

Title	Objective assessment of scatter and size effects in the Euro fracture toughness data set
Author(s)	Wallin, Kim R.W.
Citation	Procedia Engineering vol. 10(2011), pp. 833-838
Date	2011
URL	<a href="http://dx.doi.org/10.1016/j.proeng.2011.04.137">http://dx.doi.org/10.1016/j.proeng.2011.04.137</a>
Rights	Copyright © (2011) Elsevier. Reprinted from Procedia Engineering. This article may be downloaded for personal use only

<p>VTT <a href="http://www.vtt.fi">http://www.vtt.fi</a> P.O. box 1000 FI-02044 VTT Finland</p>	<p>By using VTT Digital Open Access Repository you are bound by the following Terms &amp; Conditions.</p> <p>I have read and I understand the following statement:</p> <p>This document is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of this document is not permitted, except duplication for research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered for sale.</p>
---	---



ICM11

# Objective assessment of scatter and size effects in the Euro fracture toughness data set

Kim R. W. Wallin a\*

*<sup>a</sup>VTT Materials for Power Engineering, P.O. Box 1000, FI-02044 VTT, Espoo, Finland*

---

## Abstract

One of the largest and best characterized fracture toughness data sets is the so called Euro data set developed in a European co-operation project some time ago. Probably the most objective analysis of the data set was made using a distribution comparison method. Recently, an improvement of the method has been achieved by combining the Rank probability estimates with Binomial probability estimates. The combination of Rank and Binomial probability estimates double the number of independent individual point estimates, making the overall estimate more accurate with respect to the true value. The strength in this statistical analysis method lies in the objectiveness of the result. This new statistical assessment method is here applied to analyze the Euro fracture toughness data set once more. The results form a new basis for micro-mechanistic modeling of cleavage fracture.

© 2011 Published by Elsevier Ltd. Selection and peer-review under responsibility of ICM11

*Keywords:* Statistical analysis; cleavage fracture; Rank probability; Binomial probability; Master Curve; Euro data set

---

## 1. Introduction

Cleavage fracture is controlled by a minute volume element of the order of one micrometer. This leads to a pronounced scatter. The assessment of fracture toughness data requires thus a statistical assessment. However, most statistical analysis methods like least-square fitting or maximum likelihood make some assumption regarding the underlying distribution. If the underlying distribution is un-known the use of some distribution free assessment method is preferable to standard fits. Here the combination of two different distribution free assessment methods is shown to produce good descriptions of the data scatter. The methods can be used to determine which kind of distribution is most appropriate for the data or they

---

\* Corresponding author. Tel.: +358-50-511-4126; fax: +358-20-722-7002.

*E-mail address:* [kim.wallin@vtt.fi](mailto:kim.wallin@vtt.fi)

can be used directly to develop statistically defined lower and upper bound estimates without requiring knowledge of the actual distribution. The two methods are the Rank probability and Binomial probability methods. The Rank probability, shown in figure 1a, corresponds to the probability that, for a data set ordered by rank, the  $i^{\text{th}}$  value is equal to or less than  $x_i$ . The Binomial probability, shown in figure 1b, corresponds to the probability that, for a data set ordered by rank, the lowest  $i$  values are equal to or less than  $x_i$ . The definition may appear similar, but they do contain two significant differences. First, the Rank probability is connected to a specific test result, whereas the Binomial probability is connected to some freely chosen criterion  $x_i$ . Second, the Rank probability exists for the values  $i = 1 \dots n$ , whereas the Binomial probability exists for the values  $i = 0 \dots n$ . These differences makes the methods well suitable to provide two independent descriptions of the data scatter as will be described in the following analysis.

## Nomenclature

$B_0$	normalisation thickness defined in standard ASTM E1921
$i, j$	order numbers
$K_{JC}$	elastic plastic fracture toughness based on J-integral
$K_{\min}$	theoretical lower bound fracture toughness defined in standard ASTM E1921
$M$	specimen measuring capacity criterion defined in standard ASTM E1921
$n$	total number of data values
$n_f$	number of fractured values
$n_s$	number of survived values
$P$	cumulative probability
$P_B$	Binomial probability
$P_{B0.5}$	median Binomial probability estimate
$P_{\text{conf}}$	confidence level
$P_{\text{rank}}$	cumulative Rank probability corresponding to $P_{\text{conf}}$
$P_{\text{rank}0.5}$	median Rank probability estimate
$P\{X=i\}$	probability that number of fractures ( $X$ ) is equal to $i$
$T$	temperature
$T_0$	transition temperature defined in standard ASTM E1921
$x_i$	parameter value corresponding to location $i$
$\delta_j$	censoring parameter
$\Sigma n$	sum of $\Sigma n_s$ and $\Sigma n_f$
$\Sigma n_f$	total number of fractured values
$\Sigma n_s$	total number of survived values

