

On Assessing Spatial Uniformity of Particle Distributions in Quality Control of Manufacturing Processes

Kin Ming Kam^a, Li Zeng^{a*}, Qiang Zhou^b, Richard Tran^c, Jian Yang^c

^a*Department of Industrial and Manufacturing Systems Engineering, The University of Texas at Arlington, 500 West First Street, P.O. Box 19017, Arlington, TX 76019, U.S.A.*

^b*Department of Industrial and Systems Engineering, The University of Wisconsin-Madison, 1513 University Avenue, Madison, WI 53706, U.S.A.*

^c*Department of Bioengineering, The University of Texas at Arlington, 500 UTA Blvd., P.O. Box 19138, Arlington, TX 76010, U.S.A.*

**Correspondence author. Tel.: +1 817 272 3150; Fax: +1 817 272 3406; E-Mail: lzens@uta.edu*

There are many situations in quality control of manufacturing processes in which the quality of a process is characterized by the spatial distribution of certain particles in the product, and the more uniform the particle distribution is, the better the quality is. To realize quality control and guide process improvement efforts, the degree of spatial uniformity of particle distributions needs to be assessed. On the other hand, many quantitative metrics have been developed in areas outside manufacturing for measuring uniformity of point patterns, which can be applied for this purpose. However, critical issues exist in applying existing metrics for quality control relating to which metrics to choose and how to use them in specific situations. To provide general guidelines on these issues, this research identifies popular uniformity metrics scattered in different areas and compares their performance in detecting nonuniform particle distributions under various practical scenarios through a comprehensive numerical study. Effects of different factors on the performance of the metrics are revealed and the best metric is found. The use and effectiveness of the selected metric is also demonstrated in a case study where it is applied to data from emerging material fabrication processes in nanomanufacturing and biomanufacturing.

Keywords: Complete spatial randomness (CSR); Particle distribution; Point patterns; Spatial uniformity; Metal matrix nanocomposite (MMNC); Tissue-engineered scaffolds

1. Introduction

There is an increasing situation in manufacturing processes where the quality of a process or product is characterized by the spatial distribution of certain particles or components, and the more uniform the particle distribution, the better the quality. This is especially the case in some emerging manufacturing branches such as nanomanufacturing and biomanufacturing. For example, in the metal matrix nanocomposite (MMNC) fabrication processes where nano-sized ceramic particles are embedded into the

metal matrix [1, 2], a key quality characteristic is the dispersion of nanoparticles, as shown in the scanning electron microscope (SEM) image in Figure 1(a). The introduction of nanoparticles can significantly strengthen the metal matrix, and the more uniformly the nanoparticles disperse, the better the composite structure and thus the mechanical properties of the produced MMNC material [3]. Another typical example in biomanufacturing is the fabrication processes of tissue-engineered scaffolds, which are porous polymer matrices serving as temporary substrates for cells in developing engineered tissues/organs. Introduction on tissue-engineered scaffolds can be found in the literature [e.g., 4, 5, 6]. Figure 1(b) shows an SEM image of the cell distribution in the scaffolds. The uniformity of cell distributions is highly required to produce homogeneous tissues/organs, which is, correspondingly, considered as an important quality indicator of the scaffolds.

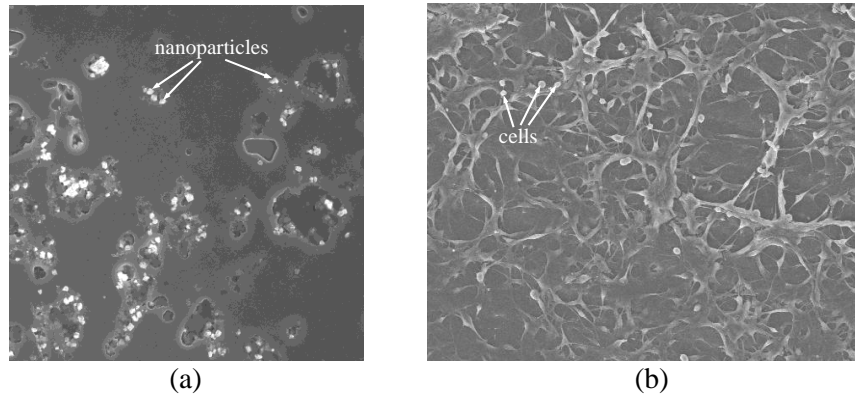


Fig. 1. SEM images of nanoparticle distribution in MMNC (a) and cell distribution in tissue-engineered scaffolds (b)

To conduct quality control and guide process improvement efforts in processes like the above examples, the degree of spatial uniformity of particle distributions needs to be assessed. However, currently, this is reported subjectively by human operators in most cases, which is neither convenient nor reliable. Obviously, quantitative metrics are needed to assess the degree of uniformity, and consistent product quality can be achieved through monitoring those metrics in the manufacturing process.

Particles like the nanoparticles and cells can be treated as dimensionless points due to their small sizes compared to the study regions, i.e., the SEM images, and their distributions are formally referred to as *spatial point patterns* in the literature [e.g., 7, 8, 9]. Quantifying the uniformity/clustering of point patterns has been an extensively studied topic in areas such as geological, ecological, environmental, and material science studies, and many different types of metrics have been developed for typical point patterns in those areas such as the distribution of certain species of plants/insects within an interested region or the mixing of different components in material processing. These metrics can be directly applied to assess the uniformity of particle distributions in quality control of manufacturing processes. However, this endeavor is faced with three critical issues: (i) *Selection of metrics*: Little knowledge on the selection

of metrics under specific situations is available in the literature, which makes it very difficult for practitioners in manufacturing to decide the suitable metric to use in their process from the vast pool of available metrics. (ii) *Specification of parameters*: The use of uniformity metrics often requires specifying some parameters. For example, to decide if the observed patterns are uniform, a benchmark of uniformity, called critical value in process monitoring, should first be established. Also, to use metrics based on a division of the study region into small grids, the number of grids should be determined beforehand. However, there is a lack of guidance on how to specify such parameters. (iii) *Robustness of metrics*: The performance of existing metrics in some special cases in quality monitoring is unknown. For example, while a large number of points, e.g., several hundreds or more, are often available in patterns in geological, environmental and material science studies, patterns with a small number of particles might be a typical situation in some manufacturing processes. Therefore, it will be helpful to study the robustness of existing metrics in such cases.

The only effort to address the above issues in the literature is the study by Zhou, *et al.* [10] which compares some uniformity metrics in assessing nanoparticle dispersion in MMNC fabrication processes through simulation. However, only a limited set of metrics is considered in that study, and the last two issues are not fully addressed. This paper will contribute by studying the performance of a broader set of metrics, encompassing those developed in various areas, and providing guidelines on all the three issues. The effects of factors that may affect the performance of the metrics are identified, and the best metric is found in terms of its performance in different scenarios of parameter setting. A case study is also provided to demonstrate the implementation of the best metric on image data from nanocomposite and tissue-engineered scaffold fabrication processes.

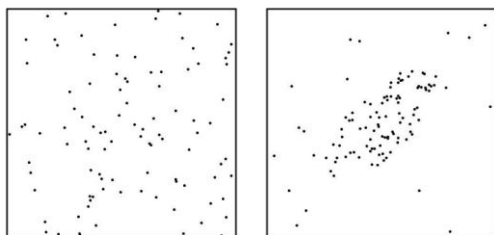
The remainder of the paper is organized as follows. In Section 2, a review of the existing methods and metrics for assessing spatial uniformity of point patterns will be presented. Section 3 will introduce the conventional uniformity monitoring procedure in manufacturing process control practice, and discuss important factors affecting the performance of the metrics. The results of the numerical study will be given in Section 4. This is followed by a case study in Section 5 and summary in Section 6.

2. Methods and Metrics for Assessing Spatial Uniformity of Point Patterns

2.1. Methods for assessing spatial uniformity

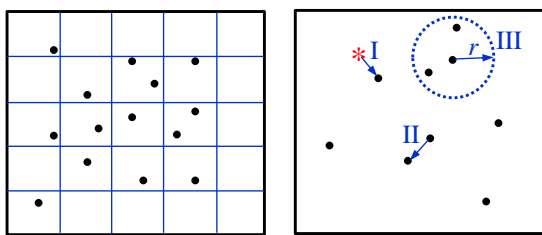
The basis for spatial uniformity assessment of point patterns is the concept of *complete spatial randomness* (CSR) represented by the pattern in Figure 2(a), which follows a homogeneous Poisson distribution. Using CSR as a reference, the degree of spatial uniformity of a pattern can be quantified by its departure from CSR, and a pattern which has a statistically significant departure will be concluded to be nonuniform in quality monitoring. As nonuniformity in a pattern is typically manifested by the

existence of clustering shown in Figure 2(b), the degree of uniformity is often inversely represented by the degree of clustering. Various methods have been developed to assess uniformity, which differ from each other in their definition of the departure from CSR and the way to calculate this departure.



