Characterisation of the shape of aggregates using image analysis

Daniel Niruban Subramaniam, Dassanayake Hitihamilage Hiruni Pubudika Dassanayake, Navakulan Ahilash, Sathushka Heshan Bandara Wijekoon & Navaratnarajah Sathiparan

INTERNATIONAL JOURNAL OF PAVEMENT ENGINEERING
2024, VOL. 25, NO. 1, 2349905
https://doi.org/10.1080/10298436.2024.2349905

Taylor & Francis
Taylor & Francis Group

Characterisation of the shape of aggregates using image analysis

Daniel Niruban Subramaniam, Dassanayake Hitihamilage Hiruni Pubudika Dassanayake, Navakulan Ahilash, Sathushka Heshan Bandara Wijekoon and Navaratnarajah Sathiparan

Department of Civil Engineering, University of Jaffna, Killinochchi, Sri Lanka

ABSTRACT

Aggregates are packed effectively packed in applications including pervious concrete, mechanical performance of which, depends primarily on packing characteristics. Size and shape distribution of aggregates affect packing. Several computational methods are used to represent shape of aggregates numerically, from analyses of digital images, yet, they have not been compared. This study aims to analyse representability of shape aspects of aggregates from different computational methods. Crushed rock-aggregates were grouped into five clusters and milled in LAAV machine for different number of revolutions (0–2000) to induce morphological changes, Samples were then sieved and aggregates from 5 to 30 mm in diameter were obtained (7191 in total). Aggregates were painted in black oil paint, laid on white sheets and digital images were obtained and analysed using an open source software ImageJ^{TM} while shapes were defined based on 14 computational methods. Statistical tests Pearson's Correlations, Principal Components Analysis and K-means Cluster analysis were used to assess shape factors. In conclusion, no shape factor singularly represented the morphological changes on aggregate particles. A combination of shape factors is required. The data matrix had three dimensions and three shape factors Circularity, Kumbrein solidity and Barksdale shape factor optimally represented aggregate shape index.

ARTICLE HISTORY
Received 7 April 2023
Accepted 26 April 2024

KEYWORDS
Aggregates; shape factor; morphology; shape distribution; aggregate surface texture

1. Introduction

The performance and efficiency of materials used in civil engineering applications in transportation, geotechnical and construction industries depend predominantly on the properties of aggregates. Aggregates are defined by properties such as specific gravity, particle size distribution (conformity and uniformity coefficients), aggregate impact value, aggregate crushing value and Los Angeles Abrasion Value (Akçaoğlu *et al.* 2004, Al-Rousan *et al.* 2005, Arasan *et al.* 2011). These characteristics determines if an aggregate qualifies to be used in subsequent applications. The skeleton matrix that gives rise to the strength in civil engineering applications such as sub-base in highways, ballast in railwat tracs and concrete in construction industry emphatically depends on the packing arrangements of aggregates particles (Evangelista and de Brito 2007, Aamer Rafique Bhutta *et al.* 2013, Ćosić *et al.* 2015, Subramaniam *et al.* 2022, Sathiparan *et al.* 2024).

Understandably, packing is affected by the size distribution (aggregates) and the shape distribution of particles. More often, aggregates in such applications are considered as of a single representative size (mono-size particles) indicated by the average size of the sample of aggregates and that they are in the shape of a sphere (spherical particles) (Dinger and Funk 2013, Chan and Kwan 2014, Pouranian and Haddock 2018, Subramaniam 2018, Chang and Deng 2020, Sajeevan *et al.* 2022). In contrast, several studies have attempted to mathematically optimise the packing model using dual-size particles (particles of two distinctive sizes), which still remain highly limited in applications (de Larrard and Sedran 1994, Gan *et al.* 2010, Das *et al.* 2011, Chang and Deng 2020). Natural aggregates come in different sizes and are represented by a particle size distribution curve, and mathematically represented by uniformity coefficient, conformity coefficient and at times together with effective diameter which corresponds to D_{10} (where 90% of the particles are bigger than the effective diameter). While these numerical quotients are employe in geotechnical engineering in soils classification as deterministic parameters, the efficacy of these numerical values representing the size distribution of particle size distribution remains debatable (Sajeevan 2023). Developing a packing optimisation model for a lump of particles with a distribution of sizes may need emphatic numerical representation, perhaps as a single or a combination of number of coefficients (Mueller 2010, Ng *et al.* 2016, Pouranian and Haddock 2018, Shen 2019).

Natural aggregates, which come in a distribution of sizes, are in addition, rarely spherical in shape as assumed in the engineering models. Defining the shape of an aggregate has been attempted by several researchers resulting in the development of various mathematical computational models to represent the shape of an aggregate (Brzezicki and Kasperkiewicz 1999, Al-Rousan *et al.* 2005, Al-Rousan *et al.* 2007, Arasan *et al.* 2011, Bangaru and Das 2012, Zhang *et al.* 2012, Ding *et al.* 2017, Xie 2017, Ge *et al.* 2018, Xu and Prozzi 2019, Bhattacharya *et al.* 2020, Ghosh *et al.* 2020, Zhao 2020). The shape of an aggregate significantly impacts the packing of particles and thereby the performance including the workability,

CONTACT Daniel Niruban Subramaniam Daniel.subramaniam@gmail.com

strength and durability of the engineering product. In geotechnical and transportation engineering applications, shape of an aggregate is represented by aspect ratio, angularity, flakiness index and elongation index that have standard methods of quantification. Aspect ratio is a measure of the length to width ratio of an aggregate particle where aggregate. While a higher aspect ratio tends to have lower packing density, a lower aspect ratio tend to have higher packing density, enhanced workability, strength and durability (Zhu *et al.* 2018, Zhang *et al.* 2019). Angularity on the other hand, is a measure of the sharpness of the corners and edges of an aggregate particle, that corresponds to the strength and frictional properties of aggregates in subsequent applications. Aggregates with higher angularity tend to have a higher interlocking strength and a better resistance to wear and abrasion while lower angularity may result in smoother surface and hence reduced strength of materials (Maerz 2004, Man and van Mier 2011). Flakiness index is a measure of the degree of flatness of the surfaces of an aggregate particle where higher flakiness index tends to have lower packing and workability and thereby reduced strength (Al-Rousan *et al.* 2005, Al-Rousan *et al.* 2007). Elongation index is a measure of the elongation of an aggregate particle and that higher elongation index indicates lower packing, workability and strength of material (Xie 2017). In addition to these shape factors, other parameters that affect the shape of aggregates include texture, roundness and sphericity of the particles (Wang 2005, Su and Yan 2018). Texture refers to the roughness of the particles, while roundness and sphericity refer to the curvature of the particles (Mahmoud *et al.* 2010, Maroof *et al.* 2020). These additional shape factors are hard to be quantified physically and are not often considered in civil engineering applications and in prediction models developed, although they have significant impact on performance. Unaccountability of the effects of these aspects of the shape of aggregates leads to higher uncertainties in the prediction models.

Image analysis is a powerful tool for analysing the shape of aggregates in construction materials and is often used to determine the size, shape and texture of individual aggregate particles as well as to quantify the distribution of these properties across a sample. Image analysis could be performed using a variety of techniques, including microscopy, digital imaging and scanning electron microscope (Fletcher *et al.* 2003, Bangaru and Das 2012, Gu *et al.* 2014, Bhattacharya *et al.* 2020). In the case of aggregate shape analysis, digital imaging is the most commonly used technique, obtained using high-resolution digital cameras and then subsequently analysed using specialised softwares. The software identifies individual particles and computes their size, shape and texture parameters. While image analysis is advantageous to analyse several individual aggregates in a sample population, subsequent analyses and representing the distribution by numerical values are often seen as challenging (Liu 2017, Krejsová *et al.* 2018).

Several researchers attempted to determine the shape of aggregates that led to devising different computational methods (Mostofinejad and Reisi 2012), more in the context of powder technology (Lin 2016) and ballast in geotechnical engineering (Kusumawardani and Wong 2020). While these studies assumed as particles were of spherical shape in modeling packing of particles, several other studies assumed the particles of rectangular shape and incorporated flatness ratio and elongation ratio in the packing models (Altuhafi *et al.* 2013, Gong *et al.* 2016, Ding *et al.* 2017). Furthermore, another study attempted to define the shape of aggregates using Digital Image Processing techniques (DIP) with six shape parameters including Flakiness ratio, Elongation ratio, Convexity ratio, Fullness ratio, Shape factor and Sphericity (Kwan *et al.* 2014). Deng *et al.* (2021) introduced the Maximally Random Jammed (MRJ) for ellipsoids, oblate spheroid and prolate spheroids. In addition to these, another study used sphericity and fullness ratio as parameters to define the shape of aggregates (Zheng 2020). Studies have considered both two-dimensional and three-dimensional representation of aggregates in quantifying the shape and size of aggregates, while some studies have approximated three-dimensional representation from two-dimensional projection of the same (Masad *et al.* 2000, Tafesse *et al.* 2012, Shuguang and Qingbin 2015, Gong *et al.* 2016).

Although several computational methods have been proposed by studies in contemporary literature to define the shape of an aggregate, they are mostly confined to the context in which the model was either developed or employed (Anburuvel and Subramaniam 2022, 2024, Wijekoon 2024). Literature has not established a universal representation of the shape of an aggregate, that could be used in a mathematical performance model, to incorporate the impact of shape of the aggregate on the performance of the material. With the increased interest in pervious concrete, the performance of which solely depends on the shape and size distribution of aggregates used, a more accurate representation of shape and size distribution of aggregates becomes critical (Sajeevan and Subramaniam 2022). Specifically, the permeability and strength of pervious concrete, which are the prime performance indicators of the concrete are significantly affected by the characteristics of the pore structure it encompasses, which in turn depends on the shape and size distribution of aggregates. This study aims to analyse widely used aggregate shape factors and computational models to analyse the efficacy of their representation of the same as a numerical coefficient.

2. Materials and methods

2.1. Aggregate properties

Crushed aggregates of gneiss origin were obtained from the North-central region of Sri Lanka, varying in size from 5 to 30 mm in diameter. Aggregates were then tested for their physical properties as explained below. The purpose of the study is to analyse the effectiveness of different computational methods to numerically represent the shape of aggregates. Although different mineralogical constituents of rocks may have different morphological changes in the aggregates during the process of milling, that dimension is out of scope of this study. And therefore, different lithological aspects of aggregates on the morphological changes in the shape of the aggregates are not considered in this study and a single type of aggregate was utilised.

2.1.1. Aggregate preparation for shape analysis

Aggregates were observed to be flaky, elongated and angular in the natural form it was obtained. The measure of LAAV (Los Angeles Abrasion Value) is done using the LAAV instrument that operates similar to a ball mill, which by abrasion deforms the aggregates. Deformation of aggregates occurs by abrasion in which the surface of the aggregates would begin to loose roughness as the milling process undergoes multiple revolutions. As the revolutions continued to increase (milling over a longer period indicated by higher number of revolutions), the aggregate then changes in shape loosing the angular corners and becoming more rounded in shape. On further milling, the aggregate would then tend to attain a spherical shape form (with the extended amount of milling). Based on this principle, the original aggregates obtained for this study was then milled for different lengths of time in LAAV machine, including 500, 1000, 1500 and 2000 revolutions (at a rate of 30 rpm). It is understandable that the aggregates particles naturally obtained for this study are flaky and elongated, which may undergo breakage and wear and tear during the process of milling. This would be advantageous for this study instill morphological changes significant, that could be analysed in subsequent analyses. The impact of different flakiness and elongation indexes on the morphological changes on the aggregates during the process of milling is beyond the scope of this study. A set of aggregates were obtained in one batch, and samples for this study were obtained from that batch, indicating uniform properties across different groups of aggregates that underwent different degrees of milling. Therefore, any morphological aspects observed to be distinctively different across the groups of aggregates are attributed to the degree of milling. Figure 1 shows the pictures of the aggregate samples that had undergone different number of revolutions of milling.

Specific amount of aggregate samples were taken from each lump of aggregates, from natural aggregates (that were not milled) and each group of aggregates (with different degrees of milling). The aggregates were then coloured using black oil paint. They were then allowed to air-dry in room temperature. After drying, those aggregate particles were spread on a white A4 sheet in row-column arrangement. An image of arranged aggregates was taken covering the complete sheet. Arranged particles were mixed together and rearranged on the sheet and an image was taken again. This procedure was repeated 12 times and 12 different images were obtained to incorporate all projections of aggregates in the two-dimensional representations. While taking the image, the background of the sheet was lit with white light, to avoid the inclusion of shadow of flash in the image adjoining the aggregates.

The images obtained were then processed using ImageJ™, an open-source image analysis tool. The image was scaled using the dimensions of the A4 sheet size and the image was then converted to binary colour scale (black and white). The pixel with black in colour bore a pixel value of 0 while the pixel with white colour had a value of 256. Collection of black pixels corresponded by a black zone in the image indicated an aggregate.

