

Extreme Value Statistics in Metal Fatigue

Statistica dei Valori Estremi e Resistenza a Fatica dei Materiali Metallici

Stefano Beretta, Politecnico di Milano

Dipartimento di Meccanica,

Politecnico di Milano, via La Masa 34, 20158 Milano, Italy

stefano.beretta@polimi.it

Clive Anderson, University of Sheffield

Dept. of Probability & Statistics,

University of Sheffield, Sheffield S3 7RH, UK

C.W.Anderson@shef.ac.uk

Riassunto: La resistenza a fatica dei materiali metallici è influenzata dalla eventuale presenza di cricche o difetti. In particolare il cedimento è determinato dalla non-propagazione delle microcricche che si formano all'apice del difetto massimo. Queste evidenze hanno suggerito l'adozione di metodi per il conteggio di inclusioni e difetti basati sulla statistica dei valori estremi. In questa memoria si espongono i concetti di base di tali applicazioni della statistica dei valori estremi, illustrando i metodi già consolidati ed i punti aperti delle ricerche in corso.

Keywords: Gumbel, Generalized Pareto, Threshold Methods, Stereology, Wicksell Transform

1. Introduction

The strength of metallic materials under cyclic stress, called 'fatigue', is dramatically affected by the presence of defects and non-metallic inclusions which act as starter for fractures that eventually lead to failures. The economical and safety implications of this issue can be easily understood considering that in U.S. some \$177 billion is spent in repairing or preventing fractures (Duga 1983). Considering that there is a definite relationship between fatigue strength and defect size, it would be worth estimating the size of maximum inclusion (or defect) which could be the origin of a prospective fatigue failure. Since in general current non destructive detection techniques are only able to reliably detect inhomogeneities in the range 0.5-1 mm (Simonen 1995), given also the unfeasibility on the complete destruction of a metal piece or component for finding its largest inclusion, the estimation of inclusions (or defects) present can be only achieved with statistical analysis, particularly extreme value analysis, coupled with a suitable sampling strategy.

The aim of this paper is to address this topic by first showing the experimental evidences which supports the application of statistics of extremes (Section 2), the current applications and new results in inclusion rating (Section 3). The attention is then paid to a detailed discussion of statistical implications, namely, the stereological problem of the estimation of a defect size in a 3D volume with the open problems (Section 4), the fatigue strength of a real component (Section 5).

2. Experimental evidences about fatigue strength

The term ‘fatigue’ refers to the fact that a cyclic loading on a component can determine its failure even if maximum applied stress is below the static strength limit of the material. The key factor is the applied stress range: the higher the applied stress, the lower is the number of cycles a component can stand. For any material there is a minimum stress range below which a material can endure for an indefinite number of cycle: this is called ‘fatigue limit’ (or ‘fatigue strength’).

A huge number of fatigue experiments carried out by Murakami and co-authors (Murakami & Endo 1986, Murakami 1993) have shown that at stress levels near the fatigue limit small cracks can be found at the tip of micro-holes introduced onto the specimens before testing. These experimental evidences support the concept that fatigue limit of materials containing defects and inhomogeneities is not a critical stress for crack initiation but the threshold stress for the nonpropagation of small cracks emanating from the original notch or inhomogeneities (Fig. 1).

Moreover, because of the presence of these small cracks, defects can be treated as cracks. Their stress intensity factor range can be calculated with the simple equation:

$$\Delta K = 0.65 \cdot \Delta S \cdot \sqrt{\pi \cdot \sqrt{area}} \quad (1)$$

where $\sqrt{area}$ (square root of the projected area of defect) is the geometric parameter expressing defect (or crack) size and ΔS is the applied stress range (Murakami 1985).

Experiments typically show that the fatigue limit increases with decreasing defect size up to the theoretical fatigue limit of the material, which corresponds the cyclic yield stress of the material (Miller 1993). Fig. 2 shows the typical relationship between fatigue limit and crack (or defect) size obtained for many different materials: this type of graph which is called Kitagawa-Takahashi diagram (Kitagawa & Takahashi 1976). Different models (Murakami & Endo 1994) are able to estimate the fatigue limit at a given defect size in terms of the cyclic stress range at which ΔK is equal to a threshold value for non propagation ΔK_{th} .