(a) CSR (b) Clustering

Fig. 2. Point patterns



(a) Quadrat methods (b) Distance methods

Fig. 3. Existing methods for assessing uniformity

In the high level, the existing methods can be roughly categorized into three classes: quadrat-based methods, distance-based methods and other methods. As illustrated in Figure 3(a), quadrat-based methods, such as those proposed by Fiser [11], David and Moore [12], Morisita [13], Lloyd [14], Douglas [15], Alemaskin *et al.* [16], and Greig-Smith [17], divide the study region into a number of small grids, called quadrats, and count the number of points falling into each grid. Consequently, the degree of uniformity can be quantified by the characteristics of these counts such as the parameters of their empirical distribution. Such methods are very popular in some areas such as ecological and environmental studies, for their simplicity and convenience in implementation. Their main drawback is that the spatial information contained in the pattern is lost since only the counts are used. In addition, the appropriate specification of the number of quadrats is case specific. Distance-based methods, illustrated in Figure 3(b), focus on the distances between points, such as those between nearest neighbors or between randomly selected locations (indicated by the asterisk in the figure) to nearest points, and quantify the degree of uniformity using the characteristics of these distances. Cressie [7] gives an excellent review of these methods. Compared with quadrat-based methods, these methods utilize more spatial information, but choosing the distances between nearest neighbors instead of those between the second, third, etc., nearest neighbors, is arbitrary, as pointed out by Cressie [7]. They may also cause the “edge effect” problem when some points or the selected location is closer to the edge of the study region than to any other points within the region. The methods for edge effect correction are summarized in Diggle [8].

Other methods for quantifying spatial uniformity include the methods based on coordinate projection proposed by Jun *et al.* [18], and Tong *et al.* [19], and the SADIE (spatial analysis by distance indices) methodology developed by Perry and Hewitt [20], and Perry [21]. The projection methods project the coordinates of the points onto a rotating axis and characterize the clustering of points by the distances between the projected values. This method is parameter-free and convenient, but it requires a relatively large number of points in the pattern. The SADIE methodology, originally developed for quantifying the

spatial variation of insect distribution in ecological studies, is essentially a hybrid of quadrat and distance methods. Basically, it divides the study region into quadrats and quantifies the uniformity by measuring the total efforts (in terms of distances moved), to rearrange the points in different quadrats to create a pattern as uniform as possible. This methodology not only bears the limitations of quadrat-based methods, but also involves complex optimization algorithms which are not convenient for quality control in practice. Therefore, the associated metrics will not be considered in this study.

Table 1. Quadrat-based metrics for assessing spatial uniformity

Name	Formula	Decision rule
Index of Dispersion	$ID = (q-1)s^2 / \bar{x}$	Upper-sided
Skewness Index	$SI = \frac{q}{(q-1)(q-2)} \sum_{i=1}^q \left(\frac{x_i - \bar{x}}{s} \right)^3$	Upper-sided
Moran's I	$MI = \frac{q}{(q-1)s^2 \omega_{..}} \sum_{i=1}^q \sum_{j=1}^q \omega_{ij} (x_i - \bar{x})(x_j - \bar{x})$	Two-sided
Geary's C	$GC = \frac{1}{2s^2 \omega_{..}} \sum_{i=1}^q \sum_{j=1}^q \omega_{ij} (x_i - x_j)^2$	Two-sided
Local Moran's I	$LMI = \max_{1 \leq i \leq q} \left[\frac{q}{(q-1)s^2} (x_i - \bar{x}) \sum_{j=1}^q \omega_{ij} (x_j - \bar{x}) \right]$	Two-sided
Local Gi	$LG = \max_{1 \leq i \leq q} \left[\sum_{j=1}^q \omega_{ij} x_j / \sum_{j=1}^q x_j \right]$	Two-sided
Global Shannon Entropy	$GSE = - \left[\sum_{i=1}^q p_i \log(p_i) \right] / \log(q), \quad p_i = x_i / \sum_{i=1}^q x_i$	Lower-sided
Global-local Shannon Entropy	$GLSE = - \left[\sum_{i=1}^q p_i \log \left(\sum_{j=1}^q \omega_{ij} p_j \right) \right] / \log(q), \quad p_i = x_i / \sum_{i=1}^q x_i$	Lower-sided

2.2. Metrics to compare

Based on the methods described above, many uniformity/clustering metrics have been developed and widely used in different areas. Their definitions, formulas and decision rules in quality monitoring are listed in Table 1 and 2. Note that in the “Decision rule” column, “Upper-sided” means that an upper-sided critical value will be applied, i.e., an observed pattern with a metric value larger than the critical value will be concluded to be nonuniform. Similarly, “Lower-sided” and “Two-sided” mean that a lower-sided critical value and two-sided critical values, respectively, will be applied.

Quadrat-based metrics

Table 1 displays the metrics resulted from quadrat-based methods. The quantity q in the formulas denotes the number of quadrats, i.e., the number of grids in Figure 3(a), x_i is the count of points in grid i , $i=1, 2, \dots, q$, and \bar{x} and s are the sample mean and sample standard deviation of the counts. These metrics can be further divided into three groups according to their principles: *simple indices based on empirical distribution of counts* (Group 1), *spatial autocorrelation measures* (Group 2), and *entropic measures*

(Group 3). The first group is based on the fact that when the pattern is not uniform, the empirical distribution of the counts $\{x_1, x_2, \dots, x_q\}$ will be different from that under CSR. Correspondingly, the degree of uniformity is quantified by the discrepancy between the distribution of counts and that under CSR. The second group is based on the fact that when there is clustering, the counts of grids within spatial vicinity will bear some correlation instead of being independent under CSR. Thus, the degree of uniformity is indicated by the degree of correlation between counts of spatial neighbors. The third group is based on the fact that when the pattern is perfectly uniform, the counts of different grids will be the same, representing the highest uncertainty of information about the point locations, or the maximal value of the Shannon entropy. Accordingly, the degree of uniformity can be quantified by the discrepancy between the observed entropy and the entropy under CSR. More details of these metrics are provided as follows.

Group 1: Simple Indices based on empirical distribution of counts

- **Index of Dispersion (ID):** This is the simplest and most widely used metric of spatial uniformity, which concerns variability of the empirical distribution of counts in the sense that when there is clustering, this variability will be larger than under CSR as quadrats near the center of the cluster will contain more points than others. Therefore, a large value of this index indicates nonuniformity or clustering. Under CSR, this index follows a χ^2 distribution with $(q-1)$ degrees of freedom.
- **Skewness Index (SI):** This index is based on a similar idea, that is, when there is clustering, the count distribution will be biased, which is indicated by the skewness of the distribution.

Group 2: Spatial autocorrelation measures

- **Moran's I (MI) and Geary's C (GC):** These are the classic measures of spatial autocorrelation which can be found in Schabenberger and Gotway [22]. The quantity ω_{ij} in the formulas is the neighbor indicator of grid i and j , and $\omega_{..}$ is the sum of all the neighbor indicators, i.e.,

$$\omega_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are neighbors} \\ 0 & \text{if } i \text{ and } j \text{ are not neighbors} \end{cases} \quad \omega_{..} = \sum_{i=1}^q \sum_{j=1}^q \omega_{ij}$$

The neighbor indicator can be defined in many ways, e.g., grid i and j are neighbors if they share a common border. Essentially, MI and GC are global correlation measures in the sense that they measure the global degree of correlation among spatial neighbors by averaging the local correlations between neighbors. As correlation has two directions, positive correlation (i.e., counts of neighbors are more similar than expected by chance) and negative correlation (i.e., counts of neighbors are more dissimilar than expected by chance), the decision rule in using these metrics is two-sided.

- **Local Moran's I (LMI) and Local Gi (LG):** To enhance the sensitivity of clustering detection, local indicators of spatial autocorrelation (LISAs) have been developed by authors such as Getis and Ord [23] to measure the degree of autocorrelation at each single grid, as represented by the quantities

in the square brackets in the formulas. These measures are designed to capture local clusters or “hot spots”. For quality control purpose, the maximum of the local measures will be monitored.

Group 3: Entropic measures

- **Global Shannon Entropy (GSE):** This is the commonly used entropic measure, where p_i is the observed probability that a point falls into grid i . Under CSR, this probability is the same over all the grids, i.e., $p_i \equiv 1/q$, and thus $GSE=1$. A small value of this measure indicates a large departure from CSR, or a low degree of uniformity. This measure has been applied by Camesasca *et al.* [24] and Alemaskin, *et al.* [25] to assess mixing quality in material processing.
- **Global-local Shannon Entropy (GLSE):** Based on the similar idea as the LISAs, Ceccato and Karlström [26] propose the GLSE which takes the local information between spatial neighbors into consideration in measuring the entropy.

