Computational Materials Science 24 (2002) 430–442

COMPUTATIONAL MATERIALS SCIENCE

www.elsevier.com/locate/commatsci

Statistics of defects in one-dimensional components

M.T. Todinov *

Reliability Engineering and Risk Management Centre, Cranfield University, Cranfield MK 43 ØAL, UK

Received 14 May 2001; accepted 20 November 2001

Abstract

Equations related to spatial statistics of defects and probability of detecting defects in one-dimensional components have been derived. The equations related to spatial statistics of defects allow to estimate the probability of existence of safe, defect-free zones between the defects in one-dimensional components. It is demonstrated that even for a moderate defect number densities, the probability of existence of clusters of two or more defects at a critically small distance is substantial and should not be neglected in calculations related to risks of failure. The formulae derived have also important application in reliability and risk assessment studies related to calculation of the probability of clustering of evens on a given time interval. It is demonstrated that while for large tested fractions from one-dimensional components, the failures are almost entirely caused by a small part of the largest defects, for small tested fractions almost all defects participate as initiators of failure. The problem of non-destructive defect inspection of one-dimensional components has also been addressed. A general equation has been derived regarding the probability of detecting at least a single defect when only a fraction of the component is examined. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Probability; Statistics; Structure; Properties; Reliability; Risk assessment; Inhomogeneous; Fracture; Defects; Non-destructive; Inspection

1. Introduction

An essential part of the statistics of the structure and properties of inhomogeneous materials is the statistics of defects in one-dimensional components such as welds, wires, fibres, rods, profiles, etc.

Important issues of the statistics of defects in one-dimensional components are (i) calculating the probability of failure controlled by defects, (ii) calculating the probability of detecting at least a single defect by inspecting part of the component and (iii) the spatial statistics of the defects.

For the probability of failure of a chain consisting of 'n' links, Weibull [1] proposed the expression:

$$\xi(z) = 1 - \exp[-n\varphi(z)] \tag{1}$$

where the only conditions $\varphi(z)$ has to satisfy are to be a positive, non-decreasing and vanishing at some value z_0. After approximating $\varphi(z)$ with the simplest function satisfying this condition: $\varphi(z) \approx [(z - z_0)/z_a]^\beta$ the cumulative Webull distribution function is obtained:

* Address: School of Industrial and Manufacturing Science, Reliability Engineering and Risk Management Center, Cranfield University, Cranfield MK 43 ØAL, UK. Tel.: +44-123-475-0111/4124; fax: +44-121-414-5232.

E-mail address: m.t.todinov@ac.uk (M.T. Todinov).

PII: S0927-0256(01)00264-6

$$\xi(z) = 1 - \exp\left[-\left(\frac{z - z_0}{z_a}\right)^{\beta}\right] \qquad (2)$$

The probability of fracture $\xi(z)$ at a load (stress) 'z' is commonly described by the cumulative distribution function (2) where z_0, z_a and 'β' are the location, scale and shape parameter respectively [2–5]. In Ref. [6], it was shown that the function $\varphi(z)$ in Eq. (1) should be of the form $\varphi(z) = |\ln[1 - \psi(z)]|$ in order to predict correctly the probability of failure of a chain.

In materials and components, failure is often triggered by defects [7–14], such as microcracks, inclusions, surface discontinuities or particles which act as links in a chain. According to the weakest link principle, fracture controlled by defects is initiated when the conditions for triggering fracture are fulfilled for at least a single defect interior to the stressed region. Thus, in Refs. [14–21] the probability of triggering fracture by defects has been quantified by equations similar to Eq. (1). In earlier work [6,22], it was demonstrated that these equations predict correctly the probability of failure for a number density of the defects calculated on the basis of large (practically unlimited) calibration length compared to the stressed length. As an alternative, a new theory pertinent to probability of fracture controlled by defects [6,22], was proposed, according to which if in a wire of length L, containing a given number of defects, a smaller length ΔL is stressed to a tensile stress σ, the probability of fracture $S(\sigma)$ is given by

$$S(\sigma) = 1 - \exp\{\mu L \ln[1 - pF(\sigma)]\} \qquad (3)$$

where $\mu = n/L$ is the number density of the defects calculated on the basis of the calibration length L, n is the number of defects in the calibration length and $p = \Delta L/L$ is the ratio of the stressed length ΔL to the entire calibration length L. $F(\sigma)$ in Eq. (3) is a cumulative probability of triggering fracture associated with the defects, i.e. the probability that a randomly chosen defect from the population will trigger fracture at a stress smaller than or equal to σ. The length L was referred to as 'calibration length', necessary to determine the defect number density μ. In this sense, the length L contains with certainty particular types and numbers of defects. An important feature of the theory is that it makes allowance for the effect of the size of the stressed region on probability of fracture.

It was also demonstrated that the predictions from the existing statistical theory of fracture deviate significantly from the exact values resulting from the new theory (Fig. 1, [22]). The discrepancies between the predictions from the new theory and the existing theory were larger the smaller the number density of the defects (Fig. 1).

If the defects trigger fracture with certainty at a specified loading stress σ ($F(\sigma) = 1$, all defects are critical) Eq. (3) becomes

