Marine Environmental Research 149 (2019) 157–169

Contents lists available at ScienceDirect

Marine Environmental Research

journal homepage: www.elsevier.com/locate/marenvrev

Using empirical and simulation approaches to quantify merits of rival measures of structural complexity in marine habitats

Saachi Sadchatheeswaran[a,*], Coleen L. Moloney[a], George M. Branch[a], Tamara B. Robinson[b]

[a] Department of Biological Sciences and Marine Research Institute, University of Cape Town, Rondebosch, South Africa
[b] Department of Botany and Zoology, Centre for Invasion Biology, Stellenbosch University, South Africa

ARTICLE INFO

Keywords:
Alien species
Benthic ecology
Coastal zone
Fractal
Rugosity
Structural complexity
Topography
Volumetric

ABSTRACT

Ecosystem engineers often affect structural complexity of habitats. There are multiple methods of quantifying complexity, variously measuring topography, surface area, volume, fractal dimension, or rugosity. We compared eight methods, four employing the 3D modelling program 'Blender' to estimate total surface area, top surface area, their ratio, and interstitial volume; and four empirically measuring interstitial volume, fractals and two indices of rugosity. We compared these using seven metrics: 1) correlations among comparable measures; 2) consistency; 3) accuracy; 4) precision; 5) discrimination among configurations of objects; 6) discernment of complexities among zones on rocky shores; and 7) practicality. Of the eight methods, the virtual volumetric method, Blender interstitial volume, performed the best. Direct measurements of three-dimensional space related more closely to patterns in biodiversity than did measurements of two-dimensional space or indirect measures of complexity like fractals. Blender interstitial volume is thus the recommended means of measuring structural complexity of benthic environments.

1. Introduction

Many marine ecological studies have demonstrated the impacts of structural complexity on organismal abundance and species diversity (Jones et al, 1994; Beck, 1998; Johnson et al., 2003; Gratwicke and Speight, 2005; Borthagaray and Carranza, 2007; Shumway et al., 2007; Gestoso et al., 2013), and the vast majority agree that an increase in complexity correlates with an increase in biological diversity. However, the methods and definitions of complexity vary, making it difficult to compare results. Difficulties are caused by 1) the use of different measures of complexity, 2) the array of different methods used even when the same measure is applied, and 3) the multiple terms and definitions used in describing complexity (McCoy and Bell, 1991; Tews et al., 2004; Kovalenko et al., 2012).

McCoy and Bell (1991) define 'habitat structure' as a combination of three parts: heterogeneity, complexity and scale, and 'complexity' as "variation attributable to the absolute abundance of individual structural components". Therefore, methods of measuring complexity can encompass Euclidean measurements of length, surface area and volume, or indirect indices of those measurements. Euclidean measurements are difficult to determine in the field, as most shapes in nature are irregular. In addition, Euclidean measurements are likely to be species-specific or object-specific (Frost et al., 2005). They are nonetheless important. Surface area affects recruitment, settlement and the turbulent exchange of materials such as water, nutrients and oxygen across the surface, whereas volume can be indicative of space available for shelter from physical stress and predators, and as feeding ground. Rugosity is a measure of the 'roughness' of the topography; high rugosity is likely to increase settlement of larvae and particles that get trapped while passively transported along the surface, whereas low rugosity decreases the amount and size of shelter (Parravicini et al., 2006). It is common for studies to utilize indices, either a ratio or a standardised variable, to measure complexity. However, these ratios of Euclidean measurements (e.g. surface area to volume ratios) are incapable of explaining the arrangement of structural elements within the space being measured (Kovalenko et al., 2012). Fortunately, the difficulties of measuring surface area or volume directly can be alleviated by assessing samples and measurements in a virtual environment.

Several papers have reviewed methods that measure complexity on rocky shores (Beck, 1998; Frost et al., 2005; Wilding et al., 2010), but so far none have compared methods in a standardised manner with controlled conditions, allowing an objective measure of their merits. Additionally, none have compared a wide array of methods that measure various aspects of complexity (e.g., surface area, volume, fractal

* Corresponding author. Department of Biological Sciences and Marine Research Institute, University of Cape Town, Private Bag X3, Rondebosch, 7701, South Africa.

E-mail address: saachi.sdc@gmail.com (S. Sadchatheeswaran).

https://doi.org/10.1016/j.marenvres.2019.03.014
Received 1 December 2018; Received in revised form 27 March 2019; Accepted 28 March 2019
Available online 04 April 2019
0141-1136/

dimension or rugosity). Using artificial, standardised spheres and mussel samples as objects, we compared eight methods of measuring complexity, three of which are novel and five that previously have been used to measure the complexity of mussel aggregations and other biotic communities on rocky shores. Our aim was to objectively and quantitatively compare the merits of alternative methods of measuring structural complexity created by rigid ecosystem engineers (e.g., barnacles, mussels or corals) in benthic habitats, in terms of a predefined set of criteria, using a combination of virtual modelling and empirical experiments. In particular, we introduce a virtual method known as 'Blender interstitial volume' that allows various measures of complexity and, as we will demonstrate, has the capacity to make retrospective estimations based on community composition.

2. Materials and methods

The eight methods we compared included four derived from the simulation program 'Blender' and four that rely on other means of measurement, as follows: 1) Blender measures of total surface area; 2) Blender measures of top surface area; 3) the ratio of total:top surface areas from Blender; 4) Blender interstitial volume, which uses the concept of virtual 'shrink-wrapping' of objects, in which a modifier is used to deform or 'vacuum-form' an object to the surface of a another, target object; 5) Displacement interstitial volume; 6) Fractal analysis; 7) Substrate rugosity index; and 8) Bidimensional rugosity index (Fig. 1; Table 1).

2.1. Description of methods

2.1.1. Measuring surface area

Three measures of surface area (total, top, ratio of total:top) were performed in Blender 2.74 (Blender Foundation, 2012), a 3D modelling program that digitally measures the first two aspects of surface area using the add-on 'Measure Panel'. Details of the program and its operation are covered in a companion methodological paper providing written and video instructions (Sadchatheeswaran et al., submitted). This add-on, which is not available in later versions of Blender, calculates the surface areas and volumes of irregular objects by measuring the scale and number of faces and points that make up the mesh of the object. Samples were modelled in Blender on a 10 cm × 10 cm plane. All the objects, except for the plane, were combined into one object prior to measurement. Total surface area of the combined object represented the total amount of substrate to which organisms may attach (Fig. 1a, Table 1). However, seaweeds and other autotrophic organisms are more likely to attach to the top of structures such as a mussel bed, and so top surface area was also calculated using Blender, by allowing Measure Panel to determine the surface area of the fraction of the object that is visible from above (Fig. 1b, Table 1). To do this, the vertices that could be viewed from the top were disconnected from the rest of the object and this top fraction was measured. The ratio of total:top surface areas was calculated as the proportion of the total substrate surface that was present on the top of the aggregation (Table 1). If the value was small, then top surface area accounted for most of the surface area, as would be the case for a unilayered mussel or barnacle bed. If the value was large, the sample was complex and multi-layered, with a proportionally small top surface area.

2.1.2. Measuring volume

Blender interstitial volume was used to calculate the volume of interstitial space between assemblages of objects, including organisms in a mussel bed, once again using the add-on 'Measure Panel' (Sadchatheeswaran et al. submitted). In short, the method involves creating a model of a sample of objects in Blender on a 10 cm × 10 cm plane (Fig. 1c, Table 1), joining all the objects into one object, calculating the solid volume of this single object and then 'wrapping' the object with the 'shrinkwrap' modifier. The volume of this shrink-

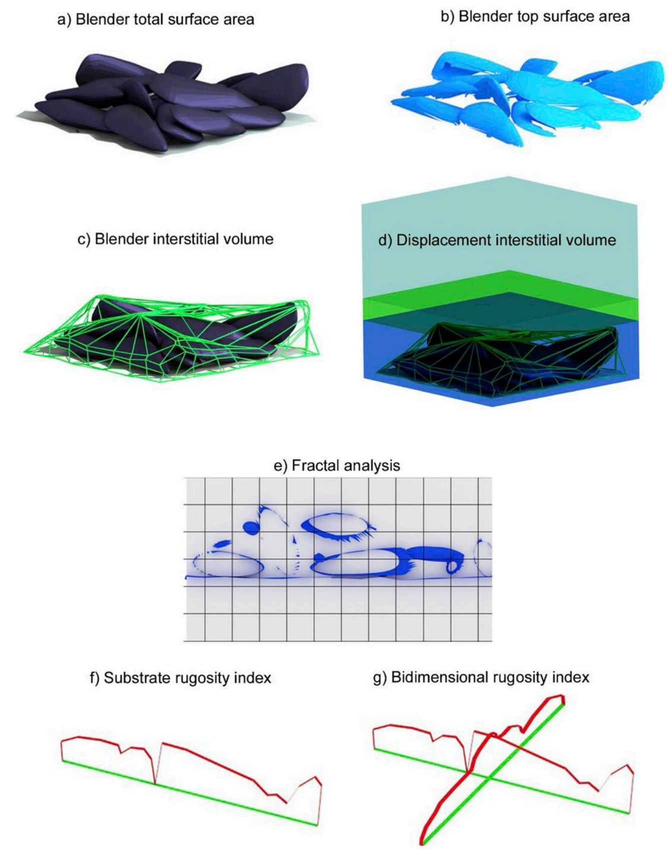

Fig. 1. Illustrations of seven of the eight methods used to calculate complexity. **a)** Blender total surface area (Method #1); **b)** Blender top surface area (#2); **c)** Blender interstitial volume (#4); **d)** Displacement interstitial volume (green water layer = displaced water volume from which the solid volume of the object is subtracted) (#5); **e)** Fractal analysis (outline of the bisected mussel sample on a grid of square cells) (#6); **f)** Substrate rugosity index (#7), **g)** Bidimensional rugosity index (#8). In (c and d), the green wireframe surrounding the sample represents the shrink-wrapped volume of an object. The two rugosity measures (f and g) are the ratio of contour length (red) to linear length (green). Figure omits ratio of total:top surface area (a/b) (Method #3). See Table 1 for more information. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

wrapped object was expanded and scaled up by 0.5× on the top and all sides within Blender (to allow inclusion of substrate) to obtain the 'expanded shrinkwrap volume'. Effectively, this amounted to multiplying the shrinkwrap volume by 1.27. The difference between the shrinkwrap volume and the solid volume was calculated to get the 'expanded interstitial volume', henceforth referred to as 'Blender interstitial volume' (Fig. 1c, Table 1).

