Proceedings of PVP2008
2008 ASME Pressure Vessels and Piping Division Conference
July 27-31, 2008, Chicago, Illinois, USA

PVP2008-61565

UNCERTAINTY ESTIMATE OF CHARPY DATA USING A 5-FACTOR 8-RUN DESIGN OF EXPERIMENTS (*)

Charles G. Interrante
Consultant
5333 Pooks Hill Road
Bethesda, MD 20814 USA
charles.interrante@gmail.com

Jeffrey T. Fong
Mathematical & Computational Sciences Division
National Institute of Standards & Tech. (NIST)
Gaithersburg, MD 20899-8910 USA
fong@nist.gov

James J. Filliben
Statistical Engineering Division
National Institute of Standards & Technology (**NIST**)
Gaithersburg, MD 20899-8980 USA
filliben@nist.gov

N. Alan Heckert
Statistical Engineering Division
National Institute of Standards & Technology (NIST)
Gaithersburg, MD 20899-8980 USA
alan.heckert@nist.gov

ABSTRACT

Scatter in laboratory data with duplicates on Charpy impact tests is analyzed by identifying several sources of variability such as temperature, manganese sulfide, initial strain, mis-orientation, and notch radius in order to estimate the predictive 95 % confidence intervals of the mean energy of absorption for each specific test temperature. Using a combination of real and virtual data on a high-strength pressure vessel grade steel (ASTM A517) over a range of temperatures from -40 °C (-40 °F) to 182 °C (360 °F), and the concept of a statistical design of experiments, we present an uncertainty estimation methodology using a public-domain statistical analysis software named DATAPLOT. A numerical example for estimating the mean, standard deviation, and predictive intervals of the Charpy energy at 48.9 °C (120 °F) is included. To illustrate the application potential of this methodology, we enhance it with formulas of error propagation to estimate the mean, standard deviation, and predictive intervals of the associated static crack initiation toughness, K_{Ic}. A discussion of the significance and limitations of the proposed methodology, and a concluding remark are given at the end of this paper.

Keywords: Applied mechanics; Charpy energy; Charpy V-notch impact test; design of experiments; engineering safety;error propagation; fracture mechanics; fracture toughness; impact resistance; mathematical modeling; materials science; mechanical properties; mechanical testing; metallurgical engineering; pressure vessels and piping; pressure vessel steels; sensitivity analysis; statistical data analysis; structural integrity; uncertainty analysis; virtual experiments.

Disclaimer: The views expressed in this paper are strictly those of the authors and do not necessarily reflect those of their affiliated institutions. The mention of names of all commercial vendors and their products is intended to illustrate the capabilities of existing products, and should not be construed as endorsement by the authors or their affiliated institutions.

(*) Contribution of the U.S. National Institute of Standards and Technology (NIST). Not subject to copyright in the U.S.

1. INTRODUCTION

In 1983, the late Dr. Spencer Bush of Pacific Northwest National Laboratory, Richland, WA, closed a keynote address at a technical session on "Flaw detection and sizing" of the International Symposium on Reliability of Reactor Pressure Components, IAEA, Stuttgart, Germany, March 21-25, with the following observation and warning [1]:

> *"There is a major problem when the NDE (nondestructive evaluation) operator does not interface with the ultimate user of his data. For example, the practitioner of* ***fracture mechanics*** *should have a close rapport and a synergism with the NDE expert so that the needs of the one are met by the test results of the other. It has been my personal experience that such rapport rarely occurs."*

In this paper, we would like to echo his sentiment by adding that the practitioner of fracture mechanics should also have a close rapport and a synergism with a statistician so that the data needs of the one are met by the modeling, analysis, computational, and predictive capability of the other.

To illustrate the need for such rapport with the statistician, we note from four examples below, that the experimentalists plots their data without reference to scatter or replicates. In Fig. 1, we show two sets of Charpy V-notch impact energy data for one of three test cylinders by Cheverton, et al [2], where an obvious scatter of data exists. In Fig. 2, we show a plot due to Loss, Gray, and Hawthorne [3] on fracture toughness, K_{Ic}, which is related to the Charpy energy. It shows upper and lower bounds on the K_{Ic} data, but no error bar on the data was included. In Fig. 3, Cheverton, et al [2] presents a summary plot of data for two fracture toughnesses, K_{Ic} and K_{Ia}, as well as ASME Section XI [4] lower bounds, but the data have no error bars. Another example existed in the literature, as shown in Figs. 4 and 5 due to deWit, Fields, and Irwin [5, 6], where the phenomenon of data scatter was obvious, but no expression of uncertainty was ever reported. The purpose of this paper is two-fold:

(1) To introduce an uncertainty estimation methodology using a small set of data and the concept of design of experiments. This serves the purposes of suggesting additional experiments to decrease the scatter and to provide error bars for data already acquired.