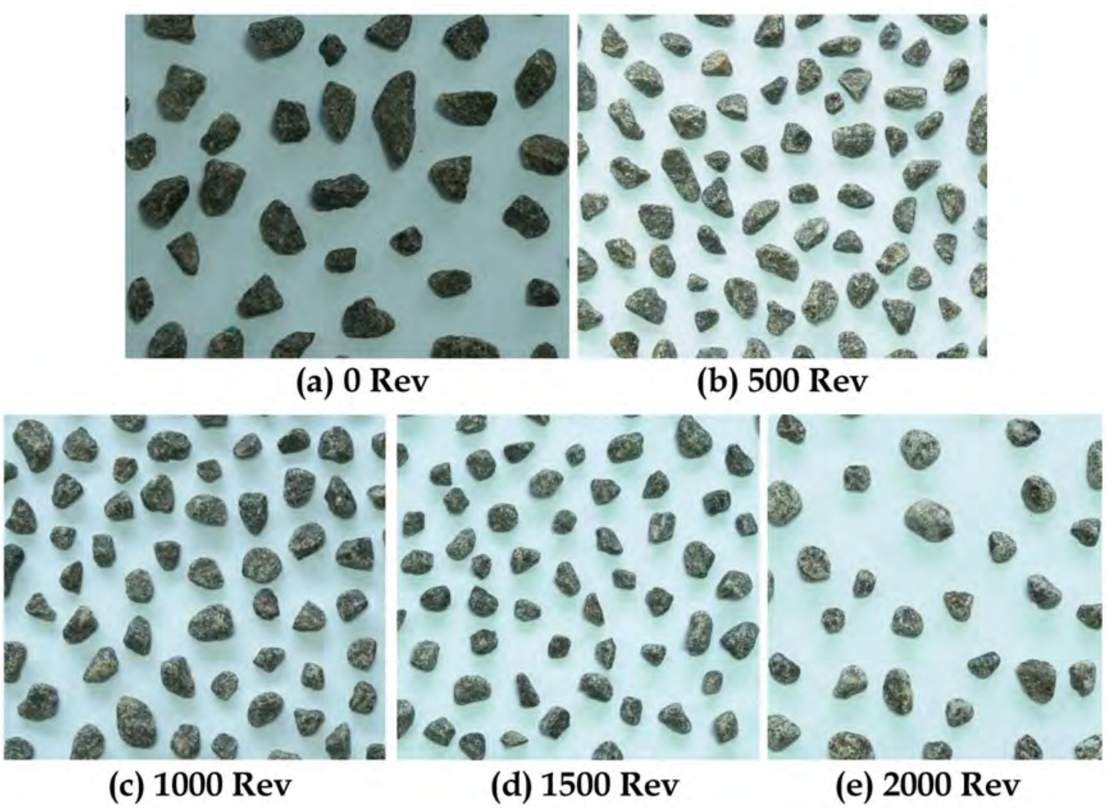

Figure 1. Pictures of aggregates that have undergone varied degrees of milling.

2.1.2. Particle size distribution analysis using image analysis

Particle Size Distribution (PSD) is the representation of the percentage of particles in each size bin (diameter), drawn to a cumulative percentage that passes through particular sieve sizes. In this analysis using image of aggregates, the area of each particle is computed individually using image analysis tool ImageJ^{TM}, and the area-equivalent circle was drawn and the diameter of which is computed and noted. The analysed aggregate is assumed to be a particle with a projected circle in shape with the computed diameter.

The particles were arranged in ascending order of corresponding diameter and bins of 1 mm were then devised and the aggregates were categorised. The total area within each bin (summation of the areas of each aggregate in each bin) was computed and recorded as the percentage of the total area (sum of areas of all aggregates). The percentages of each bin were then computed as the cumulative percentage of aggregate area up to the bin in consideration (from the lowest bin-diameter) and the graph was plotted against the bin-diameter. The plot is the PSD of aggregates in the sample analysed using image analysis tool. The mean of the percentages across all repetitive samples were obtained to represent the mean PSD of the sample population of aggregates belonging to a particular group.

2.1.3. Computation of shape factors

Computational models to quantify the shape factors require the quantification of the following dimensional parameters.

- Bounding rectangle – The smallest rectangle enclosing the selection. This dimension is called breadth and height in ImageJ as shown in Figure 2(a).
- Fit Ellipse – The best fitting ellipse for the area is obtained. From the ellipse, primary axis and secondary axis lengths were computed, which are indicated as Minor and Major axis in ImageJ as shown in Figure 2(b).
- Feret's Dimensions – There are two Feret's dimensions. It measures the size of an area along a specified direction as shown in Figure 2(c).
- Feret – The longest distance between any two points of the section boundary. (Maximum Caliper)
- MinFeret- The smallest distance between any two points of the section boundary. (Minimum Caliper).

The following 14 different computational models employed to quantify the shape of aggregate in contemporary literature are considered for analyses in this study, the details of which are given in Table 1.

2.2. Statistical analyses

Several statistical tools were employed in this study to verify the efficacy of shape factors in characterising the aggregates based on shape, which is effected by the milling process. Pearson Correlation: is a measure of how well two continuous variables are linearly correlated, the correlation coefficient that results from the test range between −1 and +1. A positive value indicates a directly proportional correlation while a negative value indicates inversely-proportional relationships. In addition, a numerical value of 1 indicates a perfect linear correlation while a numerical value close to zero indicates the absence of a linear correlation. The test also returns a significant value, whereas a value of less than 0.05 indicates more than 95% confidence in the level of linear correlation indicated by the correlation coefficient. The point to consider in this test is that the test only tests if the variables are linearly correlated and to assess non-linear correlations, the variable would need to be correspondingly transformed and then tested again. This test is employed in this study to assess if any of the shape factors correlated to the degree of milling (if any of the shape factors directly linearly predict the shape variance), and also to understand if any of these shape factors are linearly correlated among them so that they can be considered surrogate parameters.

Principal Component Analysis (PCA): It is considered a pattern recognition tool and not a definitive statistical test, and is often employed to reduce the dimensions of the data matrix. This test reorganises the variance of data into several principal components that are orthogonal to each other and computes how much of variance is explained by each of the principal components. A rotated matrix about varimax lists the variables enlisted in the matrix that are predominantly

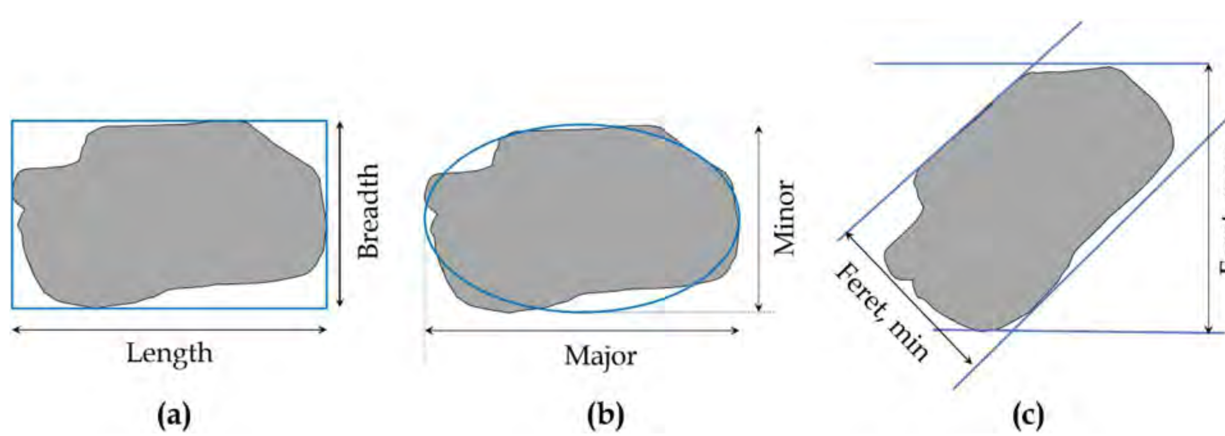

Figure 2. Dimensional parameters used to compute aggregate shape factors.

Table 1. Description of shape factors.

	Shape factor	Computation	Reference
J.C	ImageJ circularity	$\frac{4\pi \text{ area}}{(\text{perimeter})^2}$	Rueden 2017
J.AR	ImageJ aspect ratio	$\frac{\text{major axis}}{\text{minor axis}}$	Rueden 2017
J.R	ImageJ roundness	$\frac{4\pi \text{ area}}{\pi \times \text{ major axis}^2}$	Rueden 2017
J.S	ImageJ solidity	$\frac{\text{area}}{\text{convex area}}$	Rueden 2017
KS	Krumbein sphericity	$\sqrt[3]{\frac{\text{breadth}}{\text{length}^2}}$	Kim *et al.* 2019
BSF	Barksdale shape factor	$\frac{\text{length}}{\text{breadth}^2}$	Barksdale *et al.* 1991
FR	Kwan and Mora fullness ratio	$\sqrt[2]{\frac{\text{area}}{\text{convex area}}}$	Mora and Kwan 2000
SF.FR	Sneed and Folk flatness ratio	$\frac{\text{Breadth}}{\text{Length}}$	Lewis and McConchie 1994
FAR	Feret's aspect ratio	$\frac{\text{Feret}}{\text{Min Feret}}$	Xu and Di Guida 2003
SF	Shape factor	$\frac{\text{Perimeter}}{\text{Calculated perimeter}}$	
PSC	Projected sphericity/circularity	$\frac{\text{D. Calculated}}{\text{Feret}}$	Cruz-Matías *et al.* 2019
ICS	Inscribed circle sphericity	$\sqrt{\frac{\text{Min Feret}}{\text{Feret}}}$	Zheng and Hryciw Roman 2016
C	Circularity	$\frac{\text{perimeter}^2}{\text{area}}$	
PS	Projected sphericity	$\frac{\text{area}}{\text{AC}}$	

aligned to each principal component indicated by a value between −1 and +1. The positive value indicates a positive impact on the principal component while a negative value indicates a negative impact on the variance of the matrix. In addition, a value close to 1 indicates a strong impact of the variable in the dimension indicated by the corresponding principal component while a value close to zero indicates the direction of the vector being orthogonal (not impacting the variance in the relevant principal component). This test is employed in this analysis to identify how many dimensions would be required to explain the variance in the data matrix and which of the shape factors dominate each principal component. The results of this study together with Pearson's correlation test would essentially be used to reduce the dimensions of the data matrix, which means, removing the redundant or surrogate shape factors in the data matrix.

Cluster analysis: is a statistical technical tool that is used to cluster similar objects (or cases) together based on predefined set of parameters that are common to all cases. It is a powerful tool for data exploration, visualisation, pattern recognition and interpretation. The technique is based on similarity or dissimilarity between objects/cases. The similarity of dissimilarity can be measured in terms of distance metrics, such as Euclidean distance, Manhattan distance and cosine similarity. K-means clustering is a non-hierarchical clustering method, that partitions data in to K distinct clusters based on similarity. K-means clustering aims to partition a set of n observations (cases/objects) into K clusters, where each observation belongs to the cluster with the nearest mean or centroid. This is a method based on interactions, where the algorithm minimises the sum of squared distances between each observation and its assigned cluster centroid. In this study, K-means clustering is adopted that computed similarity between aggregates based on several combinations of shape factors, the similarity computed based on Euclidean distance.

3. Results and discussion

The purpose of the study is to evaluate the credibility of the shape factors used, to represent the shape of an aggregate quantitatively and to analyse the sensitivity of the shape parameters used in contemporary literature. The results and discussion section of this article includes the following,

1. Visual identification: This section contains the visual inspection and character identification, especially owing to the shape characteristics including surface texture, angularity, flakiness and elongation (micro and macro shape characteristics) of aggregate particles.
2. Descriptive statistics: This section contains statistical descriptions of the data matrix of dependent and independent variables, such as the projected area, the perimeter and shape factors of aggregate particles. Statistical descriptions discussed in this section include mean 95% confidence interval, quartiles and median and frequency distributions (histogram) and the characteristics of distribution patterns.
3. Pattern recognitions: This section discusses the patterns in the data variables considered in this study, to statistically identify the visual character identifications. Pearson's correlation coefficient is used to identify the existence of linear

correlations among dependent and independent variables to understand the dependencies of shape factors on the size distribution of aggregate particles, and between shape factors to identify the presence of surrogate shape factors. The second tool used in this section is Principal Components Analysis (PCA) with which the dimensions of the data matrix could be reduced depending on the explication of the total variance of data matrix. Third tool used in this section is the process of clustering (K-means) which could be used to identify the effectiveness of shape factors and combination of shape factors that can cluster aggregates based on the shape that is computable using the computational models analysed in this study.

3.1. Visual inspection of aggregates

Figure 1 shows the pictures of aggregates that have undergone different degrees of milling (up to 2000 revolutions) where 0 rev indicates the original crushed aggregates. It is evident from the figure that the morphological characteristics of aggregates significantly differ between the degrees of milling, particularly, the change in texture and shape.

The surface texture relates to highly frequent irregularities on the surface of the aggregates where the shape would be indicated by less frequent irregularities as per Fourier analysis of the aggregate image (Su and Xiang 2020). Specifically, when milling is more intense (prolonged), the aggregate tends to become more rounded (spherical in three dimensional) while a lower degree of milling corresponded to smoothing of the surface, essentially changing the texture of the aggregate surface. Shape factors may intend to quantify any or all of the changes in the shape characteristics of the aggregates ranging from texture to macro-shape. The analysis in this study aims to identify which of the shape factor computational models retrieved from the literature quantify the morphological change visually observed in the aggregates shown in the figure.

