Brazilian Journal of Biometrics (2022), **40**, 428–441
doi:10.28951/bjb.v40i4.614

ARTICLE

Statistical characterization of spatial and size distributions of particles in composite materials used in the manufacturing of biomedical instruments [1]

João Domingos Scalon *[1] , Victor Ferreira da Silva [2], Wélson Antônio de Oliveira [2] and Mateus Santos Peixoto [2]

[1]Department of Statistics, Institute of Exact and Technological Sciences, Federal University of Lavras, Lavras, Brazil
[2]Postgraduate Program in Statistics and Agricultural Experimentation, Institute of Exact and Technological Sciences, Federal University of Lavras, Lavras, Brazil
*Corresponding author. Email: scalon@ufla.br

(Received: June 1,2022; Revised: December 14,2022; Accepted: December 14,2022; Published: December 30,2022)

Abstract

Aluminium is extensively used in many manufacturing processes because of its intrinsic properties like soft, ductile, high electrical conductivity and highly corrosion resistant. Unfortunately, pure aluminium cannot give a required tensile strength, whereas by adding some other materials like particles of silicon carbide can give a proper strength and converted into a composite with adequate properties which is most suitable in the manufacturing of some specific biomedical instruments. It is well known that size and spatial distributions of particles are both influential in determining the mechanical properties of composite materials and, therefore, statistical characterization of these distributions is of prime importance if we wish to control the quality of the manufacturing processes for these materials. Many researchers have considered quantitative analysis of these features separately, but here we investigate the relationship between size and spatial distribution of the particles over an aluminium matrix. We have considered the actual sizes simply as 'large' or 'small' and, consequently, the characterization of the particle distribution patterns in the aluminium matrix can be carried out using statistical methods based on the theory of bivariate spatial point processes. We have applied this statistical approach to a sample of an aluminium alloy reinforced with silicon carbide particles. It is shown that the methods provide a complete characterization on the spatial interaction between small and large silicon carbide particles and it can be successfully used in a quality control step for the production of particulate composite materials used in the manufacturing of biomedical instruments.

Keywords: Al-SiC; Marked point processes; Monte Carlo simulation; Gibbs model.

1. Introduction

Aluminium alloys reinforced with silicon carbide particles (Al-SiC) constitute an aluminum-based material with improved mechanical properties compared to traditional aluminium alloys, such as high strength, modulus, wear and fatigue resistances (Chawla et al., 2001; Davison, 1993; Torralba et al., 2003). Due to these properties of the composite material, it has been one of the important applications in many industrial sectors, including biomedical instruments fabrication (Dwivedi et al., 2021; Salernitano et al., 2003). However, it is well known that there

is a close relationship between these mechanical properties of the composite and the spatial distribution of the SiC particles over the aluminium matrix (Hong et al., 2004). For example, there is evidence that cracking is more prevalent in composites with strongly clustered particles (Lewandowski et al., 1989) and ductility is inversely related to degree of clustering (Murphy et al., 1998). Thus, materials researchers need methods for the analysis of the spatial distribution of second-phase particles in composite materials.

If the centers of particles are used as the objects of study, suitable statistical tools that consider the spatial distribution of particles can be based on the theory of spatial point process. These methods have been used in the analysis of spatial distributions of second-phase particles in many papers and their efficacy has been proved (Cetin et al., 2009; Ghosh et al., 1997; Pyrz, 1994; Scalon et al., 2003; Stoyan et al., 1990).

The sizes of the particulate phase are, however, also important. For example, increasing the second-phase particle size leads to worsening the mechanical properties due to lower work hardening and higher damage accumulations rates (Chawla et al., 2001; Davison, 1993; Fathy et al., 2014; Narayanasamy et al., 2009; Willians, 2002). Statistical methods have been also applied in the analysis of particle size distributions, with most researchers relying on descriptions of volume fraction, calculation of sample moments of particle sizes and visual inspection of histograms of the particle size distribution (Chawla et al., 2001; Spitzig et al., 1985). Scalon et al. (2003a) proposed and advocated the use of a more coherent statistical approach that involves the employment of a distribution model for particle sizes in composite materials such as the log-skew-Laplace.

In this paper, we consider a question of key interest in particulate composite materials: Is the size of the reinforcing particles related to their location? The problem is to find specific methods to provide a quantitative measure of the interaction between the size of the particles and their spatial distribution.

Here we consider that the spatial distribution of the second-phases in particulate composite materials is determined by both the particle positions and their sizes and, therefore, a quantitative analysis of the interaction between size and spatial distributions of particles can be founded on marked point process theory. We advocate, initially, categorizing the sizes simply as ‘large’ or ‘small’, so we can use the simpler bivariate point processes for describing and modelling the size-spatial arrangement of reinforcing particles.

In spite of the relatively long history of marked point process theory as a useful tool in many areas such as ecology, forestry, epidemiology and cytology (Diggle, 2003; Baddeley et al., 2016), its application to materials has been quite limited. Most of the studies in materials science rely on simply describing the interaction between the size and location of the second-phase particles through either mark correlation or variogram functions (Ghosh et al., 1987; Stoyan et al., 1990; Pegel, 2009; Stoyan et al., 1991), though model-based methods are also applied (Ghosh et al., 1987; Stoyan et al., 1991; Stoyan et al., 1985; Wincek et al., 1993; Derr et al., 2000; Stoyan, 2002). To the best of our knowledge, methods based on the theory of bivariate point process have never been used to address the quantitative analysis of reinforcing particles in biocomposite materials.

The present paper is an attempt to introduce the use of the theory of bivariate point processes as a simpler and yet effective tool for characterizing the size and spatial arrangement of second-phase particles. In putting together various methods into a coherent strategy, we intend to fill gaps, extend ideas and contribute to the knowledge of the relationship between spatial and size distributions of particles in biocomposite materials.

2. Data Description

The aluminium alloy reinforced with SiC particles was synthesized by spray-forming in the Department of Engineering Materials, University of Sheffield, Sheffield, England. The SiC particles represented a volume fraction of 11%. They were in contact with the semi-molten alloy metal just for a brief time and, therefore, chemical reactions were minimal.

