Journal of Testing and Evaluation

Enrico Lucon[1] and Raymond L. Santoyo[2]

DOI: 10.1520/JTE20220198

Analyzing the NIST Charpy Program Database: Influence of Impact Hammer Type (C versus U) on Test Results

Journal of Testing and Evaluation

doi:10.1520/JTE20220198 available online at www.astm.org

Enrico Lucon[1] and Raymond L. Santoyo[2]

Analyzing the NIST Charpy Program Database: Influence of Impact Hammer Type (C versus U) on Test Results

Manuscript received April 20, 2022; accepted for publication May 9, 2022; published online July 20, 2022.

1 Applied Chemicals and Materials Division, National Institute of Standards and Technology (NIST), 325 Broadway, Boulder 80305, CO, USA (Corresponding author), e-mail: enrico.lucon@nist.gov, https://orcid.org/0000-0002-3021-4785

2 Applied Chemicals and Materials Division, National Institute of Standards and Technology (NIST), 325 Broadway, Boulder 80305, CO, USA

Reference

E. Lucon and R. L. Santoyo, "Analyzing the NIST Charpy Program Database: Influence of Impact Hammer Type (C versus U) on Test Results," *Journal of Testing and Evaluation* https://doi.org/10.1520/JTE20220198

ABSTRACT

The NIST Charpy Machine Verification Program has been supplying ASTM E23-18, *Standard Test Methods for Notched Bar Impact Testing of Metallic Materials* certified Charpy reference specimens to customers all over the world since 1989, when the program was transferred from the US Army at Watertown Arsenal, Massachusetts. Starting in 1993, customer test results have been recorded in the form of absorbed energy values in an electronic database, which currently contains more than 70,000 records. In this study, data from the period 2000–2021 have been analyzed to investigate the influence of machine type/hammer configuration (C-type versus U-type) on Charpy test results at three absorbed energy levels. The significance of the trends observed has been assessed statistically, and compared to the trends observed for the three NIST reference machines used for the certification of Charpy reference specimens. The current line-up of NIST machines (one C-type and two U-type) has been found to provide the highest percentage of successful verifications with respect to other combinations of reference machines.

Keywords

ASTM E23, Charpy reference specimens, C-type Charpy hammer, NIST Charpy Program, NIST Charpy reference machines, U-type Charpy hammer

Introduction

Since 1989, the NIST Charpy Machine Verification Program[1] has supplied reference Charpy specimens to thousands of customers all over the world to be used for indirectly verifying impact machines in compliance with the ASTM E23-18, *Standard Test Methods for Notched Bar Impact Testing of Metallic Materials*, standard.[2] Originally, the production and distribution of certified Charpy reference specimens was undertaken by the US Army, starting in the late 1960s.[3] NIST took over the program in 1989, and the three ASTM E23-18 reference impact machines used to certify specimens were transferred to the NIST Charpy laboratory in Boulder, Colorado. Two of these machines are still currently used in the certification process, whereas the third has been replaced in the mid-2000s with a similar machine having a slightly larger capacity or available impact energy.

Although the Charpy Program was originally set up to primarily support the US industry, the worldwide application of ASTM E23-18 as a test standard progressively expanded the international scope of the program. Between 2014 and 2019, the number of companies served by the program increased from 1,140 to 1,239 and the number of machines verified increased from 1,253 to 1,380. In 2019, more than three quarters (78 %) of the companies served were outside the US, scattered in 59 different countries.

Since the early 1990s, all Charpy test results returned by program customers have been entered into an electronic database, which can be used to investigate general trends in Charpy testing over the last 30 years. Although the earliest records in the database date back to October 1990, it was only in 1993 that all customer test results were routinely entered into the database. Between 1993 and 2019, the number of verification sets (each typically consisting of five absorbed energy values) analyzed and recorded more than doubled between 1993 and 2019, from 1,539 to 3,182. All the test results contained in the NIST database were obtained on impact machines equipped with strikers having an 8-mm edge radius (8-mm strikers).

The focus of this study is to investigate the influence of Charpy machine type, or pendulum hammer configuration, on the results of the verification tests. The overwhelming majority of Charpy machines in the world are of the two following types, based on the shape of the lower part of the swinging hammer:

- C-type (**fig. 1*A***), and
- U-type (**fig. 1*B*** and **1*C***).

A few machines of a third type (Z-type, **fig. 2**) are also featured in the database. Other types of Charpy machines (O, D, P, etc.) have been reported by customers through the years, but in the absence of photographs or drawings, their actual configuration is uncertain. In 2019, C-type and U-type machines combined accounted for 99.5 % of the machines indirectly verified by NIST.

Besides the different shape of the swinging hammer, C-type and U-type machines tend to be structurally different. With reference to the Charpy reference machines shown in **figure 1**, one can see that in the C-type machine (**fig. 1*A***) the pendulum swings between two columns, whereas both U-type machines (**fig. 1*B*** and **1*C***) only have one support column. As a consequence, the structure of a C-type machine tends to be more "rigid," or

FIG. 1

The NIST Charpy reference machines: (*A*) C-type and (*B*), (*C*) U-type.

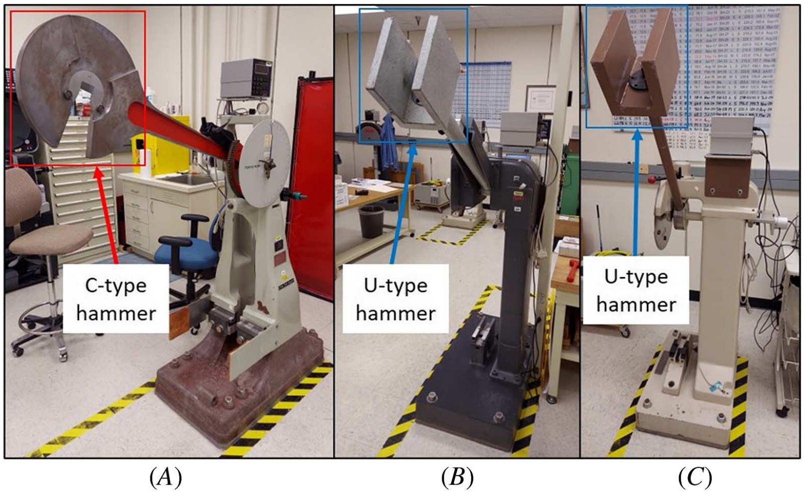


FIG. 2

Z-type NIST Charpy machine.

less compliant. In the case of the NIST reference machines, this was confirmed by compliance measurements,[4] which yielded an approximately 30 % lower compliance for the C-type machine shown in figure 1*A* with respect to both U-type machines shown in figure 1*B*. The compliance of the Z-type machine shown in figure 2 was found to be similar to that of the C-type machine.

Moreover, the "saddle-shaped" configuration of the striking end of a U-type hammer facilitates the occurrence of multiple interactions between the broken specimen halves and the hammer (secondary impacts). In contrast, both C-type and Z-type hammers have no material on either side of the striking bit, and therefore the broken halves of the specimen can fly away from the striker after the fracture event and avoid secondary impacts. However, it has been observed that in C-type machines ductile specimens often drop down and ride on the lower part of the hammer during part of the swing, thus adding some energy (1–2 J) to the recorded values.

Details on the NIST Charpy Program and the AISI 4340 steel used for low-energy and high-energy reference specimens can be found in McCowan et al.[1] Information on the AISI 9310 steel used for super-high energy reference specimens can be found in Lucon.[5] Note that T-200 maraging steel (18Ni) was used for super-high energy specimens until the mid-2000s, whereas side-grooved specimens of AISI 9310 (3Ni) steel is currently being used at this energy level.