Table 2. Distance-based metrics for assessing spatial uniformity

Name	Formula	Functional summary	CSR reference	Decision rule
F index	$\int_0^D [\hat{F}(r) - F_{CSR}(r)]^2 dr$	$F(r) = P(d \leq r)$ d : distance from a selected location to nearest point	$F_{CSR}(r) = 1 - \exp\left(-\frac{\pi r^2}{\text{total area}} n\right)$	Lower-sided
G index	$\int_0^D [\hat{G}(r) - G_{CSR}(r)]^2 dr$	$G(r) = P(d \leq r)$ d : distance from a selected point to nearest point	$G_{CSR}(r) = 1 - \exp\left(-\frac{\pi r^2}{\text{total area}} n\right)$	Upper-sided
J index	$\int_0^D [\hat{J}(r) - J_{CSR}(r)]^2 dr$	$J(r) = \frac{1 - G(r)}{1 - F(r)}$	$J_{CSR}(r) = 1$	Lower-sided
L index	$\int_0^D [\hat{L}(r) - L_{CSR}(r)]^2 dr$	$L(r) = \sqrt{\frac{N(r)}{\lambda\pi}}$ $N(r)$: number of points within distance r of a selected point	$L_{CSR}(r) = r$	Upper-sided
g index	$\int_0^D [\hat{g}(r) - g_{CSR}(r)]^2 dr$	$g(r) = \sqrt{\frac{N'(r)}{2\pi\lambda}}$ $N'(r)$: the derivative of $N(r)$	$g_{CSR}(r) = 1$	Upper-sided

Distance-based metrics

Distance-based methods described in Section 2.1 have produced a series of functional summaries of the distances between points which are given in the third column of Table 2, where $r>0$ denotes the distance. The corresponding theoretical functions under CSR are given in the next column in the table, where “total area” denotes the total area of the study region, and n is the total number of points in the observed pattern. Details of these functions are given as follows.

- **F Function:** $F(r)$, represented by case I in Figure 3(b), is the cumulative distribution function of distances from a randomly selected location within the study region to its nearest point. Such distances are often referred to as empty space distances or void distances.
- **G Function:** $G(r)$, represented by case II in Figure 3(b), is the cumulative distribution function of distances from a randomly selected point to its nearest neighbor, called nearest neighbor distances.
- **J function:** $J(r)$ is a function of $F(r)$ and $G(r)$. Under CSR, $F(r)$ and $G(r)$ are the same, and thus $J(r)$ is the constant 1.

- **L Function:** $L(r)$, represented by case III in Figure 3(b), is related with the ratio of $N(r)$, the number of points within distance r of a randomly selected point, and the intensity λ in the study region. Under CSR, $N(r) = \lambda \cdot \pi r^2$, and thus $L(r)$ is r .
- **g Function:** $g(r)$ is called pair correlation function which is related with the ratio of $N'(r)$, the first derivative of $N(r)$, and $r\lambda$. It can be roughly explained as the probability of observing a pair of points with the given distance r . Under CSR, this probability is the same for any value of r .

As quality monitoring is typically based on univariate statistics, single-number summaries of the above functions need to be found. A common choice is the Cramer-von Mises statistic which measures the overall discrepancy between the observed function and its CSR reference

$$\int_0^D [\hat{H}(r) - H_{CSR}(r)]^2 dr$$

where D is the maximum distance to be considered, $\hat{H}(r)$ is the observed function, and $H_{CSR}(r)$ is the corresponding CSR reference.

Other metrics

- **Projection Index (PI):** The detailed procedure for calculating this index can be found in Appendix A. Its value under CSR is 1, and larger values indicate a higher degree of clustering.

3. Uniformity Monitoring in Quality Control of Manufacturing Processes

3.1. Uniformity monitoring procedure

Monitoring of spatial uniformity in manufacturing processes is typically based on image data like the SEM images shown in Figure 1, each representing an observed point pattern. The central goal is to determine whether an observed pattern is uniform or not. For this purpose, the critical value for the pattern must be first found based on the statistical (null) distribution of the metric under CSR. Given the specified type I error probability α , the critical value is the 100α and $100(1-\alpha)$ percentile of the null distribution for lower-sided and upper-sided decision rule, respectively. Then the metric value of the pattern will be calculated and compared with the critical value. Decision will be made based on the corresponding decision rule of the metric. The performance of the monitoring can be measured by the detection power, i.e., the probability that a nonuniform pattern is correctly detected.

In this study, both the critical value and detection power of each metric will be determined by simulation. The use of simulations here is mainly due to the lack of asymptotics of uniformity metrics. In fact, simulation methods have been a popular way to evaluate uniformity of spatial patterns which is a very complex phenomenon [8, 9]. Specifically, to find the critical value for an observed pattern, a large number of CSR patterns will be generated, and the distribution of metric values for these simulated patterns will be used as the null distribution. To evaluate the detection power, a large number of

nonuniform patterns will be generated, and the percentage of these patterns that are correctly determined to be nonuniform will be used as estimate of the detection power.

3.2. Factors affecting performance of uniformity metrics

In general, the performance of uniformity metrics is affected by two sets of factors:

(1) *Characteristics of observed patterns*. These include the number of points in the pattern, type of nonuniformity, and degree of nonuniformity. Patterns with various characteristics may be produced in manufacturing processes. For example, the number of particles may vary a lot from process to process, and from image to image; a nonuniform pattern may contain a single cluster, multiple clusters or line-shape clusters; and different degrees of nonuniformity may be resulted, from very serious to weak. Accordingly, the performance of metrics should be evaluated for typical patterns that exist in the practice of manufacturing processes.

(2) *Parameters of the metrics*. These include the parameters in calculating the metrics, i.e., the number of quadrats, q , for quadrat-based metrics and the maximal distance allowed, D , for distance-based metrics, and the critical value for each observed pattern in decision making. There are two methods to determine the critical value, depending on how CSR patterns are generated in finding the null distribution:

- “*fixed λ* ” method: the CSR patterns will be generated from a homogeneous Poisson distribution with intensity $\lambda=n/W$, where n is the number of points in the observed pattern, and W is the area of the study region. This method is commonly used in testing spatial uniformity, but it may not work well for some metrics when the number of points in the observed pattern, i.e., n , is very small. The reason is that under a small n and thus λ , a pattern generated from the homogeneous Poisson distribution may contain a very small number of points, e.g., below 10, which makes the calculated metrics meaningless or impossible. For example, if a CSR pattern only contains, say, 3 points, the counts of most grids in Figure 3(a) will be zero, and thus the metrics based on the empirical distribution of these counts does not make much sense.
- “*fixed n* ” method: the CSR patterns will be generated from a Binomial distribution with n points. Intuitively, this is equivalent to comparing the observed pattern with what it should be if the n points in it are arranged uniformly. Since this method does not involve the generation of numbers following a Poisson distribution, it is easier to implement and does not have the issue when n is very small.

4. Numerical Study

A comprehensive numerical study is done to evaluate and compare the performance of metrics described in Section 2.2. To show the effects of factors mentioned in Section 3.2, different cases of the observed patterns and settings of metric parameters are considered. The scenario design and computation

procedure in the simulations will be described in Section 4.1, specific concerns to address in this study will be given in Section 4.2, and results will be presented in Section 4.3 and summarized in Section 4.4.

4.1. Scenario design and computation procedure

Table 3 displays the scenarios considered in the simulations. The area of patterns is set to be 1×1 for convenience. The specific setting of the two sets of affecting factors is as follows:

(1) *Characteristics of observed patterns*: 2 levels of average intensity, 100 and 30, are considered, representing the cases with a large and small number of points in each pattern; 4 types of nonuniform patterns are considered, as shown in Figure 4, including the single-cluster patterns with one single cluster in the center, Matern-cluster patterns with multiple clusters, Type I line patterns with three line-shape clusters, and Type II line patterns with multiple line-shape clusters. These patterns are popular in manufacturing practices. For example, the line patterns are a typical phenomenon in nanocomposite fabrication due to the intrinsic boundary effect of nanocomposites [3], while the multi-cluster patterns are common in scaffold fabrication. For each type of nonuniform patterns, 5 levels of the degree of nonuniformity are considered, with a decreasing trend from L5 to L1. Definition of these levels for each type of patterns can be found in Appendix B.

(2) *Parameters of the metrics*: 5 different settings are considered for the number of quadrats q , while 4 settings are considered for the maximum distance D . The two methods, “fixed λ ” and “fixed n ”, are used to determine the critical values.

Table 3. Scenarios considered in the numerical study

Characteristics of observed patterns	
(1) <i>Average intensity</i> :	High (100)/Low (30)
(2) <i>Type of nonuniformity</i> :	Single/Matern/Line I/Line II
(3) <i>Degree of nonuniformity</i> :	L5/L4/L3/L2/L1 (serious \rightarrow weak)
Parameters of metrics	
(4) <i>Parameters in calculation</i> :	$q=9/16/25/36/49$ (quadrat methods) $D=0.04/0.06/0.08/0.1$ (distance methods)
(5) <i>Critical value</i> :	Fixed λ /Fixed n

In the study, the detection power of each metric under each scenario in Table 3 is obtained through the following steps:

Step 1: Generate 10000 patterns under this scenario. The detailed procedure to generate the patterns is given in Appendix B. Examples of the generated patterns are shown in Figure 4.

Step 2: For each pattern, determine the critical value by the chosen method with type I error probability $\alpha=0.05$. Then calculate the value of the metric, and determine if the pattern is uniform by the decision rules listed in Table 1 and 2.

Step 3. The detection power is calculated by $w/10000$, where w is the number of patterns among the 10000 simulated patterns which are correctly determined to be nonuniform.