The specimen of the Al-SiC composite material was split into smaller pieces (samples) using a precision cut-off machine. The cutting process resulted in metallographic samples of approximately 192 μm × 288 μm. The samples were placed on a computer-controlled optical microscope stage analyser (Polyvar) which allowed fully automatic adjustment, focusing, positioning and scanning of each sample. The two-dimensional digital images were then analyzed using image processing techniques to extract the size (radius of the equivalent disc) and location (co-ordinates of the equivalent disc center) of each SiC particle within the image.

We choose, randomly, only one sample to illustrate the statistical analysis. The sample mean of the recorded particle sizes was used as a cut-off point. Thus, if the particle size was less than or equal to the sample median (1.37 μm), the particle was classified as being 'small'. Otherwise, the particle was classified as being 'large'. Figure 1 shows the transformation of the microscope image of the polished surface of the Al-SiC composite material sample to the resulting data that we shall analyze; that is, a bivariate point pattern.

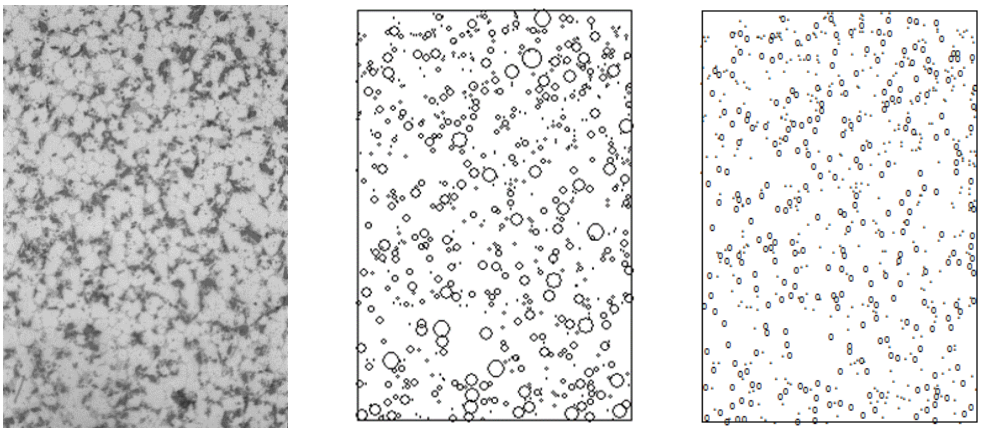

Figure 1. Microscope image of the metallographic sample with area 192 × 288 μm (left), locations of the 730 SiC particles and their sizes, represented as radii of the equivalent disc (middle) and, the bivariate spatial point pattern with 435 small particles (points) and 295 large particles (circles) (right).

We use the software R (R core team, 2022) and its library "spatstat" (Baddeley et al., 2005) to perform all analyses conducted in this work. Data and R codes used to perform statistical analyses of this paper are available upon request from the first author.

3. Describing trend and interaction between small and large particles

Exploratory statistical analysis allows a first look at the characteristics of our data. In our sample, we recorded 730 SiC particles, representing an estimated volume fraction of 11%. This value is the same as the actual volume fraction used in the process production of the material, supporting the hypothesis that the sample preparation and selection have been carried out without bias. We have computed the sample mean (me), median (md) and the standard deviation (sd) of the radius of the equivalent disc (sizes) for both types of particles and obtained the following results. Small: me = 0.73 μm, md = 0.70 μm, sd = 0.19 μm and large: me = 1.99 μm, md = 1.74 μm, sd = 0.79 μm.

The positive difference between the mean and the median of the particle size data, for both types of particles, indicates that the size distribution is skewed to the right (positive). Other researchers have shown clear evidence that the particle size distribution in Al-SiC composite material deviate from the conventional symmetric (normal) to the positive skew distribution (Scalon *et al*., 2003a). We have observed that large particles have sizes with variability greater than small particles.

The nature of a spatial point pattern may be thought of as comprising two components: intensity and dependence (or interaction) between the points of the patterns. Thus, we have to describe the stochastic structure of the bivariate point process in terms of its first and second-order properties.

The first order property of a spatial point process describes the local intensity, or the mean number of particles per unit area, as a function of position. Local intensity may be modelled, parametrically or non-parametrically, using coordinates and/or useful covariates. If such covariates are not available, as in our case, it is possible to adopt the procedure of estimating intensity by smoothing of the available data (Diggle, 2003; Baddeley, 2016; Baddeley et al., 2006).

Diggle (2003) presents a non-parametric approach for obtaining a spatially smooth intensity of a spatial point pattern based on a quartic kernel estimator. Let $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be the spatial locations of *n* particles in a bounded study region |A|. The local intensity $\lambda_\tau(\boldsymbol{u})$ at any location $\boldsymbol{u}$ is estimated by

$$\hat{\lambda}_\tau(\boldsymbol{u}) = \sum_{i=1}^{n} \frac{3}{\pi\tau^2}\left(1 - \frac{h_i^2}{\tau^2}\right)^2, \qquad (1)$$

where h_i is the Euclidean distance between an arbitrary point $\boldsymbol{u}$ and the observed particle at location $\boldsymbol{x}_i$ and $\tau > 0$ is the bandwidth that determines the amount of smoothing. We have adopted a bandwidth that minimizes the mean-square error of a cross-validation process (Diggle, 2003; Baddeley et al., 2005).

If the pattern is a realization of a homogeneous Poisson process, it is expected that the intensity estimate should be constant over matrix and approximately equal to the "average intensity" estimated by number of points/area.

Baddeley et al. (2006) suggest using the ratio of the two univariate intensity estimates for verifying whether the two point processes present the same first order properties. A constant local ratio suggests that both small and large particles may share the same degree of homogeneity (or inhomogeneity) and thus, the image plot of this intensity ratio provides a first exploratory approach for checking the interaction between the two patterns (Baddeley et al., 2006). Although, Cetin and Kalkanli (2009) and Scalon et al. (2003) have successfully utilized the kernel smoothing estimator for characterizing first order properties of univariate point patterns of second-phase particles in composite materials, they have not used the ratio estimator in such analyses.

Figure 2 shows (right and center panels) that although the local intensity appears to be approximately constant over the matrix and equal to the average intensity 0.0066, it is not obvious that the hypothesis of stationarity is realistic for the spatial distribution of both types of particles. We can identify some high and reduced local concentrations. Thus, these images may suggest slightly non-stationarity in the spatial distribution of both small and large particles. The constancy of the ratio small-to-large over the sample area suggests that the same slight inhomogeneity may affect both types of particles.