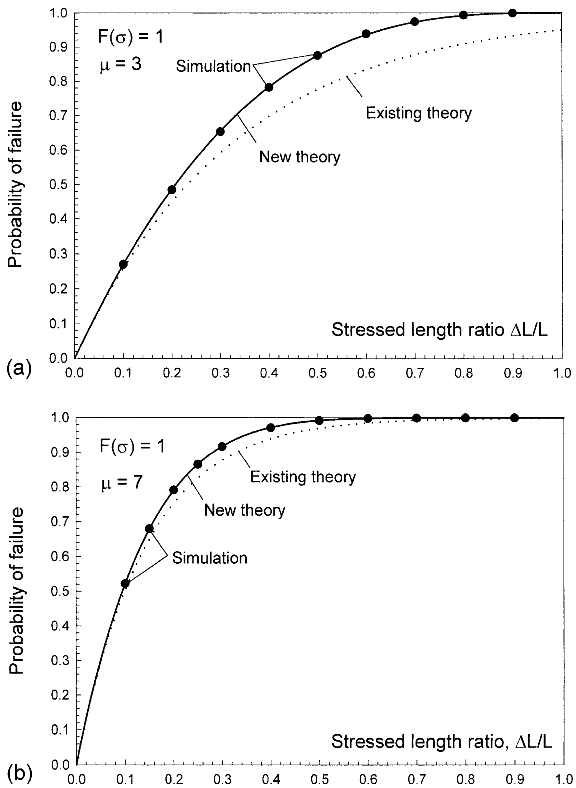

Fig. 1. Probabilities of failure of a piece of wire calculated from a Monte Carlo simulation (●), from the existing statistical theory and the new theory. Two different number densities of the defects were used: (a) three defects per 1 m length and (b) seven defects per 1 m length. The results from the new theory are in excellent agreement with the simulation results. The discrepancies between the simulation results and the results from the new theory on one hand and the results from the existing theory on the other hand are larger the smaller the number density of the defects.

$$S(\sigma) = 1 - \exp\{\mu L \ln[1 - p]\} \tag{4}$$

For M groups (types) of defects (inclusions of various kind, particles, microcracks, etc.), Eq. (3) can be generalised to

$$S(\sigma) = 1 - \exp\left(L \sum_{i=1}^{M} \mu_i \ln[1 - pF_i(\sigma)]\right) \tag{5}$$

where $F_i(\sigma)$ are the individual probabilities of triggering fracture and μ_i are the defect number densities characterising the different groups of defects $(i = 1, M)$.

It is commonly assumed that the largest defects cause the greatest number of failures. In order to verify this statement, the relationship between the size of the defects and the failure frequency characterising this size needs to be analysed. It will be demonstrated that whether the largest defects cause the majority of fractures depends significantly on the ratio of the stressed length to the whole calibration length of the examined component.

The distribution of the defects in the wire is often approximated [23] by a Poisson process existing under the following conditions: (i) the probability of more than 1 defect in a very small subinterval is zero; (ii) the probability of existence of a defect in a particular subinterval is the same for all subintervals and proportional to the length of the subinterval and (iii) the number of defects in each subinterval is independent of other subintervals. Then if the mean number density of the defects in an unit length L is $\mu = n/L$, where n is the number of defects, the number of defects x in a specified length 'd' is given by the Poisson distribution:

$$f(x) = \frac{e^{-\mu d}(\mu d)^x}{x!} \tag{6}$$

The lengths of the distances d_{ij} between the separate adjacent defects then follow exponential distribution [23]. Indeed, the probability of existence of a defect-free zone with length at least d, between the first two defects for example (the defects with the smallest coordinates, Fig. 2a), can be obtained directly from Eq. (6) by setting $x = 0$:

$$\Pr(d_{12} > d) = e^{-\mu d} \tag{7}$$

Correspondingly, the probability that the distance d_{12} between the first two defects is smaller than d is given by $\Pr(d_{12} \leqslant d) = 1 - e^{-\mu d}$ which is the cumulative distribution function of the exponential distribution $\mu e^{-\mu d}$.

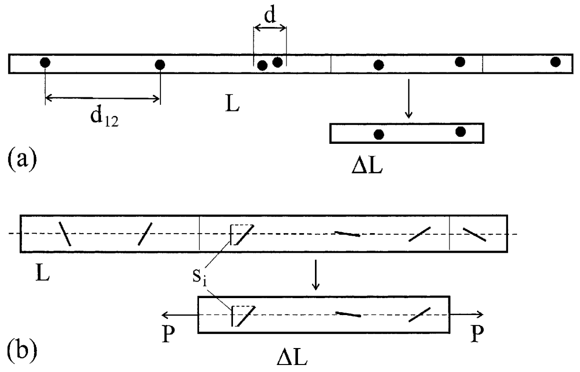

Fig. 2. (a) Clusters of two or more defects close to one another cause premature failure. (b) Modelling the influence of the defect size on the probability of failure, on one-dimensional strip with cuts (defects) of the same length, each oriented at a different angle regarding the axis.

For a component with length 1 m and $\mu = 4$ defects per 1 m length and for a distance $d = 0.55$, Eq. (7) results in $\Pr(d_{12} > d) = e^{-4\times 0.55} = 0.11$. The Monte Carlo simulation however, yields a probability $\Pr(d_{12} > d) = 0.0415$ which is almost three times smaller than that calculated from Eq. (7). Later it will be demonstrated that the reason for this discrepancy is that Eq. (7) is not exact for finite lengths L and describes well the probability of existence of defect-free distance d only in case of a small ratio d/L. Accordingly, an exact equation needs to be derived which describes correctly the probability of existence of a specified defects-free distance between any two adjacent defects, for all possible ratios d/L.

Furthermore, equations are needed for determining the probability of existence of clusters of two or more defects lying within a short distance of one another and regarding the probability of existence of minimum 'safe' distances between the adjacent defects. If the defects are too close to each other (Fig. 2a) their fields of stress concentration superpose which may result in a larger stress concentration zone. For thin wires and fibers this in turn increases the risk of failure. If the defects are at safe distances from one another, their zones of stress concentration do not interact and the load-bearing section is not reduced. The local stress magnitude, raised in the vicinity of a defect, attenuates significantly before another zone of stress concentration (belonging to another defect) is encountered and as a result, the risk of failure is reduced. Thus, calculating the probability of existence of safe distances between the adjacent defects is an important reliability calculation while evaluating the risk of poor properties.

Closely associated with the spatial statistics of defects in one-dimensional components is the problem of inspecting for defects one-dimensional components, known to contain defects with certainty, using a non-destructive testing technique. This is always associated with a certain probability of detecting a flaw (defect) of a particular kind when it is present in the component. Radiographic non-destructive testing, in particular, is better suited to detection of flaws perpendicular to the surface of one-dimensional components than to detection of flaws parallel to the surface. In other words, uncertainty is associated with the detection technique on the one hand and the kind, shape and orientation of the defect being detected on the other hand. Accordingly, it is important to know the probability of detecting at least a single defect from examining part of the component ΔL (Fig. 2a), when the probability with which the defects are detected by the non-destructive technique is known. Another important question is regarding the minimum fraction of the component that needs to be examined in order to make the probability of detecting at least a single defect greater than a specified value. These are important problems from reliability and risk assessment.

