QUANTITATIVE ANALYSIS OF CARBON BLACK AGGLOMERATES MORPHOLOGY IN ELASTOMER COMPOSITES BASED ON X-RAY TOMOGRAPHY BY MEANS OF NUMERICAL CLUSTERING

Jesbeer Kallungal, Laurent Chazeau, Jean-Marc Chenal, Jåérôme Adrien, Eric Maire, Claire Barrès

Jesbeer Kallungal, Laurent Chazeau, Jean-Marc Chenal, Jåérôme Adrien, Eric Maire, et al.. QUANTITATIVE ANALYSIS OF CARBON BLACK AGGLOMERATES MORPHOLOGY IN ELASTOMER COMPOSITES BASED ON X-RAY TOMOGRAPHY BY MEANS OF NUMERICAL CLUSTERING. Rubber Chemistry and Technology, 2023, 96 (1), pp.90-113. 10.5254/rct.22.77979 . hal-04697140

HAL Id: hal-04697140
https://hal.science/hal-04697140v1

Submitted on 13 Sep 2024

Quantitative analysis of carbon black agglomerate morphology in elastomer composites based on X-ray Tomography by means of numerical clustering

Jesbeer. Kallungal[1,2,3,*], Laurent. Chazeau[1,*], Jean-Marc. Chenal[1], Jérôme. Adrien[1], Eric. Maire[1]

[1] *Univ Lyon, INSA Lyon, CNRS, MATEIS UMR5510, F-69621, Lyon, France*

Claire. Barrès

[2] *Univ Lyon, INSA Lyon, CNRS, IMP UMR5223, F-69621, Lyon, France*

Bernard. Cantaloube, Patrick. Heuillet

[3] *LRCCP Vitry-sur-Seine, France*

Corresponding authors: jesbeer.kallungal@insa-lyon.fr and laurent.chazeau@insa-lyon.fr

Permanent adresse: Laboratoire MATEIS (PVMH), Bâtiment Blaise Pascal, 5° étage, 7 Avenue Jean Capelle, 69621 Villeurbanne cedex.

Keywords: defects, x-ray tomography, elastomers, microstructural analysis, carbon black agglomerates.

ABSTRACT

In this paper, a novel methodology for characterizing the morphology distribution of filler agglomerates in Elastomer composites is presented, based on laboratory sourced X-ray Tomography. Various feature extraction methods (via e.g. Image processing filters, segmentation) and selection tools (Spearman's rank correlation coefficient) combined with K-means unsupervised clustering algorithm were developed for identifying the distinct morphological classes in model materials (carbon filled ethylene propylene diene monomer rubber). The interest of this methodology was demonstrated by precisely differentiating the materials compounded with different processing parameters. For instance, in this example, thanks to this analysis, it was found that introducing the filler before the elastomer in internal mixer tends to favor more structured agglomerates.

1 INTRODUCTION

Elastomers are used in many applications such as load-bearing components, anti-vibration mounts, seals etc… The optimisation of properties for these applications requires the addition of many ingredients in the pristine polymers. The ingredients used are fillers (Carbon black CB, Silica), processing aids (paraffin wax), crosslinking agents (Sulphur, Peroxide), cure activators (zinc Oxide ZnO, Stearic Acid) and accelerators (to control the cure rate). A good dispersion of these ingredients, especially fillers, is achieved through their mixing, in an internal mixer and a two roll mill machine. Studies have shown that the choice of processing parameters such as fill factor, shear rate, ram pressure applied in the internal mixer, mixing energy etc… impacts the morphology (e.g size, shape factor), the spatial distribution and the volume fraction of the filler agglomerates (hereafter referred as CB_{aggl})[1–5]. Apart from them, other microstructural heterogeneities can also exist in the rubber compounds such as unreacted ZnO inclusions, voids etc..[6] Studies have demonstrated the potential impact of all these so called “flaws” or “defects” on crack mechanisms during monotonic and dynamic solicitation of the material[7–16]. Different research works emphasize the importance of quantifying their morphology. For instance, Gent *et al.* proved the dependence of the initiation of cavities in silicone elastomer on the size and spatial distribution of model flaws (using spherical glass beads)[17]. In addition, Chow *et al.* have shown theoretically the importance of the aspect ratio of rigid flaws on these mechanisms, as it impacts the stress concentration at their poles[18].

In literature, the characteristics of filler agglomerates of carbon black (CB_{aggl}) in an elastomer are obtained from the roughness measurement of their micro-tomed sections (ASTM D2663), or from the light reflexion analysis of a freshly cut sample surface with a Dispergrader technique or from high resolution electron microscopy such as SEM or TEM[10,19]. However, the main drawback of these techniques is that they can only provide 2D information and therefore a full description of the morphology or spatial distribution of CB_{aggl} cannot be ascertained. X-ray Tomography can be used to obtain 3D information on the polymer at the micron scale[20–24]. Due to the similar attenuation contrast of filler agglomerates (CB_{aggl}) and elastomer matrix, to the authors' knowledge, only one paper reports their detection and morphological characterization using non-synchrotron sourcing X-ray tomography in an elastomer composite[25]. To do so, the authors had to adapt the laboratory Tomography parameters. However, a methodology is missing to extract, from the obtained data, the pertinent descriptive features, which can help in relating the agglomerates morphology and spatial dispersion to the material mechanical properties. This is the motivation of the work presented here.

Such methodology could use different tools that have been developed for the classification of filler aggregates. Medalia *et al.* [26–29], through a series of publications, used various Euclidean geometric features, such as bulkiness factor, anisometry, area and structuring factor, to analyse TEM images. Differently, Bourrat *et al.*[30] proposed to use mass fractal analysis, which also applied to the same type of images. This mass fractal approach was extended later by Rieker *et al.*[31] for N330 carbon black dispersed in a polymer matrix, characterized by small-angle X-ray scattering (SAXS) which provides information in the bulk. It was also used by Baeza *et al*.[32] for analyzing the morphology of silica aggregates in an elastomer. Herd *et al* [33] have shown, when analyzing the morphology of nineteen different carbon blacks dispersed in an elastomer, that the results of both approaches, Fractal *vs* Euclidean geometry, are correlated[22]. However, Euclidean geometry approach appears to be the most pertinent for specific shape classification and for this reason, will be preferred in this paper.

The outline of this article is as follows. The proposed methodology for analyzing filler agglomerates in elastomers is developed using experimental data obtained from model materials. After the description of the model materials and testing methods in section 2, feature extraction and selection techniques are presented (section 3). Following the review of clustering techniques for morphological clustering (section 4), the results and discussion on the classifications are then presented in section 5 and 6 respectively. Such classification is performed for the agglomerates morphology in the different model materials. Our methodology enables to differentiate two materials with identical composition, not only by the volume fraction of CB_{aggl}, but also through various morphological features. It also provides a valuable tool to understand the impact of operating parameters during the compounding process on the morphology of CB_{aggl}.

2 MATERIALS AND METHODS

2.1 Composition and materials processing

Model elastomers have been fabricated using a non-crystallizing Ethylene Propylene Diene Monomer Rubber (EPDM), and an N300 series carbon black, N326 as reinforcing filler. Table I presents the compound composition, which was kept constant, and the compounding process conditions, which were varied. The crosslinking agent is unsupported Bis(α,α-dimethylbenzyl) peroxide. The processing of the samples was performed in a Haake Rheomix 600 OS mixer (chamber volume: 120 cm^3) using Banbury rotors (42 cm^3). All the components, except peroxide, were mixed in the internal mixer (IM) for 2 min, and then passed 10 times in the two roll-mill where the peroxide was added. The rotor speed of the internal mixer is specified for each material in Table I, as well as the sequence for introducing the ingredients. USD protocol refers to "upside down", i.e. the elastomer was added after carbon black. The resulting model materials from this procedure are referenced as 30_rpm_usd and 60_rpm_usd. TD protocol refers to "top down", i.e. carbon black was introduced after the elastomer in the internal mixer, and the resulting model materials are referenced as 30_rpm_td and 60_rpm_td. After the mixing steps, the samples were molded as 2mm thick films by hot pressing at 170°C under 190 MPa for 12 min. Cure time was estimated from the torque measurements, performed using a Monsanto rheometer analyzer.

