The NE Atlantic Marine Biological
Analytical Quality Control Scheme

www.nmbaqcs.org

Macroalgae/Angiosperms Percentage Cover

Component Report

Ring Test OMC RT17 2026

Søren Pears &
Georgina Brackenreed-Johnston

APEM Ltd.
8th May 2026

E-mail: nmbaqc@apemltd.co.uk

Contents

List of Figures

1. Introduction

1.1 Background

Quality control (QC) of biological data is a high priority to enable correct water quality classification and good management decision-making. This extends through all biological elements including macroalgae and seagrass. Good QC 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 seventeenth year in which percentage cover estimations of macroalgae have been included as an element of the NMBAQC scheme and the fifteenth 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 largely followed that of previous years (RT03 - RT16) but with the replacement of standard Z-scores with modified Z-scores in the current year in an effort to find a more robust metric for identifying outliers. The rationale for this change is discussed in Section 2.8. 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 RT17 Bulletin Report. This bulletin included the modified Z-score based ‘Pass’ and ‘Fail’ flags assigned to each result to highlight deviation from sample medians and image analysis values. The current report describes the results in more detail and should be read in conjunction with the OMC RT17 Bulletin.

1.2 Participating Laboratories

Ten laboratories were issued macroalgae/seagrass component percentage cover test material. All ten laboratories submitted results for at least one the exercises, with a total of 22 individual participants. All ten laboratories that submitted results were government organisations. Due to the nature of the exercises, there was no upper limit on the number of participants per laboratory that could submit results.

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 National Marine Monitoring Plan (NMMP) 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 the current year, the methodology has been updated to use modified Z-scores in an effort to find a metric that is more resistant to differences in standard deviation and applicable to smaller sample sizes. This is discussed in more detail in Section 2.8.

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, including details of distribution and logistics, procedures for estimation of percentage cover, completion of test result forms and full analysis and comparison of the 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 2026. 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 a shared Google Drive 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 tests 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 (32), and the final two digits representing the laboratory. 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 thirty-two would be recorded as MA3212c.

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 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 0.25 m² quadrat, allowing the participant to estimate the percentage cover in the 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 0.25 m² 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 less than 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 0.25 m^2 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. For RT01 this was attempted using GIS software, but this did not provide a fully independent analysis of percentage cover. From RT02 onwards ImageJ image analysis software has been used to calculate percentage cover. This 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 definitive measure of percentage cover.

Prior to analysis each quadrat photo was edited using Photoshop, cropping each image to the exact 0.25 m^2 inside area of the quadrat and increasing the green colour saturation to ensure a substantial contrast between the seagrass or macroalgae against the background substrata. 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 cm^2 and this was converted to a percentage by dividing the result by 2500 (i.e. the total quadrat area in cm^2) 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 were retained and stored confidentially. 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 and Modified Z-scores

For the previous Ring Tests RT01 to RT16 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:

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

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

$\boldsymbol{\sigma}$ was 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. Z-score values exceeding ± 2.00 were considered to be outside an acceptable limit of deviation from the mean and this cutoff point was used determine a 'Fail' or 'Pass' flag on the submitted data.

One of the limitations of Z-scores that has been noted in the reports from previous scheme years is that they are easily skewed by the very outliers they are aiming to detect, particularly with smaller datasets. The maximum z-score is directly limited by the sample size (i.e. the number of participant results for a given test). Based on the equation from Shiffler (1988) the maximum Z-score for a given sample size *n* is: $(\boldsymbol{n}-\mathbf{1})/\sqrt{\boldsymbol{n}}$. Using this equation, if there are less than five laboratory results for a given test it would not be possible for any Z-score fails as the maximum Z-score would be 1.79.

Following the guidance in Iglewicz & Hoaglin (1993), the methodology for RT17 has been updated to use 'modified' Z-scores based on median values to provide a more robust identification of outliers. Modified Z-scores are calculated using the following equation:

$$M_i = \frac{0.6745(x_i - \tilde{x})}{MAD}$$

Where:

$\boldsymbol{M_i}$ is the modified Z-score.

$\boldsymbol{x_i}$ is the participant weight value.

$\tilde{x}$ is the median of participant weight values.

$\boldsymbol{MAD}$ is the median of absolute deviations about the median participant value or ImageJ result (as applicable).