2. Relationship between the size of the defects and the probability of failure

In order to model the influence of the defect size on the probability of failure, a one-dimensional strip with length $L = 1$ was considered (Fig. 2b). Suppose, identical cuts are distributed along the strip, each cut oriented at a different angle regarding the axis of the strip, as shown in Fig. 2b. Next, part of the strip of length ΔL is 'cut out' and 'loaded' until failure. According to the weakest link concept, during slow loading from zero load, and subsequent increase of the load P until failure, the defect (cut) that causes failure is the defect (cut) with the largest projection (size) s_i on a plane perpendicular to the axis of the strip (Fig. 2b).

Let $x = \Delta L/L$ be the size ratio. Without loss of generality, suppose that all cuts (defects) in the strip have been ordered in descending order $s_1 > s_2 > s_3 > \cdots > s_n$ according to the size of their projections. The probability that the cut with the largest projection s_1 (referred to as the 'largest defect') will cause failure is $\Pr(s_1) = x = \Delta L/L$ since the probability of cutting out the defect with the largest projection p_1 (the largest defect) with the test piece ΔL is $\Delta L/L$. In a similar fashion, the probability that the second largest defect will initiate failure can be obtained: $\Pr(s_2) = (\Delta L/L) \times (L - \Delta L)/L = x(1 - x)$. This is the probability of the compound event that the second largest defect will be in ΔL and the largest defect will not. The two events that constitute the compound event are independent and their probabilities have been multiplied. As a result, the probabilities that the first, the second, the third, ..., the nth defect will initiate failure are given by

$$\Pr(s_1) = x$$

$$\Pr(s_2) = x(1 - x)$$

$$\Pr(s_3) = x(1 - x)^2$$

$$\vdots$$

$$\Pr(s_n) = x(1 - x)^{n-1}$$

Since $S = \sum_{i=0}^{n-1} x(1 - x)^i = 1 - (1 - x)^n$, the fraction of failures caused by the largest defect is $x/S = x/(1 - (1 - x)^n)$, the fraction of failures caused by the second largest defect is $x(1 - x)/S = x(1- x)/(1 - (1 - x)^n)$, etc.

The fraction of failures caused by the 'k' largest defects $k \leqslant n$ is then given by

$$F = \sum_{i=0}^{k-1} \frac{x(1 - x)^i}{1 - (1 - x)^n} = \frac{1}{1 - (1 - x)^n} \sum_{i=0}^{k-1} x(1 - x)^i = \frac{1 - (1 - x)^k}{1 - (1 - x)^n} \quad (8)$$

Suppose that the number of defects n is relatively large. Then the term $(1-x)^n$ in the denominator of Eq. (8) can be neglected and as a result, the fraction of failures caused by the k largest defects is given by

$$F \approx 1-(1-x)^k \tag{9}$$

3. Spatial statistics of defects uniformly distributed in one-dimensional components

3.1. General equation regarding the spatial distribution of the distances between the defects

Suppose that n defects are randomly distributed in one-dimensional component with length L (Fig. 3a) (one-dimensional Poisson process). The distribution of the coordinate x of each defect is uniform, with probability density $1/L$, i.e. each defect has the same chance $\Delta L/L$ of being in a piece with length ΔL.

The probabilities $\Pr(d_{12} \geqslant d_1)$, $\Pr(d_{23} \geqslant d_2), \ldots,$ $\Pr(d_{kk+1} \geqslant d_k)$ that the distances between any specified pairs of adjacent defects for example $d_{12} = |x_2 - x_1|$, $d_{23} = |x_3 - x_2|$ and $d_{kk+1} = |x_{k+1} - x_k|$ will be larger than or equal to the specified distances $d_i(\sum_{i=1}^k d_i < L)$ will be determined first (Fig. 3a).

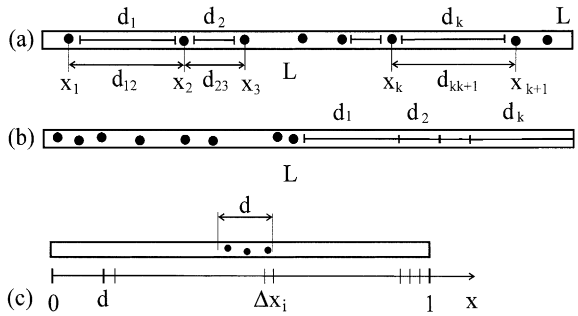

Fig. 3. A one-to-one correspondence between (a) the successful realisations where the distances between the selected pairs of defects are at least $d_1, d_2, \ldots, d_k$ and (b) the successful realisations of 'n' randomly generated defects falling only over the piece with length $L - \sum_{i=1}^k d_i$. (c) The event 'n defects are located within a distance d is an union of the mutually exclusive events: the defect with the maximum coordinate x falls into the ith subinterval and the rest of the '$n-1$' defects fall on the left, within a distance d, for all subintervals Δx_i.

Assume that the locations of the defects are 'generated' along the component and only 'successful' realisations are counted, i.e., realisations where the distances between the specified adjacent defects are greater or equal than the specified $d_1, d_2, d_3, \ldots, d_k$. In all such 'successful realisations' Fig. 3a, the minimum distances $d_1, d_2, d_3, \ldots, d_k$ can be cut out of the actual defect-free intervals d_{12}, $d_{23}, \ldots, d_{kk+1}$ between the selected k pairs of defects and 'moved' towards the end of the component, as shown in Fig. 3b. Now, there exists a one-to-one correspondence between the successful realisations, in Fig. 3a (where the distances between the selected pairs of defects are at least $d_1, d_2, \ldots, d_k$) and the successful realisations in Fig. 3b of n defects randomly generated over the length L and all falling only over the length $L - \sum_{i=1}^k d_i$ (Fig. 3b).