Displacement interstitial volume is a physical method developed by Tsuchiya and Nishihara (1985) and adapted by Van Dover and Trask (2000) (Fig. 1d). Following this method, a physical sample was taken, wrapped tightly in plastic and then lowered into a plastic graduated cylinder (10 ml gradations) filled with 1000 ml of water. The volume at the raised water level was recorded when the sample settled to the bottom of the cylinder. The displaced volume was the volume after introduction of the sample, minus 1000 ml. The sample was then unwrapped and measured again in the same manner to determine the 'solid volume' of the sample. The difference between the plastic-wrapped displaced volume of the sample and the sample's solid volume was the interstitial volume (Table 1).

2.1.3. Measuring fractal dimension

Fractals were first applied and defined by Mandelbrot (1977, 1983)

Table 1
Methods used to measure sphere configurations, and the measurements, description and formulae required for each. In the empirical surface area and volumetric measurements, n_L is the number of large spheres (radius of 2.5 cm), n_S is the number of small spheres (radius of 1.5 cm), x_L is the fraction of large sphere that was visible from above while in configuration, x_S is the same but for small spheres, n_r is the number of rectangular prisms required to make shrinkwrap, $l_r w_r h_r$ are the length, width and height respectively of each rectangular prism in shrinkwrap. For each grid used to calculate fractal dimension, C is the number of cells that intersect with the cross-sectioned outline of the configuration, when cut diagonally between two opposite corners, δ is the length of cell that makes up the grid. For rugosity measures where diagonals are between two opposite corners of the configuration, Ch is the contour length of diagonal cross section of configuration (cm), Ta is the linear length of the diagonal (cm), b is the weight of objects (g), and n is the number of objects. See Fig. 1 for more information.

Method	Measurement	Description of measurement	Formula
1) Blender total surface area	Total surface area (cm^2)	Total surface area of the absolute number of large and small spheres in configuration	$n_L(4\pi(2.5)^2) + n_S(4\pi(1.5)^2)$
2) Blender top surface area	Top surface area (cm^2)	The surface area of the configuration, limited to what can be viewed when looking at the top of the configuration from above	$[(n_L(4\pi(2.5)^2))x_L] + [(n_S(4\pi(1.5)^2))x_S]$
3) Blender total:top surface area	Total/top surface area (unitless)	Ratio of total and top surface area of configuration	$\frac{[n_L(4\pi(2.5)^2) + n_S(4\pi(1.5)^2)]}{[(n_L(4\pi(2.5)^2))x_L] + [(n_S(4\pi(1.5)^2))x_S]}$
4) Blender interstitial volume	Shrinkwrap volume (cm^3)	The volume of the least number of rectangular objects required to encase configuration tightly	$\sum_{r=1}^{n_r} l_r w_r h_r$
	Interstitial volume (cm^3)	The collective volume of gaps and spaces among spheres within a configuration	$(\sum_{r=1}^{n_r} l_r w_r h_r) - \left[n_L\left(\frac{4}{3}\pi(2.5)^3\right) + n_S\left(\frac{4}{3}\pi(1.5)^3\right)\right]$
	Expanded interstitial volume (cm^3)	Interstitial volume subtracted from shrinkwrap volume, in addition to a small amount of space immediately surrounding the spheres to account for substrate attachment. Standard measurement of Blender interstitial volume utilized in this paper	$1.27(\sum_{r=1}^{n_r} l_r w_r h_r) - \left[n_L\left(\frac{4}{3}\pi(2.5)^3\right) + n_S\left(\frac{4}{3}\pi(1.5)^3\right)\right]$
5) Displacement interstitial volume	Interstitial volume (cm^3)	The collective volume of gaps and spaces within the configuration	$(\sum_{r=1}^{n_r} l_r w_r h_r) - \left[n_L\left(\frac{4}{3}\pi(2.5)^3\right) + n_S\left(\frac{4}{3}\pi(1.5)^3\right)\right]$
6) Fractal analysis	Fractal dimension (unitless)	The negative slope of the ratio of the number of cells in a square grid intersected by the bisected outline of the configuration, and the side length of one cell in a grid, summed over a set of grids (cm) (Sugihara and May 1990)	$-1\left(\text{slope}\left(\frac{\sum logC}{\sum log\delta}\right)\right)$
7) Substrate rugosity index	Substrate rugosity index (unitless)	Ratio of diagonal contour length and diagonal linear length of configuration. Lengths are measured in cm (Risk, 1972)	$\frac{Ch}{Ta}$
8) Bidimensional rugosity index	Bidimensional rugosity index (unitless)	Ratio of substrate rugosity index using ratio of both diagonal lengths of the configuration (in cm) to the total weight and number of objects. (Gestoso et al., 2013)	$\frac{((Ch1xCh2)/(Ta1xTa2))}{(b/n)}$

as complex geometric shapes that are composed of self-replications of one shape at, ideally, infinite scales. Over time, Fractal analysis has been used to measure the fractal dimensions of habitats and organisms in multiple fields of ecology, including mussel beds and rocky intertidal shores (Snover and Commito, 1998; Commito and Rusignuolo, 2000; Kostylev and Erlandsson, 2001; Frost et al., 2005). Unfortunately, there is no single preferred way to create profiles of samples to calculate fractal dimensions. In one study, aerial photographs were taken of the top of each sample and then the outlines of the mussels in each photograph were traced onto paper for later analysis (Snover and Commito, 1998). In another study a 10-cm thick layer of plaster of Paris, mixed with seawater, was poured over mussel beds at low tide and allowed to harden and cure in an oven; the surface of the plaster cast was coated in black graphite before the samples were cut into cross-sections 1.30 cm thick so that the profile of the mussel beds was easily visible (Commito and Rusignuolo, 2000). Because of time and financial constraints, and a decision to examine cross-sections of samples rather than just the top surface, virtual samples were created for Fractal analysis in our study (Fig. 1e).

To calculate the fractal dimension of a sample, two diagonal cross-sections of each sample were analysed in a series of sixteen 10 cm × 10 cm square grids created in Photoshop CS. From the top view, the first cross-section bisected the sample from the top left corner to the bottom right corner, and the second cross-section bisected the sample along the opposite diagonal. These grids were made up of an increasing number of square cells (Fig. 1e). The number of cells (y) in each grid was $y = 2^{2x}$ where x ranged from 0 to 15, and so the length of each cell (δ) per grid decreased with an increase in the number of cells. Each cross-section was placed behind each grid and the number of cells (C) that intersected with the outline of the sample was counted manually. Fractal dimension was calculated by finding the ratio of the total sum of $logC$ over the 16 grids, and the total sum of $log\delta$ of the 16 grids, calculating the slope of this ratio and multiplying the result by −1 (Table 1).

2.1.4. Measuring rugosity

Risk (1972) developed a widely used measure of complexity, the Substrate rugosity index, otherwise known as the 'chain and tape' method. It calculates rugosity as a ratio of the contour length and the shortest distance between the beginning and the end of the contour length (Fig. 1f, Table 1). Gestoso et al. (2013) adapted this method into the Bidimensional rugosity index, the product of two Substrate rugosity indices from samples of mussels, standardised by the mean weight of the dried mussel shells (Fig. 1g, Table 1).

2.2. Calculating control measures of complexity

'Control samples' from which absolute measures could be obtained were built with virtual spheres of defined sizes in Blender 2.74, so that the surface areas, volume, fractal dimension and rugosity for the control measures of complexity could be calculated exactly using formulae (Fig. 2). In these control samples, two aspects were considered: (1) the effects of homogeneous or heterogeneous mixes of objects, and (2) the packing of objects. The first aspect was addressed by creating five configurations that fitted on 10 cm × 10 cm panels, drawn from combinations of large spheres (radius 2.5 cm) and small spheres (radius 1.5 cm). The configurations were named 'Large' (comprising six large spheres), 'Small' (25 small spheres), 'Mix 1' (a mix of six large spheres and six small spheres), 'Mix 2' (three large spheres and 15 small spheres) and 'Mix 3' (three large spheres and 18 small spheres) (Fig. 2). The rationale for the choice of configurations was based on the arrangements of small and large organisms on the shore, as found in surveys on the rocky intertidal shores of Marcus Island, on the west coast of South Africa (Robinson et al., 2007; Sadchatheeswaran et al., 2015, 2018), which recognised seven zones on the shore, termed 1, 2, 3a, 3b, 4, 5 and 6 from the top to the bottom of the shore. In this paper, we consider two high-shore zones (Zones 2 and 3a) in which there were numerous small (0–30 cm) mussels (mimicked by 'Small'), and a third zone at the bottom of the shore (Zone 6) where there were aggregations

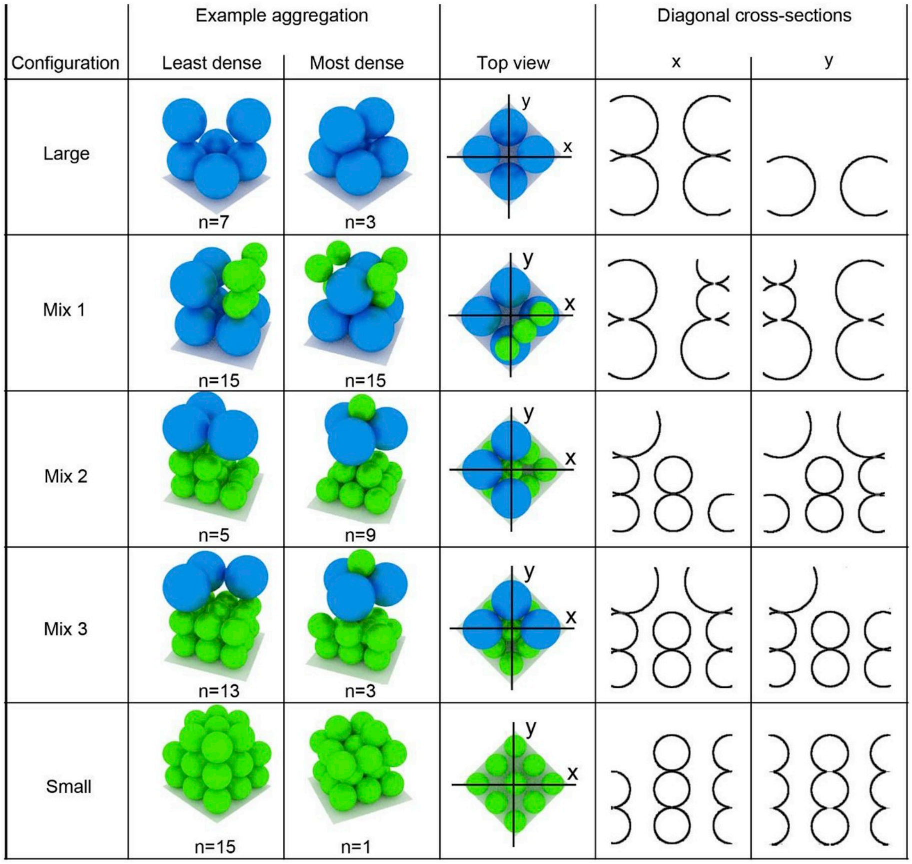

Fig. 2. Control configurations (on 10 cm × 10 cm panels shown by grey shading) illustrating examples of samples packed in the least and most dense aggregations. Top view of least dense example shown with diagonal axes for cross-sections (x,y). Cross-sections are presented as outlines that fit onto a 10 cm × 10 cm plane, and are examples of what were used to calculate the contour lengths for rugosity measures, and the fractal dimensions for Fractal analysis. The numbers of samples (n) for least and most dense aggregations are indicated.

of few but equal numbers of large and small mussels (equating to 'Mix 1'). The 'Large' configuration was included for its relevance to areas where only large mussels are found. Mix 2 and Mix 3 were variations of Mix 1 and were included to assess if methods could discriminate differences in complexity among superficially similar heterogeneous aggregations.

