A Continuous Latent Spatial Model for Crack Initiation in Bone Cement

Elizabeth A. Heron\(^1\) Cathal D. Walsh\(^2\)
\(^1\)Department of Statistics
University of Warwick, Coventry CV4 7AL, UK
\(^2\)Department of Statistics
Trinity College, Dublin 2, Ireland

Abstract

Hip replacements provide a means of achieving a higher quality of life for individuals who have, through aging or injury, accumulated damage to their natural joints. This is a very common operation, with over a million people a year benefiting from the procedure. The replacements themselves fail mainly as a result of the mechanical loosening of the components of the artificial joint due to damage accumulation occurring. This damage accumulation consists of the initiation and growth of cracks in the bone cement which is used to fixate the replacement in the human body.

The data come from laboratory experiments designed to assess the effectiveness of the bone cement in resisting damage. In this paper we examine the properties of the bone cement, with the aim being to estimate the impact that both observable and unobservable spatially varying factors have on causing crack initiation. To do this, an explicit model for the damage process is constructed taking into account the tension and compression at different locations in the specimens. A Gamma random field is used to model any latent spatial factors that may be influential in crack initiation. Bayesian inference is carried out for the parameters of this field and related covariates using Markov chain Monte Carlo techniques.

Keywords: Bone cement, crack initiation, Gamma random field, Inverse Lévy Measure algorithm, spatial Poisson process.

1 Introduction

In many statistical applications the data presented for analysis contain spatial information, an obvious example being that the data might contain the precise locations of the events of interest. Statistical models that take into account available spatial information are common in many areas of applied statistics and several different approaches to modelling the various types of spatial data that are encountered, are available, see for example Cressie (1993) and Diggle (2003).

Together with the locations of events being available, it is often the case that spatial covariate information is also presented. It is, of course, important to incorporate this spatial covariate information in any model being proposed. It is also sometimes the case that influential, but unobserved, covariates are known to vary across the region under consideration. Information regarding these unobserved phenomenon should be included, as carrying out analysis without taking such information into account may not result in accurate inferences.

In this paper the data we analyse come from laboratory experiments carried out to examine damage accumulation of bone cement in orthopaedic hip replacement specimens. Damage accumulation consists of the initiation and growth of cracks in the acrylic polymer (bone cement) used to fixate the components of the prosthesis to the bone. The data available consists of \(x\) and \(y\) coordinates of the crack locations in each of five specimens, together with stress measurements obtained through a finite element analysis.

Previous analysis carried out on damage accumulation in bone cement has tended to focus on investigating lifetimes and reliability of replacements. For example, the fitting of a Weibull distribution in a survival analysis is carried out in Malchau and Herberds (1998). In the reliability
analysis carried out by Wilson (2005), the lifetime of a replacement is considered as a function of material properties of the cement, with the reliability of a specimen being a function of the number and length of the cracks in the bone cement. The models fitted for the initiation and the growth are based on previous analysis carried out on the same data by McCormack et al. (1998). The type of models fitted are necessarily dependent on the nature of the available data. Unlike the previous data analysed by Wilson (2005) and by McCormack et al. (1998), spatial information regarding both the crack locations and stress levels is readily available in the data presented here for analysis.

We aim to fit a spatial Poisson identity-link regression model that incorporates the observed (stress measurements) covariate information. Together with this observed information available in the data it is known that pores (air bubbles) in the bone cement are influential in the initiation of cracks. However, the spatial distribution of such pores is not known and cannot easily be obtained experimentally. The model we propose will attempt to incorporate information regarding unobserved spatially varying influential factors through a Gamma random field.

Gamma random fields offer a means by which we can model uncertainty about both the location and the size of factors that we believe have an influence on the formation of cracks. Wolpert and Ickstadt (1998a) introduced the idea of using Gamma random fields in their class of Bayesian hierarchical models used to analyse spatially dependent count data. As an illustrative example, they modelled the density and spatial correlation of hickory trees. The incorporation of a Gamma random field and Gaussian kernel in order to model latent spatial covariates was also used in the analysis of the effect of traffic pollution on respiratory disorders in children (Best et al. 2000a). It has also been used in the analysis of origin/destination trip data (Ickstadt and Wolpert 1999).

We carry out the analysis using a Bayesian approach with Markov chain Monte Carlo (MCMC) techniques being employed in order to make inferences regarding the parameters of the model. The analysis will also necessitate the simulation of Gamma random fields and this is carried out using the Inverse Lévy Measure (ILM) algorithm.

The paper is structured as follows. The next section gives a detailed description of the orthopaedic hip replacement together with the experimental model that gave rise to the data which form the basis of our analysis. The proposed model is then introduced in Section 3. In Section 4 the model fitting is discussed. In Section 5 we present the results from the analysis carried out, and we discuss model validation. In Section 6 we draw some conclusions.

