

A. CIAŚ\*, A. CZARSKI\*

**THE USE OF WEIBULL STATISTICS TO QUANTIFY PROPERTY VARIABILITY IN Fe-3Mn-0.8C SINTER-HARDENED STRUCTURALLY INHOMOGENEOUS STEELS****ZASTOSOWANIE STATYSTYKI WEIBULLA DO ILOŚCIOWEGO OPISU ZMIENNOŚCI WŁASNOŚCI STRUKTURALNIE NIEJEDNORODNEJ STALI Fe-3Mn-0.8C WYTWORZONEJ TECHNIKĄ SINTER-HARDENING**

Low carbon ferro-manganese and graphite powders were admixed to Höganäs sponge, NC100.24, and water atomised, ABC 100.30 and ASC 100.29, iron powders – to produce three variants of sintered Fe-3Mn-0.8C steel. These were pressed into tensile and bend specimens at 660 MPa, sintered in semi-closed containers for 1 hour in dry nitrogen or hydrogen at 1120 or 1250°C and cooled at 64°C/min. Both tensile strength and transverse rupture strength were examined using Weibull statistics. This paper presents the results of a study to develop and evaluate goodness of fit tests for the two- and three-parameter Weibull distributions. The study was initiated because of discrepancies in published critical values for two-parameter Weibull distribution goodness of fit tests and the lack of general three-parameter Weibull distribution goodness of fit tests for properties of PM steels.

*Keywords:* strength of sintered steels, Weibull statistics, goodness of fit tests

Aby uzyskać trzy warianty spiekanej stali Fe-3Mn-0.8C wytworzono mieszanki proszków: niskowęglowego żelazomanganu, grafitu oraz proszków żelaza firmy Höganäs: gąbczastego – NC100.24 i rozpylanych – ABC 100.30 oraz ASC 100.29. Z proszków tych zostały sprasowane, przy ciśnieniu 660MPa, kształtki próbek wytrzymałościowych przeznaczonych do wykonania prób rozciągania i zginania, spieczone następnie w półhermetycznych pojemnikach, w czasie 1 godziny, w suchym azocie lub wodorze, w temperaturze 1120 lub 1250°C, po czym chłodzone z szybkością 64°C/min. Zbadano wytrzymałość na rozciąganie i zginanie otrzymanych próbek przy zastosowaniu statystyki Weibulla. W artykule zaprezentowano wyniki badań mających na celu wykorzystanie i ocenę testów zgodności dla dwu- i trzyparametrowych rozkładów Weibulla. Badania podjęto z powodu rozbieżności publikowanych krytycznych wartości testów zgodności dla dwuparametrowego rozkładu Weibulla oraz ogólnego braku testów zgodności dla trzyparametrowego rozkładu Weibulla, w odniesieniu do własności stali wytworzonych techniką metalurgii proszków.

**1. Background and Introduction**

Increased iron powder compressibility, resulting in higher green and sintered densities of identically processed powder metallurgy (PM) steels of the same chemical composition, generally results in better mechanical properties. This is exemplified in numerous reports, including manufacturers' data sheets for Höganäs powders used in this investigation. These data should be contrasted with results on sinter-hardened manganese steels, for which higher strengths and ductilities have been reported to result from the use of iron sponge, to which ferro-manganese and graphite were added [1]. Accordingly it was decided to reinvestigate this "anomalous" behaviour of PM manganese-containing steels, paying particular attention to the role of manganese vapour in sintering, especially when using semi-closed containers and flowing atmosphere [2].

The mechanical strength of sintered steels has always been an important issue in PM. When analyzing PM steels

mechanical strength data, our goal is simple: we wish to make the strongest possible conclusion from limited amounts of data. To do this, we need to overcome two problems:

- important differences can be obscured by material properties scatter;
- differences can be obscured by experimental imprecision.

This makes it hard to distinguish real differences from random variability [3, 4].

Scientists care about small differences and are faced with large amounts of variability. Statistical methods are necessary. Standardized tests are performed to determine the various types of strength, for example tensile or transverse rupture strength. Statistical analyses are most useful when we are looking for differences that are small compared to experimental imprecision and scatter of property. Our natural inclination (especially with our own data) is to conclude that differences are real, and to minimize the contribution of random variability. Statistical rigor prevents us from making this mistake.

\* AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY, FACULTY OF METAL ENGINEERING AND INDUSTRIAL COMPUTER SCIENCE, DEPARTMENT OF PHYSICAL METALLURGY AND POWDER METALLURGY, AL. A. MICKIEWICZA 30, 30-059 KRAKÓW, POLAND

Transverse rupture strength (TRS) and ultimate tensile strength (UTS) of PM sinter-hardened manganese steels typically show scatter that is assumed to be caused by flaws (pores, inclusions and other microstructural imperfections) and a Weibull distribution can be used to model the variation [5, 6, 7].

Weibull distribution is usually used to describe the mechanical properties of brittle materials. Mechanical properties' data for wrought metals and alloys is most often fitted to a Gaussian distribution, and subsequently represented by the mean and standard deviation. While the Gaussian distribution is often taken as the accepted statistical distribution for failure strength of structural low-alloy steels, there is no theoretical or experimental justification for this situation, because Gaussian distributions have tails that reach to infinite values, inconsistent with the behaviour of most metals [8, 9]. The Weibull cumulative distribution is sigmoidal like a Gaussian distribution, but is skewed. Use of the Weibull distribution provides accurate failure analysis and risk predictions with extremely small samples using simple and useful graphical plots [7]. Solutions are possible at the earliest stage of a problem without the requirement to "crash a few more". Small samples also allow cost-effective component testing [10].

