Abstract:
One approach to modelling spatial distributions of fractures in rock masses is by stochastic marked point processes. This paper describes the application of parametric modelling techniques to a rock fracture dataset. In particular, maximum likelihood and least square methods are used to fit a non-homogeneous Poisson model, a cluster model and a Cox model to the data. The fitted models are then tested and verified by using the summary statistics of inter-event distance, nearest-event distance and $K$-function with the help of Monte Carlo simulations.

Key words: point process, parametric modelling, Poisson process, Cox model, maximum likelihood estimation.

1. Introduction
One way of modelling spatial distributions of fractures in rock masses by stochastic geometry, and in particular, marked point processes, see contributions of Lee (1990), Bear (1993), Wen (1996) and Zhang (2000). There are three broad steps in the approach: conversion of the fractures into a marked point dataset, modelling using Poisson point (or other stochastic) processes and the transformation of the results back to the original context. Modelling is the most important and challenging step.

The non-homogeneous Poisson process and the Poisson cluster process are two fundamental stochastic models for point processes. They describe, respectively, two basic forming mechanisms of point patterns: inhomogeneous realisations and parent-daughter realisations. The Cox process, on the other hand, has properties of both models and is a two-stage random process, which is inhomogeneous in nature but mathematically equivalent to a cluster process (see Diggle, 1983, Cressie, 1993 and Stoyan, 1995).

The investigation reported here deals with a point pattern derived from a fracture dataset. An optimal non-parametric model (Xu et al., 2003) is used to suggest a parametric form for the non-homogeneous modelling of the point pattern and the parameters of the model are estimated by marginal maximum likelihood. As there is an appreciable amount of clustering in the dataset a cluster model, using least-squares, is also proposed (Diggle, 1983). The noise between the optimal non-parametric model and the fitted parametric model is modelled by a Gaussian Cox process. The fitted models are then assessed using Monte Carlo simulations.

2. Example dataset and the proposed model
The example dataset used in this research is derived from a fracture trace image of a rock outcrop, as shown in Figure 1(a). The data were introduced by Lee (1990) for fracture modelling for slope engineering. Figure 1(b) is the derived point dataset using the midpoints of the fracture traces.
The most important part in parametric modelling is the choice of a suitable parametric model for the configuration. Ideally, the choice should be informed by detailed knowledge of the dataset and its physical context, such as the forming mechanism of the points. In geological applications, this involves an understanding and interpretation of the geological and geotechnical history of the sampling site. In the absence of such information techniques such as non-parametric intensity estimation can be used.

The dataset was analysed by optimal non-parametric estimation using kernel functions as described in Xu et al. (2003). Figure 2 shows the final non-parametric model estimated using the optimal bandwidth ($h=1.0$). Because of the small-scale variations a suitable parametric form for this optimal model is not immediately obvious. A smooth model, however, can be conjectured to describe the major variation (trend). The two high intensity areas are the direct result of the point aggregations evident in Figure 1(b). Because of the small proportion of points in the aggregations the dataset as a whole can be modelled by a non-homogeneous Poisson process.

Using a Gaussian function to model the two modes centred at ($13.4, 17.3$) and ($16.4, 13.4$), leads to the following parametric model:

$$
\lambda(x, y) = \begin{cases} 
\frac{1}{c}e^{-\frac{(x-13.4)^2}{d}}e^{-\frac{(y-17.3)^2}{d}} + \frac{1}{v}e^{-\frac{(x-16.4)^2}{d}}e^{-\frac{(y-13.4)^2}{w}}, & (x, y) \in \mathbb{R}' \\
0, & \text{otherwise}
\end{cases}
$$

where $\mathbb{R}'$ is a sub-region as defined in Figure 2. This sub-region is necessary to control the validity of the proposed model. $a$ is the background density; $b$ and $u$ are the respective strengths of the two modes; $c$, $d$, $v$ and $w$ are parameters for describing the influence ranges and anisotropies of the two modes.