(2) To illustrate the methodology with an application in estimating fracture toughness using data of Charpy energy and yield strength at comparable temperatures. The data consists of two types: (a) Real -- results of physical experiments. (b) Virtual -- values conceived from experience and judgment of the experimentalist.

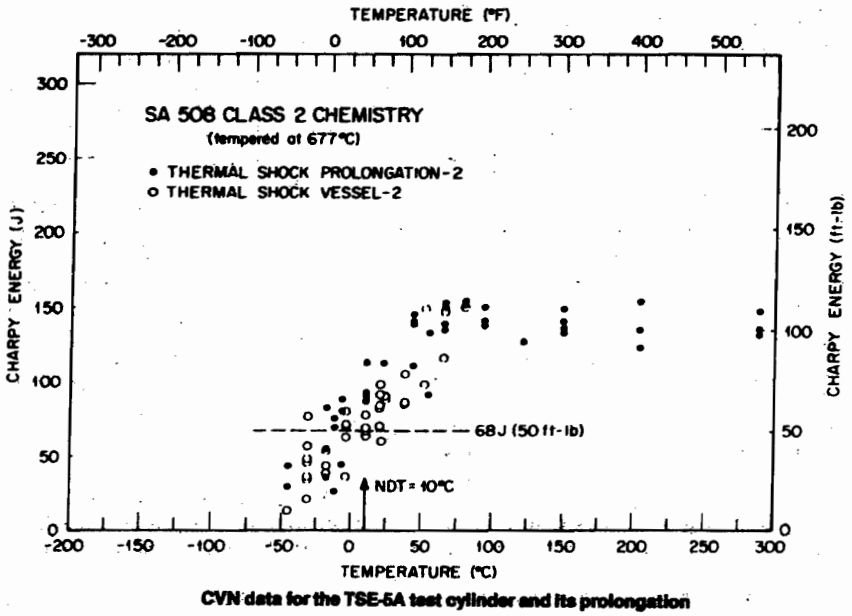

Fig. 1. Charpy V-Notch data for a thermal shock experiment (TSE) test cylinder TSE-5A for two batches of specimens: (1) Hollow circle were test data of materials taken from TSE-5A after it sustained thermal shock. (2) Solid circles were pretest data of materials from a prolongation of the same SA 508 Class 2 chemistry steel forging from which TSE-5A was made by tempering at 677 °C for 4 hours. The room temperature yield strength of TSE-5A cylinder was somewhere between 500 MPa (72.5 ksi) and 700 MPa (101.5 ksi) as estimated for steel forgings tempered at a temperature of under 594 °C and at 704 °C, respectively. Note the significant scatter of data in both batches over a broad range of temperatures from -50 °C to 300 °C. (after Cheverton et al [2]).