Two-parameter (2-p) and three-parameter (3-p) Weibull distributions are occasionally used to represent the strength distribution of PM structural parts and engineering-designed sintered steels subassemblies. In its 3-p form, the cumulative probability of failure of a material subjected to a stress  $\sigma$  can be represented as:

$$P_f = 1 - P_s = 1 - \exp\left[-\left(\frac{\sigma - \sigma_u}{\sigma_r}\right)^m\right] \text{ for } \sigma \geq \sigma_u \quad (1)$$

and  $P_f = 0$  for  $\sigma < \sigma_u$ ,

where:

$P_f$  is the failure probability – cumulative distribution function (*cdf*),  $P_s$ , the survival probability,  $\sigma$ , the failure stress,  $\sigma_r$ , a material parameter,  $m$ , the Weibull modulus and  $\sigma_u$  is the threshold stress below which the failure probability is zero.

The 3-p Weibull distribution converts failure data by taking  $(\sigma_i - \sigma_u)$  of each point and plotting the transformed data. This distribution can be expressed also in the form:

$$P_f = 1 - P_s = 1 - \exp\left[-\left(\frac{\sigma - \sigma_u}{\sigma_0 - \sigma_u}\right)^m\right] \text{ for } \sigma \geq \sigma_u \quad (2)$$

and  $P_f = 0$  for  $\sigma < \sigma_u$ ,

where:

$$\sigma_0 - \sigma_u = \sigma_r = \frac{E - \sigma_u}{\Gamma(1 - 1/m)} \quad (3)$$

and  $\Gamma$  is the Euler gamma function [10],  $\sigma_0$ , a characteristic stress at which a fraction  $1/e$  of specimens survives and  $E$  is the expected value of the distribution function.

With respect to the 3-p form, it is necessary to select a threshold stress that represents minimum failure strength for a given alloy. This was considered e. g. by Newkirk and Thakur for PM parts [11] and Biery et. al. for cast TiAl alloys [12]. The former authors used statistical procedures to assign the threshold stress value, whilst the latter took it to be the 0.2% offset yield strength. For less ductile PM steels this can be insufficiently conservative.

If the threshold stress  $\sigma_u$  is taken to be zero (the most conservative Weibull analysis and the conventional procedure for ceramic, fibrous and composite materials), the analysis becomes the well-established 2-p form, where, for uniaxially and uniformly stressed tensile specimen, the failure probability is:

$$P_f = 1 - \exp\left[-\left(\frac{\sigma}{\sigma_0}\right)^m\right] \text{ for } \sigma > 0 \text{ and } P_f = 0 \text{ for } \sigma \leq 0. \quad (4)$$

The threshold value could be used to provide a minimum property for the design of PM parts and would allow also a transfer of strength data of laboratory tensile or bend specimens to components, where the stress distribution is much more complex. As sintered steel construction practices in the EU and US are revised from deterministic to reliability-based design procedures, assessing the goodness of fit of these Weibull distributional forms becomes increasingly important. Goodness of fit tests for the two-parameter Weibull distributions have received considerable attention. Despite this extensive literature, we encountered difficulties when we studied the goodness of fit of two- and three-parameter Weibull distributions to numerous data sets consisting of 30 to 100 observations of various PM steels' strength properties. For the 3-p Weibull tests, critical values are not published for the sample sizes involved, and it is not clear how estimated shape parameters would affect the critical values derived by assuming a known shape parameter. For the 2-p Weibull tests, these difficulties included a lack of published critical values for PM steels sample sizes larger than 30 and some apparent inconsistencies in published critical values.

The 2-p Weibull distribution is adequate for a majority of Weibull analysis scenarios. However, if the transformed failure data plot has a curved rather than a straight line appearance, or if Weibull modulus is found to be greater than 6.0, then a third parameter may be needed to adequately model the data. The third parameter effectively shifts the entire distribution to the right. In practice, this can be interpreted as the lowest possible stress at which failure may occur. Of course, it may never be larger than the value of the lowest failure stress from the data set. Articles [13] and [14] provides guidance on fitting a three-parameter Weibull model.

In this paper we develop extensive and definitive 2-p and 3-p Weibull distribution goodness of fit critical values for Anderson-Darling (A-D) statistic to yield a smoothed estimate of the critical values for the statistics, and extend this statistics for use.

## 2. Experimental materials and procedure

Höganäs sponge, NC100.24, and water atomised ABC 100.30 and ASC 100.29 iron powders were the starting materials in this investigation. Typically 0.8% of carbon was introduced as fine Höganäs CU-F graphite and 3% of manganese as Elkem low carbon ferro-manganese of weight % composition 80Mn-1.3C-0.2O-balance Fe. Double-cone mixing and die compaction of 120 ISO 2740 dog bone specimens at 660 MPa, using die lubrication, were followed by sintering in a semi-closed stainless steel container in a horizontal laboratory furnace. The dew point of the sintering atmospheres,

hydrogen or nitrogen, was  $-60^{\circ}\text{C}$  (15 ppm moisture). Compacts were heated to the sintering temperature at a rate of  $75^{\circ}\text{C}/\text{min}$ , held at 1120 or  $1250^{\circ}\text{C}$  for 60 minutes and cooled, in the temperature range of 1100- $500^{\circ}\text{C}$ , at approximately  $64^{\circ}\text{C}/\text{min}$ . After sintering all the specimens were tempered at  $200^{\circ}\text{C}$ . Standard PN-EN ISO 2740 tensile specimens were tested for the ultimate tensile strength (UTS) on an MTS 810 servo-hydraulic machine at a strain rate of  $5 \times 10^{-4} \text{ s}^{-1}$ . The yield strength was measured at the 0.2% offset strain. The same specimen types were tested in three-point bending to determine the apparent rupture strength, TRS, for specimens of beam depth  $t = 6 \text{ mm}$  and width  $w = 6 \text{ mm}$ , using span length  $l$  of 28.6 mm. A plasticity correction was made to calculate the bend strength [7].