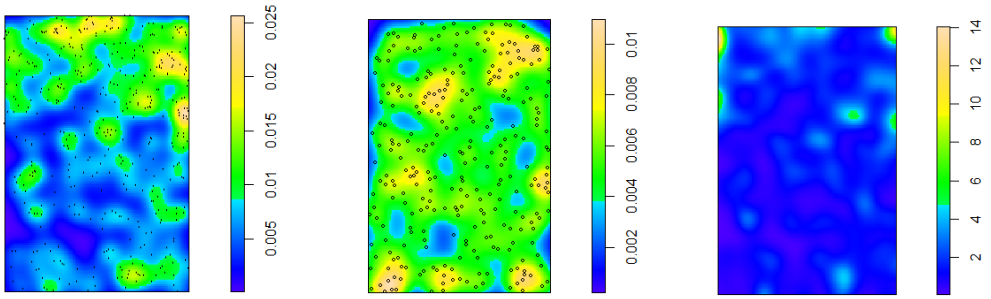

Figure 2. Image plots of the intensity estimates for small (left), large (center) and ratio of small-to-large (right) particles.

Once we have characterized the intensity, the next step might be to assess dependence (or interaction) between the points of the patterns in terms of their second-order properties. The simplest manifestation of such interaction consists of either attraction (aggregation or clustering) or repulsion (regularity) in the pattern. A standard function for exploratory analysis of the second-order properties of a stationary spatial point process is given by the *K*-function, defined by Ripley (1976) as $K(r) = \lambda^{-1}E$[number of events within distance r of an arbitrary event], where λ is the intensity of the process, and *E*[.] denotes expectation. In practice, we do not have knowledge of the actual *K*-function of the process and, therefore, we must estimate it from the spatial point pattern under consideration.

Numerous *K*-function estimators have been proposed for stationary spatial point patterns. Most of the usual estimators are weighted and renormalized empirical distribution functions of the pairwise distances among events (Diggle, 2003, Ripley, 1976). Unfortunately, many researchers ignore the fact that when the pattern is non-stationary and, the use of the usual estimators of the *K*-function is inappropriate. Since our previous analysis suggested slightly nonstationarity in the spatial distribution of particles, we should use an estimator that goes some way to alleviating this problem.

Baddeley et al. (2000) proposed an estimator of the *K*-function for nonstationary point processes (inhomogeneous *K*-function), where the second-order intensity of two points is divided by their respective local intensity is stationary. If the process is actually stationary, then the local intensity is constant and the inhomogeneous *K*-function reduces to the usual (homogeneous) *K*-function.

Møller et al. (2003) have extended the estimator proposed by Baddeley et al. (2000) for allowing estimating second-order intensity for non-stationary bivariate spatial point processes. Let n_S and n_L denote the number of particles in types small and large, respectively, in a bounded study region |A|, and set $\boldsymbol{x}_S = \{\boldsymbol{x_{S1}}, \ldots, \boldsymbol{x_{Sn_1}}\}$ and $\boldsymbol{x}_L = \{\boldsymbol{x_{L1}}, \ldots, \boldsymbol{x_{Ln_2}}\}$, where $\boldsymbol{x}_{ij}$ denotes the locations (coordinates) of the *j*th particle for types *i*. Also, let $\|\boldsymbol{x_{Si}} - \boldsymbol{x_{Lj}}\|$ be the Euclidean distance between particles $\boldsymbol{x_{Si}}$ and $\boldsymbol{x_{Lj}}$. Then, a suitable edge-corrected unbiased estimator for the inhomogeneous bivariate *K*-function for a particular distance *r* is given by

$$\widetilde{K}_{SL}(r) = \frac{|A|}{n_S n_L} \sum_{i=1}^{n_S} \sum_{j=1}^{n_L} \frac{I(\|\boldsymbol{x_{Si}} - \boldsymbol{x_{Lj}}\| \leq r)}{\hat{\lambda}(\boldsymbol{S_i})\hat{\lambda}(\boldsymbol{L_j}) w_{ij}}, \tag{2}$$

where the $\hat{\lambda}(.)$ terms are estimates of the local intensity for the patterns of small and large particles given by equation (1), w_{ij} is a weighting factor, proposed by Ripley (1976), that represents the proportion of the circumference of the circle centered at the *i*th small particle, passing through the *j*th large particle that lies within |A| and *I*(.) is an indicator function equal to 1 if $\|\boldsymbol{x_{Si}} - \boldsymbol{x_{Lj}}\| \leq r$, and 0 otherwise.

An analogous estimator for $K_{LS}(r)$ is defined by interchanging the roles played by the two types of particles. Theoretically, $K_{SL}(r) = K_{LS}(r)$ but this will not necessarily be the case for their corresponding estimates and, therefore, it is prudent to combine the estimators $\widetilde{K}_{SL}(r)$ and $\widetilde{K}_{LS}(r)$ into a single estimator. Lotwick et al. (1982) show that the most efficient estimator of the

bivariate *K*-function is the linear combination of $\widetilde{K}_{SL}(r)$ and $\widetilde{K}_{LS}(r)$. Following the same idea, Møller and Waagepetersen (2003) suggest an efficient estimator of the inhomogeneous bivariate *K*-function given by

$$\widehat{K}(r) = (n_S + n_L)^{-1}\{n_L\widetilde{K}_{SL}(r) + n_S\widetilde{K}_{LS}(r)\}. \quad (3)$$

The basic idea in interpreting the bivariate *K*-function is that under the hypothesis of independence between the two types of particles, the locations of large particles should be random with respect to those of small particles, regardless of whether the spatial distribution of either small or large particles is clustered, regular or random when considered separately. Thus, if the hypothesis of independence is true, it can be shown that $K_{SS}(r) = K_{LL}(r) = K_{SL}(r) = K_{LS}(r) = \pi r^2$ (Baddeley et al., 2000; Møller et al., 2003; Lotwick et al., 1982).

A well-recommended transformation of the *K*-function is the *L*-function $\hat{L}(r) = \sqrt{\frac{\widehat{K}(r)}{\pi}} - r$, which transform the bivariate *K*-function to the straight horizontal line at vertical height zero. This transformation seems to stabilize the variance of the estimator, making graphical visual assessment of interactions easier (Baddeley et al., 2000; Møller et al., 2003). Values of $\hat{L}(r)$ higher than zero are characteristic of positive interaction (attraction) at the distance *r*, whilst values smaller than zero are found when there are repulsion (regularity) between particles. In addition, pointwise simulation envelopes can be computed in order to conclude that there is a statistically significant difference between $\hat{L}(r)$ and zero at any value of *r*.