Indeed, according to the earlier discussion, from all successful realisations of n randomly generated defects over the length L (Fig. 3a) with the specified minimum distances between the selected adjacent defects we get all successful realisations of n random defect locations falling in the shorter length $L - \sum_{i=1}^k d_i$ by cutting out the minimum distances $d_1, d_2, d_3, \ldots, d_k$ between the corresponding adjacent pairs of defects and 'sticking' the cuts.

Conversely, from all successful realisations of n randomly generated defects falling only in the shorter length $L - \sum_{i=1}^k d_i$ (Fig. 3b), we can obtain all successful realisations of n random defects in the length L (Fig. 3a) (with the specified minimum distances between the selected pairs of defects) simply by inserting the minimum distances d_1, $d_2, \ldots, d_k$ between the corresponding adjacent defects on the shorter length $L - \sum_{i=1}^k d_i$ (Fig. 3b).

As a result, the initial problem transforms into an equivalent, much simpler problem. Correspondingly, the probability of existence of the specified minimum distances between the selected pairs of defects can be determined by calculating the probability of n defects generated over the length L actually all falling in the shorter length $L - \sum_{i=1}^k d_i$. Since the latter probability is

$$P = \left(1 - \frac{1}{L}\sum_{i=1}^{k} d_i\right)^n \tag{10}$$

this is also the probability of the event that the distances between the specified k adjacent pairs of defects will be equal or greater than the specified minimum distances $d_i, i = 1, k$.

3.2. Special cases of the general equation (10)

The probability that there will be at least two adjacent defects among the n defects, the distance between which is smaller than a specified distance d can be obtained by subtracting from unity the probability of the complementary event that 'the distance between any two adjacent defects among the n defects will be at least d. Since from Eq. (10) this probability is

$$P = \left(1 - \frac{(n-1)d}{L}\right)^n \tag{11}$$

the probability of at least two defects being at a distance smaller than d is

$$P = 1 - \left(1 - \frac{(n-1)d}{L}\right)^n \tag{12}$$

The probability that there will exist a defect-free zone of minimum length d between two specified adjacent defects 'i' and '$i+1$' is

$$P = \left(1 - \frac{d}{L}\right)^n \tag{13}$$

which is also a special case of Eq. (10). The probability of existence of a minimal defect-free length d (the indexes of the defects are not specified) can be also be obtained from the general equation (10). Suppose that at most k defect-free distances with minimal length d can be 'accommodated' between the n defects: ($kd \leqslant L$, $(k+1)d > L$). Let $P_1, P_2, \ldots, P_k$ denote the probabilities of existence of exactly $1, 2, \ldots, k$ defect-free distances d between the specified pairs of defects. These probabilities are given by Eq. (10). The probability of existence of exactly 1 defect-free distance is given by $\binom{n-1}{1}P_1$, where $\binom{n-1}{1}$ are all distinct choices (combinations) of the position of one defect-free distance among $n-1$ possible positions. $P_1 = (1 - (d/L))^n$ is given by Eq. (13) which is a special case of Eq. (10). The probability of existence of exactly 2 defect-free distances of length d is $\binom{n-1}{2}P_2$, where $\binom{n-1}{2}$ are all distinct choices (combinations) of the positions of two defect free distances among $n-1$ possible positions. $P_2 = (1 - (2d/L))^n$ is given by Eq. (10).

The probability of existence of at least a single defect-free distance d is equal to the probability of existence of one, two, …, or k defect-free distances. Since more than one defect-free distance can exist at the same time (at most k distances), the formula for the probability of sum of events is applied:

$$P = \binom{n-1}{1}P_1 - \binom{n-1}{2}P_2 + \cdots + (-1)^{i-1}\binom{n-1}{k}P_k \tag{14}$$

where the probabilities P_k are given by Eq. (10): $P_k = (1 - (kd/L))^n$.

The probability that all of the adjacent defects will be at distances between each other smaller than a specified distance d is equal to unity minus 'the probability of existence of at least two adjacent defects the distance between which is larger than d' (given by Eq. (14)):

$$P = 1 - \binom{n-1}{1}P_1 + \binom{n-1}{2}P_2 - \cdots - (-1)^{i-1}\binom{n-1}{k}P_k \tag{15}$$

The probability of a configuration where all of the defects are located within a distance d for example, can also be determined. Let the length L be unity ($L = 1$) and the interval (d,1) be divided into very small subintervals Δx_i (Fig. 3c). Then the event: n defects are located within a distance d is an union of the following mutually exclusive events corresponding to all subintervals Δx_i: the defect with the maximum coordinate 'x' falls into the 'ith' subinterval Δx_i and the rest of the $n-1$ defects fall on the left, within a distance d.

For a specified coordinate x, the probability that the maximum coordinate $X_{\max}$ will be smaller than x is given by the product of the cumulative probability functions: $\Phi(X_{\max} \leqslant x) = F(X \leqslant x)F \times (X \leqslant x) \cdots F(X \leqslant x) = F^n(X \leqslant x)$. Since $(F(X \leqslant x) = x)$, is the uniform cumulative distribution function, $\Phi(X_{\max} \leqslant x) = x^n$. Differentiating the latter with respect to x gives the probability density distribution $f(x) = nx^{n-1}$ of the maximum coordinate. Considering also that the probability of the rest of the $n-1$ defects being within a distance d left from the defect with maximum coordinate is given by $(d/x)^{n-1}$, the probability that all defects will lie within a distance d when the defect with maximal coordinate is within the interval $[d, 1]$ is given by

$$\int_d^1 nx^{n-1}(d/x)^{n-1} = nd^{n-1}(1-d) \qquad (16)$$

In order to obtain the probability that the defects will lie within a distance d over the entire length $L = 1$ of the component, the probability d^n that the defect with the maximal x-coordinate will lie within the interval $[0, d]$, must be added to the probability given by Eq. (16). Finally, the probability that all defects will lie within a distance d becomes