The constant of **0.6745** is used because the expected value of *MAD* = 0.6745σ (where σ is standard deviation) for large sample numbers. Based on the guidance in Iglewicz & Hoaglin (1993) modified Z-score cutoff values of ± 3.5 were used to determine pass/fail flags for RT17. The modified Z-score method provides a more resistant estimator to identify outliers as it is not unduly affected by a few extreme observations.

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 for the macroalgae test, there were eight participants that completed method A, seven that completed method B and ten that completed method C. For the seagrass exercise 12 completed method A, seven completed method B and ten completed method C. Three participants completed all three macroalgae tests and two 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 eight participants and had results ranging between 3% and 20% per quadrat. Six quadrats had ranges of less than 10% cover between participants and a further seven quadrats had percentage cover ranges of between 10 and 20%. The largest ranges were recorded for quadrats 5 and 15, both with 20% cover. These ranges were similar to those recorded for macroalgae Test A in RT16.

None of the percentage cover values were deemed ‘fails’ when using the modified Z-scores calculated from the median percentage cover value, giving an overall pass rate of 100% using Z-scores calculated from mean participant values. In comparison, when using Z-scores calculated from the ImageJ analysis percentage cover values there were a two ‘fails’, both for quadrat 7. The overall pass rate of modified Z-scores against ImageJ results for method A was 98.33%.

Deviation from the median varied between participants, ranging from 1.63% to 4.23% taken as an average across all quadrats. Deviation from ImageJ was slightly higher ranging between 2.73% and 6.68%. The average deviation of the median participant percentage cover from ImageJ calculated values was -0.58%, indicating a slight tendency towards underestimation of percentage cover.

3.2.2 Macroalgae Test B (5 x 5 Gridded Quadrat)

Test B was the least popular method for macroalgae RT17, with seven participants choosing this method. The ranges of percentage cover estimates were similar to those of Test A, varying from 1% to 22%. Six of the quadrats had ranges of less than 10% cover between participants. Eight quadrats had ranges between 10% and 20% cover and the remaining quadrat (quadrat 11) had a percentage cover range of 22%. These were very similar to the ranges that were recorded for method B in RT16.

There were three 'failed' quadrats when comparing modified Z-scores against the median, giving a 97.14% pass rate for this Test Component. In comparison, there were no 'fails' when compared with image analysis values, giving a pass rate of 100%. The average deviation per participant compared to the median ranged from 1.30% to 6.10%. Average deviation from ImageJ analysis was slightly higher, ranging from 3.01% to 7.45%. The average deviation of the median participant percentage cover from ImageJ calculated values was 0.02%, showing no clear pattern of over- or underestimation.

3.2.3 Macroalgae Test C (9 x 9 Crosshairs Quadrat)

Test C was the most popular of the three methods for macroalgae and was completed by ten participants. The range of results was by far the widest of the three methods, varying from 2% to 97.5% per quadrat. This extreme range is primarily due to the exceptional difference between the percentage cover for quadrat 6 submitted by laboratory 3204 (98.5%) and both the ImageJ value of 3.11% and median participant value of 1.20%, which can only be assumed to be a miscalculation of the area not covered rather than the algal cover. Of the other quadrats, three had ranges of less than 10% cover, ten had ranges between 10 and 20% cover and one quadrat had a range between 20% and 30% cover.

Nine of the percentage cover values were deemed 'fails' when using the Z-scores calculated from the median percentage cover value, four of which were attributed to one participant. The overall pass rate was 94% which was lower than for methods A and B. Comparisons against ImageJ analysis resulted in only a single 'fail', giving an 99.33% pass rate.

The average deviation from the median across all quadrats ranged from 1.76% to 9.64%. Average deviation from the ImageJ results ranged from 3.15% to 10.12%, which is higher than for methods A and B. The average deviation of the mean participant percentage cover from ImageJ calculated values was 2.48%, 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 was the most popular method with 12 respondents. The results submitted showed more variation than macroalgal Test A, with ranges between 8% and 60% per quadrat. The largest range of 60% was recorded for quadrat 11 with estimates varying between 20% and 80%. A further five quadrats had percentage cover ranges above

40%, two had ranges between 30% and 40% and four had ranges between 20% and 30%. Only one quadrat had a range of less than 10% cover between participants. The average range across all participants and quadrats was 32.00%, which is noticeably higher than in the RT16 results.

