Biom. J. 37 (1995) 4, 393 – 408 Akademie Verlag

# Analysis of Bacterial Spatial Patterns at the Initial Stage of Biofilm Formation

M. A. HAMILTON
Department of Mathematical Sciences

K. R. JOHNSON
Department of Biology

A. K. CAMPER
Department of Civil Engineering

P. STOODLEY
Center for Biofilm Engineering

G. J. HARKIN
Department of Computer Science

R. J. GILLIS
Department of Microbiology and

P. A. SHOPE
Department of Computer Science

all with appointment to the Center for Biofilm Engineering, Montana State University, Bozeman, MT 59717, U.S.A.

*Summary*

Using sophisticated microscopy techniques, we observed the spatial pattern of bacteria colonizing a sterile 316L stainless steel coupon as bulk water containing bacteria flowed across the coupon. The experiments used stainless steel of differing roughness and surface chemistry. The ultimate goal was to identify surface features which influence bacterial adsorption. The immediate statistical goal was to distinguish patterns consistent with complete spatial randomness from patterns showing regularity or aggregation. This goal was accomplished by using modified analyses of distance functions commonly applied in field ecology. The method protected against a potential multiple comparisons problem. For the null value of the distance function, we calculated tolerance envelopes such that the tolerance level was simultaneous for all distances of concern. Computer simulation experiments showed that the nominal level was accurate. The methodology was effective for detecting and describing patterns of colonization known not to be completely spatially random.

*Key words:* Spatial point pattern analysis; Bacterial adsorption; Biofilm; Spatial randomness; Multiple comparisons; Tolerance envelopes; Simulation.

## 1. Introduction

Field ecologists use various statistical techniques for describing the spatial arrangement of individuals in a plant or animal community. Quantitative descriptors of patterns can provide insight into the biological and environmental factors which structure the community (DIGGLE, 1983). Some of the more powerful methods are based on distances between individuals, or between randomly sampled points and individuals. The calculations require a map of the locations of the individuals. This paper shows how statistical distance functions, developed primarily for field ecology research, are applicable to microscale data arising in bacterial adsorption experiments. The map is provided by microscopy and image analysis.

The microscale problem is important to research on bacterial contamination and bacterial biofilms. Surfaces in flowing systems are subject to colonization by bacteria (MATTYSSE, 1992). Adsorbed bacteria are a potential source of contamination for any material which contacts the surface, a consideration of importance in ultrapure water systems such as required for pharmaceutical production and computer chip manufacturing. After bacteria adsorb to the surface, they can multiply, produce exopolymeric material, and eventually become embedded in that gelatinous (slimy) substance. The resulting aggregate is called a biofilm. The consequences of biofilm formation can be either harmful or beneficial. Biofilms can have deleterious effects on industrial process systems; e.g., by creating fouling or corrosion problems in pipelines. Biofilms also cause significant adverse health effects; e.g., they commonly occur on catheters, surgical implants, and other medical devices, and have led to aggressive bacterial infections (LAPPIN-SCOTT, COSTERTON, and MARRIE, 1992). On the other hand, biofilms can have beneficial effects. For example, bacteria may metabolize a toxic compound and reduce it to harmless products such as carbon dioxide and water. By passing the toxic compound through a packed column, where the packing material acts to increase surface area and enhance biofilm formation of selected bacterial species, one can safely dispose of the toxic material (CHARACKLIS and MARSHALL, 1990). Similar strategies are employed in the bioremediation of contaminated water or soil (THOMAS et al., 1992).

This paper focuses on data associated with the initial events in biofilm development. The experimental goal is to determine environmental surface factors that affect the rate and pattern of bacterial adsorption. Greater understanding of these factors could lead to improved control of bacterial contamination and biofilm formation in engineered systems. Decreased adsorption means

slower biofilm formation and subsequent reduced maintenance and replacement costs. On the other hand, enhanced adsorption could have a beneficial effect in a fixed film bioreactor as used, for example, in bioremediation.

Although our statistical approach does not depend on any specific phenomenological model, there are several quantitative theories for bacterial adsorption. In brief, these theories discriminate between two types of bacterial adsorption: (1) physical adsorption and (2) chemisorption or adhesion (CHARACKLIS and MARSHALL, 1990; MARSHALL, 1985). Physical adsorption is a reversible or equilibrium adsorption involving primary physical forces. Chemisorption is generally irreversible and is the result of short-range forces, including chemical bonds, dipole interactions, and hydrophobic bonding.

Some theoretical structures permit the possibility that an attached bacterium either inhibits or enhances the adsorption of others nearby. These two opposite effects have been termed "blocking" and "positive cooperativity," respectively (BUSSCHER, DOORNBUSCH, and VAN DER MEI, 1992; DOYLE, NESBITT, and TAYLOR, 1982). Under blocking, initial adhering bacteria would position themselves in a regular pattern, with few near neighbors. Under cooperativity, initial adhering bacteria would be arrayed in an aggregated pattern with many near neighbors. Some recent research on bacterial adsorption has been directed at measuring these two effects by observing the spatial patterns of surface-colonizing bacteria (SJOLLEMA et al., 1990; SCHMIDT, 1992; QUIRYNEN et al., 1991).

For the study of spatial patterns, we have discovered that statistical Spatial Point Pattern Analyses (SPPA) is especially helpful. To our knowledge, statistical SPPA based on distance functions has never been applied in bacterial surface colonization studies. Researchers may well be unaware of the SPPA statistical literature which is strongly associated with field ecology. Although some investigators have used distance function approaches (SJOLLEMA et al., 1990; ESCHER and CHARACKLIS, 1990), they have not employed conventional statistical reasoning. There seems a need for statistical approaches that can make sense of smaller data sets and that protect against mistaking chance events for meaningful effects. It should be noted that only recently did advances in image analysis make it possible to map the locations of adhering bacteria.