The second aspect of complexity, packing arrangement, was addressed in the control configurations by creating several, rigidly defined packing arrangements within 10 cm × 10 cm x 10 cm boxes using Blender. The control spheres were aggregated in several ways to allow variability (standard errors) to be calculated for the estimates (mean values) of the complexity measures, but the variation in packing was limited by a set of rules: 1) the top layer of spheres needed to be at least partially in the box; 2) samples were symmetrical around at least one axis of the box; 3) all samples within each configuration had to be unique; and 4) spheres had to touch at least one other sphere. Spheres were packed in either a 'least-dense' or a 'most-dense' aggregation to create variation in as many of the complexity measures as possible. Least-dense samples included spheres that rested on the apex of the sphere below, whereas the most-dense aggregations were packed so spheres were as close to each other as possible (Fig. 2). Without violating any of the rules above, each of the five configurations had up to 15 repetitions of least dense samples and 15 of most dense samples, for a possible maximum total of 30 samples per configuration (Fig. 2). To avoid false repetition of complexities during control calculations, equal sample sizes among different configurations could not be achieved. In particular, there was only one configuration possible for 'Small, most dense' that followed the four rules listed above. Sample sizes are listed (per configuration and split between least and most-dense) in Fig. 2.

All eight methods of measuring complexity were applied to the five sets of sphere configurations. For the four methods that used Blender to calculate interstitial volume, top and total surface areas (and their ratio), and for the displacement method, complexity measures could be calculated directly using the formulae presented in Table 1. Fractal dimension was particularly laborious to calculate and so, for logistical reasons, only three samples per configuration were calculated. Bidimensional rugosity index required a mean dry weight of each object in the sample. Since the virtual spheres in Blender have no weight, the relative volume of Large to Small spheres (4.63:1) was used to generate weights. The Small spheres were given a weight of 20 g, which was the weight of a clay sphere 1.5 cm in radius. The Large spheres were given a weight of 92.60 g (4.63 × 20 g).

2.3. Criteria for assessing methods of measuring complexity

Seven criteria were employed to quantify or rank the usefulness of the different measures of complexity: 1) correlation among methods dealing with common measures (surface area, volume, fractal or rugosity); 2) consistency between control calculations and test measurements; 3) accuracy and 4) precision of test measurements; 5) ability to discriminate among different configuration samples; 6) ability to discern complexity among configurations of objects such as mussel samples from different zones, and 7) practicality (time, cost, difficulty) of the method.

2.3.1. Correlation among methods

On the premise that methods yielding comparable measures (i.e., different measures of surface area, volume, fractal indices or rugosity) should have outputs that are correlated, Pearson product-moment correlations between methods were calculated in R, and their significance determined. The Holm-Bonferroni correction calculator ver. 1.2 by Gaetano (2013) was utilized to avoid Type 1 errors. The percentage of significant correlations for each method was taken as an index of correspondence among methods. As Fractal analysis was the only method used to measure fractal dimension, the two diagonal cross-sections per sample were compared against each (with three samples

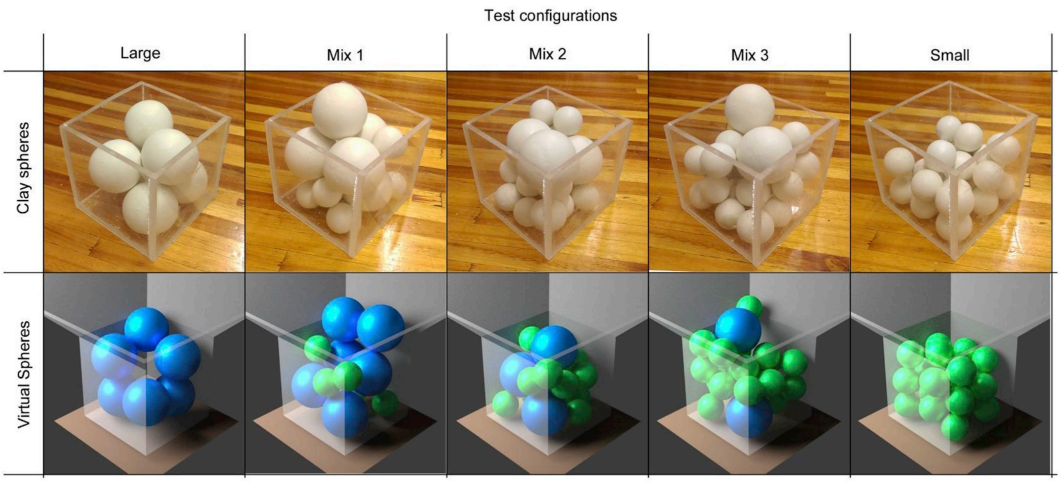

Fig. 3. Examples of test samples using clay and virtual spheres. The clay spheres were used to measure Displacement interstitial volume and both rugosity indices, whereas the virtual sphere samples were used for all other methods.

per cross-section).

2.3.2. Consistency

The methods were each applied to randomized 'test' samples of Large, Mix 1, Mix 2, Mix 3 and Small configurations. To randomize samples for Fractal analysis and for methods performed in Blender, the appropriate numbers and sizes of virtual spheres corresponding to each configuration were tumbled into the same 10 cm × 10 cm x 10 cm box used in the control configuration (Fig. 3). To randomize samples for methods performed in the real world (Displacement interstitial volume, Substrate rugosity index and Bidimensional rugosity index), clay spheres with an aluminium and glass core were created with the correct radii. The clay spheres were tumbled in a bag and deposited into a 10 cm × 10 cm x 10 cm box made of Plexiglas (Fig. 3). All methods, real and virtual, were performed on 15 test samples. Once again, the fractal dimensions of the two diagonal cross-sections per sample were compared against each other for Fractal analysis.

Assuming that differences in complexity among configurations emerging in the controls should be repeated in the tests, Pearson product-moment correlations, calculated in R, were calculated between control and test samples for each method, and the magnitude of the correlation taken as an index of consistency of each method.

2.3.3. Accuracy

The differences in values generated for complexity of sphere aggregations between control and random test samples for each method were compared among the different configurations, using univariate analyses (see *2.4* below). Averaged across the five configurations, mean (±SE) differences between control and test measures were taken as an index of accuracy. These measures indicated the percentage divergence (D) between results, with negative differences reflecting under-estimation of values by the test relative to the control, and positive values indicating over-estimations. Small values reflected good agreement. As all other means of assessing methods yielded large values for 'good' results, values of divergence were converted to ones of convergence ($D' = 100 - D$) to make final comparisons among the various assessments.

2.3.4. Precision

Mean precision of both control and test measurements was assessed by calculating the coefficients of variation (CV) for each method. Large CVs reflected a lack of precision in measurements, so CV was converted to a measure (CV') that increased with precision: $CV' = 100 - CV$. This facilitated comparisons among the various assessments in an overall evaluation of methods.

2.3.5. Discrimination among configurations

It was not possible to test the power of different methods statistically because of the small number of methods involved. Instead, comparisons were made of the capability of each method to distinguish among configurations in post-hoc tests. For each of the control and test runs, five configurations were examined, so that ten combinations could be examined to explore whether each method statistically distinguished between pairs of configurations. The percentage of cases (out of 10) that successfully distinguished pairs as being significantly different was taken as an index of 'discriminatory power'. This approach is not a formal assessment of statistical power but does provide an objective means of ranking the relative capabilities of the various methods in terms of their ability to discriminate among the complexities of different configurations.

2.3.6. Discernment among zones

Based on the fact that most studies have found a correlation between complexity and biotic diversity (referenced in the Introduction), we compared measures of diversity at different shore heights and tested the extent to which different methods of measuring complexity yielded correlations between complexity and diversity. In field-based assessments at Marcus Island on the west coast of South Africa, six 10 cm x 10 cm quadrats were randomly placed in areas with 100% cover of the Mediterranean mussel, *Mytilus galloprovincialis,* in each of three intertidal zones (Zones 2, 3a and 6; see above for definitions). In previous surveys of these zones (Sadchatheeswaran et al., 2015, 2018) greater abundance and species diversity were found in the low shore (Zone 6) – in contrast to the zones on the high shore (Zones 2 and 3a), which had statistically comparable lower values of abundance and diversity. Zones were therefore taken as a proxy for diversity, and the various methods were examined to see if they 1) had the capability to statistically discern among the complexities of Zones 2 and 3a versus Zone 6, and 2) did so in a manner reflecting the increased abundance and diversity in Zone 6 relative to Zones 2 and 3a.

The contour and linear lengths of each sample were recorded for both rugosity measures (Rugosity index and Bidimensional rugosity index). For these methods, a chain of sufficiently small chain links (2 mm) was used to avoid under-estimation of topography height relative to linear dimensions. The quadrats were then stripped, and the samples returned to the laboratory, where measurements of complexity were undertaken using the Displacement interstitial volume method. The mussels were then counted and dried in an oven for 48 h at 60 °C, opened and cleaned of mussel flesh and epifauna, and weighed to provide additional variables necessary to calculate the Bidimensional rugosity index.

The samples were recreated in Blender, using abundance and size data derived from the sampled mussels. Total surface area, top surface area, the ratio of total:top surface area, and interstitial volume were measured in Blender. Finally, the virtual samples were bisected to create two diagonal cross-sections for Fractal analysis (see *2.1.3 Measuring fractal dimensions*). Fractal analysis therefore had 12 samples per zone, whereas all other methods had six.