2 Description of the Orthopaedic Hip Replacement and the Data

2.1 Orthopaedic Hip Replacement and the Bone Cement

The deterioration of a structure that is subjected over time to an external loading, resulting in the inability of the structure to carry the intended loading, is known as fatigue. Crack initiation and crack growth are regarded as the basic causes of fatigue damage accumulation and ultimate fatigue failure (Sobczyk and Spencer 1992).

Orthopaedic joint replacements are used to replace human joints that no longer function as they should. In particular, an orthopaedic hip replacement is used to replace the ball and socket components of the hip joint. The replacement of a hip joint is a very common procedure with up to one million hip replacements being carried out annually (Huiskes and Verdonschot 1997). Typically the hip replacement consists of a prosthesis, usually metallic, being inserted into the medullary cavity of the femur bone which has been hollowed out. The metal prosthesis is a replacement for the “ball” portion of the joint. The “socket” portion of the joint is replaced with an artificial cup (typically this is made from ultra high molecular weight polyethylene; UHMWPE). The prosthesis is held in place by an acrylic polymer (polymethylmethacrylate; PMMA) cement mantle that interlocks the prosthesis and the bone. See Figure 1.

The polymer cement used in the fixation of the joint replacement is often referred to as bone cement. The cement does not form any sort of chemical bond with either the metal prosthesis or the bone. The fixation occurs instead through a mechanical interlocking that arises between
the surface of the implant and interdigitation with cancellous bone. An analogy given by Lennon (2002) for the function of the bone cement is that it performs a similar role to that of a “grout”. The bone cement is prepared in the operating theatre several minutes before the components of the joint replacement are inserted. It is prepared as a self-curing, dough-like resin thus ensuring that it is possible to insert the cement into the prepared cancellous bone and implant the prosthesis into the cement. The cement hardens within 10-15 minutes of initial preparation allowing enough time for the insertion. When hard, the bone cement has a similar composition to Perspex or Plexiglas used in industrial settings.

2.2 Reasons for Failure

There are two main reasons for failure in cemented hip replacements: infection and mechanical (aseptic) loosening of the components. Due to improvements in surgical conditions, infection has been almost eliminated (cumulative revision rate for deep infection after ten years is 0.3%, Malchau (2000)) and aseptic loosening of the components is now the dominant mode of failure, in particular aseptic loosening of the femoral stem (Malchau (2000)).

According to Lennon and Prendergast (2001), the aseptic loosening is usually caused by fatigue failure of the cement mantle under cyclic loading. Maintaining the cement mantle is not simply a matter of reducing the peak stress in the mantle. Lennon and Prendergast (2001) highlight the situation where the influence of cement porosity may dominate the effect of stress to such an extent that failure may occur, not at the location of peak stress but instead where the pores are largest. During preparation, pores (air bubbles) become trapped in the cement. Retrieval studies carried out show that a link exists between the porosity in the bone cement in the hip replacement and damage accumulation (Jasty et al. (1991) and Culleton et al. (1993)). The amount of porosity varies depending on the way in which the cement is mixed. Improvements in the mixing of the cement such as mixing under a vacuum and centrifuging as opposed to manual mixing with a bowl and spoon, decrease the amount of air bubbles trapped but do not eliminate all pores.

Other factors that are influential in determining the lifetime of a replacement do exist. For example, the femoral stem design, the surgeon’s skill, biological and non-biological inclusions in the cement, will all have an impact. However, under controlled laboratory conditions, variability exists in the damage accumulation in the bone cement and according to Lennon and Prendergast (2001) this variability is large enough that it might dominate.
2.3 The Experimental Model

The experimental model was designed in such a way as to retain the most important physical features of the femoral hip replacement, ensuring that it would behave in the laboratory in a similar fashion to a hip replacement in the human body. It was also necessary to design the model in such a way that measurements of damage accumulation could be made, i.e., that cracks could be observed and located.

The model consists of a femoral stem encased between layers of cement and strips of cancellous bone, see Figure 2. These are then held in two aluminium side plates which offer support similar to that which would be given by the cortical bone in the human body. The side plates contain windows (one medial, one lateral) in which the cement is exposed and therefore available for observation. The cement was mixed by hand and the particular type of cement that was used is translucent, enabling cracks to be stained and viewed by light transmission.

A stress loading was then applied to the physical model and with the addition of a dye penetrant to the cement layers of each of the specimens, a magnified image of the cement surface could then be projected onto a screen using an optical comparator. Crack locations were identified and traced onto acetate transparencies. The transparencies were digitally scanned and image analysis was carried out in order to obtain the position of each crack.

3 The Model

In proposing a model for crack initiation all available information, whether observed or unobserved, regarding the factors that are influential should be incorporated. As well as individual crack locations, the data provided also include a finite element analysis of the stress measurements at particular locations within the cement. Both compression and tension (negative and positive stresses, respectively) have an effect on crack initiation and we would like to estimate the effect that each has.

Another factor that has an impact on the initiation of cracks is the distribution of pores in the bone cement. In the experiment from which our data come, it was not possible to locate pores within the cement. Thus we do not have any idea as to the distribution (spatial or otherwise) of the pores in each of the specimens, but we would like to incorporate in our model the knowledge that these factors do influence crack initiation.