We calculated the simulation envelopes by generating *s* independent simulations under the null hypothesis of independence between the two patterns with the same estimated intensities in the study region (Diggle, 2003; BADDELEY et al., 2016). Compute the estimated *L*-functions for each of these realizations, that is, $\hat{L}^{(k)}(r)$, for *k* =2,…,*s*. Obtain the pointwise upper and lower envelopes of these simulated functions, $l(r) = \min_k \hat{L}^{(k)}(r)$ and $u(r) = \max_k \hat{L}^{(k)}(r)$, respectively. If the data came from two independent point processes, then $\hat{L}(r)$ and $\hat{L}^{(1)}(r)$,..., $\hat{L}^{(2)}(r)$ are statistically equivalent and independent, thus, the probability that $\hat{L}(r)$ is contained inside [*l*(*r*), *u*(*r*)] is equal to 2/(*s*+1) by symmetry (BADDELEY et al., 2016). Figure 3 shows the estimated *L*-function against distances *r* together with the 99% pointwise envelopes.

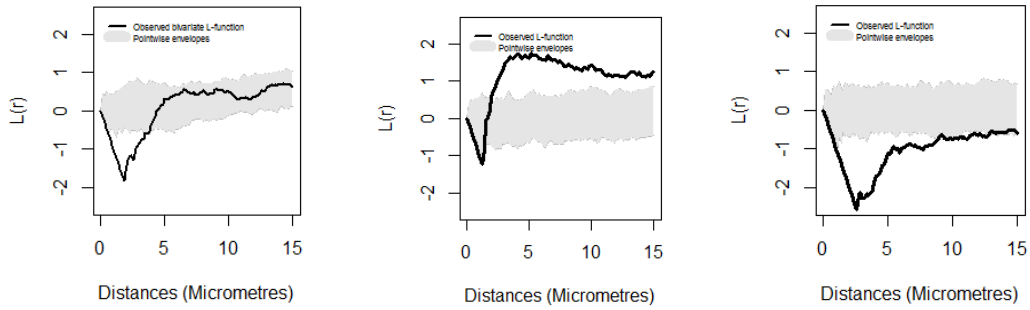

Figure 3. Estimated *L*-functions (solid line) with lower and upper envelopes from 99 independent simulations under the null hypothesis of independence between the patterns (region between the envelopes is shaded) for small (left), large (center) and both small and large (right) particles.

Figure 3 gives information about the spatial distribution of the particles over the matrix regarding not only to the behavior of each of the component patterns, but also to the interaction between the two types of SiC particles. Firstly, the difference between the curves corresponding to small and large particles suggests that, at least, with regard to second-order properties, the two component patterns are generated by two completely different stochastic mechanisms. More particularly, the values of the *L*-function for the large SiC particles are negative at distances up to 12 μm and, therefore, the curve suggests the presence of a strong regular mechanism among

the large SiC particles. On the other hand, examination of the values of the *L*-function for the small SiC particles reveal a hint of attraction (or clustering) at distances above 2.5 μm. Applied works corroborate these findings. For example, Fathy et al (2014) show that small size SiC particles (< 8 μm) leads to a non-homogeneous reinforcement distribution and to the formation of SiC clusters in a SiC-Al composite produced by a powder metallurgy technique of cold pressing at 500 MPa followed by hot extrusion at 580 ^{0}C. Cetin and Kalkanli (2009) and Fathy et al. (2014) have shown that in composites containing large SiC particles, the distribution of reinforcement tends to be uniform (regular) in the matrix.

We can also observe from Figure 3 (right) that $L(r)$ lies below the lower simulation envelope at short distances. Thus, the two types of particles exhibit negative association, resulting from repulsion between small and large particles at short distances (less than 2.8 μm). We have observed that the nearest distance between two particles is 1.35 μm. This distance is bigger than the radius of the equivalent discs of the smallest particles in the sample, that is, 0.52 μm. Thus, this repulsive behavior can be interpreted as more than a consequence of the obvious requirement for non-overlapping particles.

4. Modelling the interaction between small and large particles

The analysis presented above indicated the presence of a repulsion mechanism (regularity), at small scale, between small and large reinforcing SiC particles held in the matrix of the biocomposite material. In order to explain the nature of this mechanism, we need a stochastic point process model.

Markov (or Gibbs) point processes have proved suitable for modelling point patterns which display some degree of regularity (Baddeley et al., 2006; Ripley et al., 1977; Diggle et al., 2006). These models can be constructed by writing down either their probability density or their conditional intensity, where the second form is considered the main tool for analyzing a Gibbs point processes (Baddeley et al., 2016).

The conditional intensity of the bivariate Gibbs point process is interpreted as the conditional probability of finding a particle with a particular mark (m) near a location $\boldsymbol{u}$, given the rest of the marked point process $\boldsymbol{y}$.

Let $A\ C\ R^2$ be the study region, and M the finite set of possible marks. Then, a marked point pattern is a set of $\boldsymbol{y} = \{(x_1, m_1), \dots, (x_n, m_n)\}, \quad x_i \in A, \quad m_i \in M, \quad n \geq 0$ of pairs (x_i, m_i) of locations x_i with marks m_i.

Since we have particles classified as small and large, we use a simplified version of these models called bivariate Gibbs point process that consider only pairwise interactions among particles and, therefore, the conditional intensity of the bivariate Gibbs point process is given by

$$\lambda_{\boldsymbol{\theta}}\big((\mu, m); \boldsymbol{y}\big) = \beta_S^{n_S}(\boldsymbol{u})\beta_L^{n_L}(\boldsymbol{u}) \prod_{i=2}^{n_S} \prod_{j=1}^{i-1} h_{SS}\big(\|x_{Si} - x_{Sj}\|\big) \prod_{i=2}^{n_L} \prod_{j=1}^{i-1} h_{LL}\big(\|x_{Li} - x_{Lj}\|\big) \prod_{i=1}^{n_S} \prod_{j=1}^{n_L} h_{SL}\big(\|x_{Si} - x_{Lj}\|\big), \quad (4)$$

where ||.|| denotes the Euclidean distance between two particles, $h(.)$ are functions determining the interaction between a pair of particles x_{Si} e x_{Lj} of small (S) and large (L) particle types, and $B_S(\boldsymbol{u})$ and $B_L(\boldsymbol{u})$ determine the intensity (trend) of the process for small and large particles, respectively, at location $\boldsymbol{u}$. Since our pattern presents some degree of inhomogeneity, we are using a first-term polynomial trend given by $\beta_L^{n_L}(\boldsymbol{u}) = bL_0 + bL_1 x + bL_2 y$, and $\beta_S^{n_S}(\boldsymbol{u}) = bS_0 + bS_1 x + bS_2 y$ for modeling trend, where bL_0, bL_1, bL_2, bS_0, bS_1 e bS_2 are (scalar) parameters to be fitted, and x and y are the Cartesian coordinates.