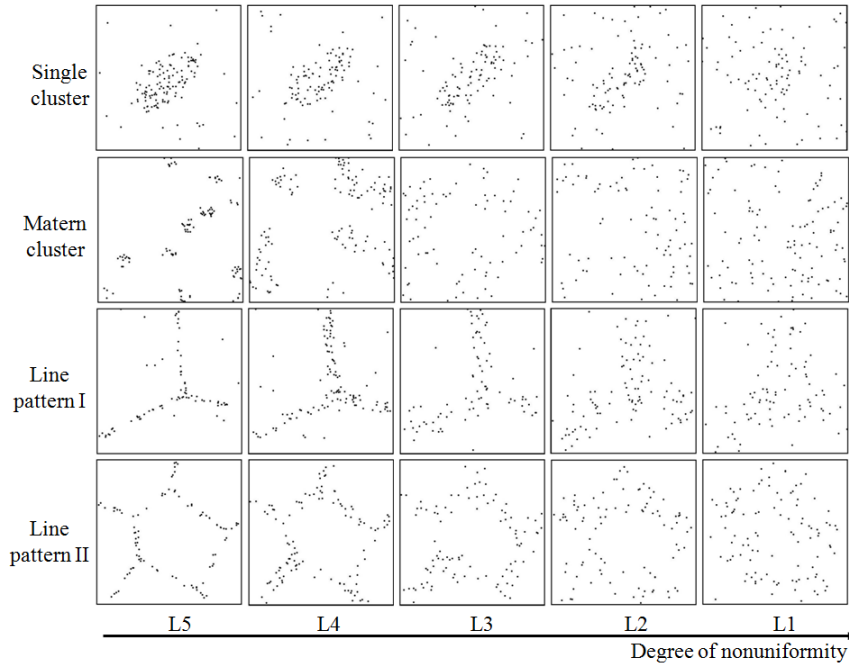


Fig. 4. Examples of nonuniform patterns generated in the simulations

4.2. Concerns to address

The obtained detection powers will be used to address the following concerns

(i) *What is the performance of each metric in uniformity monitoring, and which metric(s) is the best?*

Since the performance of the metrics depends on the value of their parameters q or D , the average and best detection powers over different settings of q or D will be used in the comparison.

(ii) *How robust is each metric to the specification of parameters, and which metric(s) is most robust?*

Such robustness is desirable in practice to provide reliable assessment of uniformity that is not significantly affected by the specification of parameters.

(iii) *What are the effects of the characteristics of observed patterns, i.e., the number of points, and the type/degree of nonuniformity, on the performance of these metrics?* Information on this will help practitioners to find metrics fitting their process.

(iv) *What are the effects of the parameters of the metrics on their performance?* In other words, what is the appropriate specification of q for quadrat-based metrics and D for distance-based metrics? Can the “fixed n ” method be used instead of the “fixed λ ” method in determining the critical values?

4.3. Results

Average and best performance over different settings of parameters

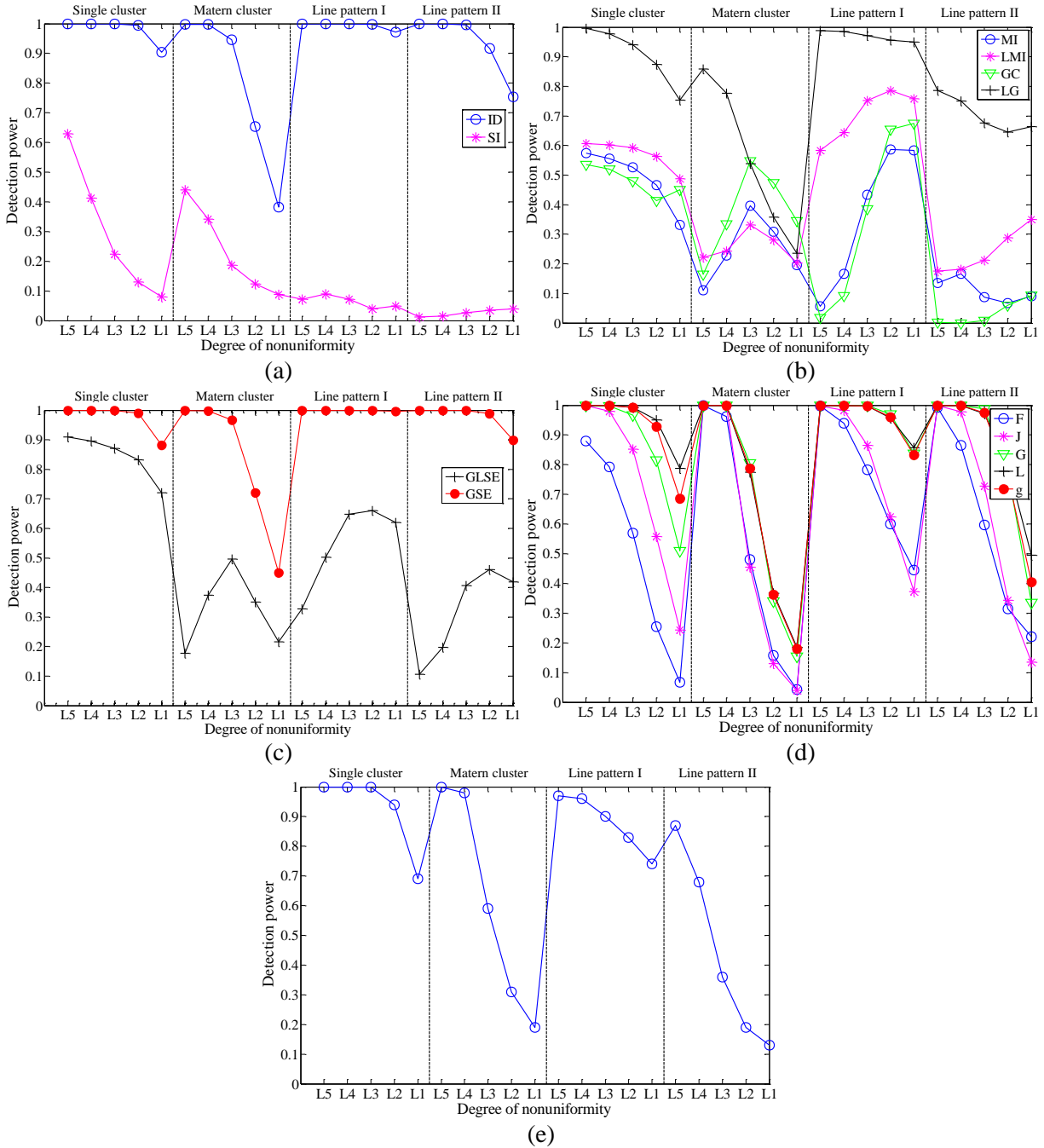


Fig. 5. Average performance of quadrat-based metrics (Group 1: (a); Group 2: (b); Group 3: (c)), distance-based metrics (d) and projection index (e)

Figure 5 displays the average detection powers of the metrics under each combination of type/degree of nonuniformity. Note that each point in Fig. 5(a)~(c) is the average over the 5 settings of q , while that in Fig. 5(d) is the average over the 4 settings of D . The projection index in Fig. 5(e) is parameter free, so the

points are not averages. The critical values used in obtaining these results are determined by the “fixed λ ” method, and the average intensity of the observed patterns is 100. Similar characteristics exist in the results obtained using the “fixed n ” method and patterns with intensity 30.

The performance of quadrat-based metrics is shown in Figure 5(a)~(c). In Group 1, the index of dispersion performs significantly better than the skewness index under all cases. In Group 2, the two local metrics, LMI and LG, exhibit better performance in most cases, and LG is better than LMI. In Group 3, the global entropy measure works much better than the global-local entropy measure. Figure 5 (d) shows the performance of distance-based metrics, where the L index works the best in all cases. The performance of the projection index shown in Fig. 5(e) is comparable to the L index, but not so excellent as ID in Fig. 5(a) and GSE in Fig. 5(c).

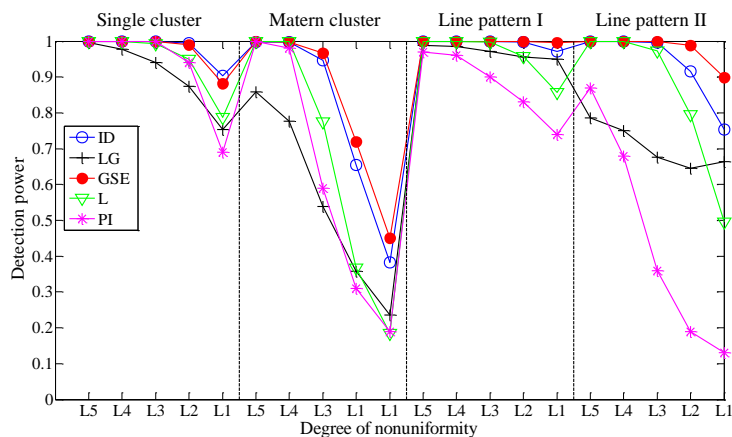


Fig. 6. Average performance of the best metrics in each group (intensity=100)

To see differences among the best metrics in each group (i.e., ID, LG, GSE, L and PI), their average detection powers are compared in Figure 6. It is clear that ID and GSE outperform other metrics uniformly, with 100% detection power in most cases; the L index outperforms the other two in most cases; and LG and PI are better than each other in some cases. Between the two best ones, GSE always outperforms ID, except bearing a slightly worse but still excellent enough detection power for weakest single-cluster patterns. Overall, we can conclude that GSE is the best in terms of average performance.