In modelling the crack locations, a natural and obvious model to consider is some form of spatial Poisson process. We would like the intensity of this process to reflect both the known and unknown spatially varying factors that are believed to influence the rate of crack formation, thus we are
considering an inhomogeneous Poisson process. The inclusion of observed and measured factors, compression and tension stresses, allows for inference and direct comparison of the influence of these important physical features. Previous models for crack initiation have not facilitated such comparisons. We also require the model to give an indication as to the spatial impact other unobserved factors (such as the pores previously mentioned) have on crack formation. We allow the Poisson intensity to incorporate a latent spatial variable through a Gamma random field. This is a flexible addition that allows investigation of spatial patterns in the crack locations that are not explained by either of the measured stresses. Modelling of such influential factors has not previously been carried out. Our choice of model, although being flexible, is specific as regards its components, but for each of these components we have physical evidence provided by engineers that merits its inclusion.

3.1 Poisson Process

Consider the $x$ and $y$ coordinates of the locations of the cracks. Let $l_{ik} \in \mathbb{R}^2$ denote the spatial coordinates of crack $k$ of specimen $i$, $i = 1, \ldots, 5$, $k = 1, \ldots, K_i$. Specimen $i$ having a total of $K_i$ cracks. This set of crack locations $\{l_{ik}\}$ can be considered as a random countable subset of some space $L \subset \mathbb{R}^2$. In particular we can model this set of crack locations as a Poisson process $L$ on $\mathcal{L}$, that is, for any disjoint measurable subsets $A_1, A_2, \ldots, A_n$ of $\mathcal{L}$, the random variables $N(A_1), N(A_2), \ldots, N(A_n)$ are independent and Poisson distributed, $(N(A) = #\{L \cap A\})$.

For each crack location, we now obtain, through kriging, the stress value at that location and denote this stress value by either $C(l_{ik}) = C_{ik}$, if it is a compression stress, or $T(l_{ik}) = T_{ik}$ if it is a tension stress. Note that one of either $C_{ik}$ or $T_{ik}$ will be zero as compression and tension cannot both be present at a given location. For each crack location $l_{ik}$ we can consider the attribute vector $a_{ik} = (C_{ik}, T_{ik})$.

In a similar fashion as we did with the set of crack locations we can consider the set of attribute vectors $\{a_{ik}\}$ as lying in some space $\mathcal{A} \subset \mathbb{R}^2$. We would like to combine all observed information, the crack locations and the attribute vectors, in one model. This can be done by considering a marked point process. By associating a vector of random variables, that is, the attribute vector, with each point of the random set $\{l_{ik}\}$, each point of the Poisson process $L$, we define a marked Poisson process.

We then consider the product space $\mathcal{X} = \mathcal{L} \times \mathcal{A}$ with the pair $x = (l, a), x \in \mathcal{X}, l \in \mathcal{L}, a \in \mathcal{A}$ being thought of as a random point in the product space $\mathcal{X}$. In turn, the set of points $\{x\}$ forms a random countable subset of $\mathcal{X}$. A fundamental result is that this set of coordinates $\{x\}$ of marked points in the product space is a Poisson process (Kingman (1995)).

Thus we have a Poisson process

$$N(dx) \sim \text{Poisson}(\Lambda(dx)), \quad (1)$$

defined on the product space $\mathcal{X}$. $\Lambda(dx)$ is an uncertain and inhomogeneous intensity measure and because of this $N(dx)$ can also be thought of as a doubly-stochastic Poisson process or Cox process.

3.2 Intensity Measure

The next stage in the construction of the model is to consider the intensity measure and how this will be modelled. We model $\Lambda(dx)$ as a product of the intensity at a point $x$ and a reference measure $\omega(dx)$ on $\mathcal{X}$. In this case we choose $\omega(dx)$ to be an area-weighted reference measure (Lebesgue measure). Thus we have

$$\Lambda(dx) = \Lambda(x)\omega(dx),$$

and the total number of cracks in $\mathcal{X}$ is given by

$$N(\mathcal{X}) \sim \text{Poisson} \left( \int_{\mathcal{X}} \Lambda(x)\omega(dx) = \Lambda(\mathcal{X}) \right)$$

We turn our attention to modelling the intensity at a point using a regression model with an identity link that incorporates both the observed (compression and tension) and unobserved factors (latent
factors, possibly representing the unknown spatial distribution of pores) that are believed to have an influence on crack initiation. We define the intensity at a point $x$ as

$$\Lambda(x) = C(l)\beta_1 + T(l)\beta_2 + X_3\beta_3, \quad \forall \ x \in X,$$

where the coefficients $\beta_1$ and $\beta_2$ are indicators of the influence of compression and tension, respectively, on causing cracks to initiate. The third term $X_3\beta_3$ in the identity link regression model represents the influence of the unobserved (latent) factors, which we will define subsequently.