The experiments discussed in this paper were designed to separate effects due to interactions between bacterial cells from those due to interactions between a becterium and the surface. Each experiment used only a single species of bacteria. The initial experiment used a topographically bacterium homogeneous surface; it provides "control" data to which experiments using heterogeneous surfaces can be compared. This approach assumes that colonization of a homogeneous surface represents the inherent biological colonization pattern for that species. By subsequently observing colonization on surfaces of increasing topographic and chemical heterogeneity, one can test for departures from the organism's inherent colonization pattern. When departures are discovered, we assume they are attributable to these surface heterogeneities, in which case, correlating

bacterial adsorption sites with physical/chemical maps of the surface may well uncover the causal factors.

The control (inherent) colonization pattern for many bacterial species is expected to be completely spatially random (CSR, defined below). Statistical SPPA based on distance functions is particularly useful for discriminating CSR from regular or aggregated patterns (RIPLEY, 1981; DIGGLE, 1983; STOYAN, KENDALL, and MECKE, 1987; CRESSIE, 1991). Because it is expensive and time-consuming to investigate surfaces associated with non-CSR patterns, it is important to control the risk of "false significance." We describe below a simultaneous statistical procedure for detecting departures from the expected distance function under CSR. The statistical procedure is an adaptation of conventional simulation methods for placing tolerance envelopes around the expected distance function.

The purposes of this paper are to describe the applicability of statistical SPPA to bacterial colonization data, to present the simultaneous tolerance evenlope method, and to show the results of an evaluation of that method.

## 2. The Experiment

We gathered bacterial colonization data on sterile 316L stainless steel coupons by following the experimental procedure described in CAMPER et al. (1994). We placed a coupon in a flow-through, flat-plate reactor and pumped a bacterial suspension through the reactor. We collected computerized images of one region of the coupon surface at each of several times. This study region consisted of a $6 \times 6$ matrix of contiguous subsquares, each subsquare approximately 50 μm on a side. Pilot experimentation indicated that the optimal magnification for observing adhering bacteria provided a $50 \times 50$ μm field-of-view. The whole study region was thus $300 \times 300$ μm and coverage required 36 microscope images, each $50 \times 50$ μm. We used in-house image analysis software to record the $x$, $y$ position of each bacterium that attached to the surface.

## 3. Spatial Point Pattern Analysis (SPPA)

### (i) Spatial Point Processes

Following standard notation and terminology (CRESSIE, 1991), let $B$ denote the region of interest; e.g., the $300 \times 300$ μm study region or a subsquare within that region. Let $N(B)$ be a random variable denoting the number of bacteria in $B$. Let $\nu(B)$ denote the area of $B$. If the stochastic process governing the positions of bacteria is a homogeneous Poisson process, then the process is said to be Completely Spatially Random (CSR). A CSR process is the "null" model for this project.

Two important first order properties of a spatial process are the mean $\mu(B) = E\{N(B)\}$ and the intensity at point $s \in B$, $\lambda(s) = \mu'(s)$, for which a rigorous definition is provided in CRESSIE (1991). Under CSR, $\lambda(s) = \lambda$ and $\mu(B) = \lambda \cdot v(B)$. In all our applications, the study region $B$ is rectangular and the density of bacteria under CSR is estimated by $\hat{\lambda} = \dfrac{N(B)}{v(B)}$.

The second order properties of a spatial point process are the second moment, $\mu_2(B_1 \times B_2) = E\{N(B_1) \cdot N(B_2)\}$, and the second order intensity at points $s_1$, $s_2$, $\lambda_2(s_1, s_2) = \dfrac{d^2 \mu_2(s_1 \times s_2)}{ds_1 \times ds_2}$ for which a rigorous definition is provided in CRESSIE (1991). For a stationary, isotropic point process, $\lambda_2$ depends only on the distance between $s_1$ and $s_2$; in which case, the distance will be denoted by $r$ and the second order intensity will be denoted by $\lambda_2(r)$.

(ii) The $K$ and $L$ functions for measuring crowding

Consider the well-known $K$ function, $K(r)$, which is a measure of relative crowding; specifically, $K(r) = \lambda^{-1} \cdot E$ {number of extra bacteria within $r$ of an arbitrary bacterium}. The function $K(\cdot)$ captures important characteristics of spatial data (RIPLEY, 1977). It is related to the second order properties of the underlying process; specifically, for a stationary, isotropic point process, $\lambda_2(r) = [\lambda^2/2\pi r]\, K'(r)$. Under CSR, $K(r) = \pi r^2$, and the standardized $K$ function, defined by $L(r) \equiv [K(r)/\pi]^{1/2} - r$, is identically zero. The function $L$ tends to be greater than zero when the true process leads to aggregation, and less than zero when the true process leads to a regular pattern (RIPLEY, 1977; CRESSIE, 1991). Under CSR, the sample estimate of $K(r)$ is subject to Poisson variation, for which the variance stabilizing transformation is the square root. Thus the sample estimate of $L(r)$ should have a homogeneous (in $r$) variance, except for large $r$ where the necessity of correcting for study region edge effects (discussed below) causes an increased variance. For these reasons, we choose to base analyses on $L(\cdot)$ instead of the unstandardized $K(\cdot)$.

To estimate $K(\cdot)$, we follow the edge correction method of RIPLEY (1976). Specifically, the estimator, denoted by $\hat{K}$, was calculated using equation (8.2.21) of CRESSIE (1991). It was calculated for each $r$ in the set $R = \{2, 3, 4, \ldots, R$ in micrometers (μm)}, where $R$ is half the narrowest dimension of $B$. Then $\hat{K}$ was transformed into $\hat{L}$, an estimate of the standardized function $L$. Distances between bacteria less than 2 μm were not of interest because the bacterial diameter could well be 2 μm and two bacteria cannot occupy the same position on the surface.

The top two panels of Figure 1 provide an example of the $L$ function approach. The left-hand panel shows the map of bacteria positions. The right-hand panel shows the realized $\hat{L}(\cdot)$, the null $L$ (that is, expectation under CSR), and three

tolerance envelopes around the null corresponding to tolerance probabilities of 0.90 (inner envelope), 0.95, and 0.99 (outer envelope). Each envelope holds simultaneously for all $r$; it was calculated using an algorithm described in Section 3(iv). Because $\hat{L}$ stays well within the envelopes, we conclude that the data do not discredit the null hypothesis. There is no evidence that the adsorption pattern is either regular or aggregated. This conclusion contradicts our initial visual impression which was that the pattern contained holes and patches. We speculate that the human mind unconsciously "connects the dots," in essence, performing many simultaneous tests for significance. Visual inspection will probably discern at least one significant feature in every adsorption pattern. This experience shows the value of using statistical reasoning when interpreting images. It is especially important to focus on the key hypotheses and to quantify the image so that one can calculate a statistic that discriminates among those hypotheses.