Modified Z-scores calculated using the population median resulted in four 'fails', three of which were from one participant. This gave a 97.78% pass rate for Test A. When comparing results against percentage cover calculated using ImageJ the number of 'fails' was also four but, in this case, spread amongst four different participants and still giving a pass rate of 97.78%.

The average deviation of results from the median and image analysis percentage cover per laboratory ranged from 2.77% to 19.63% and 4.76% to 16.14% respectively. These deviations are much higher than those recorded for macroalgae Test A. The average deviation of the median participant percentage cover from ImageJ calculated values was -3.50%, indicating a tendency towards underestimation.

3.3.2 Seagrass Test B (5 x 5 Gridded Quadrat)

Test B was the least popular method, with a total of seven sets of results for the 5 x 5 square grid quadrat method. The range of results for this test was smaller than for Test A, varying between 4% and 41%. This range is also smaller than that recorded for Test B in RT16. As with Test A, Quadrat 11 had the greatest range of results, with estimations varying between 27% and 68% cover. Two quadrats had ranges of less than 10%, four had ranges between 10% and 20% cover, three ranged from 20% to 30% cover and the remaining five quadrats varied between 30% and 40% between participants.

Comparing modified Z-scores against median percentage cover resulted in three 'fails', giving an overall pass rate of 97.14%. In comparison, there were no 'fails' when comparing against ImageJ results giving an overall pass rate of 100%, which is higher than for the Test A and C comparisons to ImageJ results.

The deviation from mean percentage cover (2.62% – 12.45%) and deviation from ImageJ analysis (4.74% - 11.36%) were both lower than for tests A and C. The average deviation of mean participant results from ImageJ analysis values was -4.65%, indicating a tendency towards underestimation of % cover compared to ImageJ results.

3.3.3 Seagrass Test C (9 x 9 Crosshairs Quadrat)

Test C was completed by a total of ten participants. This method had the widest percentage cover ranges, varying from 6% to 66% per quadrat, which was also higher than the ranges of seagrass Test C in RT16. Only one quadrat had a range of less than 10% cover (quadrat 7) and one (quadrat 1) had a range between 10% and 20% cover. Two quadrats had ranges between 20% and 30% cover, another two ranged from 30% to 40% cover, and nine quadrats had ranges above 40% between participants. The widest range was recorded for quadrat 5, with results varying between 16% and 82% cover.

Comparison of modified Z-scores calculated from the median resulted in three 'fails', divided between two participants and giving a total pass rate of 98%. Comparing results against the ImageJ calculated values also gave three 'fails', in this case all for the same participant and again giving a pass rate of 98%, which was similar to Test A but lower than for Test B.

Deviation from median percentage cover varied between 1.97% and 27.44%, which was higher than for seagrass Tests A and B. The deviation from the ImageJ analysis values were slightly higher, ranging from 5.99% to 29.40. The average level of deviation between percentage cover estimates and image analysis across all quadrats and participants was 1.96%, 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 a slightly modified version of the methodology established in previous ring tests, OMC RT17 used both the population median and the 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.

ImageJ 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 the way the selection threshold is defined. In keeping with RT13 to RT16, 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.

Modified Z-scores were used to establish a level of acceptance for results submitted by participants following the same principle used for standard Z-scores in previous ring tests. Separate modified Z-scores were calculated using both the median 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. The updated approach to pass/fail criteria produced far fewer 'fails' than the previously used Z-score methodology but provides a more useful output in terms of identifying the true outliers in the datasets that are less dependent on standard deviation.

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 much 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. In Figure 1 the Y-axis has been adjusted to exclude the extreme range for quadrat 6 Test C, which is assumed to be due to a calculation error rather than a genuine result. These scatter graphs indicate that for both macroalgae and seagrass Test C was most likely to have larger ranges than Tests A and B. Overall, the broad range of percentage cover estimates submitted by participants remains concerning in terms of consistency between laboratories as well as within laboratories, particularly for the seagrass tests.