$$P = nd^{n-1} - (n-1)d^n \qquad (17)$$

4. Non-destructive inspection of one-dimensional components containing defects with certainty

Suppose that from a component with length L containing with certainty n defects, a certain percentage $p = \Delta L/L$ (a smaller piece with length ΔL) is being examined for defects (Fig. 2a). The probability that at least a single defect will be detected can be determined by subtracting from unity the probability of the event that no defect will be detected in the examined length ΔL. It is assumed for simplicity that defects from a single type are present, each of which is characterised by a probability F of being detected. The probability that no defects will exist in the examined fraction 'p' is $(1-p)^n$; the probability of existence of exactly one defect is $np(1-p)^{n-1}$; the probability of existence of exactly 'r' defects is $\binom{n}{r} p^r (1-p)^{n-r}$. The probability $P^0_{(r)}$ of the compound event that exactly r defects exist in the examined fraction and none of them has been detected is given by

$$P^0_{(r)} = \binom{n}{r} p^r (1-p)^{n-r} [1-F]^r \qquad (18)$$

The probability $P^0_{(r)}$ is a product of the probabilities of two independent events: (i) exactly r defects are present in the examined fraction, the probability of which is given by the binomial formula $\binom{n}{r} p^r (1-p)^{n-r}$ and (ii) the event that none of the defects is detected. The probability $[1-F]^r$ of the second event is a product of the probabilities of the events that neither the first, nor the second, nor the third, ..., nor the rth defect are detected.

The event P^0 that no defect will be detected is a union of the disjoint events $P^0_{(r)}$, that no defect will be detected if the examined fraction contains $r = 0, 1, 2, \ldots, n$ defects. Consequently, the probability of the event P^0 is a sum of the probabilities defined by Eq. (18):

$$P^0 = \sum_{r=0}^{n} P^0_{(r)} = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} [1-F]^r$$

which can be simplified to

$$P^0 = [p(1-F) + (1-p)]^n = [1-pF]^n \qquad (19)$$

The probability D that at least a single defect will be detected in the examined fraction p is given by $1 - P^0$:

$$D = 1 - [1-pF]^n = 1 - \exp\{n \ln[1-pF]\} \qquad (20)$$

If $\mu = n/L$ denotes the number density of the defects in the calibration length L, Eq. (20) can also be presented as

$$D = 1 - \exp\{\mu L \ln[1-pF]\} \qquad (21)$$

If the examined length ΔL coincides with the whole calibration length L $(p = \Delta L/L = 1)$, the probability of detecting at least a single defect becomes

$$D = 1 - \exp\{\mu L \ln[1-F]\} \qquad (22)$$

If the examined fraction is very small $(p = \Delta L/L \ll 1)$, the product pF is very small $(F \leqslant 1)$ and Eq. (21) can be approximated by

$$D = 1 - \exp\{-\mu \Delta L F\} \tag{23}$$

since for small pF, $\ln(1 - pF) \approx -pF$.

The fraction $p_{\min}$ from the calibration length L that needs to be examined in order to guarantee a minimum probability $D_{\min}$ of detecting at least a single defect can be calculated from Eq. (20). Solving Eq. (20) regarding p, gives the fraction $p_{\min}$ that needs to be examined:

$$p_{\min} = (1/F)\{1 - \exp[(1/(\mu L))\ln(1 - D_{\min})]\} \tag{24}$$

Consider N types of defects where μ_i and F_i, $i = 1, N$ are the defect number densities and the probability of detecting a defect from the ith group, correspondingly. Using Eq. (20) for the probability that no defect will be detected in any of the groups, the expression

$$P^0 = \prod_{i=1}^{N} \exp\{\mu_i L \ln[1 - pF_i]\} = \exp\left(L \sum_{i=1}^{N} \mu_i \ln[1 - pF_i]\right) \tag{25}$$

is obtained.

Thus, if more than one group of defects are present, each characterised by a distinct probability F_i of being detected by the non-destructive technique, the probability of detecting at least a single defect during inspecting fraction p from the component is

$$D = 1 - \exp\left(L \sum_{i=1}^{N} \mu_i \ln[1 - pF_i]\right) \tag{26}$$

In case of numerous types of defects present with certainty in the component, the fraction $p_{\min}$ that needs to be examined in order to guarantee a specified minimum probability $D_{\min}$ of detecting a defect can be calculated by solving Eq. (26) regarding p, numerically.

5. Results

All equations that have been derived regarding the probability of detecting defects existing with certainty were validated using Monte Carlo simulations. During the simulation of non-destructive testing for defects for example, the locations of the defects were generated randomly over a calibration length $L = 1$ m. Random trials (100 000) were carried out, during which a part with length ΔL was selected randomly from the calibration length L. At each random trial, all defects belonging to ΔL were identified and for each defect, a random number γ was generated uniformly distributed in the interval $[0, 1]$. A defect from group i is considered of having been detected during inspecting the length ΔL, if the condition $\gamma \leqslant F_i$ is fulfilled. The number of trials in which a defect has been detected was divided by the total number of trials to estimate the probability of determining at least a single defect during inspection.

For three groups of defects with number densities $\mu_1 = 3$, $\mu_2 = 42$ and $\mu_3 = 5$, characterised by probabilities $F_1 = 0.6$, $F_2 = 0.42$, $F_3 = 0.27$ of being detected, the simulated probability of detecting at least a single defect in part of length $\Delta L = 0.01$ m was 0.11. This probability was in excellent agreement with the theoretical probability (0.11) calculated from Eq. (26).

For other tested lengths, the empirical probabilities from the simulation were again very close to the theoretical probabilities calculated from Eq. (26) (Table 1).

For a single type of defects characterised by a probability $F = 0.8$ of detecting the separate defects by the non-destructive inspection technique, the minimum fractions $p_{\min}$ that need to be examined to guarantee a minimum probability $D_{\min}$ of detecting at least a single defect are plotted in Fig. 4 for three different number densities of the defects.

Table 1
Comparison between the empirical probabilities of detecting defects in length ΔL cut from a calibration length $L = 1$ m

	Tested length, ΔL, (m)			
	0.01	0.05	0.1	0.2
Empirical probability	0.11	0.45	0.71	0.93
Theoretical probability	0.11	0.46	0.72	0.93

Three groups of defects are present with certainty, with number densities $\mu_1 = 3$, $\mu_2 = 42$ and $\mu_3 = 5$ and characterised by probabilities $F_1 = 0.6$, $F_2 = 0.42$ and $F_3 = 0.27$ of detecting the separate defects in the groups.