*1. **Introduction** (Continued)*

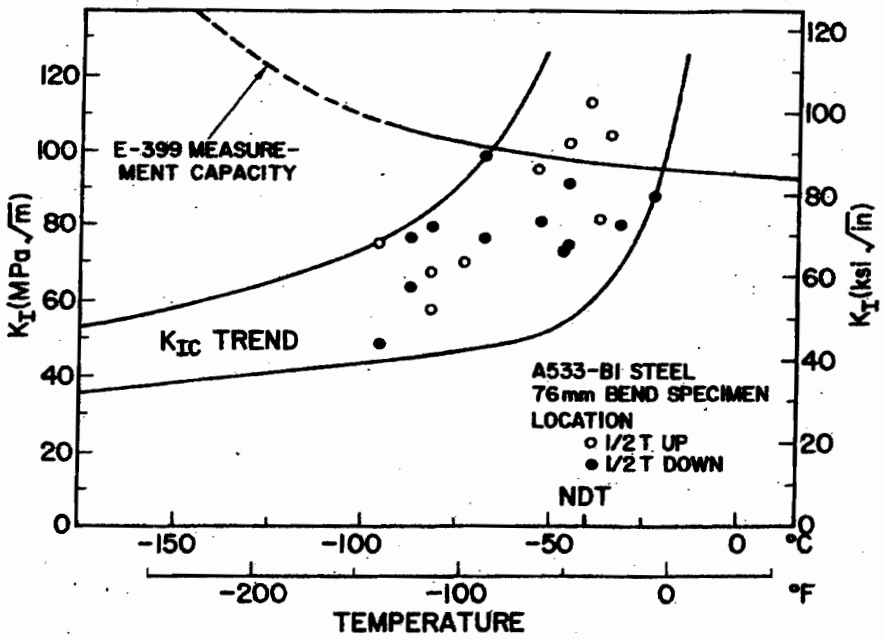

Fig. 2. Static crack initiation toughness (K_{Ic}) data for a material characterization test program in an investigation to validate the benefits of warm prestress in limiting crack extension in a nuclear reactor vessel wall during a simulated loss-of-coolant accident followed by operation of the emergency core-cooling system [3]. Note the significant scatters within and between specimens when cut across plate thickness.

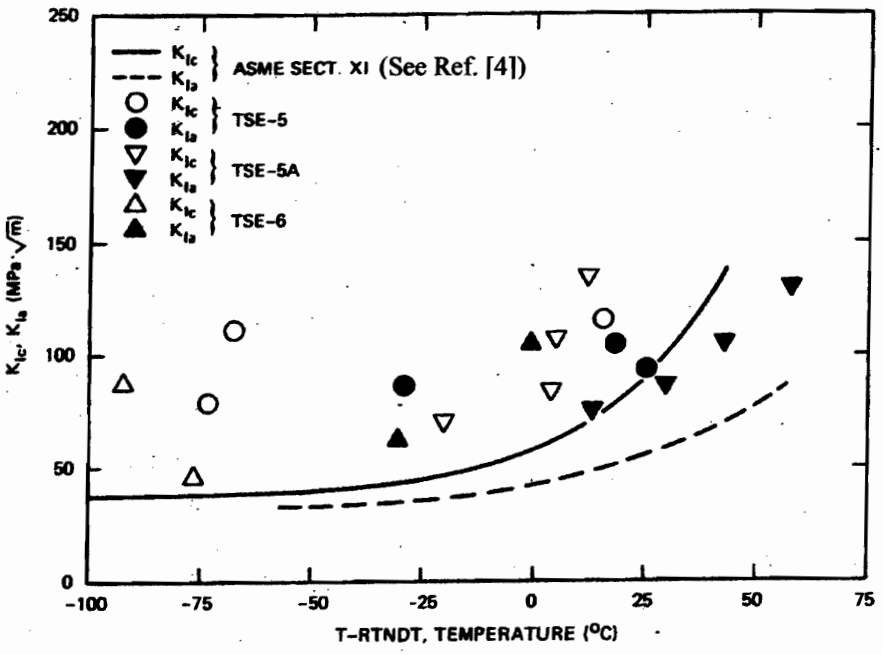

Fig. 3. Comparison of static crack initiation toughness (K_{Ic}) and static crack arrest toughness (K_{Ia}) data deduced from three thermal shock experiment (TSE) test cylinders, TSE-5, 5A, and 6, and normalized with nil-ductility temperature shift, T - RTNDT , with ASME Section XI K_{Ic} and K_{Ia} curves (after Cheverton, et al. [2]). Note the significant scatters within and between specimens of all three test cylinders.