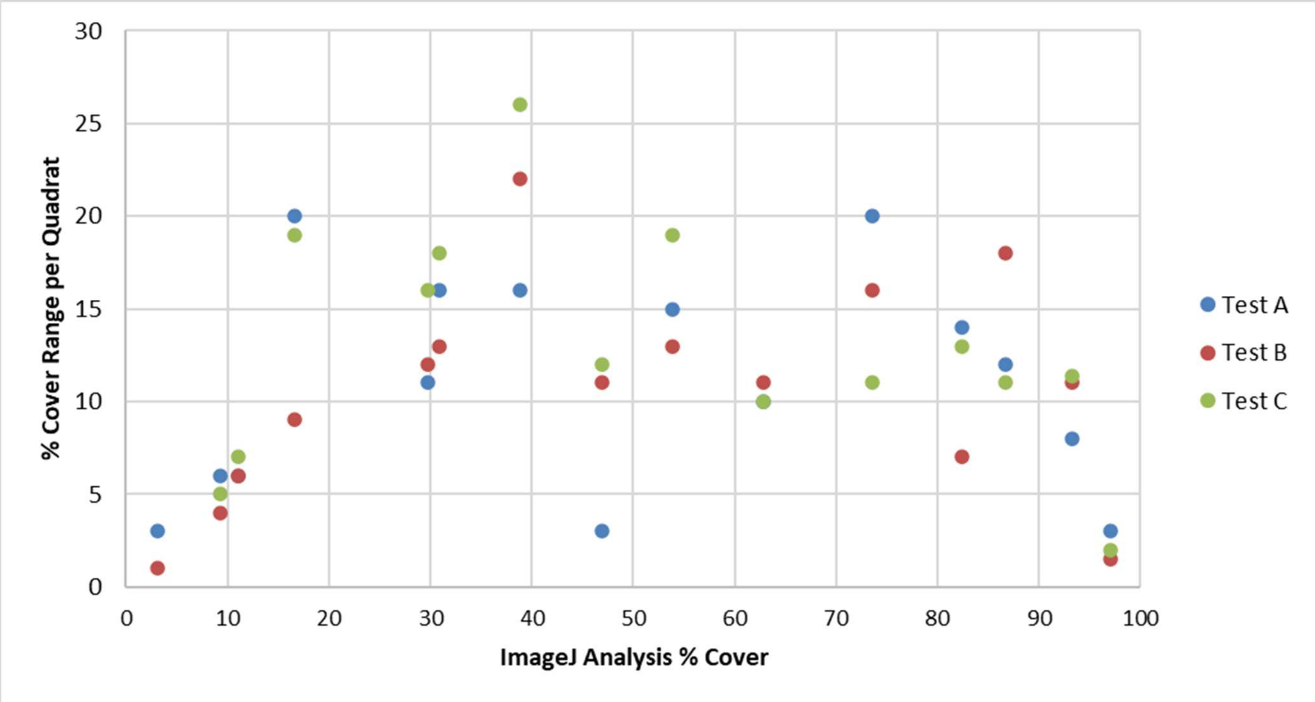

Figure 1 **Scatter graph showing the range of percentage cover results per quadrat across all three opportunistic macroalgae test methods**