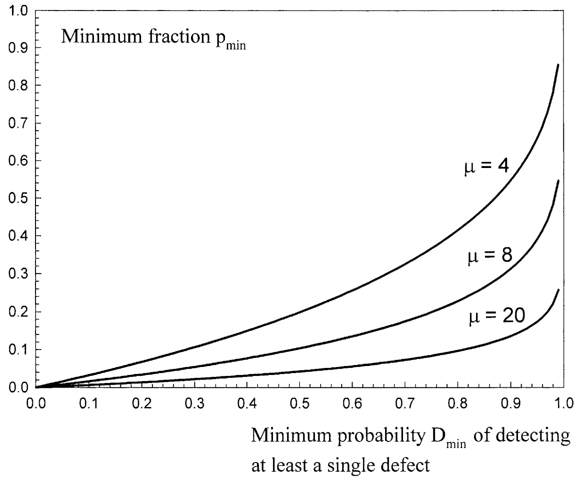

Fig. 4. Minimum fractions p that need to be examined to guarantee a minimum probability D_{min} of detecting at least a single defect. For a single type of defects characterised by a probability $F = 0.8$ of being detected by the non-destructive inspection technique, the minimum fractions are plotted for three different number densities of the defects: $\mu = 4$, $\mu = 8$ and $\mu = 20$.

The equations regarding the spatial statistics of defects were verified by Monte Carlo simulations consisting of generating uniformly distributed defect locations on a segment with length 1 m and checking the distances between them. Thus, for $n = 4$ randomly generated locations on the segment, the empirical (Monte Carlo) probability that there will be a defect-free distance of at least $d = 0.55$ between the first two adjacent defects, yielded 0.0415 which was very close to the probability calculated from Eq. (13): $P = (1 - d)^4 = 0.041$.

For the probability that between some of the defects there will be a defect-free distance greater than $d = 0.2$ m, the Monte Carlo simulation yielded 0.866. The probability, calculated from Eq. (14) yielded

$$P = 3(1 - d)^4 - 3(1 - 2d)^4 + (1 - 3d)^4 = 0.865$$

which agreed well with the probability from the simulation.

For the probability that all distances between the defects will be smaller than $d = 0.2$ m, the Monte Carlo simulation yielded 0.134 which agreed very well with the probability 0.135 calculated from Eq. (15).

For the probability that at least two defects will be at a distance smaller than 0.01 ($d < 0.01$), the Monte Carlo simulation yielded empirical probability of 0.116 which was in good agreement with the theoretical probability calculated from Eq. (12): $P = 1 - (1 - 3d)^4 = 0.115$. The probability that between each pair of adjacent defects there will be a defect-free distance of at least $d = 0.01$ is given by Eq. (11): $P = (1 - 3d)^4 = 1 - 0.115 = 0.885$.

For the probability that the four defects will be at a distance smaller than $d = 0.2$, Eq. (17) yields $P = 4d^3 - 3d^4 = 0.0272$ which is very close to the empirical probability 0.0274 from the Monte Carlo simulation.

The graph of Eq. (12) giving the probability of existence of a cluster of two or more defects at three different critically small distances $d = 0.01$ m, $d = 0.001$ m and $d = 0.0001$ m and a calibration length $L = 1$ m is given in Fig. 5.

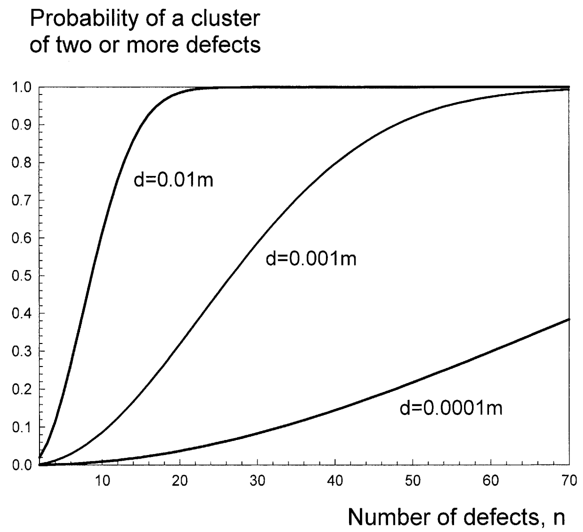

Fig. 5. Graph of Eq. (12) giving the probability of existence of a cluster of two or more defects at three different critically small distances $d = 0.01$ m, $d = 0.001$ m and $d = 0.0001$ m. Even for moderate defect numbers, the probability of existence of clusters of two or more defects at a critically small distance is substantial and should not be neglected in calculations of the risks of failure.

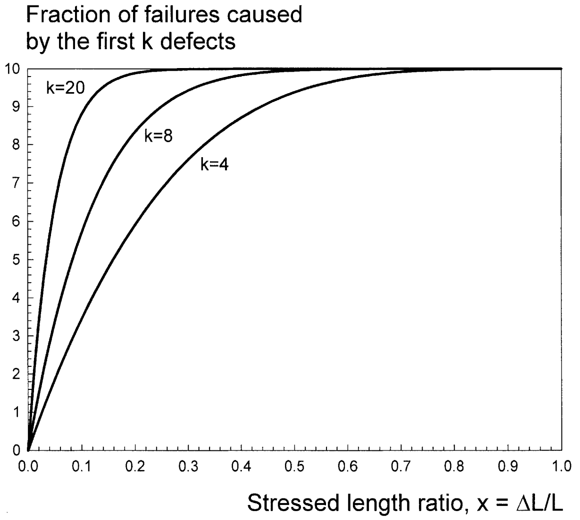

Fig. 6. Graph of Eq. (9) giving the fraction of failures caused by the k largest defects for $k = 20$, $k = 8$ and $k = 4$. For a large length ratio of the tested part a small fraction of the largest defects is accountable for most of the failures; for small length ratios however, nearly all defects participate on an equal basis as initiators of failure.

The graph of Eq. (9) giving the fraction of failures caused by the k largest defects is given in Fig. 6, for $k = 20$, $k = 8$ and $k = 4$.