3. Maximum likelihood estimation of the parameters

The likelihood function for a non-homogeneous Poisson process can be derived by assuming that the $N$ events (points) in the region $\mathbb{R}$ are independent random samples from a distribution with probability proportional to the density $\lambda$ at their individual locations $X_i$, $i=1,2,...N$ (e.g., Cressie, 1993):

$$
\ell(\theta) = \prod_{i=1}^{N} \lambda_{Q}(X_i) \cdot e^{-\int_{\mathbb{R}} \lambda_{Q}(u)\nu(du)}
$$

where $\theta$ is the parameter vector, in this case $\theta = \{a, b, c, d, u, v, w\}$. The maximum likelihood estimator, $\hat{\theta}$, is the one that maximises the likelihood functions, $\ell(\theta)$.

Direct application of the technique for our model proves difficult in terms of computing cost because of the number of parameters involved (7 in total). We have, therefore, used a marginal approach, i.e., $\ell(\theta)$ is maximised for each individual parameter in turn, i.e.,

$$
\hat{p}_i = \max_{\arg} \left\{ \ell(p_i, p_j) | p_j = \hat{p}_j, j = 1, 2,..., j \neq i \right\}
$$

where $p_i$ and $p_j$ refer to parameters $a$, $b$, $c$, $d$, $u$, $v$ or $w$. The process is repeated several times and the iteration stops when the maximum $\ell(\theta)$ value stabilizes. The estimated parameters are listed in Table 1 and the maximum likelihood non-homogeneous model is shown in Figure 3.

<table>
<thead>
<tr>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
<th>$u$</th>
<th>$v$</th>
<th>$w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.62</td>
<td>8.1</td>
<td>4.1</td>
<td>0.45</td>
<td>4.2</td>
<td>2.5</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Table 1 Estimated model parameters derived by ML.
Note that the major part of the region $\mathcal{R}$ is dominated by the background density $a = 0.62$ and the model, as expected, gives a continuously smooth density within the region. The maximum likelihood estimation shows that $l(\theta)$ is sensitive to the choices of parameters $d$, $u$ and $w$, but not to the other parameters. In this case $d$ and $w$ are the ranges covered by the two modes in the $y$ direction. The differences between $e$ and $d$ and between $v$ and $w$ suggest anisotropic behaviour in the data with a longer range of continuity in the $x$ direction.

4. Least square estimation of parameters for a cluster model

In the original form proposed by Diggle (1983), the least squares estimator $\hat{\theta}$ is that which minimises the discrepancy:

$$D(\theta) = \int_0^t \left[ \left( \hat{K}(t) \right)^2 - \left( K(t, \theta) \right)^2 \right] \cdot dt$$

(4)

There $\hat{K}(t)$ is Ripley’s $K$-function calculated from the data and $K(t, \theta)$ is the theoretical $K$-function with parameter vector $\theta$, if available, or the averaged $K$-function based on a number of Monte Carlo simulations using parameter $\theta$. $t_o$ and $c$ are tuning constants and, in general, values of $t_o = 0.25$ (for unit square region), $c = 0.5$ for more regular point patterns and $c = 0.25$ for more aggregated patterns are suggested.

As there is no unique relationship between the $K$-functions and the point patterns (see Baddeley 1984, 2000), the least squares technique based on the $K$-function is not applicable for parameter estimation for non-homogeneous Poisson processes. The technique, however, is widely used in modelling Neyman-Scott processes when the parent process is a stationary (homogeneous) Poisson process and, therefore, location (absolute or relative) parameters are not considered, see Diggle (1983), Cressie (1993) and, more recently, papers in Lawson et al. (2002).

Two parameters $(\rho, \sigma)$ are required if the Neyman-Scott process is a Thomas process, in which case the theoretical $K$-function can be written as:

$$K(t, \sigma, \rho) = \pi t^2 + \frac{1}{\rho} \left[ 1 - e^{-\frac{\rho}{4\sigma^2}} \right]$$

(5)

where $\rho$ is the intensity of the parent process and $\sigma^2$ is the dispersion variance of the bi-normal daughter distribution. The derivation of this equation, however, does not account for relationships between daughter points from different clusters and it is, therefore, valid only for ranges up to the size of the cluster, which depends on the daughter distribution variance and can be estimated as $4\sigma$. In this case, the tuning constant, in our opinion, should be set to $t_o = 4\sigma$ and $c = 1$ to maximise the goodness of fit within the valid range of the theoretical $K$-function. Care should be taken, however, for large $\rho$ and...
As demonstrated in Figure 4, for large \( \rho \) or \( \sigma \) values, the validity of the above theoretical solution is questionable.