2.2 X-ray tomography

The X-ray Tomography device is EasyTom from RX solutions. The Tungsten target current is 45 µA and a LaB_6 cathode filament is used, along with a CCD detector. Two sample specimens extracted from each material is used for the experiments (only one was used for morphological clustering analysis). Cutting of the specimen is performed with a scalpel. Specimen length and width is adapted for monitoring within the field of view of X-ray Tomography: The volume analyzed is in the order of 4-6mm^3. The sample is rotated over 360 degrees in front of the X-ray beam. Image acquisition is performed every 0.09° (4000 projections in total). Each projection step lasted 0.45s resulting in a total time per scan of 30 minutes. The voxel resolution in these experiments is 1μm^3. Imaging was performed at a tube voltage of 40kV for an optimised attenuation contrast between the elastomer and agglomerates as discussed in a previous paper[25] .

3 FEATURE EXTRACTION

3D images obtained from X-ray Tomography show different flaws like carbon black agglomerates, voids, etc... The feature extraction methodology involves applying different algorithms of image processing to selectively display CB_{aggl} from the acquired tomography scans. Thereby, various statistical information related to their morphology can be obtained. A methodology to extract the CB_{aggl} spatial distribution was also applied on the same samples. It is described in Appendix A along with the obtained results.

3.1 Image processing

Various functions such as geometry offset, ring filter, and phase contrast are applied to the raw images to reduce artefacts and to improve the contrast between the matrix and the filler particles. Image stacks are then built using a filtered back projection reconstruction algorithm. Several Python modules are then used for image processing on the 3D reconstructed volume. The general approach of feature segmentation by segregating the greyscale cannot be applied in our case, since there is no significant difference in the attenuation contrast between the matrix and the CB_{aggl}. Therefore, various image filters are applied again on the image stack (Figure 1a) to optimise the contrast and to remove additional noise. Typical denoising filters such as the median filter are not effective. For this reason, a different workflow of image processing is devised, which is inspired from the works of Emmanuelle Gouillart[34,35]. After intensity rescaling (to improve the brightness and contrast) (Figure 1b), bilateral (Figure 1c) and non-local means (Figure 1d) filters are applied on the image stack. In bilateral filter, the intensity of each pixel in the image is replaced with a weighted average intensity of nearby pixels for noise reduction and edge preservation. In non-local means filter (effective for images with fine-grained texture), each target pixel value is replaced by an average of all the neighbourhood pixels value, weighted by how similar these pixels are to the target pixel[36]. Subsequently, segmentation of the image is carried out, which is a crucial step for obtaining pertinent information. Here we are using a supervised segmentation algorithm called Random Walker for which was developed by Grady to reduce manual intervention and user bias[36,37]. The segmentation is directly carried out on 3D images. The different image treatments have been implemented through the scikit-image module in Python[38]. The output of the image processing and segmentation is presented in Figure 1e. The final process of removal of the particles touching the edges of the stack, as these may intervene in the morphological classification of CB_{aggl}, is performed using ImageJ © free software. A sample sub volume of the reconstructed 3D image is shown in Figure 2 and Appendix B.

3.2 Morphological features extracted

Each agglomerate is labelled in the binary image after its segmentation from the matrix. The X ray tomography resolution being 1µm (i.e one voxel), and given the possible error in the identification of agglomerates border (at least 1 voxel), agglomerates with volume below 5x5x5 (i.e. 125 voxels) have not been considered. This way, we reduce the errors in the calculation of the morphological parameters of the small agglomerates (note also that this volume roughly corresponds to the minimum cutoff volume in the D2663 ASTM method for evaluating carbon black dispersion in rubber*)*. Geometric features such as Eccentricity E, Anisometry Q, Bulk factor B, Volume V, Equivalent Diameter Eq_{diam}, Length of major principle axis a, and Length of minor principle axis b of each CB_{aggl} are extracted using the *regionprops* module in Python[38].

The volume V of a CB_{aggl} is calculated by counting the number of voxels it contains. The Equivalent diameter (Eq_{diam}) represents a diameter of a fictive spherical inclusion having the same volume as the analyzed CB_{aggl}. The Eq_{diam} of CB_{aggl} with a volume of 125µm^3 (1 voxel =1µm^3) is equal to 6.2µm. Length of major (a) and minor principal axis (b) are those of an *ellipsoid* that has similar normalized 2nd central moment[36] as the CB_{aggl}. Eccentricity (E) of an ellipse is defined as the ratio between its distance of two foci and the major principal axis length.

Here, E is the eccentricity of the centered elliptic section of the ellipsoid, perpendicular to the c axis (with $a > c > b$):

$$E = \sqrt{1 - b^2/a^2} \quad (1)$$

For E=0, the morphology is a sphere and for E close to 1 the morphology get closer to that of a 1D or 2D object. Anisometry (Q) is the ratio of radii of gyration and corresponds to anisometry of the agglomerates as calculated by Medalia *et al.* [26]:

$$Q = b/a \quad (2)$$

Bulk factor (B) corresponds to the ratio of the free space (volume in this case) between the agglomerates and the convex hull constructed around the agglomerate (ΔV) to the total volume of the convex hull (V_{ch}), as shown in Figure 3:

$$B(\Delta V) = 1 - V/V_{ch} \quad (3)$$

This definition is slightly different from the approach used by Medalia *et al.* [26] who proposed to express bulkiness as a ratio of the area of radius-equivalent ellipse drawn around the object to its projected area.

3.3 Feature selection

Using all the features V, Eq_{diam}, B, E, b, a, and Q for clustering presents some disadvantages. Firstly, if the two relevant features chosen are strongly correlated to each other, they add more weight to the classification process and thereby generate biased clusters. Secondly, the computation complexity of clustering increases as the number of features increases. Therefore, feature selection techniques have to be applied to evaluate the relevance of each one. Such step is very critical when defining the morphology classes by unsupervised pattern recognition, since we lack prior knowledge of each morphology class features. First, the relevance of each feature was checked using Laplacian Score technique (described in Appendix C). Then Spearman's rank correlation coefficient (ρ) has been used to discard the redundant feature. ρ characterises the monotonic relations (linear or not) between two features. It varies between -1 to 1. A value close to |1| implies that they are correlated. Here, if two features have a coefficient value below |0.7|, they will be considered to be distinct. All the calculated ρ values are reported in Figure 4 along with their distribution. All data from all studied materials have been used for the analysis. They show that a, b, and Eq_{diam} are correlated to each other and also to V. As expected, E is negatively correlated to Q (aspect ratio). B is only correlated to a and feature a is correlated to other features (V, Eq_{diam} and b). Feature B has been retained because it can provide information on the fact whether the agglomerates are structured or not and the two other relevant retained features are V (size) and E (Eccentricity).

The distributions of V, E and B for the 4 model materials are plotted in Figure 5. V distribution is similar for all the materials, with a slight difference in the domain $1.25*10^3$-$2.25*10^3$ µm^3. The relative frequency is lower for the materials processed with the upside-down protocol (USD) in this domain (cf Figure 5a, b). The distribution curves of B (structure factor) have a similar shape for all the materials (Figure 5c). The curves also indicate that the agglomerates of the 60_rpm_td are the less structured, and those in the 30_rpm_usd are the most structured. The E distributions of the 4 materials (cf Figure 5d) have similar Gaussian shape with mean values around 0.75 for all the materials, indicating that their agglomerates are not spherical.

To sum up, there are only slight differences between the 4 materials, regarding their E, B, and V distributions. Moreover, as they can only be discussed separately, it is impossible to know for instance, if the largest number fraction of agglomerates with a low structure factor in the 60 rpm_td is related to the smallest or the biggest agglomerates. In other words, these distributions do not provide a clear picture of the differences in the agglomerates morphology. This motivates the use of the clustering methodology described in the following.

4 CLUSTERING

Manual clustering of all agglomerates using all the relevant features is a cumbersome task. Some authors have recently used unsupervised clustering techniques (Machine learning algorithms) such as hierarchical clustering, or principle component analysis [39] to solve this problem[39–41] . For practical reasons, we chose K-means unsupervised clustering algorithm. Classification using other clustering algorithms (agglomerative clustering, hierarchical clustering etc...) did not provide interpretable results. K-means algorithm was implemented using scikit-learn machine learning module in Python[42] .