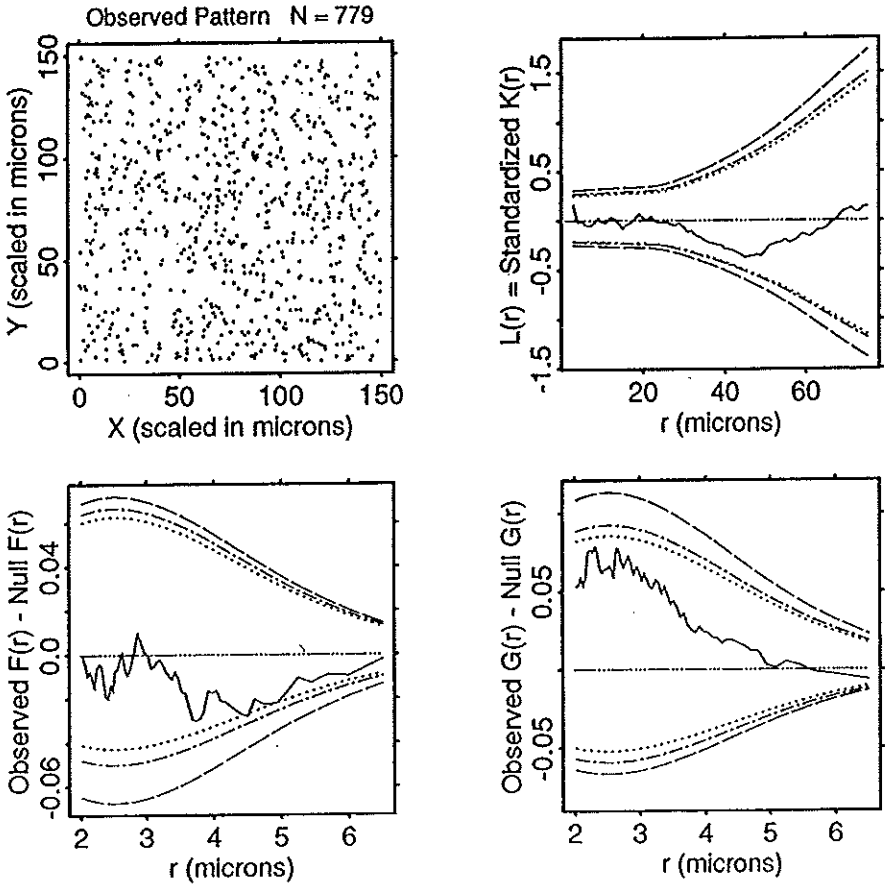


Fig. 1. Pattern of bacterial adsorption and the associated statistical SPPA; *Pseudomonas aeruginosa* on a 316L stainless steel surface at 99 minutes. The steel surface was electropolished and then hand-polished with 0.3 μm alumina grit. Under CSR, the theoretical values of the three plotted functions are identically zero. For each plotted function under CSR, the three tolerance envelopes would completely enclose a realization of that function with probabilities 90%, 95%, and 99%, respectively.

(iii) Conventional tolerance envelope for $L$

As is illustrated in CRESSIE (1991) and DIGGLE (1983), the conventional method for placing an envelope around the null function $L(r) \equiv 0$ is based on simulated spatial patterns. Simulation is required because the sampling distribution of $\hat{L}$ is unknown. Figure 2 shows three conventional envelopes for the data of Figure 1 where 779 adsorbed bacteria were observed. The envelopes correspond to tolerance probabilities of 0.90, 0.95, and 0.99. At each $r$ in $R$, the outer boundary of the 0.99 envelope was formed by the smallest and largest of 200 simulated $\hat{L}(r)$ values. At any chosen distance $r$, the chance that $\hat{L}(r)$ from a CSR process will fall outside that envelope is approximately 1%. The 0.90 and 0.95 envelopes were similarly formed using appropriate quantiles of the simulated $\hat{L}(r)$ values. The conventional simulation envelopes have a ragged appearance due to the inherent variability affecting the extreme quantiles of 200 simulated $\hat{L}$ functions. As is typical, the lower boundary of each envelope in Figure 2 is closer to the null function than is the upper boundary.

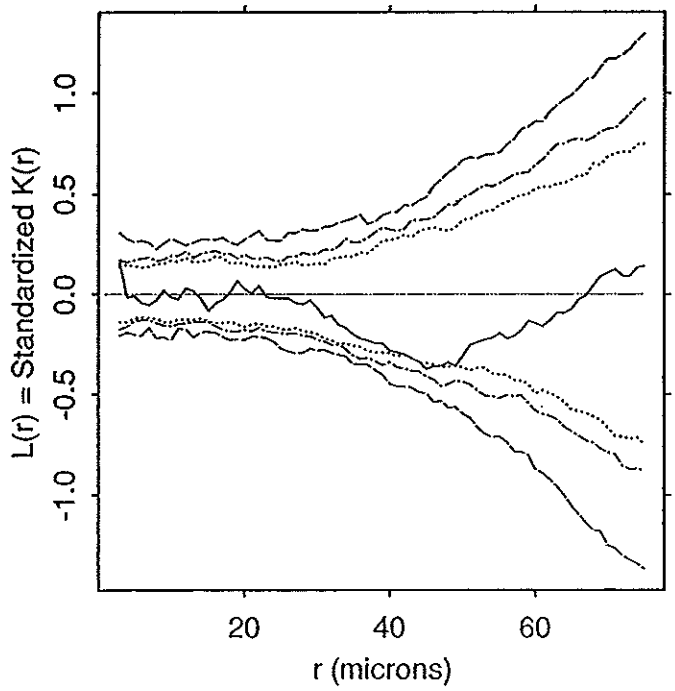


Fig. 2. Same data as in Figure 1 showing 90%, 95%, and 99% conventional simulation bands for the $L$ function under CSR.