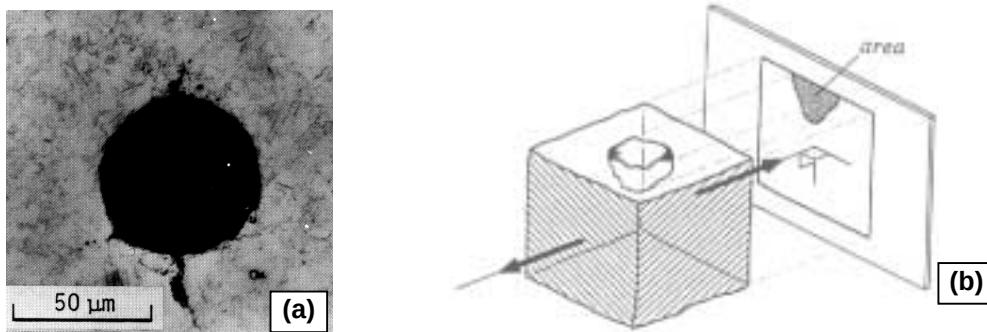


Figure 1: Fatigue limit in presence of defects: a) non-propagating cracks observed at fatigue limit of a specimen containing a micro-hole; b) the $\sqrt{area}$ concept (Murakami & Endo 1994).

The crack size a_0 , also called ‘El-Haddad parameter’, which corresponds to the kink in the diagram of Fig.2, gives an idea of the defect sizes affecting the fatigue strength of different metals: it ranges from 60 μm for high strength steels to 300 μm for Al alloys.

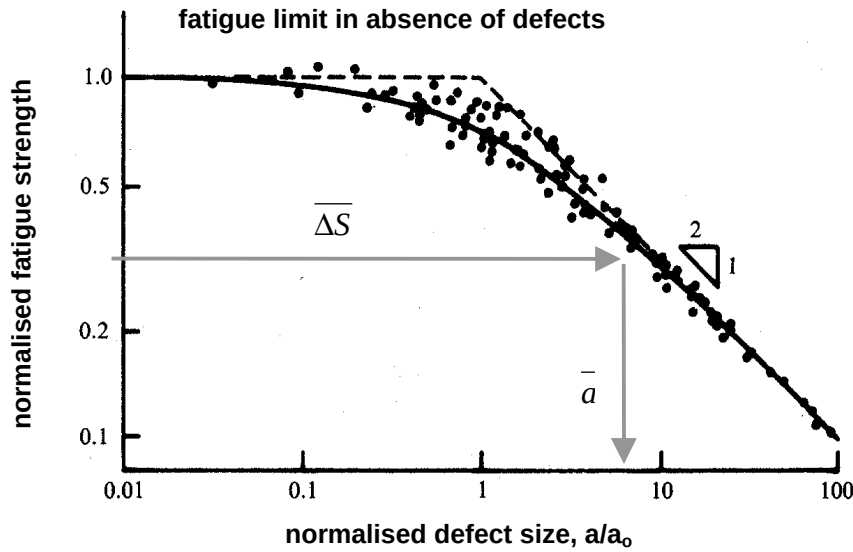


Figure 2: *The relationship between fatigue limit and crack size.*

Considering Fig. 2, it can be easily seen that at a given stress level ΔS the defects (or cracks) smaller than the limit value $\bar{a}$ (also called ‘mechanical threshold’) do not cause failure. It is then natural to think that, under uniform applied stress, is the maximum defect (or crack) present which determines fatigue failure.

3. The Ratings Problem

3.1 The problem

The ratings problem is the problem particularly facing steel manufacturers and their customers. It is how to assess the quality of production batches of steel with respect to the sizes of inclusions. An important requirement of any method of rating steels is that it should rely on measuring techniques that can be routinely implemented in the factory and it should be simple to understand and apply.