The differences among the complexities of mussel samples found in the three zones were compared for each method using univariate analyses, and it was assumed that higher values calculated with each method represented greater complexity. Thus, methods that correctly revealed that Zone 2 and 3a mussel samples were similar to each other in complexity (as expected from patterns in the field) but did not distinguish either of them from Zone 6 mussel samples were scored 1/3 or 33%. Methods that yielded the same outcome for Zone 2 and 3a but incorrectly indicated that they were significantly more complex than Zone 6 samples, were scored 2/3 or 67%. Lastly, methods that correctly rated mussel samples from Zone 2 and 3a as being similar to each other, and also correctly rated them as less complex than Zone 6 mussel samples were given a 3/3, or 100%.

2.3.7. Practicality

To determine the relative practicality of methods, the time and cost required to perform each method were compared, as well as the difficulty of performing each method on the shore or in the laboratory. As difficulty was a subjective measure, it was determined by the authors individually and then a collective view was obtained by consensus. From qualitative measures, the practicality of each method was then ranked out of 100%, with a score of 100 being most practical.

2.4. Statistical analyses

Univariate analyses were conducted using STATISTICA Version 12 with α set at 0.05. Prior to analyses, normality was assessed with normal probability plots, and homogeneity of variances was considered acceptable if the ratio of the largest and smallest variance was ≤ 4 (Quinn and Keough, 2002). Data that met assumptions of normality were analysed with a single-factor ANOVA with Tukey post-hoc comparisons, whereas for non-normal data, Kruskal-Wallis with multiple comparisons of mean ranks was used.

3. Results

3.1. Correlations among methods

The measures of complexity generated by each of the eight methods for the control samples were significantly correlated in only eight of the possible 36 combinations (Fig. 4). However, methods addressing comparable measures (surface area, interstitial volume, fractals or rugosity) exhibited strong, positive correlations (Fig. 4). Within the cluster of 'surface area' measurements, both total and top surface areas were significantly correlated with the ratio of total:top, the first positively and the second negatively. Both are logical outcomes anticipated from their relationships. However, total and top surface area calculations were not significantly correlated. The two interstitial volumetric measures (Blender and Displacement) had a strong positive relationship ($r = 0.99$). The fractal dimensions of the two cross-sections per sample were strongly and significantly correlated ($r = 0.99$), as were the two rugosity measures ($r = 0.80$). Curiously, the fractal dimension of one of the cross-sections per sample was strongly correlated with total surface area, whereas the other was not. Substrate rugosity index was positively correlated with Blender top surface area while Bidimensional rugosity index had a significant correlation with total surface area (Fig. 4).

In short, within different categories of measurement (surface area, volume, fractals or rugosity) there were strong correlations, but among these categories, there were few.

3.2. Consistency

For each method, average values of complexity for both control and test analyses are summarised in Table 2. Of the eight methods, only three exhibited significant (positive) correlations ($p < 0.0001$ to $p = 0.04$) between control and test results: Blender total surface area, Blender total:top surface area and Blender interstitial volume (Table 2, Fig. 5a, c, d). Although univariate analyses among configurations could not be conducted for Blender total surface area due to a lack of variance in the data, it was clearly the most consistent method, with test results being almost indistinguishable from the control calculations (Fig. 5a). The Bidimensional rugosity index (Fig. 5i) yielded the fourth greatest correlation, which was, however, non-significant ($p = 0.07$). Other methods all had low, non-significant correlations, and the Substrate rugosity index (Fig. 5h) was unique in having a negative (but also non-significant) correlation ($r = -0.60$, $p = 0.29$) between control and test results (Table 2).

3.3. Accuracy

The most accurate methods were Blender total surface area, Blender interstitial volume and Fractal analysis (Fig. 6). The two remaining Blender surface-area measurements were moderately accurate, and the most inaccurate methods were Displacement interstitial volume and both rugosity measures (Fig. 6).

3.4. Precision

The most precise methods (those with lowest CVs) were Blender total surface area, and Fractal analysis (Fig. 7). The moderately precise methods were the remaining Blender surface area measurements, Blender interstitial volume and Substrate rugosity index, and the most imprecise method was Bidimensional rugosity index (Fig. 7).

3.5. Discrimination among configurations

The methods that were best able to statistically discriminate complexity among configurations were Blender total surface area (100% discrimination among all combinations) and total:top surface area (80%) (Fig. 5a, c respectively). The moderately capable methods were Blender top surface area (Fig. 5b), Blender interstitial volume (Fig. 5d), Displacement interstitial volume (Fig. 5e) and Bidimensional rugosity index (Fig. 5i). Least capable were Fractal analysis (Fig. 5f and g for the two cross-sections) and Substrate rugosity index (Fig. 5 h).

3.6. Discernment among zones

All methods indicated that mussel samples from Zones 2 and 3a had statistically comparable complexities, as predicted *a priori* (Fig. 8, Table 3), and were given a score of 1/3 based on this. Blender total:top surface area (Fig. 8c) and Bidimensional rugosity index (Fig. 8h) additionally indicated that Zone 6 was significantly different in complexity from Zones 2 and 3a, raising their scores to 2/3 (67%). Only Blender interstitial volume upheld the *a priori* prediction that the complexity in Zone 6 would be significantly greater than that in Zones 2 and 3a (Fig. 8d), increasing its score to 3/3 (100%).

3.7. Practicality

Table 4 summarises the criteria used to rate the practicality of the different methods and the scores assigned to them. Integrating time, cost and difficulty, the most practical method was Substrate rugosity index, as it took the least amount of time and equipment to perform; it was scored 100% (Table 4). Although 3D modelling experience is required for all methods using Blender, the techniques required were relatively simple to learn, the time required to complete the method

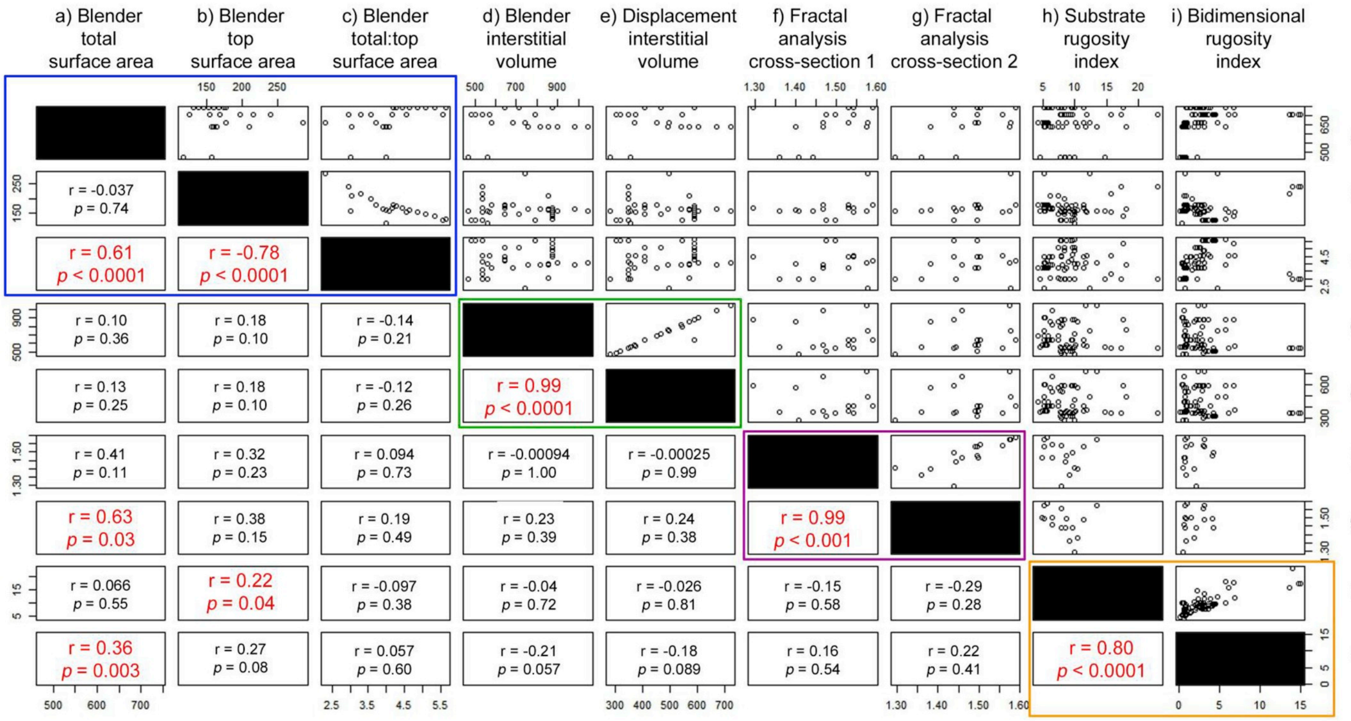

Fig. 4. Scatterplot matrix of control calculations of surface area (top left outline), volume (middle left outline), fractal dimension (middle right outline) and rugosity indices (bottom right outline) methods (based on all configurations pooled). Pearson product-moment correlations (r) and p-values shown for all correlations, with significant correlations noted in large font.

was minimal, and Blender 2.74 is a free-to-download program, and so all the Blender methods were given a score of 80%. These methods also have the added advantage of being retrospective in the sense that they can be applied to historical information, provided data on community composition exist. Bidimensional rugosity index was also given a score of 80%, despite the relatively high cost, due to the ease of accomplishing the method. Displacement interstitial volume was given a slightly lower score (70%) due to the loss of information about sample packing and increased possibility of human error. Fractal analysis was given the lowest score (40%) due to the time and cost required for determining the fractal dimensions of even virtual (and therefore easily manipulated) samples. In addition, there are multiple methods of determining fractal dimension, which means that the results of different Fractal analysis methods are often not comparable.

3.8. Summary of merits of alternative methods

Fig. 9 summarises the relative merits of the methods, scored according to the seven criteria. Among all methods, Blender interstitial volume was the most successful. Although Blender total surface area scored 100% in consistency, precision and discrimination among configurations, and 99% in accuracy, this method did not correlate well with top surface area and did not discern well among mussel samples measured in different zones. However, among all surface area methods, Blender total surface area performed the best, with total:top surface area performing second best.

Of the seven criteria used to compare the methods, Blender interstitial volume scored over 90% for five criteria, 80% for practicality, but only a modest 45% for its ability to discriminate among the different configurations of spheres. In combination, its average score of 86% was 24% more successful overall than Displacement interstitial volume, which only scored over 80% for precision and correlation among comparable methods.

Fractal analysis scored slightly better than Displacement interstitial volume, mostly due to its high accuracy and precision. However, pairs of measurements taken from the same sample did not correlate well, and the Fractal procedure proved the second least capable of discriminating among sphere configurations or among mussel samples taken from different zones, and was the least practical method.