### 3.3 Latent Spatial Covariate

If unobserved covariates vary continuously over the space, it is important to include them in the analysis and, as we have reason to believe that the distribution of pores varies continuously, we incorporate a latent spatial covariate to account for this. In order to define this covariate suppose we initially introduce a set of $M$ random locations $\{s_m\}_{m \in M}$ in $S \subset \mathbb{R}^2$ where $S$ is some region such that $\mathcal{L} \subset S$ and with each of these we associate a set of random latent magnitudes $\{\gamma_m\}_{m \in M}$, not necessarily all equal. We choose to model the influence, that the latent magnitudes $\{\gamma_m\}_{m \in M}$ have on causing cracks to form, with a Gaussian kernel, $k(l, s_m) = (1/(2\pi \rho^2)) \exp(-|l - s_m|^2/2\rho^2)$, depending on Euclidean distance, for any location $l \in \mathcal{L}$. Thus we are modelling the unobserved factors that influence the formation of cracks as point sources of not necessarily equal magnitudes, whose influence decreases with increasing distance from the point source, and the rate at which the influence decreases is determined by a Gaussian kernel. We consider the following

$$\sum_{m=1}^{M} k(l, s_m)\gamma_m. \quad (2)$$

The magnitudes $\{\gamma_m\}$ together with the locations $\{s_m\}$ are an approximation to any unobserved spatially varying latent covariate. Suppose now we increase the set, $\{s_m, \gamma_m\}$, in the limit this leads to a random field, which we denote by $\Gamma(ds)$,

$$\Gamma(ds) = \sum_m \gamma_m \delta_{s_m}(ds).$$

### 3.4 Gamma Random Field

In line with work in Wolpert and Ickstadt (1998a), Wolpert and Ickstadt (1998b) and Best et al. (2000a) we now introduce a particular type of random field, namely the Gamma random field. A random field $\Gamma(ds) \sim \text{Gamma}(\alpha(ds), b(s))$ is said to be a Gamma random field with shape measure $\alpha(ds)$ and scale function $b(s)$ over some set $S$ if

- $A \subset S$, $\Gamma(A) \sim \text{Gamma}\left(\int_A \alpha(s)ds, b(s)\right)$;
- If two sets $A, B \subset S$ are disjoint, then $\Gamma(A)$ and $\Gamma(B)$ are independent, (independent increments).

Note that the distribution is exact if $b(s)$ is constant on $A$, otherwise it is an approximation. Gamma random fields offer a means by which we can model uncertainty about both the location and the size of factors (for example, the pores in the cement) that we believe have an influence on the formation of cracks. The influence of all latent spatial point sources on a point $l \in \mathcal{L}$, Equation 2, now has the following integral form

$$\int_S k(l, s)\Gamma(ds).$$

The intensity at a point $x$ for the marked Poisson process then becomes

$$\Lambda(x) = C(l)\beta_1 + T(l)\beta_2 + \int_S k(l, s)\Gamma(ds). \quad (3)$$
For notational simplicity we let \( X_1(x) \beta_1 = C(l) \beta_1, \) \( X_2(x) \beta_2 = T(l) \beta_2, \) and \( X_3(x) \beta_3 = \int_S k(l, s) \Gamma(ds) \). The intensity can now be written as

\[
\Lambda(x) = \sum_{k=1}^{3} X_k(x) \beta_k.
\]

4 Analysis

Our analysis will be carried out in a Bayesian framework and in the following we present the joint likelihood and specify prior distributions for the parameters of interest.

4.1 Likelihood

The Poisson regression model with identity-link (see Equations 1 and 3) is the model we have chosen: \( N(X) \sim \text{Poisson}(\Lambda(X) = \Lambda) \). The joint likelihood for all crack locations for all five of the specimens may be written as follows

\[
P\left( \{ N_{ij} \} | \beta_1, \beta_2, \{ \Gamma_{ij}(ds) \}, \rho \right) = \prod_{ij} \frac{\exp(-\Lambda_{ij})^\Lambda N_{ij}}{N_{ij}!},
\]

where the indexing sets are specimen: \( i = 1, 2, \ldots, 5 \) and window: \( j = 1, 2 \) and \( N_{ij} = N_i(X_j) \). We model the lateral and medial windows separately since there is no physical link in the laboratory model between the two windows, see Figure 2. Thus \( N_{ij} \) represents the total number of cracks in window \( j \) of specimen \( i \). The intensity is given by

\[
\Lambda_{ij} = \Lambda_i(X_j) = \beta_1 \int_{X_j} X_1(x) \omega(dx) + \beta_2 \int_{X_j} X_2(x) \omega(dx) + \int_{S_j} \int_{S_j} k(x, s) \Gamma_{ij}(ds) \omega(dx),
\]

where \( X_j \) is the lateral (\( j = 1 \)) or the medial (\( j = 2 \)) window and \( S_j \) is a rectangular region containing the lateral window (\( j = 1 \)) or the medial window (\( j = 2 \)).

