which are derived independently from the population mean. These results, along with those from previous years, still require further examination to improve the methodologies employed and the means in which the percentage cover is calculated both by field methods and ImageJ analysis.

In previous years it was observed that quadrats with either a very high or low percentage cover have been easier to accurately estimate total cover, whereas quadrats with a percentage cover in the middle range (30 – 70%) generally result in a higher level of deviation with a much broader range of results. This trend continues in the current year and could be attributed to the patchier coverage of opportunistic macroalgae and seagrass in some quadrats which is much harder to estimate accurately. There is also a broader range of percentage cover estimations and deviations for both the mean and ImageJ analysis for seagrass than for macroalgae. Seagrass displays a much patchier nature of growth; its thin long strands often make it difficult to estimate percentage cover leading to a broader range of results and high levels of deviation.

Figure 1 and Figure 2 (overleaf) show this range of percentage cover results for macroalgae and seagrass, respectively. These scatter graphs indicate that in general Test B for both macroalgae and seagrass had smaller ranges of results across participants, whilst Test A generally had the greatest ranges. Overall, the broad range of percentage cover estimates submitted by participants is concerning in terms of consistency between laboratories as well as within laboratories.

There are noticeable differences both between seagrass and macroalgae results and between the different methods of estimation used and the resulting number of ‘fails’ (Figure 3). The macroalgae results for comparisons against the mean had a slightly higher number of ‘fails’ for Test A compared to Test B and C, whereas for comparisons against ImageJ highest number of fails was for Test C and the lowest for Test A. The seagrass results had the lowest numbers of ‘fails’ for Test B and highest number for Test A for comparisons against the mean but had the opposite pattern for comparisons against ImageJ, with the highest number of ‘fails’ for Test C and the lowest number for Test A.

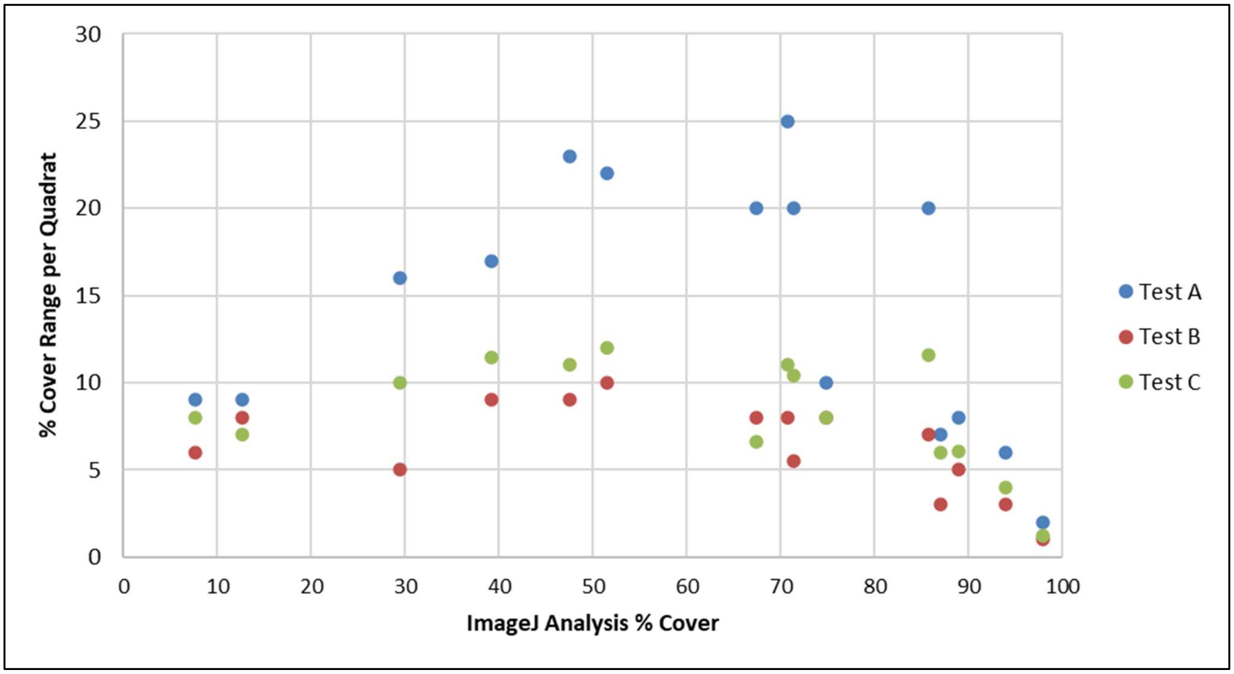

Figure 1 **Scatter graph showing the range of percentage cover results per quadrat across all three opportunistic macroalgae test methods**