Current rating methods are in general based on assessing the quality of a steel by comparing inclusions observed on polished plane sections of the steel with predetermined maps or quality indexes. The problems of these methods essentially are: i) quality indexes have been elaborated on given steel qualities and they so unable to appreciate the cleanliness of modern super-clean steels; ii) standard concepts of metallurgical quality are related to average content of contaminants (e.g. oxygen, sulphur) but this has no direct relationship with the dimension of extreme particles; iii) it has been proven that current rating practices are not related with the resulting fatigue properties of a steel (Murakami 1993).

Because of these inabilities in recent years there has been an effort to establish new ‘fatigue oriented’ rating methods based on the statistics of extremes. Murakami (Murakami 1994) was the first who suggested basing the rating of steels on observed characteristics of the largest inclusion in a plane section of fixed area (the control area S_{ref}). A sample of observations of maximum sizes (taken as $\sqrt{area}$) of inclusions from multiple control areas is obtained and is found to fit well to the Gumbel extreme value distribution.

Specifically, the extreme value distributions arise as the possible limit distributions for the maximum M_n of n independent random variables under linear normalization. It can be demonstrated that for suitable normalizing constants $a_n > 0$ and b_n :

$$P((M_n - b_n)/a_n \leq x) \rightarrow G(x), \text{ as } n \rightarrow \infty \quad (2)$$

where G is a proper distribution function, called Generalized Extreme Value distribution, whose expression is:

$$G(x) = \exp(-(1 + \xi(x - \lambda)/\delta)^{-1/\xi}) \quad (3)$$

with parameters $-\infty < \xi < \infty$, $\delta > 0$ and λ , and where the range of x is $-\infty < x < \lambda - \delta/\xi$ when $\xi < 0$, $\lambda - \delta/\xi < x < \infty$ when $\xi > 0$, and $-\infty < x < \infty$ when $\xi = 0$. The special case when $\xi = 0$ is the Gumbel distribution whose distribution function is:

$$G(x) = \exp(-\exp(-(x - \lambda)/\delta)), -\infty < x < \infty. \quad (4)$$

A probability plot (called a Gumbel plot) to assess the fit of the Gumbel distribution to data is obtained by plotting $-\ln(-\ln(i/(n+1)))$ against the i_{th} data value, $x_{(i)}$ say, arranged in increasing order ($x_1 \leq x_i \leq x_n$). A straight line indicates a good fit to the distribution (1), the slope of the line giving a graphical estimate of $1/\delta$ and the intercept of $-\lambda/\delta$. More objective maximum likelihood estimates of the parameters λ and δ are straightforward to find numerically: statistical properties and recommended minimum sample sizes have been discussed by Beretta & Murakami (1998).

The analysis of inclusions is usually applied for extrapolating measurements obtained on small areas S_{ref} to estimate the size of maximum defects in much bigger areas. The *characteristic size of the maximum inclusion* (or defect) which can be found in a prospective area S_p is:

$$x_{\max}(S_p) = \lambda - \delta \cdot \ln(-\ln(1 - 1/T)) \quad (5)$$

where $T = S_p/S_{ref}$ is the return period of the maximum defect in S_p .

By adopting the statistics of extremes concepts it is simple to discriminate among different steels or to determine the acceptability of different batches of the same material (Fig. 3).

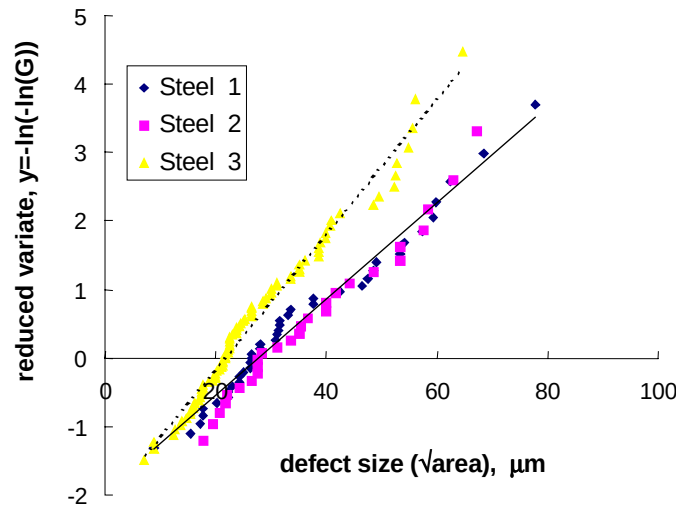


Figure 3: Gumbel probability plots of the maximum inclusions in three different spring steels (Beretta et al. 2001).