6. Discussion

Eq. (3) giving the probability of fracture controlled by defects describes in fact failure of a chain containing $n = \mu L$ links, each of which fails with probability $F(\sigma)$ at the stress level σ. Thus, a loaded one-dimensional component with length L, known to contain a certain number of defects can be regarded as a 'chain' with links corresponding to the separate defects. The individual probabilities characterising the capability of the defects to trigger fracture are the probabilities with which the links of the chain fail under load.

Eq. (3) acknowledges the circumstance that only part of the defects will initiate failure at the loading stress σ, i.e., at any stress level, a certain probability of triggering fracture is associated with the defects due to factors that increase or hinder the capability of a defect to trigger fracture. Such factors can be the orientation of the defects regarding the direction of the maximum stress, or the presence of microstructural or other features, such as weak bonds with the matrix, sharp edges acting as stress raisers, shape variations, tessellation stresses, etc. Because of the inevitable variations in the microstructure, orientation, the bond with the matrix, etc., among the defects which cannot sustain an applied load P without triggering fracture, the fracture will start from the 'weakest' defect, with the most unfavourable combination of size, shape, microstructural features and orientation.

If for example the strip in Fig. 2b is loaded with a force P and the tensile strength of the material of the strip is σ, a particular defect will trigger failure whenever the stress acting on the effective load-bearing section exceeds the strength σ of the material of the strip in the vicinity of the cut. Since the size of the load-bearing section depends on the orientation of the cut (on the angle which the cut subtends with the horizontal axis), the cuts (defects) will be associated with a certain probability of triggering failure, i.e., they will be associated with a certain probability of having unfavourable orientation and causing failure under the load P (since all of the cuts have identical shape). Among the unfavourably oriented cuts which cannot sustain load P, the rupture of the strip will be triggered by the cut with the most unfavourable orientation according to the weakest link principle.

Among the unfavourably oriented defects (cuts) in Fig. 2b that cannot sustain load P without triggering rupture, the cut with the largest projection (the largest defect) will in fact initiate the rupture.

As can be seen from the graph of Eq. (9), presented in Fig. 6, for large length ratios $x = \Delta L/L$, F quickly tends to unity, even for small k. Thus, for $x = 0.6$, four of the largest defects are accountable for most of the failures. This means that for large tested size ratios, the Pareto principle regarding defects and failures holds: the majority of failures are caused by a small fraction of the largest defects.

For small length ratios $x = \Delta L/L$ however, larger fraction of the defects participates as initiators of failure. For a very small size ratio x, almost all defect sizes participate on an equal basis as failure initiators. Consequently, the widespread opinion that all failures are caused by a tiny fraction of the largest defects is correct only if the tested size ratio is large. For small size ratios this statement is not correct.

The statistical estimation of the distribution of the largest defects is important for commercial applications since a small fraction of the largest inclusions often controls the fatigue or fracture resistance. Suppose that the random samples of defect sizes (determined by image analysis technique for example) are sufficiently large and come from a distribution with an unbounded upper tail, decaying at least as fast as an exponential distribution. Under these assumptions, the distribution of the maximum defect size x (the largest defect) can be approximated by a type I extreme value distribution of greatest values, also known as the Gumbel [24] or double exponential distribution. The type I extreme value distribution has a cumulative distribution function

$$F(x) = \exp\{-\exp[-(x - \xi)/\theta]\}$$

where the parameter ξ is the mode of the distribution and θ is a scale factor proportional to the standard deviation.

There exists a very strong analogy between the model proposed for calculating the probability of failure of a one-dimensional component (Eq. (3)) and the model related to a defect inspection by a non-destructive technique. Indeed, in order to obtain the probability of detecting at least a single defect during inspecting a fraction of the component, it suffices to substitute the probability of triggering failure $F(\sigma)$ by the defects in Eq. (3) with the probability F of detecting a defect by the non-destructive inspection technique. The probability of detecting a defect varies according to the inspection method, the size and the orientation of the defect. Eq. (21) is particularly useful in cases where it is impossible or very expensive to examine the whole length of a component which contains defects with certainty.

If the calibration length L is sufficiently large, the number of defects detected on that length will vary little from length to length. Accordingly, it can be assumed that the calibration length of that size certainly contains a particular number of defects and Eq. (21) can be applied. The graphs in Fig. 4 show that in case of large number densities, a specified minimum probability $D_{\min}$ of detecting a defect can be achieved by testing a relatively small fraction $p_{\min}$. The minimum fraction $p_{\min}$ that needs to be examined however increases rapidly with decreasing the number density of the defects. Furthermore, at a specified number density of the defects, probabilities $D_{\min}$ smaller than 0.8 require a relatively small minimum fraction to test. Minimum probabilities $D_{\min}$ larger than 0.8 however, require relatively large minimum fractions $p_{\min}$.

The computer simulations in Section 5 illustrate that the formulae derived regarding the spatial statistics of randomly distributed defects in one-dimensional components are correct. They allow to estimate the probability of existence of safe distances between adjacent defects and also the probability of existence of clusters of two or more defects at a critically small distance. These probabilities, which can be obtained as special cases of the general Eq. (10), are important in cases where clusters of two or more defects cause of failure of the loaded components. The probability given by Eq. (14) is important for estimating the probability of existence of a specified continuous defect-free zone in a one-dimensional component.

One unexpected corollary from the general equation (10) is that the probability of existence of a minimum defect-free distance of a given size between two adjacent defects does not depend on the indices of the adjacent defects but on the magnitude of the distance and the number of defects only. Furthermore, another unexpected corollary from Eq. (10) is that the probability of existence of a particular set of minimum defect-free distances between any specified adjacent defects is equal to the probability of existence of any other set of distances between any other specified adjacent defects, as long as the sums of the distances in the two data sets are equal.

Besides, this means for example that the probability of existence of several minimum defect-free distances between any specified adjacent defects is equal to the probability of existence of a single defect-free distance (between any two specified defects) which is equal to the sum of the specified minimum defect-free distances.