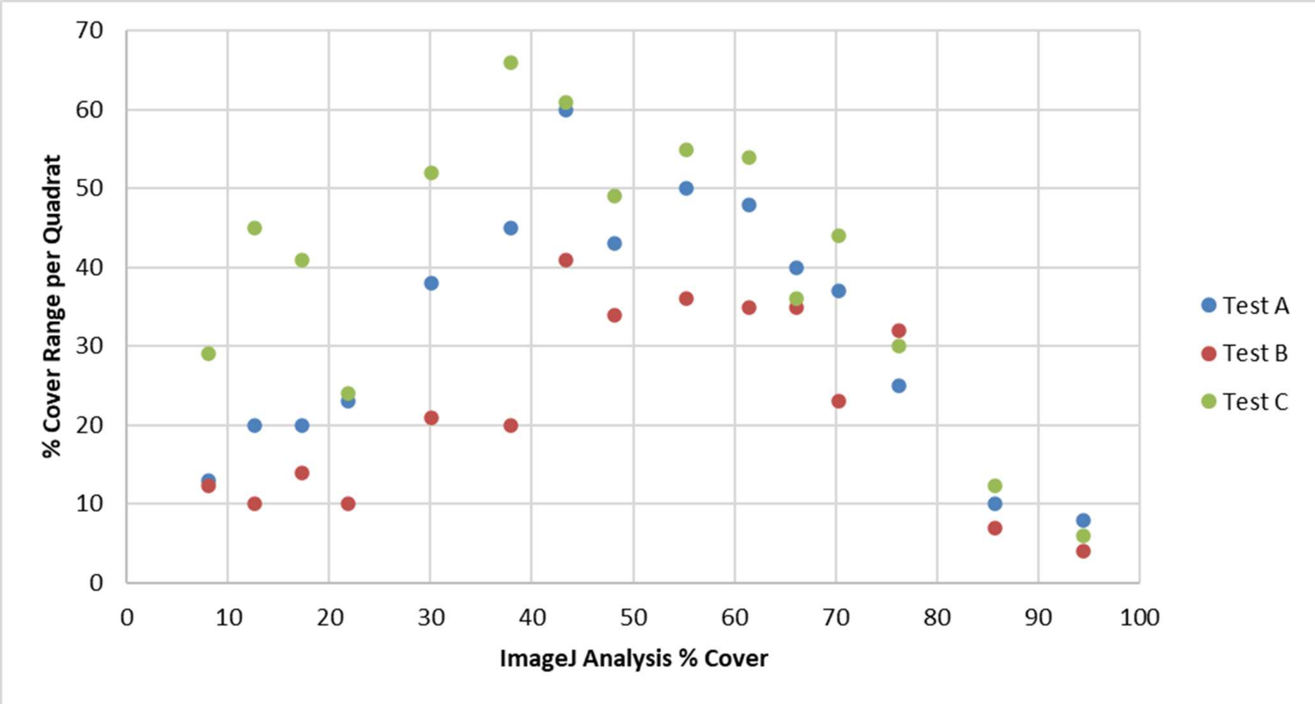

Figure 2 **Scatter graph showing the range of percentage cover results per quadrat across all three seagrass test methods**

There were noticeable differences both between seagrass and macroalgae results and between the different methods of estimation used and the resulting number of 'fails'. Since the numbers of participants varies both between seagrass and macroalgae tests and between the different methodologies, the chart in Figure 3 presents the numbers of fails as a percentage of the submitted results for each test to make them more comparable.

The macroalgae results show that for comparisons against the median the percentage of 'fails' was highest for Test B and lowest for Test A. Conversely, the comparisons against ImageJ results had the highest percentage of 'fails' for Test A and the lowest for Test C. The seagrass results had the lowest percentage of 'fails' for Test B and the highest percentage of 'fails' for Test A for both comparisons against the median and comparisons against the ImageJ results. The percentage of 'fails' was slightly higher for Tests A and C for comparisons against ImageJ compared to 'fails' against the median, whilst the percentages for Test B were the same for both comparisons. This contrasts with the previous z-score methodology where the number of 'fails' for the ImageJ comparisons were always much greater than comparisons against the mean.

The chart in Figure 3 also shows that the percentage of fails against ImageJ was consistently higher for macroalgae than for seagrass in Tests A and B. One possible reason for this is that the macroalgae results were generally more consistent for the macroalgae results than the seagrass, making it more evident when there are genuine outliers.