Figure 7 shows the best detection powers of the metrics over different parameter settings. Note that the settings of q or D that achieved the best performance might be different under different scenarios. For example, for the local Moran’s I index in Figure 7(b), $q=49$ is the best setting for the cluster patterns, while $q=9$ is that for the line patterns. Again, the conclusion is that GSE is the best metric.

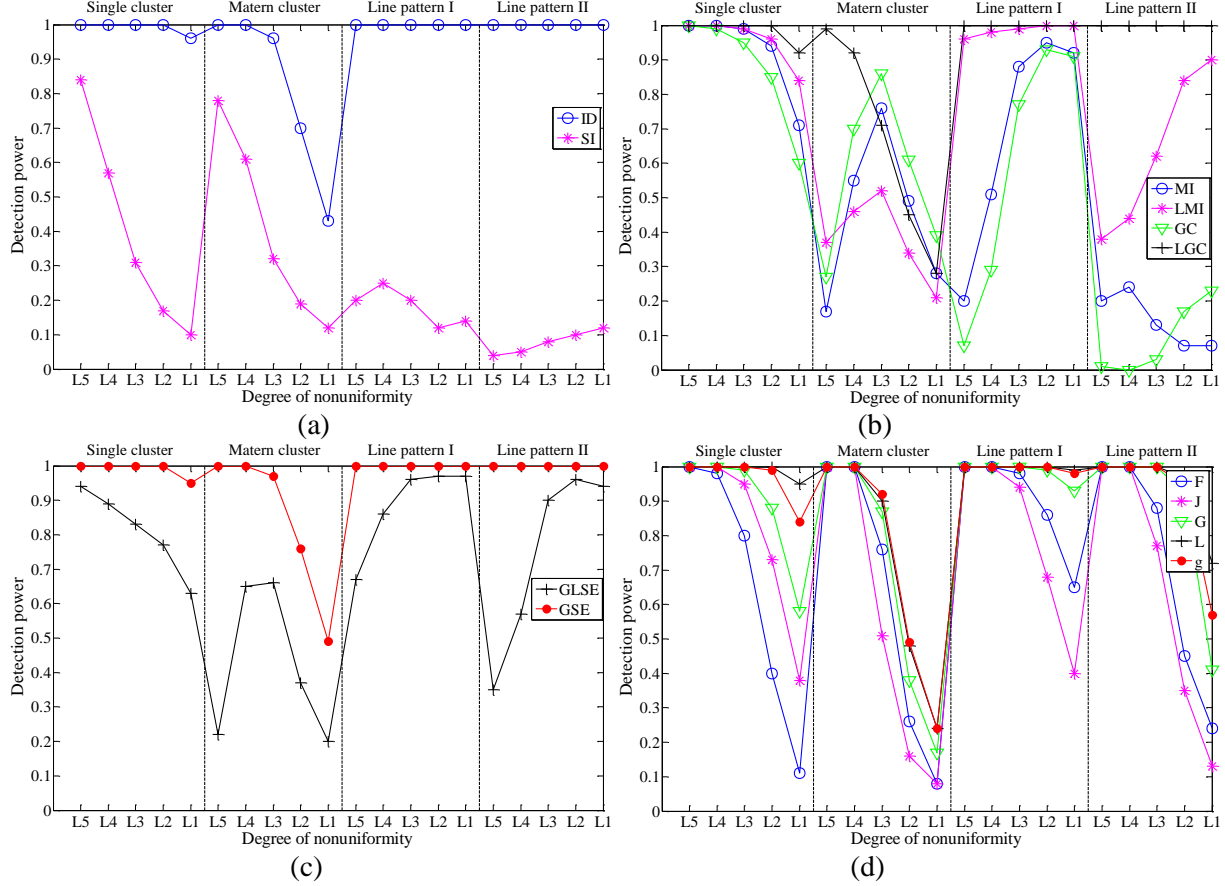


Fig. 7. Best performance of quadrat-based metrics (Group 1: (a); Group 2: (b); Group 3: (c)), distance-based metrics (d) and projection index (e)

Robustness to parameter specification

The robustness of a metric to parameter specification, i.e., the setting of q for quadrat-based metrics and that of D for distance-based metrics, is inversely indicated by the variability of the detection power of this metric over different settings of the parameters. As there are only 5 settings of q and 4 settings of D considered in the study, the range, i.e., difference between the largest and smallest detection power over the settings, will be used as the measure of robustness. Values of this measure are shown in Figure 8. Note that the robustness issue does not exist for the parameter-free projection index.

From Figure 8(a), we can see that ID is more robust than SI in that its detection power varies less under different settings of q in most cases. In Group 2 of quadrat-based metrics, the performance of all the metrics varies dramatically under different parameter settings, and the range of detection power of some metrics can be as large as close to 1 in some cases, meaning that these metrics can detect nonuniform patterns 100% surely when q is appropriately specified, but not so at all under other settings of q . In other words, they are very sensitive to parameter specification. In Group 3, the global entropy measure shows a much smaller variability and thus a much higher robustness than the global-local entropy measure. The robustness of GSE is even better than ID in Group 1. In Fig. 8(d), the G index and J index are the ones

with the best robustness to the specification of D among distance-based metrics, but in most scenarios, they are not so robust as GSE. Overall, GSE is the most robust metric to parameter specification.

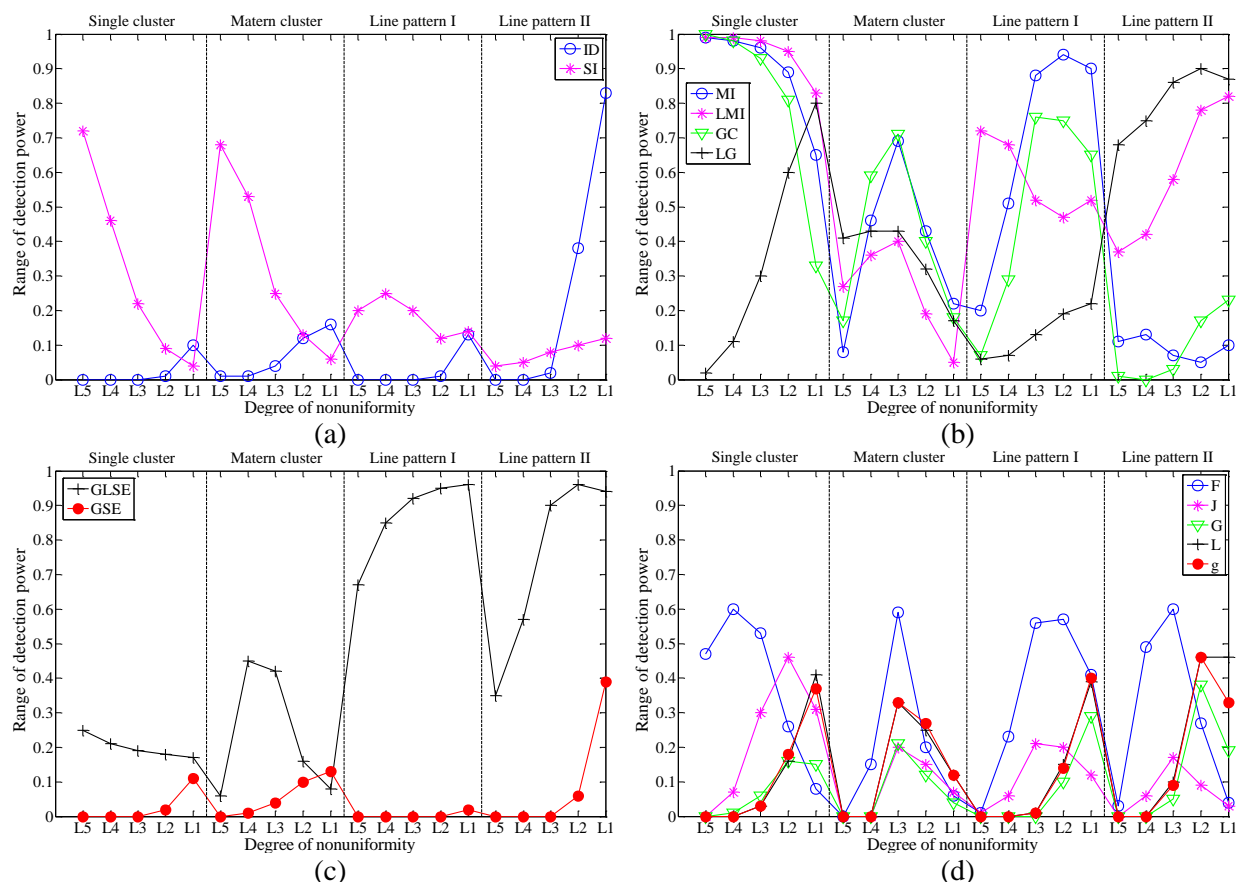


Fig. 8. Robustness of quadrat-based metrics (Group 1: (a); Group 2: (b); Group 3: (c)), and distance-based metrics (d)

Effects of characteristics of observed patterns

The effects of the characteristics of observed patterns on the performance of these metrics can be seen from Figure 5~7. To make the features clearer, the average performance of the best metrics in each group when the intensity of the patterns is 30 is given in Figure 9, which is the counterpart of Figure 6 where the intensity is 100. The following are the findings from the results in Fig. 6 and 9:

- (1) *Effect of the number of points:* For patterns with a smaller number of points, the detection power of all the metrics decreases. This is expected as clustering will not be manifested by a smaller number of points so clearly as by a large number of points, and thus is likely to be missed in the monitoring. Comparatively, GSE is still the best one in terms of average performance in this case.
- (2) *Effect of type of nonuniformity:* Matern-cluster patterns and Type II line patterns are more difficult to detect than the single-cluster patterns and Type I line patterns. Intuitively, this is because the Matern-cluster patterns, which bear multiple clusters, and the Type II line patterns, which contain multiple line-

shape clusters, have a wider spread within the study region and thus are closer to CSR. This can be seen clearly from the plots in Figure 4 where the Matern patterns and Type II line patterns are very similar to CSR especially when the degree of nonuniformity is weak.