Although the Substrate rugosity index and Bidimensional rugosity index were easy to implement in practice, they were not very successful at measuring complexity. This was mainly due to the low accuracy of both methods, low precision of the Bidimensional rugosity index, and the inability of Substrate rugosity indices to discriminate among different sphere configurations and mussel samples taken from different zones.

4. Discussion

4.1. Facets of measuring complexity

Despite several studies that compare methods of measuring habitat complexity (Frost et al., 2005; Kostylev et al., 2005; Wilding et al., 2010), a lack of generalization exists (Beck, 1998; Gestoso et al., 2013; Loke et al., 2014). The problem does not seem to be the use of synonyms, like 'structural complexity' and 'surface availability', but rather when one word, 'complexity', is used to describe multiple facets of habitat structure (Beck, 1998). This practice impedes comparisons among studies, masks interesting trends and reduces the capacity to recognise physical structure as an important environmental factor (McCoy and Bell, 1991; Warfe et al., 2008). Our paper set out to identify the most useful facet of complexity by quantitatively demonstrating which of the methods that measure complexity are most consistent, accurate, precise, correlate well with other methods under various conditions, and are practical.

There were very few correlations among methods, but strong, significant correlations did exist between methods considering comparable measurements, (i.e. those that respectively measured surface area, volume, fractal dimension or rugosity). Since methods for comparable measurements were sensibly correlated, none could be ruled out on

Table 2
Complexity (mean ± SE) of each sphere configuration (control and test) measured by each method (integrated across least and most dense aggregations). Consistency of each method is presented as Pearson's correlation between control and test results. Main effects of either ANOVA or Kruskal-Wallis (K-W) tests indicate statistical differences among configurations for both control and test results. Blender total surface area could not be analysed with univariate statistics because of zero variance. Unless indicated, measurements are unitless.

Method	Config.	Large	Mix 1	Mix 2	Mix 3	Small	Consistency (Pearson's r)	Main effects
Blender total surface area (cm^2)	Control	471 ± 0	641 ± 0	659 ± 0	744 ± 0	706 ± 0	r = 1.00,	N/A
	Test	467 ± 0	638 ± 0	654 ± 0	739 ± 0	701 ± 0	$p < 0.0001$	N/A
Blender top surface area (cm^2)	Control	153 ± 3.7	168 ± 4.7	198 ± 10	157 ± 3.3	151 ± 6.7	r = 0.31,	K-W, $H_{4,86} = 36.5$, $p < 0.001$
	Test	147 ± 1.3	164 ± 1.2	144 ± 1.2	144 ± 1.9	117 ± 0.36	$p = 0.61$	ANOVA, $F_{4,70} = 159$, $p < 0.0001$
Blender total:top surface area	Control	3.10 ± 0.09	3.85 ± 0.09	3.45 ± 0.14	4.78 ± 0.10	4.90 ± 0.17	r = 0.89,	ANOVA, $F_{4,81} = 25.0$, $p < 0.001$
	Test	3.18 ± 0.03	3.90 ± 0.03	4.55 ± 0.04	5.15 ± 0.00	6.00 ± 0.02	$p = 0.04$	ANOVA, $F_{4,70} = 17.9$, $p < 0.001$
Blender interstitial volume (cm^3)	Control	541 ± 12	886 ± 22	638 ± 16	748 ± 28	527 ± 9.8	r = 0.92,	K-W, $H_{4,86} = 68.4$, $p < 0.0001$
	Test	618 ± 20	757 ± 18	682 ± 19	663 ± 12	627 ± 11	$p < 0.0001$	ANOVA, $F_{4,70} = 10.7$, $p < 0.001$
Displacement interstitial volume (cm^3)	Control	343 ± 9.3	596 ± 17	415 ± 13	504 ± 22	340 ± 7.7	r = 0.64,	K-W, $H_{4,86} = 67.0$, $p < 0.0001$
	Test	122 ± 9.6	214 ± 8.9	179 ± 8.9	193 ± 10	198 ± 6.6	$p = 0.24$	ANOVA, $F_{4,70} = 14.7$, $p < 0.001$
Fractal dimension cross-section 1	Control	1.40 ± 0.02	1.48 ± 0.05	1.53 ± 0.03	1.55 ± 0.01	1.50 ± 0.02	r = 0.85,	K-W, $H_{4,16} = 4.29$, $p = 0.39$
	Test	1.39 ± 0.02	1.52 ± 0.01	1.52 ± 0.01	1.55 ± 0.01	1.59 ± 0.01	$p = 0.07$	K-W, $H_{4,75} = 42.8$, $p < 0.001$
Fractal dimension cross-section 2	Control	1.37 ± 0.04	1.47 ± 0.06	1.52 ± 0.03	1.53 ± 0.02	1.50 ± 0.03	r = 0.82,	K-W, $H_{4,16} = 5.84$, $p = 0.21$
	Test	1.45 ± 0.02	1.51 ± 0.02	1.54 ± 0.01	1.55 ± 0.01	1.61 ± 0.01	$p = 0.10$	ANOVA, $F_{4,70} = 17.3$, $p < 0.001$
Substrate rugosity index	Control	9.01 ± 0.77	8.93 ± 0.94	7.45 ± 0.70	8.87 ± 0.71	10.24 ± 0.67	r = −0.60,	ANOVA, $F_{4,81} = 1.90$, $p = 0.12$
	Test	2.68 ± 0.08	3.05 ± 0.11	3.08 ± 0.16	3.39 ± 0.13	2.49 ± 0.05	$p = 0.29$	ANOVA, $F_{4,70} = 17.2$, $p < 0.01$
Bidimensional rugosity index	Control	0.83 ± 0.17	1.64 ± 0.37	1.93 ± 0.34	2.85 ± 0.38	5.17 ± 0.68	r = 0 85,	K-W, $H_{4,86} = 44.6$, $p < 0.0001$
	Test	0.20 ± 0.02	0.32 ± 0.03	0.41 ± 0.02	0.40 ± 0.02	0.47 ± 0.04	$p = 0.07$	K-W, $H_{4,86} = 41.5$, $p < 0.0001$

grounds of disagreement among methods. The fact that few correlations existed among methods that did not share common measures is interesting, as these different types of measurements probably have different kinds of biological significance. Measures of surface area reflect attachment points on primary or secondary substrate (Crooks, 2002; Munguia et al., 2011). Measures of interstitial volume indicate space – for shelter from physical stresses and predators, and for feeding (Tsuchiya and Nishihara, 1985; Firstater et al., 2011). Rugosity is an index of 'roughness' that might influence small-scale water movements, influencing nutrient and oxygen transfer and propagule settlement (Parravicini et al., 2006). Fractals reflect the level of irregularity present in nature and organic structures, and this irregularity can lead to size preferences for individuals occupying the substrate or space (Sugihara and May 1990; Kostylev et al., 2005). Thus, different measures reflect various types of physical structures, and therefore are relevant to different types of biological responses. Selecting the optimal measurement will depend in part on what biological question is being addressed.

4.2. Evaluation of methods

The most accurate and precise method by far was one of the novel methods examined, Blender total surface area (Fig. 9). Only two papers have previously attempted to measure the surface area of rocky shores: one required the use of a proxy equation that calculated surface area from the squared average profile lengths obtained using a contour gauge (Kostylev et al., 2005); the other was similar but used triangulation to procure a three-dimensional measure of rugosity for 36 vertices measured per quadrat (Parravicini et al., 2006). Euclidean measurements overall, however, are not popular in the literature as indices or ratios. This is possibly because irregular objects, which are common in nature, cannot easily be measured using regular geometric equations. By using virtual samples and measurement tools available in Blender, it was feasible and quick to calculate the total and top surface areas of an aggregation. Furthermore, when this method of calculating the total surface area was applied to aggregations of spheres, it proved to be exact, and highlighted the fact that packing of objects in an aggregation is not a characteristic that can be detected by total surface area. Between the Blender top and total:top surface area methods, the latter was the preferred method for calculating complexity of two-dimensional spaces as it exhibited a strong positive correlation with total surface area, was precise, practical and was also statistically consistent (Fig. 9). Total:top surface area did not, however, correlate well with expectations driven by biodiversity on the rocky shores.

Displacement interstitial volume has been used previously to measure the three-dimensional complexity of mussel beds in the form of interstitial volume (i.e., the volume of the gaps between objects within aggregations) (Tsuchiya and Nishihara, 1985; Van Dover and Trask, 2000). Blender interstitial volume, a novel method to measure interstitial volume within a collection of objects, was introduced in Sadchatheeswaran et al. (2015). The difference in the two measurements is that interstitial volume is expanded in Blender to include a small amount of space surrounding the aggregation, to which epifauna can attach. For all but one of the criteria used to compare methods, Blender interstitial volume performed as well or better than Displacement interstitial volume. The relatively poor performance of Displacement interstitial volume was unexpected, given that this method involved the direct, real-world measurement of irregular volumetric gaps using water; it was expected to be at least as accurate and precise as Blender interstitial volume derived from a virtual state. Yet, unlike Blender interstitial volume, Displacement interstitial volume involved a greater number of steps, each of which could have been affected by human error. Blender was far more capable of measuring the exact interstitial volume of samples to a high number of significant digits. In addition, Blender interstitial volume was the only method that yielded greater complexities of mussel samples in the low shore compared to