4.1 K-means unsupervised clustering

K-means clustering algorithm attributes a set of data points x_i (x_i is defined by p coordinates) describing n agglomerates $(1 \leq i \leq n)$ by p features (V, B, E)) into k clusters C= $\{C_1, C_2,..,C_k\}$. Each cluster can be described by the mean μ_j of the observations in the cluster, commonly called as cluster centroid. The algorithm tries to find a centroid that minimizes the inertia or the variance of the squared Euclidean distance within the cluster, $||x_i - \mu_j||^2$

$$\sum_{j=1}^{k} \sum_{x \in C_j} \left\| x_i - \mu_j \right\|^2 \qquad (4)$$

As the number of clusters k is not known *a priori*, 2 validation criteria were used to determine the optimal number of clusters.

- Calinski Harabasz (*CAH*): this ratio is sometimes called variance ratio criterion.

  $$CAH = (SS_B/SS_W) \times (N - k) / (k - 1) \qquad (5)$$

  Where SS_B is the overall “between-cluster” variance, SS_w is the overall “within cluster” variance, k is the number of clusters and N is the total number of data points (i.e. agglomerates). *CAH* metric works very well for clustering algorithm based on squared Euclidean distances such as K-means[43] . The ideal number of clusters corresponds to the highest *CAH* score.

- The Davies and Bouldin index (*DBI*) can be used to infer the average similarity of a cluster with its most similar cluster[44] . In this context, it is given by the equation as follows:

  $$DBI = \frac{1}{k}\sum_{i=1}^{k} \max_{j \neq i}\left(\frac{s_i + s_j}{d_{ij}}\right) \quad (6)$$

  s_i is the average distance between each data points in cluster i and its centroid (similarly for s_j) , d_{ij} is the average distance between the clusters centroids i and j. *DBI* index is the lowest for the ideal number of clusters.

The workflow deployed in this article to have a generalized morphological cluster analysis is schematized in Figure 6. Across-over sampling between model materials 30_rpm_usd, 60_rpm_usd, 30_rpm_td and 60_rpm_td has been applied. 30% of the overall data points (163907 data points) have been randomly chosen to train the clustering algorithm. This step is repeated 10 times to remove the effect of sampling on the determination of the number of clusters. The trained model is later on applied to all the data of each material to build their final clustering.

4.2 Statistical analysis of geometric features

As shown in the relative frequency distribution of V (cf. Figure 5), the 75% quartile and mean values indicate that most of the data lie close to the lower side of the distribution. In the case of *B* and *E*, they are more uniformly distributed. Since these features have a different order of magnitude, clustering on the raw data of *V* might bring biased clusters. For this reason, we applied standardization calculation to each feature *f* as follows:

$$f_{standard} = \frac{f(x) - f_{mean}}{\sigma} \quad (8)$$

which scales the data to unit variance and 0 mean. f_{mean} and σ are the mean and standard deviation of the distribution of the observations of the feature *f*. Other standardization techniques exist, such as rescaling the data to [0, 1], scaling the data according to the quantile range, or transforming data from various distributions to a normal distribution. These techniques were discarded as it was found that they do not provide relevant results. The normalized distribution for each feature is presented in Appendix Figure D.

4.3 Number of clusters

The number of clusters is chosen from the calculations of the metrics *DBI* and *CAH*. As said previously, they have been calculated from 10 random samplings of data. These samplings have a very small influence on the found metric values. Their average is given in Table II for *DBI*, and reported in Figure 7 for *CAH*. Note that a third criteria, Silhouette coefficient (S_{avg}) has also been used to determine the optimal number of clusters[45] . However, this metric does not seems relevant for the choice of the cluster number (Appendix E). For a cluster number of 9, *DBI* and CAH metrics show a local minimum and maximum respectively, suggesting that 9 is the optimal number of clusters.

5 RESULTS

5.1 Classification of the agglomerates

Morphology features of the centroids (V_C, E_C, B_C) of the 9 classes are presented in Table III. They can also be visualised in (V, B) or (E, B) plots (cf. Figure 8).

The agglomerates in classes #1, #2, #3 are the smallest and most similar in size (Appendix Figure Fa and Figure 8a). They differ from each other by their E_C values and E distributions (Figure 9a). These distributions are well separated from each other with small overlap. The class #3 agglomerates have the lowest eccentricity (most spherical) and structure (lowest bulk factor), whereas class #2 has the highest eccentricity ($E_C = 0.804$). The representative B_C and B distributions (Appendix Figure Fb and Figure 8b) for the 3 first classes suggest that they contain low structured agglomerates (low B_C values). Class #1 and #3 have similar B distribution, whereas B values for class #2 is higher (meaning its CB_{aggl} are more structured).

The agglomerates belonging to classes #2, #4 and #6 have an intermediate size (even though the class #6 agglomerates are significantly bigger than those from classes #2 and #4) (Appendix Figure Ga). Their eccentricity is the largest suggesting that they are the most elongated in one or both directions (cf Table III, Appendix Figure Gb). The main difference between these three classes lies in their structure factor *B*, as shown in Figure 9b, the class #6 agglomerates are the most structured, followed by the class #4 agglomerates.

The agglomerates belonging to the #7, #8 and #9 classes are the biggest of all agglomerates. The differentiating parameter between these classes is the size of the agglomerates: the higher the class, the bigger the agglomerates. Their eccentricity distribution is very similar. Moreover, the bigger the agglomerates, the more structured they are, as shown on Figure 8a. Finally, the class #5 agglomerates have an intermediate size and a *B* distribution close to that of class #4 agglomerates (cf. Figure 9b), and an *E* distribution close to that of agglomerates in class #1 (cf. Appendix Figure Gc).

5.2 Application of the methodology to the model materials

Model materials presented in Materials and Testing section have been processed differently. Two rotor speeds (30 rpm vs 60 rpm) along with 2 ways of introducing the fillers and rubber into the internal mixer (Upside-down and Top-down approach) have been tested. The analysis has been performed on a volume which actually could have been eight times lower than the used volume, as shown by the spatial clustering analysis described in Appendix A. Note also that the same experiment and volume fraction analysis have been made on a second sample, and gave the same results (∓0.3%). Cumulated volume fraction ϕ_{cumul} of CB_{aggl} in the model mixes is clearly correlated to the rotor speed (30_rpm_usd: 10.4%, 30_rpm_td: 11.6%, 60_rpm_usd: 6.0%, 60_rpm_td: 5.6%). Note that the cumulated volume fraction of the agglomerates of size below 125 voxels, which have not been taken into account in the morphological analysis, is around 0.4% for all the materials. From ϕ_{cumul} , we can estimate a dispersion index[46] (ϕ_{cumul} divided by the CB volume fraction, without any factor accounting for the possible elastomer content in the agglomerates) equal to 0.54, 0.6, 0.31, 0.29 for 30_rpm_usd, 30_rpm_td, 60_rpm_usd and 60_rpm_td respectively. Deeper analysis of the agglomerates morphology and spatial distribution is required to understand the differences induced by the USD and TD approaches. Actually, as described in Appendix A, the consequences of changing from TD to USD protocol on the CB_{aggl} spatial distribution is small. An analysis of the morphology classes of the agglomerates appears to be much more insightful and is presented in the following.

5.2.1 Differentiation of model mixes using CB_{aggl} morphology clustering

CB_{aggl} volume fraction in each morphological class for the 4 model mixes is presented in Figure 10. For a given filler addition protocol, in most of the classes, the agglomerate volume fraction decreases when the rotor speed is increased, especially for classes with the biggest agglomerates and/or the most structured ones. Interestingly, the contrary is found for classes #1 and #3, which correspond to relatively small and low structured agglomerates (that have a larger eccentricity than in class #2).

As mentioned earlier, ϕ_{cumul} of CB_{aggl} in the model mixes is not significantly changed when the filler addition protocol changes. Compared to TD protocol, Upside Down protocol on model mixes processed at 30_rpm leads to larger volume fractions of agglomerates in the classes #4 and #6, which correspond to the agglomerates with the largest eccentricity. In the other morphological classes, the agglomerates volume fraction is lower when the USD protocol was applied. The impact of changing from USD to TD protocol is different for the materials processed at 60_rpm. With this rotor speed, the USD protocol leads to a higher volume fraction of CB_{aggl} in the classes #4 and above, corresponding to agglomerates with V and B distribution which spread over a domain of values larger than in classes #1, #2 and #3.