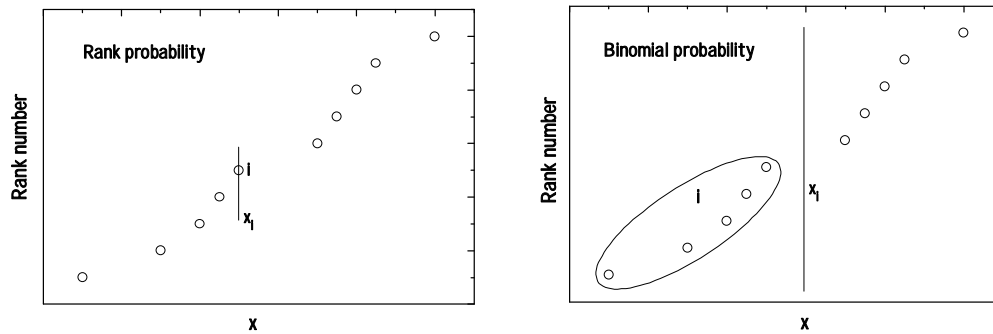


Fig. 1. (a) Rank probability represents the probability that the  $i^{\text{th}}$  rank ordered value is equal to or less than  $x_i$ ; (b) Binomial probability represents the probability that the lowest  $i$  values are equal to or less than  $x_i$ .

### 1.1. Rank probability

Rank probability is a popular way of analyzing intermediate size data sets visually. Since all test results represent individual random probabilities, they follow the rules of order statistics (See e.g. [1]). When test results are ordered by rank, they can be designated Rank probabilities, which describe the cumulative probability distribution. Each data point corresponds to a certain cumulative failure probability with a certain confidence. This can be expressed in a mathematical form based on the binomial distribution as discussed e.g. in [2]. The estimation requires the solving of  $P_{\text{rank}}$  for a specific  $P_{\text{conf}}$  and this makes the estimation somewhat cumbersome. Due to the slight inconvenience in using the exact solution, people usually prefer to use simple approximations of the median ( $P_{\text{conf}} = 0.5$ ) or the mean Rank probability estimate. One of the most accurate analytical simple median Rank probability estimates has the form given in Eq. (1) (See e.g. [2]).

$$P_{\text{rank}0.5} \approx \frac{i - 0.3}{n + 0.4} \quad (1)$$

Eq. (1) can only be used, as such, for data sets where all results correspond to failure. It can also be used with data sets where all values above a certain value have been censored e.g. due to non-failure or exceeding the measuring capacity limit, but in this case the data set size,  $n$ , must refer to the total data set including the censored data.

If the data set contains non-censored failure results at higher values than the lowest censored value, a method of random censoring (often called the suspended items concept) is needed. In this case the order number,  $i$ , in the rank estimation do not remain an integer. The effective order number can be expressed in the form of Eq. (2) [2]. The censoring parameter  $\delta_j$  is zero for censored data and one for non-censored data. Even though Eq. (2) is used on all values, only the non-censored values may be used in the resulting analysis.

$$i_j = \frac{(n + 1 - j) \cdot i_{j-1} + (n + 1) \cdot \delta_j}{n + 1 - j + \delta_j} \quad (2)$$

## 1.2. Binomial probability

The binomial distribution shown in Eq. (3) is often used in proof type testing, where a certain fraction of results fail a certain value. It gives the probability that there are exactly  $i$  events in a set of size  $n$ , when the discrete probability of the event is equal to  $P_B$ .

$$P\{X = i\} = \binom{n}{i} \cdot P_B^i \cdot (1 - P_B)^{n-i} \quad (3)$$

The problem with Eq. (3) is that the probability of the event ( $P_B$ ) is assumed to be known. In a situation where  $n$  tests have been made and  $r$  events have been found, the question is reversed to what the discrete probability ( $P_B$ ) may be with some confidence ( $P_{\text{conf}}$ ). A cumulative probability expression for the confidence can be written, similarly to the Rank probability expression.

The median Binomial probability estimate  $P_{B0.5}$  can also be expressed in a simple analytical form analogous to the median Rank probability estimate in the form of Eq. (4).

$$P_{B0.5} \approx \frac{i + 0.684}{n + 1.368} \quad (4)$$

Censored data values can be used as un-censored when the censored value is higher than the criterion used for  $P_B$ , otherwise the value is disregarded. This is described in more detail in the following analysis.

## 2. The Euro fracture toughness data set

One of the best documented cleavage fracture toughness data sets is the so called Euro fracture toughness data set [3]. The data, excluding the inhomogeneous sub-plate X9, is shown in figure 2.