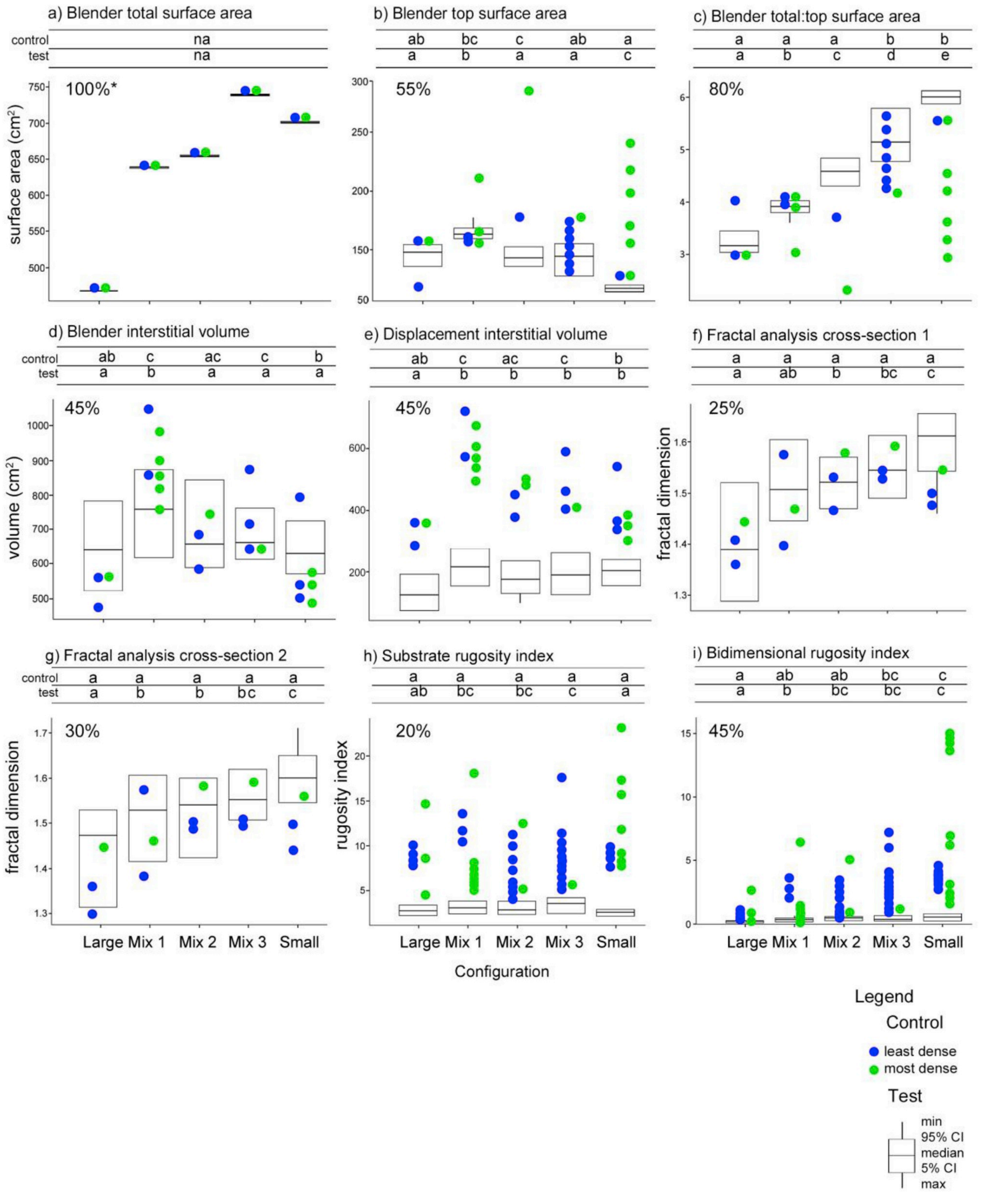

Fig. 5. Measurements of complexity recorded as scatterplots for control results, overlain with boxplots of test results. The complexities of the five configurations were compared against each other within each method. Letters differ where complexities were significantly different ($p < 0.05$) in post-hoc Tukey analyses (ANOVA) or multiple comparisons of ranks (Kruskal-Wallis). Percentages (top left corner of each graph) show the discrimination scores for the combined test and control configurations. *Blender total surface area had no variance, and so all configurations were effectively discriminated. See Table 2 for associated data.

the high shore on Marcus Island, and thus correlated with patterns of biodiversity. However, the results of both Blender and Displacement interstitial volume under control conditions were highly correlated, and so researchers working solely in the field may prefer to use the Displacement method to compare habitable space in highly three-dimensional, physically complex ecosystems.

Fractal analysis was very accurate and precise but was not as successful as direct measurements of surface area or volume. This, too, was

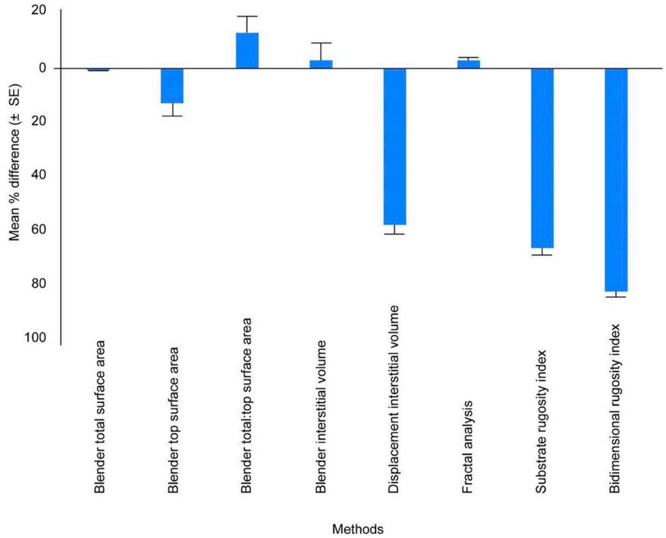

Fig. 6. Accuracy of methods, as indicated by mean (+SE) % difference between control and test results, integrated across all configurations.

a surprise, as Fractal analysis has been used on numerous occasions to measure the complexity of rocky shores and mussel aggregations (Beck, 1998; Lawrie and McQuaid, 2001; Kostylev et al., 2005) and is also commonly used in other fields of marine ecology (McCormick, 1994; Aronson and Precht, 1995; Brokovich et al., 2006). However, multiple papers that have used Fractal analysis to measure mussel aggregations do not agree on the actual fractal dimensions of such aggregations, most likely due to the variations in methods used to calculate the dimension (Kostylev et al., 2005). In our study fractal dimensions of the same sphere or mussel aggregation spanned a range of 0.2, which is a fifth of the range of fractal dimension (1.0–2.0). This diminished the precision of the method but boosted its accuracy. Moreover, Fractal analysis was the most difficult method to use in practise.

Another unexpected result was the poor performance of the Substrate rugosity index. This method, developed by Risk (1972), is most likely the oldest method for quantifying complexity. It was by far the most practical method to use in the field, and is particularly appropriate when sampling must be non-destructive and is time-limited, but it substantially underestimated rugosity when compared with control values. Bidimensional rugosity index, despite taking longer to perform, had one major advantage over Substrate rugosity index. When applied to test samples, it had greater powers of discrimination among configurations of spheres and demonstrated that aggregations made of many small spheres were more complex than aggregations made of a few large spheres, whereas the Substrate rugosity index failed to detect any difference between them. When applied to mussel samples collected in different zones on the shore at Marcus island, Bidimensional rugosity index also detected significant difference in complexity between Zone 6 and Zones 2/3a, which the Substrate rugosity index failed to do, but yielded lower values of complexity for the lowshore (Zone 6), than for the highshore (Zones 2 and 3a), contrary to the expectation of higher complexity values in Zone 6, based on patterns of biodiversity. Therefore, between the two rugosity measures, Bidimensional rugosity index was the more consistent method, but did not correlate well with higher indices of species richness and diversity in Zone 6 in comparison to Zones 2 and 3a.

As noted above, we recognise that different measures of complexity such as surface area and interstitial volume will be relevant to different biological and physical processes. Indeed, our measures allow determination of these various aspects of complexity, which can then be prioritised depending on the context and questions that are being addressed.

The approach we advocate works most easily with objects that are rigid, such as mussels, barnacles and corals. It is more difficult to apply to objects such as seagrasses, kelps or algae that are flexible. Movement of such objects will alter complexity, which will be very different if, say, seagrass is flattened by currents or held erect in calm water. However, this is a problem that is inherent in all methods of measuring complexity, so complexity is intrinsically easier to measure for rigid objects. It can be resolved by taking measurements of flexible objects in a range of states and then either expressing the results for specific states or, if sufficient simulations or measurements or done, providing average results. The practicalities of doing this will be formidable in methods that rely on field measurements, but would be easier to achieve in the simulations that are possible with Blender.

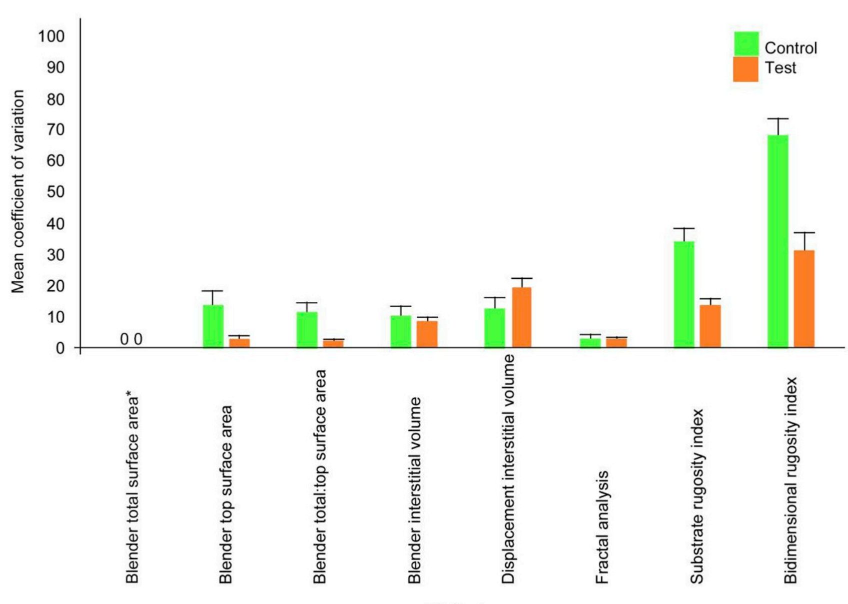

Fig. 7. Precision of each method, as indicated by the mean (+SE) coefficient of variation (CV) of control and test results, integrated across all configurations. Precision is inversely related to CV. *Blender total surface area had zero variance for both control and test results.