Early Research on the Influence of Machine Type on Charpy Absorbed Energies

Through the first half of the 20th century, the Charpy test suffered the negative reputation of being an "unreliable" test because of excessive scatter observed for test results obtained from different machines on the same materials. During the 1950s, however, D. Driscoll from the Watertown Arsenal showed that much of that scatter could be attributed to poor machine maintenance and excessively large tolerances in specimen machining and finishing.[6,7] His investigations facilitated the formulation of tighter procedures for Charpy testing and specimen machining, and ultimately led to the development of the Charpy Machine Verification Program.

Around the same time, J. Bluhm[8] was the first to suggest that the basic design of an impact machine, and in particular its flexibility, could have a significant influence on the results of Charpy testing, in terms of absorbed energy. Specifically, he claimed that a decrease in stiffness of the pendulum could "give rise to fallacious energy absorption values," thus implying that part of the energy recorded with a "flexible" (compliant) machine could be attributed to energy dissipated by vibration of the hammer. His study was limited to specimens with absorbed energies lower than 20 J, thus similar to the NIST low-energy indirect verification specimens.

In the mid-1990s, NIST researchers published a study[9] that systematically compared absorbed energy values obtained from low-energy Charpy specimens (15–17.5 J) tested at −40°C using one C-type and two U-type machines—namely, the ASTM reference machines located in Boulder. Besides absorbed energy, other test variables were considered, such as lateral expansion, height of the shear lips, and features of the striker marks created by the interactions of the specimens with the machine anvils during fracture (first-strike marks) or with other parts of the machine after fracture (second- and third-strike marks).

The most relevant conclusions from the study were the following:

(a) The average value of absorbed energy for the C-type machine was about 9 % lower than the average of the two U-type machines. For an average energy absorption of 17 J, this corresponds to a difference of about 1.5 J.
(b) The striker of the C-type machine (fig. 1*A*) is encased in a massive disk of steel, which tends to reduce longitudinal vibrations, whereas the strikers in the U-type machines (fig. 1*B* and 1*C*) are attached to short and relatively slender cantilever beams. As a consequence, much less dissipation of hammer energy though vibrations could be envisaged for the C-type hammer.
(c) The examination of striker marks did not provide evidence of an effect of machine type on their number, size, or orientation.

A few years later, Manahan and Stonesifer published two articles[10,11] in which they investigated the differences between Charpy absorbed energies measured by means of an optical encoder and using an instrumented striker. Although their investigations focused on the factors affecting the differences between the energies returned by the machine encoder and calculated from the force-displacement curve obtained from the instrumented striker, some of their findings provided important insight on how machine design can affect the measurement of absorbed energy, particularly at low energy level (below 20 J).

By definition, absorbed energy in a Charpy test is given by the work spent by the pendulum to fracture the specimen. The most straightforward approach is to measure absorbed energy as the difference between initial and final (residual) potential energy of the hammer. These quantities can be calculated from the hammer mass and the heights of fall and rise before and after the test, which are provided by a digital encoder that records the fall and rise angles. The actual fracture energy is obtained by subtracting from the calculated energy the following extraneous components:

(1) *Friction and windage losses*. These can be easily measured by swinging the hammer without a specimen in place, and scaling the losses on account of the actual angle of wing.
(2) *Vibrational energy in the pendulum*. These were found to be in the range 0.5 to 1.5 J for NIST certified low-energy specimens, and up to 3 J for ductile specimens (such as high-energy and super-high-energy specimens).[10]
(3) *Energy losses occurring as a result of radial play at the bearings*, induced by hammer vibrations. These losses were quantified to be in the range 0.12 to 0.6 J for super-high-energy specimens, and much smaller for low-energy specimens.[10]
(4) *Secondary impacts* between the broken specimen halves and the swinging pendulum. These interactions become less likely if the broken halves are ejected in a direction opposite to the pendulum swing, as in the case of low-energy specimens tested at −40°C.

For most of these components, the magnitude is sensitive to the Charpy machine design, and specifically to the hammer type (C-type or U-type). Namely, vibrational energy losses increase as the stiffness of the hammer decreases, and are therefore expected to be larger (i.e., close to the upper limit of the 0.5–1.5 J range) for the more compliant U-type machines. Conversely, for a stiffer C-type hammer, one could expect them to be closer to the lower limit (0.5 J).

As far as secondary impacts are concerned, intuitively their number and severity also depend on the machine design. With reference to the NIST reference machines shown in figure 1, it is apparent that the C-type machine has a much more "open" design, which should allow broken halves to fly away more freely after fracture without interacting with the swinging hammer. On the other hand, the U-type hammers have a more "closed" structure, although shrouds are being used to prevent broken specimens from rebounding against the pendulum (fig. 3).

FIG. 3
Shrouds used for the two NIST U-type reference machines.

TABLE 1
Results of low-energy tests performed at room temperature on low-energy specimens

Machine Type	Shrouds?	Absorbed Energy, J		Number of Secondary Impact Marks per Specimen	
		Mean	Standard Deviation	Mean	Standard Deviation
C-type	N/A	15.577	1.4616	15.2	3.27
U-type	YES	16.904	1.2529	13.2	3.11
U-type	YES	17.562	1.7374	10.2	4.38
	NO	17.180	1.6364	20.4	3.58

To quantify the possible effect of secondary impacts on the absorbed energies recorded by the NIST reference machines, low-energy specimens were tested at room temperature (21°C).[*] Five specimens were tested on each of the three machines, and an additional five were tested on one of the U-type machines after removing the shrouds. On each of the broken specimen halves, we counted the number of visible marks/nicks generated by secondary impacts between the halves and various parts of the machines (not just the hammer). The results are provided, in terms of average values and standard deviations, in Table 1.

Data shown in Table 1 do not support the claim that the specimens tested on the C-type machine experience fewer secondary impacts. Indeed, the opposite was observed when comparisons are made with U-type machines with shrouds installed. When shrouds were removed, the number of secondary impact marks significantly increased. Obviously, it's impossible to establish how many marks resulted from interactions with the hammer (which affect absorbed energy) or with other parts of the machine (which do not). Overall, the impression is that secondary impacts have a marginal influence on the differences between the NIST reference machines.

In Table 1, absorbed energy differences between C-type and U-type machines range between 1.3 and 2 J. The study by Manahan and Stonesifer[11] concluded that a 20 % change in striker/hammer stiffness can lead to an absorbed energy difference of about 9 % for low-energy Charpy specimens. The reason for stiffer machines yielding lower fracture energies was given as follows: increasing the stiffness of the striker/hammer system increases the time frequency of the inertial peaks during the elastic portion of the specimen response. This higher frequency causes the critical fracture force to be reached at smaller displacements, therefore giving rise to lower absorbed energies.

As previously mentioned, compliance measurements on the NIST Charpy machines showed an approximately 30 % higher stiffness for the C-type machine with respect to the U-type machines. Using the same simple two mass–two spring dynamic model, the expected difference in absorbed energy would amount to 13.5 % of the

[*]Low-energy specimens tested at room temperature are ejected from the front of the machine (in the direction of pendulum swing), which increases the likelihood of secondary impacts with respect to −40°C tests, where broken halves fly in the opposite direction as a consequence of rebounding against the anvils after fracture under extreme brittleness conditions.

certified value, or 2 J.[11] Therefore, it appears that the energy differences shown in Table 1 could be fully explained in terms of different machine stiffness, affecting vibrational energy losses in low-energy specimens.