Chemical analyses for oxygen and carbon in the iron starting powders and in the sintered alloys, also for nitrogen, were carried out on Leco apparatus, TC-336 and CS-125, respectively. For all iron powders these were  $\sim 0.2\%$  O and  $\sim 0.01\%$  C. An aim of the investigation was to attain in all sintered alloys C $\sim 0.6\%$ . Metallographic and fractographic examinations were finally carried out, but this is not the subject matter of this paper.

### 3. Results

Within the (small) experimental error, the densities of Fe-Mn-C compacts remained unchanged on sintering. For ABC 100.30 based steel it was  $\sim 7.1 \text{ g}/\text{cm}^3$ .  $>0.2 \text{ g}/\text{cm}^3$  larger than for the sponge-based alloy, Table 1, which also lists the results of the C and N analyses of sintered samples. Carbon content was generally near 0.6%; but in the specimens sintered at  $1120^{\circ}\text{C}$  in nitrogen, however, carbon content was  $\sim 0.7\%$ .

#### 3.1. Mechanical properties

The values of UTS and TRS were determined for at least 15 specimens of each batch. Using the assumption of a normal distribution allows a mean and standard deviation of a random variable to be determined (Table 2).

The standard deviation is a "natural" measure of statistical dispersion if the centre of the data is measured about the mean. However, if the distribution of failure loads fits a Weibull distribution, it is different from that predicted by a Gaussian distribution, so there is no proper standard deviation.

TABLE 1

Densities, carbon and nitrogen contents after sintering of Fe-3Mn-0.6C steel

Sintering	Carbon and nitrogen contents in % and sintered density in g/cc						
	NC 100.24			ABC 100.30		ASC100.29	
	Carbon content, wt. %	Nitrogen content, wt. %	Sintered density $\text{gcm}^{-1}$	Carbon content wt. %	Sintered density $\text{gcm}^{-1}$	Carbon content wt. %	Sintered density $\text{gcm}^{-1}$
1120°C, H <sub>2</sub>	0.62	0.0036 $\pm 0.0001$	6.89 $\pm 0.03$	0.59	7.11 $\pm 0.02$	0.63	6.93 $\pm 0.02$
1120°C, N <sub>2</sub>	0.72	0.0332 $\pm 0.005$	6.91 $\pm 0.02$	0.71	7.09 $\pm 0.06$	0.70	6.94 $\pm 0.01$
1250°C, H <sub>2</sub>	0.52	0.0032 $\pm 0.0005$	6.91 $\pm 0.03$	0.54	7.09 $\pm 0.04$	0.54	6.93 $\pm 0.03$
1250°C, N <sub>2</sub>	0.57	0.0324 $\pm 0.001$	6.91 $\pm 0.01$	0.59	7.09 $\pm 0.04$	0.58	6.92 $\pm 0.02$

$\pm$  – standard deviation measured on 15 samples.

TABLE 2

UTS and TRS of Fe-3Mn-0.8C steel

Sintering	Stress in MPa					
	NC 100.24		ABC 100.30		ASC100.29	
	UTS	TRS	UTS	TRS	UTS	TRS
1120°C, H <sub>2</sub>	551 $\pm$ 52	1120 $\pm$ 138	524 $\pm$ 74	1073 $\pm$ 111	475 $\pm$ 39	959 $\pm$ 141
1120°C, N <sub>2</sub>	683 $\pm$ 50	1357 $\pm$ 113	603 $\pm$ 74	1128 $\pm$ 89	565 $\pm$ 38	1057 $\pm$ 152
1250°C, H <sub>2</sub>	713 $\pm$ 68	1337 $\pm$ 150	650 $\pm$ 76	1250 $\pm$ 196	578 $\pm$ 64	996 $\pm$ 93
1250°C, N <sub>2</sub>	732 $\pm$ 53	1354 $\pm$ 121	651 $\pm$ 67	1244 $\pm$ 184	626 $\pm$ 45	1109 $\pm$ 134

$\pm$  – standard deviation measured on 15-26 samples.

Fracture stress, also known as fracture strength, is the minimum tensile stress that will cause fracture. It is well-recognized that the value TRS can exceed that of UTS, of the same PM material, identically processed, by a factor up to ~2, although both these parameters relate to the tensile stress causing fracture [8].

#### 4. Parameters estimation

There are number methods for determination of Weibull parameters from mechanical strength measurements, Zanakis [15] documents 17 different methods of obtaining the parameters of the 3-p Weibull distribution, but only two are in common use. Two common parameters estimation methods used in engineering are the least-squares linear regression (LR) and maximum likelihood method (MLE), which are the methods considered in this paper.