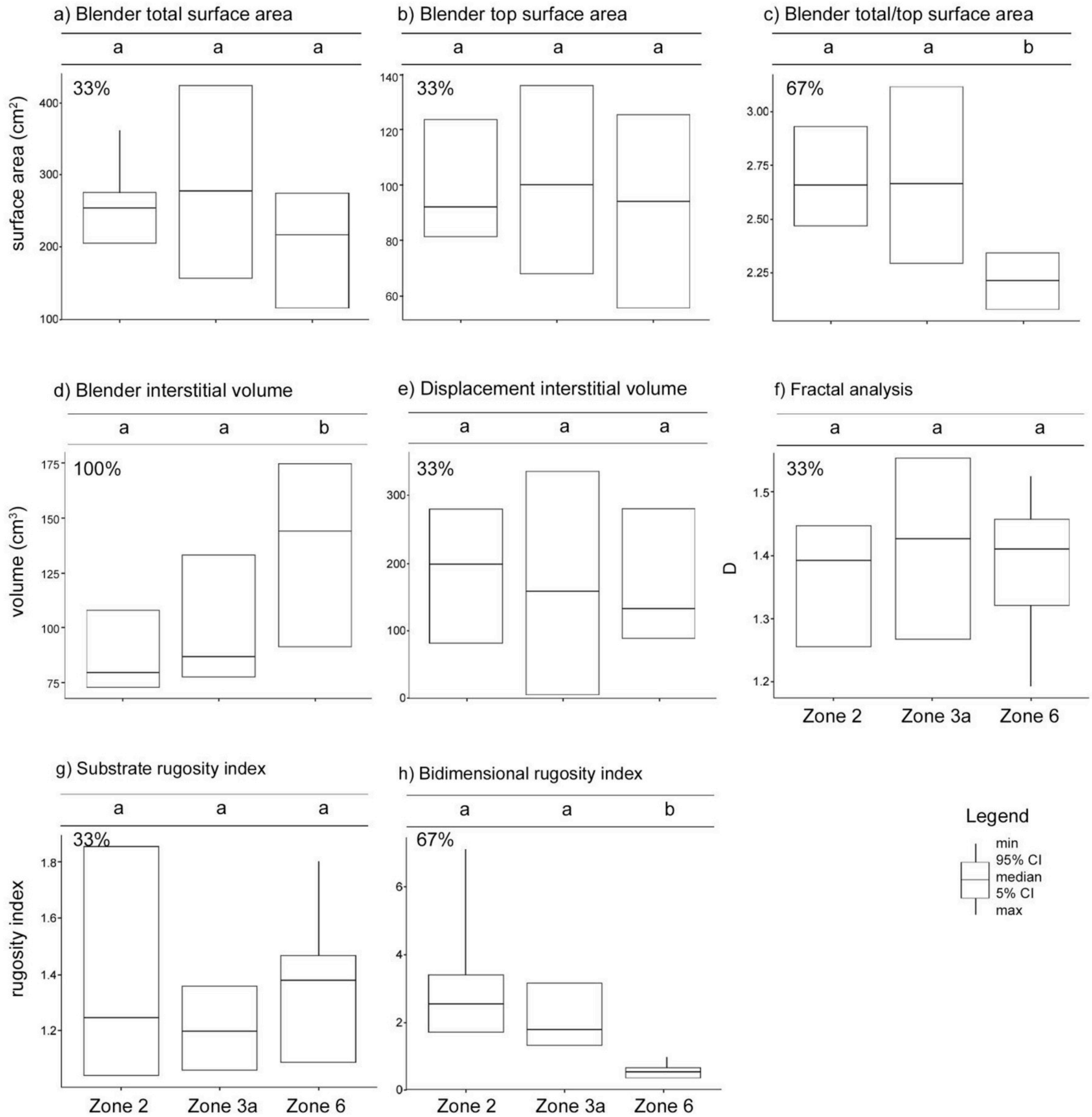

Fig. 8. Boxplots showing spread of complexity values measured by the different methods for mussel aggregations in three zones of Marcus Island. Letters differ where zones were found to be significantly different from each other in complexity ($p < 0.05$) based on post-hoc Tukey calculations (ANOVA) or multiple comparisons of ranks (Kruskal-Wallis). Percentages (top left corner of each graph) show discernment scores among the three zones.

5. Conclusion

Blender total surface area was the best method to measure the surface area of a sample, even though it did not correlate well with expectations based on diversity patterns. Bidimensional rugosity index was also statistically consistent, and therefore the preferred rugosity measure, although it scored the same as Substrate rugosity index when averaged across all seven criteria applied to the evaluation of methods. Bidimensional rugosity index was also the only recommended method that could be reliably performed on real samples on the shore.

Table 3
Mean (±SE) complexity of mussel aggregations as measured by the eight methods applied to three zones on Marcus Island. Main effects of statistical differences among zones are presented for each method and are bold if there were significant differences. Unless indicated, the measurement was unitless. See Fig. 8 for more information.

Method	Zone 2	Zone 3a	Zone 6	Main effects
Blender total surface area (cm^2)	261 ± 21.1	277 ± 23.1	200 ± 19.9	ANOVA, $F_{2,12} = 1.87$, $p = 0.18$
Blender top surface area (cm^2)	96.9 ± 5.77	101 ± 6.67	90.1 ± 8.65	ANOVA, $F_{2,12} = 0.40$, $p = 0.68$
Blender total/top surface area	2.67 ± 0.07	2.67 ± 0.09	2.21 ± 0.03	ANOVA, $F_{2,12} = 8.72$, $\boldsymbol{p = 0.0031}$
Blender interstitial volume (cm^3)	84.4 ± 4.89	94.7 ± 4.00	135 ± 10.9	ANOVA, $F_{2,12} = 7.14$, $\boldsymbol{p = 0.0067}$
Displacement interstitial volume (cm^3)	189 ± 31.5	163 ± 42.4	157 ± 26.4	ANOVA, $F_{2,12} = 0.16$, $p = 0.85$
Fractal analysis	1.38 ± 0.02	1.42 ± 0.02	1.39 ± 0.02	ANOVA, $F_{2,12} = 0.68$, $p = 0.51$
Substrate rugosity index	1.32 ± 0.11	1.20 ± 0.04	1.39 ± 0.05	ANOVA, $F_{2,12} = 1.03$, $p = 0.38$
Bidimensional rugosity index	3.14 ± 0.76	2.02 ± 0.25	0.57 ± 0.04	$K\text{-}W_{2,15} = 12.11$, $\boldsymbol{p = 0.0023}$

Table 4
Practicality of each method on the shore. Cost is in Euros (€) and does not include the cost required to get to the study area for field methods, the cost of a computer for virtual methods, or the cost of a drying oven (approx. 618 €). Each sample took 20 min to collect, count and measure on the shore. The methods performed in Blender include the time taken to recreate each sample. The time to complete Bidimensional rugosity index does not include time to dry mussels in an oven (48 h) or the extra 20 min per sample needed to clean the mussel shells.

Method	Time (min)	Equipment Cost (€)	Difficulty	Score
Blender methods			Pros ● Can obtain Euclidean measures of irregular objects easily ● Possible to perform multiple measurements on one sample very quickly and easily ● Able to select only part of the sample (example top surface) to measure it independently ● Decreased possibility of human error ● Samples could be saved and measured at leisure (i.e. could be used retroactively) ● Can be applied retrospectively Cons ● Requires some modelling experience ● Cannot measure actual samples on shore; recreated in virtual environment	
Blender total surface area (cm^2)	7.17 ± 0.50	Blender ver 2.74: 0.00 Total: 0.00		80%
Blender top surface area (cm^2)	8.17 ± 0.50	Blender ver 2.74: 0.00 Total: 0.00		80%
Blender total/top surface area (unitless)	8.17 ± 0.50	Blender ver 2.74: 0.00 Total: 0.00		80%
Blender interstitial volume (cm^3)	8.50 ± 2.17	Blender ver 2.74: 0.00 Total: 0.00		80%
Displacement interstitial volume (cm^3)	2.15 ± 0.32	paintscraper: 5.85 plastic bag: 0.16 1000 ml beaker: 13.00 Total: 19.01	Pros ● Very easy and requires very little material Cons ● Destroys packing of sample to make the measurement ● Human error; difficult to be accurate with standard equipment	70%
Fractal analysis (unitless)	24.1 ± 3.00	Blender ver 2.74: 0.00 Photoshop CS: 1235.00 Image J: 0.0 Total: 1235.00	Pros ● Used by multiple fields in ecology so universally accepted ● Cross-sections can be saved and measured at leisure ● Possible to collect sample and save its packing, or generate virtual samples in 3D-modelling program Cons ● Even virtual and automated, method took considerable time ● No standard method to acquire the images to be analysed ● Extremely tedious so even when partially automated, too many chances for human error	40%
Substrate rugosity index (unitless)	2.17 ± 0.75	Chain: 6.50 Ruler: 1.79 Total: 8.29	Pros ● Very easy and requires the least amount of material ● Can measure real samples on the shore	100%
Bidimensional rugosity index (unitless)	11.0 ± 1.50	Chain: 6.50 Ruler: 1.79 Digital balance: 71.5 Total: 79.8	Pros ● Easy to accomplish ● Can measure real samples on the shore Cons ● Takes very long (2 days drying in the oven, plus 20 min per sample to clean and weigh mussel shells) ● Requires laboratory equipment, so it cannot be applied solely in the field	80%

However, the results of both Blender and Displacement interstitial volume under control conditions were highly correlated, and so researchers working solely in the field may prefer to use the Displacement method to compare habitable space in highly three-dimensional, physically complex ecosystems.

Benthic environments such as intertidal rocky shores are complex, with various microhabitats influenced by various types of ecosystem engineers. These range from simple, unilayered beds of barnacles, to complex, multi-layered towers of some mussel species, and these differences in architecture sometimes require retrospective analysis. Blender interstitial volume scored higher than any of the other methods when assessed by multiple criteria, was the only method to correspond well with patterns of biodiversity, and together with other Blender methods, could be applied retrospectively, so it emerged as the most versatile and recommended method to measure complexity on rocky shores.

Animal ethics

Animals were collected under a permit issued by the Department of Agriculture, Forestry and Fisheries, and conformed to national and institutional ethics guidelines.

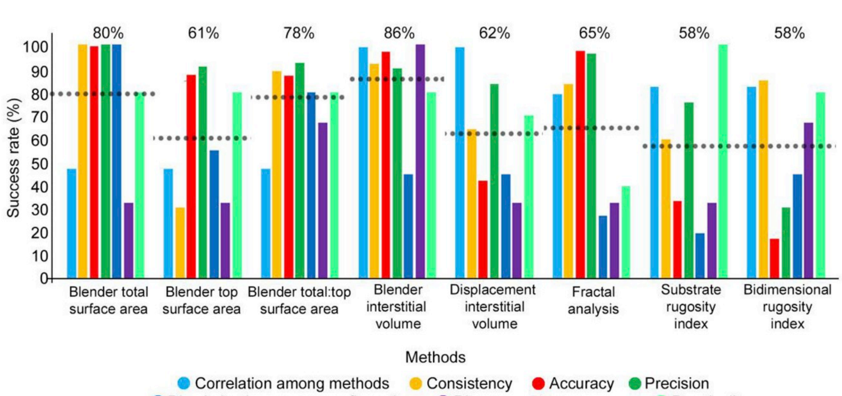

Fig. 9. Comparison of the success scores (%) of each method measuring complexity, integrated across configurations, with 100% being the maximum score for each of seven criteria: 1) mean correlations among like methods multiplied by 100; 2) consistency (r-values between control and test results) multiplied by 100; 3) accuracy; 4) precision; 5) discrimination among configurations; 6) discernment among zones; 7) practicality. Dotted horizontal lines along with the percentages above the bars indicate mean success rate over all seven criteria for each method.