3.2. Descriptive statistics of shape factors

Table 2 shows the details of statistics of aggregate shape characteristic parameters including projected area, perimeter and various shape factors analysed in this study. The statistics presented here are for a total of 7191 aggregate particles, belonging to groups corresponding to different degrees of milling. It could be observed that the area and perimeter have a very high standard deviation (more than 50% of the mean) and that the skewness is also significantly high. Simultaneously, several of the shape factors indicate a higher variation among the aggregate particles analysed. Kurtosis indicates the roundness or sharpness of the distribution at the peaks, which again is observed to vary significantly among various shape factors.

Analysis of the variations of these parameters (aggregate shape characteristics) within each group of aggregates that have been grouped based on degrees of milling would be essential for further analysis. Figure 3 shows the mean and 95% confidence interval of the distribution of J.C, KS, J.R and BSF within different groups of aggregates grouped

Table 2. Descriptive statistics details of variables of aggregates characteristics analysed in this study.

		Statistics															
		P	Area	J.C	J.AR	J.R	J.S	KS	BSF	FR	SF.FR	FAR	SF	PSC	ICS	C	PS
N	Valid	7191	7191	7191	7191	7191	7191	7191	7191	7191	7191	7191	7191	7191	7191	7191	7191
	Missing	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mean		43.12	134.43	0.77	1.35	0.76	0.96	0.41	0.12	0.98	0.85	1.37	1.14	0.83	0.86	16.38	0.69
Median		36.61	81.04	0.78	1.29	0.77	0.97	0.41	0.12	0.98	0.87	1.32	1.13	0.84	0.87	16.07	0.70
Mode		28.93	33.35	0.81	1.21	0.80	0.97	0.48	0.12	0.99	1.00	1.13	1.06	0.54	0.69	14.15	0.30
Std. Deviation		19.09	128.46	0.06	0.25	0.12	0.01	0.06	0.05	0.01	0.11	0.22	0.05	0.06	0.06	1.50	0.10
Variance		364.42	16502.68	0.00	0.06	0.01	0.00	0.00	0.00	0.00	0.01	0.05	0.00	0.00	0.00	2.24	0.01
Skewness		1.37	2.05	−1.54	1.82	−0.53	−1.39	−0.29	0.76	−1.44	−0.80	1.76	2.97	−0.74	−0.86	3.90	−0.51
Std. Error of Skewness		0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Kurtosis		1.37	4.15	4.86	5.54	0.09	4.49	−0.62	1.42	4.86	0.32	5.10	20.40	0.79	0.89	35.76	0.23
Std. Error of Kurtosis		0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
Percentiles	25	29.22	53.23	0.75	1.18	0.69	0.96	0.37	0.08	0.98	0.79	1.22	1.11	0.80	0.83	15.47	0.63
	50	36.61	81.04	0.78	1.29	0.77	0.97	0.41	0.12	0.98	0.87	1.32	1.13	0.84	0.87	16.07	0.70
	75	51.46	159.65	0.81	1.46	0.85	0.97	0.45	0.15	0.99	0.94	1.47	1.16	0.87	0.90	16.87	0.76

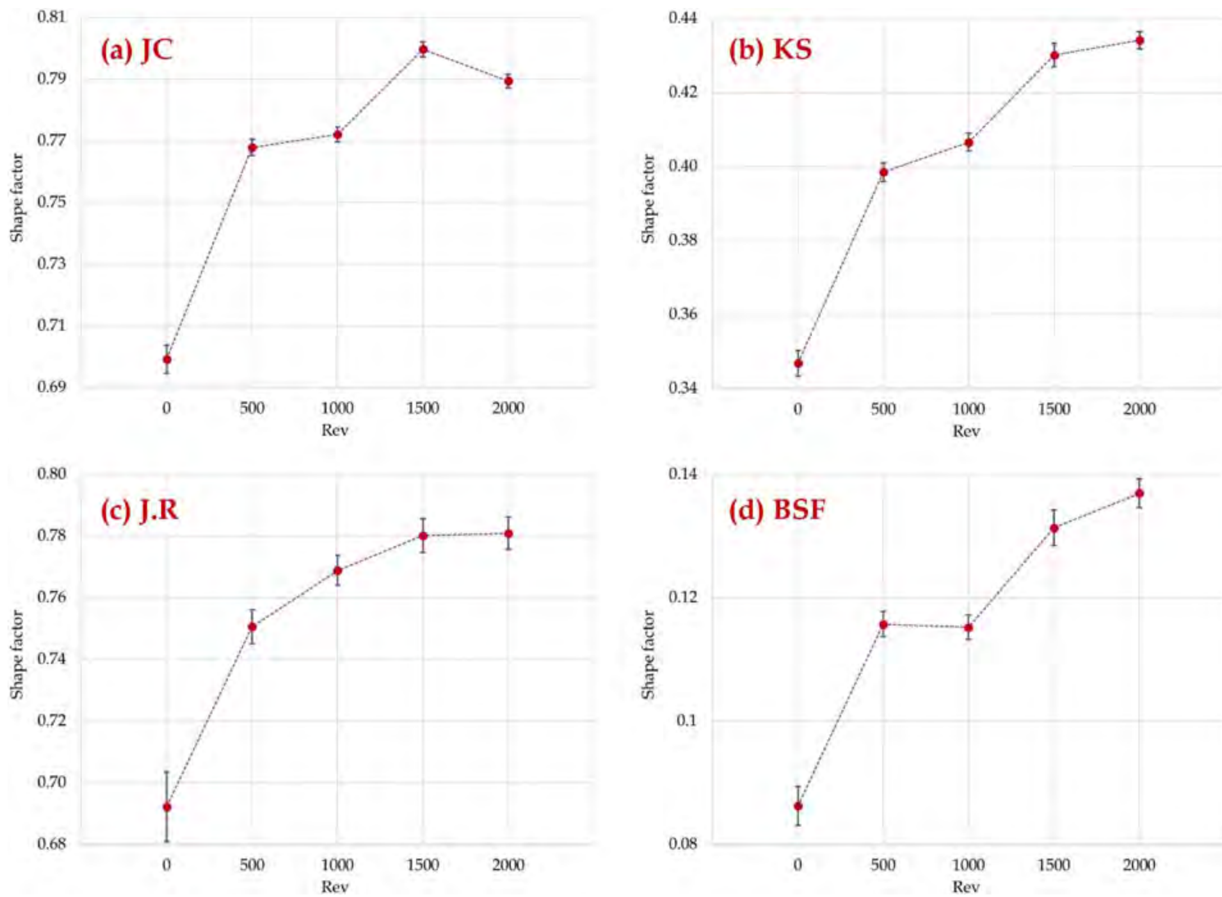

Figure 3. Mean and 95% confidence interval of the distribution of J.C, KS, J.R and BSF within different groups of aggregates.

based on degrees of milling. The four shape factors that showed significant variations among degrees of milling have been reported in the figure. It could be observed that the shape factors J.C, KS, J.R and BSF have shown an increasing trend with the degree of milling, where a rapid increase was observed between 0 and 500 Rev compared to the difference between a higher degree of milling. Other than BSF, the other three shape factors tend to saturate with the degree of milling within the scope of the study, attaining saturation at least by 1500 revolutions.

Further analysis of the distribution of shape factors within each group revealed that a higher variation could be observed within the group with no milling (natural crushed aggregate particles), compared to other groups of aggregates. Specifically, J.R reveals a much higher variation, compared to other shape factors shown in the figure within the first group of aggregates.

The shape factors are mostly based either on the dimensions specifically, length, breadth, perimeter and/or projected area of aggregate particles. The quantified shape factors would therefore change based on the variations among the aggregates when proportions among these dimensions changed. More specifically, the primary shape aspects of aggregates such as surface texture, angularity, flakiness and elongation, would effect a change in the above-mentioned dimensions. A higher variation within a group of aggregates indicates the presence of highly varying above-mentioned aspects, while a narrower distribution indicates a more uniform presence of the same. The narrower distribution of a parameter indicates that the aggregate belonging to that group has a more uniform distribution of the parameter, while a wider distribution indicates the contrary.

For a given area of an aggregate particle, the more deformed the shape is (due to surface irregularities, angularity or flaky/elongated nature), the perimeter would be expected to increase. The proportion of area/perimeter would therefore be lower (increased perimeter for the same area). The process of milling is expected to effectively smoothen the surface and braking of flaky or sharp corners in lower number of revolutions due to abrasion and impact. Milling further corresponding to higher number of revolutions is expected to erode the corners further and make them more rounded, effectively increasing the area/perimeter proportions. This process is further expected to make the particles more uniform in characteristics.

Considering shape factor J.R it could be observed that the distribution is significantly wider for the zero Rev group and that the distribution remained rather consistent throughout the other groups that had experienced milling, irrespective of the degree of milling. This indicates that the process of milling with a degree of fewer than 500 revolutions had instigated a morphological change that removed random irregularities that naturally occurred in the crushed aggregate particles, more particularly the surface texture and angular and flaky corners. Simultaneously, it could be concluded that J.R has significantly quantified such morphological shape aspects of aggregate particles, compared to other shape factors such as KS, that showed rather trivial differences between the distributions among the groups.

Figure 4 shows the box and whiskers plot of the distribution of shape factors, J.R, J.C, KS and SF, among the group of aggregates grouped based on degree of milling. The difference between the plot given in Figure 3 is that in the box and

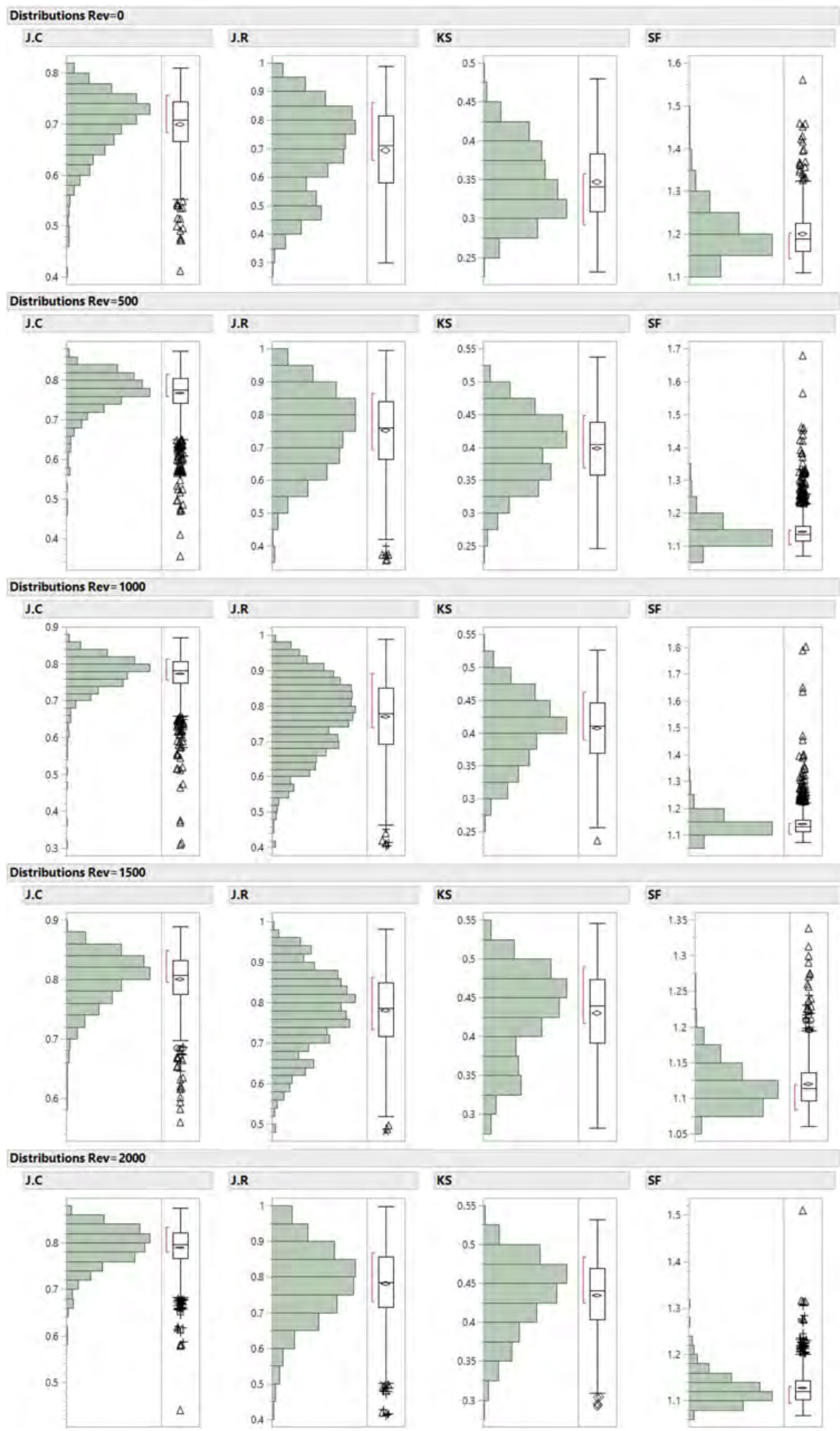

Figure 4. Distribution of shape factors among groups of aggregates.

whiskers plot, the quartiles (with median at the centre) and maximum and minimum are indicated by the box boundaries and the outliers are also shown whereas in the previous plot, it was the 95% confidence shown with mean where outliers are also included. It is evident from Figure 4 that several shape factors such as J.C and SF have a higher number of outliers, distant from the median, while other shape factors such as KS had less number of outliers, for the same aggregate particles analysed. This indicates that several shape factors can identify different aspects of the shape characteristics of aggregate particles.