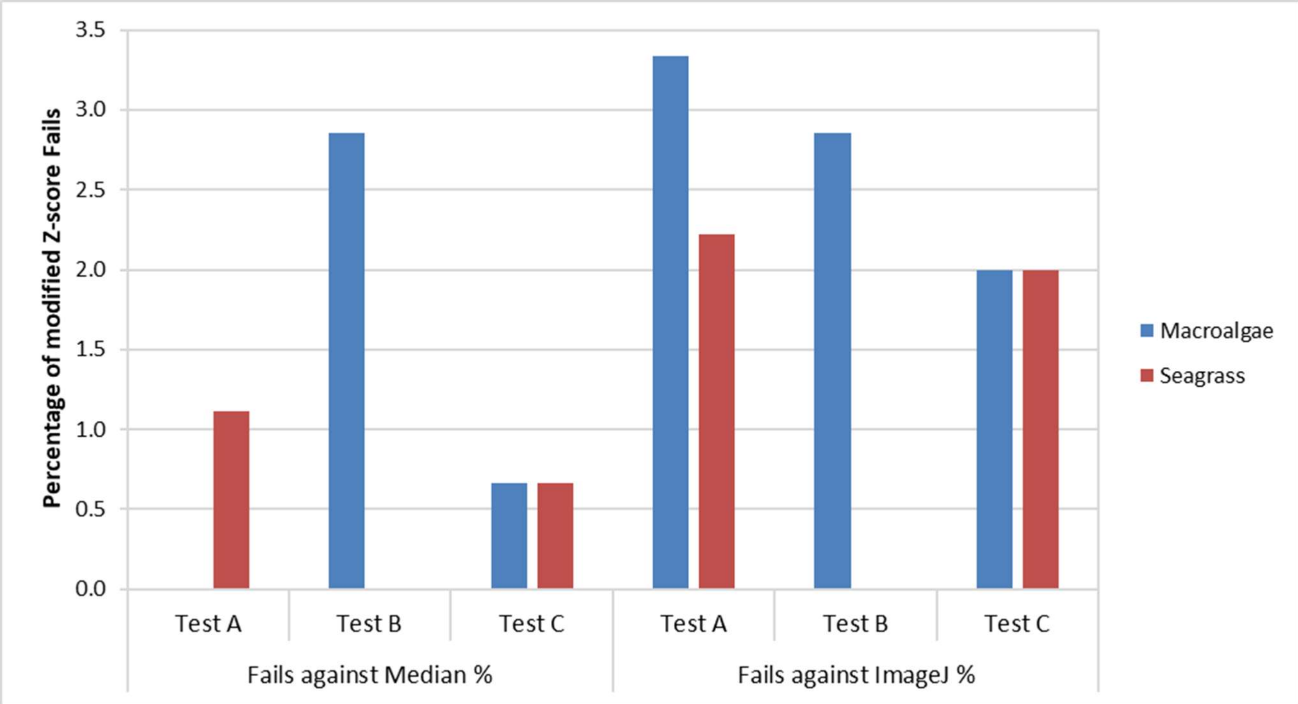

Figure 3 **Percentage of modified Z-score 'fails' recorded in each test for macroalgae and seagrass quadrats for OMC RT17.**

Direct comparisons between years are confounded by the change in pass/fail criteria for RT17, along with the inconsistent numbers of participants both between methodologies and between years. However, Figure 4 shows the results from the OMC RT16 using the previous Z-score methodology to highlight the differences in the percentages of 'fails' recorded. The percentage of 'fails' against the mean were two to three times higher for both macroalgae and seagrass in RT16 compared to percentage of fails against the median in RT17. The percentage of 'fails' against ImageJ for RT16 compared to percentage of fails against ImageJ in RT17.

The ranges of results provided for the macroalgae tests in RT17 were similar to those of RT16, with the exception of Test C which was skewed by the presence of one extreme outlier. In the seagrass results, the ranges for Test A and C were higher in RT17 than those recorded in RT16, whilst Test B ranges were slightly lower. The ranges are mostly higher than the recommended ± 5% between surveyors, with some quadrats having estimated ranges of up to 66% between participants (excluding the extreme range of 97.5% in the macroalgae, which is assumed to be the result of a calculation error). The average range per test varied between 10.37% and 18.5% for macroalgae and between 22.28% and 40.29% for seagrass.