Acknowledgements

Financial contributions from the Marine Research Institute, University of Cape Town (BASICS Programme), the Andrew Mellon Foundation, the South African Research Chair Initiative (funded through the South African Department of Science and Technology (DST) and administered by the South African National Research Foundation (NRF)), and the DST-NRF Centre of Excellence for Invasion Biology are gratefully acknowledged. None of the funders were involved in the design or execution of the research. A warm thank you to West Coast National Parks, Pierre Nel and Carmen Gagiano, for facilitating frequent trips to Marcus Island, as well as to Ignacio Gestoso for answering key methodological questions. Martin Emanuel and Stewart Norman are thanked for their help in the field. Thanks to reviewers for their constructive improvements.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marenvres.2019.03.014.

References

Aronson, R.B., Precht, W.F., 1995. Landscape patterns of reef coral diversity: a test of the intermediate disturbance hypothesis. J. Exp. Mar. Biol. Ecol. 192, 1–14. https://doi.org/10.1016/0022-0981(95)00052-S.

Beck, M.W., 1998. Comparison of the measurement and effects of habitat structure on gastropods in rocky intertidal and mangrove habitats. Mar. Ecol. Prog. Ser. 169, 165–178. https://doi.org/10.3354/meps169165.

Blender Foundation, 2012. Blender Ver 2.74. http://www.download.blender.org, Accessed date: 13 October 2012.

Borthagaray, A.I., Carranza, A., 2007. Mussels as ecosystem engineers: their contribution to species richness in a rocky littoral community. Acta Oecol. 31, 243–250. https://doi.org/10.1016/j.actao.2006.10.008.

Brokovich, E., Baranes, A., Goren, M., 2006. Habitat structure determines coral reef fish assemblages at the northern tip of the Red Sea. Ecol. Indicat. 6, 494–507. https://doi.org/10.1016/j.ecolind.2005.07.002.

Commito, J.A., Rusignuolo, B.R., 2000. Structural complexity in mussel beds: the fractal geometry of surface topography. J. Exp. Mar. Biol. Ecol. 255, 133–152. https://doi.org/10.1016/S0022-0981(00)00285-9.

Crooks, J.A., 2002. Characterizing ecosystem-level consequences of biological invasions: the role of ecosystem engineers. Oikos 97, 153–166. https://doi.org/10.1043/j.1600-0706. 2002.970201.x.

Firstater, F.N., Hidalgo, F.J., Lomovasky, B.J., Ramos, E., Gamero, P., Iribarne, O.O., 2011. Habitat structure is more important than nutrient upwelling in modifying mussel bed assemblage in an upwelling area of the Peruvian coast. Helgol. Mar. Res. 65, 187–196. https://doi.org/10.1007/s10152-010-0214-3.

Frost, N.J., Burrows, M.T., Johnson, M.P., Hanley, M.E., Hawkins, S.J., 2005. Measuring surface complexity in ecological studies. Limnol. Oceanogr. Methods 3, 203–210. https://doi.org/10.4319/lom.2005.3.203.

Gaetano, J., 2013. Holm-Bonferroni Sequential Correction: an EXCEL Calculator (1.2) [Microsoft EXCEL Workbook]. Retrieved May 28, 2017 from. https://www.Researchgate.Net/publication/242331583_Holm-Bonferroni_Sequential_Correction_An_EXCEL_Calculator_-_Ver._1.2.

Gestoso, I., Arenas, F., Rubal, M., Veiga, P., Pena, M., Olabarria, C., 2013. Shifts from native to non-indigenous mussels: enhanced habitat complexity and its effects on faunal assemblages. Mar. Environ. Res. 90, 85–95. https://doi.org/10.1016/j.marenvres.2013.05.015.

Gratwicke, B., Speight, M.R., 2005. The relationship between fish species richness, abundance and habitat complexity in a range of shallow tropical marine habitats. J. Fish Biol. 66, 650–667. https://doi.org/10.1111/j.1095-8649.2005.00629.x.

Johnson, M.P., Frost, N.J., Mosley, M.W.J., Roberts, M.F., Hawkins, S.J., 2003. The area-independent effects of habitat complexity on biodiversity vary between regions. Ecol. Lett. 6, 126–132. https://doi.org/10.1046/j.1461-0248.2003.00404.x.

Jones, C.G., Lawton, J.H., Shachak, M., 1994. Organisms as ecosystem engineers. Oikos 69, 373–386. https://doi.org/10.1007/978-1-4612-4018-1_14.

Kostylev, V., Erlandsson, J., 2001. A fractal approach for detecting spatial hierarchy and structure on mussel beds. Mar. Biol. 139, 297–506. https://doi.org/10.1007/s002270100597.

Kostylev, V.E., Erlandsson, J., Ming, M.Y., Williams, G.A., 2005. The relative importance of habitat complexity and surface area in assessing biodiversity: fractal application on rocky shores. Ecol. Complex. 2, 272–286. https://doi.org/10.1016/j.ecocom.2005.04.002.

Kovalenko, K.E., Thomaz, S.M., Warfe, D.M., 2012. Habitat complexity: approaches and future directions. Hydrobiologia 685, 1–17. https://doi.org/10.1007/s10750-011-0974-z.

Lawrie, S.M., McQuaid, C.D., 2001. Scales of mussel bed complexity: structure, associated biota and recruitment. J. Exp. Mar. Biol. Ecol. 257, 136–261. https://doi.org/10.1016/S0022-0981(00)00290-2.

Loke, L.H.L., Jachowski, N.R., Bouma, T.J., Ladle, R.J., Todd, P.A., 2014. Complexity for artificial substrates (CASU): software for creating and visualising habitat complexity. PLoS One 9, e87990. https://doi.org/10.1371/journal.pone.0087990.

Mandelbrot, B.B., 1977. Fractals: Form, Chance and Dimension. WH Freeman and Company, San Francisco.

Mandelbrot, B.B., 1983. The Fractal Geometry of Nature. W.H. Freeman and Company, New York.

McCormick, M.I., 1994. Comparison of field methods for measuring surface topography and their associations with a tropical reef fish assemblage. Mar. Ecol. Prog. Ser. 112, 87–96. http://www.jstor.org.ezproxy.uct.ac.za/stable/24847640.

McCoy, E.D., Bell, S.S., 1991. Habitat structure: the evolution and diversification of a complex topic. In: Bell, S.S., McCoy, E.D., Mushinsky, H.R. (Eds.), Habitat Structure: the Physical Arrangement of Objects in Space. Chapman and Hall, London, pp. 3–27.

Munguia, P., Osman, R.W., Hamilton, J., Whitlatch, R.B., Sajac, R., 2011. Changes in habitat heterogeneity alter marine sessile benthic communities. Ecol. Appl. 21, 925–935. https://doi.org/10.1890/09-2398.1.

Parravicini, V., Rovere, A., Donato, M., Morri, C., Bianchi, C.N., 2006. A method to measure three-dimensional substratum rugosity for ecological studies: an example from the date-mussel fishery desertification in the north-western Mediterranean. J. Mar. Biol. Assoc. U. K. 86, 689–690. https://doi.org/10.1017/S0025315406013579.

Quinn, G., Keough, M., 2002. Comparing groups or treatments – analysis of variance. In: Quinn, G., Keough, M. (Eds.), Experimental Design and Data Analysis for Biologists. Cambridge University Press, New York, pp. 173–188.

Risk, M.J., 1972. Fish diversity on a coral reef in the Virgin Islands. Atoll Res. Bull. 193, 1–6.

Robinson, T.B., Branch, G.M., Griffiths, C.L., Govender, A., Hockey, P.A.R., 2007. Changes in South African rocky intertidal invertebrate community structure associated with the invasion of the mussel *Mytilus galloprovincialis*. Mar. Ecol. Prog. Ser. 340, 163–171. https://doi.org/10.3354/meps340163.

Sadchatheeswaran, S., Branch, G.M., Moloney, C.L., Robinson, T.B., 2018. Impacts of alien ecosystem engineers overwhelm inter-annual and seasonal shifts in rocky-shore community composition on Marcus Island, South Africa. Afr. J. Mar. Sci. 137–147. https://doi.org/10.2989/1814232X.2018.1462729.

Sadchatheeswaran, S., Branch, G.M., Robinson, T.B., 2015. Changes in habitat complexity resulting from sequential invasions of a rocky shore: implications for community structure. Biol. Invasions 17, 1799–1816. https://doi.org/10.1007/s10530-014-0837-4.

Sadchatheeswaran, S., Moloney, C.L., Branch, G.M., Robinson, T.B., submitted for publication. Blender interstitial volume: a novel virtual measurement of structural complexity applicable to marine benthic habitats. Method X.

Shumway, C.A., Hofmann, H.A., Dobberfuhl, A.P., 2007. Quantifying habitat complexity in aquatic ecosystems. Freshw. Biol. 52, 1065–1076. https://doi.org/10.1111/j.1365-2427.2007.01754.x.

Snover, M.L., Commito, J.A., 1998. The fractal geometry of *Mytilus edulis* L. spatial distribution in a soft bottom system. J. Exp. Mar. Biol. Ecol. 223, 53–64. https://doi.org/10.1016/S0022-0981(97)00148-2.

Sugihara, G., May, R.M., 1990. Applications of fractals in ecology. Trends Ecol. Evol. 5, 79–86. https://doi.org/10.1016/0169-5347(90)90235-6.

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M.C., Schwager, M., Jeltsch, F., 2004. Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. J. Biogeogr. 31, 79–92. https://doi.org/10.1046/j.0305-0270.2003.00994.x.

Tsuchiya, M., Nishihara, M., 1985. Islands of *Mytilus* as a habitat for small intertidal animals: effect of island size on community structure. Mar. Ecol. Prog. Ser. 25, 71–81. http://www.jstor.org.ezproxy.uct.ac.za/stable/24816880.

Van Dover, C., Trask, J., 2000. Diversity at deep-sea hydrothermal vent and intertidal mussel beds. Mar. Ecol. Prog. Ser. 195, 169–178. https://doi.org/10.3354/meps195169.

Warfe, D.M., Barmuta, L.A., Wotherspoon, S., 2008. Quantifying habitat structure: surface convolution and living space for species in complex environments. Oikos 117, 1764–1773. https://doi.org/10.1111/j.1600-0706.2008.16836.x.

Wilding, T.A., Palmer, E.J.L., Polunin, N.V.C., 2010. Comparison of three methods for quantifying topographic complexity on rocky shores. Mar. Environ. Res. 69, 143–151. https://doi.org/10.1016/j.marenvres.2009.09.005.