4.2 Priors

We would like to make inferences regarding the unknown parameters \( \beta_1, \beta_2, \rho \) and also the Gamma random fields \( \Gamma_{ij}(ds) \). For the coefficients of compression and tension, \( \beta_1 \) and \( \beta_2 \) respectively, we choose independent Gamma prior distributions. One reason for choosing Gamma priors is that the mean of the Poisson must be non-negative; another is our belief that each of these corresponding factors positively influences the formation of cracks. We also have reason to believe from communications with engineers that tension stresses have a greater impact on crack initiation than do compression stresses and our priors should reflect this knowledge. We choose the following priors:

\[
\begin{align*}
\pi(\beta_1) &\sim \text{Gamma}(\alpha_1 = 1, b_1 = 0.1), \ E(\beta_1) = 10, \\
\pi(\beta_2) &\sim \text{Gamma}(\alpha_2 = 3, b_2 = 0.1), \ E(\beta_2) = 30.
\end{align*}
\]

The parameter \( \rho \) indicates the distance over which the latent spatial variables have an influence. Best et al. (2000a) used this type of kernel in their identity link spatial regression model which related the prevalence of respiratory illness in children in Huddersfield, UK, to NO\(_2\) concentrations and unmeasured factors. The authors treated \( \rho \) as fixed or certain. Several different fixed values for \( \rho \) were considered and the value for \( \rho \) that was chosen was the one that gave results that were most consistent with the data. Best et al. (2000b), again examining the prevalence of respiratory illness in children using an identity link regression model, chose a Lognormal prior for \( \rho \). The mean and variance of this Lognormal prior reflected the fact that spatial effects on a small scale would not be detectable and those on a much larger scale would appear as large-scale trends. Wolpert
and Ickstadt (1998a) analyse the density and spatial correlation of hickory trees and they also use a Gaussian kernel in their analysis, again with Lognormal prior with appropriate mean and variance parameters. For \( \rho \) we choose a Lognormal prior. Engineering intuition suggests that the influence of the latent factors would be of a localised nature. For a given value of kernel will lie within a radius of \( 2\rho \), thus the influence of a latent point source will not extend much beyond this radius. We choose the following parameters for the Lognormal prior distribution for \( \rho \):

\[
\pi(\rho) \sim \text{Lognormal}(\mu = 1, \sigma = 1.4),
\]

giving a prior mean of \( \approx 7.2 \). This allows for a reasonable range of influence given the size of a window. Later, we consider other prior distributions for \( \rho \) in order to examine the sensitivity of the prior on our analysis.

For the prior for the Gamma random fields \( \Gamma_{ij}(ds) \) we choose a Uniform shape measure \( \alpha(ds) = 0.6 \) on a rectangular region surrounding each of the lateral and medial windows separately and we choose a constant scale parameter \( b = 1 \). This allows for a vague, apriori Gamma random field with no prior indication as to the location of the unobserved factors.

### 4.3 The Posterior Distribution

The joint posterior distribution for all the data can be written as follows:

\[
\mathbb{P}(\beta_1, \beta_2, \{\Gamma_{ij}(ds)\}, \rho|\{N_{ij}\}) \propto \prod_{ij} \left\{ \frac{\exp(-A_{ij})(A_{ij})^{N_{ij}}}{N_{ij}!} \pi(\Gamma_{ij}(ds)) \right\} \pi(\beta_1)\pi(\beta_2)\pi(\rho).
\]

### 4.4 Inference

We want to perform inference on the parameters \( \beta_1, \beta_2, \rho \), and \( \{\Gamma_{ij}(ds)\} \). In order to do this we employ MCMC techniques. The first of which is data augmentation (Tanner and Wong (1987)). For each crack \( (k) \) in window \( (j) \) of specimen \( (i) \) we have a point \( x_{ijk} = (l_{ijk}, a_{ijk}) \) in the product space \( X_j \). During each iteration of the algorithm, we assign to each of these points an indicator \( I_{ijk} \in \{1, 2, 3\} \), where \( \mathbb{P}(I_{ijk} = n) \propto X_n(x_{ijk})\beta_n \), \( n = 1, 2, 3 \). And we introduce a new set of random variables \( N_{ijn}(X_j) = N_{ijn} \) defined as

\[
N_{ijn} = \#\{k : I_{ijk} = n\}, \quad N_{ij} = \sum_{n=1}^{3} N_{ijn}
\]

The idea is that we are breaking up the count \( N_{ij} \) of cracks into a sum of counts, where each of these new counts represents the number of cracks attributable either to compression, tension, or latent factors, \( N_{ij1}, N_{ij2}, N_{ij3} \), respectively. The inclusion of this data augmentation step facilitates the use of the Gibbs sampler in order to draw samples from the full conditional distributions for \( \beta_1 \) and \( \beta_2 \). The conditional distributions for \( \beta_1 \) and \( \beta_2 \) are as follows:

\[
\mathbb{P}(\beta_1|\beta_2, \rho, \{\Gamma_{ij}\}) \sim \text{Gamma} \left( \sum_{ij} N_{ij1} + \alpha_1, I \sum_{j} \int_{X_j} X_1(x)\omega(dx) + b_1 \right),
\]

and

\[
\mathbb{P}(\beta_2|\beta_1, \rho, \{\Gamma_{ij}\}) \sim \text{Gamma} \left( \sum_{ij} N_{ij2} + \alpha_2, I \sum_{j} \int_{X_j} X_2(x)\omega(dx) + b_2 \right),
\]

where \( I = 5 \), the total number of specimens.