The threshold value was calculated using maximum likelihood estimation (MLE). For known experimental failure data  $\sigma_i$  ( $i = 1, 2, \dots, N$ ), the parameters  $\sigma_u$ ,  $\sigma_r$  and  $m$  were determined by maximisation of the likelihood probability density function:

$$L = \prod_{i=1}^N f(\sigma_i; \sigma_u; \sigma_r; m) \quad (5)$$

This method requires solving the three simultaneous equations. For calculation of the  $\sigma_u, \sigma_r$  and  $m$  from the 3-p Weibull distribution following equations were used:

$$N\sigma_0^m - \sum_{i=1}^N (\sigma_i - \sigma_u)^m = 0 \quad (6)$$

$$\frac{N}{m} - \sum_{i=1}^N \left[ \left( \frac{\sigma_i - \sigma_u}{\sigma_r} \right)^m \ln \left( \frac{\sigma_i - \sigma_u}{\sigma_r} \right) \right] - \sum_{i=1}^N \ln \left( \frac{\sigma_i - \sigma_u}{\sigma_r} \right) = 0 \quad (7)$$

$$\left( 1 - \frac{1}{m} \right) \sigma_r^m \sum_{i=1}^N \frac{1}{\sigma_i - \sigma_u} - \sum_{i=1}^N (\sigma_i - \sigma_u)^{m-1} = 0 \quad (8)$$

Equations (6-8) were solved for  $m$  with an iterative Newton-Raphson method. Estimation of the Weibull parameters was performed under the constraints that estimators  $m > 1$  and  $0 < \sigma_u < \min \{\sigma_1, \dots, \sigma_N\}$ . The second condition stems from the observation that, within the setting of the theory, a failure stress smaller than  $\sigma_u$  is absolutely impossible.

The maximum likelihood principle for parameter estimation is intuitively appealing, but for small sample sizes it has been shown that the MLE method gives a biased estimate of the Weibull modulus [15]. No information is available about the bias of the MLE estimates of the parameters of the Weibull distribution, describing the strength of PM steels. However, assuming that known results about the bias of the three-parameter Weibull distribution also apply to the present study, the reference value is expected to be biased only slightly, and the Weibull modulus  $m$  is underestimated.

For the sake of completeness, use will also be made of the less complex 2-p Weibull analysis. For uniaxially and uniformly stressed tensile specimen the failure probability is given by the Eq. (4).

Although the parameters for the 2-p Weibull distribution could be determined using a least-squares fit with a weight

function on the linearized Weibull equation (linear regression – LR), the MLE method is generally used. This method for the best estimate of parameters  $\sigma_0$  and  $m$  shows the smallest coefficients of variation (the ratio of the standard deviation and mean of a random quantity). The maximum likelihood method was used to determine parameters for the 2-p Weibull distribution setting  $\sigma_u = 0$  and only using two equations. This method finds values of  $m$  and  $\sigma_0$  and predicts with the highest probability the measured distribution of strengths. Although this approach has the advantage that it gives the minimum estimation error when the highest and lowest values of strength completely predominate in the analysis, it can lead to serious errors in  $m$  values. Since abnormal low or high values of strength can easily arise in concentrations in grips, local friction in bending tests, etc., this represent a serious drawback. The likelihood of a given probability density function is defined as  $L = \prod_{i=1}^N f(\sigma_i; \sigma_0; m)$  and thus its log-likelihood function is  $\ln L = \sum_{i=1}^N f(\sigma_i; \sigma_0; m)$ , where  $N$  is the number of strength experiments (specimens). Thus, estimates of these parameters can be found by maximising the log-likelihood function. For the 2-p Weibull distribution, the equation for determining  $m$  from  $N$  measured strengths  $\sigma_i$  is:

$$\frac{\sum_{i=1}^N \sigma_i^m \ln \sigma_i}{\sum_{i=1}^N \sigma_i^m} = \frac{1}{m} + \frac{1}{N} \sum_{i=1}^N \ln \sigma_i \quad (9)$$

where  $m$  can be obtained by an iterative procedure. and then  $\sigma_0$  is calculated by

$$\sigma_0^m = \frac{1}{N} \sum_{i=1}^N \sigma_i^m \quad (10)$$

The second method of estimating the parameters of the Weibull 2-p distribution that be of use is LR. For a constant tested volume (specimen size and gauge length) it is often calculated using a LR with a weight function on the linearized Weibull equation:

$$\ln(\ln(1/1 - P_f)) = m \ln \sigma - m \ln \sigma_0 = \ln(\ln(1/P_s)) = m \ln \sigma - k \quad (11)$$

The Weibull modulus can be determined by plotting  $\ln(\ln(1/1 - P_f))$  against  $\ln \sigma$ . It can be obtained directly from the slope term in Eq. (11), and scale parameter can be deducted from the intercept term. Since the true value of  $P_f$  for each  $\sigma_i$  is not known, a prescribed probability estimator has to be used as the  $P_i$ -value, where  $P_i$  is the probability of failure for the  $i^{\text{th}}$ -ranked stress datum. The probability estimator has significant effect on the estimation precision of the Weibull parameters in the LR method. Several expressions are applied to define the probability estimator. The relative merits of these estimators have been investigated by several authors with actual computer-generated strength data [16, 17]. It has been shown that the optimal probability estimator determined varies with the sample size. It has been shown also that the MLE method results in the highest precision of estimation with a lower safety than the LR [18]. In this study the survival probability,  $P_s$ , and failure probability,  $P_f$ , were estimated by Bernard's formula:  $P_s = 1 - P_f = 1 - [(i-0.3)/(N+0.4)] =$

$(N - i + 0.7)/(N + 0.4)$  where  $n$  is the total number of specimens and  $i$  is the rank number. We use Bernard's median rank because it shows the best performance and it is the most widely used to estimate the probability of failure. The weight function used was  $w_i = [(1 - P_f) \ln(1 - P_f)]^2$ . In statistics, there are many methods of measuring goodness to fit, but some authors prefer the simple correlation coefficient. It is ideal for testing the goodness of fit to a straight line. The Pearson's correlation coefficient –  $R$  is intended to measure the strength of a linear relationship between two variables of the Weibull plot. As Weibull plots always have positive slopes, they will always have positive  $R$  [7, 8]. Whatever the bias of the estimates may be, the plots of the estimated failure probability of the UTS and TRS specimens fitted the experimental data quite well, e.g. Fig. 1.