The NIST Charpy Program Customer Database

Test results obtained by NIST Charpy Program customers (in terms of absorbed energy) have been regularly entered in a database, which at the time of writing contains almost 70,000 records. Each record corresponds to 3 to 5 absorbed energy values and contains information about the customer company (name, address, contact person information) and the Charpy machine (serial number, manufacturer, hammer type, capacity, typical testing range). The oldest record in the database dates back to 1990, but it was only in 1993 that results were consistently entered into the database by the Charpy Program Coordinator.

In 2019, the Charpy Machine Verification Program served 1,237 companies in 58 countries located all over the world, for a total of 1,380 Charpy machines indirectly verified.

Analyzing the NIST database therefore offers a unique opportunity to examine general trends that occurred in the international Charpy testing community over the last 30 years. In this investigation, records between 1/1/2000 and 9/30/2021 have been analyzed, corresponding to a total of 56,893 sets of results (each set comprising between 3 and 5 Charpy tests).

A couple of general trends that can be noted by the analysis of the NIST database are shown in figures 4 and 5.

Figure 4 illustrates the number of Charpy machines indirectly verified by NIST between 2000 and 2019, which almost doubled over a period of 20 years. This clearly reflects a sharp increase in the number of machines in service around the world. The same figure also demonstrates a steady increase of the average machine capacity from 362 J in 2001 to 430 J in 2020, prompted by the increase in strength and toughness of steel in the last 20 years.

During the same time span, the percentage of C-type machines verified by NIST went from one-third in 2000-2001 to almost half in 2019–2021 (fig. 5). Z-type machines were “invented” in the mid-2000s, and the first

FIG. 4 Number and average capacity of Charpy machines verified by NIST in the period 2020–2021.

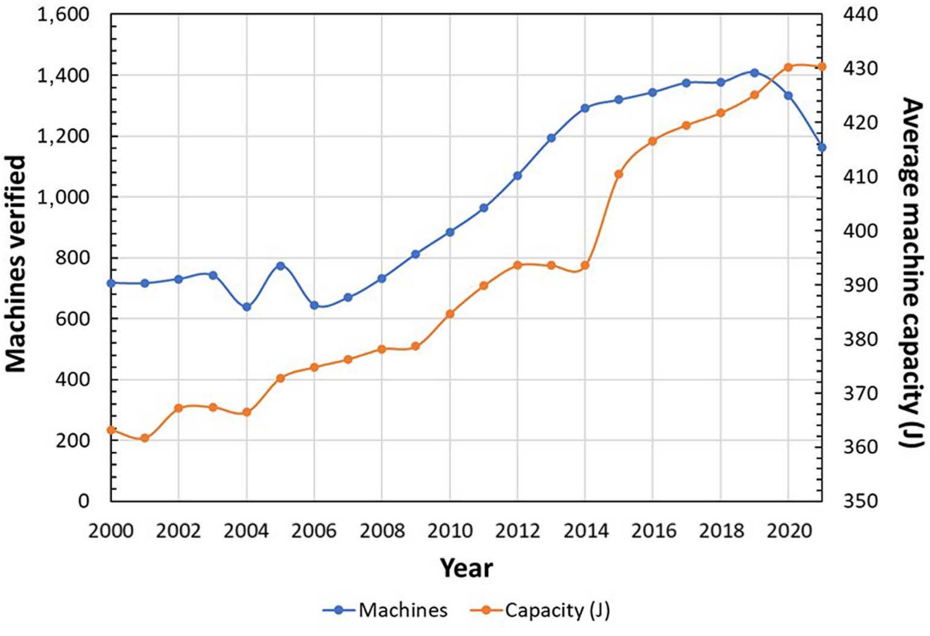
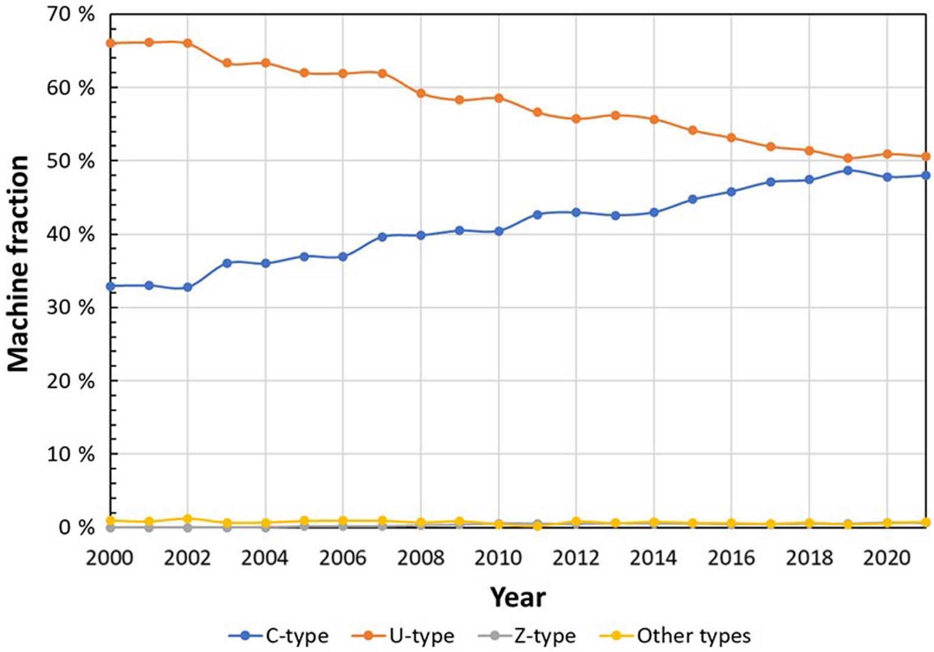


FIG. 5 Percentages of different types of Charpy machines verified by NIST in the period 2020–2021.



one was verified by NIST in 2005. Their share remains well below 1 % (0.5–0.7 % between 2019 and 2021). Other types of machines, as reported by customers, represent about 1 % of the total.

Influence of Machine Type on Charpy Absorbed Energy

CHARPY RESULTS FROM THE NIST REFERENCE MACHINES

When the Charpy Machine Verification Program was initiated by the Watertown Arsenal in the late 1960s, a decision was made to certify specimen lots[*] by conducting tests on three machines, one of C-type and two of U-type. At the time, the ratio 1:2 most probably closely matched the ratio between C-type and U-type Charpy machines in the world. This was still the case between 1989, when NIST took over the program, up to approximately 2002, as shown in figure 5. Eventually, the fraction of C-type machines kept steadily increasing, currently approaching a 1:1 ratio.

Starting in the mid-to-late 1990s, the NIST Charpy Program started recording the average values of absorbed energy obtained from the encoders of each reference machine during the certification of indirect verification lots. This allows investigating, over a period of more than 25 years, the influence of machine type on data obtained from very controlled and homogeneous Charpy specimens tested on accurately maintained machines at three different energy levels (low, high, and super-high).

Data are plotted in figures 6, 7, and 8 for low (112 lots), high (119 lots), and super-high (35 lots)[†] energy specimens, respectively. Each data point in the three figures corresponds to the average absorbed energy of 50 specimens from a specific lot.

[*]A "lot" is a group of Charpy specimens (typically between 1,000 and 2,000) that have been heat treated and machined all at the same time. Between 10 and 15 specimen lots of three energy levels (low, high, and super-high) are certified at NIST every year.

[†]During a period of 10 years (2008–2018), super-high-energy Charpy verification lots were unavailable in the NIST Standard Reference Materials catalog.

FIG. 6 Mean absorbed energy values obtained during the certification of low-energy Charpy specimens at NIST.

