Designation: E 2283 – 03

Standard Practice for Extreme Value Analysis of Nonmetallic Inclusions in Steel and Other Microstructural Features[1]

This standard is issued under the fixed designation E 2283; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This practice describes a methodology to statistically characterize the distribution of the largest indigenous nonmetallic inclusions in steel specimens based upon quantitative metallographic measurements. The practice is not suitable for assessing exogenous inclusions.

1.2 Based upon the statistical analysis, the nonmetallic content of different lots of steels can be compared.

1.3 This practice deals only with the recommended test methods and nothing in it should be construed as defining or establishing limits of acceptability.

1.4 The measured values are stated in SI units. For measurements obtained from light microscopy, linear feature parameters shall be reported as micrometers, and feature areas shall be reported as micrometers.

1.5 The methodology can be extended to other materials and to other microstructural features.

1.6 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:* [2]

E 3 Methods of Preparation of Metallographic Specimens

E 7 Terminology Relating to Metallography

E 45 Test Methods for Determining the Inclusion Content of Steel

E 178 Practice for Dealing with Outlying Observations

E 456 Terminology Relating to Quality and Statistics

E 768 Practice for Preparing and Evaluating Specimens for Automatic Inclusion Assessment of Steel

E 883 Guide for Reflected-Light Photomicrography

E 1122 Practice for Obtaining JK Inclusion Ratings Using Automatic Image Analysis

E 1245 Practice for Determining the Inclusion Content or Second-Phase Constituent of Metals by Automatic Image Analysis

3. Terminology

3.1 *Definitions*—For definitions of metallographic terms used in this practice, refer to Terminology, E 7; for statistical terms, refer to Terminology E 456.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 A_f—the area of each field of view used by the Image Analysis system in performing the measurements.

3.2.2 A_o—control area; total area observed on one specimen per polishing plane for the analysis. *A_o is assumed to be 150 mm^2 unless otherwise noted.*

3.2.3 N_s—number of specimens used for the evaluation. N_s is generally six.

3.2.4 N_p—number of planes of polish used for the evaluation, generally four.

3.2.5 N_f—number of fields observed per specimen plane of polish.

$$N_f = \frac{A_o}{A_f} \quad (1)$$

3.2.6 N—total number of inclusion lengths used for the analysis, generally 24.

$$N = N_s \cdot N_p \quad (2)$$

3.2.7 *extreme value distribution*—The statistical distribution that is created based upon only measuring the largest feature in a given control area or volume **(1,2)**.[3] The continuous random variable x has a two parameter (Gumbel) Extreme Value Distribution if the probability density function is given by the following equation:

$$f(x) = \frac{1}{\delta}\left[\exp\left(-\frac{x-\lambda}{\delta}\right)\right] \times \exp\left[-\exp\left(-\frac{x-\lambda}{\delta}\right)\right] \quad (3)$$

and the cumulative distribution is given by the following equation:

---

[1] This practice is under the jurisdiction of ASTM Committee E04 on Metallography and is the direct responsibility of Subcommittee E04.09 on Inclusions.

Current edition approved Nov. 1. 2003. Published December 2003.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] The boldface numbers in parentheses refer to the list of references at the end of this standard.

$$F(x) = \exp(-\exp(-(x - \lambda) / \delta)) \quad (4)$$

As applied to this practice, x, represents the maximum feret diameter, Length, of the largest inclusion in each control area, A_o, letting:

$$y = \frac{x - \lambda}{\delta} \quad (5)$$

it follows that:

$$F(y) = \exp(-\exp(-y)) \quad (6)$$

and

$$x = \delta\, y + \lambda \quad (7)$$

3.2.8 λ—the location parameter of the extreme value distribution function.

3.2.9 δ—the scale parameter of the extreme value distribution function.

3.2.10 *reduced variate*—The variable y is called the reduced variate. As indicated in Eq 6, y is related to the probability density function. That is $y = F\ (P)$, then from Eq 6, it follows that:

$$y = -\ln(-\ln(F(y))) = -\ln(-\ln(P)) \quad (8)$$

3.2.11 *plotting position*—Each of the N measured inclusion lengths can be represented as x_i, where $1 \leq i \leq N$. The data points are arranged in increasing order such that:

$$x_1 \leq x_2 \leq x_3 \leq x_4 \leq x_5 \ldots \leq x_N$$

Then the cumulative probability plotting position for data point x_i is given by the relationship:

$$P_i = \frac{i}{N + 1} \quad (9)$$

The fraction (i / (N + 1)) is the cumulative probability. $F\ (y_i)$ in Eq 8 corresponds to data point x_i.