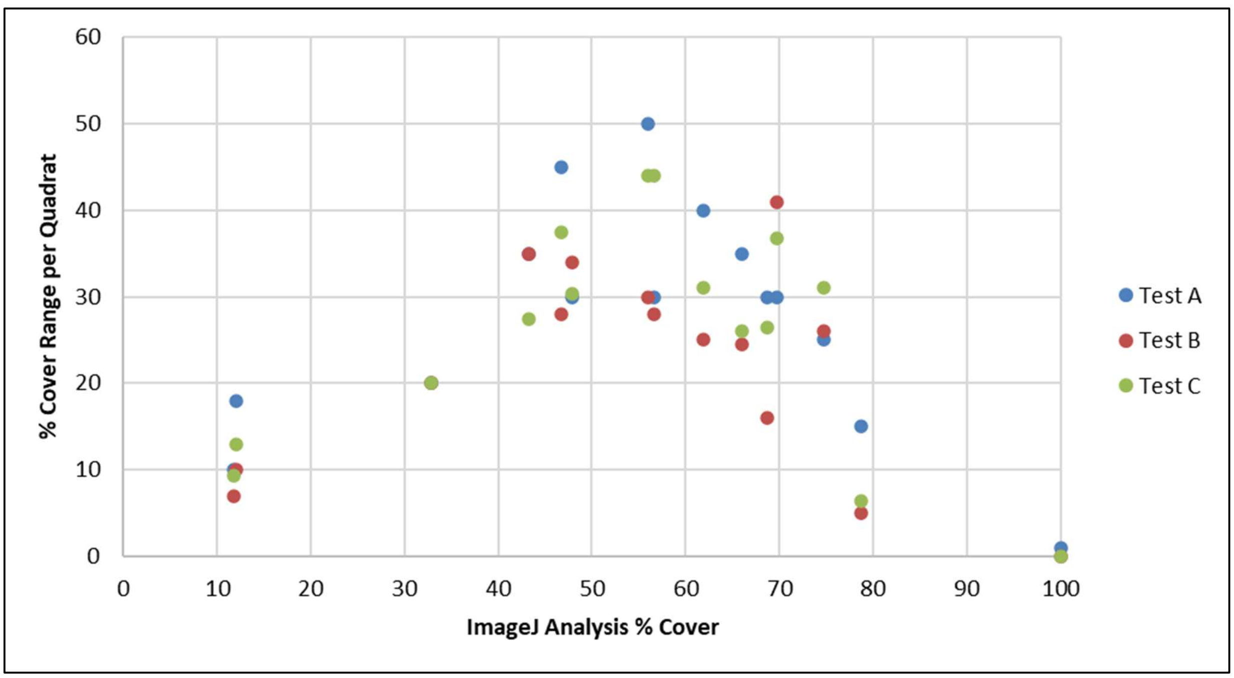

Figure 2 **Scatter graph showing the range of percentage cover results per quadrat across all three seagrass test methods**

There do not appear to be consistent patterns of variation in results between years, although direct comparisons are hampered by the variations in numbers of participants each year.

Figure 4 shows the results from the OMC RT13 for comparison. The number of ‘fails’ against the mean were mostly lower in RT13 than the current ring test, with the exception of Test C for macroalgae. The number of ‘fails’ against ImageJ were all lower in RT13 than in the current year, although there were also lower numbers of participants for all the tests in RT13 compared to the current year, which is likely to have been a significant contributing factor in the differences between years.

There were differences in the results for the different methods, with a higher number of ‘fails’ against the mean for Test A than Tests B or C for both seagrass and macroalgae in the current year. The comparisons against ImageJ showed the that for macroalgae the lowest numbers of ‘fails’ were for Test A, followed by Test B and the highest number were for Test C, whereas for the seagrass the highest number of ‘fails’ was for Test B, followed by Test C.