(iv) Simultaneous tolerance envelope for $L$

Table 1 shows the steps we used in calculating the new, simultaneous tolerance envelope. Our approach is different from the conventional method in three respects. First, the probability level $1-\alpha$ is simultaneous for all $r$ in $\boldsymbol{R}$. That is, if the null hypothesis were true and the same number of adsorbed bacteria were observed, then with probability $\alpha$, the realized $L$ function will cross the $100(1-\alpha)\%$ envelope one or more times. Simultaneous inference is required because we are using SPPA as a screening technique. The SPPA analysis identifies surfaces on which the colonizing bacteria form a pattern significantly different from the CSR pattern observed on smooth, chemically homogeneous surfaces. We can choose a tolerance probability and, if the observed $\hat{L}$ function crosses the tolerance envelope anywhere, we will expend time and resources attempting to determine why the pattern is not CSR. The conventional envelope is inadequate for this type of application because looking simultaneously at all $r$ in $\boldsymbol{R}$ poses a serious multiple comparisons problem. The proper simultaneous error rate is achieved by basing the tolerance envelope on a maximum deviation statistic similar to the Kolmogorov-Smirnov procedure for placing a tolerance envelope around a cumulative distribution function (MOOD, GRAYBILL, and BOES, 1974). Similar maximum deviation approaches have previously been proposed (RIPLEY, 1988; STOYAN, KENDALL, and MECKE, 1987; KOEN, 1991).

Table 1

Steps in the simulation technique for placing a 95% tolerance envelope around the $L$ function under the null hypotheses of CSR.

1. Observe $N$ events in region $B$.
2. Calculate $\hat{K}(r)$, $r \in R$, which is the observed $K$ function, and the standardized $\hat{K}$ function, $\hat{L}(r) = [\hat{K}(r)/\pi]^{1/2} - r$.
3. Start computer simulation; set $i = 1$.
4. Simulate $N$ cell location events from a uniform distribution on $B$.
5. Calculate $\hat{K}(r)$ for the simulated events; denote the result by $K_i(r)$.
6. If $i = 200$, go to 7; otherwise, set $i = i+1$, then go to 4.
7. Find $L_i(r) = [K_i(r)/\pi]^{1/2} - r$, $i = 1, \ldots, 200$
8. $\hat{V}_L(r) = \mathrm{Var}\{L_i(r), i = 1, \ldots, 200\}$
9. Using LOWESS with span = 20 μm, smooth $\hat{V}_L(r)$. Denote the smooth by $V_{LS}(r)$.
10. $\mathrm{Max}_i = \max\{L_i(r)/[V_{LS}(r)]^{1/2}, r \in R\}$
11. UPPER = 97.5%-tile of $\{\mathrm{Max}_1 \ldots, \mathrm{Max}_{200}\}$
12. $\mathrm{Min}_i = \min\{L_i(r)/[V_{LS}(r)]^{1/2}, r \in R\}$
13. LOWER = 2.5%-tile of $\{\mathrm{Min}_1 \ldots, \mathrm{Min}_{200}\}$
14. Upper Boundary $(r) = [V_{LS}(r)]^{1/2} \cdot$ UPPER
15. Lower Boundary $(r) = [V_{LS}(r)]^{1/2} \cdot$ LOWER
16. $\hat{L}(r)$ departs significantly from CSR if it crosses a boundary at any $r \in R$. (0.05 level of significance).

The second way in which the envelope differs from the conventional is that each boundary is smooth, a feature achieved by applying the LOWESS smoother (CLEVELAND, 1979) to the variance function. Although we know of no formula for the variance of $\hat{L}(r)$ when the process is CSR, we expect the variance to be smooth in $r$. The simulation estimate of the variance, $\hat{V}_L(r)$, is based on only 200 simulated patterns and is relatively rough, having an appearance similar to the envelope in Figure 2. In addition, $\hat{V}_L(r_1)$ and $\hat{V}_L(r_2)$ will be correlated for $r_1$ near $r_2$. There is an opportunity to improve the variance estimate by averaging nearby $\hat{V}_L(r)$ values. For these reasons, the simulation technique smooths $\hat{V}_L(r)$. Experience indicates that the LOWESS smoother works well when the window width is set at 20 μm (10 bacterial diameters).

The third difference is that the upper envelope boundary and lower boundary are calculated separately because of the asymmetry of deviations from the null $L$ illustrated in Figures 1 and 2. Let $V_{LS}(r)$ denote the LOWESS smoothed estimate of the variance of $\hat{L}(r)$ for the observed number of adsorbed bacteria under CSR. The upper boundary depends on the distribution of the statistic Max, the maximum of the standardized deviations; specifically, $\text{Max} = \max\{\hat{L}(r)/[V_{LS}(r)]^{1/2}, r \text{ in } \boldsymbol{R}\}$. Similarly, the lower boundary is based on $\text{Min} = \min\{\hat{L}(r)/[V_{LS}(r)]^{1/2}, r \text{ in } \boldsymbol{R}\}$. The simulation technique described in Table 1 is used to find $V_{LS}(r)$ and to estimate the quantiles of the distributions of Max and Min. A simulation study described below (Section 4) indicates that 200 simulated spatial point patterns provides sufficient accuracy for our application. It is obvious how to adjust the simulation technique to achieve any specified probability level for the envelope. We have chosen to plot 90%, 95%, and 99% envelopes routinely.

(v) The $F$ and $G$ functions for analyzing small scale patterns

The $K$ function is but one of a variety of distance methods suggested in the literature. Two additional functions, denoted by $F(r)$ and $G(r)$, have proven informative in our bacterial colonization research. They provide detailed information about local (small $r$) patterns of bacterial positions. The function $F(r)$ is the cumulative distribution function (cdf) of the distance between a randomly chosen point and the nearest adhering bacterium. It is particularly successful at detecting holes characteristic of aggregated patterns. The function $G(r)$ is the cdf of the distance between a randomly chosen cell and its nearest neighbor. It is particularly successful at detecting an absence of short nearest neighbor distances characteristic of regular patterns. These functions are discussed in CRESSIE (1991) and DIGGLE (1983).