Owing to the different distributions and presence of outliers, it is also worth to make observations on the histogram of the parameter values for the group of aggregates considered. Figure 4 also shows the histogram (frequency distribution) of shape factors J.C, J.R, KS and SF, among different groups of aggregates. Both J.C and J.R show a uniformly distributed parameter in the first group of aggregates (0 Rev) compared to other groups, which shows a log-normal distribution with a high frequency mean and skewness. On the other hand, SF shows a rather narrow, but more normal distribution. Shape factor KS has a similar trend to that of SF, yet there is also a trend of reducing range within even higher degree milling groups. This observation reinstates the discussion of texture and angular corner morphological changes and the effectiveness of each shape factor in quantifying the aspects. Further statistical tests are required to analyse the effectiveness of each shape factor in quantitatively representing the shape aspect of aggregate particles. Most of the other shape factors had a similar trend as shown and discussed in this section.

3.3. Pattern recognition among aggregate characteristics

Prior to analysing data with statistical tests, it is vital to analyse the trends and patterns among different parameters pertaining

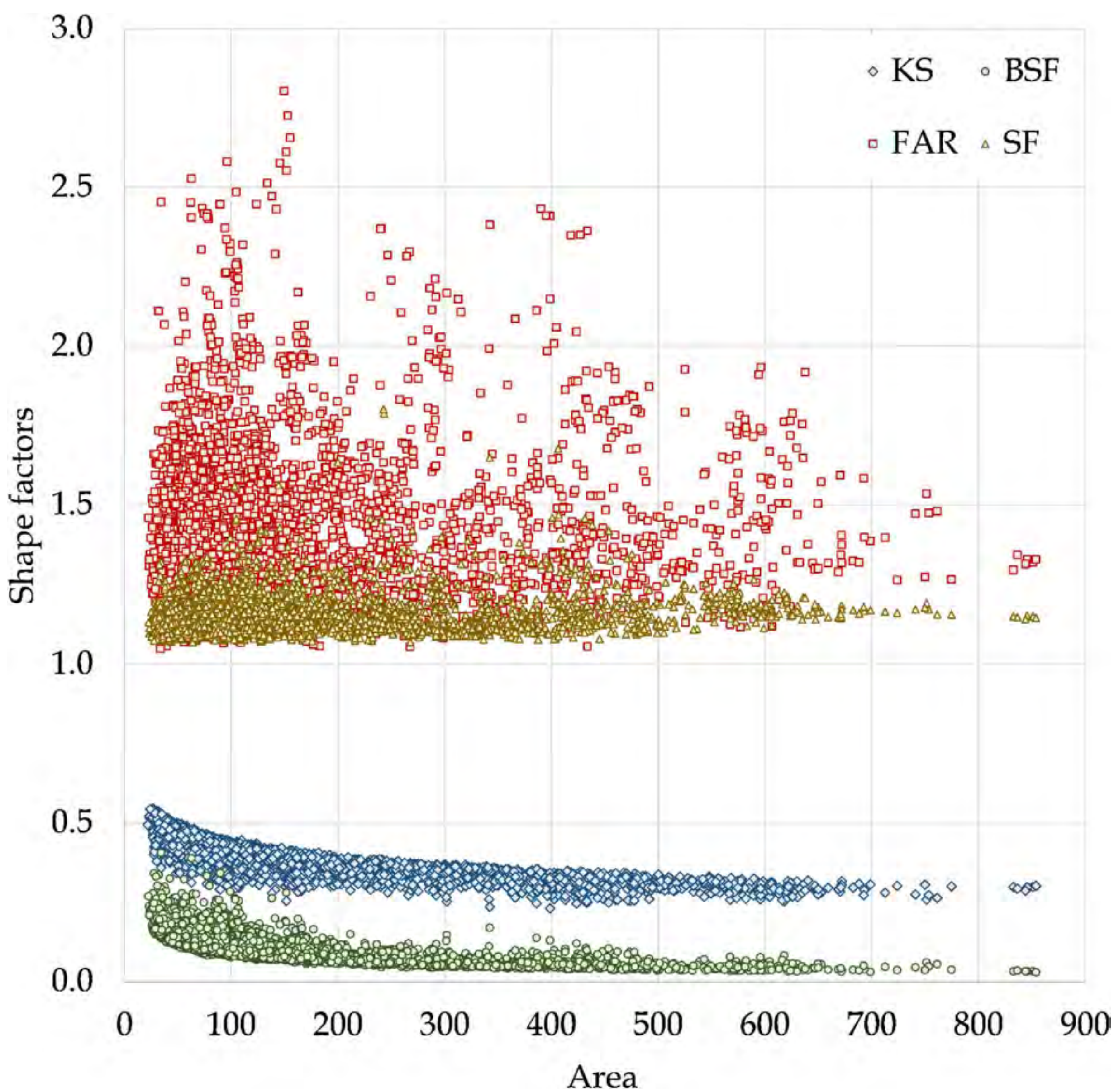

Figure 5. Variation of shape factors with projected area of aggregate particles.

to aggregate shape characteristics. As computation of most shape factors are based on the dimensions of aggregates, the variations of shape parameters with the two primary parameters considered in this study area and perimeter are analysed. Figure 5 shows the variation of shape factors KS, BSF, FAR and SF with the projected area of aggregate particles. A distinct pattern is observed in both KS and BSF with the area of aggregates, where the shape factor quantified decreases with the increase in area. Furthermore, in BSF and KS, a boundary is observed (a top boundary for KS and a bottom boundary for BSF), indicating a restriction to the parameter with the area. This could be due to a shape aspect that the factors quantify, which approaches an asymptote with respect to the projected area. Shape factors FAR and SF on the other hand do not show any correlation to the projected area, yet, do exhibit a boundary (lower boundary) as observed in KS and BSF. All other shape factors not shown in the plot had no significant impact from the projected area and followed a similar distribution but did not show a distinct boundary as observed in these cases.

3.3.1. Pearson's correlations

A Pearson's correlation analysis was conducted between the 14 shape factors and revolutions (milling degree). Seven cases were analysed, the descriptions of the cases are given below (Table 3). The cases were analysed for seven groups of aggregates, grouped based on the size of aggregates (projected area).

Table 4 shows the Pearson's correlation coefficient of Perimeter (P), projected area (A) and 14 shape factors for 7 cases, with degrees of milling (number of revolutions). The size of aggregates indicated by P and A has not shown any significant correlations with number of revolutions, indicating that any bias due to the size of aggregates with the number of revolutions is insignificant in the subsequent analysis. However, C06 shows a higher correlation between P and Rev, whereas, such an anomaly is not reflected in the correlation between A and Rev. Therefore, the distribution of aggregates of all sizes could be considered uniform across the number of revolutions.

For the case of C01 where aggregates of all sizes were analysed, none of the 14 shape factors has shown any significant linear correlation with Rev, indicating none could directly quantify the morphological change induced by different degrees of milling. The highest correlation coefficients however were obtained for shape factors KS, SF and C with 0.4, 0.35 and 0.35, respectively, indicating a degree of representation compared to other shape factors. Further analysis across the cases C02 – C07 where the size of aggregates was considered, correlations coefficients of several shape factors varied significantly with Rev across the cases. Assuming a significant relationship exists with a coefficient above 0.5, seven shape factors including J.AR, J.R, KS, BSF, SF.FR, FAR and ICS revealed insignificant relationships across cases whereas the other seven factors showed significant relationships with Rev as the aggregate size varied.

Figure 6 shows the variation of Pearson's correlation coefficients of seven shape factors with Rev, across seven cases analysed (C01–C07). While C01 was the case with aggregates of all sizes, which could be considered as the mean representation of the relationship irrespective of aggregate sizes, C02 till C07 represented an increasing trend in aggregate sizes. From the figure, shape factor J.C initially had a negative correlation (even though it was not significant) with aggregates less than 35 mm^2 in area, it swiftly rose to a positive correlation when aggregates up to 50 mm^2 were considered, yet remaining insignificant as the coefficient was close to zero. However, as the aggregates increased in size across C04 – C07, the linear correlation between J.C and Rev became significant with the coefficient assuming a value close to 0.8 for C06. A general trend indicates a rapid increase in the impact of Rev on J.C between aggregates of size up to 150 mm^2 and then saturated, indicated by rather constant coefficients. Simultaneously, JF, FR, PCS and PS showed a similar trend of rapid increase up to 150 mm^2 and then saturating at a significant linear

Table 3. Descriptions of grouping of aggregates for analysis.

Case	Aggregates size range
C01	Aggregates of all sizes ranging from 20 to 800 mm^2
C02	Aggregates of all sizes up to 35 mm^2
C03	Aggregates of all sizes up to 50 mm^2
C04	Aggregates of all sizes ranging between 150 and 200 mm^2
C05	Aggregates of all sizes ranging between 300 and 350 mm^2
C06	Aggregates of all sizes ranging between 500 and 550 mm^2
C07	Aggregates of all sizes above 600 mm^2

Table 4. Pearson's correlation coefficient of variables with degree of milling (revolutions).

	C01	C02	C03	C04	C05	C06	C07
P	−0.382	−0.073	−0.131	−0.316	−0.291	−0.769	−0.491
Area	−0.321	−0.178	−0.117	0.119	−0.049	−0.364	−0.267
J.C	0.369	−0.104	0.062	0.556	0.387	0.806	0.59
J.AR	−0.211	0.16	0.046	−0.414	−0.335	−0.191	0.306
J.R	0.187	−0.169	−0.042	0.41	0.352	0.199	−0.279
J.S	0.185	−0.106	0.085	0.482	0.579	0.761	0.681
KS	0.397	−0.05	0.039	0.332	0.3	0.288	0.002
BSF	0.274	0.115	0.076	−0.272	−0.188	0.081	0.128
FR	0.185	−0.107	0.085	0.479	0.578	0.759	0.68
SF.FR	0.123	−0.084	−0.017	0.304	0.26	0.091	−0.101
FAR	−0.233	0.115	0.016	−0.425	−0.373	−0.274	0.211
SF	−0.348	0.103	−0.056	−0.52	−0.309	−0.782	−0.559
PSC	0.249	−0.117	−0.001	0.493	0.407	0.572	0.029
ICS	0.225	−0.114	−0.006	0.43	0.396	0.295	−0.187
C	−0.336	0.102	−0.054	−0.504	−0.277	−0.773	−0.548
PS	0.246	−0.115	0.002	0.494	0.411	0.586	0.041

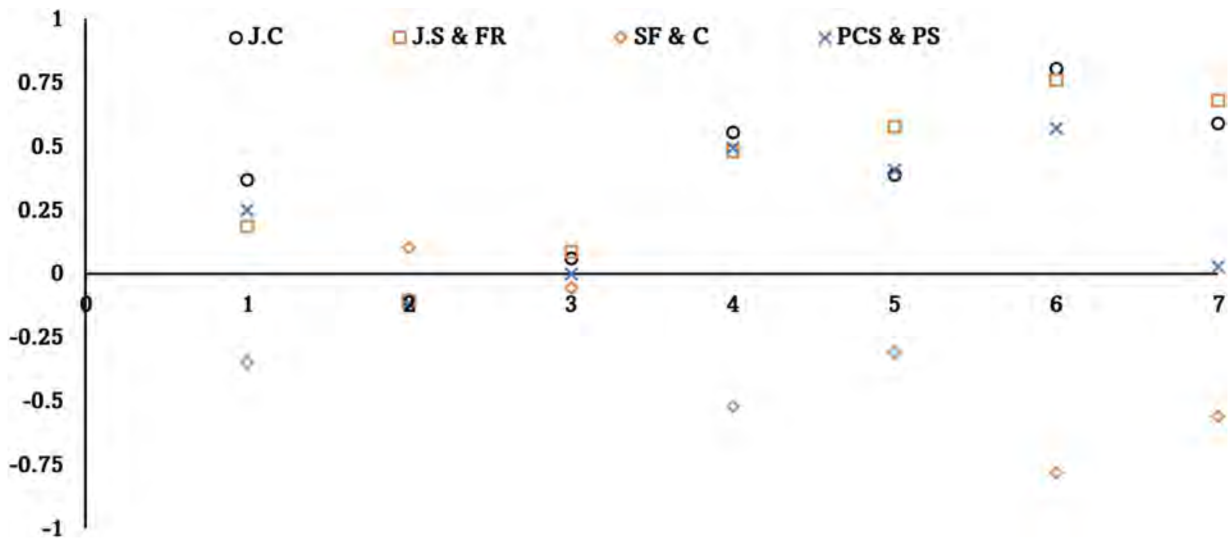

Figure 6. Variation of Pearson's correlation coefficients across cases considered.

correlation. Shape factors SF and C on the other hand showed a similar yet negative correlation. This indicates that these seven shape factors can represent a specific shape characteristic of aggregate, given that the aggregate is of sufficient size. They may not be able to quantify those aspects of the shape of the aggregate when the aggregates are smaller in size.