Knowing the distribution of each feature in each morphological class, the differentiation between the model mixes can be analyzed even more accurately. One can use for instance, the cumulative frequency distribution of the features *B* and *E*, shown in Figure 11. Increasing the rotor speed produces agglomerates with lower structure (with a shift in the cumulative probability distribution of B towards the lowest values). This effect is prominent for the morphological classes #7 and #8 (Figure 11a, b) (as seen in Appendix Figure H, this effect is negligible for the other classes), which were the classes for which the differentiating parameter was the size. Similarly, there is a decrease in *E* of the agglomerates in class #8 (which corresponds to the biggest agglomerates) due to an increase in the rotor speed, as shown in Figure 11c.

6 DISCUSSION

As shown above, the methodology we developed, provides insightful information (and because of the use of only 3 features, they were easy to visualize, in Figure 8) to evaluate the influence of the processing parameters on the morphology of the CB_{aggl}. By doubling the rotor speed, the total shear deformation undergone by the compound in the internal mixer was also doubled. This promotes more rupture and erosion of agglomerates[47,48]. As expected, this leads to a lower cumulative volume fraction of agglomerates with B parameter above 0.2 (10.4% vs 6.0%), corresponding to the classes #4 and above, i.e. this decreases the volume of the large and more structured agglomerates. Moreover, the increase in the volume fraction of the less structured small agglomerates (classes #1 and #3) with the smallest eccentricity (compared to class #2) suggests that the increase in the rotor speed promotes the creation of more spherical agglomerates from the erosion of the big ones. This is confirmed by the evolution of the *E* and *B* distributions within the classes #7 and #8.

In the comparison of TD vs USD protocol, interesting differences can be found through the clustering methodology. The consequences of the carbon black addition protocol seems to depend on the rotor speed. At 30 rpm, when changing from TD to USD protocol, one observes in particular that the volume fraction of the big agglomerates (classes #7, #8 and #9) decreases. These agglomerates seem to be broken down to lower size agglomerates with large eccentricity corresponding to an increase in the CB volume fraction in the classes #4 to #6. At 60 rpm, the trend seems different with a slight increase in the volume fraction of the biggest agglomerates (classes #7 and #8). In general, TD protocol leads to a more rapid incorporation of CB in comparison to USD approach, as shown by the higher fraction of remaining CB, not incorporated after the IM step (Appendix Table I), i.e. prior to the roll mill step. This is actually in contradiction with the previous observations for the model materials mixed at 30 rpm. This means that the final CB agglomerates dispersion and morphology are the result of a very complex interplay between the CB dispersion at the end of the IM step with the following roll mill step. Only supplementary studies, including a precise CB dispersion characterization as a function of time during the different mixing steps may help to understand our model material microstructure. For this, the use of the methodology developed in this paper is required.

7 CONCLUSION

Very few studies have focused on the morphologies of carbon black agglomerates in industrial rubber compounds. This is mainly due to the limitations in the characterization techniques available at laboratory scale. Here, we demonstrated that even when using laboratory sourced X-ray Tomography, a good attenuation contrast between agglomerates and matrix is attainable provided that a LaB_6 X ray source operated at 40 kV is used. The attenuation contrast between phases can also be increased with the help of various image processing algorithms, thereby the extraction of various features becomes efficient. Like with carbon black aggregates, various morphological features can be used to characterize CB_{aggl}. Out of the 7 features extracted from image analysis, only 3 were retained using feature selection technique (Spearman's rank correlation coefficient). The relevant features appeared to be the volume, the bulk factor and the eccentricity. Morphological classification based on these features was performed using K-means unsupervised clustering algorithm.

Thus, we developed a complete methodology to obtain a morphological classification on CB_{aggl}. This provides many insights on the peculiarity of the morphology of agglomerates in the studied model materials, and enables to distinguish samples with the same measured agglomerates volume fraction. The main advantage of our approach is that it enables to understand the impact of small changes in the materials processing protocol. Notably in the current study, the impact of two processing parameters, (influence of rotor speed and filler addition protocol) can be precisely characterized. This paves the way for a better understanding of the incorporation and desagglomeration mechanisms of carbon blacks during the process. This can also provide useful information to identify in the final materials, the critical morphology of carbon black agglomerates for crack initiation and propagation. More generally, the above presented methodology and the obtained results bring out new questions and prospects for further works. Not only we may study the influence of different processing parameters, e.g. mixing time, milling steps in roll mill etc.., on the carbon blacks dispersion, but - as long as we obtain sufficiently contrasted tomography images - we may also apply the methodology to various type of elastomer filled with various types of fillers (e.g. glass beads, clay, calcium carbonate, recovered carbon black from tire pyrolysis, silica etc..).

8 ACKNOWLEDGEMENTS

This work was supported by LRCCP and ANRT (CIFRE N°2018/0429)

9 REFERENCES

[1]G. R. Cotton. *Rubber Chem. Technol.* **57**, 118 (1984). DOI: 10.5254/1.3535988.

[2]M. Astruc, Thesis, *l'École des Mines de Paris*, (2008).

[3] E. S. Dizon. *Rubber Chem. Technol.* **49**, 12–27 (1976). DOI: 10.5254/1.3534941.

[4] W. M. Hess, R. A. Swor, E. J. Micek. *Rubber Chem. Technol.* **57**, 959 (1984). DOI: 10.5254/1.3536052.

[5]V. Collin, I. Boudimbou, E. Peuvrel-Disdier. *J. Polym. Sci.* **127**, 2121 (2013). DOI: 10.1002/app.37769.

[6]S. Horiuchi, H. Dohi. *Langmuir.* **22**, 4607 (2006). DOI: 10.1021/la052308f.

[7]B. Huneau, I, Masquelier, Y. Marco, V. Le Saux, S. Noizet, C. Schiel, P. Charrier. *Rubber Chem. Technol.* **89**, 126 (2016). DOI: 10.5254/rct.15.84809.

[8]G. Weng, G. Huang, H. Lei, L. Qu, Y. Nie, J. Wu. *Polym. Degrad. Stab.* **96,** 2221 (2011). DOI: 10.1016/j.polymdegradstab.2011.09.004.

[9]R. Liu, E. Sancaktar. *Int. J. Fatigue.* **111**, 144 (2018**).** DOI: 10.1016/j.ijfatigue.2018.01.033.

[10]E. S. Dizon, A. E. Hicks, V. E. Chirico. *Rubber Chem. Technol.* **47**, 231-249 (1974). DOI: 10.5254/1.3540429.

[11]G. J. Lake, P. B. Lindley. *J. Polym. Sci.* **8***,* 707 (1964). DOI: 10.1002/app.1964.070080212.

[12]F. Abraham. In *Constitutive models for rubber VII.* S. Jerrams, N. Murphy, Eds. CRC Press Inc, London, **pp 331** (2012).

[13]J. -B. Le Cam, B. Huneau, E. Verron. *Int. J. Fatigue* **52***,* 82 (2013). DOI: 10.1016/j.ijfatigue.2013.02.022.

[14]A. N. Gent, P. B. Lindley. *Proc. Math. Phys. Eng.* **249***,* 195 (1959). DOI: 10.1098/rspa.1959.0016.

[15]A. N. Gent, P. B. Lindley, A. G. Thomas. *J. Polym. Sci.* **8***,* 455 (1964). DOI: 10.1002/app.1964.070080129.

[16]G. J. Lake, P. B. Lindley. *J. Polym. Sci.* **9**, 1233 (1965). DOI: 10.1002/app.1965.070090405.

[17]A. N. Gent, B. Park. *J. Mater. Sci.* **19** 1947 (1984). DOI: 10.1007/BF00550265.

[18]T. S. Chow. *J. Mater. Sci.* **15**, 1873 (1980). DOI: 10.1007/BF00550613.