Let $\hat{F}(r)$ and $\hat{G}(r)$ be estimators of $F$ and $G$ based on HANISCH's (1984) proposal for edge correction; specifically, $\hat{F}(r)$ and $\hat{G}(r)$ are calculated using CRESSIE's (1991) equations (8.4.17) and (8.4.16), respectively. The data leading to $\hat{F}$ were created by generating $N(B)$ points from a uniform distribution over

$B$ and, for each random point, finding the distance to the nearest bacterium. For both estimators, a distance to the nearest bacterium entered the calculation only if it was less than the distance to the nearest boundary of the study region. Denote the numbers of distances used in the calculations of $\hat{F}(r)$ and $\hat{G}(r)$ by $M_F$ and $M_G$, respectively, where $M_F \leqq N(B)$ and $M_G \leqq N(B)$.

Let the functions $F_0$ and $G_0$ denote the expectations of $\hat{F}$ and $\hat{G}$ under a CSR process. It is known that $F_0(r) = G_0(r) = 1 - \exp(-\lambda\pi r^2)$. We tabulate $\hat{F}(r)$, $\hat{G}(r)$, $F_0$, and $G_0$ at the 100 values of $r$ in $R_D = \{r\colon\ 1 - \exp(-\hat{\lambda}\pi r^2) = i/101, i = 1, \ldots, 100\}$. Then we plot the difference functions $D_F(r) = \hat{F}(r) - F_0(r)$ and $D_G(r) = \hat{G}(r) = G_0(r)$. Both $D_F(r)$ and $D_G(r)$ are identically zero under CSR.

Table 2

Steps in the simulation technique for placing a 95% tolerance envelope around the $D$ function (either $D_F$ or $D_G$) under the null hypothesis of CSR.

| | |
|---|---|
| 1. Observe $N$ events in region $B$. | 8. $\text{Max}_i = \max\{D_i(r)/[V_D(r)]^{1/2}, r \in R_D\}$ |
| 2. Calculate the Observed minus Null function $D(r)$, $r \in R_D$. | 9. UPPER = 975%-tile of $\{\text{Max}_1 \ldots, \text{Max}_{200}\}$ |
| 3. For each $r \in R_D$, calculate $V_D(r) = [1 - \exp(-\hat{\lambda}\pi r^2)] \exp(-\hat{\lambda}\pi r^2)/M$, where $M$ is the number of nearest-event distances allowed by the edge-correction rule for calculating $D$. | 10. $\text{Min}_i = \min\{D_i(r)/[V_D(r)]^{1/2}, r \in R_D\}$ |
| | 11. LOWER = 2.5%-tile of $\{\text{Min}_1, \ldots, \text{Min}_{200}\}$ |
| 4. Start computer simulation; set $i = 1$. | 12. Upper Boundary $(r) = [V_D(r)]^{1/2} \cdot$ UPPER |
| 5. Simulate $N$ cell location events from a uniform distribution on $B$. | 13. Lower Boundary $(r) = [V_D(r)]^{1/2} \cdot$ LOWER |
| 6. Calculate $D(r)$ for the simulated events; denote the result by $D_i(r)$. | 14. $D(r)$ departs significantly from CSR if it crosses a boundary at any $r \in R_D$ (0.05 level of significance). |
| 7. If $i = 200$, go to 8; otherwise, set $i = i + 1$, then go to 5. | |

Table 2 shows the calculation steps for placing tolerance envelopes around $D_F$ and $D_G$. The method is the same as that for the $L$ function with one exception, the variance estimate. For the remainder of this section suppress the subscripts on $D$ and $M$. Suppose $D$ was based on the empirical distribution function of $M$ independent nearest neighbor distances, then the variance of $D(r)$ under CSR would be $[1 - \exp(-\lambda\pi r^2)] \exp(-\lambda\pi r^2)/M$, which can be estimated by substituting $\hat{\lambda}$ for $\lambda$. We used this estimated variance, denoted by $V_D(r)$, rather than a simulated variance to form the simultaneous tolerance envelope. Note that the rationale for the variance formula is somewhat incorrect in two ways. First, the quantity $M$ is itself a random variable determined by the number of times the distance to the nearest bacterium is smaller than the corresponding distance to the region boundary. Second, the nearest neighbor distances entering $\hat{G}(r)$ are

possibly dependent. For cases we have checked, however, the method produces an envelope with the correct tolerance coefficient (see Section 4). The envelopes are smooth and the associated probabilities are simultaneous over the 100 values of $r$ in $R_D$. Examples of the envelopes for $D_F$ and $D_G$ are shown in the lower two panels of Figure 1 (also Figure 4).

(vi) Examples

Figures 1–4 show examples of $F$, $G$, and $L$ analyses for data from our experiments. The stainless steel surface used in the experiment associated with Figures 1 and 3 was very smooth. There is no evidence that the bacteria adhere in a non-CSR pattern. Figure 3 shows the time sequence of colonization of one 50 μm subsquare, as summarized by the $L$ function. The envelopes become increasingly narrower as the number of adsorbed bacteria increases and the observed $L$ functions stay well within the envelopes at all observation times. Note that a map for an early time is not necessarily a subset of the map for a later time. In the interim some bacteria may detach from the surface while others may "glide" along the surface to new positions.