It is also relevant to note, that several of the shape factors indicated very similar trends with other shape factors, indicating a possibility that they may be quantifying the same aspects of the shape of an aggregate particle. Figure 7 shows the Pearson's correlation coefficients between the 14 shape factors analysed. It could be observed that several shape factors revealed very high linear correlations with others while some shape factors did not correlate with any other factors. FR-J.S, C-SF and PS-PSC showed perfect correlations indicating they had quantified the same aspects of the shape of an aggregate and could be considered surrogate parameters. Simultaneously, several others shown in green reveal very high correlation (proportional and inversely proportional indicated by positive and negative coefficients), indicating that they may be quantifying similar aspects of the shape of an aggregate but not the same aspects. Shape factors including SF.FR, BSF and KS showed no correlation to any other shape factors, indicating they quantify some shape aspects of aggregates that are not assessed by the remaining shape factors. This observation vouches for further analysis of dimensionality of the data matrix considered for this study.

3.3.2. Principal components analysis (PCA)

Principal Component Analysis (PCA) would be required to understand the dimensionality of the data matrix with regard to the total variance in the same. Considering all sizes of aggregates (C01), a KMO and Bartlett test was carried out which returned a Kaiser-Meyer-Olkin value of 0.697. This indicates that the subsequent PCA would be credible for interpretation. From the PCA, only three PCs had an eigenvalue of more than 1.00, whereas other PCs were considered insignificant that contained eigenvalues less than 1. Table 5 shows the percentage of the total variance in the data matrix explained by each principal component derived from the PCA. The first PC explains approximately half of the variance observed in the data matrix (46.50%) while all three PCs together explain more than 90% of the total variance. This indicates that 90% of the variance of the data matrix that had 14 dimensions could be explained with just three dimensions. Considering the noise in the data and assuming 10% could be attributed to the noise, the total data matrix could be reduced to 3 dimensions from the 14 dimensions that it represents.

Table 6 shows the rotated components (about varimax) from the PCA. First PC is dominated by J.R, J.AR, FAR, SF.FR, PSC, ICS and PS (above 0.85) indicating that these parameters could be considered as a single dimension. However, loading of these variables may be on either side of the PC1, indicating similar but not exact surrogate parameters, which may get an indication from Pearson's correlation. The second PC is dominated by J.C, J.S, FR, SF and C, where these too could be considered as a single dimension when the total variance of the data matrix is considered. Out of the 14 shape factors analysed, 12 factors have been aligned to two orthogonal principal components, where the remaining two shape factors KS and BSF lay along the third PC. Their loading along the third PC is not as affirmed as several other variables along first two PCs. In addition, KS, SF.FR and BSF have been observed not to correlate linearly with any other shape factors, which may indicate that they could well be individual dimensions. However, considering the total variance observed and the noise in the data should also be considered, the individual dimensions these shape factors may represent may not be very significant in the shape variations observed in the aggregates used in this study.

PCA among the cases C01–C07 was conducted and the rotated components in each case are given in Table 7. It could be observed that the PCs and the variables dominating in each PC for each case do not vary significantly except for the KS, BSF and SF.FR. All other variables have similar dominations along the principal components, and their impact on

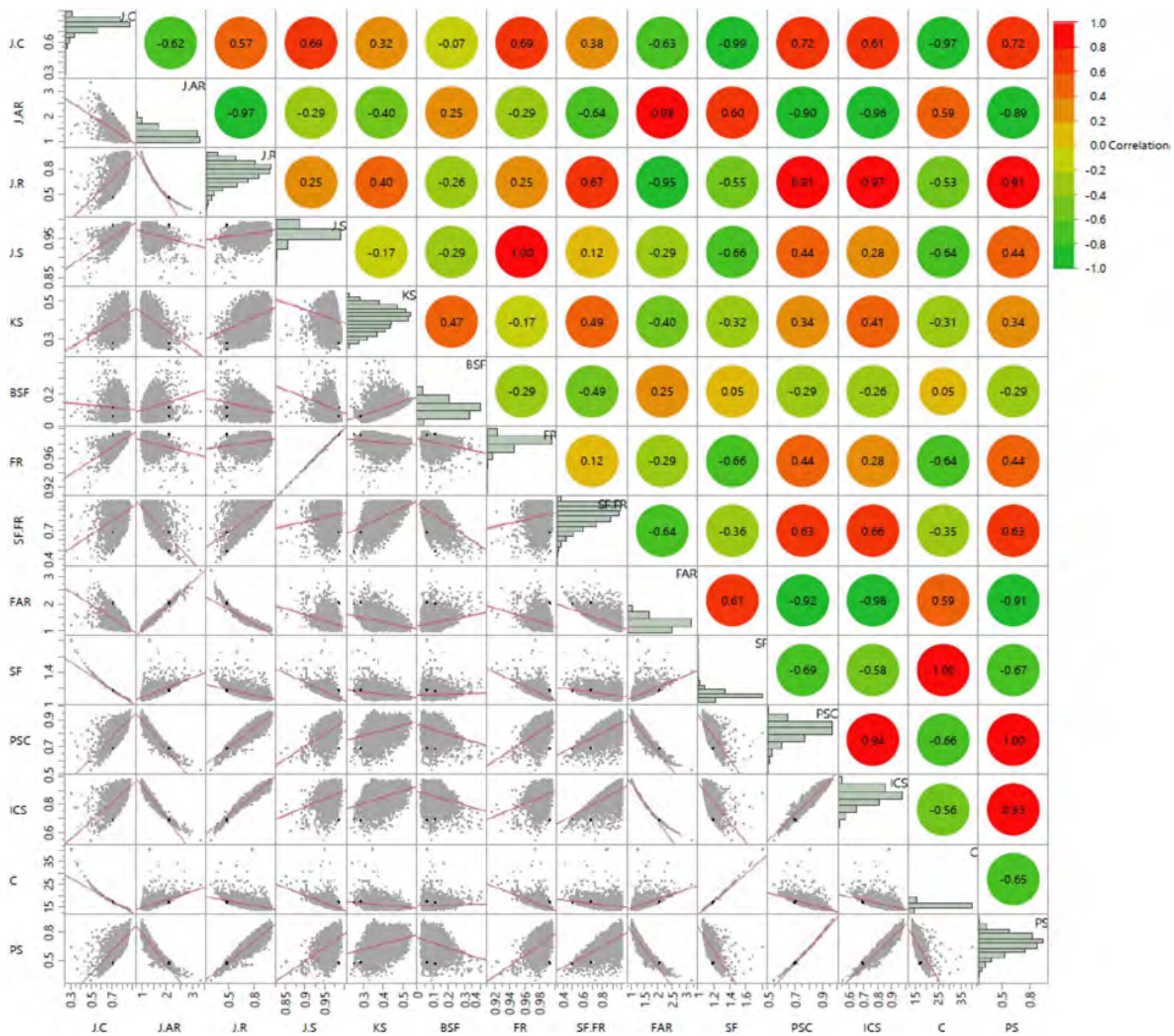

Figure 7. Pearson's correlation coefficients between shape factors.

the PCs also remained rather consistent across different aggregate sizes. In cases C02 and C05, only two PCs had eigenvalues more than 1, indicating that the 14 dimensions of the matrix could be reduced to just two dimensions. In those instances (C02 and C05), BSF, KS and SF.FR had significant loading along the first PC. How effective these shape factors are in classifying the aggregates based on similarities of shape will need to be assessed.

3.4. Cluster analysis from conventional statistical methods

A cluster analysis is done on all aggregate samples, ranging in sizes from 20 to 800 mm^2 of projected area. The aggregates belonged to five different degrees of milling as explained above, 0, 500, 1000, 1500 and 2000 Revs. The first analysis in clustering was done to restrict clustering to five clusters with the expectation that each group of aggregates with a different degree of milling would get clustered based on the 14 shape factors computed. The cluster analysis was done on the z-scores of the shape factors, to avoid bias due to the computed shape aspect values.

Table 5. Variation explained by each PC from PCA.

PC	% of Variance	Cumulative %
1	46.505	46.505
2	31.059	77.564
3	12.49	90.054

Table 6. Rotated (varimax) components and variables.

	PC1	PC2	PC3
J.C	0.437	0.850	0.197
J.AR	−0.924	−0.262	−0.060
J.R	0.945	0.202	0.038
J.S	0.072	0.900	−0.317
KS	0.446	−0.025	0.796
BSF	−0.360	−0.038	0.869
FR	0.072	0.900	−0.315
SF.FR	0.796	0.013	−0.080
FAR	−0.930	−0.266	−0.061
SF	−0.410	−0.847	−0.223
PSC	0.871	0.419	−0.001
ICS	0.948	0.239	0.045
C	−0.394	−0.838	−0.231
PS	0.868	0.411	−0.009

Table 7. Rotated components (varimax) for cases C01–C07g (orange and green colour indicate significantly high negative and positive correlations).

	PC1							PC2							PC3						
	C01	C02	C03	C04	C05	C06	C07	C01	C02	C03	C04	C05	C06	C07	C01	C02	C03	C04	C05	C06	C07
J.C	0.44	0.40	0.38	0.48	0.40	0.28	0.33	**0.85**	**0.87**	**0.88**	**0.82**	**0.86**	**0.93**	**0.92**	0.20		0.15	0.20		-0.02	0.10
J.AR	**-0.92**	**-0.87**	**-0.90**	**-0.87**	**-0.88**	**-0.94**	**-0.94**	-0.26	-0.37	-0.18	-0.30	-0.39	-0.11	0.02	-0.06		-0.33	-0.33		-0.24	-0.25
J.R	**0.95**	**0.89**	**0.92**	**0.87**	**0.90**	**0.93**	**0.93**	0.20	0.30	0.13	0.27	0.32	0.12	-0.04	0.04		0.32	0.35		0.27	0.26
J.S	0.07	0.13	0.04	0.16	0.08	0.05	-0.20	**0.90**	**0.92**	**0.95**	**0.95**	**0.90**	**0.95**	**0.93**	-0.32		0.02	0.09		0.08	0.07
KS	0.45	0.84	0.44	0.44	0.92	0.39	0.27	-0.02	0.26	0.14	0.22	0.14	0.21	0.13	0.80		0.76	0.84		0.87	0.91
BSF	-0.36	-0.86	-0.31	-0.27	-0.85	-0.21	-0.22	-0.04	-0.05	-0.07	-0.09	-0.09	0.14	-0.11	0.87		-0.86	-0.91		-0.94	-0.93
FR	0.07	0.13	0.04	0.16	0.08	0.05	-0.20	**0.90**	**0.92**	**0.95**	**0.95**	**0.90**	**0.95**	**0.93**	-0.31		0.02	0.09		0.08	0.07
SF.FR	0.80	0.87	0.39	0.35	0.90	0.30	0.21	0.01	0.12	0.09	0.14	0.05	0.03	0.11	-0.08		0.91	0.92		0.95	0.97
FAR	**-0.93**	**-0.87**	**-0.90**	**-0.89**	**-0.88**	**-0.95**	**-0.96**	-0.27	-0.37	-0.21	-0.28	-0.40	-0.14	-0.05	-0.06		-0.33	-0.33		-0.23	-0.21
SF	-0.41	-0.40	-0.38	-0.46	-0.40	-0.28	-0.34	**-0.85**	**-0.87**	**-0.88**	**-0.83**	**-0.82**	**-0.93**	**-0.91**	-0.22		-0.16	-0.19		0.03	-0.12
PSC	**0.87**	**0.81**	**0.83**	**0.81**	**0.87**	**0.81**	**0.91**	0.42	0.49	0.39	0.44	0.39	0.50	0.24	0.00		0.30	0.31		0.19	0.11
ICS	**0.95**	**0.89**	**0.91**	**0.89**	**0.89**	**0.94**	**0.96**	0.24	0.34	0.18	0.28	0.36	0.16	0.06	0.05		0.33	0.34		0.26	0.22
C	-0.39	-0.40	-0.37	-0.45	-0.39	-0.28	-0.34	**-0.84**	**-0.86**	**-0.88**	**-0.84**	**-0.80**	**-0.93**	**-0.90**	-0.23		-0.16	-0.18		0.04	-0.13
PS	**0.87**	**0.80**	**0.83**	**0.81**	**0.87**	**0.80**	**0.91**	0.41	0.48	0.39	0.43	0.38	0.51	0.24	-0.01		0.30	0.31		0.20	0.11

Figure 8 shows the cluster membership as a percentage of the total number of aggregates in each group based on degree of milling. For example, approximately 30 and 31% of the aggregates that had 0 Revs of milling got clustered in the first and second cluster, respectively, while another 17% was in cluster 4. It is therefore apparent that the five clusters did not cluster the aggregates based on degree of milling. This indicates that the distance between the aggregates computed based on 14 shape factors did not solely rely on the variations induced by the process of milling, and that they were either assessing several other aspects of the shape of aggregate that had been distributed within each group of aggregates grouped based on number of revolutions or that the morphology of aggregates in the 0 Rev group was highly distributed (as observed in the previous distribution frequency analysis). On the other hand, it could be observed that a significantly higher percentage of aggregates that had been milled got clustered in the 5th cluster as opposed to other clusters. This indicates that the 14 shape factors are identifying an aspect that was deformed by the process of milling. The fifth cluster had less than 10% of the 0 Rev aggregates while it had 34, 37, 51 and 46% of aggregates that had 500, 1000, 1500 and 2000 Revs. A closer observation reveals a slight increase in the percentage of aggregates from 1500 and 2000 Revs present in the 5th cluster, where the lower degree milled aggregates' presence was significantly lower. Observations on the 4th cluster reveal a higher presence of lower-degree-milled aggregates compared to higher-degree-milled aggregates, indicating a significant process of identification of degree of milling.