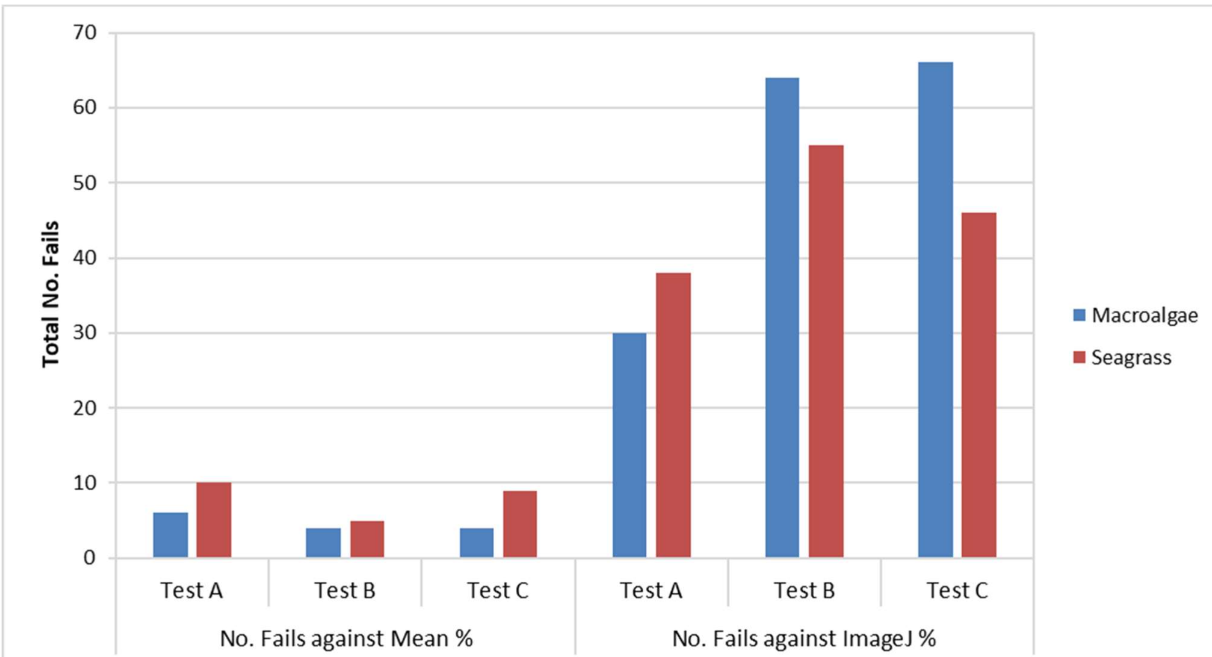

Figure 3 **Number of fails recorded in each test for macroalgae and seagrass quadrats for OMC RT14.**

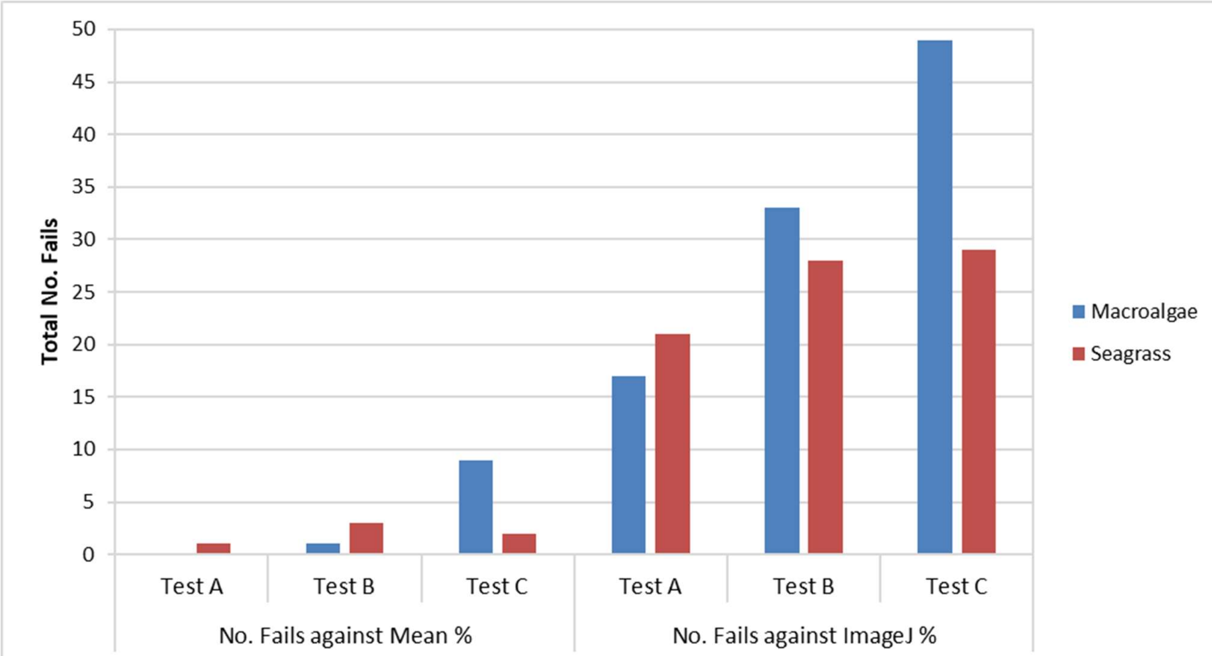

Figure 4 Number of fails recorded in each test for macroalgae and seagrass quadrats for OMC RT13.

In contrast to previous years, Test A appeared to be the preferred methodology in the current year, with the joint highest number of participants for macroalgae and the highest number of participants for the seagrass exercises. The numerical preferences of the varying methodologies will largely have been determined by the methods used in the field by the respective laboratories. Some participants cited time restraints and other work priorities preventing them repeating the tests using alternative methodologies.