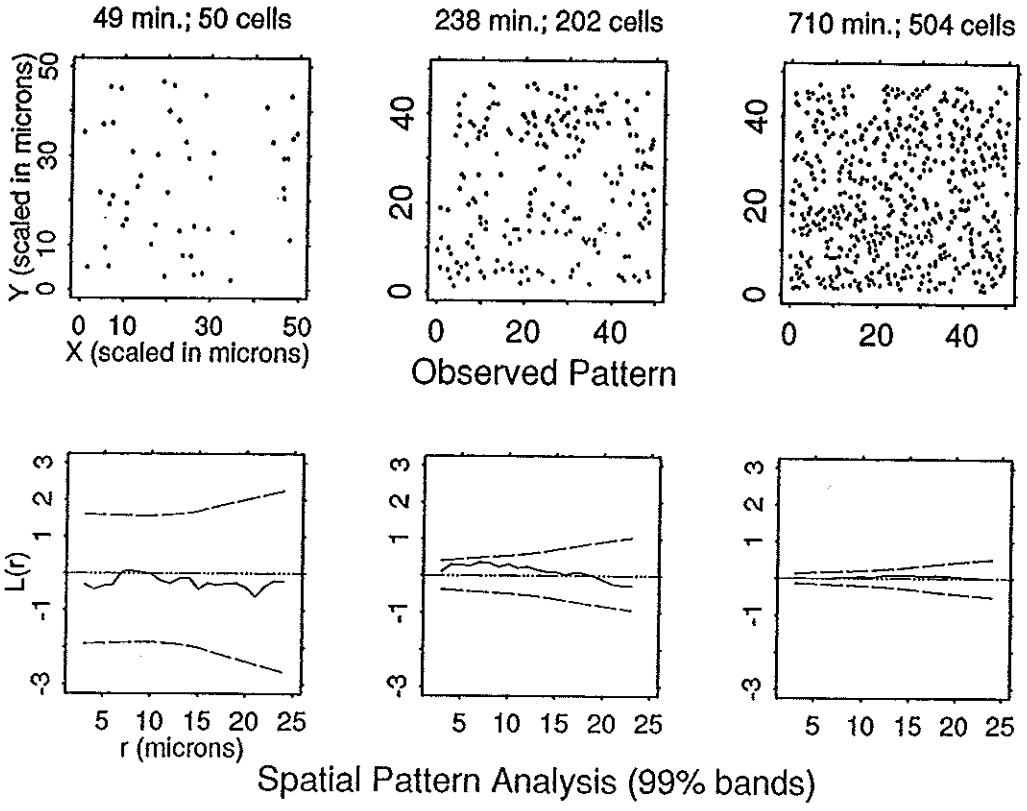


Fig. 3. Time series of adsorption patterns and associated $L$ function analyses; same experiment as for Figure 1. Area of interest is one of the 50 μm × 50 μm squares that make up the area shown in Figure 1.

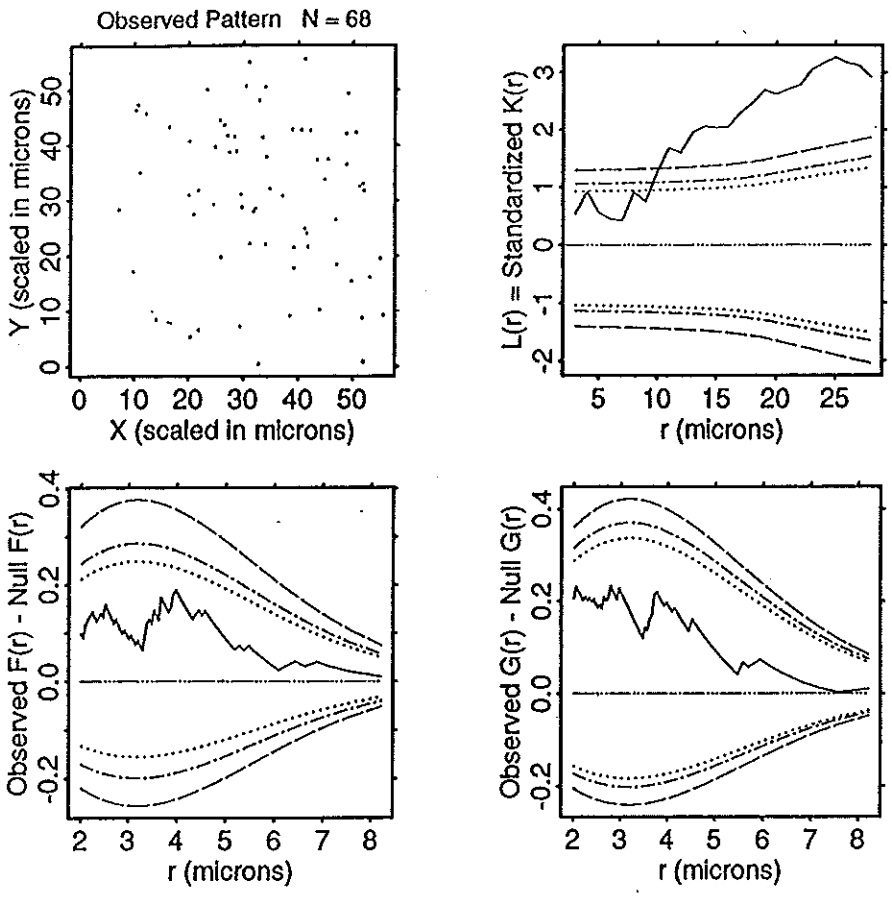


Fig. 4. Pattern of bacterial adsorption and associated statistical SPPA functions with 90%, 95%, and 99% tolerance envelopes; *Citrobacter freundii* on 316L stainless steel at 16 days. The steel surface had an industrial grade 2B finish. Data courtesy of Rick Gillis.

For the experiment associated with Figure 4, it is known that the stainless steel surface is heterogeneous in topography and chemistry. Furthermore, it is known that the bacterium *Citrobacter freundii* adheres in a pattern determined by those heterogeneities (GILLIS, 1993). Figure 4 shows that the $L$ function was successful in detecting the non-CSR pattern.

## 4. Simulation Study of Tolerance Envelope Error Rates

We have conducted two simulation experiments to check the accuracy of the nominal error rates attached to the tolerance envelopes. In the first experiment, we had the computer simulate 1000 CSR patterns, each of $N = 50$ bacteria positions on a $50 \times 50$ μm region ($\lambda = 0.02$). The second was identical, except it simulated $N = 100$ bacteria positions on the region ($\lambda = 0.04$). For each of the simulated patterns, the computer calculated the associated realizations of $\hat{F}(r)$, $\hat{G}(r)$, and $\hat{L}(r)$ along with the tolerance envelopes. For each function and each $\alpha = 0.1$, 0.05, 0.01, the program recorded an "S" for Significant whenever the

estimated function crossed the associated $100(1-\alpha)\%$ tolerance envelope boundary; otherwise, it recorded an "NS" for nonsignificant. The results, which are shown in Table 3, indicate that the nominal error rates are reasonably accurate for the two CSR models simulated.