[19]S. Otto, O. Randl, O. Goncalves, B. Cantaloube. *KGK Kautschuk Gummi Kunststoffe.* **58,** 390

(2005).

[20]S. C. Garcea, Y. Wang, P. J. Withers. *Compos. Sci. Technol.* **156**, 305 (2018). DOI: 10.1016/j.compscitech.2017.10.023.

[21]E. Maire, P. J. Withers. *Int. Mater. Rev.* **59**, 1 (2014). DOI: 10.1179/1743280413Y.0000000023.

[22]S. Robin, T. Alshuth. In *Constitutive Models for Rubber VIII*. N. Gil-Negrete, A. Alonso. Eds.; Taylor & Francis, (2013).

[23]M. El Yaagoubi, D. Juhre, J. Meier, N. Kröger, T. Alshuth, U. Giese. *Int. J. Fatigue.* **112**, 341 (2018) DOI: 10.1016/j.ijfatigue.2018.03.024.

[24]O. Gehrmann, M. El Yaagoubi, H. El Maanaoui, J. Meier. *Polym. Test* **75**, 229 (2019) DOI: 10.1016/j.polymertesting.2019.02.025.

[25]J. Kallungal, L. Chazeau, J. -M. Chenal, J. Adrien, E. Maire, C. Barres, B. Cantaloube, P. Heuillet. In *Constitutive Models for Rubber XI*; CRC Press, pp 77 (2019). DOI: 10.1201/9780429324710-14.

[26]A. I. Medalia, F. A. Heckman. *Carbon.* **7**, 567 (1969). DOI: 10.1016/0008-6223(69)90029-3.

[27]A. I. Medalia, F. A. Heckman. *J. Colloid Interface Sci.* **36**, 173 (1971) DOI: 10.1016/0021-9797(71)90162-7.

[28]A.I. Medalia. *J. Colloid Interface Sci.* **32**, 115 (1970). DOI: 10.1016/0021-9797(70)90108-6.

[29]A.I. Medalia. *J. Colloid Interface Sci.* **24**, 393 (1967). DOI: 10.1016/0021-9797(67)90267-6.

[30]X. Bourrat, A. Oberlin, H. Van Damme, C. Gateau, R. Bachelar. *Carbon.* **26**, 100 (1988). DOI: 10.1016/0008-6223(88)90016-4.

[31]T. P. Rieker, M. Hindermann-Bischoff, F. Ehrburger-Dolle. *Langmuir* **16**, 5588 (2000). DOI: 10.1021/la991636a.

[32]G. P. Baeza, A.- C. Genix, C. Degrandcourt, L. Petitjean, J. Gummel, M. Couty. J. Oberdisse. *Macromolecules.* **46**, 317 (2013). DOI: 10.1021/ma302248p.

[33]C. R. Herd, G. C: McDonald, W. M. Hess. *Rubber Chem. Technol.* **65**, 107 (1992). DOI: 10.5254/1.3538594.

[34]E. Gouillart, J. Nunez-Iglesias, S. van der Walt. *Adv. Struct. Chem.* **2**, 18 (2017). DOI: 10.1186/s40679-016-0031-0.

[35]E. Gouillart *Github*. https://github.com/emmanuelle/image-processing-tutorials/blob/master/tomo_segmentation.ipynb.

[36]E. Gouillart, M. J. Toplis, J. Grynberg, M. -H. Chopinet, E. Sondergard, L. Salvo, M. Suéry, M. Di Michiel, G. Varoquaux, *J. Am. Ceram. Soc.* **95**, 1504 (2012). DOI: 10.1111/j.1551-2916.2012.05151.x.

[37]L. Grady. Random walks for image segmentation *IEEE Trans. Pattern Anal. Mach. Intell.* **28,** 11 (2006). DOI: 10.1109/TPAMI.2006.233.

[38]S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu. *PeerJ* **2**, e453 (2014). DOI: 10.7717/peerj.453.

[39]E. P. Denis, C. Barat, D. Jeulin, C. Ducottet. *Mater. Charact.* **59**, 338, (2008). DOI:10.1016/j.matchar.2007.01.012.

[40]M. del M. Fernández-Arjona, J. M. Grondona, P. Granados-Durán, P. Fernández-Llebrez, M. D. López-Ávalos. *Front. Cell. Neurosci.* **11**, 235 (2017). DOI: 10.3389/fncel.2017.00235.

[41]A. Madra, J. Adrien, P. Breitkopf, E. Maire, F. Trochu. *Compos. Part A Appl. Sci.* **102**, 184 (2017). DOI: 10.1016/j.compositesa.2017.07.028.

[42]F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É.Duchesnay. *J Mach Learn Res* **12**, 2825 (2011). DOI: 10.48550/arXiv.1201.0490.

[43]J. Hämäläinen, S. Jauhiainen, T. Kärkkäinen. *Algorithms*. **10**, 105 (2017). DOI: 10.3390/a10030105.

[44]D. L. Davies, D. W. Bouldin. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-1*, **vol PAMI-**

1, no. 2, (1979). DOI: 10.1109/tpami.1979.4766909.

[45]P. J.Rousseeuw. *J. Comput. Appl. Math.* **20**, 53 (1987). DOI: 10.1016/0377-0427(87)90125-7.

[46]T. Villmow, P. Pötschke, S. Pegel, L. Häussler, B. Kretzschmar. *Polymer.* **49**, 3500 (2008). DOI: 10.1016/j.polymer.2008.06.010.

[47]S. P. Rwei, I. Manas-Zloczower, D. L. Feke. *Polym Eng Sci.* **31**, 558 (1991). DOI: 10.1002/pen.760310804.

[48]V. Collin. *Thesis,. École Nationale Supérieure des Mines de Paris*. (2009).