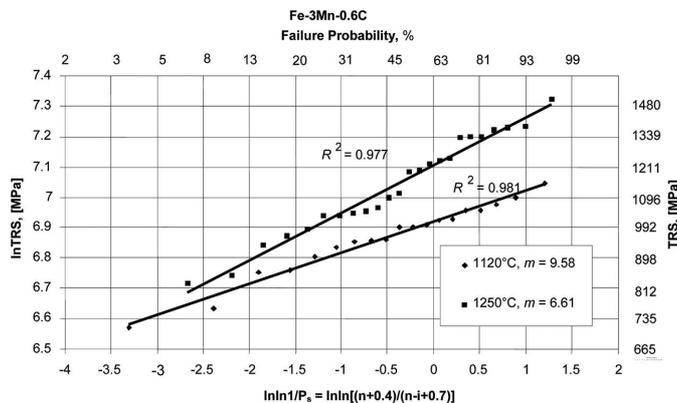


Fig. 1. 2-p Weibull plots;  $m$  (given by slope),  $R^2$  (Pearson's correlation coefficient), and  $\sigma_0$  1211 MPa (1250°C) and 1002 MPa (1120°C) were obtained using regression tool and Bernard's formula:  $P_s = 1 - P_f = 1 - [(i - 0.3)/(n + 0.4)] = (n - i + 0.7)/(n + 0.4)$

The value of Weibull modulus, which can be calculated from a test group of specimens, enables the 'dependability' of a material to be evaluated numerically. Discontinuities present in the Weibull graphs are expected to be linked to different defect populations.

Description of data sets parameters for 2-p and 3-p Weibull are given in Table 3.

## 5. Goodness of fit

The goodness of fit of a statistical model describes how well it fits a set of observations. Many methods, such as the Kolmogorov-Smirnov (K-S), Chi-square, the Anderson-Darling (A-D) tests, exist for determining the goodness to fit of a probability distribution to a set of data. Measures of goodness of fit typically summarize the discrepancy between observed values and the values expected under the model in question.

The A-D test was used to test the hypothesis that a random sample  $X_1, \dots, X_i, \dots, X_N$ , with strength empirical distribution  $F_n(\sigma)$ , comes from a continuous population with distribution function  $F(\sigma)$  where  $F(\sigma) = F_0(\sigma)$  for some completely specified distribution function  $F_0(\sigma)$ . The A-D test was chosen for this study as it is more sensitive to the tail behaviour [19] and

has been recommended for statistical analysis of strength of materials. The sensitivity to the tail behaviour is particularly useful in structural engineering applications, where the tail is important in computing the mechanical reliability [20, 21]. If a sample of data came from a population with a specific distribution, typically the A-D test is used. It is a modification of the K-S test and gives more weight to the tails than does the K-S test.

The A-D statistic ( $A_N^2$ ) is defined as:

$$A_N^2 = N \int_{-\infty}^{+\infty} \frac{[\tilde{F}_N(\sigma) - F_0(\sigma)]^2}{F_0(\sigma)[1 - F_0(\sigma)]} dF_0(\sigma) \quad (12)$$

where  $\tilde{F}_N(\sigma)$  is a step function that jumps at the order statistics of  $\sigma$ , and  $\tilde{F}_0(\sigma)$  is the hypothesized continuous *cdf*. The A-D statistic is a measure of the square of the error between the data and the hypothesized distribution weighed so that the tails of the data are more important than the central portion.

In this and study we develop calculations using the A-D test for testing the 2-p and 3-p Weibull assumption. The A-D goodness to fit test for normality [18] has the functional form:

$$A_N^2 = \sum_{i=1}^N \left[ \frac{1 - 2i}{N} \{ \ln[F_0(\sigma_{(i)})] + \ln[1 - F_0(\sigma_{N+1-i})] \} \right] - N \quad (13)$$

where  $F_0$  is the assumed (Weibull) distribution with the assumed or sample estimated parameters.  $\sigma(i)$  is the  $i$ -th sorted, standardized, sample value (the  $i$ -th order statistic of the data set; "N" is the sample size and subscript "i" runs from 1 to N).

For the Weibull distribution, an observed significance level (OSL) is obtained as follows [20]:

$$OSL = \frac{1}{1 + \exp[-0.10 + 1.24 \ln AD^* + 4.48 AD^*]} \quad (14)$$

in which

$$AD^* = \left[ 1 + \frac{0.2}{N} \right] A_N^2 \quad (15)$$

The OSL is the probability of obtaining a value of the test statistic at least as large as that obtained from the data if the hypothesis that the data are actually from the distribution being tested is true. Typically, a 5% significance level is used. The convention is to argue that any event or observation, whose likelihood of happening by chance alone is less than five times in 100 is statistically significant. This is expressed as a probability of 0.05 or as a percentage value of 5%. So, for any experiment we perform or observation study we carry out, if the probability of an observed event falls below the 5% significance level, we can argue a statistically significant effect. The null hypothesis, that the true distribution is  $F_0$  with the assumed parameters, is only rejected if the OSL is less than 0.05 for sample size  $n$ .

The OSL was obtained for each of the distributions for each data set. P-values and likelihood ratio test (LRT) results were obtained also for each of the distributions for each data set. The results are shown in Table 3.