(3) *Effect of degree of nonuniformity*: In general, the higher the degree of nonuniformity, the higher the detection power of the metrics. When the degree of nonuniformity is very serious (e.g., at level L5), most of them exhibit satisfactory performance with detection power over 90% even when the number of points is small, while when the degree of nonuniformity is very weak (i.e., at level L1), even the most powerful ones, e.g., GSE, do not perform very well. In addition, there are several metrics which show opposite trends under some scenarios, that is, their detection power may increase as the degree of nonuniformity decreases. For example, in Fig. 5(b), the power of MI, LMI and GC shows such a trend for Type I line patterns. A possible explanation is that when the points form tight lines (i.e., at level L5), the correlation among spatial neighbors is actually weak because some neighbors may have no points falling in it, and thus the correlation measure will be low.

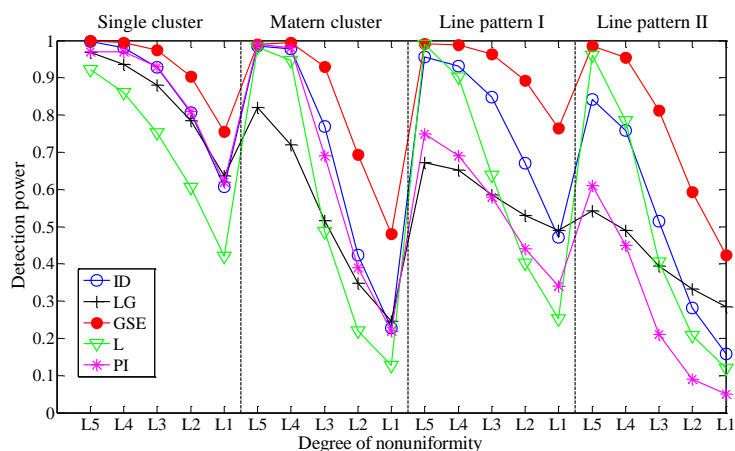


Fig. 9. Average performance of the best metrics in each group (intensity=30)

Effects of parameters of metrics

(1) *Effect of q and D* : Figure 8 implies that the performance of most metrics depends substantially on the value of their parameters, i.e., the number of quadrats, q , for quadrat-based metrics, and the maximum distance, D , for distance-based metrics. Thus, these parameters need to be chosen appropriately to achieve good performance. Based on the simulation results, it is found that the appropriate choices of q and D vary among metrics and scenarios. Table 4 gives the settings of q or D that lead to the highest detection power for the best metrics in each group. Note that when more than one setting of the parameters yield the highest detection power, the smallest setting is reported in the table.

Table 4. Appropriate settings of parameters (q and D) for the best metrics in each group

Name	Intensity	Single cluster					Matern cluster					Line pattern I					Line pattern II				
		L5	L4	L3	L2	L1	L5	L4	L3	L2	L1	L5	L4	L3	L2	L1	L5	L4	L3	L2	L1
ID	100	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
	30	9	9	9	9	9	9	9	9	9	9	16	16	16	16	16	25	25	25	25	25
LG	100	49	49	49	49	49	49	49	49	49	49	9	9	9	9	9	9	9	9	9	9
	30	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	9	9	9	9	9
GSE	100	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
	30	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	25	25	25	25	25
L	100	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	30	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

The results in Table 4 suggest that for ID and GSE, a small number of quadrats, such as 9, is adequate to yield excellent performance when the number of points in the pattern is large, and a little larger value, such as 16, should be chosen when the number of points is small. In fact, considering the robustness of these two metrics as indicated in Fig. 8(a) and (c), the setting of q will not have significant impact on their performance. For LG, the appropriate choice of q varies dramatically for different types of patterns. Since the type of an observed pattern is unknown in quality monitoring, there will be some difficulty in specifying the value of q in using LG. For the L index and g index, a large value of D should always be chosen, which provides convenience in using these two metrics in practice.

(2) *Effect of methods to determine critical value:* It is found that the metrics form three groups in terms of their performance under the two methods to determine critical values: For ID, SI, GSE, GLSE, LMI and LG, the “fixed λ ” method always yields a higher detection power, especially when the number of points in the observed patterns is small. As an example, Figure 10(a) shows the detection powers of GSE (intensity of patterns=30, single-cluster patterns, $q=25$) under these two methods. For the F index and J index, the situation is opposite. Figure 10(b) shows the detection powers of F index under the two methods. For the remaining, their performance under the “fixed λ ” method and that under the “fixed n ” method are similar.

To find more details on the difference of metrics in the first two groups, their critical values for different values of n are obtained, as shown in Fig. 10(c) and (d). According to Fig. 10(c), the critical values of GSE determined by the two methods bear no considerable difference when the number of points in a pattern is large (>50), whereas the critical values resulted from the “fixed λ ” method are higher than those from the “fixed n ” method when the number of points is small, thus leading to a higher detection power as shown in Fig. 10(a) by the lower-sided decision rule used for GSE. For the F index, the critical values resulted from the “fixed λ ” method are always lower than those from the “fixed n ” method, leading to a uniformly lower detection power as shown in Fig. 10(b).

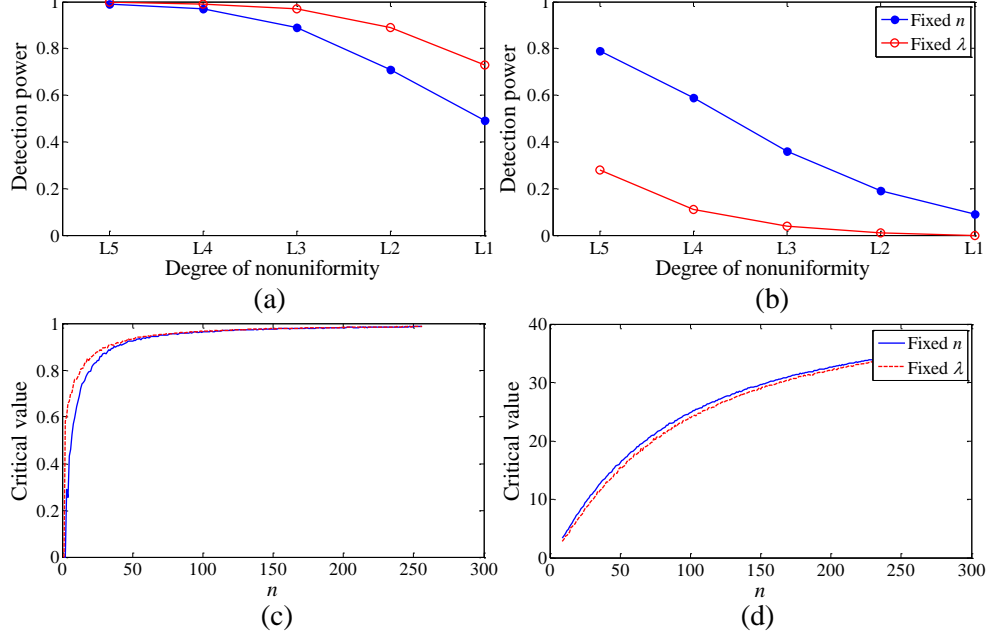


Fig. 10. Comparison of the two methods for determining critical values: detection power (a) and critical values (c) of GSE, and detection power (b) and critical values (d) of F index

4.4. Summary of results and general guidelines

The results of the numerical study provide the following answers to the concerns listed in Section 4.2:

(i) The global Shannon entropy (GSE) measure is the best metric in terms of the average and best performance over different parameter settings.

(ii) GSE is also the most robust metric to the specification of parameters.

(iii) The performance of these metrics is significantly affected by the characteristics of the observed patterns. In general, a pattern with a larger number of points will be detected with a higher probability; Multiple-cluster patterns are more difficult to detect than single-cluster patterns; a pattern with a higher degree of nonuniformity will be captured more easily.