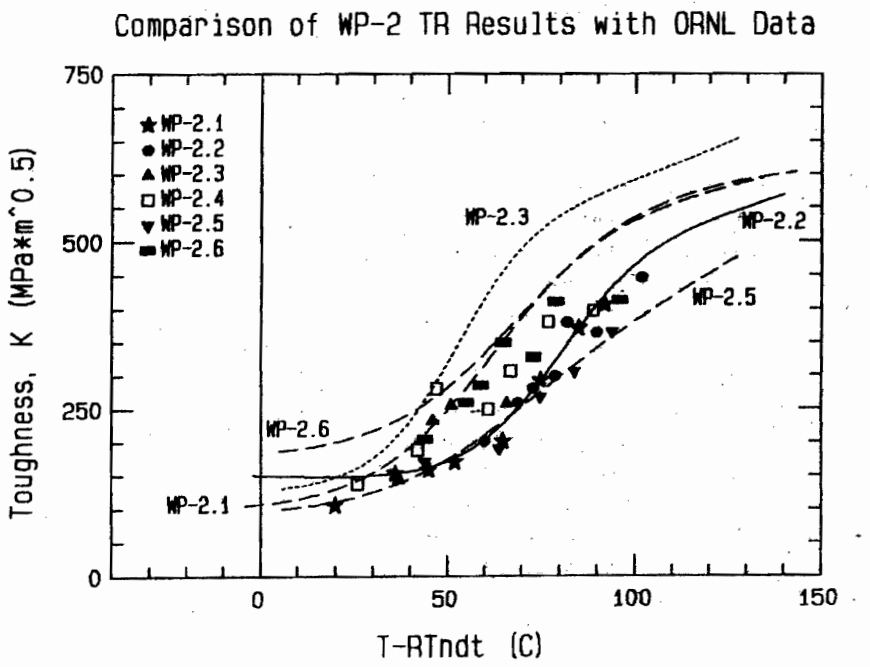

Fig. 4. A summary comparison of static crack arrest toughness (K_{Ia}) data (curves) estimated from thickness reduction (TR) measurements and a smooth *K*-from-TR technique [5, 6] for six specimens (WP-2.1 through 2.6) fabricated from a low upper-shelf base material, with Oak Ridge National Laboratory (ORNL) two-dimensional finite element analyses using ADINA/VPF dynamic crack analysis code (points). Note the significant between-specimen scatters in both experiments (curves) and analyses (points) over a broad range of temperature shifts.

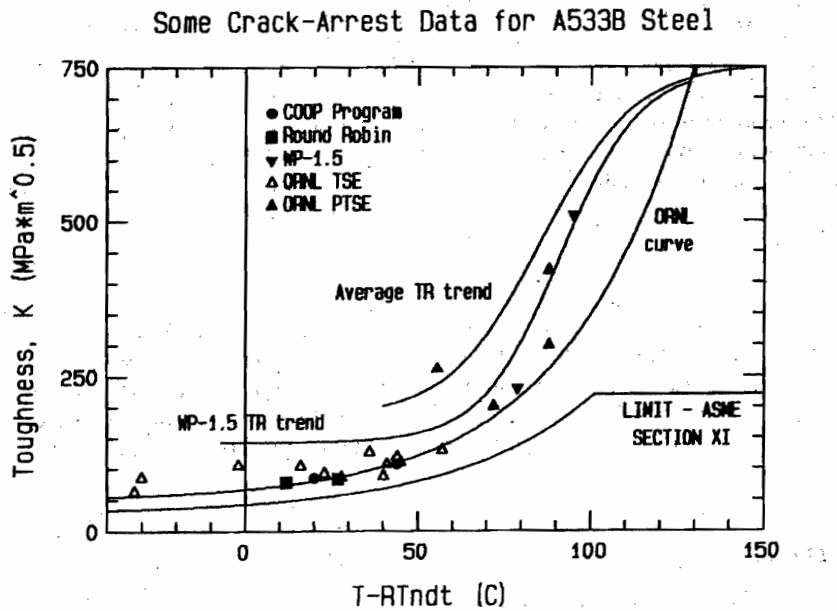

Fig. 5. Comparison of static crack arrest toughness (K_{Ia}) data vs. nil-ductility temperature shift, T - RTNDT , for A533 Grade B Class 1 steel from five different sources, with ASME Section XI K_{Ia} curves (after deWit, Fields, and Irwin [6] and ASME [4]). Note the significant scatter between specimens of different sources over a broad range of temperature shifts from 0 °C to 100 °C.

2. STATEMENT OF THE PROBLEM

Long before the arrival of fracture mechanics in the 1970s as the appropriate tool to characterize the toughness of a material to resist sudden fracture due to crack initiation and propagation, engineers have relied on several simple and inexpensive tests, one of which is the well-known Charpy V-notch impact test.

Following ASTM [7] and books on fracture mechanics such as Kanninen and Popelar [8], Dowling [9], etc., we show in Fig. 6 the specimen geometry and loading configuration for such a test, where a notched three-point bend specimen is impacted with a pendulum. The fracture energy is equated to the energy lost by the pendulum during the impact, and is known as the energy absorption, or Charpy V-notch energy (*CVN*).