For the parameter \( \rho \) we propose updating using a random walk Metropolis step.

We now examine the full conditional distribution for \( \Gamma(ds) \), the Gamma random field over the space \( S \). The following method and results (Wolpert and Ickstadt (1998a)) apply also to \( \Gamma_{ij}(ds) \), the Gamma random field over the space \( S_j \) for all \( i \) and \( j \).
Consider $N_3(dx)$, it is a finite integer-valued measure on $\mathcal{X}$ and as such can be represented as the sum of a random number of unit point masses at points $x_n$, which need not necessarily be distinct. Again we use the technique of data augmentation in order to obtain the full conditional distribution of $\Gamma(ds)$ in known form. For each of these $x_n$’s select an additional random variable $s_n \in S$, i.e., a point in the auxiliary space $S$, where

$$P(s_n) = \frac{k(x_n, s_n)\Gamma(ds_n)}{\sum_n k(x_n, s_n)\Gamma(ds_n)}.$$ 

We now have pairs of points $(x_n, s_n)$, $x_n \in \mathcal{X}, s_n \in S$. We introduce a new random measure $Z$ on $\mathcal{X} \times S$ such that

$$Z(dx, ds) = \sum_n \delta_{(x_n, s_n)}(dx, ds).$$

Note that $Z(\mathcal{X} \times S) = N_3$. Let $Z_1(dx) = Z(dx \times S) = N_3(dx)$, i.e. $Z_1(dx)$ recovers the unaugmented data and let $Z_2(ds) = Z(\mathcal{X} \times ds)$, i.e. $Z_2(ds)$ recovers the augmented data. The following result now holds:

$$P(\Gamma(ds)|N_1, N_2, Z_2(S), \beta_1, \beta_2) \sim \text{Gamma} \left( \alpha(ds) + Z_2(ds), b(s) + \int_{\mathcal{X}} k(x, s)\omega(dx) \right).$$  \hfill (6)

Thus we augment each crack location and attribute vector $(x = (l, a))$, that we attribute to the latent source, with a point in the space $S$ and this augmentation step allows us to obtain the full conditional distribution for the Gamma random field. This full conditional distribution is also of Gamma form and all that is now required is a means of simulating such a Gamma random field.

### 4.5 Inverse Lévy Measure Algorithm

A Gamma random field can be simulated using the Inverse Lévy Measure (ILM) algorithm, see Wolpert and Ickstadt (1998a) and Wolpert and Ickstadt (1998b) for full details. The ILM algorithm is based on an idea regarding characteristic functions of infinitely-divisible distributions and particular positive measures which are termed “Lévy measures” and on the idea that a Gamma process can be constructed from a Poisson process. The algorithm can be used to draw random samples from Gamma and other non-negative independent-increment random fields.

Suppose we want to sample from the Gamma random field $\Gamma(ds) \sim \text{Gamma}(\alpha(ds), b(s))$, over the set $S$. A single realization of this Gamma random field will be discrete, and will consist of countably many point masses of random magnitudes $\gamma_m$ at locations $s_m \in S$. The theory allows for the sample locations to be drawn from any distribution $\Pi$, provided that whenever a set $A$
exists such that $\alpha(A) > 0$ then $\Pi(A) > 0$. For example it is possible to exploit information about where points associated with the latent variables would be expected to lie and to sample heavily from those areas. If such information is not available then the Uniform distribution on $S$ is a good choice from which to sample. The following is the ILM algorithm to sample from a Gamma random field $\Gamma(ds) \sim \text{Gamma}(\alpha(ds), b(s))$:

1. Set $M$ to be some large integer.
2. Choose a distribution $\Pi(ds)$ on $S$ from which it is easy to sample.
3. Generate $M$ independent identically distributed draws $\{\sigma_m\}$ from $\Pi(ds)$.
4. Generate the first $M$ jump times $\{\tau_m\}$ of a standard Poisson process; to do this simulate $M$ independent exponential random variables $\{\epsilon_m\}$ and set $\tau_m = \sum_{i=1}^{m-1} \epsilon_i$.
5. Set $\gamma_m = E_1^{-1}\{\tau_m/\alpha(\sigma_m)\}b(\sigma_m)$, for $m = 1, \ldots, M$, where $E_1(t) = \int_t^\infty e^{-u}u^{-1}du$ denotes the exponential integral function, see Abramowitz and Stegun (1964), pg. 228. See Wolpert and Ickstadt (1998b) for details on how to approximate this function and its inverse.
6. $\Gamma(ds) \approx \Gamma_M(ds) = \sum_m \gamma_m \delta_{\sigma_m}(ds)$.