(iv) The performance of these metrics is also significantly affected by the specification of parameters. In general, appropriate specifications vary from one metric to another and depend on the characteristics of the observed patterns as well. The exceptions include ID and GSE whose performance is quite stable in most cases, and the L index and g index for which a large value of the parameter, D , is always preferred. The two methods to determine critical values, “fixed λ ” and “fixed n ”, lead to similar performance for some metrics, and different performance for others. For ID, SI, GSE, GLSE, LMI and LG, the two methods yield similar performance when the number of points in the observed pattern is relatively large; otherwise the “fixed λ ” method will produce a higher detection power. For the F index and J index, the “fixed n ” method always yields a better performance.

Considering all the above aspects, ***the global Shannon entropy measure will be recommended for uniformity monitoring***. This metric performs the best or among the best for all types of patterns, and is most robust to parameter specification. Such good features are not surprising since it is well known that entropic measures are powerful in characterizing information contained in random variables. However, several new findings regarding their performance are obtained in our study: the global Shannon entropy measure works better than the global-local entropy measure in uniformity monitoring; the global entropy measure, which utilizes only count information, performs better than the distance-based metrics which take the location information into consideration; and the performance of the entropy measure is not sensitive to the specification of the number of grids. These findings will add to the general understanding of entropy measures.

To achieve the best performance when using the global Shannon entropy measure in uniformity monitoring, the following general guidelines are provided on the specification of its parameters: (1) When the number of points in the observed pattern is large, a small value can be specified for q ; when the number of points is small, choose a relatively large value for q . (2) When the number of points is large, the “fixed n ” method can be used to determine the critical value since it is simpler than the “fixed λ ” method while yields similar performance; when the number of points is small, the “fixed λ ” method should be used; when the number of points is extremely small and the “fixed λ ” method fails, the “fixed n ” method can be used as a substitute.

5. Case Study

To demonstrate the use and effectiveness of the metrics on uniformity monitoring, they are applied to two sets of image data from manufacturing processes. The first set, as shown in the upper panel of Figure 11(a) and (b), is the SEM images from an ultrasonic-based aluminum nanocomposite fabrication process. Details of this process can be found in Yang *et al.* [27], and Yang and Li [28]. This process uses ultrasonic waves to distribute nanoparticles in order to achieve uniform particle distribution, which is a novel way for MMNC fabrication. The nanocomposite matrix used in this study was made of 99.8% (weight) aluminum and 0.2% titanium. To the alloy, 1% aluminum-oxide nanoparticles was added. These images have been used in the study of Zeng *et al.* [3] for assessing the boundary effect of the nanoparticles. The second set, as shown in the upper panel of Figure 11 (c) and (d), is the SEM images from a tissue-engineered scaffolds fabrication process based on urethaned-doped polyester elastomers (CUPEs). Details of this process can be found in Dey *et al.* [29], and Yang *et al.* [30]. The CUPEs is a novel class of biomaterial which has excellent materials properties to be used for fabricating scaffolds of soft tissues such as blood vessels and cardiac tissues. The white dots in the images are NIH 3T3 fibroblast cells in the scaffold.

Before applying the metrics, some image preprocessing procedures are applied to grayscale SEM images to obtain point patterns as shown in the lower panel of Figure 11. The preprocessing procedures consist of three main steps: (1) converting grayscale images to binary images, (2) morphological operations for particle separation, and (3) identifying mass centers of particles. These procedures have been extensively studied and numerous well-established methods are readily available in the area of computer image analysis. After preprocessing the SEM images, the particles in the images are denoted by their mass centers, and then they can be modeled as spatial point patterns. The calculated values of the global Shannon entropy measure are given in Table 5, where the critical value, C , of Image (a), (b) and (c) is determined by the “fixed λ ” method with $\alpha=0.05$, and that of Image (d) is determined by the “fixed n ” method due to the small number of points in that image. A lower-sided decision rule is used here, which means that values of this metric smaller than the critical value will be concluded to be nonuniform.

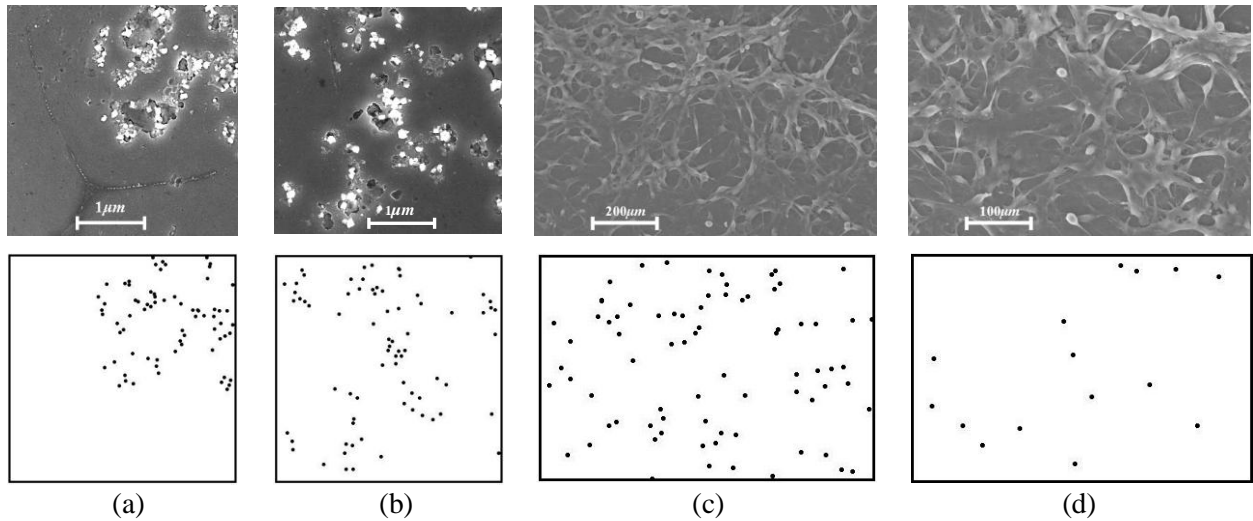


Fig. 11. SEM images in nanocomposite fabrication ((a), (b)) and scaffold fabrication ((c), (d))

Table 5. Calculated values of the global Shannon entropy measure

	n	$q=9$		$q=16$		$q=25$		$q=36$		$q=49$	
		C	GSE	C	GSE	C	GSE	C	GSE	C	GSE
(a)	72	0.95	0.61	0.94	0.70	0.92	0.67	0.90	0.72	0.88	0.70
(b)	75	0.96	0.92	0.94	0.88	0.92	0.87	0.90	0.86	0.88	0.82
(c)	77	0.79	0.98	0.74	0.95	0.69	0.94	0.65	0.91	0.62	0.91
(d)	15	0.75	0.88	0.70	0.78	0.66	0.74	0.62	0.70	0.59	0.65

Since Image (a)~(c) contains a relatively large number of points, results under $q=9$ is reliable enough. Image (d) contains a small number of points, so results under $q=36$ and 49 will be used. Between the images from the nanocomposite fabrication process, (a) represents a case of serious nonuniformity, while (b) represents a case of weak nonuniformity. Consistent with this, the corresponding GSE value of (a) is

considerably smaller than the critical value, and that of (b) is slightly smaller than the critical value. Between the images from the scaffold fabrication process, (c) is quite uniform, and this is validated by the corresponding GSE value which is much larger than the critical value. Image (d) exhibits weak nonuniformity, but the corresponding value of GSE is a little larger than the critical value, meaning that the nonuniformity is not captured. According to the numerical study, this is understandable since the number of points in the pattern is so small.

6. Summary

The need for assessing spatial uniformity of particle distributions widely exists in quality control of many manufacturing processes. The existing uniformity metrics that have been developed in many other areas can be applied for this purpose. To provide guidelines on the use of these metrics in uniformity monitoring, this study compares the performance of existing metrics in detecting nonuniform particle distributions under different practical scenarios through simulation. It is found that the performance of these metrics depends significantly on the characteristics of the observed patterns as well as the parameters of these metrics. The global Shannon entropy measure performs the best under all scenarios in terms of its average and best detection power over different parameter settings and robustness to parameter specification. Guidelines on the use of this metric are also provided. A case study is done in which this metric is applied to image data from two emerging manufacturing processes to demonstrate its use and effectiveness.

Acknowledgment

The authors gratefully appreciate Professor Xiaochun Li in the University of Wisconsin-Madison, Madison, WI, for the access of the SEM images from the metal matrix nanocomposite fabrication process.

References

- [1] Akio K, Atsushi O, Toshiro K, Hiroyuki T. Fabrication process of metal matrix composite with nano-size SiC particle produced by vortex method. *Journal of Japanese Institute of Light Metals* 1999; **49**(4):149-154.
- [2] Mussert KM, Vellinga WP, Bakker A. A nano-indentation study on the mechanical behavior of the matrix material in an AA6061-Al₂O₃ MMC. *Journal of Materials Science* 2002; **37**(4):789-794.
- [3] Zeng L, Zhou Q, De Cicco M, Li X, Zhou S. Quantifying boundary effect of nanoparticles in metal matrix nanocomposite fabrication processes. *IIE Transactions* 2011, in press.
- [4] Saltzman WM. *Tissue Engineering: Principles for the Design of Replacement Organs and Tissues*. Oxford University Press: New York, 2004.
- [5] Ma PX, Elisseeff J. *Scaffolding in Tissue Engineering*. Taylor & Francis: Boca Baton, FL, 2006.