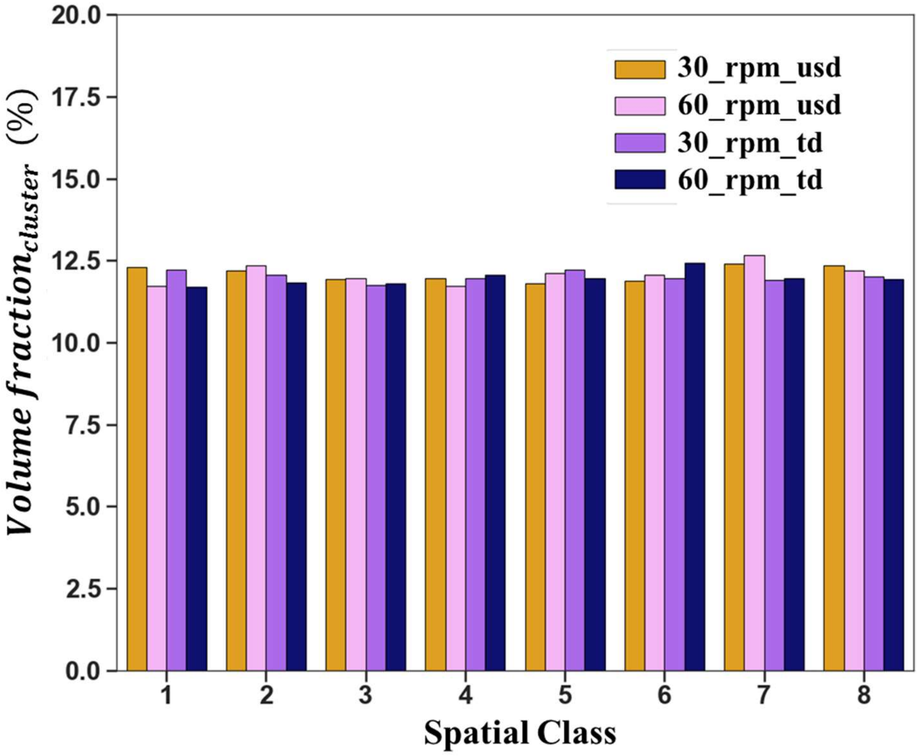
10 APPENDIX

Appendix A: spatial distribution

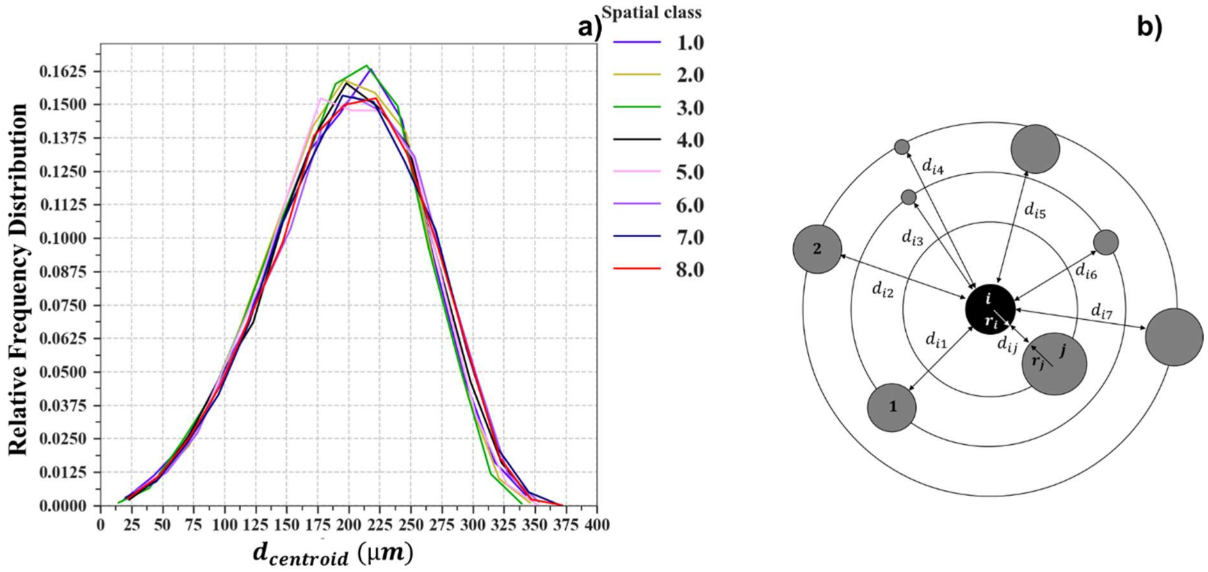
In addition to the agglomerates morphologies, the spatial distribution of agglomerates in a mix is another important information to characterise the quality of the filler dispersion. To understand the aggregation of agglomerates at the micron level, it is interesting to calculate the distances to the nearest neighbours of each CB_{aggl}. In order to do it, a matrix with the distances between each agglomerate and all other agglomerates in the model material has to be evaluated, based on their position. Since the number of agglomerates in each model material is very high, it requires huge computation power and memory storage. To reduce this computation and the memory requirement, spatial clustering was performed on each model material using the same K-means algorithm applied for morphological clustering, except that the data points are the spatial coordinates of the agglomerates. All the agglomerates that are close to each other (centroids of the agglomerates) in the model material are clustered into one spatial class. Here 8 classes have been identified as the optimal number, which translates to around 12% of the volume fraction for each spatial class in the stack (Appendix Figure A1). The distribution of the spatial distance of each agglomerate in each class to the class centroid $D_{centroid}$ ($D_{centroid} = (x_i - \mu_k)^{0.5}$, where x_i is the spatial coordinates of agglomerate center and μ_k is the centroid of the kth spatial class ($1 \leq k \leq 8$)) gives insight on the agglomerates spatial distribution. The distributions found for the 8 spatial classes are very close to each other (Appendix Figure A2a) indicating that each class can be considered as a representative volume of the model mix. As mentioned earlier, we are more interested in the dispersion at lower scale. The clustering (equivalent to a reduction of the sample size) being completed, the distances between the 20 nearest neighbours for each agglomerate in one class can therefore be calculated. 20 has been arbitrarily chosen, knowing that a parallelepiped in which an agglomerate can be bounded, counts 26 neighbors (in contact through the faces, vertices and edges). These distances d are the minimal ones between the surfaces of 2 agglomerates, as depicted in Appendix Figure A2b, and are calculated as follows:

$$d_i = \sqrt[2]{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} - \left(r_i + r_j\right) \quad (1 \leq j \leq 20)(1 \leq i \leq M) \quad (9)$$

Where x, y, z are the spatial coordinates of the agglomerate centre, r its radius calculated from its Eq_{diam}, the index i is for the agglomerate considered, j for one of its 20 closest neighbours and M is the total number of agglomerates in the spatial class considered for the analysis . In Appendix Figure A3, we have chosen the spatial class n°1 for all the model materials.

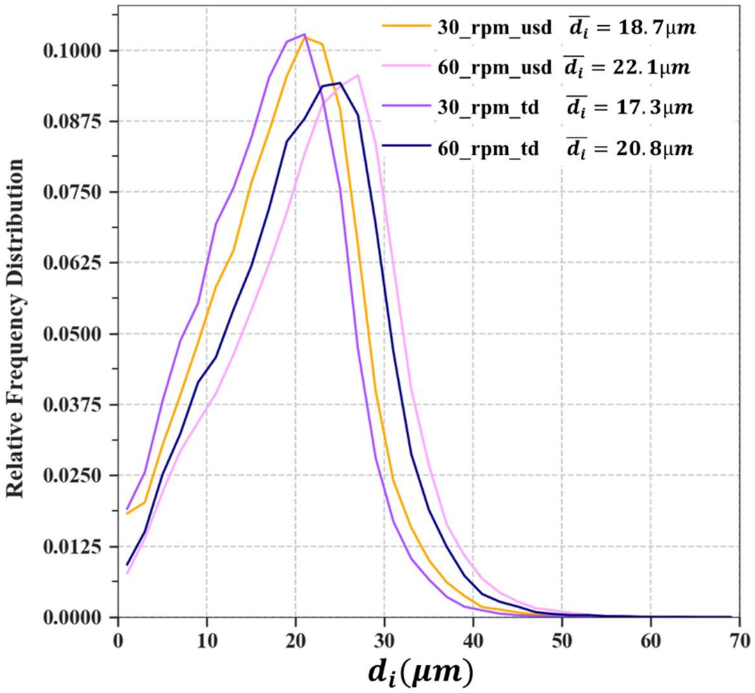


Appendix Figure A1: Volume fraction of each spatial class for 4 model materials.



Appendix Figure A2: a) Distribution of the distances of each agglomerate to its class centroid. b) Illustration of the nearest-neighbour distance

Result: Spatial Distribution of CB_{aggl}



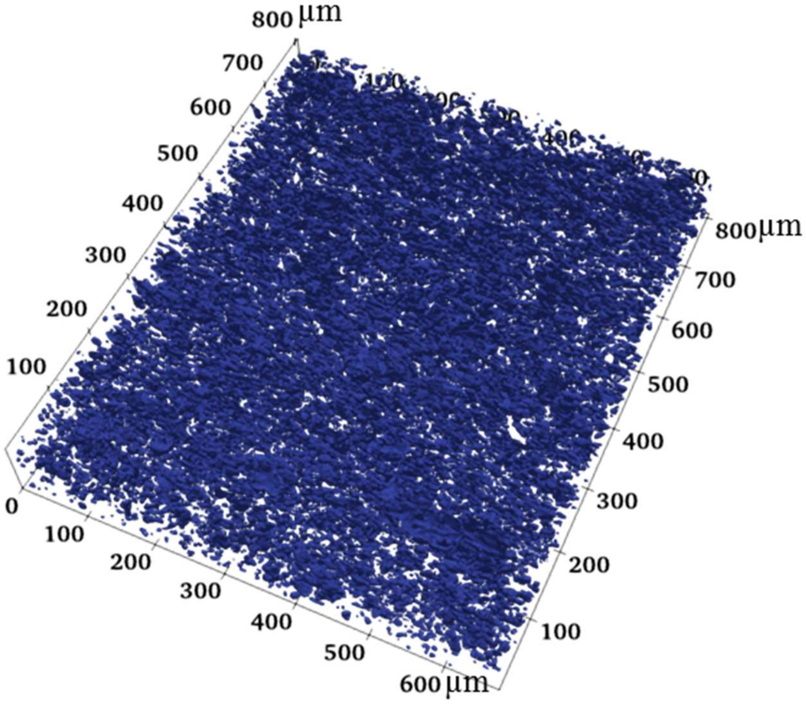
Appendix Figure A3: Distribution of nearest-neighbour distance for various model mixes along with the mean value for each mix. $\overline{d_\iota}$ corresponds to the mean of the distribution

Model mixes	**WMH (µm)**
30_rpm_usd	18.9
60_rpm_usd	20.9
30_rpm_td	18.1
60_rpm_td	20.1

Appendix Table A : Value of the Width at MidHeight of the peak of the nearest neighbor distance distribution for the model mixes

The distributions of the distances to the nearest neighbours are presented in Appendix Figure A3 for all the model mixes. There are slight differences between the different materials. Increasing the rotor speed leads to a shift in the distribution towards larger values. The Width of the peak distribution at its MidHeight (cf Appendix Table A1) is also slightly increased. In addition, we know that the agglomerates volume fraction is reduced. Thus, the increase in the rotor speed improves the CB deagglomeration, reduces the size and volume fraction of the agglomerates, and logically increases the distance between them.