Let us consider for example the distributions of maximum inclusions found in three types of steel, almost equivalent in terms of fatigue properties, which were candidate for producing large train suspension springs (Beretta et al 2001). Traditional inclusion ratings were not able to support the choice of one of the steels. Extreme value inclusion rating solved the problem: if apparently the maximum sampled defects have a similar average, it appears that on large areas Steel 1 will have smaller maximum defects (Fig. 3) and consequently better fatigue properties.

An alternative to the above is to measure all inclusions larger than a fixed threshold size in an area (*Peak Over Threshold* method), and to fit a Generalized Pareto distribution to these measurements (Shi, Atkinson, Sellars and Anderson 1999). The assumption on which it rests is that if X denotes the size of a randomly selected inclusion then

$$P(X \leq x | X > x_0) = 1 - \{1 + \xi(x - x_0)/\sigma\}^{-1/\xi} \quad (6)$$

for a suitably high threshold x_0 and for parameters σ and ξ . The distribution in (6) is the *Generalized* Pareto distribution. It reduces to an exponential distribution when the parameter $\xi = 0$. There is a close relation to the extreme value distributions in (3): if defects of size greater than x_0 occur according to a Poisson process of rate ν and the sizes of defects are distributed independently according to (6), then the distribution of the largest defect in a given area or volume is of Generalized Extreme Value form (3). The parameter ξ takes the same value in both distributions, and so the Gumbel distribution corresponding to the value $\xi = 0$ in (3) corresponds to an exponential GP distribution in (6).

The fitting of a Generalized Pareto distribution (6) can therefore yield estimates of the distribution of the size of the largest inclusion in an area, and is able to utilise more information in doing so, as shown by Anderson et al. (2000) and by Shi et al. (2001). It does however require a greater effort in measurement.

3.2 A Complication

Whilst the approaches above work well in many cases, in some steels they appear to fail. Let us consider two sets of observations collected on a carbon steel using two different control areas $S_1 = 0.384 \text{ mm}^2$ and $S_2 = 66.37 \text{ mm}^2$ by Beretta & Murakami (2001): the two data sets appear to belong to Gumbel distribution but the slope of data in the Gumbel plot is different (Fig. 4.a).

The distribution function of maximum inclusions in S_2 can be derived from distribution of maximum particles in S_1 with the transformation $G_{S_2} = (G_{S_1})^{S_2/S_1}$. The estimated distribution of maximum inclusions from S_1 data is absolutely different from real observations on S_2 (Fig. 4.a). Alternatively, if data on S_2 are transformed to S_1 areas, the Gumbel plot shows a peculiar kink (Fig. 4.b). Further investigation revealed two chemically-different types of maximum particle found on the two control areas.

How can such data be described? Are there plausible models for inclusion occurrence and size that can explain the finding that only one kind of defect is observed when small control areas are inspected, yet a larger type of inclusion is seen when larger control areas are inspected?