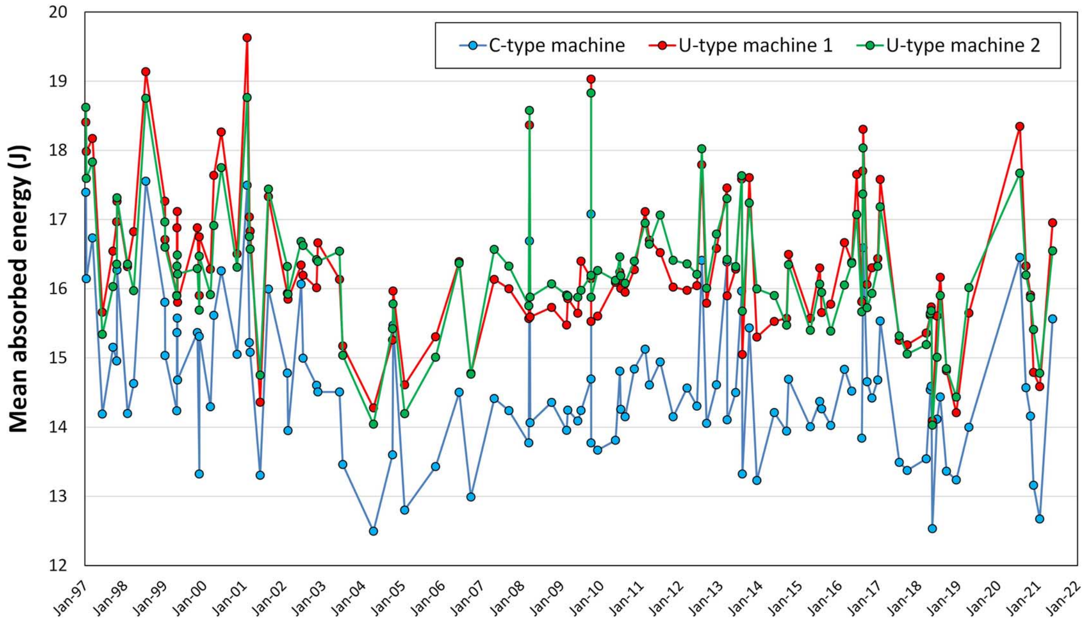
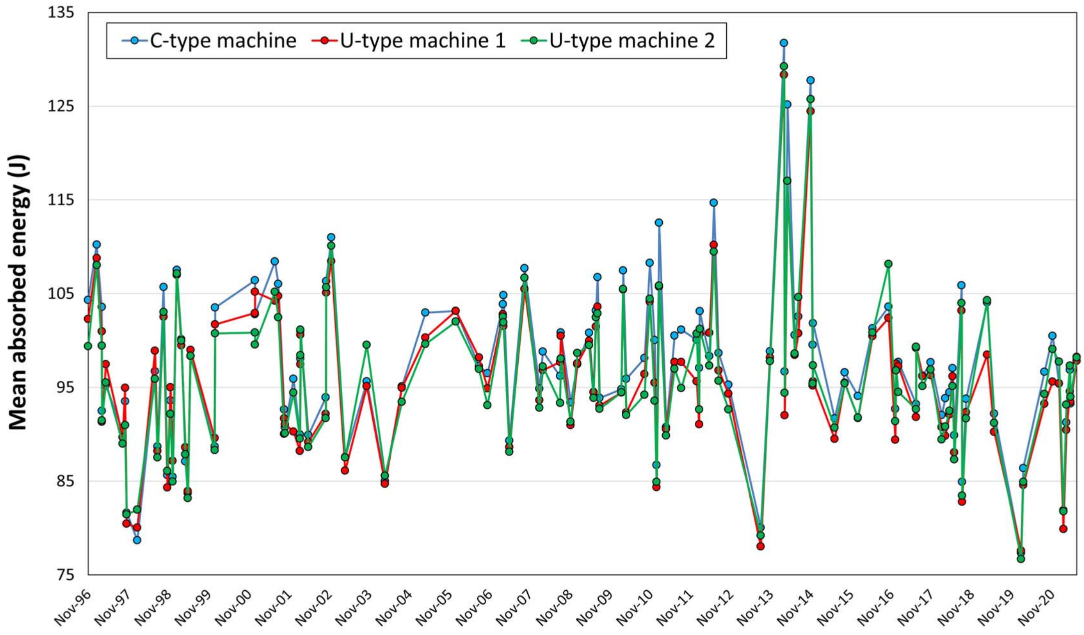
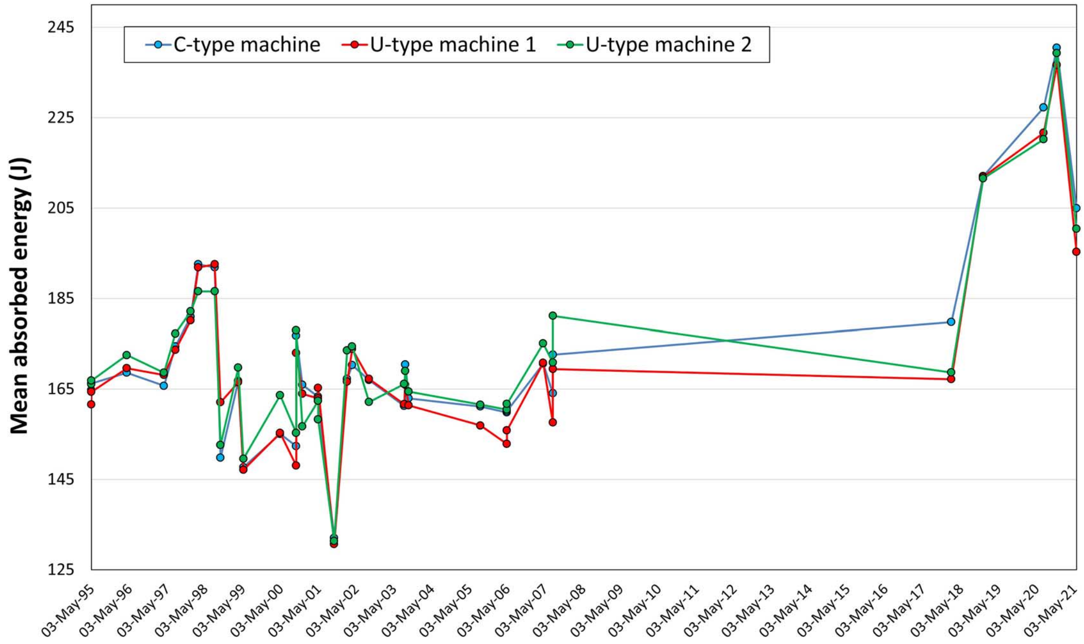


FIG. 7 Mean absorbed energy values obtained during the certification of high-energy Charpy specimens at NIST.



For low-energy specimens (fig. 6), there is a clear and consistent trend for the C-type machine to provide lower energy values than both U-type machines. Considering differences between C-type values and the average between the two U-type values, they range from −0.40 to −2.47 J. The mean difference is −1.64 J.

FIG. 8 Mean absorbed energy values obtained during the certification of super-high-energy Charpy specimens at NIST.



Mean values for the three machines were statistically analyzed using the nonparametric Friedman test,[12] a nonparametric alternative to the Repeated Measures ANOVA that is used to determine whether or not there is a statistically significant difference between the means of three or more groups in which the same subjects (reference machines) show up in each group. Extremely significant probability values $p < 0.0001$ were obtained for both comparisons (C-type versus U-type 1 and C-type versus U-type 2), indicating that the variations among the sets of mean absorbed energy values are significantly greater than expected by chance. Averages from the two U-type machines were not found significantly different by the same test.