According to Figure 9, cluster 3 has cases that lay on the extremes of each shape factor (either tail, indicated by positive and negative z-scores). Most of the z-scores of the shape factors of the cases belonging to cluster 3 are above ±2, the cases lay beyond 2 standard deviations from the mean of each shape factor. This is evident from very few numbers of aggregates as could be seen in Figure 8, and they are mostly from unmilled aggregates. 2nd cluster has the next highest z-scores of shape factors indicating that they are also distinctively distant (different) from other clusters, and it corresponds to aggregates specifically pertaining to original

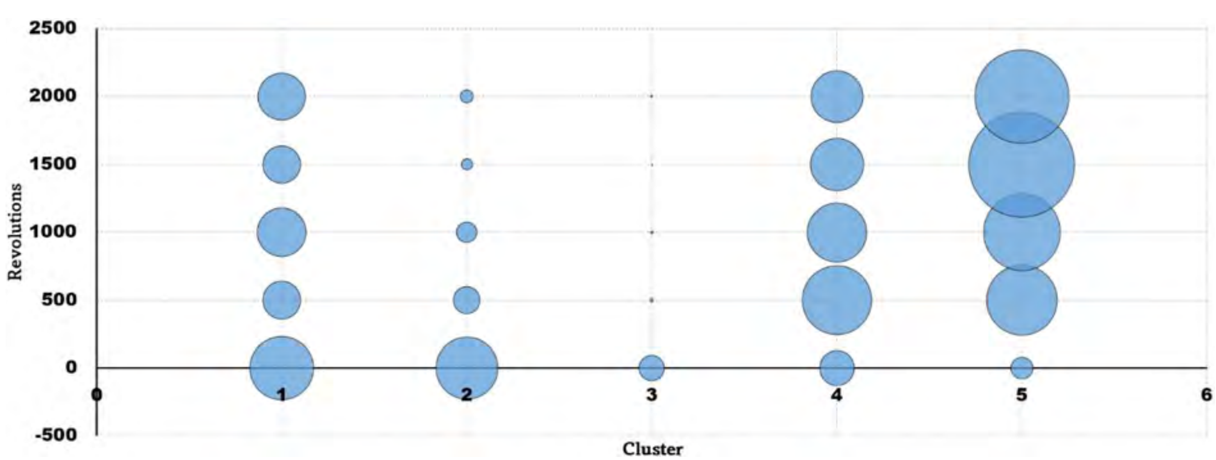

Figure 8. Cluster information on membership of aggregates among five clusters after KNN cluster analysis using all 14 shape factors.

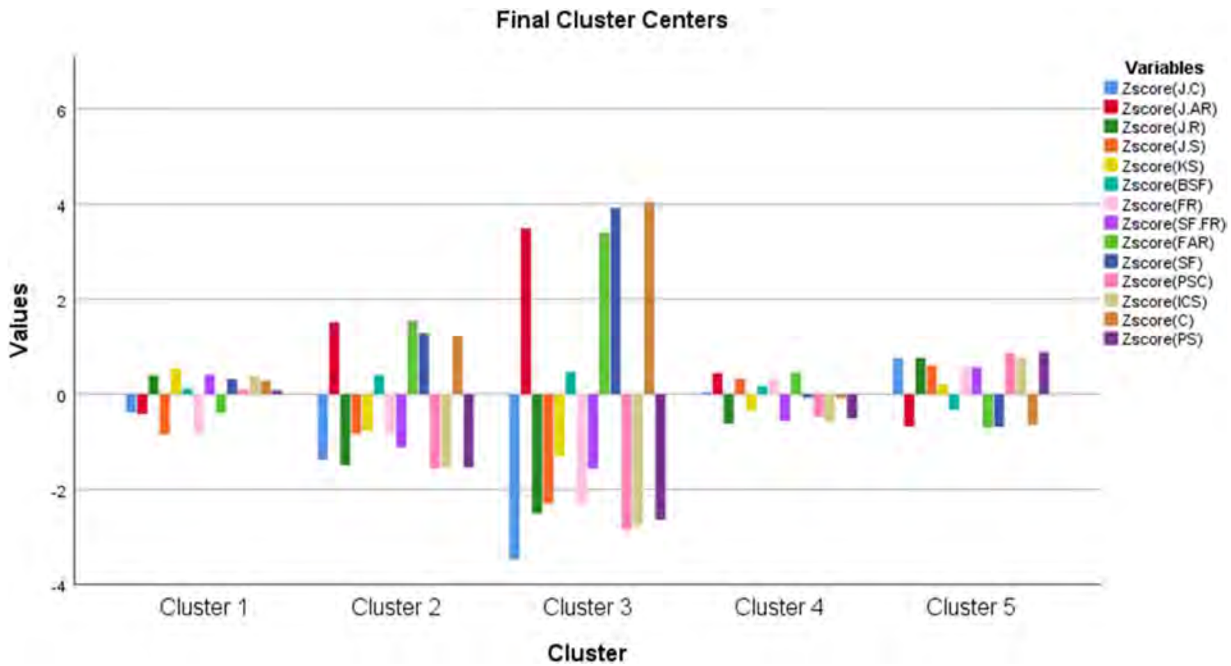

Figure 9. Final cluster centre information of the KNN cluster analysis.

crushed aggregates. Although the number of aggregates is less and they are mostly from unmilled aggregates group, aggregates from different degrees of milling are also present (approximately 8–14%), that increased steadily with degree of milling. The third cluster had the next distant centres based on shape factors which corresponded to aggregates that had a higher degree of milling. Another important observation is that the shape factors that had high positive z-scores in clusters 3 and 2 have now negative values and vice-versa in cluster 5. This indicates a significant shift in the shape factors relating to the process of milling.

To analyse the impact of dimension reduction on the clustering analysis, several combinations of shape factors were used in the KNN cluster analysis. From several trials, it could be observed that two shape factors, J.C and J.R were sufficient to obtain similar clustering as obtained using all 14 shape factors. Figure 10 shows the cluster membership percentage of each aggregate group across different combinations of shape factors. It could be observed that the cluster membership significantly changes across the combinations and that the impact of addition of different shape factors indicates that they quantify different aspects of shape that are varied by

JC/JR

Percentage of aggregates with Revs and Clusters

	1	2	3	4	5
0	39	30	17	11	3
500	17	16	3	33	32
1000	18	12	2	32	36
1500	11	6	0	31	51
2000	14	7	1	31	47

JC/JR/SF

Percentage of aggregates with Revs and Clusters

	1	2	3	4	5
0	36	35	20	7	1
500	41	18	3	38	0
1000	44	14	2	40	0
1500	35	7	1	57	0
2000	37	9	1	53	0

JC/JR/KS/BSF

Percentage of aggregates with Revs and Clusters

	1	2	3	4	5
0	44	29	19	6	1
500	23	6	23	31	18
1000	19	3	26	34	17
1500	6	1	25	46	22
2000	5	1	16	52	25

JC/JR/SF/PS

Percentage of aggregates with Revs and Clusters

	1	2	3	4	5
0	36	35	13	9	7
500	11	13	37	2	37
1000	13	10	36	1	40
1500	7	5	32	0	55
2000	9	7	32	1	52

JC/JR/KS/BSF/SF

Percentage of aggregates with Revs and Clusters

	1	2	3	4	5
0	61	17	11	8	3
500	25	2	32	10	31
1000	21	2	37	7	33
1500	6	1	39	6	49
2000	5	0	31	9	55

JC/JR/SF/PS/BSF/KS

Percentage of aggregates with Revs and Clusters

	1	2	3	4	5
0	48	40	6	4	3
500	29	11	37	1	22
1000	30	6	41	1	22
1500	16	2	55	0	27
2000	11	3	57	0	28

Figure 10. Summary of cluster membership percentages across different combinations of shape factors.

degree of milling. For example, the addition of SF to J.R and J.C grouped several aggregates in one group that had an almost equal percentage of all degree of milling (cluster 01) as opposed to other combinations of shape factors. Of all the combinations tested, the combination of J.R, J.C, KS and BSF yielded a distinct cluster membership where 0 Rev and 2000 Rev had been significantly isolated to the highest accuracy, while 500, 1000 and 1500 Revs had been distributed among the clusters.

An important observation made in this analysis is that the combination of shape factors was congruent to the analysis from PCA and Pearson's correlation. J.R was dominant in PC1, while J.C was dominant in PC2. In addition, KS and BSF were predominantly laid in PC3, and they were not linearly correlated. The other shape factors were dominant in either PC1 or PC2 and simultaneously were observed to linearly correlate to either J.C or J.R (proportional or inversely proportional), yielding a higher correlation coefficient. They could hence be considered surrogate parameters, and they did not significantly contribute to the clustering of aggregates. In addition, including redundant shape factors as such increased the blur in cluster membership, effectively reducing the distinct clustering of aggregates, specifically 0 Rev and 2000 Rev. Although SF was linearly correlated to J.C (inversely proportional), it could be observed from Figure 10, the addition of SF to the combination of 4 shape factors (J.C, J.R, KS and BSF) significantly enhanced the distinction in cluster membership.

From the trials, it was also observed that another combination of three shape factors, J.C, J.S and J.R also yielded a higher distinction of cluster membership, as shown in Figure 11. However, J.S is dominant in PC2, while it does not have a significant correlation to J.C or J.R or SF. Although it looks very similar to the membership observed where all 14 shape factors were used, the percentages are significantly better in the analysis with three factors alone. 50 and 55% of 2000 Rev and 1500 Rev group aggregates got clustered in the 5th cluster that only had 7% of 0 Rev aggregates.

When the number of clusters is set to 2, the combinations discussed have clustered more than 75% of 0 Rev aggregates to cluster 1 and more than 75% of all other aggregates to cluster 2 as shown in Figure 12. The membership of aggregates with different degrees of milling was observed to shift gradually between the two clusters as shown in the figure.

Several aspects may cause a blur in distinction between the clusters. Firstly, even though all aggregates were milled over an equal number of revolutions for a group based on the degree of milling, all aggregates may not have received equal treatment, owing to the milling action itself. The uncertainty in the treatment may have resulted in noise in the data. Secondly, not all original aggregates are of the same size and shape that may have undergone different morphological changes when equal treatment was given in the milling process. Thirdly, the aggregates may have a shape aspect that was not affected by the milling process, and that remained distributed across the aggregate groups with different degrees of milling. Fourthly, the shape factors considered in this study may not have quantified some shape aspects of the aggregates, that have undergone morphological change in the milling process. Fifthly, the process of clustering may not effectively capture the different degrees of impact each shape factor may have imparted on quantifying the shape of an aggregate. All these uncertainties may contribute to noise in the variance of data used in this study, reducing the effectiveness of clustering distinctively.

It could therefore be concluded that the degree of accuracy obtained in this analysis is significant and that the dimension reduction and correlation between shape factors are significant and valid. In essence, two combinations of shape factors could be considered statistically significant in defining the shape of an aggregate;

1. J.C, J.R, SF, BSF and KS
2. J.C, J.R, J.S, BSF and KS

A further reduction in dimensions of the data matrix may be considered, with J.C, J.R and either SF or J.C, with a compromise on the distinction between the cluster membership. To understand the effectiveness of the shape factor in clustering within two groups of aggregates, the KNN cluster analysis was conducted between two groups of aggregates (0 Rev–500 Rev, 0 Rev–1000 Rev and 0 Rev–2000 Rev). Figure 13 shows the results of two cluster analyses using different combinations of shape factors, indicated by the percentage of each group of

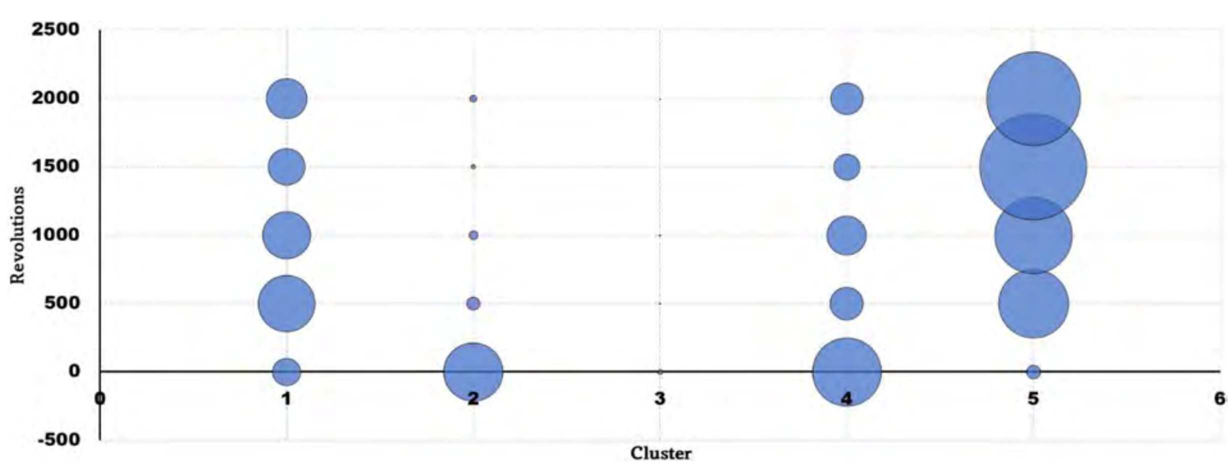

Figure 11. Cluster information on membership of aggregates among five clusters after KNN cluster analysis using J.C, J.S and J.R shape factors.