3.2.12 *mean longest inclusion length*—$\bar{L}$ is the arithmetic average of the set of N maximum feret diameters of the measured longest inclusions.

$$\bar{L} = \frac{1}{N} \sum_{i=1}^{i=N} L_i \quad (10)$$

3.2.13 *standard deviation of longest inclusion lengths*—Sdev is the standard deviation of the set of N maximum feret diameters of the measured longest inclusions.

$$\text{Sdev} = [\sum_{i=1}^{N} (L_i - \bar{L})^2 / (N - 1)]^{0.5} \quad (11)$$

3.2.14 *return period*—the number of areas that must be observed in order to find an inclusion equal to or larger than a specified maximum inclusion length. Statistically, the return period is defined as:

$$T = \frac{1}{1 - P} \quad (12)$$

3.2.15 *reference area, A_{ref}*—the arbitrarily selected area of 150 000 mm^2. A_{ref} in conjunction with the parameters of the extreme value distribution is used to calculate the size of the largest inclusion reported by this standard. As applied to this analysis, the largest inclusion in each control area A_o is measured. The Return Period, T, is used to predict how large an inclusion could be expected to be found if an area A_{ref} larger than A_o were to be evaluated. For this standard, A_{ref} is 1000 times larger than A_o. Thus, T is equal to 1000. By use of Eq 12 it would be found that this corresponds to a probability value of 0.999, (99.9 %). Similarly by using Eq 6 and 7, the length of an inclusion corresponding to the 99.99 % probability value could be calculated. Mathematically, another expression for the return period is:

$$T = \frac{A_{ref}}{A_o} \quad (13)$$

3.2.16 *predicted maximum inclusion length, L_{max}*—the longest inclusion expected to be found in area A_{ref} based upon the extreme value distribution analysis.

4. Summary of Practice

4.1 This practice enables the experimenter to estimate the extreme value distribution of inclusions in steels.

4.2 Generally, the largest oxide inclusions within the specimens are measured. However, the practice can be used to measure other microstructural features such as graphite nodules in ductile iron, or carbides in tool steels and bearing steels. The practice is based upon using the specimens described in Test Method E 45. Six specimens will be required for the analysis. For inclusion analysis, an area of 150 mm^2 should be evaluated for each specimen.

4.3 After obtaining the specimens, it is recommended that they be prepared by following the procedures described in Methods E 3 and Practice E 768.

4.4 The polished specimens are then evaluated by using the guidelines for completing image analysis described in Practices E 1122 and E 1245. For this analysis, feature specific measurements are required. The measured inclusion lengths shall be based on a minimum of eight feret diameter measurements.

4.5 For each specimen, the maximum feret diameter of each inclusion is measured. After performing the analysis for each specimen, the largest maximum feret diameter of the measured inclusions is recorded. This will result in six lengths. The procedure is repeated three more times. This will result in a total of 24 inclusion lengths.

4.6 The 24 measurements are used to estimate the values of δ and λ for the extreme value distribution for the particular material being evaluated. The largest inclusion L_{max} expected to be in the reference area A_{ref} is calculated, and a graphical representation of the data and test report are then prepared.

4.7 The reference area used for this standard is 150 000 mm^2. Based upon specific producer, purchaser requirements, other reference areas may be used in conjunction with this standard.

4.8 When required, the procedure can be repeated to evaluate more than one type of inclusion population in a given set of specimens. For example, oxides and sulfides or titanium-carbonitrides could be evaluated from the same set of specimens.

5. Significance and Use

5.1 This practice is used to assess the indigenous inclusions or second-phase constituents in metals using extreme value statistics.

5.2 It is well known that failures of mechanical components, such as gears and bearings, are often caused by the presence of large nonmetallic oxide inclusions. Failure of a component can often be traced to the presence of a large inclusion. Predictions related to component fatigue life are not possible with the evaluations provided by standards such as Test Methods E 45, Practice E 1122, or Practice E 1245. The use of extreme value statistics has been related to component life and inclusion size distributions by several different investigators **(3-8)**. The purpose of this practice is to create a standardized method of performing this analysis.

5.3 This practice is not suitable for assessing the exogenous inclusions in steels and other metals because of the unpredictable nature of the distribution of exogenous inclusions. Other methods involving complete inspection such as ultrasonics must be used to locate their presence.

6. Procedure

6.1 Test specimens are obtained and prepared in accordance with E 3, E 45 and E 768.

6.2 The microstructural analysis is to be performed using the types of equipment and image analysis procedures described in E 1122 and E 1245.