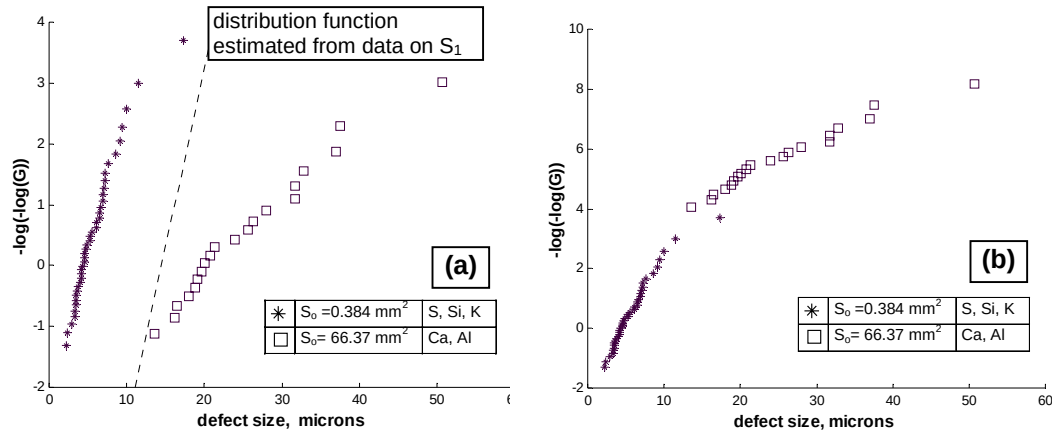


Figure 4: Defects collected on a carbon steel: a) Gumbel plot of inclusions data collected on control areas $S_1 = 0.384 \text{ mm}^2$ and $S_2 = 66.37 \text{ mm}^2$; b) the two data sets transformed at control area S_1 (Beretta & Murakami 2001).

Recent research (Anderson, Beretta & Murakami 2002) has considered two alternative models, based respectively on mixed distributions and competing risks, to account for the empirical findings outlined above.

The mixture model, first proposed for the analysis of defects by Beretta & Murakami (2001), is suggested by the form of the Gumbel plot in Fig. 4.b, and by there being two types of inclusion. It supposes that the distribution function F_{mix} of the maximum size M_{ref} of inclusion observed in a reference area S_{ref} is

$$F_{mix}(x) := P(M_{ref} \leq x) = (1-p)G_1(x) + pG_2(x) \quad (9)$$

where $G_i(x) = \exp(-\exp((x - \lambda_i)/\delta_i))$ for $i=1,2$. The G_i are interpreted as the distribution functions of the maximum sizes of inclusions of type i observed in area S_{ref} , and p is a mixing fraction. The distribution function of the size, M_0 say, of the largest inclusion observed in a different area S_0 is then taken to be:

$$P(M_0 \leq x) = (F_{mix}(x))^{S_0/S_{ref}} \quad (10)$$

In the competing risks model it is assumed that the largest inclusion of type 1 within a reference area S_{ref} , M_1 say, follows a Gumbel distribution G_1 , and the largest inclusion of type 2, M_2 say, within the same area, a Gumbel distribution G_2 . If occurrence and sizes of the two types of inclusion are independent, then the size M_{ref} of the largest inclusion of either type observed in S_{ref} , being the larger of M_1 and M_2 , has distribution function:

$$F_{crisk}(x) := P(M_{ref} \leq x) = G_1(x)G_2(x) \quad (11)$$

where, as before, $G_i(x) = \exp(-\exp((x - \lambda_i)/\delta_i))$ for $i=1,2$.

For a different area S_0 , the distribution function of the size M_0 of the largest inclusion observed in this area is, as before, taken to be

$$P(M_0 \leq x) = (F_{crisk}(x))^{S_0 / S_{ref}} \quad (12)$$

A useful interpretation of the parameters in the second model may be derived from a simple description of the random occurrence of defects of the two types. Under the hypotheses that defects of two types: i) occur at the points of independent spatial Poisson processes with mean numbers of inclusions per unit area ρ_1 and ρ_2 respectively; ii) have sizes which are independent exponentially distributed variables with means δ_1 and δ_2 respectively, then the distribution function of inclusion size in S_{ref} is:

$$\begin{aligned} F_{crisk}(x) &= \exp(-\exp(-(x - \delta_1 \ln \rho_1) / \delta_1)) \cdot \exp(-\exp(-(x - \delta_2 \ln \rho_2) / \delta_2)) \\ &= G_1(x)G_2(x) \end{aligned} \quad (13)$$

$$\text{where: } \lambda_i = \delta_i \ln \rho_i \quad \text{for } i = 1, 2. \quad (14)$$

It is easily seen that both models can account for the observed data and give very similar results with almost identical precision (Fig. 5), increasing confidence in their robustness. Next steps on this topic will be: i) to derive POT analyses able to account for the presence of two defect types; ii) to suggest guidelines for the control areas to use in measuring, and to compare different steels.