There are many ways to express $h(.)$ in order to create different kinds of interaction behavior. In this paper we are using

$$h(.) = \begin{cases} 0 & if \quad \|.\| < H \\ \gamma & if \quad h \leq \|.\| < R. \\ 1 & if \quad \|.\| > R \end{cases} . \tag{5}$$

We are termed the conditional intensity function given by equation (4) with interaction function given by equation (5) as "inhomogeneous multitype Strauss-hard-core model". The hard-core distance parameters ($H_{..}$) specifies the radius around a particle in which no other particles cannot occur. The interaction distance parameters ($R_{..}$) determines the radius around particles in which a spatial interaction occurs and must be $R_{..} > H_{..}$. The interaction parameter ($\gamma_{..}$) specifies the strength and direction of the interaction. For distances between $H_{..}$ and $R_{..}$, the interaction parameter is interpreted as a positive interaction (attraction) when $\gamma_{..} > 1$, no interaction if $\gamma_{..} = 1$, and a negative interaction (repulsion) if $0 \leq \gamma_{..} < 1$. The hard-core distances, interaction distances, and interaction parameters are all symmetric, for example, $\gamma_{SL} = \gamma_{LS}$.

The estimation of the model parameters ($\boldsymbol{\theta}$) may be difficult because we do not have an analytic expression in a closed form for the likelihood of equation (4). To work around this problem, the model given by equation (4) can be fitted by maximum pseudolikelihood. Baddeley et al. (2016) show that the log pseudolikelihood is

$$log\ PL(\theta; y) = \sum_{i=1}^{n(y)} log\big((\boldsymbol{x}_i, m_i); \boldsymbol{y}\big) - \sum_{m \in M} \int_W \lambda_\theta\big((u, m); \boldsymbol{y}\big)\, d\boldsymbol{u}\ . \tag{6}$$

You may observe that Equation (6) is not a log of the likelihood, but the analogue of the score equation $\frac{d}{d\theta} \log PL\ (\theta) = 0$ is an unbiased estimating equation. Thus, Baddeley et al. (2006) argue that the maximum pseudolikelihood estimator is asymptotically unbiased, consistent and asymptotically normal under appropriate conditions.

The main disadvantage of the maximum pseudolikelihood estimator is that the conditional intensity of the model $\lambda_\theta\big((\mu, m); \boldsymbol{y}\big)$ must be loglinear in the parameter, that is, $\lambda_\theta\big((\mu, m); \boldsymbol{y}\big) = \theta S(u, \boldsymbol{y})$, where $S(u,\boldsymbol{y})$ is a real-valuated function of locations $\boldsymbol{u}$ and configuration y. The problem is that some parameters of the vector $\boldsymbol{\theta}$ do not appear in the loglinear form. They are called "irregular parameters". For instance, in our model we have six irregular parameters (H_{SS}, H_{LL}, H_{LS}, R_{SS}, R_{LL} e R_{LS}), and nine "regular parameters" that appear in the loglinear form (bL_0, bL_1, bL_2, bS_0, bS_1, bS_2, γ_{SS}, γ_{LL} e γ_{LS}). Thus, we have adopted a three-step procedure to perform inference on the model parameters.

In the first step, we have used the maximum likelihood method to estimate hard-core parameters ($H_{..}$), which corresponds to the minimum observed distance between two small particles (Baddeley et al., 2016).

In the second step, we have used the profile pseudo-likelihood method to estimate the interaction distance parameters ($R_{..}$). Baddeley et al. (2006, 2016) show that this method finds the interaction distances with the maximum pseudo-likelihood within a set of hard-core distances.

Once the estimates of the irregular parameters ($H_{..}$ and $R_{..}$) are set, in the third step, we have used the maximum pseudo-likelihood, proposed by Turner and Baddeley (2000), to perform inference on the regular parameters.

By using this approach, we have got the following estimates: $\widehat{H}_{SS} = \widehat{H}_{LS} = \widehat{H}_{SL} = 1.35$ μm, $\widehat{H}_{LL} = 3.08$ μm, $\widehat{R}_{SS} = 3.38$ μm, $\widehat{R}_{LL} = 3.22$ μm, $\widehat{R}_{SL} = \widehat{R}_{LS} = 1.36$ μm, $b_{L0} = -5.118$, $b_{L1} = 0.0008$, $b_{L2} = 0.0001$, $b_{S0} = -5.639$, $b_{S1} = 0.0022$, $b_{S2} = 0.0023$, $\hat{\gamma}_{SS} = 2.21$, $\hat{\gamma}_{LL} = 0.455$, and $\hat{\gamma}_{SL} = 0.00003$.

These results tell us that the estimated intensity of large particles in the composite would increase by a factor of *exp*(-5.118+0.008+0.0001) = 0.00604 if the slopes increased to 1. The

largest slope values in data are (192, 288), at which stage the predicted intensity has risen from 0.00599 to 0.00719 particles/μm^2. Basically, the same trend behavior would be observed for small particles and, therefore, the model managed to capture the slight non-stationarity of the two point processes.

These results also show that the shortest distance between any two particles in the material is 1.35 μm, however allowing a radius around particles between 1.36 and 3.38 μm in which a spatial interaction may occur.

Evidence for dependence between particle of different types is quantified by the interaction parameter ($\hat{\gamma}_{SL}$). This is near zero, suggesting a strong negative interaction (repulsion). The estimate of within-large particles interaction ($\hat{\gamma}_{LL}$) is below unity, suggesting inhibition. The estimate of within-small particles interaction ($\hat{\gamma}_{SS}$) is much greater than unity, suggesting aggregation.