6.3 Determine the appropriate magnification to use for the analysis. For accurate measurements, the largest inclusion measured should be a minimum of 20 pixels in length. For specimens containing relatively large inclusions, objective lens having magnifications ranging from 10 to 20× will be adequate. Generally, for specimens with small inclusions, an objective lens of 32 to 80× will be required. The same magnification shall be used for all the specimens to be analyzed.

6.4 Using the appropriate calibration factors, calculate the area of the field of view observed by the image analysis system, A_f. For each specimen, an area of 150 mm^2 shall be evaluated. Using Eq 1, the number of fields of view required to perform the analysis is $N_f = A_o / A_f = 150 / A_f$. N_f should be rounded up to the next highest integer value; that is, if N_f is calculated to be 632.31, then 633 fields of view shall be examined.

6.5 *Image Analysis Measurements*:

6.5.1 In this practice, feature specific parameters are measured for each individual inclusion. The measured inclusion lengths shall be based on a minimum of eight feret diameters.

6.5.2 For each field of view, focus the image either manually or automatically, and measure the maximum feret diameter of each detected oxide inclusion. The measured feret diameters are stored in the computer's memory for further analysis. This procedure is repeated until an area of 150 mm^2 is analyzed.

6.5.3 In situations where only a very few inclusions are contained within the inspected area, the specimen can first be observed at low magnification, and the location of the inclusions noted. The observed inclusions can then be remeasured at high magnification.

6.5.4 After the specimen is analyzed, using the accumulated data, the maximum feret diameter of the largest measured inclusion in the 150 mm^2 area is recorded. This procedure is repeated for each of the other five specimens.

6.5.5 The specimens are then repolished and the procedure is repeated until each specimen has been evaluated four times. This will result in a set of 24 maximum feret diameters. For each repolishing step, it is recommended that at least 0.3 mm of material be removed in order to create a new plane of observation.

6.5.6 The mean length, $\bar{L}$, is then calculated using Eq 10.

6.5.7 The standard deviation, Sdev, is calculated using Eq 11.

6.6 The 24 measured inclusion lengths are sorted in ascending order. An example of the calculations is contained in Appendix X1. The inclusions are then given a ranking. The smallest inclusion is ranked number 1, the second smallest is ranked number 2 etc.

6.7 The probability plotting position for each inclusion is based upon the rank. The probabilities are determined using Eq 9: $P_i = i / (N + 1)$. Where $1 \leq i \leq 24$, and $N = 24$.

6.8 A graph is created to represent the data. Plotting positions for the ordinate are calculated from Eq 8: $y_i = -\ln(-\ln(P_i))$. The variable y in this analysis is referred to as the Reduced Variate (Red. Var.). Typically the ordinate scale ranges from −2 through +7. This corresponds to a probability range of inclusion lengths from 0.87 through 99.9 %. The ordinate axis is labeled as Red. Var. It is also possible to include the Probability values on the ordinate. In this case, the ordinate can be labeled Probability (%). The abscissa is labeled as Inclusion Length (mm); the units of inclusion length shall be micrometers.

6.9 *Estimation of the Extreme Value Distribution Parameters*:

6.9.1 Several methods can be used to estimate the parameters of the extreme value distribution. Using linear regression to fit a straight line to the plot of the Reduced Variate as a function of inclusion length is the easiest method; however, it is the least precise. This is because the larger values of the inclusion lengths are more heavily weighted than the smaller inclusion lengths. Two other methods for estimating the parameters are the method of moments (mom), and the method of maximum likelihood (ML). The method of moments is very easy to calculate, but the method of maximum likelihood gives estimates that are more precise. While both methods will be described, the maximum likelihood method shall be used to calculate the reported values of δ and λ for this standard. (Since the ML solution is obtained by numerical analysis, the values of δ and λ obtained by the method of moments are good guesses for starting the ML analysis.)

6.9.2 *Moments Method*—It has been shown that the parameters for the Gumbel distribution, can be represented by:

$$\delta_{mom} = \frac{\text{Sdev}\sqrt{6}}{\pi} \tag{14}$$

and

$$\lambda_{mom} = \bar{L} - 0.5772 \cdot \delta_{mom} \tag{15}$$

where the subscript mom indicates the estimates are based on the moment method.