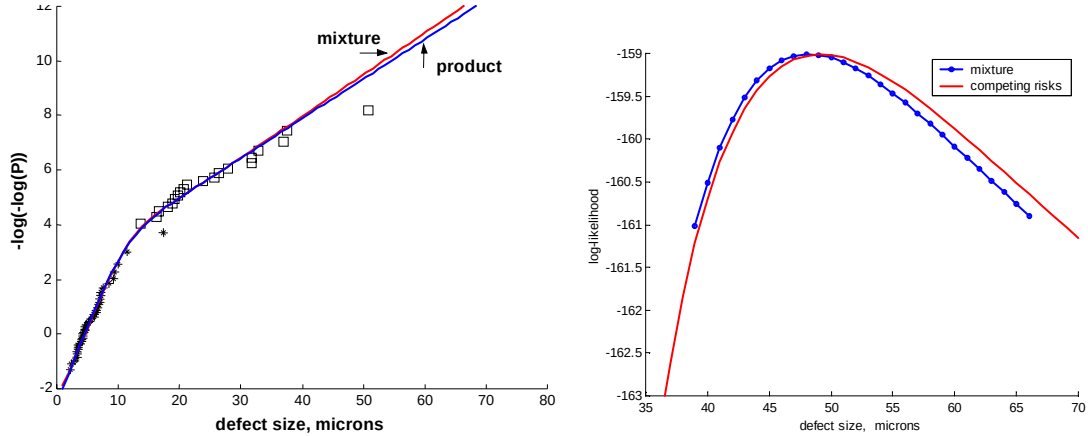


Figure 5: Comparison of mixture and competing risks models on carbon steel data: a) distribution function plotted onto Gumbel plot; b) likelihood profiles for the defect with $T=10000$ (Anderson, Beretta & Murakami 2002).

4. The Three Dimensional Problem

The above yields inferences and practical procedures for the maximum size of 2-dimensional sections of inclusions. However, fatigue theory in fact suggests that it is the inclusion with the largest projected area ($\sqrt{\text{area}}$) that is crucial for fatigue properties. Section areas and projected areas are not necessarily the same, so if we wish to use 2-dimensional section measurements to make inferences about large inclusions (large in the sense of having large projected areas) we need to take the stereological relationship

between 2- and 3-dimensional measurements into account. Moreover the aim is to make inferences about the largest inclusion in a large volume of the steel, so there is an extrapolation issue too. Several authors have studied this problem [Drees & Reiss (1992), Takahashi & Sibuya (1996, 1998, 2001), Anderson & Coles (2002), Joosens & Beirlant (2001)]. We outline two broad approaches.

4.1 Direct Approach

Experimental evidence by Murakami (1994) suggests that the difference between the distribution of the maximum section size and maximum projection size is small. It is argued too that observations of inclusions revealed in a plane section of steel are equivalent in some sense to the result of sampling a slice of the steel corresponding to a slab of a certain thickness below the measured plane surface. Murakami (1990) suggested taking this thickness to be the average, $\bar{d}$ say, of the sizes (square root areas) of the largest inclusions observed in a sample of control areas. The resulting effective volume sampled is $V_0 = \bar{d} \times S_{ref}$, where S_{ref} is the control area. In a larger volume V of steel the largest inclusion, M_v say, is the maximum of the separate maxima in sub-volumes of size V_0 , and so (assuming independence and homogeneity) will have distribution function