A more effective way to understand the fitted interaction parameters is to produce a plot of the fitted interaction function like shown in Figure 4. The function $h(r)$ in each panel is the (exponentiated) pairwise interaction between points of the given types.

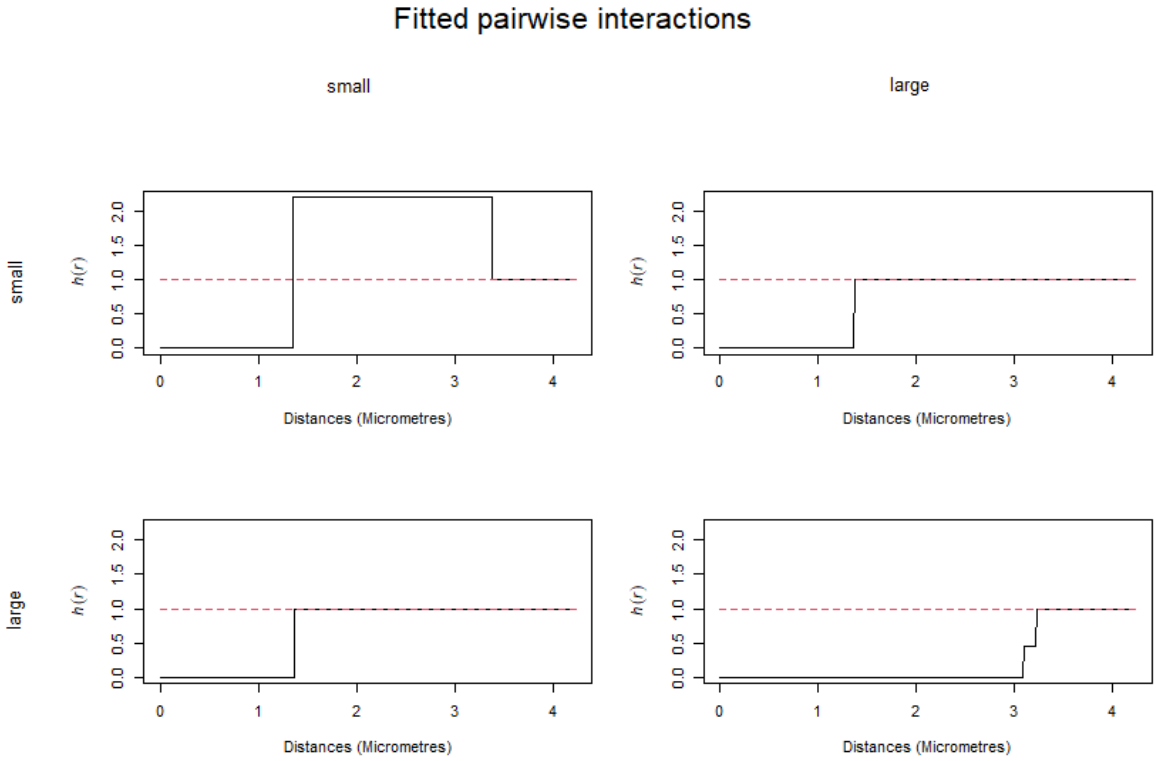

Figure 3. Interpoint interaction functions for the non-stationary multitype Strauss-hard-core model fitted to the observed point pattern marked by type (small, large).

The diagonal panels (large-small and small-large) of Figure 3 show evidence for dependence (inhibition) between particle of different types up to 1.3 μm, while the large-large panel suggest there is a quite strong inhibition between points of large particles up to 3 μm. The panel small-small suggests evidence for inhibition up to 1.3 μm and evidence for aggregation between 1.3 and 3.4 μm.

From a physical point of view, this singular behavior (attraction and repulsion) can be interpreted as a consequence of the actual requirement for non-overlapping particles that creates this repulsive effect on one hand, and of a constant pressure of the small particles to fill the large gaps among the large particles on the other. In other words, we can say that the small particles are aggregated in the gaps between the large particles but they tend to repel each other when they are very close together.

5. Model validation

Having fitted a point process model to data, it is important to validate it to check that the model provides a good fit to the data. Baddeley *et al*. (2016) argues that there is little theory available to support goodness-of-fit testing and model validation for a fitted Gibbs point process model and, therefore, goodness-of-fit methods often rely on summary functions such as *F*, *G*, *J*, *K* and *L*.

Baddeley *et al.* (2016) suggests using the same Monte Carlo approach that we have applied to test the null hypothesis of independence between the two processes. We simply have to generate simulated realizations from the fitted model, compute the *L*-function for each simulated realization, and construct the pointwise upper and lower simulation envelopes. The result is shown in Figure 4.

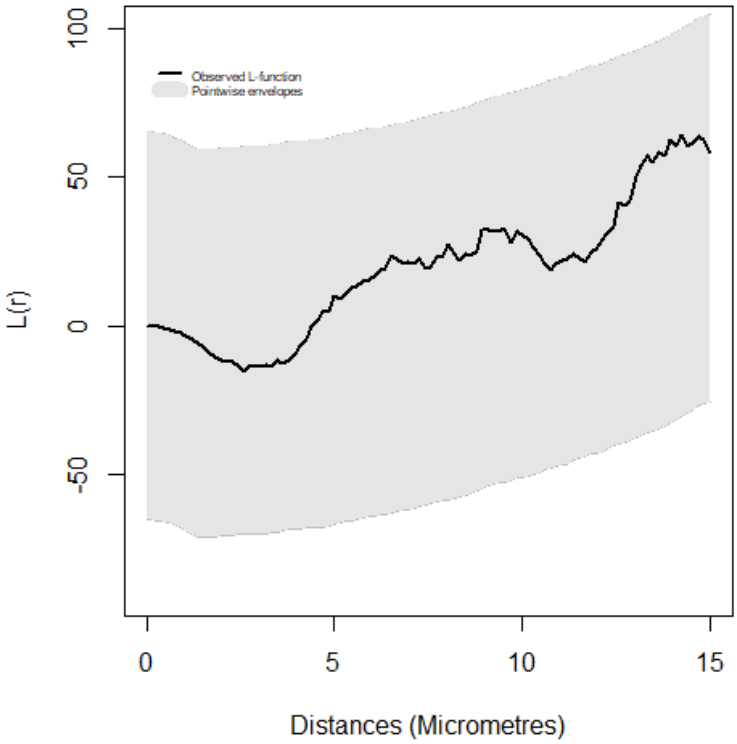

Figure 4. Estimated *L*-functions (solid line) with lower and upper envelopes from 99 independent simulations under the null hypothesis of the non-stationary (inhomogeneous) multitype Strauss-hard-core model.