6.9.3 *Maximum Likelihood Method*—This method is based on the approach that the best values for the parameters δ and λ are those estimates that maximize the likelihood of obtaining the measured set of inclusion lengths. Since the extreme value distribution is based on a double exponential function, the maximization process is easiest to perform on the log of the distribution function. That is for the given set if measurements:

$$LL = \sum_{i=1}^{n} \ln(f(x_i, \lambda, \delta)) \quad (16)$$

$$= \sum_{i=1}^{n} \ln\left(\frac{1}{\delta}\right) - \left(\frac{x_i - \lambda}{\delta}\right) - \exp\left(-\frac{x_i - \lambda}{\delta}\right) \quad (17)$$

The maximization of LL is best performed by numerical analysis. This can be done via a spreadsheet or an appropriate computer analysis program. The values of δ and λ that are determined from Eq 17 are referred to as δ_{ML} and λ_{ML}. An example of the maximization process is described in Appendix X1. Having determined the best estimates for δ_{ML} and λ_{ML}, it follows that:

$$x = \delta_{ML}(\text{Red. Var.}) + \lambda_{ML} \quad (18)$$

or

$$x = \delta_{ML} \ln(-\ln(P)) + \lambda_{ML} \quad (19)$$

In terms of the return period:

$$x = -\delta_{ML} \ln\left(-\ln\left(\frac{T-1}{T}\right)\right) + \lambda_{ML} \quad (20)$$

6.9.4 *Outlying Observations*—Practice E 178 shall be used to deal with outlying observations. As applied to this standard, an upper significance of 1 % shall be the governing criterion. The recommended criteria for single sample rejections is described in Section 4 of Practice E 178. If a data point is concluded to be an outlier, then in accordance with Practice E 178, section 2.3, it shall be rejected. The specimen containing the outlier shall then be repolished, and the analysis repeated. Examples of outlier calculations are described in Appendix X1.

6.9.5 The standard error, SE, for any inclusion of length x based upon the ML method is:

$$SE(x) = \delta_{ML} \cdot \sqrt{(1.109 + 0.514 \cdot y + 0.608 \cdot y^2) / n} \quad (21)$$

6.9.6 *95 % Confidence Intervals*—In practice, very large return periods are used in predicting how large an inclusion will be present is a particular area of steel. Thus the results of the extreme value analysis shall be presented with confidence limits. The approximate 95 % confidence intervals are:

$$95\,\%\ CI = \pm 2 \cdot SE(x) \quad (22)$$

6.10 *Predicted Longest Inclusion, L_{max}*—The return period is used to predict how large an inclusion would be expected to be found if an area much greater than A_o were to be examined. As previously defined, 3.2.15, this area is referred to as A_{ref} = 150 000 mm^2. Thus using the calculated values of δ_{ML} and λ_{ML} from the maximum likelihood method, Eq 17, and P = 0.999, L_{max} is calculated.

6.11 *Comparison of Different Lots of Steel*—Using the methodology described herein, the following procedure can be used to compare the differences in sizes of large nonmetallic inclusions in two steels designated A and B.

6.11.1 For steel A, δ_A, λ_A, are calculated from Eq 17. The SE for steel A is calculated based upon the value of L_{max} for steel A by using Eq 21. The same parameters are calculated for steel B.

6.11.2 The approximate 95 % confidence interval for L_{max} (A) − L_{max} (B) is:

$$CI = L_{max}(A) - L_{max}(B) \pm 2 \cdot \sqrt{SE_{ref}(A)^2 + SE_{ref}(B)^2} \quad (23)$$

6.11.3 If the lower to upper bounds of the 95 % CI include 0, then conclude that there is no difference in the characteristic sizes of the largest inclusions in heat A and B.

6.11.4 If the value 0 is less than the bounds of the confidence interval, then conclude that characteristic size of the largest inclusion in heat A is greater than that in heat B.

6.11.5 If the value 0 is greater than the bounds of the confidence interval, then conclude that characteristic size of the largest inclusion in heat B is greater than that in heat A.

7. Report

7.1 The report shall consist of a graphical representation of the data, information discussing how the data was measured and the results of the statistical analysis.

7.2 The graphical analysis shall contain the data points used for the analysis, the best-fit line as determined by the maximum likelihood method, and the 95 % confidence intervals for the data. The ordinate of the graph may be the Reduced Variate or the probability values. The abscissa will be Inclusion Length in micrometers. The control area, A_0 shall be included on the graph.

7.3 For this practice, the accompanying report shall contain the following:

7.3.1 Name of the person performing the analysis.
7.3.2 Date the analysis was completed.
7.3.3 Material Type.
7.3.4 Specimen location and size of material.
7.3.5 Microscope objective magnification.
7.3.6 Image Analysis Calibration Constant.
7.3.7 A_f [μm^2].
7.3.8 A_o [μm^2].
7.3.9 N_f.
7.3.10 $\bar{L}$.
7.3.11 Sdev.
7.3.12 δ_{ML} (to 3 decimal places).
7.3.13 λ_{ML} (to 3 decimal places).
7.3.14 L_{max}.