Figure 3 shows a single realisation of a Gamma random field.

### 4.6 MCMC Algorithm

The following is a brief outline of the MCMC algorithm used to sample from the posterior distribution in order to obtain samples for the parameters: $\beta_1, \beta_2, \rho$ and $\Gamma(ds)$.

1. Initialise $\beta_1^{(0)}, \beta_2^{(0)}, \rho^{(0)}, \Gamma^{(0)}(ijk) \approx \int_{S_i} k(l_{ijk}, s) \Gamma^{(0)}(ij(s))$ for each crack $k$ of window $j$ of specimen $i$, and set the iteration counter $r = 1$.
2. For each crack set $A^{(r)}_{ijk} = C_{ijk}\beta_1^{(r-1)} + T_{ijk}\beta_2^{(r-1)} + \Gamma^{(r-1)}(ijk)$.
3. For each crack, using Bernoulli random variables, indicate whether the crack is attributable to compression, tension, or latent factors. In so doing $N_{ij1}, N_{ij2}$ and $N_{ij3}$ are obtained.
4. For each crack attributable to latent factors simulate a corresponding location $\gamma_{ij} \in S$.
5. For each $\gamma_{ij}$ carry out a random-walk Metropolis step proposing a new location $\gamma_{ij\text{test}}$. Set $s_{ijk} = \gamma_{ij\text{test}}$ if the new step is accepted, otherwise $s_{ijk} = \gamma_{ij}$, $k = 1, \ldots, N_{ij3}$.
6. Using the ILM algorithm, simulate the full conditional distributions for each of the Gamma random fields. The full conditional distribution (Equation 6) has a shape parameter composed of both continuous ($\alpha(ds)$) and discrete ($Z_2(ds)$) parts. For the continuous component simulate $\gamma_{ijm}$ as in part (e) of the ILM algorithm. For the discrete component $\gamma_{ij1} = e_1 b(\gamma_1)$, $\gamma_{ijm} = (e_m - e_{m-1})b(\gamma_m)$, $m = 2, \ldots, N_{ij3}$.
7. Simulate $\beta_1^{(r)}$ and $\beta_2^{(r)}$ from their full conditional distributions (Equations 4 and 5).
8. Carry out a random-walk Metropolis step for $\rho$, proposing $\rho_{\text{test}}$, if accepted $\rho^{(r)} = \rho_{\text{test}}$, otherwise $\rho^{(r)} = \rho^{(r-1)}$.
9. $r = r + 1$. Repeat steps (b) through (i).

Step (c) is optional. Without this extra random-walk Metropolis step it is highly likely that the same augmentation points $\{\gamma_{ij}\}$ in the space $S$ will be continuously chosen and this can slow the convergence of the algorithm.
Table 1: Median and (5%, 95%) quantile estimates for the parameters $\beta_1$, $\beta_2$ and $\rho$ for various prior distributions for $\rho$.  

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\pi(\rho) \sim \text{Lognormal}(1, 1.4)$</th>
<th>$\pi(\rho) \sim \text{Lognormal}(0.05, 1.2)$</th>
<th>$\pi(\rho) \sim \text{Lognormal}(2, 1.2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0.015 (0.003, 0.033)</td>
<td>0.016 (0.004, 0.038)</td>
<td>0.017 (0.004, 0.036)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.031 (0.023, 0.038)</td>
<td>0.03 (0.022, 0.04)</td>
<td>0.03 (0.023, 0.039)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>1.78 (1.48, 2.04)</td>
<td>1.77 (1.48, 2.06)</td>
<td>1.78 (1.49, 2.05)</td>
</tr>
</tbody>
</table>

5 Results

We now present the results obtained from carrying out inference on the unknown parameters of the model using the MCMC algorithm detailed above. Table 1 shows median and quantile estimates and priors for $\beta_1$, $\beta_2$, and $\rho$. The estimates for these parameters are based on sample values obtained from the posterior distribution by running the MCMC algorithm. We computed 6,000 iterations of the program; the first 1000 of these iterations were attributed to burn-in. For each parameter we examined a trace plot of its chain and from this inspection there was no evidence for lack of convergence. We also ran multiple independent chains from various starting points and again, after an initial burn-in, there was no reason to believe the chains had not converged.

In Figure 4 we present in image form, the posterior mean of the Gamma random field for Specimen 1 and in Figure 5 we show the posterior mean intensity for Specimen 1 as an image plot, here the crack locations are also indicated. The corresponding plots for Specimens 2-5 may be found in Figures 8 and 9.

The median estimates for $\beta_1$ (0.015) and $\beta_2$ (0.031) are both smaller than the prior estimates ($E(\beta_1) = 10$, $E(\beta_2) = 30$) but are still consistent with the belief that tension, positive stress has a higher impact on crack initiation than does compression (negative stress). Prior to this analysis, the range over which the latent factors, such as pores, would have been thought to be influential was small. In our analysis we did allow the prior for $\rho$ to have a reasonably large range but the analysis suggests that the influential range is approximately 4mm, ($2\rho \approx 4$). The results for the Gamma random fields for each specimen are represented graphically. A first observation is that there is large specimen variation between the Gamma random fields. This is important since all specimens were identical (as much as was physically possible) and all specimens were subjected to the same stress loading in laboratory conditions, suggesting that factors other than those observed have contributed to this spatial variability. The Gamma random fields themselves give a spatial indication as to where crack initiation causing factors may lie. This information would allow for a closer re-examination of the specimens in order to investigate the presence of, for example, pores.