$$P(M_v \leq x) = (P(M_{V_0} \leq x))^{V/V_0} \quad (15)$$

4.2 Threshold Approach

In this approach we assume that the conditional upper tail of the distribution of 3-dimensional projection sizes is of Generalized Pareto form (4). We assume also that observations are available on all 2-dimensional sections of inclusions larger than a threshold size in a given area of polished section. This is a natural assumption, but it introduces the difficulty for inference that the observations are not themselves the subject of the model assumptions. The link between the two distributions (of section-size and projection-size) is provided by Wicksell's formula (Wicksell 1925) if inclusions may be supposed spherical. Let G denote the distribution function of the projection diameter, V say, of inclusions, and let F denote the distribution function of the diameter of sections, S say, of inclusions cut by a random plane. Then Wicksell's formula is:

$$P(S > s) = 1 - F(s) = \frac{1}{E(V)} \int_s^\infty (v^2 - s^2)^{1/2} dG(v). \quad (16)$$

From (10) it is simple to show that the conditional distribution of section sizes S given that they are larger than a value v_0 may be expressed in terms of conditional distribution function G_{v_0} of projection diameters given they are larger than v_0 as:

$$P(S \leq s | S > v_0) = 1 - \frac{\int_s^\infty (v^2 - s^2)^{1/2} dG_{v_0}(v)}{\int_{v_0}^\infty (v^2 - v_0^2)^{1/2} dG_{v_0}(v)}. \quad (17)$$

If a Generalized Pareto form (6) is assumed for G_{V_0} , (17) leads to a likelihood function based on observations of large section sizes, and hence inference. Numerical maximum likelihood fitting is feasible, though computationally heavy. An attractive alternative (Anderson & Coles 2001) is a Bayesian MCMC approach, which has the advantage that with little extra effort it produces predictive distributions of variables of interest, for example the projection size of the largest inclusion in a specified large volume. The predictive distribution embodies both estimation uncertainty and uncertainty due to the random nature of the occurrence of inclusions.

Other open points that need further research are:

- The effect of non-sphericity on the inference, since inclusions are rarely spherical;
- The Wicksell formula shows that maxima of projection and section sizes cannot both follow Gumbel distributions, so the assumptions made in §3 and §4 are logically inconsistent but perhaps they represent reasonable numerical approximations in most practical situations.
- The methods developed to date have assumed that inclusions occur randomly and homogeneously throughout the material. However, inclusions are sometimes observed to cluster, and models to represent this feature are of much interest.

5. Component Design

Given the geometry of a steel component and knowledge of the stress it will be subject to in use, what is the probability that an inclusion within it will lead to fatigue failure? Internal stresses within the component will generally vary with position, and a fatal crack will develop if any inclusion within the component is too large to sustain the local stress where it is located. It will not necessarily be the largest inclusion on the component at which failure occurs; indeed it would be an unlucky coincidence if the highest stress happened to occur at the weakest point within the component. A practical 'rule of thumb' in such problems is to calculate the $csmi$ in the most stressed volume, the so called '90% volume' where the stress S is $0.9 \cdot S_{\max} < S < S_{\max}$ (Murakami 1994).

However, knowledge of the size distribution of all large inclusions (not just the largest particle in the component) together with knowledge of the relation between the stress sustainable by defects of different sizes (the Kitagawa-Takahashi diagram) enables estimates of failure probability to be made for any given internal stress field. Yates, Shi, Atkinson, Sellars & Anderson (2002) show that the calculation may be expressed rather simply by formulating the model as a marked spatial Poisson process of large defects, marks corresponding to size. They use the Generalized Pareto distribution for sizes exceeding a threshold. The resulting methodology is of great potential value to designers, since it allows the consequences of changes in geometry or material specification to be explored, and, by appropriate choice of sampling scheme, it can take into account the rare dangerous defects described in §3.2.

References

- Anderson C.W. and Coles S.G. (2001) The largest inclusion in a piece of steel. submitted to *Extremes*.
- Anderson C.W., Shi G., Atkinson H.V. and Sellars C.M. (2000) The precision of methods using the statistics of extremes for the estimation of maximum size of inclusions in clean steels, in: *Acta Materialia*, 48, 4235-4246.