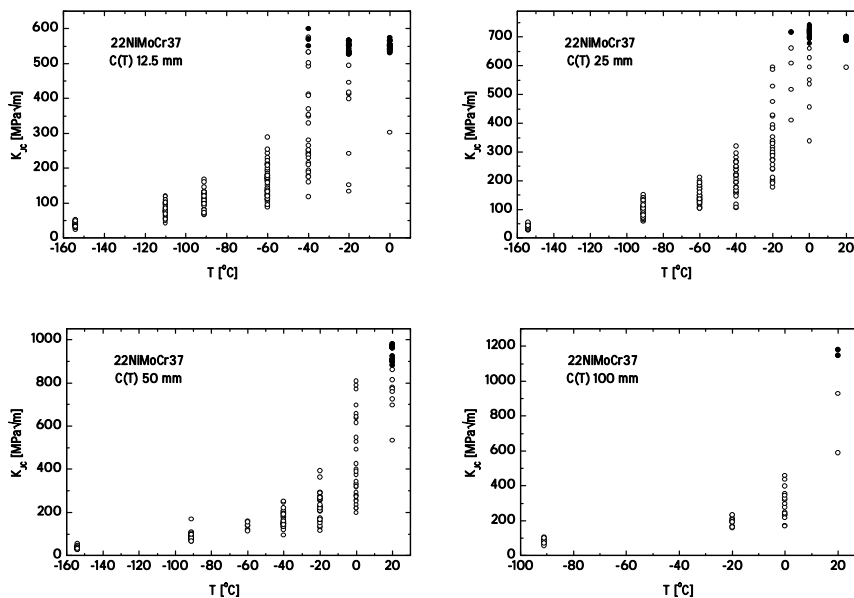


Fig. 2. The Euro fracture toughness data set excluding specimens from sub-plate 9X [6].

The Rank probability analysis is straightforward. The data for each separate temperature is simply ordered by rank and the Rank probability is obtained from Eq. (1). The resulting Rank probability description is shown for the 12.5 mm specimens in figure 3a. For each temperature it is from the figure simple to determine  $K_{JC}$  values corresponding to desired probabilities. This way one obtains  $T - K_{JC}$  data pairs corresponding to a specific probability, without having to make any assumptions regarding the underlying distribution. It is of course also possible to fit a specific distribution to the whole data set shown in figure 2 or to find a relation that would collapse the different probability traces into one curve. An exercise like this has been previously done for the EURO data set [4]. In this work only three different probability levels are considered (5 %, 50 % and 95 %).

The estimation of the Binomial probability is more complicated than of the Rank probability. For the estimation it is best to write the data in the form of  $\Sigma n_s$  including all results above the selected fracture toughness level for the specific temperature or temperatures below,  $\Sigma n_f$  including all values corresponding to fracture toughness below the selected fracture toughness level for the specific temperature or temperatures above and  $\Sigma n$  as the sum of  $\Sigma n_s$  and  $\Sigma n_f$ .  $\Sigma n_f$  represents  $i$  and  $\Sigma n$  represents  $n$  in Eq. (4). The treatment of censored values is simple. The censored value may contribute to  $\Sigma n_s$  but not to  $\Sigma n_f$ . The resulting probability diagram for the 12.5 mm specimens is presented in figure 3b. Similarly as for the Rank probability diagram, it is possible to fit a specific distribution to the data or to try to collapse the data into a single trend. Here, only three different probability levels are considered (5 %, 50 % and 95 %).

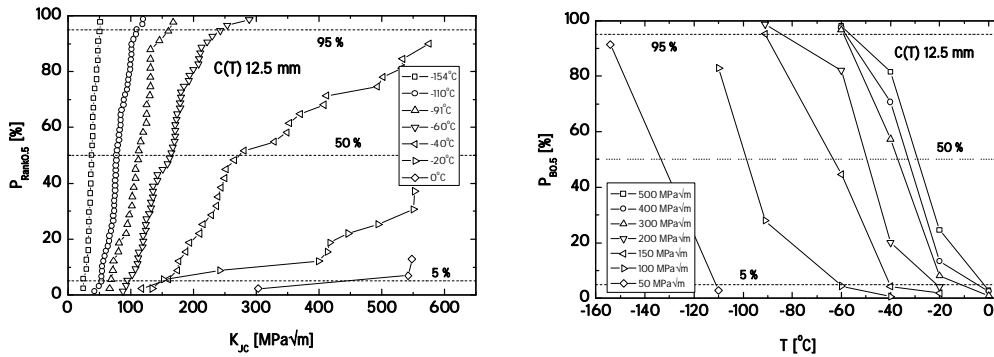


Fig. 3. Rank (a) and Binomial (b) analysis of the 12.5 mm C(T) data.