7.4 The length of any outlier measurements that were rejected shall be reported.

7.5 When possible, the report should contain the steel Oxygen, Silicon, Aluminum and Calcium contents.

7.6 Any other information deemed necessary shall be based upon purchaser-producer agreements.

8. Keywords

8.1 extreme value statistics; inclusion length; maximum inclusion length; maximum likelihood method

APPENDIX

(Nonmandatory Information)

X1. EXAMPLE CALCULATION

X1.1 The data contained in Table X1.1 represents the largest maximum feret diameters, inclusion lengths, measured in a group of specimens. The specimens are numbered one through six, and the four planes of polish are A through D respectively. The mean length, $\bar{L}$, of 51.75 µm is the arithmetic mean of the 24 measurements, Eq 10. The Sdev of these lengths is 18.86 µm, Eq 11.

X1.2 After obtaining the 24 measurements, the data from Table X1.1 is pasted into a spreadsheet. The inclusion data is then sorted in ascending order; that is, the smallest inclusion length is first, etc. The sorted data is the first column (A) in Table X1.2.

X1.3 The ranking for each inclusion is then assigned. The smallest inclusion is number 1, the next smallest is number 2 etc., Table X1.2, column B.

X1.4 The probability plotting position for each inclusion is next calculated using Eq 9, Table X1.2, column C. For example consider the inclusion having a length of 40.29 µm. The rank of this inclusion is 9. The probability position for the inclusion is:

$$P_i = \frac{i}{N+1} = \frac{9}{24+1} = 0.36 \qquad \text{(X1.1)}$$

X1.5 Using the probability plotting positions, the Reduced Variate for each position is calculated using Eq 8, Table X1.2, column D. For example the probability value for inclusion 9, having a length of 40.29 µm is 0.36; hence, from Eq 8 it follows that:

$$y = -\ln(-\ln(P_9)) = -\ln(-\ln(0.36)) = -\ln(1.022) = -0.021 \qquad \text{(X1.2)}$$

X1.6 Using the Inclusion Length data in column A, the Mean inclusion length and the standard deviation if the inclusion lengths are calculated, Eq 10 and 11 respectively. These values appear in column B above the inclusion data.

X1.7 The mean inclusion length and the standard deviation are used to calculate δ_{mom} and λ_{mom} using Eq 14 and 15 respectively. The results of these calculations are: δ_{mom} = 14.71 and λ_{mom} = 43.26. These results are listed above the inclusion measurements in Table X1.2, column E.

TABLE X1.1 Largest Inclusion Lengths Measured from 24 Polishing Planes from Steel Z

Specimen	A	B	C	D
1	40.29	30.73	73.48	78.91
2	37.24	37.43	44.79	46.53
3	29.03	35.00	70.87	94.28
4	52.46	44.82	59.83	49.15
5	62.21	66.13	22.18	82.39
6	33.98	48.55	64.32	37.43
Mean Length = 51.75 (µm)			Sdev = 18.86	

X1.8 *Maximum Likelihood Method for* δ *and* λ:

X1.8.1 In order to evaluate δ and λ by the maximum likelihood method, the natural logarithm of the probability density of the extreme value function, Eq 3, must first be determined. This function must then be evaluated for each data point. The function is the terms following the summation symbol in Eq 17. For simplicity it will be identified as $\ln(f(x_i, \delta, \lambda))$. The values of δ and λ that maximize the sum of these values is the maximum likelihood solution. The solution is determined as follows:

X1.8.2 As a first guess, assume the values of δ_{mom} and λ_{mom} are the solution. These values are copied into column H just above the inclusion data.

X1.8.3 The value of $\ln(f(x_i, \delta, \lambda))$ is evaluated for each measured inclusion length. For the first calculation, the values of δ_{mom} and λ_{mom} in column H are used.

X1.8.4 The summation of each value of $\ln(f(x_i, \delta, \lambda))$ is denoted SUM (LL). In Table X1.2, it is at the bottom of column F.

X1.8.5 The maximization of the sum of the terms in column F is determined by numerical analysis. For this example, using an EXCEL spreadsheet, the SOLVER function is used for this process. SOLVER is used by maximizing the SUM(LL) by determining the proper values of δ and λ. For this example, the solution set is δ_{ML} = 14.981 and λ_{ML} = 43.056.

NOTE X1.1—Other types of spreadsheets or analytic software programs can be used to perform the calculations.