As far as high-energy lots (fig. 7) are concerned, trends are less obvious, but a tendency of C-type energies to be highest in most cases can be noticed. Differences (C-type minus mean U-type) range between −2.29 and 8.13 J, with an average difference of 1.57 J. The number of lots for which the C-type average was the lowest among the reference machines is only 16 % (19 out of 119).

Friedman test provided extremely significant probability values $p < 0.0001$ when comparing C-type machine values to either U-type machine values. No difference was found between the two U-type machines.

Finally, data for super-high-energy lots (fig. 8) appear more difficult to interpret. The mean difference between C-type and mean U-type energy averages is 0.64 J, with a minimum of −7.56 J and a maximum of 11.84 J. C-type averages turned out to be higher for 60 % of the lots (21 out of 35).

Friedman test indicates, however, that differences are not significant ($p > 0.05$) between the C-type and either of the U-type machines; surprisingly, the difference between the mean values of the two U-type machines is fairly significant ($p < 0.01$).

In the light of the literature review previously reported, the results presented can be interpreted as follows.

(a) For low-energy lots, the significantly lower vibrational energy of the much stiffer C-type hammer justifies the consistently lower absorbed energy. Furthermore, the fact that low-energy specimens tested at −40°C are ejected backward at high velocity, as well as the more “open” design of the C-type pendulum, make the occurrence of significant secondary impacts between broken specimens and pendulum more unlikely.

(b) For fully ductile specimens (high and super-high-energy), the difference in vibrational behavior between machines is negligible, given the magnitude of the energies involved (100–200 J). Specimens exit the machine forward and, in some cases, do not even break, and therefore there can be prolonged contact between specimens and striker (specimens "riding" on the swinging hammer).[*] This becomes the predominant source of machine differences.

CHARPY RESULTS FROM NIST CUSTOMERS

As the influence of machine type is known to be most significant for energies below 20 J, most of the data analyzed from the NIST database belong to low-energy lots. At high and super-high energies, a more limited number of lots have been considered for the assessment of machine type influence.

Statistical Assessment of Machine Type Influence

For each of the lots investigated, the statistical significance of the recorded differences between C-type and U-type machine results was assessed by means of the following statistical tests:

- If both C-type and U-type test results were normally distributed, the unpaired *t* test[13] was used.
- If either or both sets of test results were not normally distributed, the Mann-Whitney,[14] or rank sum, test[†] was employed.

The normality of test results was evaluated by means of the Anderson-Darling test.[15] Most of the C-type machine results (86 %) turned out to be normally distributed, whereas the same was true only for 18 % of the U-type machine results.

The final judgment on how significant the difference was between C-type and U-type results for a specific lot was based on the calculated two-tailed[‡] probability (p) value, according to the following subjective (but reasonable) guidelines:

- $p < 0.0001$: extremely significant;
- $0.0001 \leq p < 0.001$: very significant;
- $0.001 \leq p < 0.01$: fairly significant;
- $0.01 \leq p < 0.05$: significant; and
- $p \geq 0.05$: not significant.

As a final statistical check, the Wilcoxon matched-pairs signed-ranks test[16] was performed to check whether the median of the differences between C-type and U-type machines differed significantly from zero.

Low-Energy Level

Fifty-six low-energy lots, certified between 2000 and 2021, were selected from the NIST database. Each of the selected lots included more than 200 verification sets[§] tested by customers using both C-type and U-type machines.

The overall ratio between C-type and U-type machines for the selected lots is 43.8 % C-type versus 56.2 % U-type. Other machine types, including Z-type, were not considered.

The average absorbed energies from C-type and U-type machines were 15.64 J and 15.99 J, respectively (difference = −0.35 J). This difference is approximately one-fifth of the average deviation recorded between the NIST C-type and U-type machines (−1.64 J).[**]

[*] On our C-type machine, some ductile/unbroken specimens "rest" on the striker during the hammer swing, and this causes a slight hammer deceleration. On the U-type machines, unbroken specimens fall below the swinging hammer because of gravity and do not experience further interactions with it.

[†] Also called Mann-Whitney-Wilcoxon test.

[‡] In this case, we used a two-tailed probability, because we wanted to allow both positive and negative differences.

[§] Verification sets normally consist of five Charpy test results, with a few cases including just three or four test results.

[**] We feel this is primarily due to a significant difference in compliance between the NIST C-type and U-type machines, which is not always the case for machines used by other companies in the world.

TABLE 2
Results of the statistical analyses performed on low-energy lots

P Value Range	Difference between Medians Is…	Number of Selected Lots	Percentage of Selected Lots
<0.0001	Extremely significant	42	84 %
Between 0.0001 and 0.001	Very significant	7	14 %
Between 0.001 and 0.01	Fairly significant	6	12 %
Between 0.01 and 0.05	Significant	1	2 %
≥0.05	Not significant	0	0 %

For the 56 low-energy lots examined, results are summarized in Table 2. As can be seen, every lot showed a statistically significant difference between C-type and U-type machine results, and for the vast majority (84 %), the difference was found to be extremely significant ($P < 0.0001$).

The calculated two-tailed P value from the Wilcoxon matched-pairs signed-ranks test was less than 0.0001, indicating that the difference between the medians of the average absorbed energies from the two machine types is statistically extremely different.

High-Energy Level

Ten high-energy lots, certified between 2000 and 2021, were selected from the NIST database. Each of the selected lots included more than 300 verification sets.

The overall ratio between C-type and U-type machines for the selected lots is 46.8 % C-type versus 53.7 % U-type. Other machine types, including Z-type, were not considered.

The average absorbed energies from C-type and U-type machines were 98.31 J and 97.77 J, respectively, corresponding to a difference of 0.54 J, which is about one-third of the average deviation recorded between the NIST C-type and U-type machines (1.57 J).

Among the 10 lots selected, C-type results were normally distributed in 6 cases (60 %), whereas U-type results were normally distributed in 7 cases (70 %).

The results of the Wilcoxon tests are provided in Table 3. For 6 of the 10 selected lots, the difference was found statistically significant, with C-type machines providing higher absorbed energy values. The effect of machine type, however, is not as evident as in the case of low energies.

Based on the Wilcoxon matched-pairs signed-ranks test, the median of the differences between C-type and U-type machines is very significantly different from zero ($P = 0.0037$).

Super-High-Energy Level

Ten super-high-energy lots, certified between 2000 and 2021, were selected from the NIST database. Each of the selected lots included more than 195 verifications.

The overall ratio between C-type and U-type machines for the selected lots is 39.2 % C-type versus 60.8 % U-type. Other machine types, including Z-type, were not considered.

TABLE 3
Results of the statistical analyses performed on high-energy lots

P Value Range	Difference between Medians Is…	Number of Selected Lots	Percentage of Selected Lots
<0.0001	Extremely significant	2	20 %
Between 0.0001 and 0.001	Very significant	1	10 %
Between 0.001 and 0.01	Fairly significant	2	20 %
Between 0.01 and 0.05	Significant	1	10 %
≥0.05	Not significant	4	40 %

TABLE 4
Results of the statistical analyses performed on super-high-energy lots

P Value Range	Difference between Medians Is...	Number of Selected Lots	Percentage of Selected Lots
<0.0001	Extremely significant	0	0 %
Between 0.0001 and 0.001	Very significant	0	0 %
Between 0.001 and 0.01	Fairly significant	0	0 %
Between 0.01 and 0.05	Significant	2	20 %
≥0.05	Not significant	8	80 %

The average absorbed energies from C-type and U-type machines were 217.84 and 217.43 J, corresponding to a difference of 0.41 J. This is approximately two-thirds of the average deviation recorded between the NIST C-type and U-type machines (0.64 J).

Among the 10 lots selected, only one C-type data set and one U-type data set were not normally distributed. Therefore, 90 % of all data sets were normally distributed.

The results of the Wilcoxon tests are provided in **Table 4**. For only 2 of the 10 selected lots, the difference was found statistically significant, but with relatively high values ($P > 0.01$). C-type machines tend to yield slightly higher absorbed energy values than U-type machines, but the effect of machine type, however, is less pronounced than for high-energy lots.

Based on the Wilcoxon matched-pairs signed-ranks test, the median of the differences between C-type and U-type machines was found to be not significantly different from zero ($P = 0.1055$).

Machine Type Influence on Pass/Fail Rates for NIST Reference Specimens

In accordance with ASTM E23-18,[2] the indirect verification of a Charpy machine is considered successful if the difference between the reference absorbed energy (KV_{ref}) and the average of the verification tests performed is within the larger of ±1.4 J or ±5 % of KV_{ref}. The absolute limit is used for low-energy specimens ($KV_{ref} < 20$ J), whereas the relative limit is used at high and super-high energy level.

For the 76 NIST Charpy lots considered in this study (56 low-energy, 10 high-energy, and 10 super-high-energy), the calculated pass/fail rates for C-type and U-type machines are shown in **Table 5**.

At every energy level, the pass rate for C-type machines is higher than for U-type machines. In particular, at the low-energy level the fail rate of U-type machines (11.2 %) is almost double than that of C-type machines (5.8 %).

Figure 9 shows that the fail rate for all machine types at low-energy level has been decreasing from 2000 to the present time (2021), which can be attributed to an improving quality and maintenance level for Charpy machines around the world.

We also looked at the ratio between customers failing "high" (i.e., with an average absorbed energy higher than the certified value plus the larger of 1.4 J or 5 % of KV_{ref}) and failing "low" (i.e., with an average absorbed

TABLE 5
Pass and fail rates for selected low, high, and super-high energy NIST verification lots

Energy Level	Lots Examined	Machine Type	Pass Rate, %	Fail Rate, %
Low	56	C	94.2	5.8
		U	88.8	11.2
High	10	C	94.6	5.4
		U	93.7	6.3
Super-high	10	C	92.3	7.7
		U	90.5	9.5

FIG. 9 Customers' fail rate at low-energy level between 2000 and 2021.

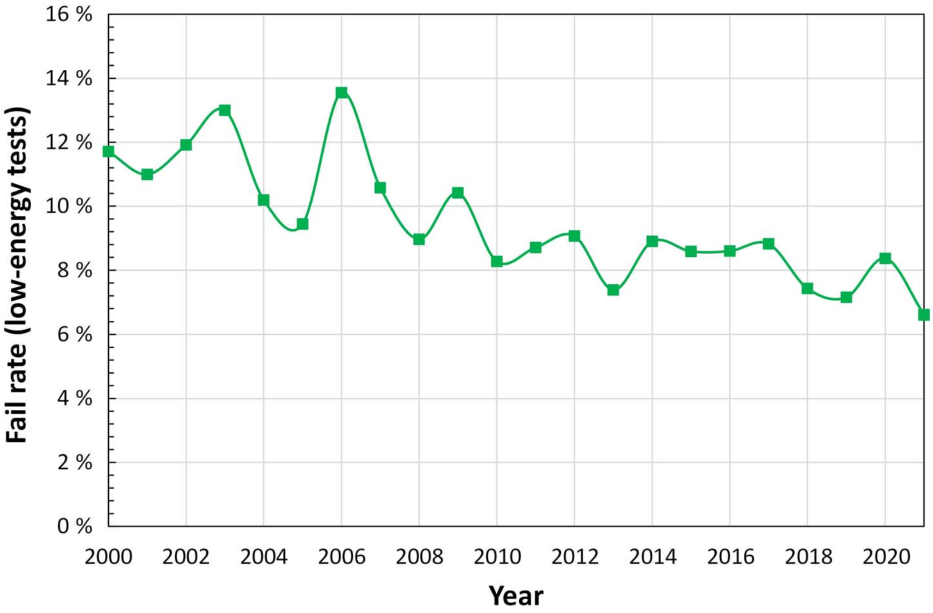


TABLE 6

"Low" and "high" failures in 10 years of machine verifications at NIST (2010–2019) as a function of machine type

			"Low" Failures		"High" Failures	
Machine Type	Energy Level	Number of Failed Verifications	#	%	#	%
C	Low	317	60	19 %	257	81 %
	High	402	63	16 %	339	84 %
	Super-high	4	2	59 %	2	50 %
	Total	723	125	17 %	598	83 %
U	Low	802	52	6 %	750	94 %
	High	478	107	22 %	371	78 %
	Super-high	16	4	25 %	12	75 %
	Total	1,296	163	13 %	1,133	87 %

energy lower than the certified value minus the larger of 1.4 J or 5 % of KV_{ref}). The results for 10 years of machine verifications (2010–2019) are compiled in **Table 6**.

Data presented in **Table 6** show that machines are much more likely to fail "high" than "low," by more than a factor 5, irrespective of the hammer design or the energy level. This can be attributed to the fact that most of the failure causes that can be detected by visual examination of the broken specimens tend to artificially increase absorbed energy, such as worn or damaged anvils, off-center specimen or striker, or bent pendulum.[17*] This tendency appears slightly more pronounced for U-type machines, arguably as a result of these being, on average, "older" than C-type machines.

*All these issues can be detected through the visual examination of the marks/imprints left by the anvils on the broken specimens.[17]

NIST Charpy Machine Verification Program: What If...?

Since the onset of the Charpy Machine Verification Program by the US Army, the pool of reference machines used to certify reference specimens has always consisted of one C-type machine and two U-type machines. The only change to the original line-up of reference machines took place in the early 2000s, when one of the U-type machines was replaced by another machine of the same type and manufacturer, but with higher capacity.

This 1:2 ratio between C-type and U-type machines closely matched the proportion among customers' machines at the beginning of the 21st century, although the ratio is now approaching 1:1 as shown in **figure 5**.

In this part of the study, we decided to assess the changes in customers' fail rates at the various energy levels and for the two machine types, under the following three hypothetical scenarios/machine line-ups (scenario 0 corresponds to the current NIST Program configuration, one C-type and two U-type machines):

- Scenario 1: one C-type machine.
- Scenario 2: two U-type machines.
- Scenario 3: one C-type machine and one U-type machine (average between existing U-type machines).

For the different scenarios, pass and fail rates were calculated after redefining KV_{ref} for each lot using only certification results from the C-type machine (scenario 1), the two U-type machines (scenario 2), or the C-type

TABLE 7
Fail rates for selected low-energy verification specimens (56 lots) under different scenarios

	All Machines			C-Type Machines			U-Type Machines		
Scenario	Fail Rate	Fail Low	Fail High	Fail Rate	Fail Low	Fail High	Fail Rate	Fail Low	Fail High
0	10 %	11 %	89 %	7 %	21 %	79 %	13 %	7 %	93 %
1	60 %	0.2 %	99.8 %	45 %	0.2 %	99.8 %	70 %	0.2 %	99.8 %
2	11 %	53 %	47 %	12 %	75 %	25 %	10 %	37 %	63 %
3	20 %	3 %	97 %	13 %	6 %	94 %	25 %	2 %	98 %

TABLE 8
Fail rates for selected high-energy verification specimens (10 lots) under different scenarios

	All Machines			C-Type Machines			U-Type Machines		
Scenario	Fail Rate	Fail Low	Fail High	Fail Rate	Fail Low	Fail High	Fail Rate	Fail Low	Fail High
0	6 %	24 %	76 %	5 %	21 %	79 %	6 %	25 %	75 %
1	16 %	40 %	60 %	15 %	39 %	61 %	17 %	39 %	61 %
2	13 %	5 %	95 %	13 %	5 %	95 %	12 %	5 %	95 %
3	9 %	26 %	74 %	8 %	23 %	77 %	9 %	28 %	72 %

TABLE 9
Fail rates for selected super-high-energy verification specimens (10 lots) under different scenarios

	All Machines			C-Type Machines			U-Type Machines		
Scenario	Fail Rate	Fail Low	Fail High	Fail Rate	Fail Low	Fail High	Fail Rate	Fail Low	Fail High
0	9 %	55 %	45 %	8 %	46 %	54 %	10 %	59 %	41 %
1	12 %	55 %	45 %	10 %	47 %	53 %	13 %	59 %	41 %
2	11 %	53 %	47 %	10 %	46 %	54 %	12 %	56 %	44 %
3	10 %	54 %	46 %	8 %	43 %	57 %	11 %	59 %	41 %

FIG. 10 Fail rates of Charpy machines (all types) at different energy levels for different configurations of the NIST Charpy Machine Verification Program (scenarios).

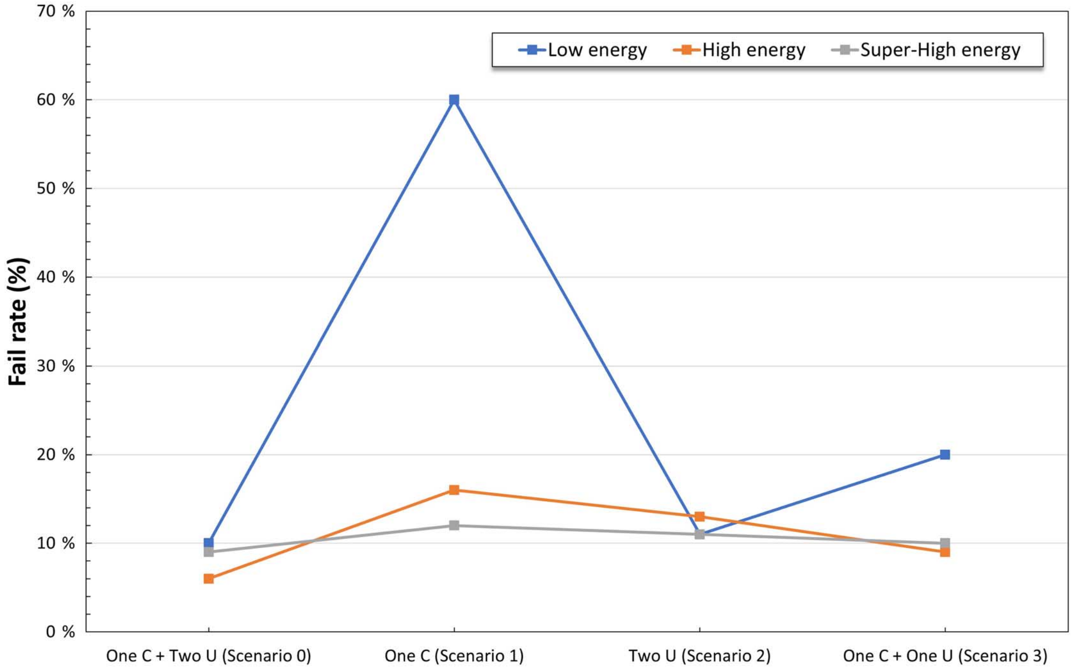


FIG. 11 Fail rates of C-type Charpy machines at different energy levels for different configurations of the NIST Charpy Machine Verification Program (scenarios).

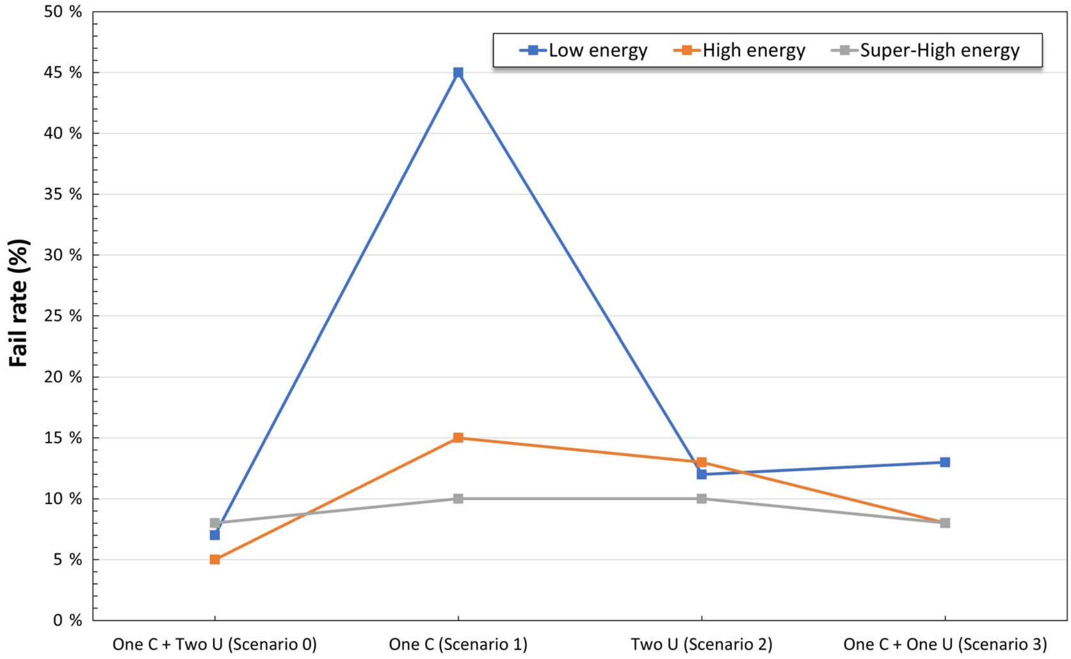
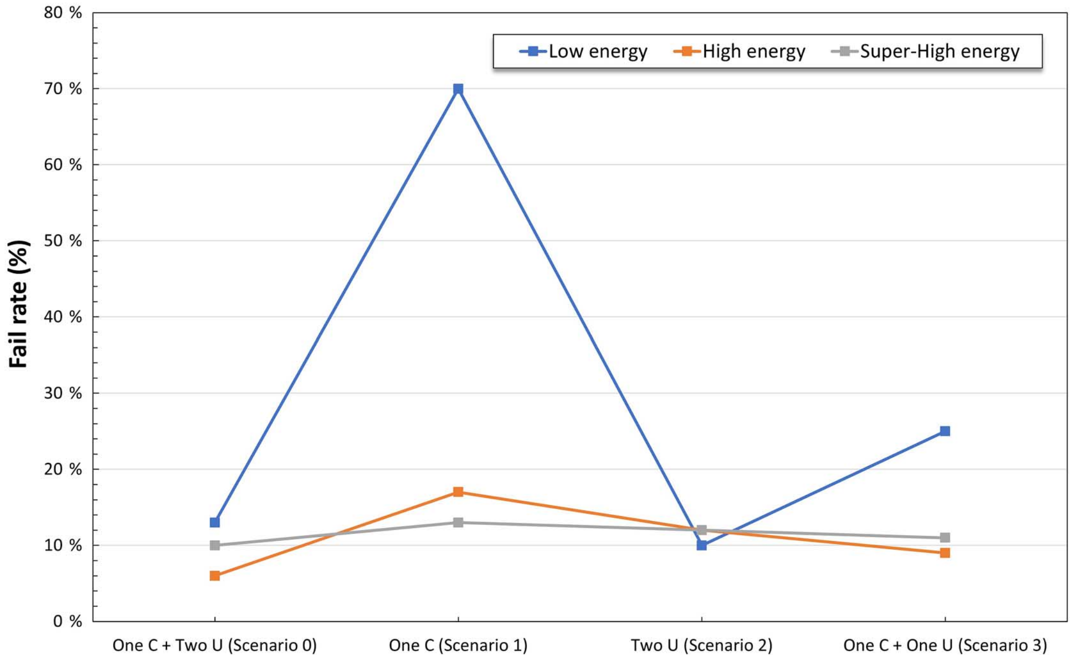


FIG. 12 Fail rates of C-type Charpy machines at different energy levels for different configurations of the NIST Charpy Machine Verification Program (scenarios).



machine and the average of the two U-type machines (scenario 3). The recalculated fail rates obtained for the selected specimen lots are summarized in Table 7 (low-energy), Table 8 (high-energy), and Table 9 (super-high energy).