Figure 4 shows that the null hypothesis would be accepted at the 1% level with any choice of distance r and, suggests an obvious adequacy in the fit of the inhomogeneous multitype Strauss-hard-core model to the bivariate point pattern.

For a more formal goodness-of-fit assessment, Diggle *et al*. (2006) suggest using the test statistic $T = \sum_{r=1}^{R}\left[\left\{\hat{K}_{SL}(r) - \overline{K}_{SL}(r)\right\}/r\right]^2$, where $\hat{K}_{SL}(r)$ are the estimates of $K_{SL}(r)$ calculated from the data and, $\overline{K}_{SL}(r)$ the corresponding mean of estimates from s independent realisations from the fitted model (null hypothesis). The corresponding values of the test statistic, say $T_i = T(\boldsymbol{x}(i))$ for $i = 2,\ldots,s$ and, the rank of T in the set of values $\{T_1,\ldots,T_m\} \cup \{T\}$, that is, $R = 1 + \sum_{i=1}^{m} 1\{T_i > T\}$ are computed. Then, under the null hypothesis, the rank R is uniformly distributed on $\{1, 2,\ldots,s\}$, assuming there are no ties. Hence, the associated *p*-value is $p = \frac{R}{s+1}$.

Based on 999 simulations of the fitted model, the attained *p*-value of the Monte Carlo test

was 0.125, and also indicating a good overall fit of the inhomogeneous multitype Strauss-hard-core model and, therefore, proving that the sizes of the second-phase reinforcing particles are related to their location in the biocomposite material.

Although, we have performed our goodness-of-fit analysis using only the bivariate L-function, it would be possible to consider nearest neighbor properties using other bivariate pattern descriptors such as the F, G and, J functions described by Diggle (2003) and Baddeley *et al*. (2016). We have also carried out the analysis using these functions (not shown here) and we have reached the same conclusions as reached by using the L-function.

6. Final remarks

One may argue that one sample is not enough to draw conclusions whether or not the sizes of the reinforcing particles are related to their location in the composite material. We certainly agree that it would be necessary analyzing more samples of the material for a conclusive characterization of the biocomposite material. Our aim here was to present a methodology that could be applied to many samples as necessary. In fact, we have applied to the full set of 37 samples of this material and we have reached the same conclusions.

Other problem related to our work would be the choice of the cut off point between large and small particles. It is possible that this choice would affect the results of the bivariate spatial analysis. We have performed the analysis using two others cut off points (not shown here), one smaller and the other larger than the actual cut off point used in this work. We observed that in both cases the analysis continued to show the same characterization of the bivariate spatial point pattern. The use of the smaller cut off point decreased the degree of negative association for short distances and increased the degree of positive association for moderate distances between the two types of particles. The use of the bigger cut off point produced opposite results. These results may suggest that the smaller the particles, the stronger the degree of positive association and that the larger the particles, the stronger the degree of negative association between the two types of SiC particles.

We now return to our decision to consider a simplified version of the size-spatial relationship, focusing only on the spatial distributions of ‘large’ and ‘small’ particles instead of using the actual sizes of particles. Of course, the full statistical analysis of a two-phase material microstructure using marked point processes theory, with quantitative marks, is physically very appealing. In this paper, we have followed an approach based on the theory of bivariate point patterns because it is the simplest case of a marked point process and, therefore, we consider this the starting point for exploring the relationship between spatial and size distributions. Methods that are more sophisticated (Stoyan *et al*., 2002) might be carried out if the simple version proves inadequate. Results show that our feasible approach works appropriately. So, it may be possible to avoid more complex methods in many cases, depending on the objective of the analysis.

It is well known that model validation to check formally that the model is a good fit to the data, and that all terms in the model are appropriate is of prime importance in statistical modelling. Unfortunately, model validation in spatial point processes is under development, and available just for simple models. For instance, residuals from the fitted model have recently been developed only for (homogeneous or inhomogeneous) Poisson processes (Baddeley et al., 2006, 2016). Statistical theory of parameter estimation and hypothesis testing for spatial point processes are rather limited, and depend on the class of models envisaged. Once more, for homogeneous (or inhomogeneous) Poisson processes, much of the classical theory of maximum likelihood and methods available for formal inference are applicable. For marked (bivariate) Gibbs processes, to the best of our knowledge, there is no statistical theory for hypothesis tests based on the pseudolikelihood. Thus, the analysis is based almost exclusively on the literal interpretation of the parameter estimates and, therefore, Monte Carlo tests based on simulations of summary functions are very popular (Baddeley et al., 2006, 2016)

Finally, it worst to point out that the described methods presented in this paper for quantitative characterization of the interaction between spatial and size distributions of second-phases in particulate composite materials can be easily adapted to a computer system in order to perform quality control in the production process of particulate metal matrix composites and, further associate these features with the mechanical properties of the material.

7. Conclusions

We advocate in this paper that the characterization of the interaction between spatial and size distributions of second-phases in particulate composite material applied for manufacturing biomedical instruments might be carried out by using methods from bivariate spatial (marked) point processes theory. We perform a characterization of a specimen of an Al-SiC composite material by conducting four steps: exploratory analysis of the first order properties, testing the hypothesis of no interaction of particles, fitting an appropriate model and testing goodness-of-fit. This approach shows that the component patterns of small and large particles are generated by two completely different stochastic mechanisms. There is evidence of significant clustering among the small particles and, on the other hand, there is a presence of a moderately regular spacing among the large particles. The analysis also showed a presence of some degree of interaction between small and large particles, suggesting that small particles are aggregated in the gaps among large particles, but they tend to repel each other at small scales.