TABLE 3

Description of data sets parameters for 2-p and 3-p Weibull distributions, P-values, OSL using the Anderson-Darling test and LRT

No	Set of specimens	Number of spec.	2-p Weibull					3-p Weibull						LRT
			<i>m</i>	$\sigma_0$ , MPa	A-D	P-v (A-D)	OSL	<i>m</i>	$\sigma_T$ , MPa	$\sigma_w$ , MPa	A-D	P-v (A-D)	OSL	
1	UTS_ABC_1120_H2	15	6.98	556	0.838	0.025	0.024*	2.60	200	345	0.561	0.125	0.131	0.087
2	TRS_ABC_1120_H2	29	12.25	1120	0.142	>0.250	0.860	3.93	422	975	0.518	>0.500	0.177	0.761
3	UTS_ABC_1120_N2	15	9.12	636	0.582	0.123	0.116	2.00	156	465	0.370	0.446	0.384	0.128
4	TRS_ABC_1120_N2	27	12.83	1059	0.361	>0.250	0.410	2.85	254	793	0.213	>0.500	0.727	0.14
5	UTS_ABC_1250_H2	15	10.72	683	0.494	0.208	0.195	9.48	607	76	0.501	0.129	0.188	0.931
6	TRS_ABC_1250_H2	25	6.74	1222	0.348	>0.250	0.435	3.02	582	622	0.328	0.478	0.476	0.343
7	UTS_ABC_1250_N2	15	12.11	680	0.467	0.234	0.228	9.43	533	146	0.480	0.147	0.212	0.895
8	TRS_ABC_1250_N2	29	8.12	1322	0.540	0.171	0.156	5.07	853	462	0.516	0.128	0.179	0.719
9	UTS_ASC_1120_H2	14	13.60	493	0.286	>0.250	0.559	4.00	149	340	0.251	>0.500	0.638	0.365
10	TRS_ASC_1120_H2	22	7.34	1020	0.396	>0.250	0.342	1.39	228	750	0.679	0.082	0.066	<b>0.029**</b>
11	UTS_ASC_1120_N2	15	17.17	583	0.654	0.079	0.075	1.67	69	503	0.575	0.142	0.121	0.122
12	TRS_ASC_1120_N2	24	6.85	1124	1.173	<0.010	0.004*	1.96	316	776	0.536	0.18	0.158	<b>0.006**</b>
13	UTS_ASC_1250_H2	15	10.22	607	0.435	>0.250	0.273	2.28	150	445	0.345	0.477	0.433	0.21
14	TRS_ASC_1250_H2	22	11.92	1037	0.307	>0.250	0.520	4.46	395	635	0.211	>0.500	0.729	0.299
15	UTS_ASC_1250_N2	14	17.44	646	0.241	0.02	0.660	2.00	93	543	0.269	0.037	0.597	0.507
16	TRS_ASC_1250_N2	20	8.70	1181	0.333	>0.250	0.463	2.18	313	843	0.168	>0.500	0.813	0.065
17	UTS_NC_1120_H2	14	13.27	574	0.241	>0.250	0.660	6.59	294	278	0.269	>0.500	0.597	0.797
18	TRS_NC_1120_H2	20	8.70	1181	0.333	>0.250	0.463	2.18	313	843	0.168	>0.500	0.813	0.065
19	UTS_NC_1120_N2	15	26.91	698	0.158	>0.250	0.829	8.13	217	480	0.177	>0.500	0.794	0.71
20	TRS_NC_1120_N2	15	13.43	1407	0.536	0.168	0.153	3.92	421	976	0.498	0.152	0.191	0.317
21	UTS_NC_1250_H2	15	12.12	743	0.273	>0.250	0.589	2.61	176	557	0.287	>0.500	0.558	0.288
22	TRS_NC_1250_H2	23	9.69	1404	0.401	>0.250	0.334	2.36	366	1013	0.243	>0.500	0.661	0.082
23	UTS_NC_1250_N2	15	15.85	756	0.373	>0.250	0.378	2.16	118	628	0.322	>0.500	0.481	0.199
24	TRS_NC_1250_N2	26	10.64	1300	0.561	0.149	0.137	2.47	309	971	0.340	0.474	0.452	<b>0.036**</b>

\* – OSL<0.05 2-p; the 2-p Weibull distribution does not provide a reasonable fit.

\*\* – LRT<0.05 (the significance level = 0.05); it is recommended that the 3-p Weibull be used to characterise PM steel properties.

The P-value is a probability, with a value ranging from zero to one. It is the answer to this question: if the populations really have the same mean overall, what is the probability that random sampling would lead to a difference between sample means as large (or larger) than we observed? P-value is “the probability, if the test statistic really were distributed as it would be under the null hypothesis, of observing a test statistic (...) the one actually observed” [19]. The smaller the P-value, the more strongly the test rejects the null hypothesis, that is, the hypothesis being tested. Intuitively, we think that P-value=0.0001 is more statistically significant than P-value=0.04. Using strict definitions, this is not correct. Once we have set a threshold P-value for statistical significance, every result is either statistically significant or is not statistically significant.

## 6. Evaluation of the results and discussion

The 2-p Weibull distribution cannot be rejected at the 5% significance (OSL>0.05) in 22 out the 24 cases. The 3-p Weibull distribution cannot be rejected at conventional significance level 0.05 in any case (Table 3). The average OSL for 2-p Weibull distribution was 0.369, with the average OSL for the 3-p Weibull distribution being 0.441. Although strictly speaking the OSL cannot be used for ranking distributions, higher values of the OSL do indicate a higher significance level. Therefore it appears that the 3-p Weibull distribution is slightly preferable to the 2-p, though the 2-p Weibull distribution cannot be rejected in most cases.

However, the likelihood ratio test is >0.05 in 21 out the 24 cases. LRT was a statistical test used to compare the fit of two models, one of which (the 2-p Weibull model) is a special case of the other (the 3-p Weibull). The test is based on the likelihood ratio, which expresses how many times more

likely the data are under one model than the other. This likelihood ratio, or equivalently its logarithm, can then be used to compute a P-value, or compared to a critical value to decide whether to reject the null model in favour of the alternative model. The likelihood ratio test rejects the null hypothesis if the value of this statistic is too small. How small is too small depends on the significance level of the test. When  $LRT > 0.05$  (the significance level = 0.05) it is recommended that the 2-p Weibull be used to characterise PM steel properties.