X1.8.6 The maximum likelihood analysis results for δ and λ are used to represent the best-fit line for the data, Eq 18:

$$x = \delta_{ML} \cdot \text{Red.Var.} + \lambda_{ML} \qquad \text{(X1.3)}$$

The points on the best-fit line are calculated using Eq 18, the ML values of δ and λ and the Red. Var. for each data point, Table X1.2, Column H.

X1.8.7 Similarly using Eq 21 and 22, the 95 % confidence interval points are determined for each data point, Columns I and J respectively.

X1.8.8 L_{max} is calculated for a return period of 1000 ($A_{ref.}$= 150 000 mm^2) using Eq 20 and δ_{ML} and λ_{ML}. That is:

$$L = -\delta_{ML} \ln\left(-\ln\left(\frac{T-1}{T}\right)\right) + \lambda_{ML} \qquad \text{(X1.4)}$$

$$L_{max} = -14.981 \ln\left(-\ln\left(\frac{1000-1}{1000}\right)\right) + 43.056$$
$$= 146.53$$

X1.8.9 95 % Confidence Interval for L_{max}. The standard error for L_{max} is based on a probability P = 99.9 %. Thus:

$$y = -\ln(-\ln(P)) = -\ln(-\ln(0.999)) = 6.61 \qquad \text{(X1.5)}$$

$$SE_{(x)} = \delta_{ML} \cdot \sqrt{(1.109 + 0.514 \cdot y + 0.608 \cdot y^2) / n}$$
$$= 14.981 \cdot \sqrt{(1.109 + 0.514 \cdot (6.91) + 0.608 \cdot (6.91)^2) / 24}$$

TABLE X1.2 Ranking, Probability Positions and Calculated Statistical Parameters for the Measured Inclusions

A	B	C	D	E	F	G	H	I	J
Mean	51.751		δ_{mom}	14.71		δ_{ML}	14.981		
Sdev	18.864		λ_{mom}	43.26		λ_{ML}	43.056		

Length (Y) Data	Rank	Prob.	Red. Var. (X) RV	ln (f(x_i, δ, λ))	X	X_low	X_high
22.18	1	0.04	-1.169	-5.342	25.54	18.5	32.6
29.03	2	0.08	-0.927	-4.321	29.18	22.6	35.7
30.73	3	0.12	-0.752	-4.161	31.80	25.5	38.1
33.98	4	0.16	-0.606	-3.934	33.98	27.8	40.2
35.00	5	0.20	-0.476	-3.881	35.93	29.8	42.0
37.24	6	0.24	-0.356	-3.793	37.73	31.6	43.9
37.43	7	0.28	-0.241	-3.787	39.44	33.3	45.6
37.43	8	0.32	-0.131	-3.787	41.10	34.8	47.4
40.29	9	0.36	-0.021	-3.725	42.74	36.3	49.1
44.79	10	0.40	0.087	-3.713	44.37	37.8	50.9
44.82	11	0.44	0.197	-3.713	46.01	39.2	52.8
46.53	12	0.48	0.309	-3.732	47.69	40.6	54.7
48.55	13	0.52	0.425	-3.767	49.42	42.1	56.8
49.15	14	0.56	0.545	-3.779	51.22	43.6	58.9
52.46	15	0.60	0.672	-3.868	53.12	45.1	61.2
59.83	16	0.64	0.807	-4.153	55.14	46.7	63.6
62.21	17	0.68	0.953	-4.264	57.33	48.4	66.3
64.32	18	0.72	1.113	-4.368	59.73	50.2	69.3
66.13	19	0.76	1.293	-4.461	62.43	52.2	72.6
70.87	20	0.80	1.500	-4.720	65.53	54.5	76.5
73.48	21	0.84	1.747	-4.869	69.22	57.2	81.2
78.91	22	0.88	2.057	-5.191	73.87	60.6	87.2
82.39	23	0.92	2.484	-5.405	80.27	65.1	95.4
94.28	24	0.96	3.199	-6.159	90.97	72.7	109.3
			SUM (LL) = −102.893				

$$SE(x) = 17.74$$

From Eq 22:

$$95\ \%\ CI = \pm 2 \cdot SE(x) = \pm 2 \cdot 17.74 = \pm 35.48 \qquad \text{(X1.6)}$$

X1.9 *Outlying Observations*:

X1.9.1 The largest inclusion. For the reported data set, the largest measured inclusion is 94.28 µm, Table X1.2, column A. Assume that this inclusion is replaced by one having a length of 125 µm. Using the new inclusion length, it is found that the new mean is $\bar{L}$ = 53.03 µm and the new standard deviation is σ = 22.56. As cited in Practice E 178, Section 4:

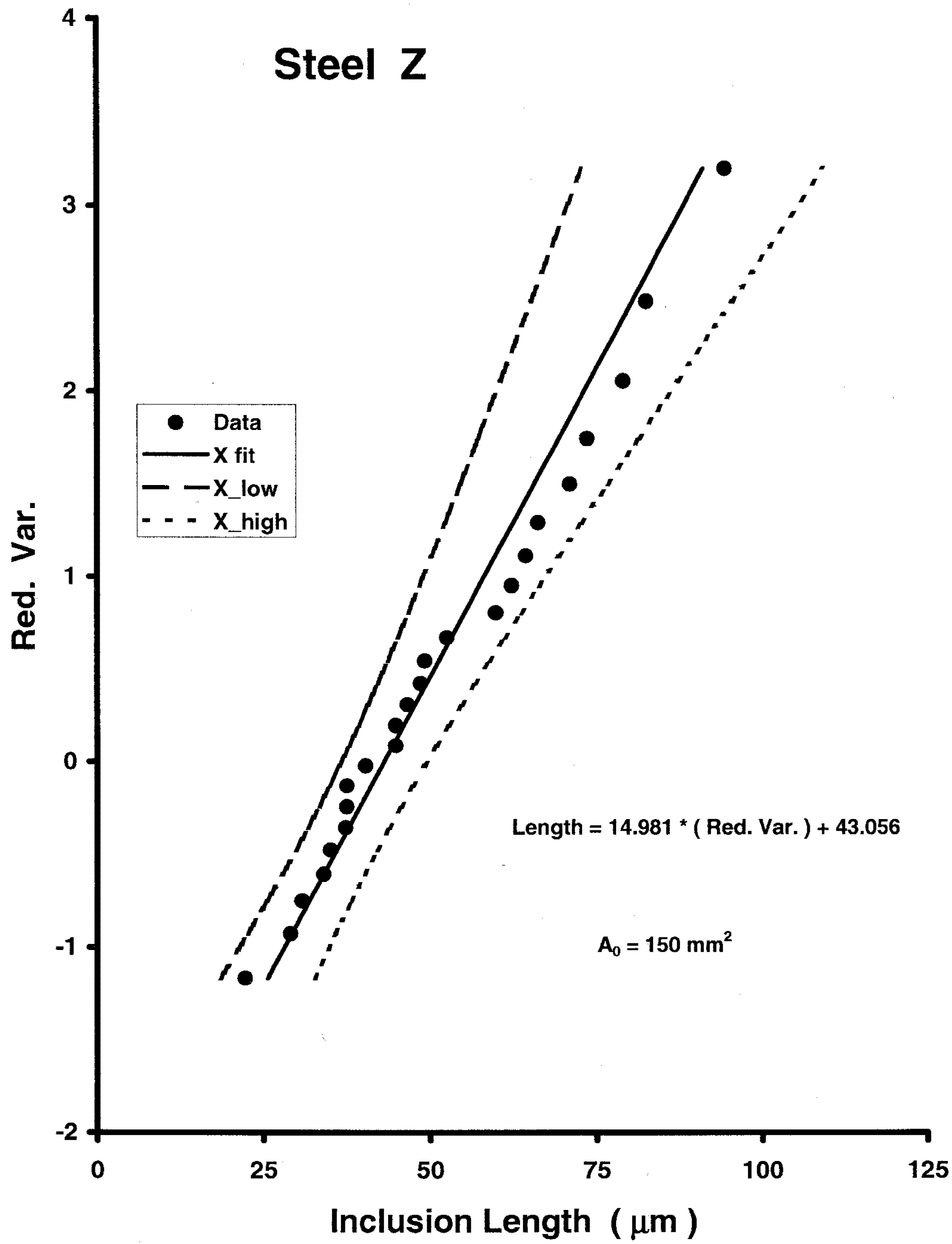
$$T_{24} = (L_{24} - \bar{L}) / \sigma = (125 - 53.03) / 22.56 = 3.19 \qquad \text{(X1.7)}$$

For the Upper 1 % confidence interval, T_{24} must be 2.987 or less, Practice E 178, Table 1. Since T_{24} for the 125 µm inclusion is 3.19, this fails the test. Hence the 125 µm inclusion is an outlier. The specimen containing this inclusion should be repolished and reevaluated for the longest inclusion.