Appendix Figure B: Sub-section volume of the model material 30_rpm_usd



Appendix Figure B: Sub-section volume of the model material 30_rpm_usd obtained using 3D Tomography visualized using ParaView © software

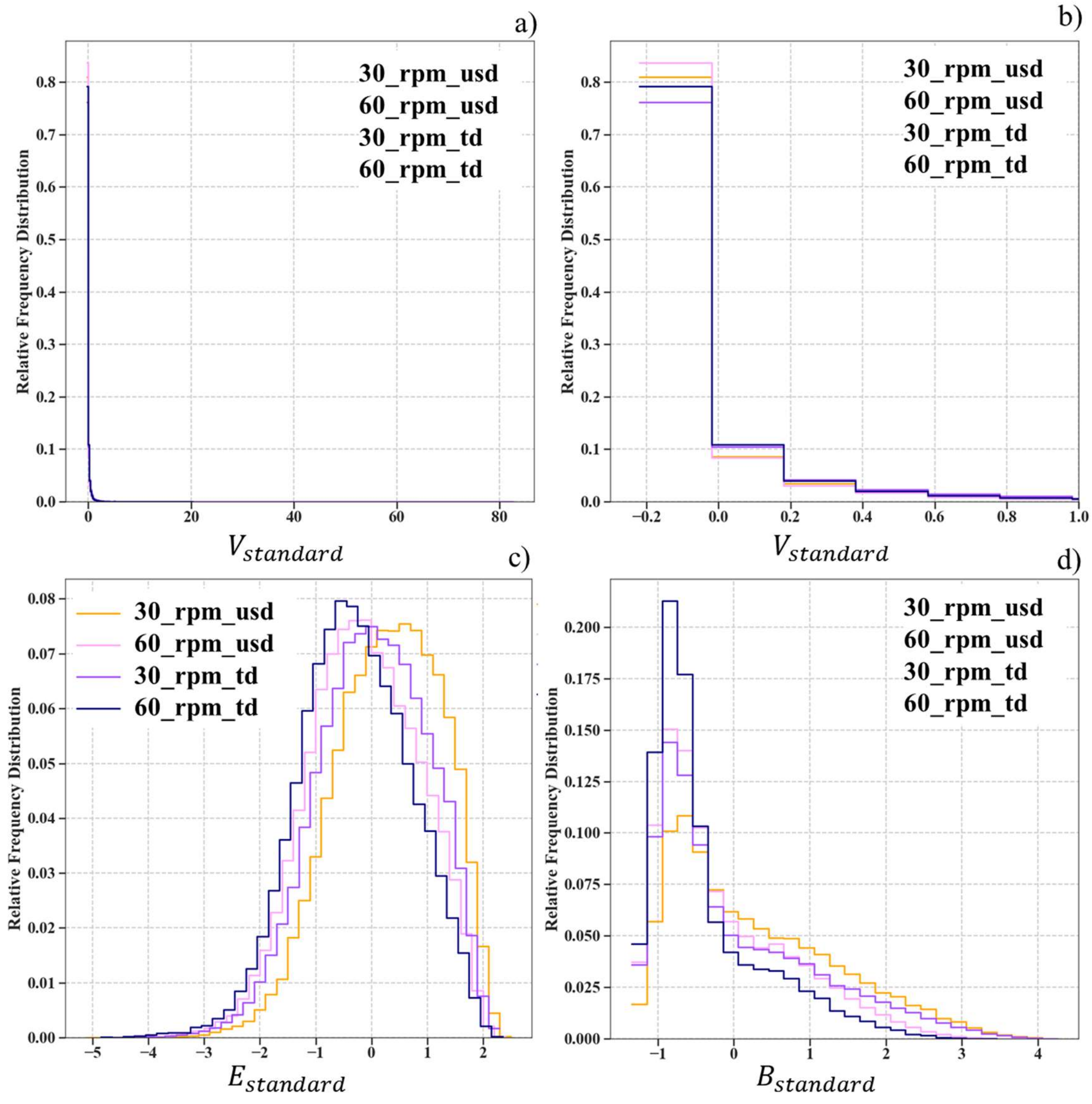
Appendix C: laplacian score

Laplacian Score (LS) works on the principle that data from the same class tend to be close to each other and each feature is evaluated based on its “locality preserving power”. LS ranges from 0 to 1, where score value close to 1 corresponds to a relevant parameter. The algorithm developed by He *et al.* [45] has been used to calculate the Laplacian score for each feature (Appendix Table C). Laplacian score shows that each feature can be used to distinguish different clusters in the data as all the features have a LS above 0.9, except B (0.85).

V	***Eq_{diam}***	***B***	***E***	***b***	***a***	***Q***
1.000	1.000	0.851±0.034	0.920±0.009	0.972±0.004	0.970±0.003	0.925±0.010

Appendix Table C: Laplacian score for each feature

Appendix Figure D



Appendix Figure D: Statistics on each standardized geometric features(using equation 8), $V_{standard}$ ((a), (b)), $E_{standard}$ (c) and $B_{standard}$ (d)- which have been retained as descriptors of the morphology of CB$_{aggl}$ for all the model mixes (30_rpm_usd, 60_rpm_usd, 30_rpm_td, 60_rpm_td) - is presented

Appendix E: silhouette coefficient

The Silhouette coefficient (S)[42] is a validation metric to check the consistency within clusters of data. It calculates the cohesion of an observation to its own cluster compared to other clusters using the following equation.

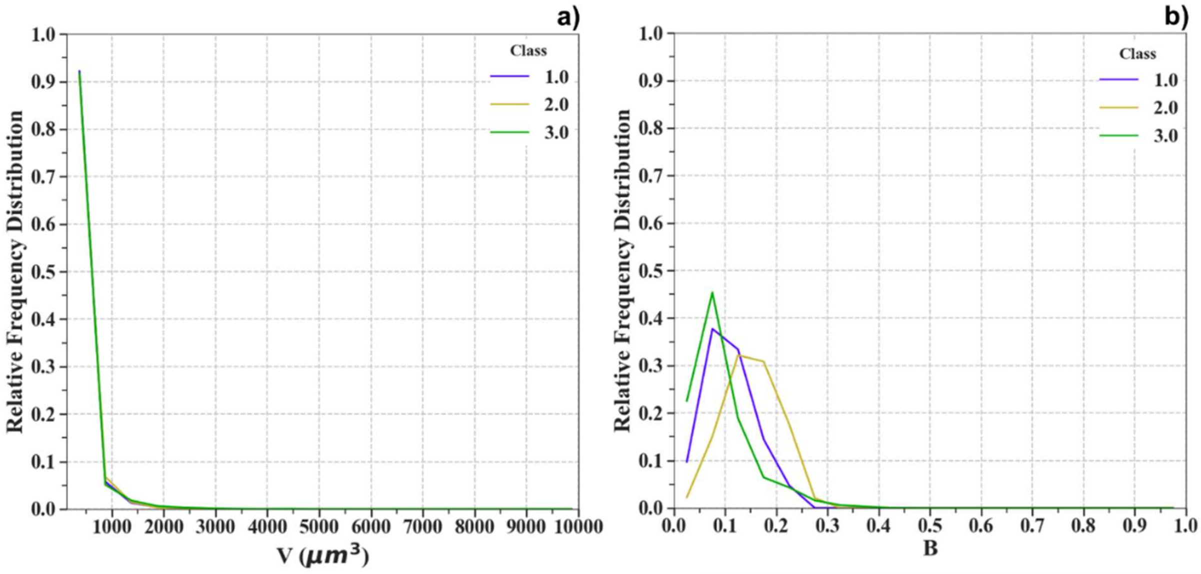
$$S(i) = \frac{b(i)-a(i)}{max\{a(i),b(i)\}} \quad (7)$$