The objectives of this paper are to present the feasibility of utilizing the Weibull distribution to predict the manganese steel strength performance from laboratory test results. Further, the proposed approach for uniaxial conditions has been extended to three point bending test stress conditions. The importance of understanding the complex interaction of processing parameters and mechanical properties increases, as powder metallurgy continues to increase in acceptance as a method for fabricating complex near net shaped parts. The Weibull theory approach has potential for use in recursive mechanistic-empirical design procedure. A very wide scatter is usually observed in laboratory strength test (tensile and transverse rupture) data of sinter-hardened Mn steel specimens due to randomness in the number, orientation and distribution of pores and micro-voids [22]. This leads to an uncertainty in choosing the representative design strength which leads to the need for a probabilistic approach to the analysis of test data.

The characterization of strength of the semi-brittle sinter-hardened steels is problematic due to the scatter of test results. Several proposals using Weibull statistics have been made, some of them in terms of UTS and others in terms of TRS; some authors employ a two parameters Weibull function, while others use a three parameters function. An analysis about the relationship between Weibull distributions expressed in terms of UTS and TRS is presented in this paper. It is shown that, if the UTS results follow a 3p-Weibull, their equivalent TRS values do not exactly fit a 3p-Weibull function obtained by means of MLE. Nevertheless, an approximated 3p-Weibull function in TRS terms is proposed in this work. It fits very well with the corrected values and their parameters are related to those expressed in TS terms.

Newkirk et al. [11] compared the 2-p and 3-p Weibull distribution for TRS of PM steel parts. They discussed the use of Weibull statistics to analyse the properties of PM parts, and suggest new ways to determine property variability for design application. The 2-p and 3-p Weibull distributions were compared on the basis of the correlation coefficient of the best fit line and the data. This is consistent with the present results in that the 3-p Weibull distribution in general provides a slightly better fit of the data than the 2-p Weibull distribution, although the 2-p Weibull distribution provides a reasonable fit. If the 3-p Weibull provides a slightly better fit to fracture properties of PM steels data, the question must be raised as to why the 2-p Weibull distribution is used instead? The best answer available seems to be that it is simply easier to use the 2-p Weibull distribution, although it may not be accurate to do so. However, sometimes it is recommended that the 3-p Weibull distribution be used to characterise PM sinter-hardened structural steels properties.

The primary basis for this recommendation is small differences in allowable loads between the 2-p and 3-p Weibull

distributions. Similar OSL and the fact that the location parameters of the 3-p distribution is near the first order statistic are other supporting reasons for the recommendation. The last reason implies there is a load near the lowest data point that can be applied to the specimen (or structural member) for which there is no chance of failure. This seems counterintuitive, as it seems reasonable that there would be some chance of failure at any load level, albeit the probability could be quite small. The 3-p Weibull analysis provides a consistent measure of variability and avoids some of drawbacks of the 2-p parameter form, but is important to note that the Weibull parameters do not capture all of the material behaviour.

When we use a three parameter Weibull the modulus,  $m$ , is less than the two parameter modulus (Table 3). The 3-p Weibull is a much more complex distribution than the two parameter and we have fixed requirements to meet before we adopt the 3-p solution. The four hard fixed rules for using 3-parameter to quantify property variability in PM inhomogeneous steel are:

- we must have  $N = 15$  or more failures, some experts say 100,
- we must be able to explain why the physics of failure support a guaranteed failure free zone,
- the 2-p plot should show curvature,
- the distribution analysis must favour the 3-p.

If we meet all these criteria above, the 3-p distribution is the best distribution and the 3-p Weibull modulus is the correct modulus.

The inherent scatter in strength and size effects generally means that reliability analysis of advanced sinter-hardened structural parts is usually more favourable than using standard safety factors. If a quasi-brittle material with an obvious scatter in tensile strength is selected for its high-strength attributes, then statistical analysis should be an integral part of the design process. But this statistical approach involves a certain risk of unacceptable performance, identified as a component's probability of failure (or alternatively, component reliability).

## 7. Conclusions

1. Ultimate tensile strength (UTS) and transverse rupture strength (TRS) test data from 24 batches of Mn sinter-hardened steel specimens have been analysed. The effect of different processing parameters on the mechanical strength of PM steels was quantified by using the Weibull analysis.
2. A Weibull analysis of the properties of several PM Mn sinter-hardened steels implies that an intrinsic flaw population of regions that accumulate premature high strain is responsible for the variability of strength for these materials. The finite capacity for plastic straining in Mn PM steels results in higher 2-p Weibull moduli than those of engineering ceramics. However, the mechanical properties (TRS and UTS) sinter-hardened manganese steels are somewhat variable, and this variability has been analysed using 2-p and 3-p Weibull analysis of distributions of failure strengths.
3. Unlike the normal distribution, which is symmetric to the mean value, the Weibull distribution is left skewed cor-

relating well with the obtained experimental strength data, which also present a similar skewed pattern. For this reason an attempt has been made to propose such an approach which employs Weibull's statistical strength theory. Data set of press and sinter-hardened Mn steels have been analysed and generally the 3-p Weibull statistics produced only slightly better fit than the 2-p Weibull statistics when describing the results of UTS and TRS testing on relatively small number of specimens.

4. The 3-p approach is needed because of nonzero minimum failure strength of the investigated materials, resulting from the finite capability for plastic strain, which makes the failure behaviour quite different from that of sintered ceramic materials.