5.1 Model Validation

In order to explore the fit of the model we examine residuals as follows. For each of the regions in Figure 6(a), we calculate the standardised residuals $r_{ij}$ for region $j$ of specimen $i$ using the following approximation

$$r_{ij} = \frac{1}{T} \sum_{t=1}^{T} \frac{(N_{ij} - \Lambda(P_{ij})^{(t)})}{\sqrt{\Lambda(P_{ij})^{(t)}}},$$

where

$$\Lambda(P_{ij}) = \int_{P_{ij}} \Lambda(x)\omega(dx),$$

and $T$ is the total number of iterations. We plot these standardised residuals against the predicted counts

$$\hat{N}_{ij} = \frac{1}{T} \sum_{t=1}^{T} \Lambda(P_{ij})^{(t)}$$

in Figure 6(b). This plot suggests that, in general, there is no unmodelled trend in the residuals, which we would hope for if the model is a reasonable fit. Although there are a small number of
large residuals. When we further examine the residuals by plotting them against both specimen and window, in Figures 6(c) and 6(d) respectively, it appears that the large residuals are associated with Specimens 1, 2, and 4. There appears to be no distinction between windows as to how well the model fits.

We also examine the predictive count for each of the specimens. A plot of this analysis may be seen in Figure 7. In general the model appears to predict the count of cracks in each of the specimens well, as all of the true counts lie within the 90% quantiles, and are all close to the median predicted values. The model does allow a lot of flexibility and so we would expect a reasonable overall fit.

A prior sensitivity analysis was also carried out for the parameter $\rho$. We considered priors that would indicate both short (mean = 2.16) and longer (mean = 15.18), ranges of influence, columns 3 and 4 in Table 1, respectively. Our model does not appear to be sensitive to the prior choice for this important parameter as indicated by the estimates for both $\rho$ and the other parameters.

6 Conclusions

We have estimated the impact that both compression and tension have on causing cracks to initiate by explicitly modelling these covariates and obtaining estimates of the parameters $\beta_1$ and $\beta_2$. From the analysis carried out, there is evidence that unmeasured, spatially varying factors have an impact on crack initiation. The influence of these unmeasured factors is modelled using the Gamma random fields. The range over which the spatially varying factors exert an influence appears to be short (< 4mm).

The identification of such important factors is helpful for the understanding of why damage accumulation is so variable. Specimens that were subjected to identical stress loads under laboratory conditions did not present the same damage accumulation patterns. The knowledge that such factors exist and the identification of them as crack causing, together with the estimation of the range over which they have an influence can lead to strategies to identify the physical causes. The graphical representations of the Gamma random fields for each of the specimens could be compared by the engineers, with the actual experimental specimens to check for evidence of unmeasured influential factors, such as pores. It may then be possible to either eliminate or reduce these factors in order to decrease the amount of damage accumulation and ultimately prolong the lifetime of the hip replacement.

A continuous model allows for more accurate estimates of the parameters but at the cost of locating each individual crack. If collecting crack counts on a discrete grid facilitates the replication of more data than would locating each individual crack, then using a model based on counts of cracks with more data would be preferable. The advantage of the continuous model is that it allows for more accurate estimates of the range over which the latent parameters exert an influence due to the fact that no aggregation of the data need take place.
Figure 4: Image plot indicating the posterior mean of the Gamma random field over the two windows of Specimen 1. The posterior mean is calculated on a $100 \times 100$ grid and the Gamma random field in each grid segment $G$ is obtained by calculating $\int_G \Gamma(ds) \approx \sum_{m \in G} \{ \gamma_m : \sigma_m \in G \}$ at each iteration of the program and averaging over the iterations.

Figure 5: Posterior mean crack density for Specimen 1 together with actual crack locations.
Figure 6: (a) The regions for which the residuals were calculated. The standardised residuals against the posterior mean crack count (b), against specimen (c), and against window (d).
Figure 7: For each specimen we show the actual total crack count (cross), and the median posterior predicted count (circle), together with 90% quantiles for the predicted counts.
Figure 8: Image plots indicating the posterior mean of the Gamma random field over the two windows for Specimens 2-5. The posterior mean is calculated on a $100 \times 100$ grid and the Gamma random field in each grid segment $G$ is obtained by calculating $\int_G \Gamma(ds) \approx \sum_{m} \{ \gamma_m : \sigma_m \in G \}$ at each iteration of the program and averaging over the iterations.
Figure 9: Posterior mean crack density for Specimens 2-5 together with actual crack locations for each specimen being indicated.
References