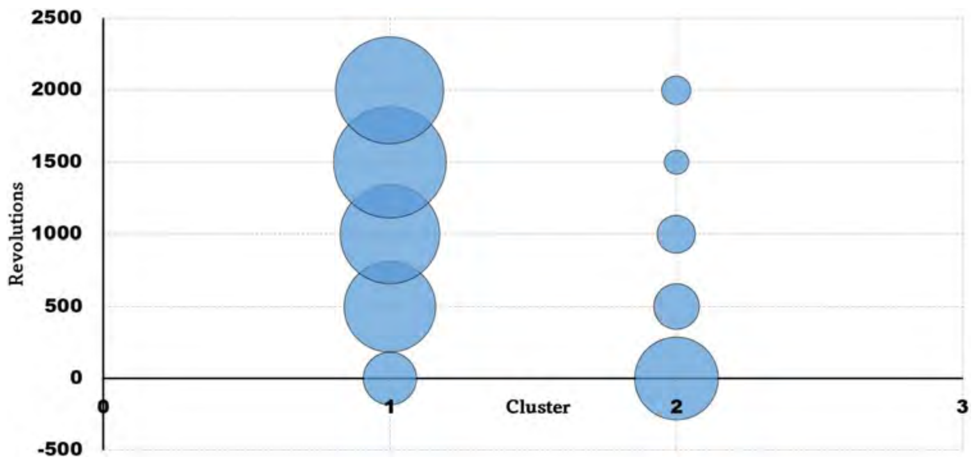

Figure 12. Cluster information on membership of aggregates among two clusters after KNN cluster analysis using J.C, J.S and J.R shape factors.

aggregates in each cluster. The combinations shown are the most effective clustering obtained over trials of various combinations. Rows 01 and 04 show the results of the analysis with 0 and 500 Rev aggregate groups, while rows 02 and 05 show that of 0 Rev and 1000 Rev and columns 03 and 06 show that of 0 and 2000 Rev. Simultaneously, the first column shows the results for the combination where all shape factors are used, while other columns show combinations of different shape factors as indicated on the top of each bar chart.

Should a combination of shape factors perfectly identify the morphological changes effected by the degree of milling, each cluster would have a representation of a single group of aggregates, indicated by just one colour for either of the bars (one be orange and the other be blue, with 100% representation of each group of aggregates). The general observation would be that no combination was perfect in clustering and that the clustering is consistently improved between 0–500 Rev, 0–1000 Rev and 0–2000 Rev. The observations from this analysis indicate, that the effectiveness of shape factors to represent morphological changes during milling does vary with degree of milling, as discussed in the descriptive statistics section. When all shape factors are considered for the analysis, even though a higher portion of milled aggregates are grouped to cluster 02, it failed to differentiate 0 Rev aggregates in significant distinction.

Considering Column 01 of Figure 13, a consistent improvement in clustering higher degree milled aggregates, the 0 Rev aggregates were equally distributed among cluster 01 and cluster 02. In addition, all other combinations shown in this figure have higher distinctions in clustering of aggregates, specifically 0 Rev aggregates, compared to the case where all shape factors are used. As suggested from earlier discussion based on Pearson's correlation analysis and Principal Component Analysis, a dimension reduction is not just adequate, but also essential.

Considering the combination of BSF/KS (column 02), a significant improvement was observed in clustering 0 Rev aggregates with 78% in cluster 01 and 22% in cluster 02, with a compromise on clustering efficiency of milled aggregates. The 0–500 Rev and 0–1000 Rev of the same combination even had not improved the milled aggregate clustering efficiency. Yet, a highly significant improvement is observed in the case of 0–2000 Rev, where 87% of 0 Rev got clustered in cluster 01 and 82% of 2000 Rev aggregates got clustered in cluster 02. Adding J.C to the combination (of BSF/KS) significantly improved all cases (column 03) where in the case of 0–2000 Rev, 92% of 0 Rev aggregates and 84% of 2000 Rev aggregates got clustered in clusters 01 and 02.

Considering the combination J.C/J.R/KS, a slight improvement in clustering milled aggregates is observed with a larger compromise on the 0 Rev aggregates clustering efficiency, compared to the combination discussed above. Although the clustering efficiency slightly increased for 0 Rev aggregates, 90% efficiency was observed in this combination for 2000 Rev aggregates. BSF and KS had a significant increment across degree of milling as opposed to other shape factors that saturated beyond 500 or 1000 Rev. This had effected a significant shift between 0-500 Rev analysis and 0-2000 Rev analysis. In addition, from the earlier discussion where J.C had a significant reduction in the range (frequency histogram) between 0 Rev and milled aggregate groups, it has contributed to improving the clustering of 0 Rev aggregates, where BSF and KS had similar distribution across all aggregate groups. On the other hand, J.R had a wider range of distribution across all aggregate groups, which negatively impacted the clustering efficiency in the third combination (column 04, rows 01-03) with the absence of contribution from BSF.

Considering the analysis with the combination of J.S and SF, it has a somewhat similar observation with the combination where all factors were used, with a significant improvement in clustering 0 Rev aggregates. This combination is observed not to contribute significantly to reduce noise and effectively cluster aggregates. However, as discussed above, the addition of BSF and KS to this combination has significantly improved the clustering efficiency and this further reinstates the quantification of a distinct shape aspect of aggregate that has undergone morphological change in the process of milling is represented in BSF and KS, distinctively. A combination of J.C, J.R, BSF and KS (column 04) showed rather less efficiency in clustering as opposed to other combinations discussed above. Although each shape factor may have a significant impact on the clustering, a combination of factors resulted in agglomerated impact that may have inhibitions from other factors in the combinations. Out of this analysis,

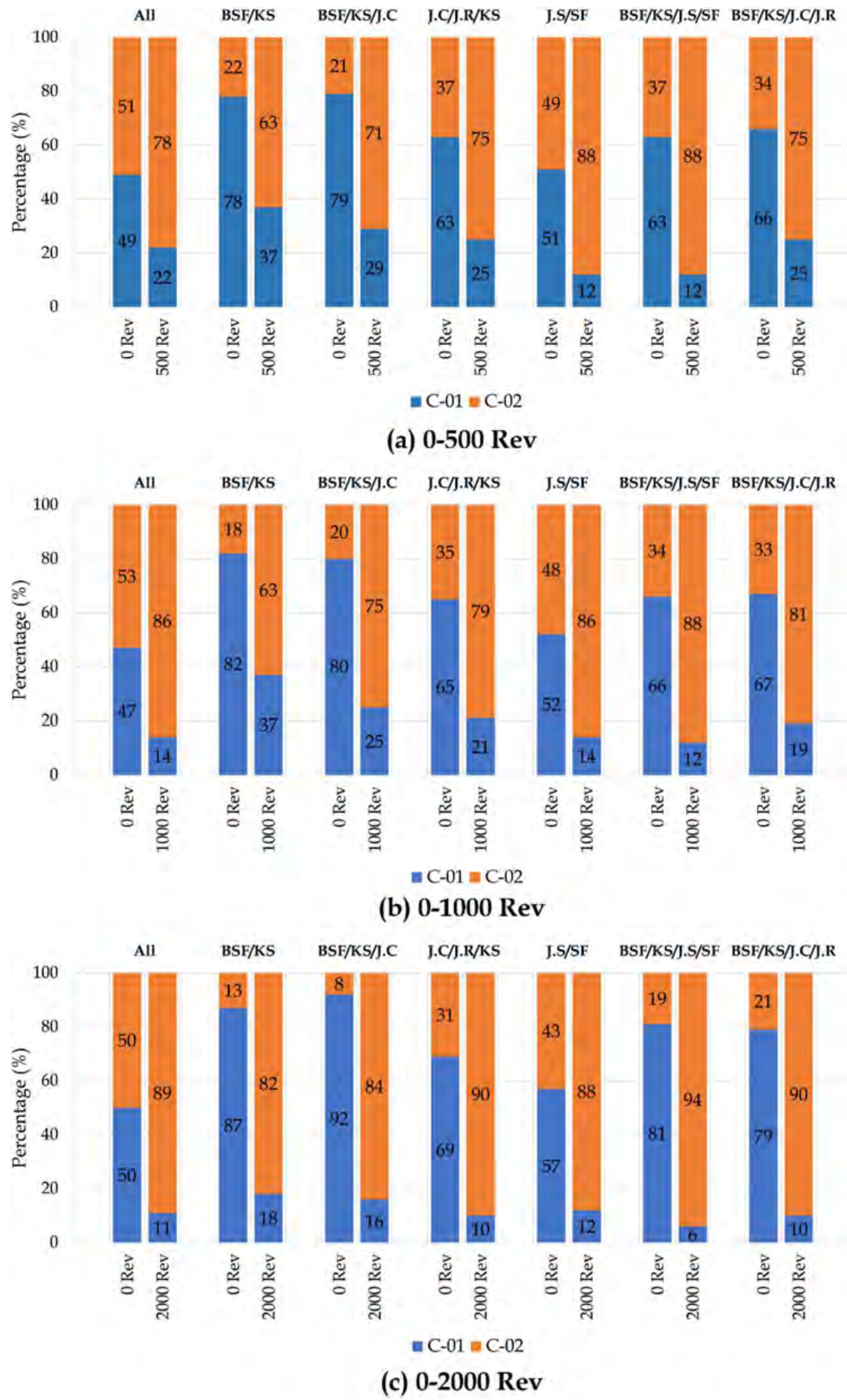

Figure 13. Two cluster analysis of various combinations of shape factors, indicated by the percentage of each group of aggregates in each cluster.

the best combinations were observed to be BSF/KS/J.C and BSF/KS/JSF/SF where BSF and KS are considered significant contributors in identifying the aspect of aggregates varied across degree of milling.

Further clustering attempts were done between 200–500 Rev, 500–1000 Rev and 1000–2000 Rev groups, using the same combinations and other combinations. None of them revealed a clustering efficiency more than 70%. This indicates the morphological changes induced by the process of milling beyond 500 revolutions were not quantifiable using the shape factors considered, or perhaps, the changes were trivial for numerical quantification based on dimensions.

Considering the distribution of shape factors within the group of aggregates, it could also imply that the other shape aspects within the group of aggregates were more pronounced than the morphological changes induced by higher degree of milling. It could also be possible that the noise in the data, or the impact of aggregate size on the computation of shape factors significantly reduced the representation of morphological aspects that were quantified by the shape factors.

4. Conclusions

This study analysed the aggregate shape factors, computed based on dimensions obtained from images of aggregates, in representing the shape aspects of aggregate particles ranging from 5 mm to 30 mm in diameter. Aggregates were obtained from naturally crushed rocks and milled in a ball mill for different number of revolutions (degree of milling) to instill micro and macro morphological shape aspects of aggregates. The following conclusions are made from the analysis.

- Shape factors had varied distribution patterns among the group of aggregates (degree of milling), that significantly changed across groups, indicating the ability to quantify different shape aspects. Most shape factors had a significant difference in mean value between 0 Rev group of aggregates and 500 Rev group, where the texture and angular flaky corners of aggregates were observed to change. Other than BSF and KS, other shape factors did not show a significant difference between the means 1000, 1500 and 2000 Rev aggregates. Histogram (frequency distribution) of shape factors showed a significant difference in the range and distribution, specifically for J.C and J.R between milled and unmilled aggregates, while KS and BSF showed similar distributions across the group of aggregates. The range of 95% confidence was significantly reduced for some shape factors such as J.R and J.AR, between milled and unmilled aggregates, indicating the effect of milling on the unique texture the crushed aggregates displayed. None of the shape factors showed a linear correlation with revolutions, indicating that none were capable of assessing the shape changes due to milling, as a single factor. A combination of factors is therefore necessary.
- Several of the shape factors had perfect or almost perfect linear correlations between them, indicated by Pearson's correlation coefficients 1 or close to 1. Shape factors in pairs such as SF/FR, J.R/J.AR, SF/C, J.S/FR, PSC/PS, SF/FR and J.R/ICS had perfect correlations and could be considered surrogate parameters. While most shape factors had significant correlations indicated by a value at least greater than 0.7 in Pearson's Correlation test, KS and BSF showed no significant correlation to any other shape factors. A Principal Components Analysis (PCA) suggested the 14-dimensional data matrix be reduced to 3 dimensions explaining more than 90% of the total variance. Considering the dominance shape factors had in defining the principal components and in congruence with the correlation between the shape factors, J.R from PC1, J.C and SF from PC2 and KS and BSF from PC3 were chosen as influencing shape factors for combinational analysis.
- A Five-cluster analysis revealed that neither of the combinations could effectively cluster aggregates congruent to the degree of milling. However, a significant trend in clustering the members across was observed with a reduced number of shape factors, the combinations as indicated by PCA and Pearson's correlations. A two-cluster analysis confirmed efficient clustering of aggregates congruent to degree of milling, specifically between milled and unmilled aggregates. The combinations BSF/KS/J.C and BSF/KS/J.S/KS as the best to cluster at an accuracy of approximately 90%. This also indicates that the BSF and KS are critical in quantifying the morphological changes instilled by degree of milling, where they were also observed as unique shape factors, that did not correlate with any other shape factors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by National Research Council Sri Lanka [grant number 19-045].