There are clear indications in the fracture dataset of point aggregations around two locations (13.4, 17.3) and (16.4, 13.4). The \( \hat{K}(t) \) function of the dataset, as shown in Figure 5, can be used to fit a cluster model. As indicated, the least squares fitting gives values of \( \rho = 1/170 \) and \( \sigma = 2.875 \). Clearly it is not a good fit even within the valid range of the theoretical \( K \)-function. The reason could be that the point pattern is predominantly driven by a non-homogeneous process rather than a cluster process. Nevertheless, this model will be tested against that estimated by the likelihood method.

5. A Gaussian Cox model

A comparison of Figure 3 and Figure 2 shows that small-scale variations are not modelled by the non-homogeneous process given in Equation 1. The difference between the two models is shown in Figure 6(a), which quantifies the errors of using the non-homogeneous model to represent the more accurate non-parametric model. Figure 6(b) shows a histogram of values obtained by independently sampling the error field, which indicates that the error field can be modelled approximately by a Gaussian random variable.
The mean and variance of the values represented in Figure 6 (a) are, respectively, 0.091 and 0.406. We therefore introduce a Gaussian Cox process as an additional model to the non-homogeneous model previously obtained. This random component can be specified as:

\[
C(x, y) = \begin{cases} 
N(0,0.4) & (x, y) \in \mathbb{R} \\
0 & \text{otherwise}
\end{cases}
\] (6)

The slightly biased mean is ignored in this model. Note the purpose of this process is solely to model the small-scale, random component apparent in the non-parametric model. The final Cox model is the summation of Equations (1) and (6).

6. Model verifications

The models are tested using Monte Carlo simulation and the summary statistics of inter-event distance \( \hat{H}(t) \), nearest-event distance \( \hat{G}(y) \) and the K-function \( \hat{K}(t) \). One hundred independent simulations were conducted for each model and their \( \hat{h}(t) \) (histogram form of \( \hat{H}(t) \)), \( \hat{g}(y) \) (histogram form of \( \hat{G}(y) \)) and \( \hat{K}(t) \) functions were averaged and compared to those of the data. The results are shown in Figurea 7 (a) – (c). In these figures, the solid line is for the dataset shown in Figure 1 (b), line 1 is the average result for the non-homogeneous model estimated by maximum likelihood, line 2 is for the cluster model given in Section 5 and line 3 is for the Cox model described in Section 5.

The analysis indicates that, for this dataset, the non-homogeneous model is, overall, better than the other two models except that the Cox model gives performs better with respect to \( G(y) \). This is due to the extra random component used in the Cox process to model small-scale variations. Note that the cluster model does not provide a good fit. This is as expected as only a small proportion of the data points are involved in forming the apparent clusters. Figure 8 shows an example of the point realisation from the Cox model described. Lee et al. (1990) proposed a hierarchical model for this dataset but, as the models they used were not fully explained in their only paper, we are unable to make a comparison.
7. Summary and conclusions
This paper describes the application of parametric modelling techniques to a point dataset derived from a set of rock fracture traces. In summary:

- The non-homogeneous model estimated by maximum likelihood provides the best performance for this dataset. Non-parametric modelling can be used to suggest a parametric form when little is known of physical forming mechanisms of a point pattern realisation.
- The Least squares technique is only suitable for the Neyman-Scott process when the parent process is a homogeneous Poisson process. Determining parameters using this technique is a simple procedure. If, however, the underlying model does not physically follow the proposed process, there is likely to be a significant difference between the summary statistics for the fitted model and those for the dataset.
- Local scale variations can be modelled by a Cox process. The final model will be the sum of the large-scale, non-homogeneous model (general trend model) and the local scale Cox model. This partial Cox model approach improves the small scale summary statistics such as the nearest-event distance $G(y)$.

8. Acknowledgement
This work is supported by EPSRC grant number GR/R94602.

9. References: