NISTIR 6652

Dynamic Force Measurement:
Instrumented Charpy Impact Testing

C. M. McCowan
J. D. Splett
E. Lucon

National Institute of Standards and Technology
U.S. Department of Commerce

NISTIR 6652

Dynamic Force Measurement: Instrumented Charpy Impact Testing

C. N. McCowan
Materials Reliability Division
Materials Science and Engineering Laboratory

J. D. Splett
Statistical Engineering Division
Information Technology Laboratory
National Institute of Standards and Technology
325 Broadway
Boulder, CO 80305

E. Lucon
Reactor Material Research Department
Institute of Nuclear Material Science
SCK•CEN (Belgian Nuclear Research Centre)
Mol, Belgium

January 2008

U.S. Department of Commerce
Carlos M. Gutierrez, Secretary

National Institute of Standards and Technology
James M. Turner, Acting Director

CONTENTS

Dynamic Force Measurement: Instrumented Charpy Impact Testing

C. N. McCowan[a] and J. D. Splett[b]
National Institute of Standards and Technology
325 Broadway
Boulder, CO 80305

E. Lucon[c]
SCK•CEN (Belgian Nuclear Research Centre)
Mol, Belgium

Maximum forces measured on impact reference materials in round robin testing are reported. These dynamic results are compared with static force calibration of instrumented strikers. One machine in the round robin was identified as an outlier by the dynamic force measurement results, indicating that dynamic testing may be a useful approach to check the performance of the instrumented machine and measurement system. Repeatability (95 %) and reproducibility limits of 0.45 kN and 2.16 kN, respectively, were determined for the high energy specimens.

Key Words: Charpy impact; dynamic force measurement; instrumented impact testing

1. Summary

The maximum forces measured by the machines in this round robin are in good agreement. This general result shows that the static force calibration of instrumented strikers is quite robust, and that the various striker designs evaluated here performed in a predictable manner. One machine in the round robin was identified as an outlier by the dynamic force measurement results, indicating that dynamic testing may be a useful approach to check the performance of the instrumented machine and measurement system.

It is not clear if differences observed between strikers with 2 mm and 8 mm radii are of practical significance. More data are needed to fully evaluate this variable.

The certified value determined for the maximum force of the low energy (LL-103) specimens is 32.97 ±0.75 kN (±2.3 %), with an expanded uncertainty of 1.84 kN (5.6 %). The specimens have a 95 % uncertainty interval of ±6 %.

---

[a] Materials Reliability Division, Materials Science and Engineering Laboratory
[b] Statistical Engineering Division, Information Technology Laboratory
[c] Reactor Material Research Department, Institute of Nuclear Material Science

The certified value determined for the maximum force of the high energy (HH-103) specimens is 24.06 ±0.28 kN (±1.2 %), with an expanded uncertainty of 0.70 kN (2.9 %). The specimens have a 95 % uncertainty interval of ±3 %.

The precision of the maximum force measured for the HH-103 specimens is higher than that measured for the LL-103 measurements. The 95 % repeatability and reproducibility limits for the HH-103 measurements were 0.45 kN and 2.16 kN, respectively (compared with 1.90 kN and 5.96 kN for LL-103).

The mean general yield values for the HH-103 and LL-103 specimens were 19.60 kN and 30.84 kN, respectively. The 95 % repeatability and reproducibility limits for the HH-103 measurements were 2.30 kN and 3.59 kN, respectively (compared with 2.09 kN and 9.60 kN for LL-103).

The low variation associated with the HH-103 specimen makes it suitable for use as a material for dynamic force verification.

2. Background

Instrumented strikers for impact machines are calibrated with statically applied loads, traceable to certified load cells. This static calibration method has been found to work quite well. However, in some cases it may be convenient to check the performance of the machine, striker, and force measurement system using a verification specimen. Currently NIST does not offer a certified reference material that can be used to verify the performance of an instrumented impact machine.

3. Test Plan

3.1. Objective

The round robin was conducted primarily to evaluate the suitability of using Charpy verification materials as dynamic force verification materials, and to establish certified reference values for the maximum force measured for these two steels in an instrumented impact test. These steels are expected to provide dynamic force verification at two levels of maximum force (around 25 and 30 kN). In addition, the precision of the data are evaluated in terms of repeatability and reproducibility according to ASTM E 691-05.

3.2. Design

Specimens from batches of high (HH-103) and low (LL-103) energy level Charpy V-notch verification specimens are tested on instrumented Charpy pendulum impact machines. The specimens have already been certified by NIST for absorbed energy using conventional Charpy pendulum impact machines. The batch sizes for these high and low energy specimens are such that, following this round robin, about 800 to 900 specimens can be certified as force verification specimens and some reserve specimens can be held for future stability testing, etc.

3.3. Reporting

This summary report on the testing will be distributed to the participants and cited on certificates developed for the specimens. The summary report is public information, and participants are encouraged to write journal articles based on the results of the testing.

3.4. Testing Details

- Materials: Two steels (LL-103 low energy and HH-103 high energy)
- Specimens: Standard size ASTM E23 Charpy V-notch specimens (10 mm x 10 mm x 55 mm)
- Anvil Geometry: The ASTM E23 or ISO 148 anvil geometry
- Striker Geometry: Testing on either 2 mm or 8 mm radius striker geometries
- Machine Type: Testing on instrumented pendulum impact machines
- Test Procedure: Testing conducted in accordance with ASTM E23 and ASTM work item WK383 (Method for Instrumented Impact Tests of Metallic Materials) or ISO 148 – 1 and ISO 14556
- Test Matrix: Ten specimens at each energy level tested for each striker geometry used
- Test Temperature: Room temperature
- Test Time: Test the specimens within 3 months of receiving them

Table 1. Materials and machines.

Material	Machine type	Striker geometry (mm)	Number of specimens
Steel 1	Pendulum	8	10
LL-103		2	10
Steel 2	Pendulum	8	10
HH-103		2	10

3.5. Participants

Table 2. Participants.

Contact	Laboratory
Wolfgang Böhme	Fraunhofer Institute for Mechanics of Materials
Harald Diem	MPI University of Stuttgart
Enrico Lucon	SCK·CEN, Institute of Nuclear Materials
Chris McCowan	National Institute of Standards and Technology
Václav Mentl	SKODA Research
Hans-Werner Viehrig	Forschungszentrum Rossendorf

4. Results and Discussion

4.1. Maximum Force

A total of 10 specimens were tested on each of 8 Charpy machines for the force verification round robin. The data were evaluated from averaged force-signals according to ISO 14556 or ASTM WK383. The raw data from the round robin are shown in Appendix A. Two sets of ten verification specimens were tested with machine #8 (identified as 8a and 8b in Appendix A). The measurements for machine #8 were combined since there is no statistical difference between the averages of the two sets and no apparent trend in the data. Machine #2 and machine #8 use an 8 mm striker, the remaining machines all use a 2 mm striker. **Figures 1** and **2** display the maximum force for low-energy and high-energy specimens, respectively.

The data appear to be fairly consistent among machines, with the exception of machine #4, which has test results that are much lower than the rest of the data for both low and high energies. Since machine #4 produces consistently low values, we exclude it from further analyses.

4.1.1. Striker Radius

Figures 3 and **4** display the force data versus striker radius for low and high energies, respectively.

Figures 3 and 4 indicate that striker radius may have some effect on maximum force. We performed a *t*-test to determine if the mean of the data from machines having an 8 mm striker is statistically different from the mean of data from machines with 2 mm strikers for each energy level. Based on the *t*-test (using the Satterthwaite method for unequal variances [1]), the 8 mm striker mean is statistically different from the 2 mm striker mean for both low and high energies ($p < 0.0001$). **Figure 5** displays the means of each group along with error bars (twice the standard error of the mean, or $2s/\sqrt{n}$).

The error bars in **Figure 5** do not overlap for either low or high energy, also indicating that there is a significant difference between 2 mm and 8 mm striker means. In addition, the figure shows that the 8 mm striker means are higher than the 2 mm striker means.

A *t*-test using combined machine data may not be the best way to examine striker differences because of the large sample sizes in each group. The *t*-test will always find a difference between groups if the sample sizes are large enough since the uncertainty of the mean decreases as the sample size increases. An alternative to the *t*-test is to determine if the two 8 mm striker machines have large influence on the consensus value [2]. If both machines with 8 mm strikers are highly influential, then there is evidence that there are striker differences. **Table 3** lists the "hat" value, standardized residual, and Cook's D for each machine and energy level.

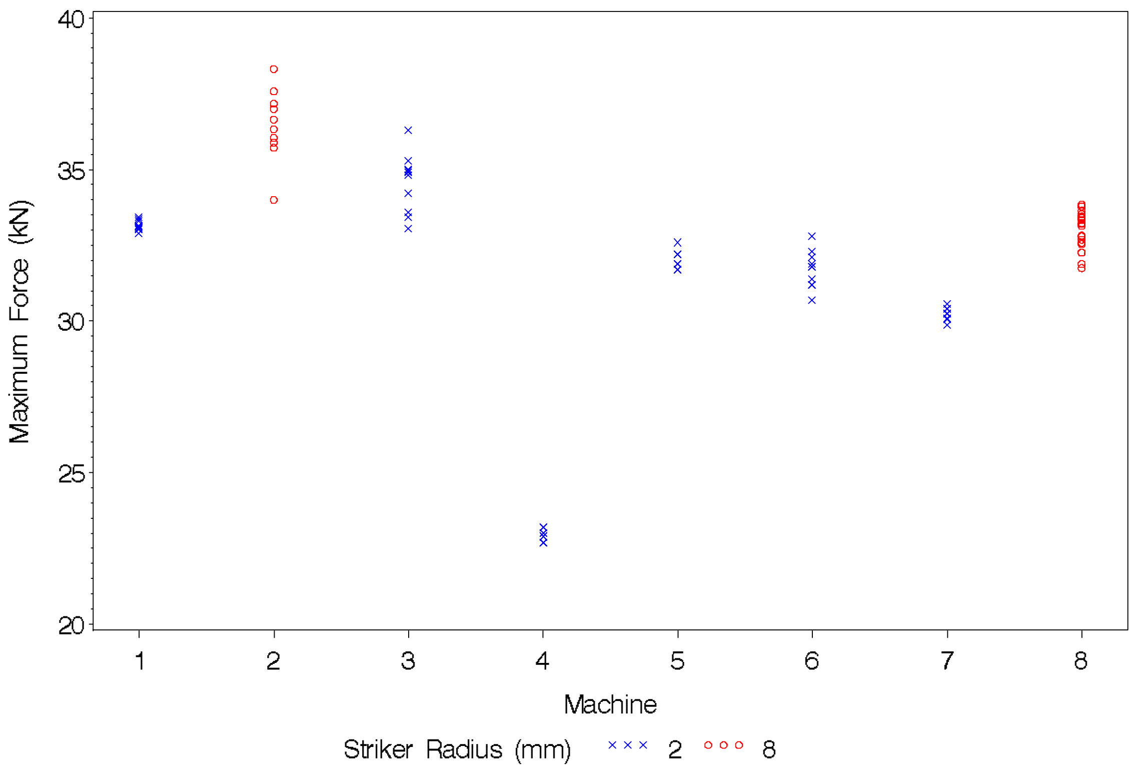

Figure 1. Low-energy maximum force for eight machines.

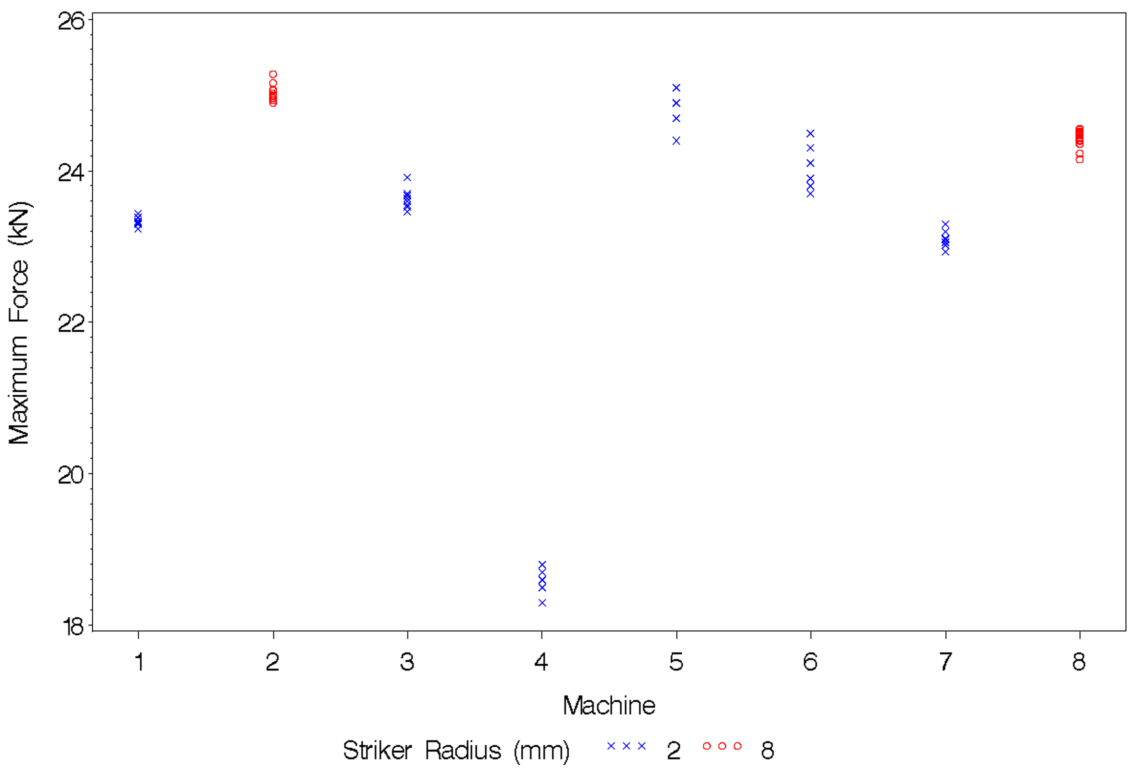

Figure 2. High-energy maximum force for eight machines.

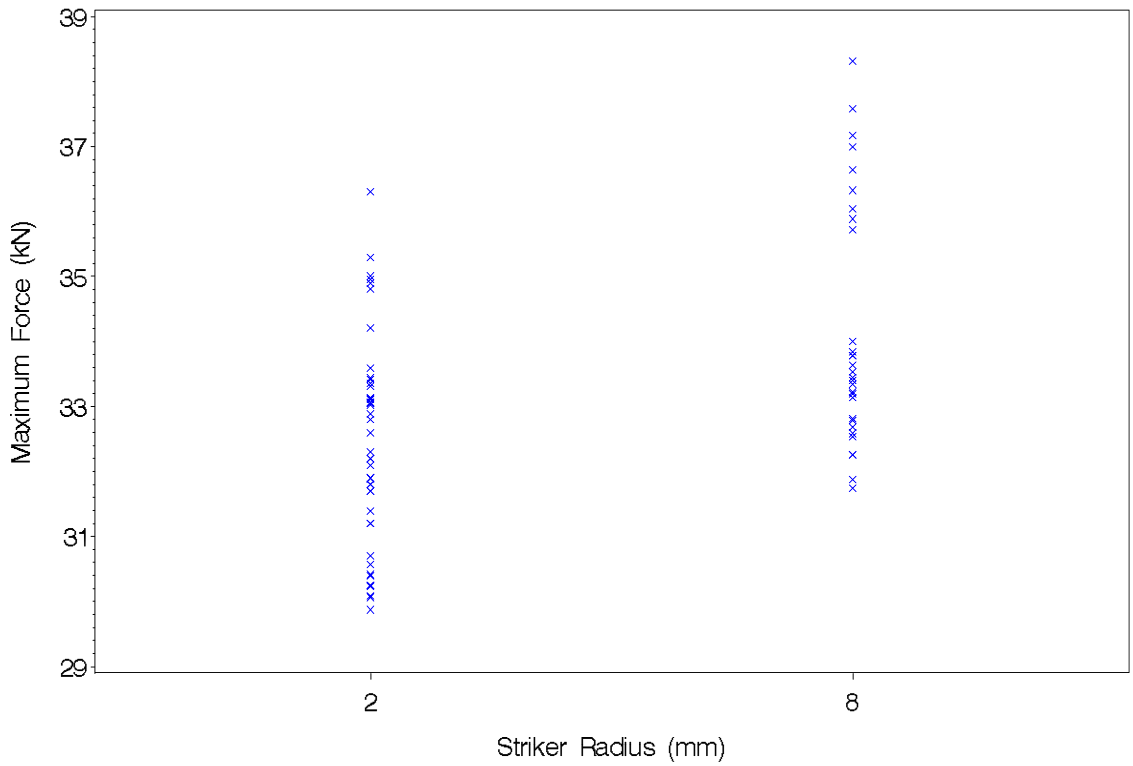

Figure 3. Maximum force versus striker radius for low-energy specimens.

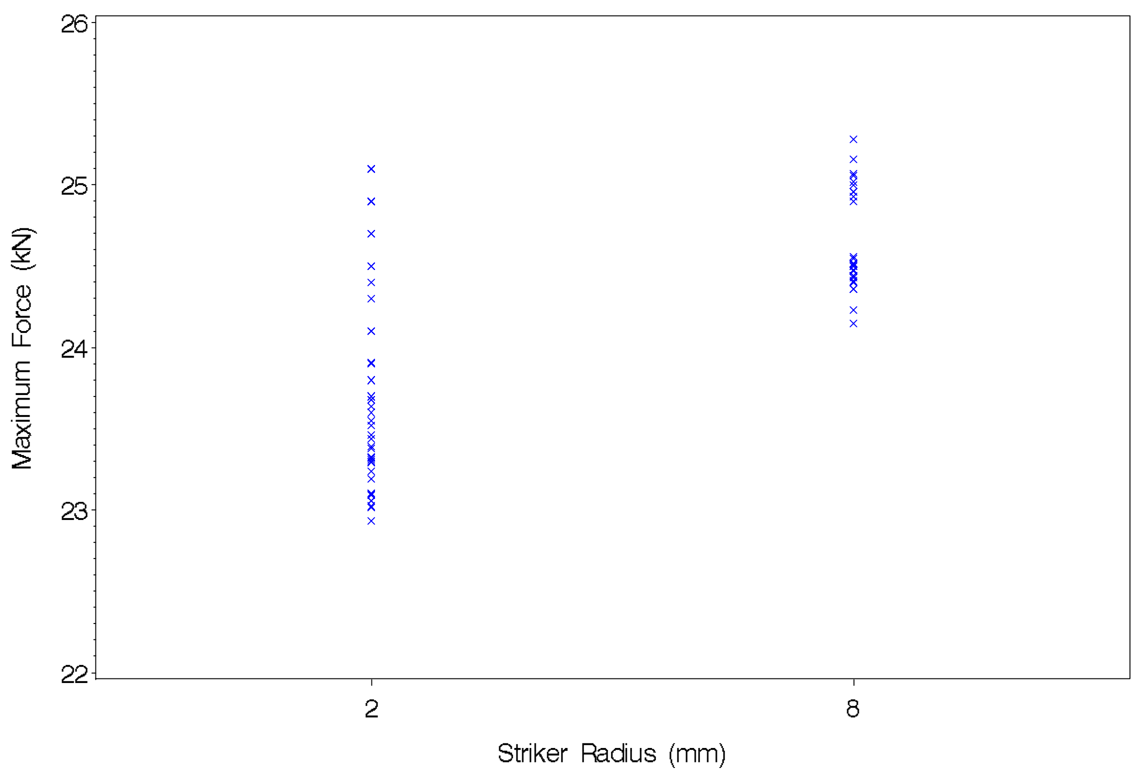

Figure 4. Maximum force versus striker radius for high-energy specimens.

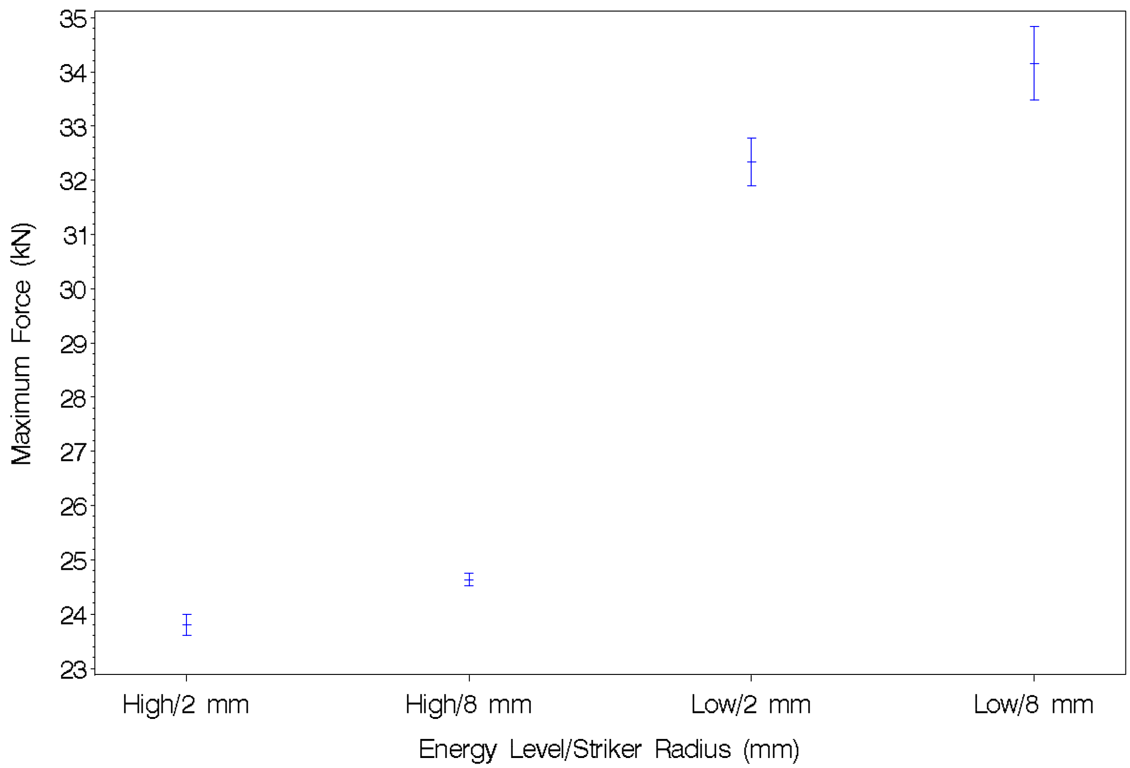

Figure 5. Maximum force means and associated error bars (twice the standard error of the mean) for both low and high energy striker means.

Table 3. Hat value, standardized residual, and Cook's D for each machine and energy level.

Machine	Low Energy			High Energy		
	Hat Value	Std Residual	Cook's D	Hat Value	Std Residual	Cook's D
1	0.14	0.1	0.001	0.14	-1.1	0.185
2	0.14	1.8	0.501	0.14	1.4	0.324
3	0.14	0.8	0.011	0.14	-0.6	0.062
5	0.14	-0.5	0.047	0.14	1.2	0.220
6	0.14	-0.7	0.078	0.14	-0.0	0.000
7	0.14	-1.5	0.372	0.14	-1.4	0.329
8	0.14	0.0	0.000	0.14	0.5	0.048

A machine is considered influential if the hat value is greater than 0.29, or if the standardized residual is larger than the absolute value of 2, or if Cook's D is larger than 0.445. None of the quantities listed in **Table 3** exceed the guidelines for detecting influential machines, indicating that it may not be necessary to calculate a separate consensus value for each striker type.

Although means for the two striker radii are statistically different for both energy levels based on a *t*-test, there is no evidence of striker differences based on an analysis of influence quantities. Thus, it is not clear if striker differences are of *practical* significance. Additional data from machines with 8 mm strikers would be helpful in deciphering the striker effect. For now, we will proceed with computing the certified reference value as if the striker radius has no effect.

4.1.2. Certified Reference Value

The certified reference value was determined by use of the Paule-Mandel method for computing a consensus value [3]. The consensus value depends on the assumption that all Charpy machines in the study are "good" and that no single machine is thought to produce better results than the others. The consensus value is a weighted mean in which the weights account for both within-machine and between-machine variation. The certified reference value (consensus value) $\widetilde{Y}$ is computed from

$$\widetilde{Y} = \frac{\sum_{i=1}^{k} \omega_i \overline{Y}_i}{\omega_i} , \tag{1}$$

where k is the number of machines, ω_i is the weight for the i^{th} machine, and $\overline{Y}_i$ is the average of the measurements for the i^{th} machine. The standard uncertainty of the certified reference value is

$$u(\widetilde{Y}) = \frac{1}{\sqrt{\sum_{i=1}^{k} \omega_i}} . \tag{2}$$

The estimated weight for the i^{th} machine is

$$\omega_i = \frac{1}{\frac{s_i^2}{n_i} + s_b^2} , \tag{3}$$

where s_i^2 is the usual sample variance of the i^{th} machine, n_i is the number of test results for the i^{th} machine, and s_b^2 is the estimated variance between machines. The between-machine variance estimate is determined by an iterative procedure [3]. We used the "consensus means" command in a statistical analysis package called DATAPLOT [4] to compute the between-machine variance. **Table 4** lists certified reference values, associated uncertainties, coverage factors, and expanded uncertainties corresponding to 95 % uncertainty intervals for low and high energy specimens.

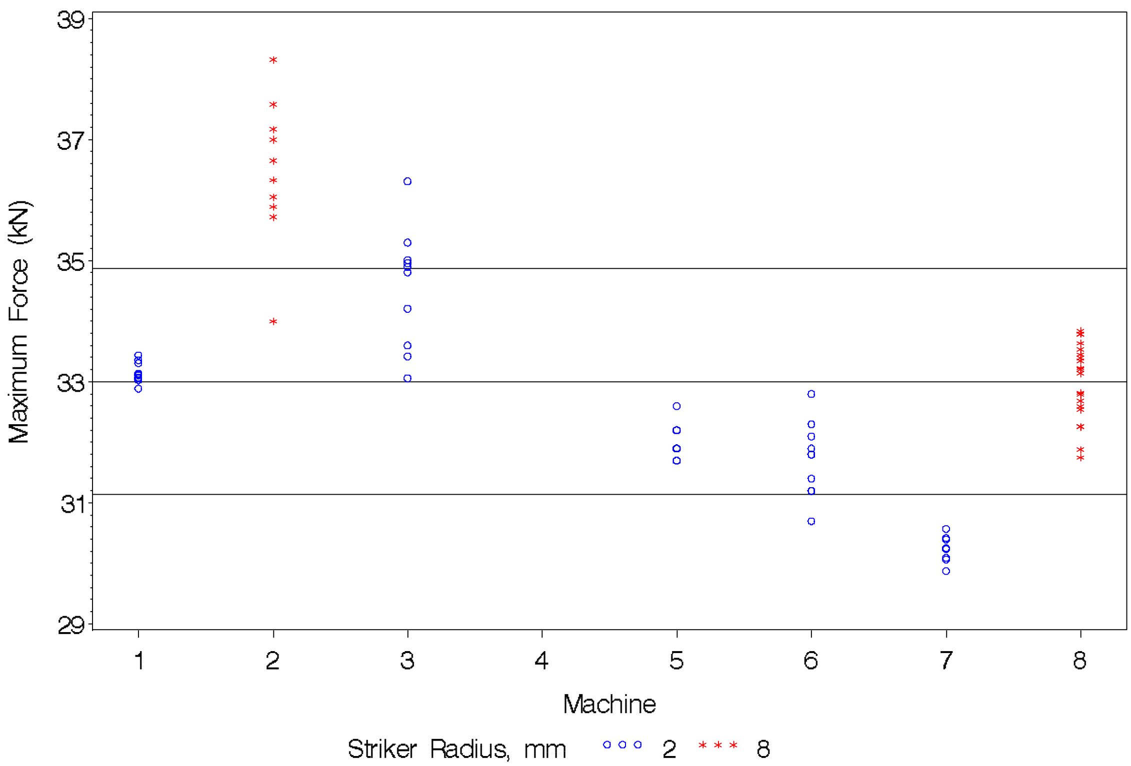

Figure 6. Low-energy maximum force for final set of machines with certified reference value (middle line). The top and bottom horizontal lines represent the 95 % uncertainty interval for the certified value.

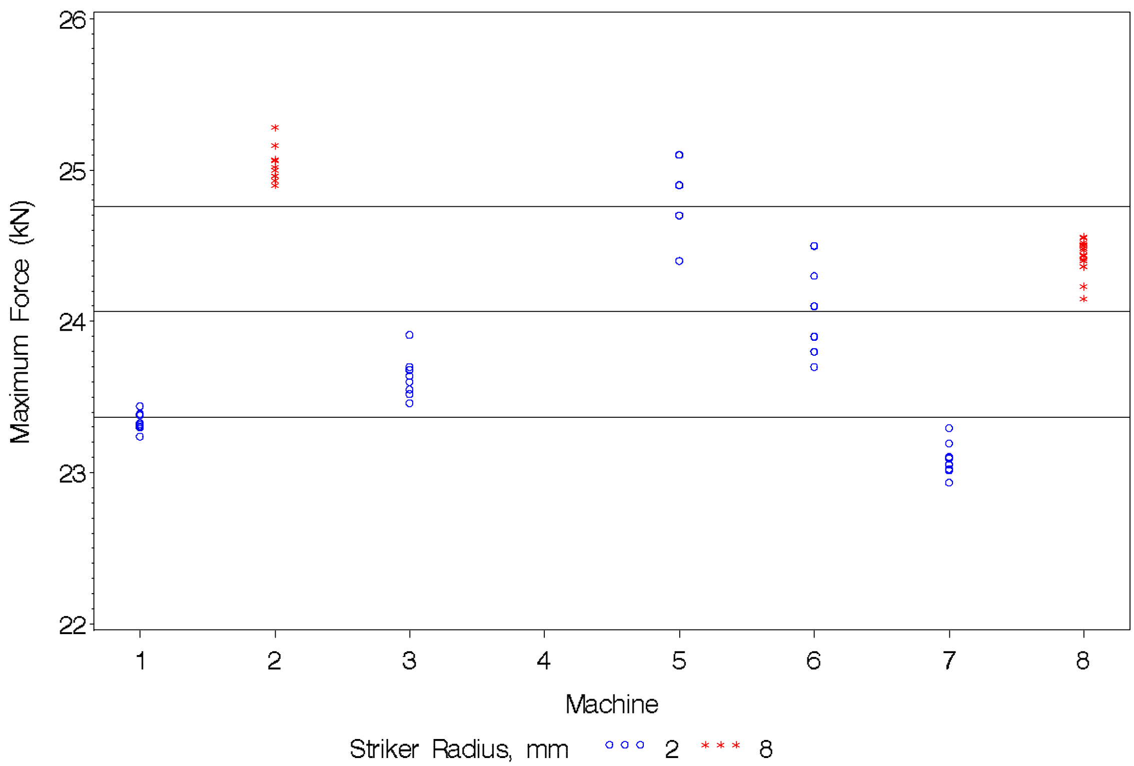

Figure 7. High-energy maximum force for final set of machines with certified reference value (middle line). The top and bottom lines represent the 95 % uncertainty interval for the certified reference value.

Table 4. Certified reference values, standard uncertainty, coverage factor, and expanded uncertainty.

Energy level	Certified reference value (kN)	Standard uncertainty (kN)	Coverage factor, k	Expanded uncertainty (kN)
Low	33.00	0.76	2.447	1.86
High	24.06	0.28	2.447	0.70

The coverage factor for both low and high energy levels is based on a *t*-table value for 7−1=6 degrees of freedom and a 95 % two-sided uncertainty interval. The data, certified reference value, and 95 % uncertainty bounds for low and high energy data are shown in **Figures 6** and **7**.

We also computed the reference value for just the machines having 2 mm strikers to compare with the reference value based on all machines. The reference values based on just the machines having 2 mm strikers were lower by 2.1 % and 1.1 % for low and high energy, respectively. Thus, the reference value is not substantially altered by including machines with 8 mm strikers.

4.1.3. Repeatability and Reproducibility

The data from this round robin are also useful in estimating the precision of maximum force measurements. The precision is evaluated according to the terms of repeatability and reproducibility defined in ASTM E 691-05. The procedures in E 691 rely on the assumption that an equal number of measurements are taken for each machine. Thus, the second set of NIST measurements was not included in the analysis. Although there are only nine measurements for machine #3 for the high energy material, this small deviation in sample size should not significantly influence the results of the analysis. The raw data for the eight machines are displayed in **Figure 1** for low-energy specimens and in **Figure 2** for high-energy specimens.

The plots of the raw data indicate that the measurements for machine #4 are not consistent with the remaining data. **Table 5** lists summary statistics for each machine and material. The deviation is the difference between the individual machine mean and the grand mean for that material. The low-energy grand mean is 31.75 kN and the high-energy grand mean is 23.38 kN. Consistency statistics for each machine are listed in **Table 6**. The “k” consistency statistic represents the consistency of variation within a laboratory, while the “h” consistency statistic summarizes the consistency between laboratories.

The critical values for h and k were taken from **Table 5** in E 691-05 for $p = 8$ and $n = 10$. An h consistency statistic is considered to be significant if its absolute value is larger than 2.15. A k consistency statistic is considered significant if it is larger than 1.55. Plots of the consistency statistics and the associated critical values are shown in **Figures 8** and **9**.

Figure 8 shows three machines with k values larger than the critical value, indicating that the within-machine variation is large compared to the other machines. Machines #2 and #3 have large variation for the low-energy material, while machine #6 has large variation for the high-energy material. Since no machine has significant k values for both materials, there is no reason to remove machines from the study based on the k statistics.

Table 5. Machine summary statistics.

Machine	Low energy (LL-103)			High energy (HH-103)		
	Mean (kN)	Standard deviation (kN)	Deviation (kN)	Mean (kN)	Standard deviation (kN)	Deviation (kN)
1	33.15	0.17	1.40	23.33	0.06	-0.05
2	36.47	1.19	4.72	25.03	0.12	1.65
3	34.56	0.99	2.81	23.64	0.13	0.26
4	22.90	0.20	-8.85	18.60	0.15	-4.78
5	32.02	0.28	0.27	24.87	0.22	1.49
6	31.72	0.61	-0.03	24.06	0.29	0.68
7	30.24	0.20	-1.51	23.09	0.10	-0.29
8	32.93	0.55	1.18	24.42	0.11	1.04

Table 6. Consistency statistics for eight Charpy machines.

Machine	Low energy (LL-103)		High energy (HH-103)	
	h	k	h	k
1	0.34766	0.26354	-0.0231	0.34751
2	1.16914	1.86445	0.80543	0.71004
3	0.69552	1.54991	0.12535	0.80409
4	-2.19078	0.31448	-2.32847	0.91881
5	0.06716	0.43727	0.72555	1.36436
6	-0.00712	0.96304	0.33101	1.79579
7	-0.37329	0.3097	-0.14263	0.61345
8	0.29171	0.85946	0.50685	0.64934

Figure 9 shows that the h consistency statistics are significant for machine #4 for both materials, indicating that the measurements for this machine are significantly different from those of the remaining machines relative to the between-machine variation. We will remove machine #4 from the study since the measurements for both low and high energy materials were significantly lower than the rest of the measurements.

Differences in variation between the high and low energy material, are due to differences in the characteristic curves for these materials. As shown in **Figures 10** and **11**, the data for high energy material, when fitted with an averaged curve, has a single maximum that defines the maximum force. The curve for the low energy material has a maximum point, but this value is likely to vary more that the high energy curve due to details of individual fracture events.

Figures 12 and **13** display the raw data for the seven machines included in the round robin for low and high energy materials, respectively.

The plots of the raw data do not indicate any inconsistencies among machines. **Table 7** lists deviations of machine means from the grand mean for each material based on seven machines. The low-energy grand average is 33.01 kN and the high-energy grand average is 24.06 kN.

The critical values for h and k in **Table 8** are based on $p = 7$ and $n = 10$. An h consistency statistic is considered to be significant if its absolute value is larger than 2.05. A k consistency statistic is considered significant if it is larger than 1.54. **Figures 14** and **15** display the consistency statistics from Table 8 in graphical form along with the critical values.

Table 8 displays h and k consistency statistics based on seven machines for both low and high energy materials.

Table 7. Deviations for seven machines.

Machine	**Low energy deviation (kN)**	**High energy deviation (kN)**
1	0.14	-0.73
2	3.46	0.97
3	1.55	-0.43
5	-0.99	0.81
6	-1.29	0.00
7	-2.77	-0.98
8	-0.09	0.36

Table 8. Consistency statistics based on seven machines.

Machine	**Low energy**		**High energy**	
	h	**k**	**h**	**K**
1	0.06905	0.24805	-0.972	0.34371
2	1.70392	1.75492	1.29185	0.70227
3	0.76133	1.45886	-0.56638	0.79529
5	-0.48921	0.41158	1.07358	1.34943
6	-0.63703	0.90646	-0.00444	1.77614
7	-1.36577	0.29151	-1.29861	0.60673
8	-0.0423	0.80896	0.47601	0.64223

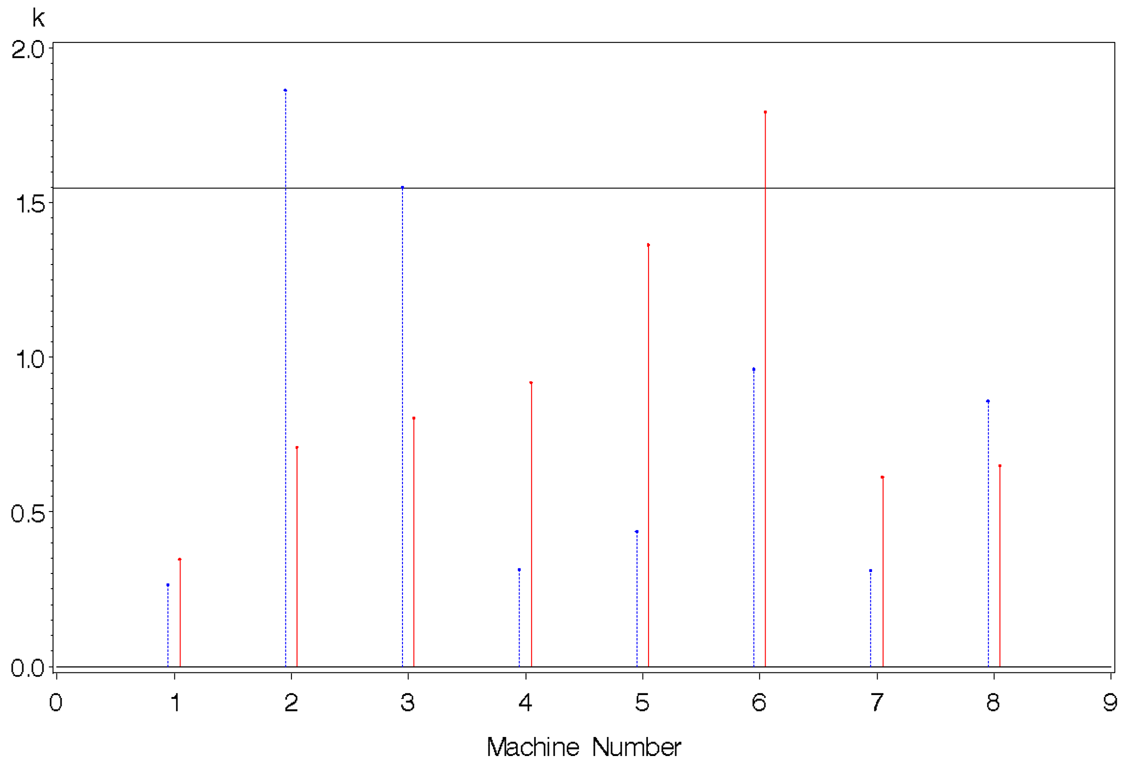

Figure 8. Within-laboratory consistency k statistics for eight Charpy machines. The dashed blue lines (left) correspond to the low energy data and the solid red lines (right) represent high energy data.

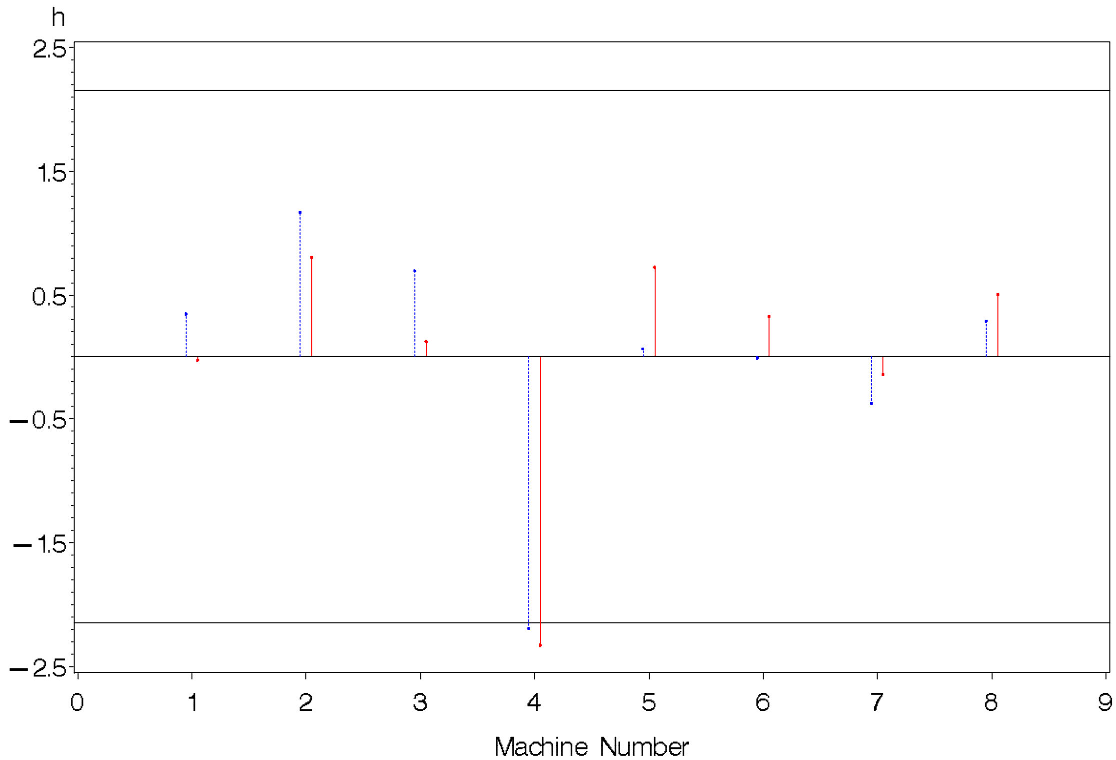

Figure 9. Between-laboratory consistency h statistics for eight Charpy machines. The dashed blue lines (left) correspond to the low energy data and the solid red lines (right) represent high energy data.

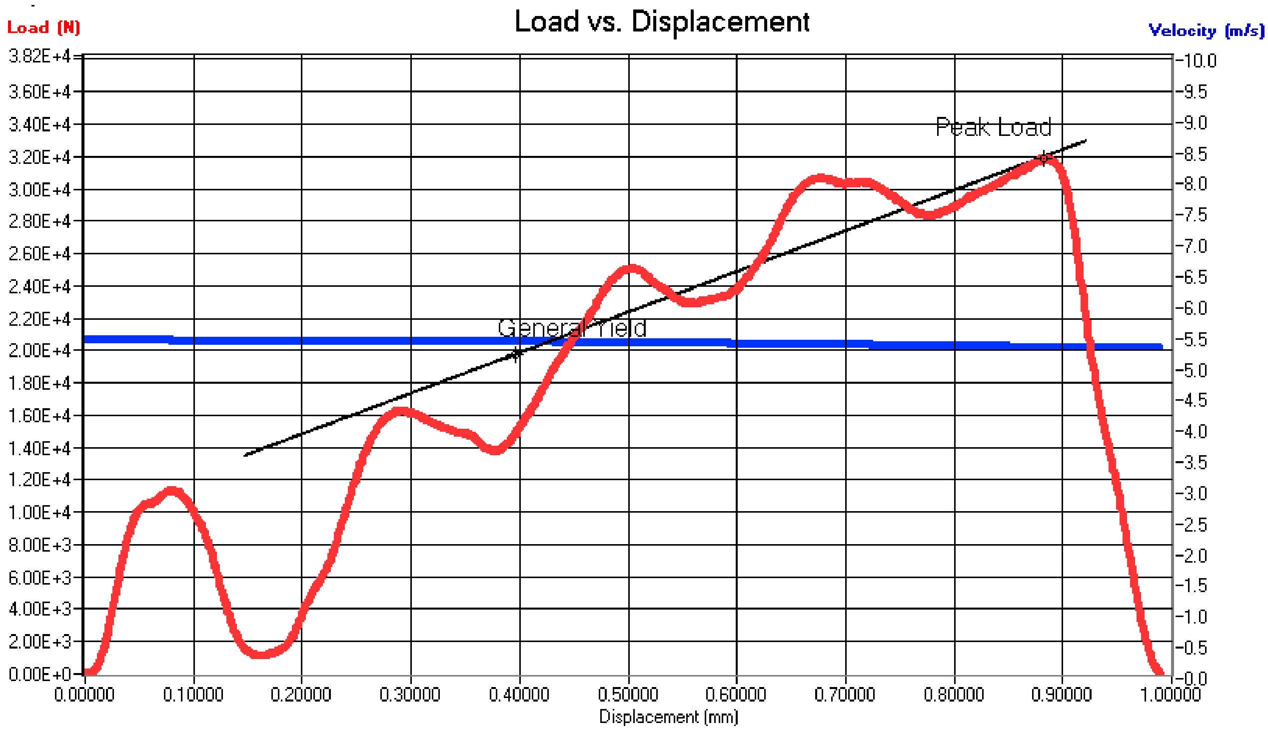

Figure 10. Low energy curve showing maximum force (peak load).

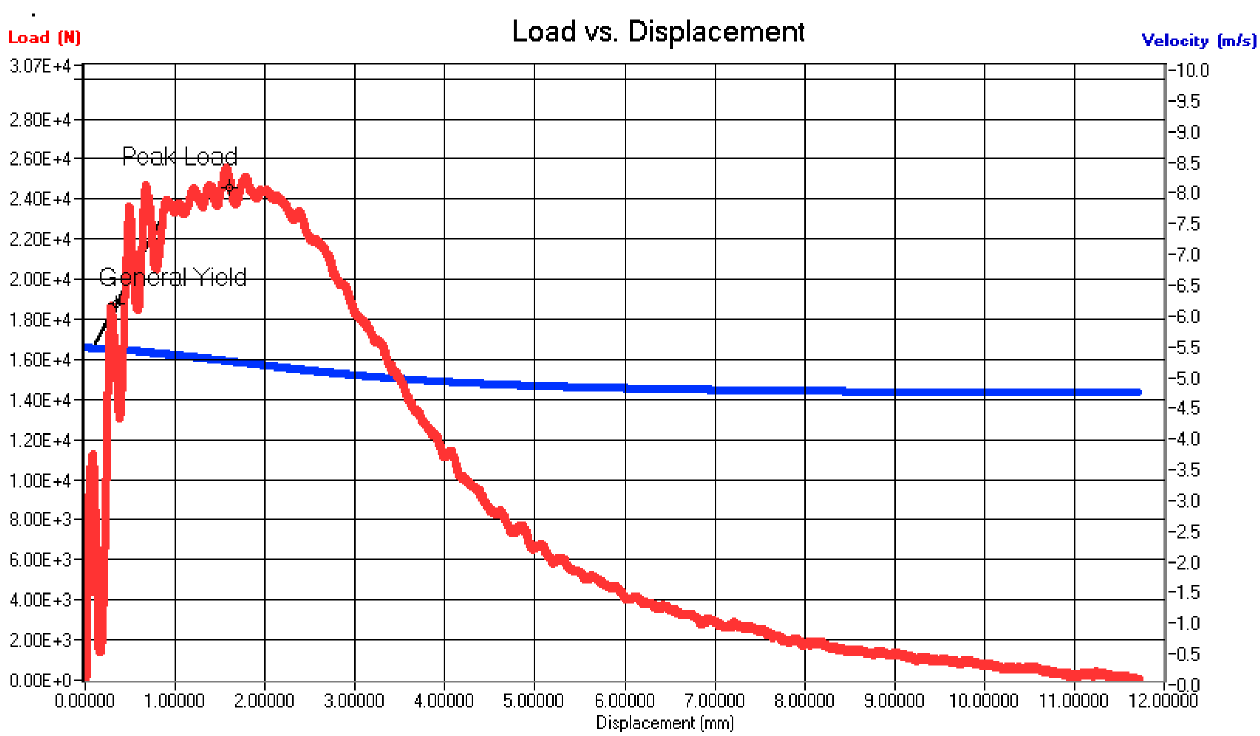

Figure 11. High energy curve showing maximum force at the maxima of the fitted curve (peak load).

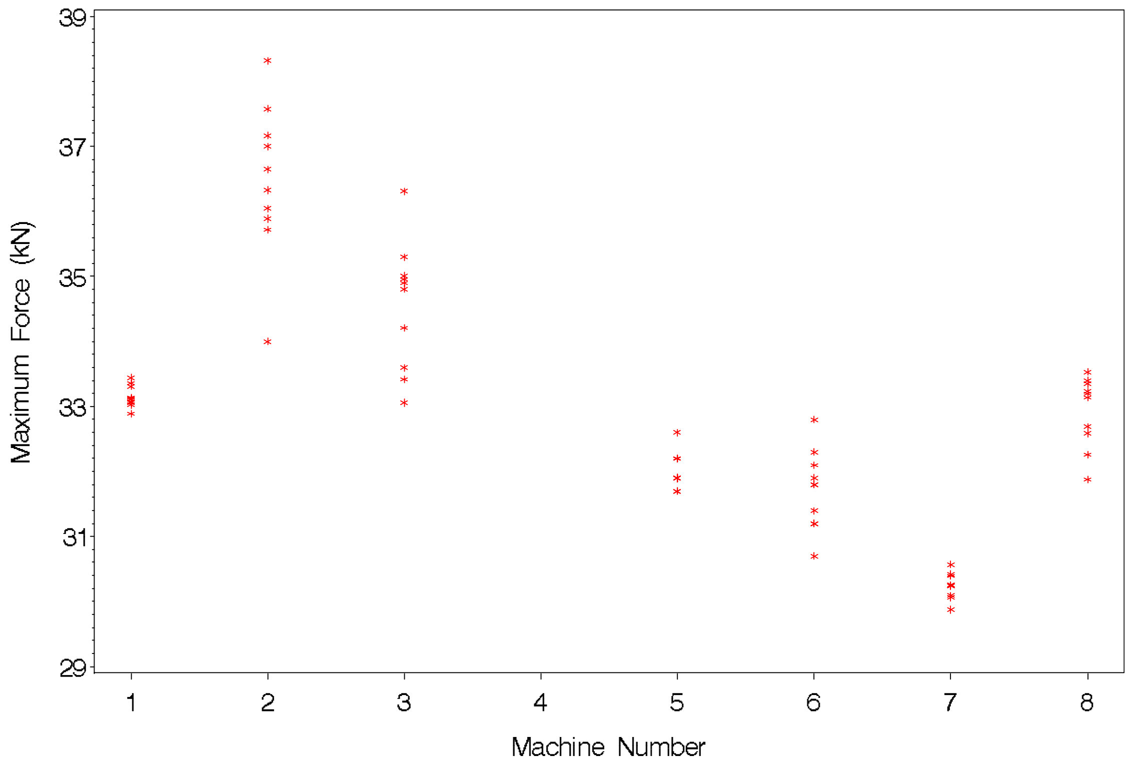

Figure 12. Maximum force for seven Charpy machines at the low energy level.

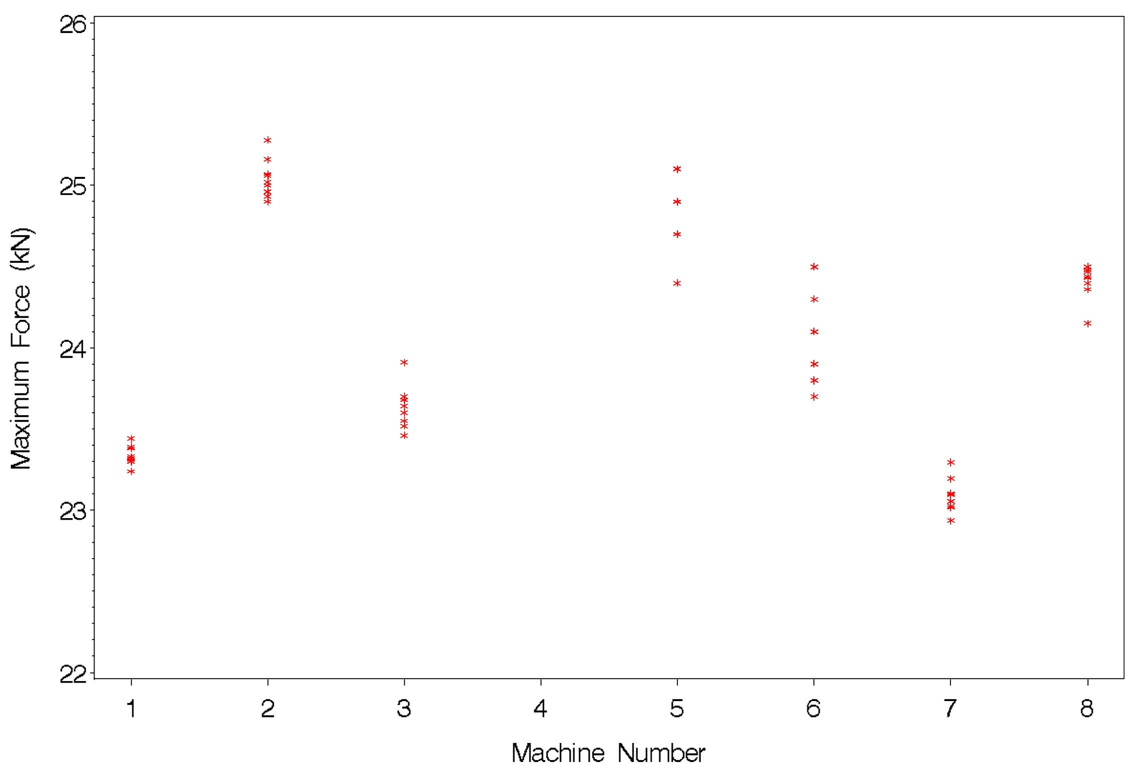

Figure 13. Maximum force for seven Charpy machines at the high energy level.

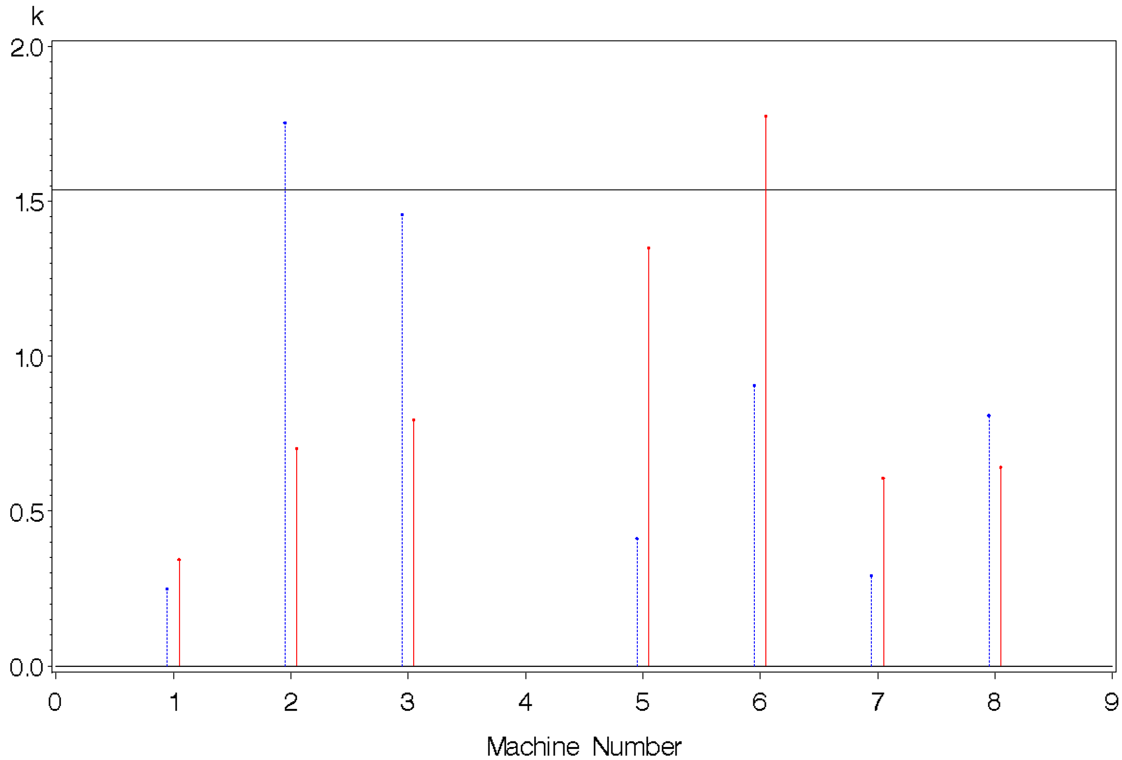

Figure 14. Within-laboratory consistency k statistics for seven Charpy machines. The dashed blue lines (left) correspond to the low energy data and the solid red lines (right) represent high energy data.

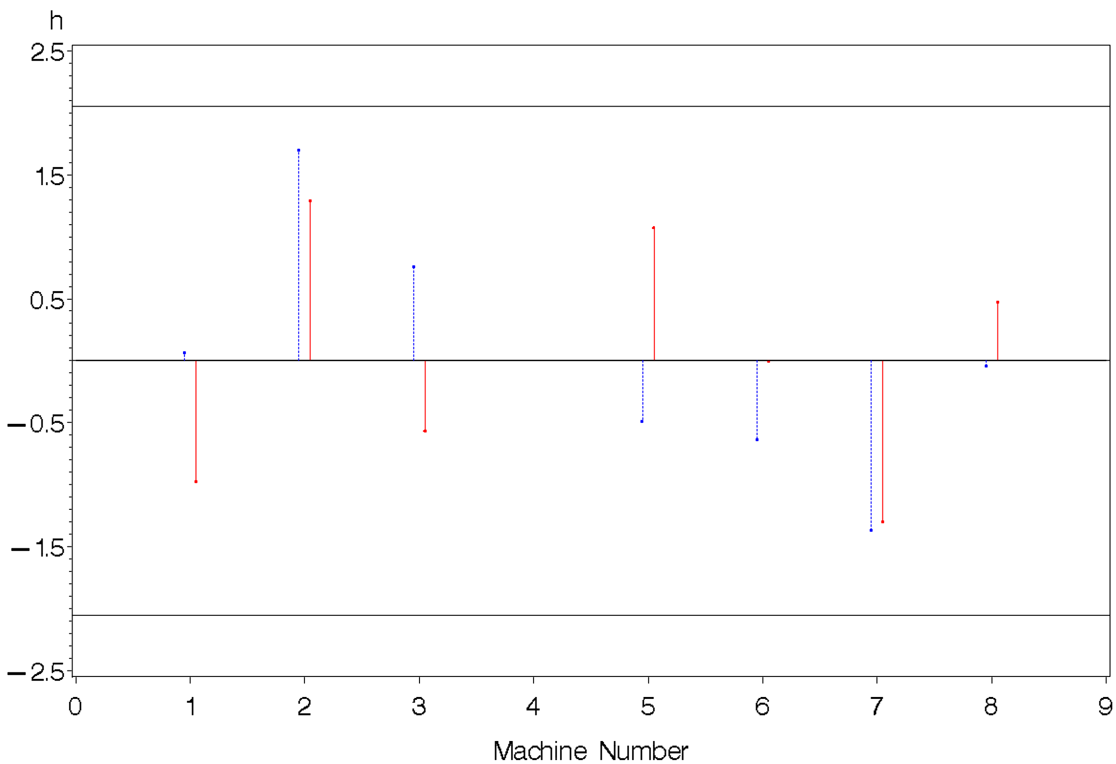

Figure 15. Between-laboratory consistency h statistics for seven Charpy machines. The dashed blue lines (left) correspond to the low energy data and the solid red lines (right) represent high energy data.

Table 9. Precision statistics based on seven machines.

ID	Grand mean (kN)	Repeatability standard deviation, S_r (kN)	95 % repeatability limit, r (kN)	Reproducibility standard deviation, S_R (kN)	95 % reproducibility limit, R (kN)
HH-108	24.06	0.16	0.45	0.77	2.16
LL- 108	33.01	0.68	1.90	2.13	5.96

The data in **Figures 14** and **15** do not indicate that any machines should be removed from the interlaboratory comparison. Significant k statistics for machines #3 and #6 are observed for only one material, and there are no significant h statistics.

Since the data from seven machines are consistent, the precision statistics for each material can be computed. **Table 9** lists the grand mean, repeatability standard deviation, reproducibility standard deviation, the 95 % repeatability limit ($2.8S_r$) and the 95 % reproducibility limit ($2.8S_R$).

The grand mean is the unweighted average of the individual machine means and is not the same as the certified reference value (which is a weighted mean) discussed earlier.

4.2. Forces at General Yield

Forces at general yield (F_{gy}) are displayed in **Figure 16** and **17** for low-energy and high-energy specimens, respectively. No results were provided for machine #4. For machine #7, forces at general yield were reported only for high-energy specimens; therefore, for low-energy specimens we have assumed that F_{gy} corresponds to F_m. Consequently, values of maximum force have been used in the analyses for machine #7 with low-energy specimens.

The data appear to be fairly consistent among machines with the exception of machine #6, which provides test results much lower than the rest of the data for both low and high energies. Data from machine #6 are therefore excluded from further analyses.

4.2.1. Striker Radius

Figures 18 and 19 display data for general yield forces versus striker radius for low and high energies, respectively.

In order to establish whether striker radius has an effect on force at general yield, we performed a *t*-test to determine if the mean of the data from machines with 8 mm striker is statistically different from the mean of data from machines with 2 mm strikers for each energy level. Based on the *t*-test (one tailed distribution, two-sample unequal variance), the 8 mm striker mean is statistically different from the 2 mm striker only for high energies ($p < 0.0001$) but not for low

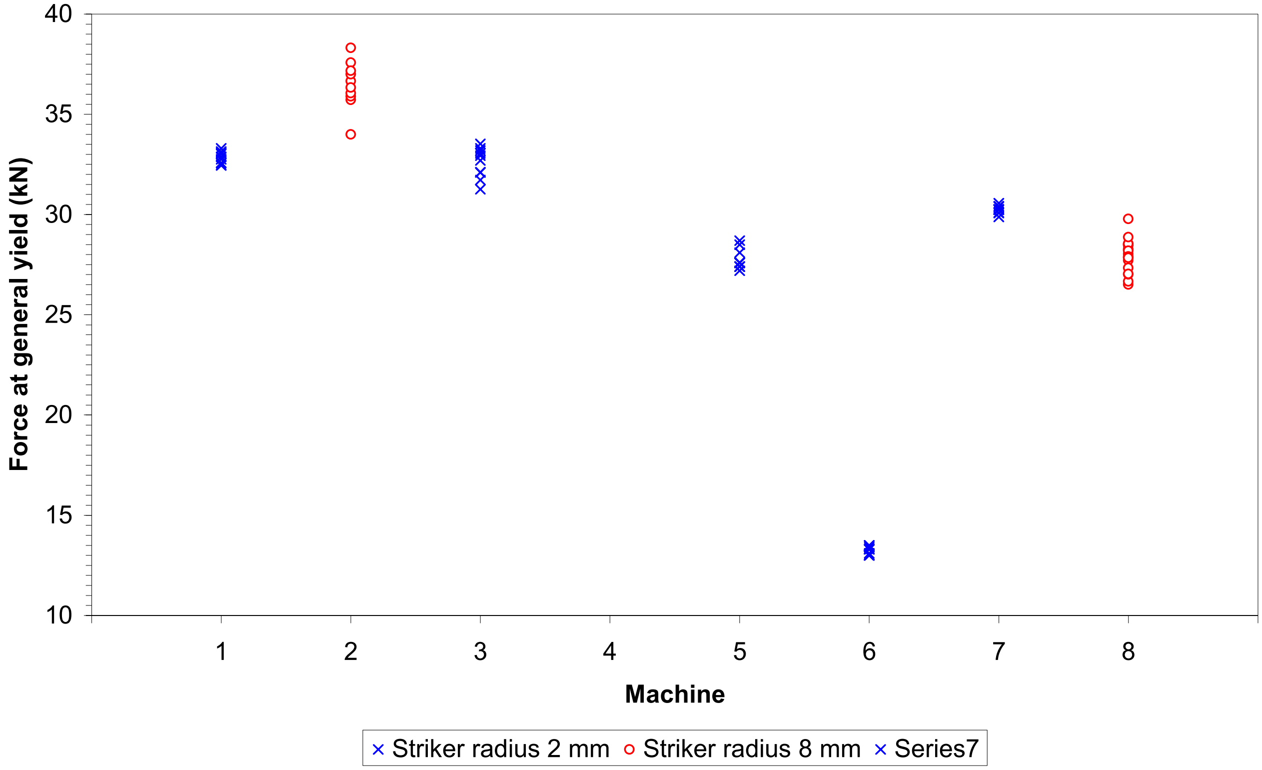

Figure 16. Low-energy force at general yield F_{gy} for seven machines.

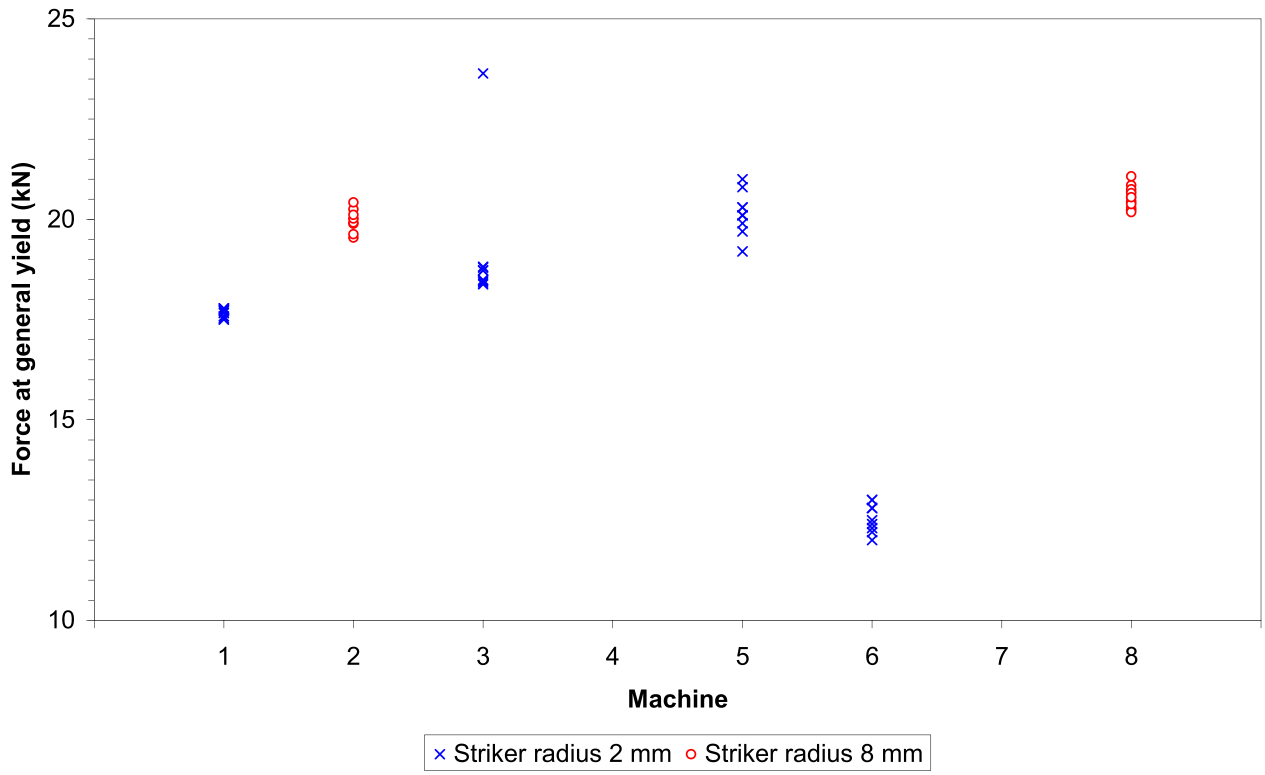

Figure 17. High-energy force at general yield F_{gy} for seven machines.

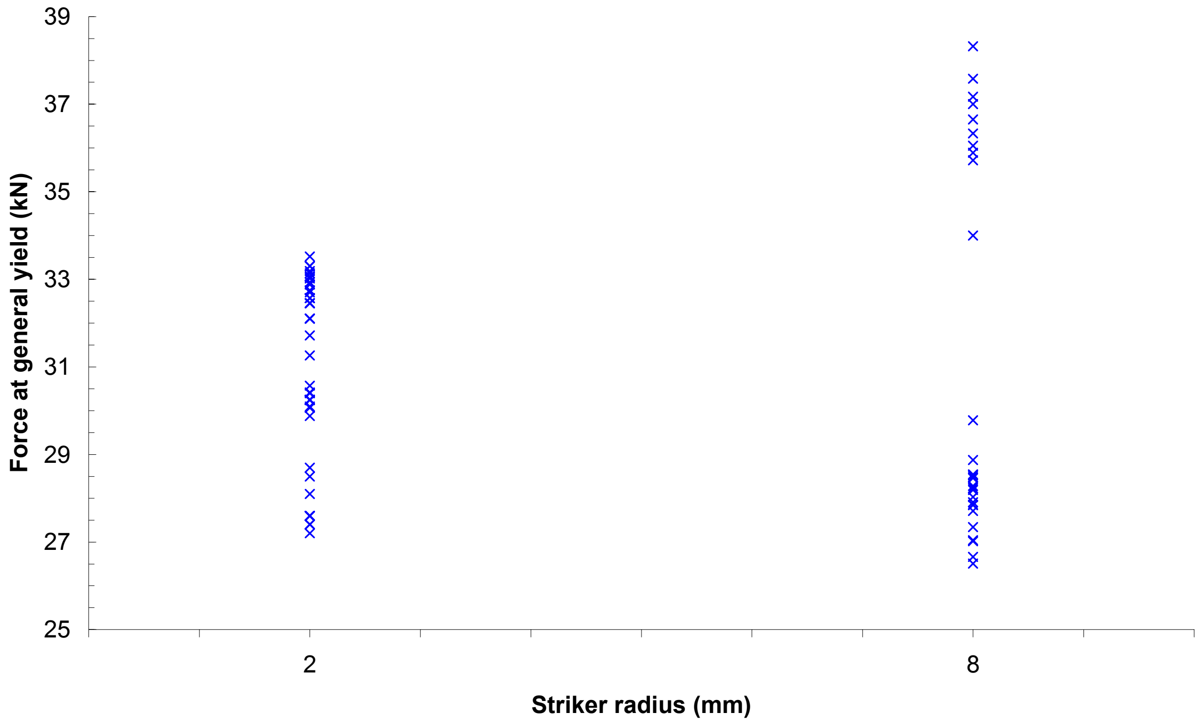

Figure 18. General yield force F_{gy} versus striker radius for low-energy specimens.

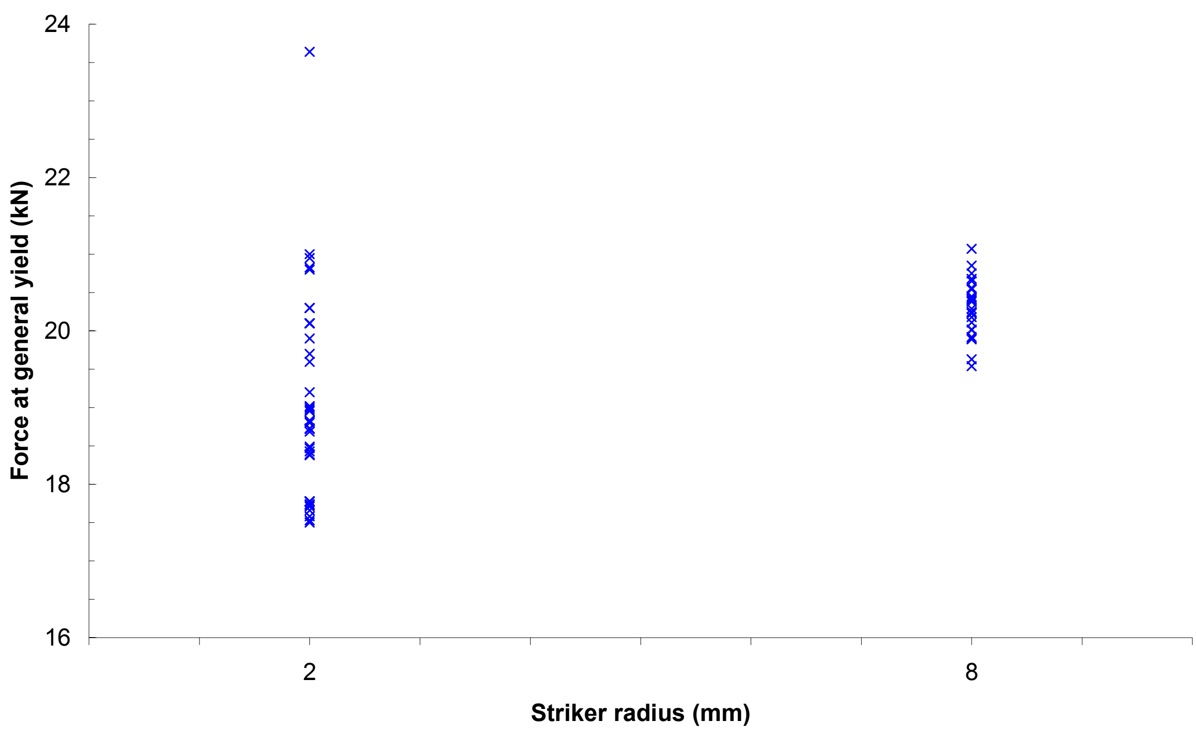

Figure 19. General yield force F_{gy} versus striker radius for high-energy specimens.

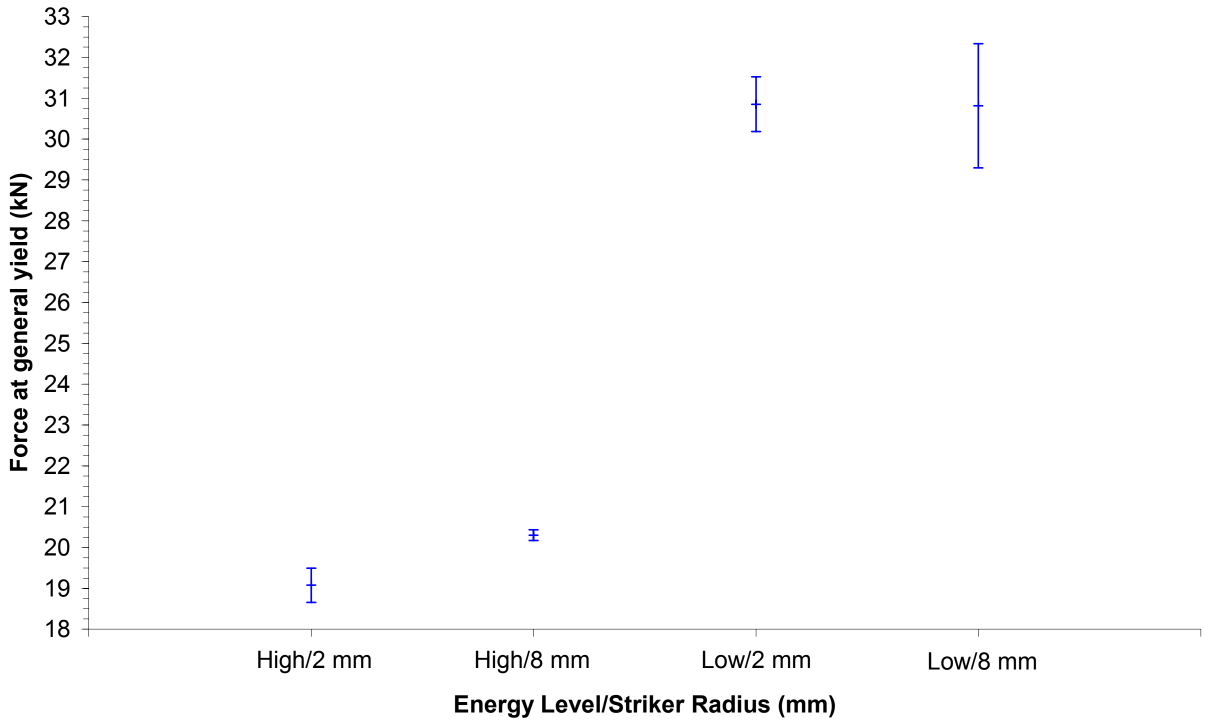

Figure 20. F_{gy} means and associated error bars (twice the standard error of the mean) for both low and high energy striker means.

Table 10. Machine summary statistics.

Machine	Low energy (LL-103)			High energy (HH-103)		
	Mean (kN)	**Standard deviation (kN)**	**Deviation (kN)**	**Mean (kN)**	**Standard deviation (kN)**	**Deviation (kN)**
1	32.82	0.30	4.11	17.68	0.11	-0.80
2	36.47	1.19	7.76	19.97	0.26	1.49
3	32.59	0.75	3.88	19.16	1.69	0.69
5	27.77	0.50	-0.94	20.15	0.51	1.67
6	13.24	0.20	-15.47	12.58	0.35	-5.90
7	30.24	0.20	1.53	19.32	0.88	0.84
8	27.83	1.00	-0.88	20.55	0.28	2.07

Table 11. Consistency statistics for seven Charpy machines.

Machine	Low Energy (LL-103)		High Energy (HH-103)	
	h	k	h	k
1	0.54992	0.43836	-0.28914	0.03817
2	1.03804	1.70618	0.53897	0.09479
3	0.51876	1.07941	0.24807	0.60973
5	-0.12555	0.71481	0.60439	0.18529
6	-2.06867	0.28938	-2.13188	0.12614
7	0.20490	0.28341	0.30543	0.31760
8	-0.11740	1.43411	0.74898	0.09975

energies ($p > 0.0001$). **Figure 20** displays the means of each group along with error bars (twice the standard error of the mean, or $2s/\sqrt{n}$).

The error bars in **Figure 20** do not overlap for high energy, also indicating that there is a significant difference between 2 mm and 8 mm striker means. On the other hand, they do overlap for low energy, thus confirming that there is no significant difference between these means.

4.2.2. Repeatability and Reproducibility

The precision of general yield force measurements is evaluated according to the terms of repeatability and reproducibility defined in ASTM E691-05.

For reasons already explained, the second set of NIST measurements was not included in the analysis. Moreover, the fact that machine #3 has only 9 measurements instead of 10, like all remaining machines, is considered to have negligible influence on the results of the analysis. The raw data for the seven machines are displayed in **Figure 16** for low-energy specimens and in **Figure 17** for high-energy specimens.

The plots of the raw data indicate that the measurements for machine #6 are not consistent with the remaining data, most likely due to an incorrect interpretation of the instrumented curves. **Table 10** lists summary statistics for each machine (including machine #6) and material. The low-energy grand mean is 28.71 kN and the high-energy grand mean is 18.48 kN.

Consistency statistics for each machine are listed in **Table 11**.

The critical values for h and k were taken from Table 5 in E 691-05 for $p = 7$ and $n = 10$. An h consistency statistics is considered significant if its absolute value is larger than 2.05. A k consistency statistic is considered significant if it is larger than 1.54. Plots of the consistency statistics and the associated critical values are shown in **Figures 21** and **22**.

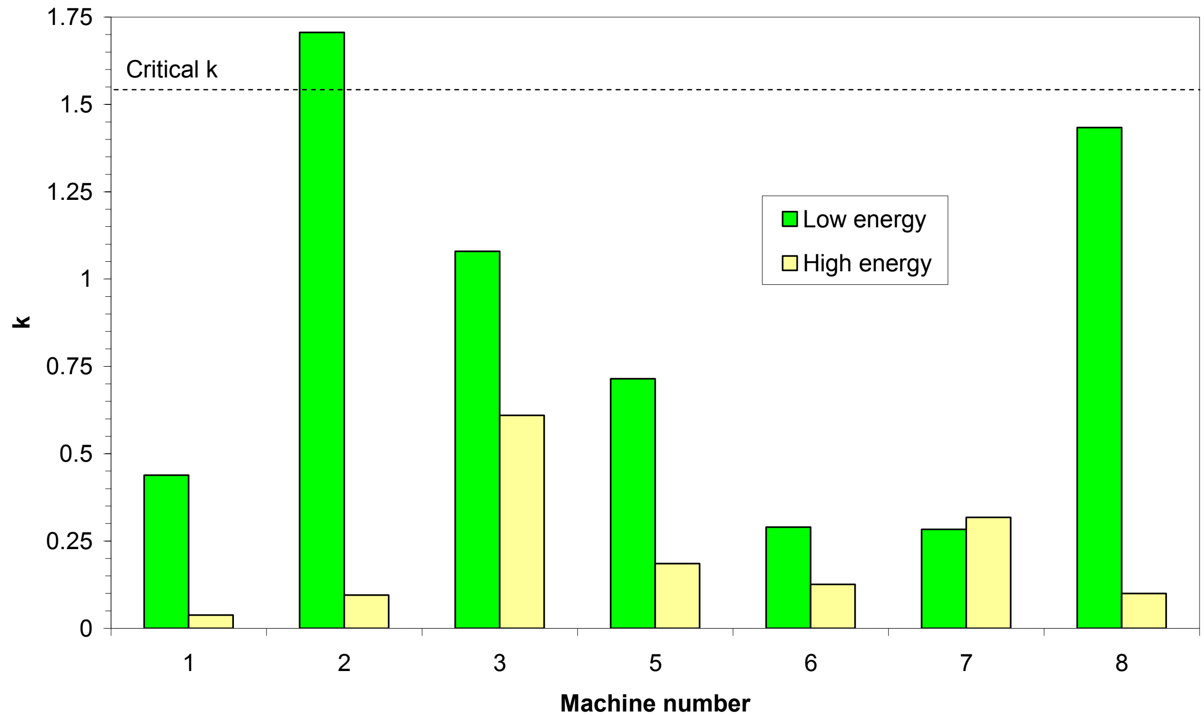

Figure 21. Within-laboratory consistency k statistics for seven Charpy machines.

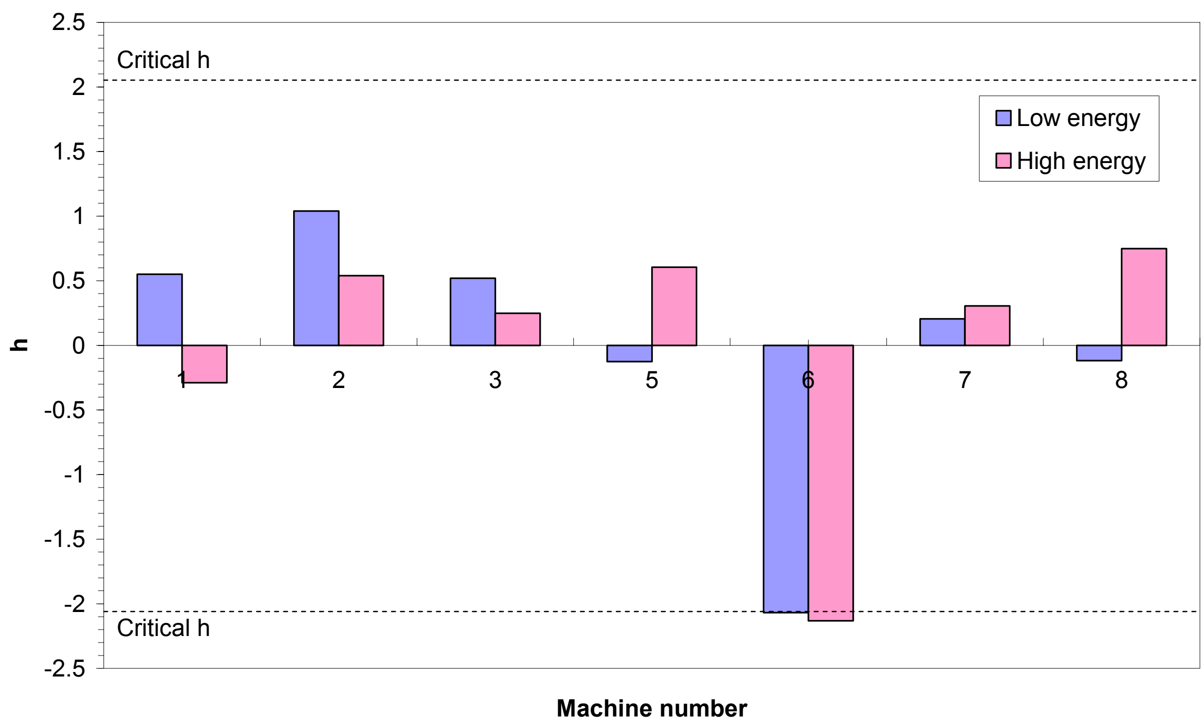

Figure 22. Between-laboratory consistency h statistics for seven Charpy machines.

One machine (#2) has a k value larger than the critical value for low energy, indicating that its within-machine variation is large compared to the other machines. Machine #8 also has a large variation for the low-energy material, but its k statistics remains below the critical value. Since none of the machines has significant k values for both materials, there is no reason to remove any of them from the study based on k statistics.

The h consistency statistics are significant for machine #6 for both materials, indicating that the measurements for this machine are significantly different from those of the remaining machines relative to the between-machine variation. We will remove machine #6 from the study since the measurements for both low and high energy materials were significantly lower than the rest of the measurements.

Figures 23 and 24 display the raw data for the six machines included in the round-robin for low and high energy materials, respectively.

The plots of the raw data do not indicate any inconsistencies among machines. **Table 12** lists deviations of machine means from the grand mean for each material based on six machines. The low-energy grand mean is 30.84 kN and the high-energy grand average is 19.60 kN.

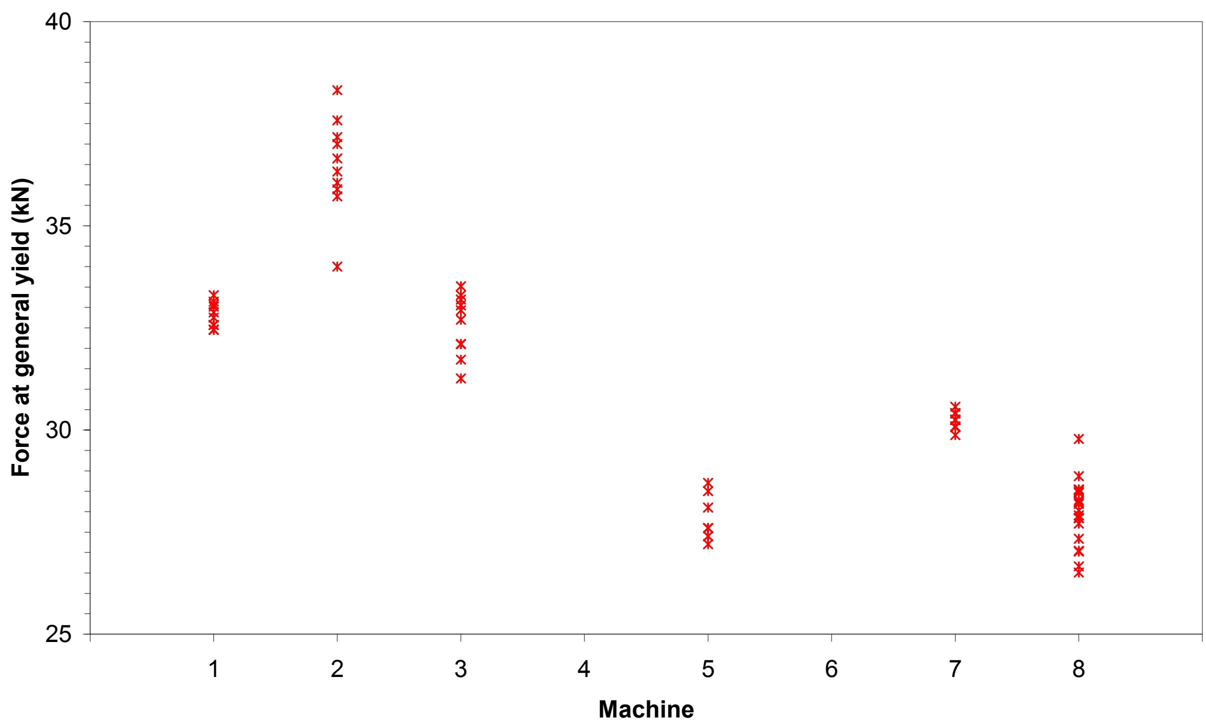

Figure 23. Force at general yield F_{gy} for six Charpy machines at the low energy level.

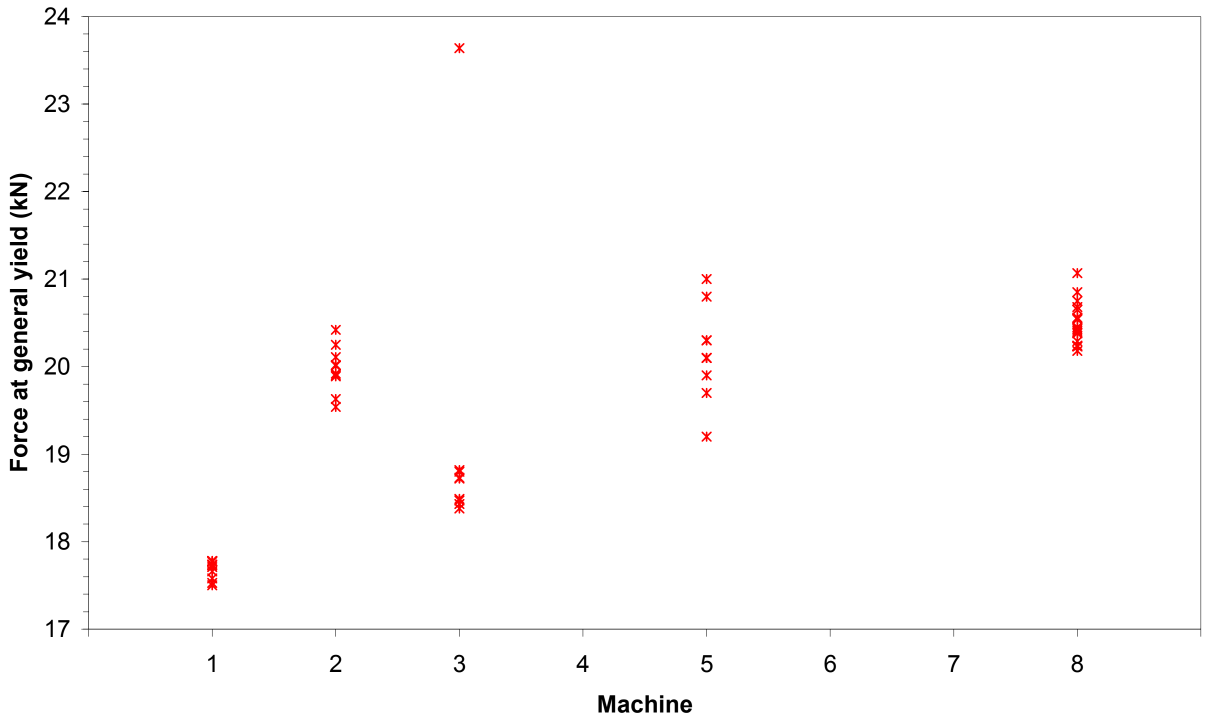

Figure 24. Force at general yield F_{gy} for six Charpy machines at the high energy level.

Table 12. Deviations for six machines.

Machine	Low energy deviation (kN)	High energy deviation (kN)
1	1.98	-1.92
2	5.63	0.37
3	1.75	-0.44
5	-3.07	0.55
7	-0.60	-0.28
8	-3.01	0.95

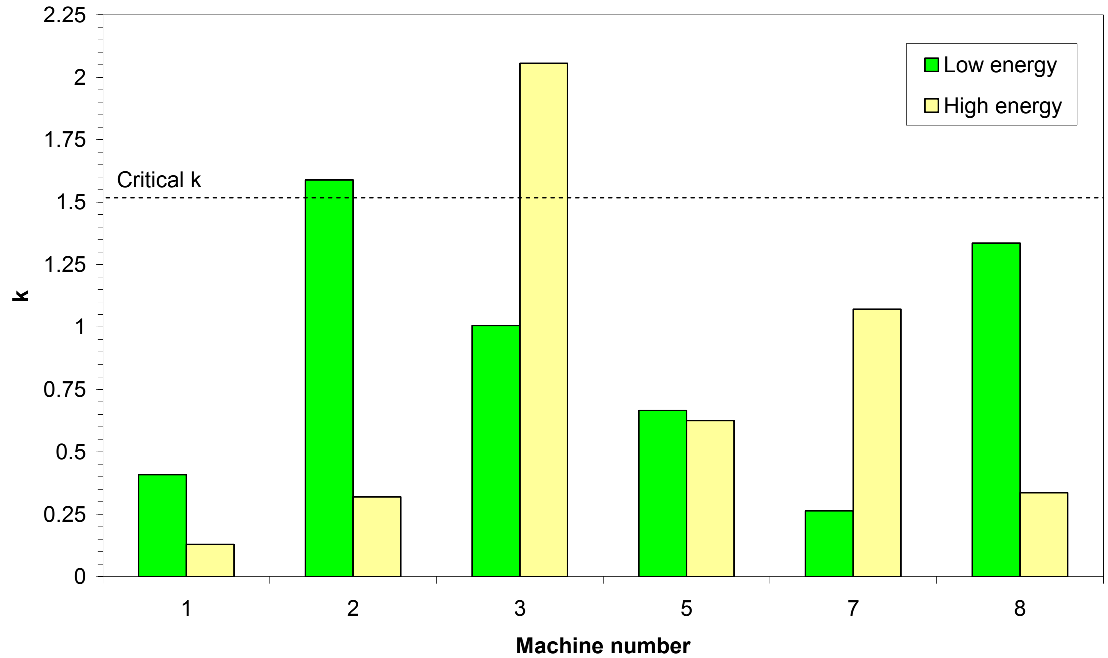

Figure 25. Within-laboratory consistency statistics for six Charpy machines.

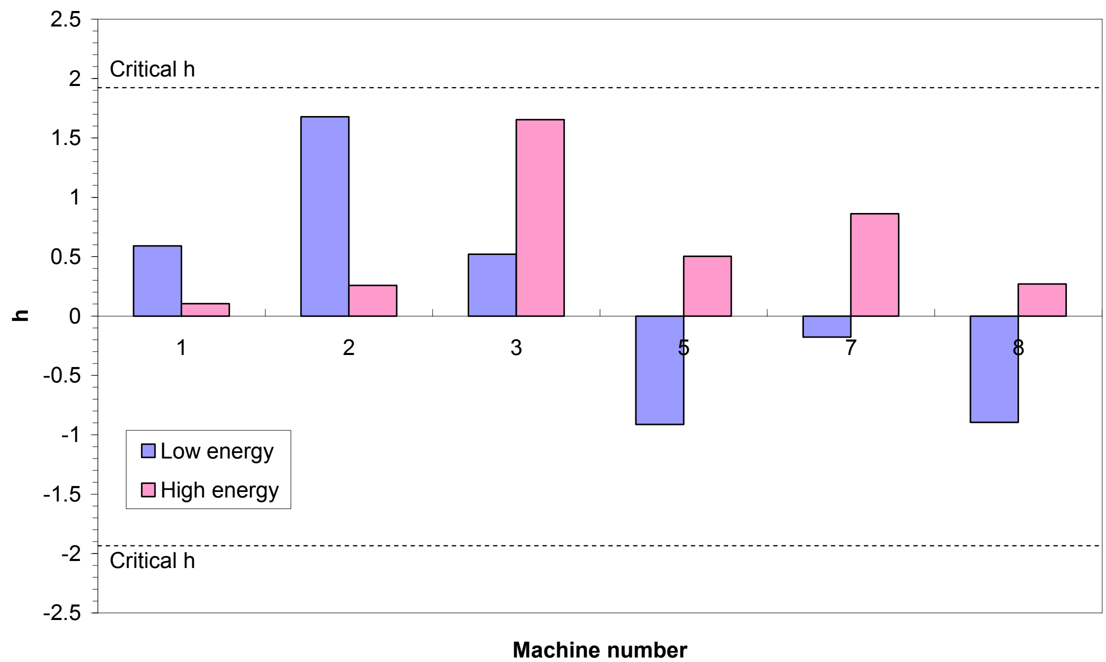

Figure 26. Between-laboratory consistency h statistics for six Charpy machines.

Table 13. Consistency statistics based on six machines.

Machine	Low energy		High energy	
	h	k	h	k
1	0.59060	0.40829	0.10351	0.12875
2	1.67808	1.58915	0.25704	0.31972
3	0.52118	1.00537	1.65342	2.05664
5	-0.91429	0.66578	0.50246	0.62500
7	-0.17808	0.26397	0.86123	1.07126
8	-0.89611	1.33574	0.27050	0.33646

Table 13 displays h and k consistency statistics based on six machines for both low and high energy materials.

The critical values for h and k in **Table 13** are based on $p = 6$ and $n = 10$. An h consistency statistic is considered to be significant if it is larger than 1.92. A k consistency statistic is considered to be significant if it is larger than 1.52. **Figures 25 and 26** display the consistency statistics from **Table 13** in graphical form along with the critical values

The data in **Figure 25** and **26** do not indicate that any machines should be removed from the interlaboratory comparison. Significant k statistics for machines #2 and #3 are observed only for one material, and there are no significant h statistics.

Since the general yield data from six machines is consistent, the precision statistics for each material can be computed. **Table 14** lists the grand mean, repeatability standard deviation, reproducibility standard deviation, the 95 % repeatability limit ($2.8S_r$) and the 95 % reproducibility limit ($2.8S_R$).

Table 14. Precision statistics based on six machines.

ID	Grand mean (kN)	Repeatability standard deviation, s_r (kN)	95 % repeatability limit, r (kN)	reproducibility standard deviation, s_R (kN)	95 % reproducibility limit, r (kN)
HH-103	19.60	0.82	2.30	1.28	3.59
LL-103	30.84	0.75	2.09	3.43	9.60

5. References

[1] F. W. Satterthwaite, An approximate distribution of estimates of variance components, Biometrics Bulletin, 2: 110-114; 1948.

[2] C. M. Wang and H. K. Iyer, Detection of influential observations in the determination of the weighted-mean KCRV, Metrologia, 42: 262-265; 2005.

[3] R. C. Paule and J. Mandel, Consensus values and weighting factors, NIST Journal of Research, 87 (5): 303-308; 1982.

[4] N. A. Heckert and J. J. Filliben, DATAPLOT reference manual, volume I: Commands, NIST Handbook 148; 2003.

Appendix A. **Raw Data from Force Verification Round Robin**

Machine	Test Number	Fm (kN) Low Energy	Fm (kN) High Energy
1	1	33.12	23.39
1	2	33.14	23.44
1	3	33.03	23.32
1	4	33.06	23.31
1	5	33.31	23.38
1	6	33.07	23.24
1	7	32.89	23.30
1	8	33.11	23.32
1	9	33.36	23.30
1	10	33.44	23.33
2	1	36.65	25.06
2	2	37.58	25.16
2	3	37.00	24.93
2	4	38.32	24.90
2	5	35.72	24.96
2	6	37.17	24.96
2	7	35.89	25.07
2	8	34.00	25.00
2	9	36.05	25.02
2	10	36.33	25.28
3	1	34.90	23.60
3	2	34.81	23.52
3	3	34.96	23.68
3	4	34.21	23.68
3	5	33.60	23.91
3	6	33.42	23.46
3	7	35.30	23.64
3	8	35.01	23.70
3	9	33.06	23.55
3	10	36.31	
4	1	22.90	18.50
4	2	22.70	18.30
4	3	22.70	18.50
4	4	23.20	18.60
4	5	22.70	18.60
4	6	23.00	18.60
4	7	22.70	18.60
4	8	23.00	18.80
4	9	23.20	18.70
4	10	22.90	18.80
5	1	31.70	24.40
5	2	31.90	24.90
5	3	32.20	24.70
5	4	31.90	25.10
5	5	32.20	24.90

Machine	Test Number	Fm (kN) Low Energy	Fm (kN) High Energy
5	6	31.90	25.10
5	7	32.60	24.90
5	8	32.20	25.10
5	9	31.90	24.70
5	10	31.70	24.90
6	1	31.90	24.30
6	2	31.80	24.50
6	3	32.80	24.50
6	4	31.40	24.10
6	5	32.10	23.90
6	6	31.20	23.80
6	7	30.70	23.70
6	8	32.30	23.90
6	9	31.80	23.80
6	10	31.20	24.10
7	1	30.42	23.30
7	2	30.24	23.10
7	3	30.09	23.02
7	4	30.40	23.11
7	5	30.25	23.03
7	6	30.57	22.94
7	7	30.07	23.06
7	8	30.25	23.19
7	9	30.25	23.10
7	10	29.88	23.06
8a	1	33.35	24.40
8a	2	33.14	24.43
8a	3	32.69	24.47
8a	4	33.40	24.44
8a	5	33.23	24.50
8a	6	32.26	24.48
8a	7	33.53	24.15
8a	8	33.20	24.50
8a	9	32.59	24.48
8a	10	31.88	24.36
8b	1	32.54	24.23
8b	2	33.84	24.51
8b	3	32.82	24.52
8b	4	32.79	24.41
8b	5	33.64	24.51
8b	6	31.75	24.36
8b	7	33.45	24.40
8b	8	33.79	24.56
8b	9	32.26	24.55
8b	10	33.79	24.44