Table 3
Proportion of significant (S) and Nonsignificant (NS) realizations of $F$, $G$, and $L$ in 1000 simulations of a CSR spatial pattern.

| | | | $N=50$ $\alpha$ | | | $N=100$ $\alpha$ | | |
|---|---|---|---|---|---|---|---|---|
| $F$ | $G$ | $L$ | 0.01 | 0.05 | 0.10 | 0.01 | 0.05 | 0.10 |
| *Rates for each Possible Pattern of Significance* | | | | | | | | |
| | None of the Three Functions Significant | | | | | | | |
| NS | NS | NS | 0.977 | 0.867 | 0.758 | 0.965 | 0.863 | 0.757 |
| | Exactly One Function Significant | | | | | | | |
| NS | NS | S | 0.007 | 0.040 | 0.065 | 0.010 | 0.031 | 0.050 |
| NS | S | NS | 0.008 | 0.038 | 0.064 | 0.010 | 0.039 | 0.070 |
| S | NS | NS | 0.007 | 0.049 | 0.072 | 0.015 | 0.049 | 0.071 |
| | Exactly Two Functions Significant | | | | | | | |
| NS | S | S | 0 | 0.002 | 0.021 | 0 | 0.010 | 0.025 |
| S | NS | S | 0 | 0.002 | 0.013 | 0 | 0.004 | 0.011 |
| S | S | NS | 0 | 0.002 | 0.007 | 0 | 0.004 | 0.013 |
| | All Three Functions Significant | | | | | | | |
| S | S | S | 0 | 0 | 0.001 | 0 | 0 | 0.003 |
| | | TOTAL | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| *Significance Rates for Each Individual Function* | | | | | | | | |
| S | – | – | 0.007 | 0.042 | 0.082 | 0.015 | 0.057 | 0.098 |
| – | S | – | 0.008 | 0.040 | 0.089 | 0.010 | 0.053 | 0.111 |
| – | – | S | 0.007 | 0.051 | 0.102 | 0.010 | 0.045 | 0.089 |

One might decide that the data show a departure from CSR if any one of the three observed functions crosses its envelope boundary. One can account for this simultaneous inference by using the Bonferroni inequality. If the tolerance envelopes for $F$, $G$, and $L$ each have associated probability of $1-\alpha$, then the chance that a CSR process would lead to one or more of the observed functions crossing an envelope boundary is no larger than $3\alpha$. The results in Table 3 show that for $3\alpha = 0.03$, 0.15, and 0.30, the proportions of simulations where one or more of $\hat{F}$, $\hat{G}$, and $\hat{L}$ crossed a boundary are 0.023, 0.133, and 0.242, respectively, for $N = 50$, and 0.035, 0.137, and 0.243 for $N = 100$.

## 5. Discussion

### (i) Other distance functions used in microbial adsorption studies

H. Busscher and his colleagues have mapped positions of adhering bacteria in flowing systems (BUSSCHER, DOORNBUSCH, and VAN DER MEI, 1992; SJOLLEMA et al., 1990; COWAN and BUSSCHER, 1993). They calculated two distance functions, which they called the "radial pair distribution function" and the "angular pair distribution function" (SJOLLEMA and BUSSCHER, 1990). The radial pair distribution function is proportional to the derivative of $K$ and therefore is proportional to the second order intensity parameter $\lambda_2(r)$ (STOYAN, KENDALL, and MECKE, 1987). The angular pair distribution function accounts for possible anisotropy. For any pair of bacteria consider the angle of the line connecting the downstream bacterium to the upstream one, where "angle" is defined relative to the direction of flow. The angular pair distribution function, which is called the "directional distribution" in STOYAN et al., (1987), is the radial pair distribution function restricted to pairs of bacteria for which the angle is within a specified range. Stoyan and Benes have suggested alternative approaches for describing anisotropies (STOYAN and BENES, 1991). BUSSCHER et al. (1991) recently proposed a general "pair distribution function" which contains information for calculating the angular pair distribution function for any angle. They recommend the pair distribution function only for patterns involving at least two thousand bacterial positions.

### (ii) Future work

The method of Table 1 for forming tolerance envelopes could easily be adapted for use with the angular and radial pair distribution functions so that those functions could be applied to maps of only a few bacteria positions. For an increased flow velocity, the spatial pattern of colonizing bacteria may well become anisotropic. It would be informative to simulate anisotropic processes for purposes of comparing the power levels of tests based on the various functions.

Recent experiments provided data discrediting the CSR hypothesis. We are presently correlating bacteria positions with topographic and chemical features of the stainless steel surface. It is anticipated that this investigation will lead to an inhomogeneous Poisson process model for bacterial adsorption in which the Poisson intensity parameter is a function of topography or chemistry. If that modeling task is successful, then we will do a simulation study with the non-CSR model to determine the power of the tolerance envelope procedure for each function ($F$, $G$, $L$, angular pair, and radial pair).

### (iii) Conclusions

Statistical SPPA can be applied to bacterial colonization data. It has been successfully used to communicate experimental results to scientists and engineers. A small simulation study indicates that the nominal percentage attached to the

simulation envelopes is accurate. The $L$ function and associated envelope are powerful enough to detect patterns known not to be completely spatially random. A formal power analysis cannot be conducted until an appropriate alternative model has been discovered.

## Acknowledgements

This research was supported in part by Cooperative Agreement EEC-8907039 between the National Science Foundation and Montana State University and by the Industrial Associates of the Center for Biofilm Engineering. We benefited from constructive suggestions by anonymous referees. Pierre Morin contributed the French language summary.

*Résumé*

En utilisant des techniques de microscopie sophistiquées, nous avons observé la distribution spatiale de bactéries colonisant un coupon en acier inoxidable stérile immergée dans une solution bactérienne. Lors de ces expériences, nous avons utilisé de l'acier de différentes surfaces en terme de chimie et de rugosité. Le but final était d'identifier les éléments influencant l'adsorbtion bactérienne. L'intêret statistique immédiat était de distinguer les distributions relevant du hasard complet des motifs réguliers ou dus a l'aggrégation bactérienne. Ce but a été atteint en utilisant des analyses modifiées de fonctions de distance couramment employées en écologie de terrain. Cette méthode nous permettait d'éviter les problèmes possibles avec les comparaisons multiples. Pour la valeur nulle de la fonction de distance, nous avons calculé une zone de tolérance de telle manière que le niveau de tolérance soit équivalent pour toure les distances concernées. Des expériences de simulation sur ordinateur ont montré que le niveau nominal était correct. La méthodologie est utilisable pour détecter et décrire les motifs de colonisation connus pour n'être pas totalement du au hasard.