X1.9.2 Consider replacing the smallest inclusion having a length of 22.18 µm by an inclusion having a length of 0.0. That is no inclusion was measured on one of the specimens. For this case, the new mean inclusion length $\bar{L}$ = 55.83, and the new standard deviation is σ = 20.82. Thus:

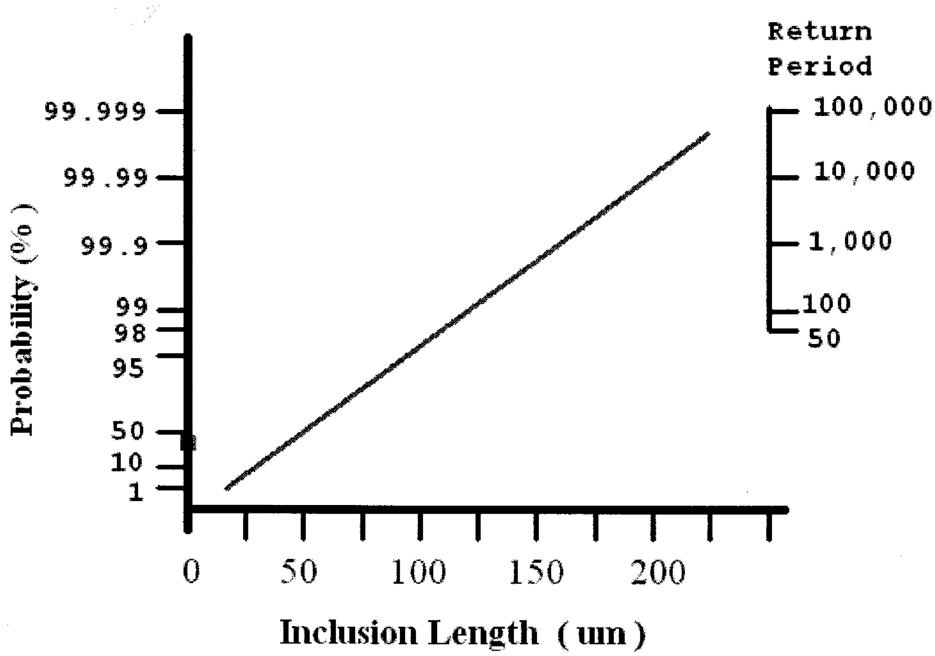
$$T_1 = (\bar{L} - L_1) / \sigma = (50.83 - 0) / 20.82 = 2.44 \qquad \text{(X1.8)}$$

Since 2.44 is less than the upper 1 % significance level of 2.987, the value of 0.0 is not an outlier.



NOTE—The ordinate is the Reduced Variate, Eq 18.

FIG. X1.1 Graphical Representation of the Extreme Value Data Analysis



NOTE—The ordinate is a probability scale based upon Eq 19.

FIG. X1.2 Graphical Representation of the Extreme Value Distribution of Steel Z

REFERENCES

(**1**) Gumbel, E. J., *Statistics of Extremes*, Columbia Univ. Press, NYC, 1958.

(**2**) Kinnison, R. R., *Applied Extreme Value Statistics*, Battelle Press, Columbus, 1985.

(**3**) Murakami, Y., et al., "Quantitative Evaluation of Effects of Non-Metallic Inclusions on Fatigue Strength of High Strength Steels. I: Basic Fatigue Mechanism and Evaluation of Correlation between the Fatigue Fracture Stress and the Size and Location of Non-Metallic Inclusions," *Int. J. Fatigue*, 11, No. 5, 1989, pp. 291-298.

(**4**) Murakami, Y., et al., "Quantitative Evaluation of Effects of Non-Metallic Inclusions on Fatigue Strength of High Strength Steels. II: Fatigue Limit Evaluation Based on Statistics for Extreme Values of Inclusion Size," *Int. J. Fatigue*, 11, No. 5, 1989, pp. 299-307.

(**5**) Sih, G., et al., "Comparison of Extreme Value Statistics Methods for Predicting Maximum Inclusion Size in Clean Steels," *Ironmaking and Steelmaking*, 26, No. 4, 1999, pp. 239-246.

(**6**) Beretta, S., et al., "Largest-Extreme-Value Distribution Analysis of Multiple Inclusion Types in Determining Steel Cleanliness," *Met Trans B*, 32B, June 2001, pp. 517-523.

(**7**) Beretta, S., et al., "Statistical Analysis of Defects for Fatigue Strength Prediction and Quality Control of Materials," *Fatigue Fract. Eng. Mat. Struct.*, 21, 1998, pp. 1049-1065.

(**8**) ESIS P11-03 (2003), "Technical Recommendations for the Extreme Value Analysis of Data on Large Nonmetallic Inclusions in Steels," GKSS, Geesthacht, Germany, ISSN 1616-2129.