- [6] Laurencin CT, Nair LS. *Nanotechnology and Tissue Engineering: the Scaffold*. Taylor & Francis: Boca Baton, FL, 2008.
- [7] Cressie NAC. *Statistics for Spatial Data. Revised edition*. John Wiley & Sons: New York, 1993.
- [8] Diggle PJ. *Statistical Analysis of Spatial Point Patterns* (2nd edition). Oxford University Press: New York, 2003.
- [9] Illian J, Penttinen A, Stoyan H, Stoyan D. *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley & Sons: Chichester, England, 2008.
- [10] Zhou Q, Zeng L, De Cicco M, Li X, Zhou S. A comparative study on clustering indices for distribution of nanoparticles in metal matrix nanocomposites. *proceeding of the 44th CIRP International Conference on Manufacturing Systems*, June, 2011, Madison, WI.
- [11] Fiser RA, Thornton HG, Mackenzie WA. The accuracy of the plating methods of estimating the density of bacterial populations. *Annals of Applied Biology* 1922; **9**:325-359.
- [12] David FN, Moore PG. Notes on contagious distributions in plant populations. *Annals of Botany* 1954; **18**:47-53.
- [13] Morisita M. Measuring of the dispersion and analysis of distribution patterns. *Memoires of the Faculty of Science, Kyushu University, Series E. Biology* 1959; **2**:215-235.
- [14] Lloyd M. Mean crowding. *Journal of Animal Ecology* 1967; **36**:1-30.
- [15] Douglas JB. Clustering and aggregation. *Sankhya B* 1975; **37**:398-417.
- [16] Alemaskin K, KManas-Zloczower I, Kaufman M. Simultaneous characterization of dispersive and distributive mixing in a single screw extruder. *ANTEC 2003 Conference Proceedings*; 268-272.
- [17] Greig-Smith P. The use of random and contiguous quadrats in the study of the structure of plant communities. *Annals of Botany* 1952; **16**:293-316.
- [18] Jun C-H, Hong Y, Kim SY, Park K-S, Park H. A simulation-based semiconductor chip yield model incorporating a new defect cluster index. *Microelectronics Reliability* 1999; **39**(4):451-456.
- [19] Tong L-I, Wang C-H, Chen D-L. Development of a new cluster index for wafer defects. *International Journal of Manufacturing Technology* 2007; **31**:705-715.
- [20] Perry JN, Hewitt M. A New index of aggregation for animal counts. *Biometrics* 1991; **47**(4):1505-1518.
- [21] Perry JN. Measures of spatial pattern for counts. *Ecology* 1998; **79**:1008-1017.
- [22] Schabenberger O, Gotway CA. *Statistical Methods for Spatial Data Analysis*. Chapman & Hall: Boca Raton, FL, 2005.
- [23] Getis A, Ord JK. The analysis of spatial association by use of distance statistics. *Geographical analysis* 1992; **24**(3):189-206.

- [24] Camesasca M, Kaufman M, Manas-Zloczower I. Quantifying fluid mixing with the Shannon entropy. *Macromolecular Theory and Simulations* 2006; **15**(8):595-607.
- [25] Alemaskin K, Camesasca M, Manas-Zloczower I, Kaufman M. Entropic measures of mixing tailored for various applications. *Proceedings of NUMIFORM 2004*; 712: 169-173.
- [26] Karlström A, Ceccato V. A new information theoretical measure of global and local spatial association. *Jahrbuch für Regionalwis-senschaft* 2002; **22**:13-40.
- [27] Yang Y, Lan J, Li X. Study on bulk aluminum matrix nanocomposite fabricated by ultrasonic dispersion of nano-sized SiC particles in molten aluminum alloy. *Materials Science and Engineering A* 2004; **380**(1-2):378-383.
- [28] Yang Y, Li X. Ultrasonic cavitation based nanomanufacturing of bulk aluminum matrix nanocomposites. *Journal of Manufacturing Science and Engineering* 2007; **129**(2):497-501.
- [29] Dey J, Xu H, Shen J, Thevenot P, Gondi SR, Nguyen KT, Sumerlin BS, Tang L, Yang J. Development of biodegradable crosslinked urethane-doped polyester elastomers. *Biomaterials* 2008; **29**:4637-4649.
- [30] Yang J, Webb AR, Ameer GA. Novel citric acid-based biodegradable elastomers for tissue engineering; *Advanced Materials* 2004; **16**:511-516.

Appendices

Appendix A: Procedure to calculate the projection index

This index is developed based on projecting coordinates of points onto a rotating axis as illustrated in Figure A.1, with a coordinate system established at the center of the study region. The value of the projection index will be obtained following the steps below:

Step 1: Rotate the x axis counterclockwise by θ , where $\theta = 0^\circ, 1^\circ, \dots, 179^\circ$.

Project all points (x_i, y_i) in the observed pattern onto the axis and obtain the projected coordinates by

$$x_{i,\theta} = x_i \cos \theta + y_i \sin \theta \quad i = 1, \dots, n$$

Step 2: Denote $v_{i,\theta}$ as the distance between each pair of adjacent points in $x_{i,\theta}$. Calculate the squared coefficient of variation for $v_{i,\theta}$

$$SCV_\theta = (s_{v,\theta} / \mu_{v,\theta})^2$$

where $s_{v,\theta}$ and $\mu_{v,\theta}$ are sample standard deviation and mean of $v_{i,\theta}$, respectively.

Step 3: Calculate all the 180 SCV_θ at different angles and obtain the index as

$$PI = \left(\sum_{\theta=0^\circ}^{179^\circ} SCV_\theta \right) / 180$$

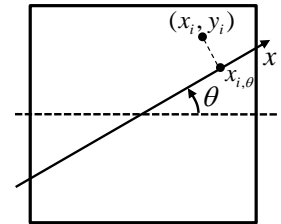


Fig. A.1. Rotating axis

Appendix B: Procedure to generate the patterns in Figure 4

(1) *Single-cluster patterns*: These patterns are generated by the superimposition of two point patterns: a Poisson distribution with a high intensity generated within an ellipse at the center of the study region, and a Poisson distribution with a low intensity generated in the entire study region. Given the average intensity (100 or 30 in the numerical study), $p\%$ of the points are generated randomly within the ellipse in the center, while the remaining $(100-p)\%$ points are generated within the entire region. 5 levels of p are considered in the simulations: 80(L5), 70(L4), 60(L3), 50(L2), 40(L1).

(2) *Matern-cluster patterns*: These patterns are generated through two steps: First, a Poisson distribution of parent points is generated with intensity κ_1 , and then each parent point is replaced by a random cluster of points of intensity κ_2 , within radius r_M of the parent point. In the simulations, $\kappa_1=\kappa_2=10$, and 5 levels of r_M are considered: 0.05(L5), 0.1(L4), 0.2(L3), 0.3(L2), 0.4(L1).

(3) *Line patterns*: These patterns are generated through two steps: First, points following a Poisson distribution are generated on the line segments, and then each point is randomly jittered within a disc of radius r_l centered at its original location. 5 levels of r_l are considered for both the Type I and Type II line patterns: 0.02(L5), 0.05(L4), 0.1(L3), 0.15(L2), 0.2(L1).

Biographies

Kin Ming Kam received his B.S. in Industrial Engineering from the City University of Hong Kong in 2010. He is currently a Ph.D. student in the Department of Industrial and Manufacturing Systems Engineering at the University of Texas at Arlington.

Li Zeng received her B.S. and M.S. degrees in Optical Engineering from Tsinghua University, Beijing, China, in 2002 and 2004, respectively, M.S. degree in Statistics, and Ph.D. degree in Industrial Engineering, both from the University of Wisconsin-Madison, in 2007 and 2010, respectively. She is currently an Assistant Professor in the Department of Industrial and Manufacturing Systems Engineering at the University of Texas at Arlington. Her research interests are statistical modeling and control of complex manufacturing and healthcare delivery systems. She is a member of INFORMS and IIE.

Qiang Zhou received the B.Eng. degree in Vehicle Engineering and the M.Eng. Degree in Mechanical Engineering from Tsinghua University, Beijing, China, in 2005 and 2007, respectively, M.S. degree in Statistics, and Ph.D. degree in Industrial Engineering from the University of Wisconsin-Madison in 2010. He is currently a research associate in the Department of Industrial and Systems Engineering at UW-Madison. He is a member of INFORMS.

Richard Tran received his B.S. in Bioinformatics from Baylor University, and Ph.D. in Bioengineering from the University of Texas at Arlington in 2010. He is currently a postdoctoral research fellow in the Bioengineering Department at UT-Arlington.

Jian Yang is an Associate Professor of Bioengineering at the University of Texas at Arlington. He received his Ph.D. from the Institute of Chemistry at the Chinese Academy of Sciences in Beijing, and has been a postdoctoral research fellow in the Biomedical Engineering Department at Northwestern University. His research interests are polymer synthesis and characterization, cell scaffold fabrication and tissue engineering. His research is sponsored by National Science Foundation, National Institutes of Health, American heart Association, etc. He is a recipient of the CAREER Award from NSF in 2010.