The fail rates shown in the three tables above are also illustrated in figure 10 (all machines), figure 11 (C-type machines), and figure 12 (U-type machines).

Scenario 1 (one C-type machine) is by far the most unfavorable, with more than half of the verified machines (60 %) failing at the low energy level. Under this scenario, basically every machine (99.8 %) would fail high.

The remaining two alternative scenarios would only increase the fail rate by a few percent, with the exception of scenario 3 (one C + one U) at the low-energy level, where the fail rate would increase significantly.

The observed trends of the fail rates support the historical and current machine line-up of the NIST Charpy Program (one C-type and two U-type machines), in that it corresponds to the highest customer pass rates at all energy levels.

Conclusions

The availability of 30 years of Charpy test results produced by international customers of the NIST Charpy Machine Verification Program, recorded as absorbed energy values in the NIST customers' database, allowed us to investigate the effect of machine type or hammer configuration (C-type versus U-type) on Charpy test results, and compare general trends with the behavior of NIST reference machines used for the certification of indirect verification reference specimens (one C-type machine and two U-type machines).

Using appropriate statistical tests, we analyzed the differences between absorbed energies measured with the two machine types at three energy levels: low (<20 J), high (~100 J), and super-high (~200 J). These differences were correlated with several factors that were identified in previous literature studies, such as machine rigidity/compliance and corresponding vibrational behavior, and secondary impacts between broken samples and swinging hammer.

At the low energy level, C-type hammers absorb less energy than U-type hammers because of their higher stiffness and more "open" design. Differences were found to be statistically extremely significant, both for the NIST reference machines and the results contained in the NIST database. At high energy, more energy tends to be absorbed by C-type hammers, and although differences are less evident, they were also found statistically significant. Around 200 J (super-high energy level), there seems to be a slight tendency for U-type machines to absorb less energy, but deviations are not statistically significant.

As far as the NIST reference machines are concerned, differences between C-type and U-type results are more pronounced that those shown by the analysis of the database record, because of the specific design of the NIST C-type machine, but the general trends are confirmed (C-type < U-type at low energies, C-type > U-type at high and super-high energies).

Finally, we considered several hypothetical scenarios, where the line-up of reference machines at NIST would be different than the actual one (one C-type and two U-type), and found that the current line-up corresponds to the highest pass rates (percentage of successfully verified machines) at all energy levels.

References

1. C. N. McCowan, T. A. Siewert, and D. P. Vigliotti, "The NIST Charpy V-Notch Verification Program: Overview and Operating Procedures," in *Charpy Verification Program: Reports Covering 1989-2002, NIST Technical Note 1500-9* (Boulder, CO: National Institute of Standards and Technology, 2003).
2. *Standard Test Methods for Notched Bar Impact Testing of Metallic Materials*, ASTM E23-18 (West Conshohocken, PA: ASTM International, approved June 1, 2018), http://dx.doi.org/10.1520/E0023-18
3. *Mechanical Testing—The Charpy V-Notch Impact Test, AMXMR-P-702-104* (Watertown, MA: Army Materials and Mechanics Research Center, 1969).
4. E. Lucon, *Determination of the Compliance of NIST Charpy Impact Machines, NISTIR 8043* (Boulder, CO: National Institute of Standards and Technology, 2015), http://dx.doi.org/10.6028/NIST.IR.8043
5. E. Lucon, "Influence of Shear Lip Symmetry on the Fracture Behavior of Charpy Specimens," *Journal of Testing and Evaluation* 47, no. 2 (January 2019): 1129–1146, https://doi.org/10.1520/JTE20180403
6. D. E. Driscoll, "The Charpy Impact Machine and Procedure for Inspection and Testing Charpy 'V' Notch Impact Specimens," *ASTM Bulletin* 191 (July 1953): 60–64.
7. D. E. Driscoll, "Reproducibility of the Charpy Impact Test," in *Symposium on Impact Testing*, ed. F. G. Tatnall (West Conshohocken, PA: ASTM International, 1956), 70–74, https://doi.org/10.1520/STP47578S
8. J. I. Bluhm, "The Influence of Pendulum Flexibilities on Impact Energy Measurements," in *Symposium on Impact Testing*, ed. F. G. Tatnall (West Conshohocken, PA: ASTM International, 1956), 84–92, https://doi.org/10.1520/STP47580S
9. A. K. Schmieder, P. T. Purtscher, and D. P. Vigliotti, "The Role of Strike Marks on the Reproducibility of Charpy Impact Test Results," in *Pendulum Impact Machines: Procedures and Specimens for Verification*, ed. T. A. Siewert and A. K. Schmieder (West Conshohocken, PA: ASTM International, 1995), 3–18, https://doi.org/10.1520/STP14653S
10. M. P. Manahan Sr. and R. B. Stonesifer, "The Difference between Total Absorbed Energy Measured Using an Instrumented Striker and That Obtained Using an Optical Encoder," in *Pendulum Impact Testing: A Century of Progress*, ed. T. A. Siewert and M. P. Manahan Sr. (West Conshohocken, PA: ASTM International, 2000), 181–197, https://doi.org/10.1520/STP14394S
11. M. P. Manahan Sr., R. B. Stonesifer, T. A. Siewert, C. N. McCowan, and D. P. Vigliotti, "Observations on Differences between the Energy Determined Using an Instrumented Striker and Dial/Encoder Energy," in *From Charpy to Present Impact Testing*, ed. D. François and A. Pineau (Amsterdam, the Netherlands: Elsevier, 2002), 229–236.
12. M. Friedman, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," *Journal of the American Statistical Association* 32, no. 200 (December 1937): 675–701, http://dx.doi.org/10.1080/01621459.1937.10503522
13. G. W. Snedecor and W. G. Cochran, *Statistical Methods*, 8th ed. (Ames, IA: Iowa State University Press, 1989).
14. H. B. Mann and D. R. Whitney, "On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other," *The Annals of Mathematical Statistics* 18, no. 1 (March 1947): 50–60, http://dx.doi.org/10.1214/aoms/1177730491
15. M. A. Stephens, "EDF Statistics for Goodness of Fit and Some Comparisons," *Journal of the American Statistical Association* 69, no. 347 (1974): 730–737, http://dx.doi.org/10.1080/01621459.1974.10480196
16. W. J. Conover, *Practical Nonparametric Statistics*, 3rd ed. (Hoboken, NJ: John Wiley & Sons, 1999).
17. D. P. Vigliotti, T. A. Siewert, and C. N. McCowan, *Installing, Maintaining, and Verifying Your Charpy Impact Machine, NIST Recommended Practice Guide, Special Publication 960-4* (Boulder, CO: National Institute of Standards and Technology, 2000).