- Anderson C.W., Beretta S., Murakami Y. (2002) Assessment of clean steels containing multiple types of inclusions, submitted to *Acta Materialia*.
- Beretta S. and Murakami Y. (1998) Statistical analysis of defects for fatigue strength prediction and quality control of materials, in: *Fatigue Fract. Engng. Mater. Struct.*, 21, 1049-1065.
- Beretta S. and Murakami Y. (2001) Largest Extreme-Value Distribution Analysis of Multiple Inclusion Types in Determining Steel Cleanliness, in: *Metall. Materials Trans.*, 32B, 517-523.
- Beretta S., Blarasin A., Clerici P., Giunti T. (2001) Metodologia 'Defect Tolerant' per la previsione del limite di fatica di acciai per molle in condizioni di carico multiassiali, in: *Proc. XXX Convegno Naz. AIAS*, Alghero.
- Drees H. & Reiss R.D. (1992) Tail behaviour in Wicksell's corpuscle problem, in *Probability & Applications: Essays to the Memory of J Mogyorodi (eds. J. Galambos & I. Katai)*, Kluwer, Dordrecht, 205-220.
- Duga J.J. et al. (1983) The economic effect of fractures in the U.S., in: *Natl. Bureau Standards Special Publication* 647-2.
- Kitagawa H. and Takahashi S. (1976) Applicability of fracture mechanics to very small cracks or cracks in the early stage, in: *Proc. 2nd Int. Conf. Mechanical Behaviour of Materials*, Boston, 627-631.
- Joossens, E. & Beirlant, J. (2001) On the estimation of the largest inclusions in a piece of steel using extreme value analysis. Pre-print, Dept of Mathematics, KU Leuven.
- Miller K.J. (1993) The two thresholds of fatigue behaviour, in: *Fatigue Fract. Engng. Mater. Struct.*, 16, 931-939.
- Murakami Y. (1985) Analysis of Stress intensity Factors of mode I, II and III for inclined surface cracks of arbitrary shape, in: *Engng. Fract. Mech.*, 22, 101-114.
- Murakami Y. (1993) *Metal Fatigue: Effect of Small Defects and Nonmetallic Inclusions*, Yokendo Ltd, Tokyo.
- Murakami Y. & Endo M. (1994) Effect of defects, inclusions and inhomogeneities on fatigue strength, in: *Int. J. Fatigue*, 16, 163-182.
- Murakami Y. (1994) Inclusion rating by statistics of extreme value and its application to fatigue strength prediction and quality control of materials, in: *J. Res. Natl. Standard*, 99, 345-351.
- Shi G., Atkinson H.V., Sellars C.M and Anderson C.W. (1999) Application of the Generalized pareto Distribution to the estimation of the size of maximum inclusion in clean steels, in: *Acta Materialia*, 47, 1455-1468.
- Shi G., Atkinson H.V., Sellars C.M., Anderson C.W. and Yates J.R. (2001) Computer simulation of the estimation of the maximum inclusion size in clean steels by the Generalized pareto distribution, in: *Acta Materialia*, 49, 1813-1820.
- Simonen F.A. (1995) Nondestructive examination reliability, in: *Probabilistic Structural Mechanics Handbook*, C. Sundararajan Ed., Chapman & Hall, New York.
- Takahashi R. and Sibuya M. (1996) The maximum size of planar sections of random spheres and its application to metallurgy, in: *Ann. Inst. Statist. Math.*, 48, 127-144.
- Takahashi R. & Sibuya M. (1998) Prediction of the maximum size in Wicksell's corpuscle problem, in: *Ann. Inst. Statist. Math.*, 50, 361-377.
- Takahashi R. & Sibuya M. (2001) Prediction of the maximum size in Wicksell's corpuscle problem, II, to appear in: *Ann. Inst. Statist. Math.*
- Yates, J.R., Shi, G., Atkinson, H.V., Sellars, C.M. & Anderson, C.W. (2002) Fatigue tolerant design of steel components based on the size of large inclusions. To appear in *Fatigue & Fracture Engineering of Materials & Structures*