ORCID

Daniel Niruban Subramaniam http://orcid.org/0000-0002-1023-1617

References

Aamer Rafique Bhutta, M., *et al.*, 2013. Properties of porous concrete from waste crushed concrete (recycled aggregate). *Construction and Building Materials*, 47, 1243–1248.

Akçaoğlu, T., Tokyay, M., and Çelik, T, 2004. Effect of coarse aggregate size and matrix quality on ITZ and failure behavior of concrete under uniaxial compression. *Cement and Concrete Composites*, 26 (6), 633–638.

Al-Rousan, T., *et al.*, 2005. New methodology for shape classification of aggregates. *Transportation Research Record*, 1913 (1), 11–23.

Al-Rousan, T., *et al.*, 2007. Evaluation of image analysis techniques for quantifying aggregate shape characteristics. *Construction and Building Materials*, 21 (5), 978–990.

Altuhafi, F., O'Sullivan, C., and Cavarretta, I, 2013. Analysis of an image-based method to quantify the size and shape of sand particles. *Journal of Geotechnical and Geoenvironmental Engineering*, 139 (8), 1290–1307.

Anburuvel, A., and Subramaniam, D. N, 2022. Investigation of the effects of compaction on compressive strength and porosity characteristics of pervious concrete. *Transportation Research Record*, 2676 (9), 513–525.

Anburuvel, A., and Subramaniam, D. N, 2024. A novel multi-variable model for the estimation of compressive strength of pervious concrete. *International Journal of Pavement Research and Technology*, 17, 720–731.

Arasan, S., *et al.*, 2011. Correlation between shape of aggregate and mechanical properties of asphalt concrete. *Road Materials and Pavement Design*, 12 (2), 239–262.

Bangaru, R. S., and Das, A, 2012. Aggregate shape characterization in frequency domain. *Construction and Building Materials*, 34, 554–560.

Barksdale, R., *et al.*, 1991. Measurement of aggregate shape, surface area, and roughnes. *Transportation Research Record*, 1301, 107–116.

Bhattacharya, S., *et al.*, 2020. Estimation of 3D sphericity by volume measurement – application to coarse aggregates. *Transportation Geotechnics*, 23, 100344.

Brzezicki, J. M., and Kasperkiewicz, J, 1999. Automatic image analysis in evaluation of aggregate shape. *Journal of Computing in Civil Engineering*, 13 (2), 123–128.

Chan, K. W., and Kwan, A. K. H, 2014. Evaluation of particle packing models by comparing with published test results. *Particuology*, 16, 108–115.

Chang, C. S., and Deng, Y, 2020. Packing potential index for binary mixtures of granular soil. *Powder Technology*, 372, 148–160.

Ćosić, K., *et al.*, 2015. Influence of aggregate type and size on properties of pervious concrete. *Construction and Building Materials*, 78, 69–76.

Cruz-Matías, I., *et al.*, 2019. Sphericity and roundness computation for particles using the extreme vertices model. *Journal of Computational Science*, 30, 28–40.

Das, N., *et al.*, 2011. Modeling granular particle shape using discrete element method. *Geo-Frontiers*, 2011, 4293–4302.

de Larrard, F., and Sedran, T., 1994. Optimization of ultra-high-performance concrete by the use of a packing model. *Cement and Concrete Research*, 24 (6), 997–1009.

Deng, W., *et al.*, 2021. Shape effects on mechanical properties of maximally random jammed packings of intersecting spherocylinders. *Powder Technology*, 383, 443–453.

Ding, X., Ma, T., and Gao, W, 2017. Morphological characterization and mechanical analysis for coarse aggregate skeleton of asphalt mixture based on discrete-element modeling. *Construction and Building Materials*, 154, 1048–1061.

Dinger, D. R., and Funk, J. E, 2013. Particle-packing phenomena and their application in materials processing. *MRS Bulletin*, 22 (12), 19–23.

Evangelista, L., and de Brito, J., 2007. Mechanical behaviour of concrete made with fine recycled concrete aggregates. *Cement and Concrete Composites*, 29 (5), 397–401.

Fletcher, T., *et al.*, 2003. Aggregate imaging system for characterizing the shape of fine and coarse aggregates. *Transportation Research Record*, 1832 (1), 67–77.

Gan, Y., Kamlah, M., and Reimann, J, 2010. Computer simulation of packing structure in pebble beds. *Fusion Engineering and Design*, 85 (10), 1782–1787.

Ge, H., *et al.*, 2018. Three-dimensional characterization of morphology and abrasion decay laws for coarse aggregates. *Construction and Building Materials*, 188, 58–67.

Ghosh, S., Dhang, N., and Deb, A, 2020. Influence of aggregate geometry and material fabric on tensile cracking in concrete. *Engineering Fracture Mechanics*, 239, 107321.

Gong, F., *et al.*, 2016. A new method for characterizing coarse aggregate morphology through a MATLAB program. *Geo-China*, 2016, 53–60.

Gu, X., Tran, Y., and Hong, L, 2014. Quantification of coarse aggregate shape in concrete. *Frontiers of Structural and Civil Engineering*, 8 (3), 308–321.

Kim, Y., Suh, H. S., and Yun, T. S, 2019. Reliability and applicability of the Krumbein-Sloss chart for estimating geomechanical properties in sands. *Engineering Geology*, 248, 117–123.

Krejsová, J., *et al.*, 2018. The influence of different aggregates on the behavior and properties of gypsum mortars. *Cement and Concrete Composites*, 92, 188–197.

Kusumawardani, D. M., and Wong, Y. D, 2020. Evaluation of aggregate gradation on aggregate packing in porous asphalt mixture (PAM) by 3D numerical modelling and laboratory measurements. *Construction and Building Materials*, 246, 118414.

Kwan, A. K. H., Ng, P. L., and Huen, K. Y, 2014. Effects of fines content on packing density of fine aggregate in concrete. *Construction and Building Materials*, 61, 270–277.

Lewis, D. W., and McConchie, D., 1994, Textures. *In:* D. W. Lewis and D. McConchie, eds. *Analytical sedimentology*. Boston, MA: Springer US, 92–129.

Lin, W., *et al.*, 2016. Development of permeability test method for porous concrete block pavement materials considering clogging. *Construction and Building Materials*, 118, 20–26.

Liu, P., *et al.*, 2017. Modelling and evaluation of aggregate morphology on asphalt compression behavior. *Construction and Building Materials*, 133, 196–208.

Maerz, N. H, 2004. Technical and computational aspects of the measurement of aggregate shape by digital image analysis. *Journal of Computing in Civil Engineering*, 18 (1), 10–18.

Mahmoud, E., Masad, E., and Nazarian, S, 2010. Discrete element analysis of the influences of aggregate properties and internal structure on fracture in asphalt mixtures. *Journal of Materials in Civil Engineering*, 22 (1), 10–20.

Man, H. K., and van Mier, J. G. M., 2011. Damage distribution and size effect in numerical concrete from lattice analyses. *Cement and Concrete Composites*, 33 (9), 867–880.

Maroof, M. A., Mahboubi, A., and Noorzad, A, 2020, July. A new method to determine specific surface area and shape coefficient of a cohesionless granular medium. *Advanced Powder Technology*, 31 (7), 3038–3049.

Masad, E., Button, J. W., and Papagiannakis, T, 2000. Fine-aggregate angularity: automated image analysis approach. *Transportation Research Record*, 1721 (1), 66–72.

Mora, C. F., and Kwan, A. K. H, 2000. Sphericity, shape factor, and convexity measurement of coarse aggregate for concrete using digital image processing. *Cement and Concrete Research*, 30 (3), 351–358.

Mostofinejad, D., and Reisi, M, 2012. A new DEM-based method to predict packing density of coarse aggregates considering their grading and shapes. *Construction and Building Materials*, 35, 414–420.

Mueller, G. E, 2010. Radial porosity in packed beds of spheres. *Powder Technology*, 203 (3), 626–633.

Ng, P.-L., Kwan, A. K.-H., and Li, L. G, 2016. Packing and film thickness theories for the mix design of high-performance concrete. *Journal of Zhejiang University-SCIENCE A*, 17 (10), 759–781.

Pouranian, M. R., and Haddock, J. E, 2018. Determination of voids in the mineral aggregate and aggregate skeleton characteristics of asphalt mixtures using a linear-mixture packing model. *Construction and Building Materials*, 188, 292–304.

Rueden, C. T., *et al.*, 2017. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, 18 (1), 529.

Sajeevan, M., *et al.*, 2023. Investigation of compaction on compressive strength and porosity of pervious concrete. *International Journal of Pavement Research and Technology*. https://doi.org/10.1007/s42947-023-00377-w

Sajeevan, M., Ahilash, N., and Subramaniam, D. N. 2022. Investigating impact of boundary layer in pervious concrete. *In:* Proceedings Road and Airfield Pavement Technology, Cham, Springer International Publishing, 431–442.

Sajeevan, M., and Subramaniam, D. N, 2022. Investigation of boundary layer impact on pervious concrete. *International Journal of Pavement Engineering*, 24 (2), 1–17.

Sathiparan, N., Jeyananthan, P., and Subramaniam, D. N, 2024. Effect of aggregate size, aggregate to cement ratio and compaction energy on ultrasonic pulse velocity of pervious concrete: prediction by an analytical model and machine learning techniques. *Asian Journal of Civil Engineering*, 25, 495–509.

Shen, P., *et al.*, 2019. The effect of curing regimes on the mechanical properties, nano-mechanical properties and microstructure of ultra-high performance concrete. *Cement and Concrete Research*, 118, 1–13.

Shuguang, L., and Qingbin, L, 2015. Method of meshing ITZ structure in 3D meso-level finite element analysis for concrete. *Finite Elements in Analysis and Design*, 93, 96–106.

Su, D., and Xiang, W, 2020. Characterization and regeneration of 2D general-shape particles by a Fourier series-based approach. *Construction and Building Materials*, 250, 118806.

Su, D., and Yan, W. M, 2018. Quantification of angularity of general-shape particles by using Fourier series and a gradient-based approach. *Construction and Building Materials*, 161, 547–554.

Subramaniam, D. N., *et al.*, 2018. Dynamics of clay particles in non-vegetated stormwater biofilters. *Water, Air, & Soil Pollution*, 229 (9), 302.

Subramaniam, D. N., Hareindirasarma, S., and Janarthanan, B, 2022. An alternative approach to optimize aggregate-to-cement ratio and compaction in pervious concrete. *Arabian Journal for Science and Engineering*, 47 (10), 13063–13071.

Tafesse, S., Fernlund, J. M. R., and Bergholm, F, 2012. Digital sieving-Matlab based 3-D image analysis. *Engineering Geology*, 137-138, 74–84.

Wang, L., *et al.*, 2005. Unified method to quantify aggregate shape angularity and texture using Fourier analysis. *Journal of Materials in Civil Engineering*, 17 (5), 498–504.

Wijekoon, S. H., *et al.*, 2024. A mathematical model to predict the porosity and compressive strength of pervious concrete based on the aggregate size, aggregate-to-cement ratio and compaction effort. *Asian Journal of Civil Engineering*, 25, 67–79.

Xie, X., *et al.*, 2017. Evaluation of morphological characteristics of fine aggregate in asphalt pavement. *Construction and Building Materials*, 139, 1–8.

Xu, R., and Di Guida, O. A., 2003. Comparison of sizing small particles using different technologies. *Powder Technology*, 132 (2), 145–153.

Xu, H., and Prozzi, J. A, 2019. Correlation of aggregate form properties with Fourier frequencies. *Transportation Research Record*, 2674 (8), 405–419.

Zhang, D., Huang, X., and Zhao, Y, 2012. Investigation of the shape, size, angularity and surface texture properties of coarse aggregates. *Construction and Building Materials*, 34, 330–336.

Zhang, S., Li, R., and Pei, J, 2019. Evaluation methods and indexes of morphological characteristics of coarse aggregates for road materials: a comprehensive review. *Journal of Traffic and Transportation Engineering (English Edition)*, 6 (3), 256–272.

Zhao, L., *et al.*, 2020. A digitalized 2D particle database for statistical shape analysis and discrete modeling of rock aggregate. *Construction and Building Materials*, 247, 117906.

Zheng, J., *et al.*, 2020. Three-dimensional particle shape characterizations from half particle geometries. *Powder Technology*, 367, 122–132.

Zheng, J., and Hryciw Roman, D, 2016. Roundness and sphericity of soil particles in assemblies by computational geometry. *Journal of Computing in Civil Engineering*, 30, 04016021.

Zhu, Z., Provis, J. L., and Chen, H, 2018. Quantification of the influences of aggregate shape and sampling method on the overestimation of ITZ thickness in cementitious materials. *Powder Technology*, 326, 168–180.