In theory test method C should provide the least subjective method of estimation, as counting the number of crosshairs under which macroalgae or seagrass lay should be a relatively straight forward method. However, there is still a large disparity in results, often much higher than for the other test methods that may suggest the method is not being used consistently between participants.

Only four participants for macroalgae and four participants for seagrass completed all three different test methods, which makes it difficult to differentiate variance between methodologies from variance between individual participants. This makes it harder to conclude which of the test methods is the most accurate and consistent as it is highly dependent upon the participant and the sample size. The broad range of results across participants and laboratories is a problem that needs to be addressed either by adopting a less subjective means of estimating percentage cover of macroalgae or seagrass, or by holding a field meeting with which to synchronise such methodologies and reduce the high degree of variation currently being recorded within the NMBAQC macroalgae and seagrass percentage cover tests. The current test methods do not provide an obvious solution and where one method may work best with macroalgae this may not be the case for seagrass, so this remains an area requiring significant further investigation.

5. Conclusions and Recommendations

1. There is still a high degree of variation in results both between tests and between participants and this may prompt the need for a specific workshop whereby methods can be discussed, and possibly percentage cover estimations compared in the field. It is not possible from the current ring test to conclude which percentage cover estimation method provides the most accurate results; however, based on the dominant proportion of the data returns, during OMC RT14 Test method A was the most favoured method for both macroalgae and seagrass, contrasting with previous years where Test C has been the most popular.

2. There are still large differences between z-scores calculated from the mean and z-scores calculated from image analysis results and given the varied levels of deviation between the two it is unclear which is the most accurate method from which to compare participants results. However, the high standard deviation across all test methods is having a significant impact on the overall results.

3. The image analysis method used during RT14 aims to provide a more objective result than skilled eye estimation. However, the precise methodology used to prepare the images and calculate the percentage cover in ImageJ is still being explored and will continue to undergo further refinement for each round of tests. It is recommended at this time that participants should continue to use the z-scores derived from comparisons with the mean if they are required for internal quality reports.

4. During this fourteenth cycle of the macroalgae percentage cover exercise eight of the eleven laboratories were able to complete the ring test within the allocated timescale. It is appreciated that conflicts with other work may prevent laboratories from meeting the deadline. However, it is important that all laboratories continue to attempt to submit results within the requested deadlines as detailed at the beginning of the exercise. This is in both their own interests and brings greater benefit to all participants in the scheme by increasing the dataset and ensuring preliminary bulletins and reports are circulated within the set timescale. In subsequent years reminders will continue to be distributed one week prior to the completion of the exercise to ensure the deadline is met. Due to the interdependence of all participant results in calculating z-scores, any results submitted outside of this deadline may not be accepted and it may not be possible to included them in the analysis. It is requested that any participants unable to meet the deadline should give prior notice of two weeks.

5. This year all participants who submitted results filled out the spreadsheets provided and removed any prior calculations, particularly with regards to Test method C. This has made the analysis process much easier and reduced the risk of error during subsequent calculations. It is requested that participants continue to exclude all calculations. Where calculations or formulas are included, there is greater chance of error when transferring data to a single spreadsheet and during subsequent data analysis.

6. This year the test material was distributed via an emailed link to a file sharing website. This method both reduces unnecessary postage and ensures arrival of the test materials on the designated start date. Several participants reported difficulties in accessing the test material this year. In some cases, this was due to access being restricted to registered contact email addresses for participating laboratories rather than individual participants and also due to changes in the file sharing security settings this year. Problems were investigated and resolved as soon as possible after they were identified, but this process will be reviewed to try and make access smoother for future tests.

7. The feedback forms indicated that there is still often insufficient time or resources to complete all three test methods. Having results for the same participants across all three test methods would allow for a more direct comparison of the methods used and the results obtained. It is still recommended that all participants try to complete all three methods where possible.

8. There was feedback suggesting that some quadrats could contain a mix of seagrass and other macroalgae (i.e. not opportunistic macroalgae) as well as other substrata such as stones to represent a more representative range of natural habitats in which these species might be found. At present the results are still highly variable there is no evidence to suggest that a combined quadrat would be of benefit and may further confuse the results. This would also create difficulties distinguishing seagrass from macroalgae in the ImageJ analysis. However, it is acknowledged that in the field there can be a mixture of seagrass and macroalgae in the same location.

If anyone has further thoughts on this, or disagrees with any of the interpretation, please forward your comments to nmbaqc@apemltd.co.uk. This ring test is now in its thirteenth year and although it has general approval, we are still very happy to receive feedback particularly suggestions on how it may be improved.

6. References

Xiong, Y., West, C.P., Brown, C.P. & Green, P.E. (2019). Digital image analysis of old world bluestem cover to estimate canopy development. *Agronomy Journal* 111(3), 1247-1253.