Acknowledgments

We dedicate this paper in memory of Nick Fieller (Sheffield University). We would like to thank Helen Atkinson (Leicester University) for providing the images of the composite material. We are indebted to Eleanor Stillman (Sheffield University), Paul Blackwell (Sheffield University), reviewer and editor for their valuable comments and suggestions which have substantially improved the manuscript. Authors are grateful to CAPES and CNPq for the financial support.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Baddeley, A., Møller, J. & Waagepetersen, R. P. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica* **54**, 329–350 (2000).
2. Baddeley, A., Rubak, E. & Turner, R. *Spatial Point Patterns: methodology and applications with R*. (CRC Press, 2016).
3. Baddeley, A. & Turner, R. Spatstat: an *R* package for analyzing spatial point patterns. *Journal of Statistical Software*, **12**, 1–42 (2005).
4. Baddeley, A. & Turner, R. Modelling spatial point patterns in R. In: A. Baddeley, P. Gregori, J. Mateu, R. Stoica, D. Stoyan (Eds.), *Case Studies in Spatial Point Pattern Modelling*. (Springer-Verlag, 2006).
5. Chawla, N. & Shen, Y. H. Mechanical behavior of particle reinforced metal matrix composites. *Advances in Engineering of Materials* **3**, 357–370 (2001).
6. Cetin, A. & Kalkanli, A. Multi-scale characterization of particle clustering in discontinuously reinforced composite. *Materials Characterization* **60**, 268–572 (2009).
7. Davison, D. L. Fatigue and fracture toughness of aluminium alloys reinforced with SiC and alumina particles. *Composites* **24**, 248–255 (1993).

8. Derr, R. & Ji, C. *Fitting microstructural models in Materials Science*. www.unc.edu/depts/stator/miscellaneous/cji/scan05.pdf. 2000.
9. Diggle, P. J. *Statistical Analysis of Spatial Point Patterns*. (Hodder Education, 2003).
10. Diggle, P. J., Eglen, S. J. & Troy, J. B. Modelling the bivariate spatial distribution of amacrine cells, In: A. Baddeley, P. Gregori, J. Mateu, R. Stoica, D. Stoyan (Eds.). *Case Studies in Spatial Point Pattern Modelling.* (Springer-Verlang, 2006).
11. Dwivedi, S. P., Maurya, M. & Chauhan, S. S. Mechanical, physical and thermal behaviour of SiC and MgO reinforced aluminium based composite material. *EVERGREEN Joint Journal of Novel Carbon Resource Sciences & Green Asia Strategy* **8**, 318–327 (2021).
12. Fathy, A., Sadoun, A & Abdelhameed, M. Effect of matrix/reinforcement particle size ratio (PSR) on the mechanical properties of extruded Al–SiC composites. *International Journal of Advanced Manufacturing Technology* **73**, 1049–1056 (2014).
13. Ghosh, S, Nowak, Z. & Lee, K. Quantitative characterization and modeling of composite microstructures by Voronoi cells. *Acta Materialia* **45**, 1215–1234 (1997).
14. Hong, S. J. *et al*. Effect of clustering on the mechanical properties of SiC particulate reinforced aluminum alloy 2024 metal matrix composites. *Material Science and Engineering A* **347**, 198–204 (2003).
15. Lewandowski, J. J., Liu, C. & Hunt Jr., W. H. Effects of matrix microstructure and particles distribution on fracture of aluminum metal matrix composites. *Material Science and Engineering A* **107**, 241–255 (1989).
16. Lotwick, H. W. & Silverman, B. W. Methods for analyzing spatial processes of several types of points. *Journal of the Royal Statistical Society B* **44**, 406-413 (1982).
17. Møller, J. & Waagepetersen, R. *Statistical Inference and Simulation for Spatial Point Processes*. Boca Raton: Chapman and Hall/CRC, 2003.
18. Murphy, A. M., Howard, S. J. & Clyne, T. W. Characterization of severity of particle clustering and its effect on fracture of particulate MMC's. *Journal of Material Science and Technology* **14**, 959–969 (1989).
19. Narayanasamy, R., Ramesh, T. & Prabhakar, M. Effect of particle size of SiC in aluminum matrix on workability and strain hardening behavior of P/M composite. *Material Science and Engineering A* **504**, 13–23 (2009).
20. Pegel, S. *et al*. Spatial statistics of carbon nanotube polymer composites. *Polymer* **50**, 2123–2132 (2009).
21. Pyrz, R. Quantitative description of the microstructure of composites. Part I: Morphology of unidirectional composite systems. *Composite Science and Technology* **50**, 197–208 (1994).
22. R core team, *R: a language and environment for statistical computing*. Viena: R Foundation for Statistical Computing, http://www.R-project.org. 2022.
23. Ripley, B. D. The second-order analysis of stationary processes. *Journal of Applied Probability* **13**, 255–266 (1976).
24. Ripley, B. D. & Kelly, F. P. Markov point processes. *Journal of the London Mathematical Society* **15**, 188–192 (1977).
25. Salernitano, E. & Migliaresi, C. Composite Materials for Biomedical Applications: A Review. *Journal of Applied Biomaterials and Biomechanics* **1**, 3–18 (2003).
26. Scalon, J. D. *et al*. Spatial pattern analysis of second-phase particles in composite materials. *Material Science and Engineering A* **356**, 245-257 (2003).
27. Scalon, J. D. et al. A model-based analysis of particle size distributions in composite materials. *Acta Materialia* **51**, 997–1006 (2003a).

28. Spitzig, W. A., Kelly & J. F. Richmond, O. Quantitative characterization of second-phase populations. *Metallography* **18**, 235–261 (1985).
29. Stoyan, D. Systems of hard particles: Models and statistics. *Chinese Journal of Stereology and Image Analysis* **7**, 1–14 (2002).
30. Stoyan, D. & Schnabel, H. D. Description of relations between spatial variability of microstructure and mechanical strength of alumina ceramics. *Ceramic International* **16**, 11-18 (1990).
31. Stoyan, D. & Stoyan, H. On one of Matern's hard-core point process models. *Mathematische Nachrichten* **122**, 205-214 (1985).
32. Stoyan, D. & Wiencek, K. Spatial correlations in metal structures and their analysis. *Material Characterization* **26**, 167-176 (1991).
33. Torralba, J. M., Costa, C. E. & Velasco, F. P/M aluminum matrix composites: an overview. *Journal of Material Procedures Technology* **133**, 203–206 (2003).
34. Turner, R. & Baddeley, A. Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics* **42**, 283–322 (2000).
35. Willians, J. J. Effect of overaging and particle size on tensile deformation and fracture of particle-reinforced aluminum matrix composites. *Metallurgical and Materials Transactions A* **33**, 3861–3869 (2002).
36. Wincek, K. & Stoyan, D. Correlations in metal structures and their analysis. II: The covariance. *Material Characterization* **38**, 47-53 (1993).