### 3. Discussion

Figure 4 shows the combined estimates for the 5 % 50 % and 95 % probabilities for all specimen sizes. As seen from the figure, the two different estimates complement each other very well. Free hand fits to the three probability estimates are compared with the result of a standard ASTM E1921 Master Curve fit to the data shown in figure 2. Within the validity window of the Master Curve, the Rank and Binomial probability based estimates have nearly perfect correspondence with the Master Curve fit. The 100 mm specimens indicate a smaller scatter than predicted by the Master Curve. This may be partly due to the scarce number of data values in the Master Curve temperature validity region. Above the measuring capacity limit the Rank and Binomial probability based estimates for 50 % and 95 % probability begin to rise above the Master Curve. This is due to a loss of constraint in the specimens. The 5 % estimates rise above the Master Curve at a lower fracture toughness level. This is indicative that in this region, the simple assumption of a constant lower bound fracture toughness, close to 20 MPa√m, is no longer valid.

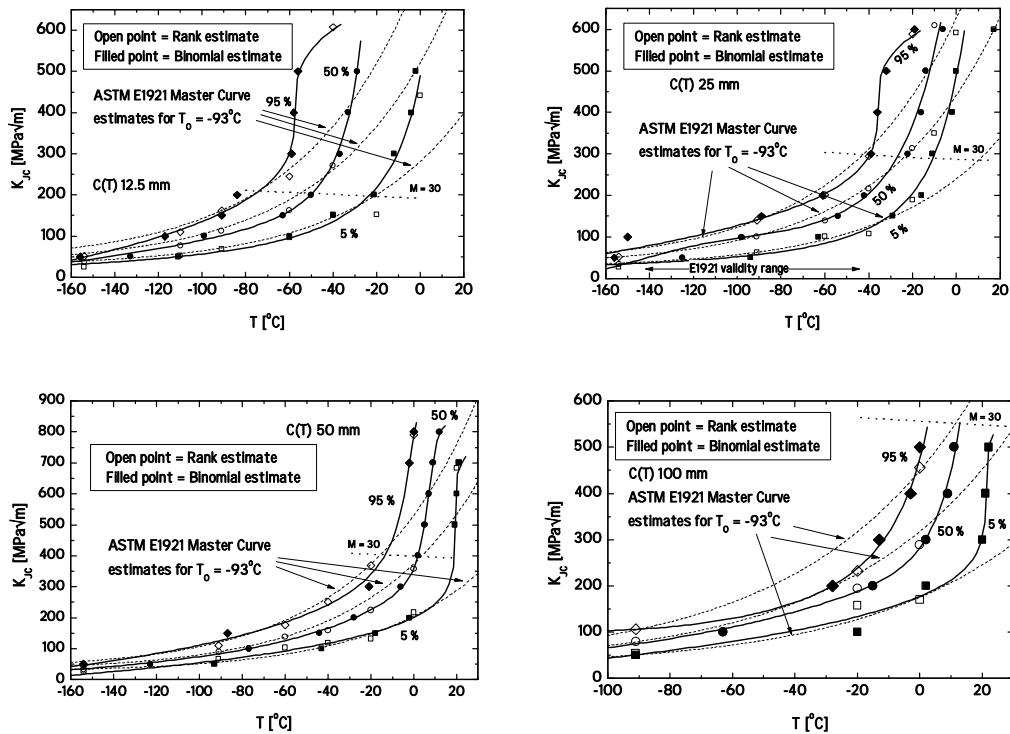


Fig. 4. Resulting combined estimates for the 12.5 mm (a), 25 mm (b), 50 mm (c) and 100 mm (d) C(T) data.

The results can be used for a direct comparison of different size/type specimens similarly to what was done using distribution comparison (Q-Q plots) for the same data set [5]. The method is not restricted to fracture toughness data. It can be applied to any data expressed in the form of two parameters.

## Acknowledgements

This work is part of the authors Academy Professorship and is funded by the Academy of Finland decision 117700.

## References

- [1] David HA, Nagaraja HN. *Order statistics*, 3rd ed. Hoboken, New Jersey: John Wiley & Sons; 2003.
- [2] Lipton C, Sheth NJ. *Statistical design and analysis of engineering experiments*. New York: McGraw-Hill; 1973.
- [3] Heerens J, Hellmann D. Development of the Euro fracture toughness dataset. *Engng Fract Mech* 2002;**69**:421–449.
- [4] Wallin K. Master Curve analysis of the "Euro" fracture toughness dataset. *Engng Frac Mech* 2002;**69**:451 - 481.
- [5] Wallin K. Quantitative experimental method for studying brittle fracture constraint loss in CT and 3PB specimens. *PVP* 2001;**423**:191–8.