The Weibull parameters obtained using these techniques can be used to obtain design strengths similar to  $-3\sigma$  and  $-6\sigma$  strengths. The distribution of failure loads that fit a Weibull distribution is different from the predicted by a Gauss distribution, so there is not proper standard deviation. Nevertheless, it is possible to derive a failure strength similar to the  $-6\sigma$  strength used by PM component designers. The  $-6\sigma$  strength designates a stress where there is only a  $3.4 \cdot 10^{-4}\%$  chance of failure, or  $P_f = 3.4 \cdot 10^{-6}$ . It is possible to solve for the  $-6\sigma$  equivalent strength using Eq. 1 [6]. For Weibull modulus,  $m$ , of 10 and a  $\sigma_r = \sigma_0 - \sigma_U = 200$  MPa, the formula is:

$$3.4 \cdot 10^{-6} = 1 - \exp\left(-\frac{\sigma - \sigma_U}{200}\right)^{10} \quad (16)$$

When solved for  $\sigma - \sigma_U$  implies that  $100 - 3.4 \cdot 10^{-4} = 99.99966\%$  of samples will have failure loads of 56.8 MPa higher than threshold value,  $\sigma_U$ . A less rigorous  $-3\sigma$  equivalent can be solved to show that 99.865% of samples will have failure loads of 103.29 MPa higher than threshold value.

5. As the proposed approach employs Weibull's parameters, goodness of fit tests have been performed to check the fitness of tests data to Weibull's distribution. It was found that the 3-p Weibull a little more accurately models the actual distributions. It allows for the use of a threshold value, consistent with mechanical properties of PM steels. The threshold value could be used to provide a minimum property for the design of parts where a reliable performance is required. When a reliable performance is required this value could be used to provide a minimum property for the design of structural members.
6. The 3-p Weibull is a much more complex distribution than the two parameter and we have fixed requirements to meet before we adopt the 3-p solution. However, the four hard fixed rules for using 3-parameter are: (a) first of all we must be able to explain why the physics of failure support a guaranteed failure-free zone; (b) we must have 15 or more failures; (c) the 2-p plot should show curvature; (d) the statistical distribution analysis must favour the 3-p. If we meet all these criteria above, the 3-p distribution is

the best distribution and the 3-p Weibull modulus is the correct Weibull modulus.

### Acknowledgements

This work was sponsored by the Polish Ministry of Science and Higher Education under Contract no 11.11.110.158. Appreciation is also expressed to Prof. A. S. Wronski of the University of Bradford for useful discussion.

### REFERENCES

- [1] A. C ia ś, S.C. Mitchell, K. Pilch, H. C ia ś, M. Su ł ow ski, A. W ro ń ski, Powder Metallurgy **46**, 2, 165-170 (2003).
- [2] M. Su ł ow ski, A. C ia ś, Archives of Metallurgy and Materials, **49**, 55-72 (2004).
- [3] M. Su ł ow ski, A. C ia ś, M. Stoytchev, T. Andreev, Materials Science Forum **534-536**, 753-756 (2007).
- [4] M. Su ł ow ski, A. C ia ś, H. Frydrych, J. Frydrych, I. Olszewska, R. Goleń, M. Sowa, Materials Science Forum **534-536**, 757-760 (2007).
- [5] P. Dudek, A. C ia ś, Archives of Metallurgy and Materials **53**, 3, 809-814 (2008).
- [6] R.M. German, Powder metallurgy science, 2nd edn, Princeton, NJ, MPIF 1994, 378.
- [7] A. C ia ś, Development and properties of Fe-Mn-(Mo)-Cr-C sintered structural Steels, ed. AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków 2004.
- [8] A. C ia ś, A.S. W ro ń ski, Powder Metallurgy **53**, 4, 328-335 (2010).
- [9] G.E. Dieter, Mechanical metallurgy, 3edn, NY, Mc Graw-Hill, 1986, 370.
- [10] R.B. Abernethy, The New Weibull Handbook, 4th edition, 536 Oyster Road, North Palm Beach, Florida, November 2000.
- [11] J.W. Newkirk, S. Thakur, Powder Metallurgy **45**, 4, 354-358 (2002).
- [12] N. Biery, M. De Graef, J. Beuth, R. Raban, A. Elliott, T.M. Pollock, C. Austin, Metallurgical and Materials Transactions A **33**, 10, 3127-3136 (2002).
- [13] J.I. McCool's, Journal of Quality Technology **30**, 2, 119-126 (1998).
- [14] A. Khalili, K. Kromp, J. Mater. Sci. **26**, 6741-52 (1991).
- [15] S.H. Zanakis, J. Stat. Comput. Simul. **9**, 101-116 (1979).
- [16] B. Bergman, J. Mat. Sci, Lett. **3**, 689-92 (1984).
- [17] D. Sullivan, P.H. Lauzon, J. Mater. Sci. Lett. **5**, 1245-47 (1986).
- [18] Wu Dingfang, Lu Guanzhong, J. Am. Ceram. Soc. **87**(9), 1799-1802 (2004).
- [19] J.F. Lawless, Statistical models and methods for lifetime data. John Wiley and sons, NY, 1982.
- [20] M. Alqam, R.M. Bennett, A-H. Zureick, Composite Structures **58**, 497-503 (2002).
- [21] S.W. Rust, F.R. Todt, B. Harris, D. Neal, M. Vamgel, Statistical methods for calculating material allowables for MIL-HDBK-17 in: Test methods for design allowables for fibrous composites ASTM STP 1003, 136-149 (1989).
- [22] A. C ia ś, S.C. Mitchell, A. Watts, A.S. W ro ń ski, Powder Metallurgy **42**, 3, 227-233 (1999).