Where *b(i)* is the mean distance of i^{th} data point to all other data points in the closest neighboring cluster to which *i* is not linked, i.e. between an agglomerate and agglomerates from other clusters (mean nearest-cluster distance) and *a(i)* is the average distance of i^{th} observation to all other observations in the same cluster (mean intra-cluster distance). *S(i)* varies between -1 and 1, where 1 corresponds to a good partition of a cluster, score near 0 indicates overlapping of clusters space and negative value signifies that observations have been partitioned wrongly. We will use S_{avg} which is the average of all the *S(i)* values (Appendix Table E).

k number of clusters	S_{avg}
2	0.448±0
3	0.450±0
4	0.371±0
5	0.365±0
6	0.347±0
7	0.358±0
8	0.351±0
9	0.344±0
10	0.327±0
11	0.322±0
12	0.326±0
13	0.323±0
14	0.315±0

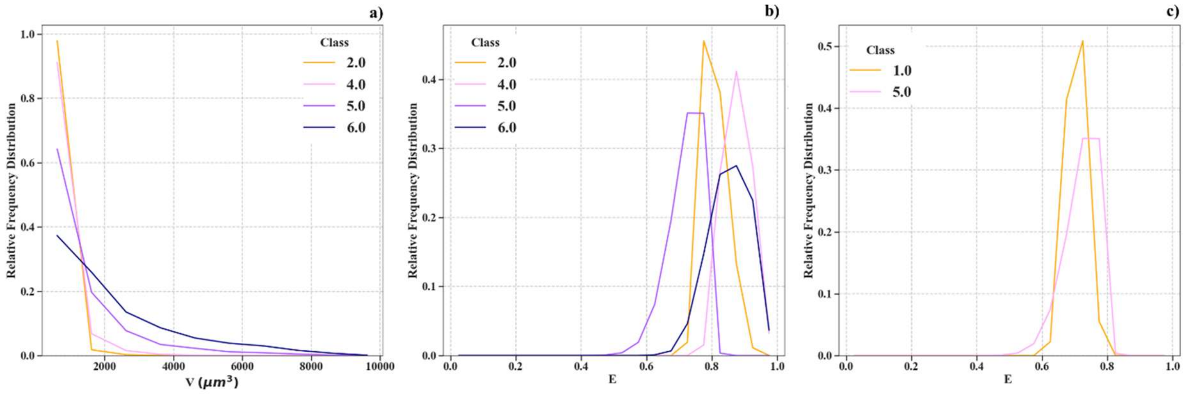
Appendix Table E: S_{avg} metric evaluation for each number of clusters, deduced from 10 random samplings

Appendix Figure F



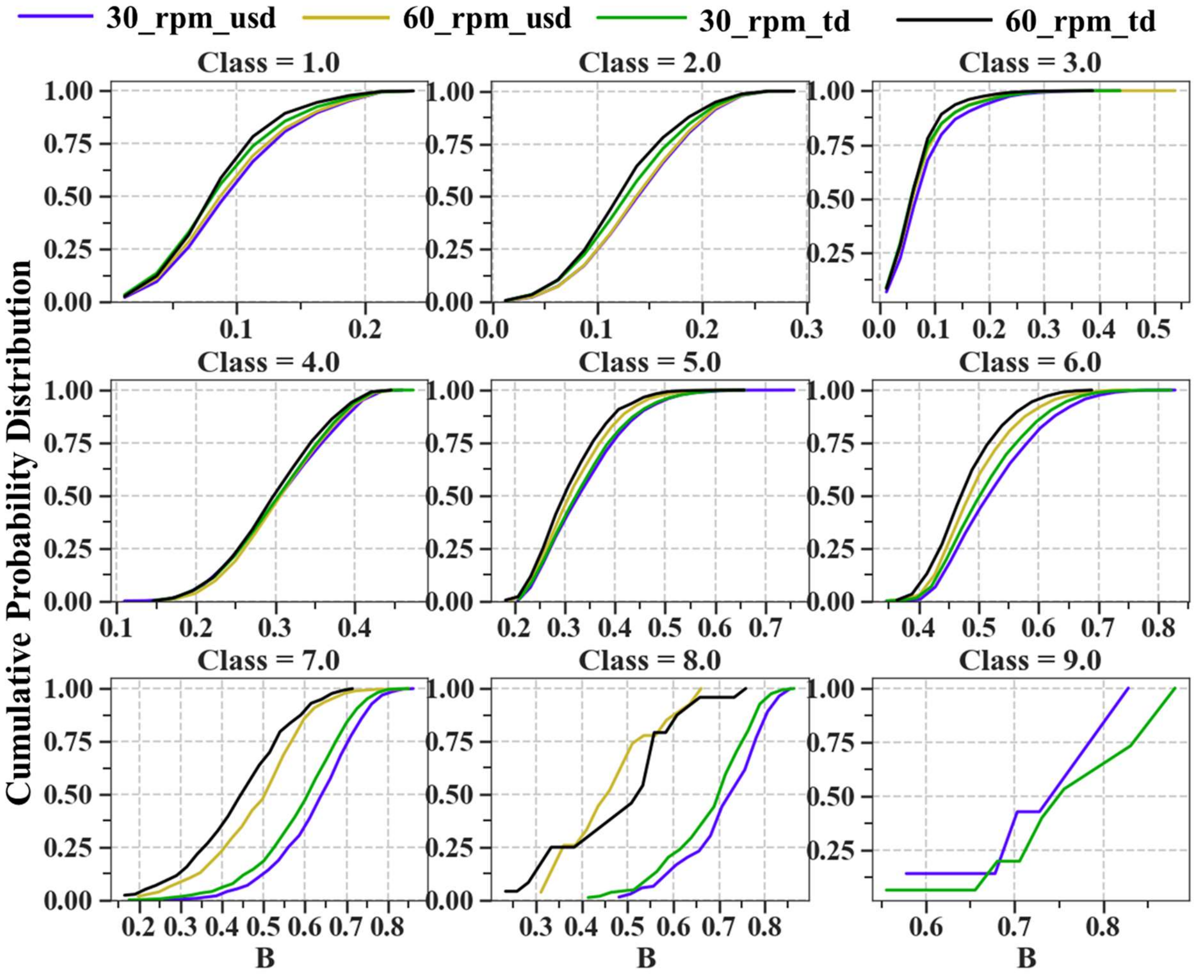
Appendix Figure F: a) and b) Show the relative frequency distribution plot of V and B for model material 30_rpm_usd corresponding to morphological classes #1, #2 and #3.

Appendix Figure G



Appendix Figure G: a) and b) Show the relative frequency distribution plot of V and E respectively corresponding to morphological classes #2, #4, #5 and #6, c) shows the relative frequency distribution plot of E corresponding to morphological classes #1 and #5 . The agglomerates considered for plotting is from model material 30_rpm_usd

Appendix Figure H



Appendix Figure H: Cumulative probability distribution of B within each morphological class corresponding to each model material.

Appendix Table I

Protocol	**Time (min)**	**Total rotation (rad)**	**Cohesive (gm)**	**Loose Carbon Black (gm)**	**Total mass (gm)**	**% of Carbon Black loose**
USD @ 60 rpm	2	753.6	53.1	1.8	54.9	3.3
TD @ 60 rpm	2	753.6	54.4	0.2	54.6	0.4
USD @ 30 rpm	2	376.8	52	2.392	54.392	4.4
TD @ 30 rpm	2	376.8	51	1.761	52.76	3.3

Appendix Table I: Amount of Carbon Black incorporated depending on the filler addition protocol and rotor speed of the internal mixer. (This information was obtained for a mix fabricated using the filler addition protocol suggested in Materials and Testing). This table demonstrates the differences in the CB incorporation in the mix depending on the time spent in the internal mixer.