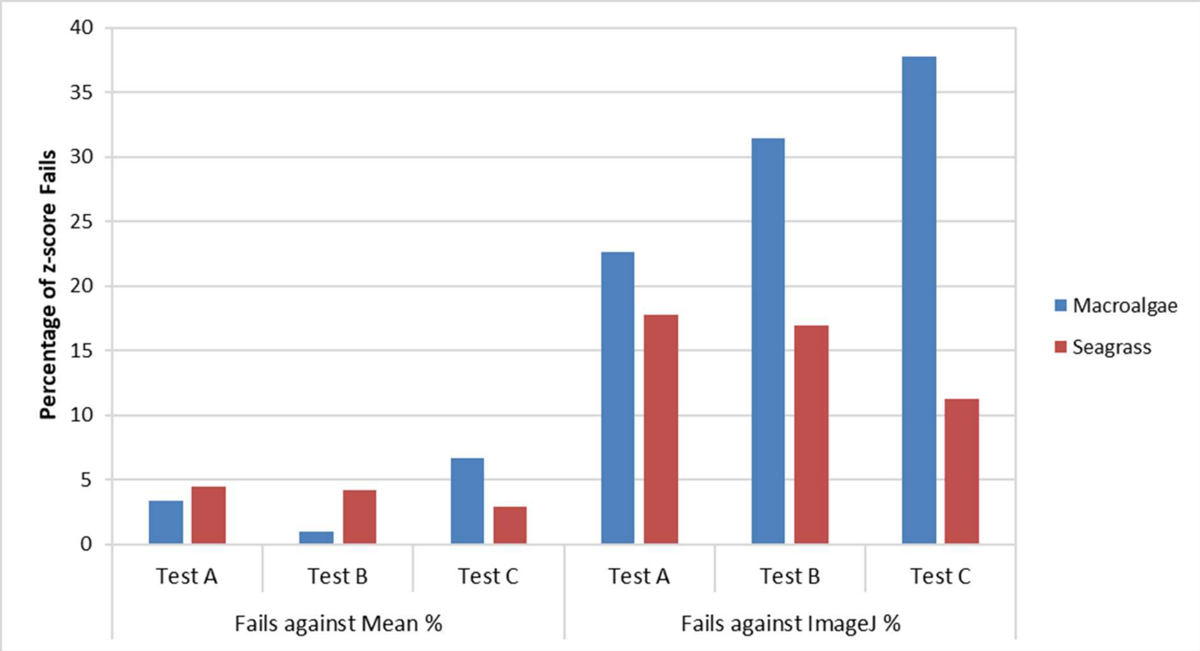

Figure 4 Percentage of Z-score 'fails' recorded in each test for macroalgae and seagrass quadrats for OMC RT16.

In RT17 Test C was the preferred methodology for the macroalgae and Test A was the most popular for the seagrass exercise. Test C has consistently been the most popular option for macroalgae since RT09, whereas seagrass shows more variation between years. 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 and this method had the greatest ranges for both macroalgae and seagrass, which suggests that this method is either not being used consistently between participants or that the use of an overlaid grid may obscure the photograph and make it more difficult to confirm whether there is algae/seagrass beneath.

Only three participants completed all three different test methods for macroalgae and two completed all three tests for seagrass. Even comparing results for these participants show no consistent patterns, with percentage cover scores and numbers of 'fails' still highly variable between individual participants and between the different methodologies for the same participant.

The broad range of results across participants and laboratories remains an ongoing problem. In-house workshop sessions within participating laboratories could prove useful to try and reduce the variability between individuals, along with trying alternative methodologies to see if one gives more consistent results than the others. Where one method may work best with macroalgae this may not be the case for seagrass.

5. Conclusions and Recommendations

1. The updated pass/fail methodology using modified Z-scores has reduced the number of ‘fails’ identified and reduced the disparity between the ‘fails’ identified using median values and those derived from the ImageJ results. The updated methodology was an effort to reduce the influence of standard deviation, which is easily skewed by one or two extreme values in small datasets. The updated methodology can be applied consistently to quadrats across the percentage cover range and provide a more robust metric to identify genuine outliers within the submitted percentage cover values that will still work in years with smaller sample sizes.

2. There is still a high degree of variation in results both between tests and between participants, particularly with the seagrass tests. It is not possible from the current ring test to conclude which percentage cover estimation method provides the most accurate results; however, method C unquestionably had the widest ranges of recorded values for both macroalgae and seagrass. This method remains the most popular for macroalgae, whereas method A was favoured for seagrass in RT17.

3. The image analysis method used during RT17 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 in case new image analysis or percentage cover apps become available. It is recommended at this time that participants should use the modified Z-scores derived from comparisons with the median if they are required for internal quality reports.

4. During this seventeenth cycle of the macroalgae percentage cover exercise nine of the ten laboratories completed 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 modified Z-scores, any results submitted outside of this deadline may not be accepted and it may not be possible to include them in the analyses. 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. This 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, particularly with regards to Test method C and to ensure that macroalgae and seagrass data are entered in the correct tab of the spreadsheet.

6. This year the test material was distributed via a shared Google Drive link. This method both reduces unnecessary postage and ensures arrival of the test materials on the designated start date. None of the participants reported access issues and all participants were able to download the required material successfully.

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 participants try to complete all three methods where time allows.

8. As with previous years, there was feedback suggesting that some quadrats could contain a mix of seagrass and opportunistic macroalgae to make the quadrats more representative of the natural conditions in which these species might be found. However, the results have frequently shown higher variability in results for seagrass estimation compared to macroalgae and this disparity would be obscured if quadrats contained mixtures of both and may even further confuse the results. 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 seventeenth 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

Iglewicz, Boris & Hoaglin, David C. (1993). How to Detect and Handle Outliers. *The American Society for Quality Control Basic References in Quality Control: Statistical Techniques Volume 16*. 85 pp.

Schiffler, Ronald. E. 1988. Maximum Z Scores and Outliers. *The American Statistician* 42 (1), 79-80.

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.