*References*

BUSSCHER, H. J., NOORDMANS, J., MEINDERS, J., and VAN DER MEI, H. C., 1991: Analysis of the spatial arrangement of microorganisms adhering to solid surfaces – methods of presenting results. Biofouling **4**, 71–79.

BUSSCHER, H. J., DOORNBUSCH, G. I., and VAN DER MEI, H. C., 1992: Adhesion of *mutans streptococci* to glass with and without a salivary coating as studied in a parallel-plate flow chamber. J. Dental Res. **71**, 491–500.

CAMPER, A., HAMILTON, M. A., JOHNSON, K. R., STOODLEY, P., HARKIN, G., and DALY, D. S., 1994: Bacterial colonization of surfaces in flowing systems: methods and analysis. Ultrapure Water **II**, 6, 26–35.

CHARACKLIS, W. G., and MARSHALL, K. C., 1990: *Biofilms.* New York: Wiley.

CLEVELAND, W. S., 1979: Robust locally weighted regression and smoothing scatterplots. J. Amer. Statist. Assoc. **74**, 829–836.

COWAN, M., and BUSSCHER, H. J., 1993: Flow chamber study of the adhesion of *Prevotella intermedia* to glass after preconditioning with mutans streptococcal species: kinetics and spatial arrangement. Microbios. **73**, 135–144.

CRESSIE, N. A. C., 1991: *Statistics for Spatial Data.* New York: Wiley.

DIGGLE, P. J., 1983: *Statistical Analysis of Spatial Point Patterns.* New York: Academic Press.

DOYLE, R. J., NESBITT, W. E., and TAYLOR, K. G., 1982: On the mechanism of adherence of *Streptococcus sanguis* to hydroxylapatite. FEMS Microbiological Letters **15**, 1–15.

ESCHER, A. R., and CHARACKLIS, W. G., 1990: Modeling the initial events in biofilm accumulation. In: *Biofilms*, pp. 445–486. Edited by CHARACKLIS, W. G., and MARSHALL, K. C. New York: Wiley.

GILLIS, R., 1993: Bacterial colonization and modification of grain boundaries on 316L stainless steel (M. S. Thesis, Department of Microbiology, Montana State University, Bozeman, MT 59717, U.S.A.).

HANISCH, K. H., 1984: Some remarks on estimators of the distribution function of nearest neighbor distance in stationary spatial point processes. Math. Operationsforsch. Statist. Ser. Statist. **15**, 409–412.

KOEN, C., 1991: Approximate confidence bounds for Ripley's statistic for random points in a square. Biom. J. **33**, 173–177.

LAPPIN-SCOTT, H. M., COSTERTON, J. W., and MARRIE, T. J., 1992: Biofilms and biofouling. In: *Encyclopedia of Microbiology* – Vol. 1, pp. 277–284. Edited by LEDERBERG, J., San Diego: Academic Press.

MARSHALL, K. C., 1985: Mechanisms of bacterial adhesion at solid-water interfaces. In: *Bacterial Adhesion-Mechanisms and Physiological Significance*, pp. 133–161. Edited by SAVAGE, D. C., and FLETCHER, M. New York: Plenum.

MATTYSSE, A. G., 1992: Adhesion, Bacterial. In: *Encyclopedia of Microbiology – Vol. 1*, pp. 29–36. Edited by LEDERBERG, J. San Diego: Academic Press.

MOOD, A. M., GRAYBILL, F. A., and BOES, D. C., 1974: *Introduction to the Theory of Statistics* – Third Edition. New York: McGraw-Hill.

QUIRYNEN, M., MARECHAL, M., VAN STEENBERGHE, D., BUSSCHER, H. J., and VAN DER MEI, H. C., 1991: The bacterial colonization of intra-oral hard surfaces in vivo: influence of surface free energy and surface roughness. Biofouling **4**, 187–198.

RIPLEY, B. D., 1976: The second-order analysis of stationary point processes. J. Appl. Prob. **13**, 255–266.

RIPLEY, B. D., 1977: Modelling spatial patterns. J. Roy. Statist. Soc. B**39**, 172–212.

RIPLEY, B. D., 1981: *Spatial Statistics*. New York: Wiley.

RIPLEY, B. D., 1988: *Statistical Inference for Spatial Processes*. New York: Cambridge University Press.

SCHMIDT, R., 1992: Aggregation kinetics of *Saccharomyces cerevisiae* on solid surfaces. Acta Biotechnologica **12**, 203–212.

SJOLLEMA, J., VAN DER MEI, H. C., UYEN, H. M., and BUSSCHER, H. J., 1990: Direct observations of cooperative effects in oral streptococcal adhesion to glass by analysis of the spatial arrangement of adhering bacteria. FEMS microbiology Letters **69**, 263–270.

SJOLLEMA, J., and BUSSCHER, H. J., 1990: Deposition of polystyrene particles in a parallel plate flow cell. 2. Pair distribution functions between deposited particles. Colloids and Surfaces **47**, 337–352.

STOYAN, D., and BENES, V., 1991: Anisotrophy analysis for particle systems. J. Microscopy **164**, 159–168.

STOYAN, D., KENDALL, W. S., and MECKE, J., 1987: *Stochastic Geometry and Its Applications*. New York: Wiley.

THOMAS, J. M., WARD, C. H., RAYMOND, R. L., WILSON, J. T., and LOEHR, R. C., 1992: Bioremediation. In: *Encyclopedia of Microbiology* – Vol. 1, pp. 369–385. Edited by LEDERBERG, J. San Diego: Academic Press.

Received, Jan. 1994
Revised, July 1994
Accepted, July 1994

M. A. HAMILTON
Center for Biofilm Engineering
409 Cobleigh Hall
Montana State University
Bozeman, MT 59717-0398
U.S.A.