The excellent agreement between the Monte Carlo simulation results and the results from Eq. (13) shows that the latter equation should be used instead of the classical equation (7) which shows significant deviations from the simulation results for components with final length L. These discrepancies can be explained by taking logarithms from the right-hand parts of Eqs. (7) and (13) which results in $n \ln(1 - d/L)$ and $-\mu d = -nd/L$, respectively. As can be verified these logarithms are equal (therefore the probabilities in the left part of the equations are equal too) only for $d/L \ll 1$. In this case $\ln(1 - d/L) \approx -d/L$ and Eq. (13) transforms into Eq. (7). If the ratio d/L is not small, however, the exact (13) should be applied, not Eq. (7).

The graphs in Fig. 5 show that the probability of existence of a cluster of two or more defects at a distance 1 mm (the middle curve) increases rapidly with increasing the number of defects. For example, even for a relatively small number of defects $n = 12$, the probability of a cluster of at least two defects closer than 1 mm is substantial: $P = 0.124$. For $n = 30$ defects, the probability is $P = 0.58$. Even for the very small critical distance of 0.1 mm, the probability of existence of such clusters is substantial if the number of defects is greater than 30 (Fig. 5, the curve corresponding to $d = 0.0001$ m). The analysis shows that contrary to the expectations, even for a moderate defect number density, the probability of existence of clusters of two or more defects at a critically small distance is substantial and should not be neglected in calculations related to risk assessments.

The equations related to the statistics of defects in one-dimensional components, are an important part of the statistics of the structure and properties of inhomogeneous materials. Moreover, they can be applied without modification to calculate the probability of clustering of events on a given time interval. In this case the one-dimensional component with length L corresponds to a time interval with length L and the defects are in fact the events occurring during this time interval. Then Eq. (12) gives the probability that at least two events will occur close to each other. This has an important application in the reliability and risk assessment studies for example when the events correspond to failures whose repair consumes resources and the resources are limited. In this case it is important to determine the probability that resources will be available for all failures which form a Poisson process (i.e. there will be no running out of resources for repair).

7. Conclusions

1. The equations related to the spatial statistics of defects in one-dimensional components are special cases of the general equation (10) and allow to estimate the probability of existence of a defect-free zone of any specified length, the probability of existence of clusters of two or more defects close to each other and the probability of existence of safe distances between the defects. The equations can also be applied to problems of clustering of events on a time interval.
2. From the general equation (10) it follows that the probability of existence of a minimum defect-free distance of given size between two adjacent defects does not depend on the indices of the adjacent defects but on the magnitude of the distance and the number of defects only. Additionally, the probability of existence of a particular set of minimum defect-free distances between any specified adjacent defects is equal to the probability of existence of any other set of distances between any other specified adjacent defects, as long as the sums of the distances in the two data sets are equal.
3. Even for moderate defect number densities, the probability of existence of clusters of two or more defects at a critically small distance is substantial and should not be neglected in calculations estimating risks of failure of one-dimensional components.
4. An equation has been derived regarding the probability of detecting at least a single defect when only a fraction of the component has been inspected. The equation has potential application in non-destructive testing of components which contain defects with certainty.

5. For a large size ratio of the tested part cut from a given calibration length containing defects, a small fraction of the largest defects is accountable for most of the failures of the tested part; for small size ratios however, nearly all defects participate on an equal basis as initiators of failure.
6. The exact equation (13) (which is a special case of Eq. (10)) instead of the classical equation (7) should be used for calculating the probability of existence of a defect-free zone of specified length between any particular pair of adjacent defects in one-dimensional components. The probability calculated from Eq. (7) shows large discrepancies from the true values when the specified length fraction of the defect-free zone is relatively large.

References

[1] W. Weibull, J. Appl. Mech. September (1951) 293–297.
[2] F.M. Beremin, Metall. Trans. A 14 (1983) 2277–2287.
[3] B. Dodson, Weibull Analysis, ASQC Quality Press, Milwaukee, 1994.
[4] R.B. Abernethy, The New Weibull Handbook, Robert B. Abernethy, North Palm Beach, 1994.
[5] B. Bergman, J. Mater. Sci. Lett. 4 (1985) 1143–1146.
[6] M.T. Todinov, Mater. Sci. Eng. A 276 (2000) 39–47.
[7] A.R. Rosenfield, in: K.S. Chan (Ed.), George R. Irwin Symposium on Cleavage Fracture, The Minerals Metals and Materials Society, Materials Park, OH, 1997, pp. 229–236.
[8] C.J. McMahon, M. Cohen, Acta Metall. 13 (1965) 591.
[9] M.A. Linaza, J.L. Romero, J.M. Rodriguez-Ibabe, J.J. Urcola, 36th MWSP Conf. Proc., ISS-AIME 32 (1995) 483–494.
[10] J.H. Chen, G.Z. Wang, Metall. Trans. A 23 (1992) 509–517.
[11] D. Brooksbank, K.W. Andrews, J. Iron & Steel Inst. 4 (1972) 246–255.
[12] G.T. Hahn, Metall. Trans. A 15 (1984) 947–959.
[13] K. Wallin, T. Saario, K. Torronen, Metal Sci. 18 (1984) 13–16.
[14] R.A. Hunt, L.N. McCartney, Int. J. Fract. 15 (1979) 365–375.
[15] L.N. McCartney, Int. J. Fract. 15 (1979) 477–487.
[16] A.G. Evans, J. Am. Ceram. Soc. 61 (1978) 302–308.
[17] A.G. Evans, T.G. Langdon, Prog. Mater. Sci. 21 (1976) 320–421.
[18] J.R. Matthews, F.A. McClintock, W.J. Shack, J. Am. Ceram. Soc. 59 (1976) 304–308.
[19] R. Danzer, J. Eur. Ceram. Soc. 10 (1992) 461–472.
[20] A.De S. Jayatilaka, K. Trustrum, J. Mater. Sci. 12 (1977) 1426.
[21] S.B. Batdorf, J.G. Crose, J. Appl. Mech. June (1974) 459–464, 171–441.
[22] M.T. Todinov, Mater. Sci. & Eng. A 302 (2001) 235–245.
[23] D.C. Montgomery, G.C. Runger, N.F. Hubele, Engineering Statistics, second ed., John Wiley & sons, New york, 2001.
[24] E.J. Gumbel, Statistics of Extremes, Columbia University Press, New York, 1958.