In Figs. 7 and 8, we show typical family of plots [10] of *CVN* vs. test temperature for several heats of the steels of ASTM A514/517 [11], having a minimum room temperature tensile yield strength of 620 MPa (90 ksi). Note the S-shaped curve of each plot with a lower-shelf for low temperatures and an upper shelf for high temperatures.

The concept of the so-called nil-ductility temperature (*NDT*) is defined as the upper limit of temperatures in the lower shelf or plateau where the fracture is brittle, i.e., the fracture surface is flat (cleavage) with little or no shear lips, and the so-called plane-strain fracture under impact exists. The upper shelf corresponds to the so-called ductile fracture regime, and the sharp rise between the two plateaus is the brittle-ductile transition. It was the common practice before 1970s to specify the minimum toughness of a material by prescribing a minimum Charpy energy at a given service temperature. As an added measure of safety, the *NDT* was required to be below the lowest anticipated service temperature so that load-bearing material is operating above the *NDT*.

As documented by ASM [12] and reproduced below for some well-known formulas, many correlations between the static K_{Ic}, or the dynamic K_{Id}, with *CVN* have been reported [13-17], where either the Young's modulus, E, or the yield strength, σ_y, is included as an additional parameter:

Rolfe-Novak — $\sigma_y \approx 100$ ksi (Ref. [13], 1970)

$$\left(\frac{K_{Ic}}{\sigma_y}\right)^2 = 5\,(\mathrm{CVN}/\sigma_y - 0.05)$$

K_{Ic} = ksi$\sqrt{\text{in.}}$
CVN = ft · lb
σ_y = ksi

Sailors-Corten (Ref. [14], 1972)

$$\frac{K_{Ic}^2}{E} = 8\,(\mathrm{CVN})$$

K_{Ic}, K_{Id} = psi$\sqrt{\text{in}}$
E = psi
CVN = ft · lb

Wullaert-Server (Ref. [16], 1978)

$$K_{Ic,d} = 2.1\,(\sigma_y\,\mathrm{CVN})^{1/2}$$

$K_{Ic,d}$ = ksi$\sqrt{\text{in.}}$
CVN = ft · lb
σ_y = ksi corresponding to approximate loading rate

Barsom-Rolfe (Ref. [17], 1979)

$$\frac{K_{Ic}^2}{E} = 2\,(\mathrm{CVN})^{3/2}$$

K_{Ic}, K_{Id} = psi$\sqrt{\text{in}}$
E = psi
CVN = ft · lb

An example of the data scatter problem is given in Fig. 7, where one observes that for a specific temperature, say, 120 °F (48.9 °C), the heat-to-heat variation of *CVN* between 10 and 80 ft-lb (13.6 and 108 N-m) was too large. Yet the statistical fit for one such heat in Figure 8 failed to take that into account.

The problem to solve in this paper is to estimate the uncertainty of *CVN* for ASTM A517 steel at 120 °F (48.9 °C), and to use that result to estimate the uncertainty of fracture toughness, K_{Ic}, based on one of the above correlation formulas.

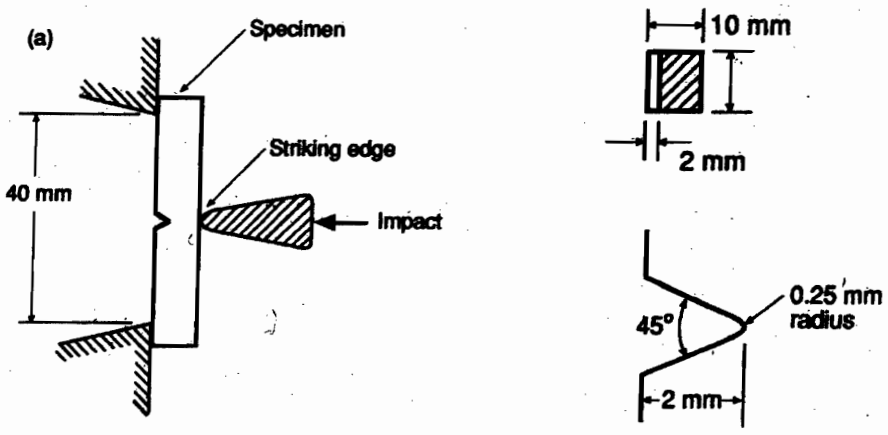

Fig. 6. Specimen and loading configuration for Charpy V-notch test (after Dowling [9, p. 186] and ASTM test standard E23-98 [7]).

2. ***Problem*** *(Continued)*

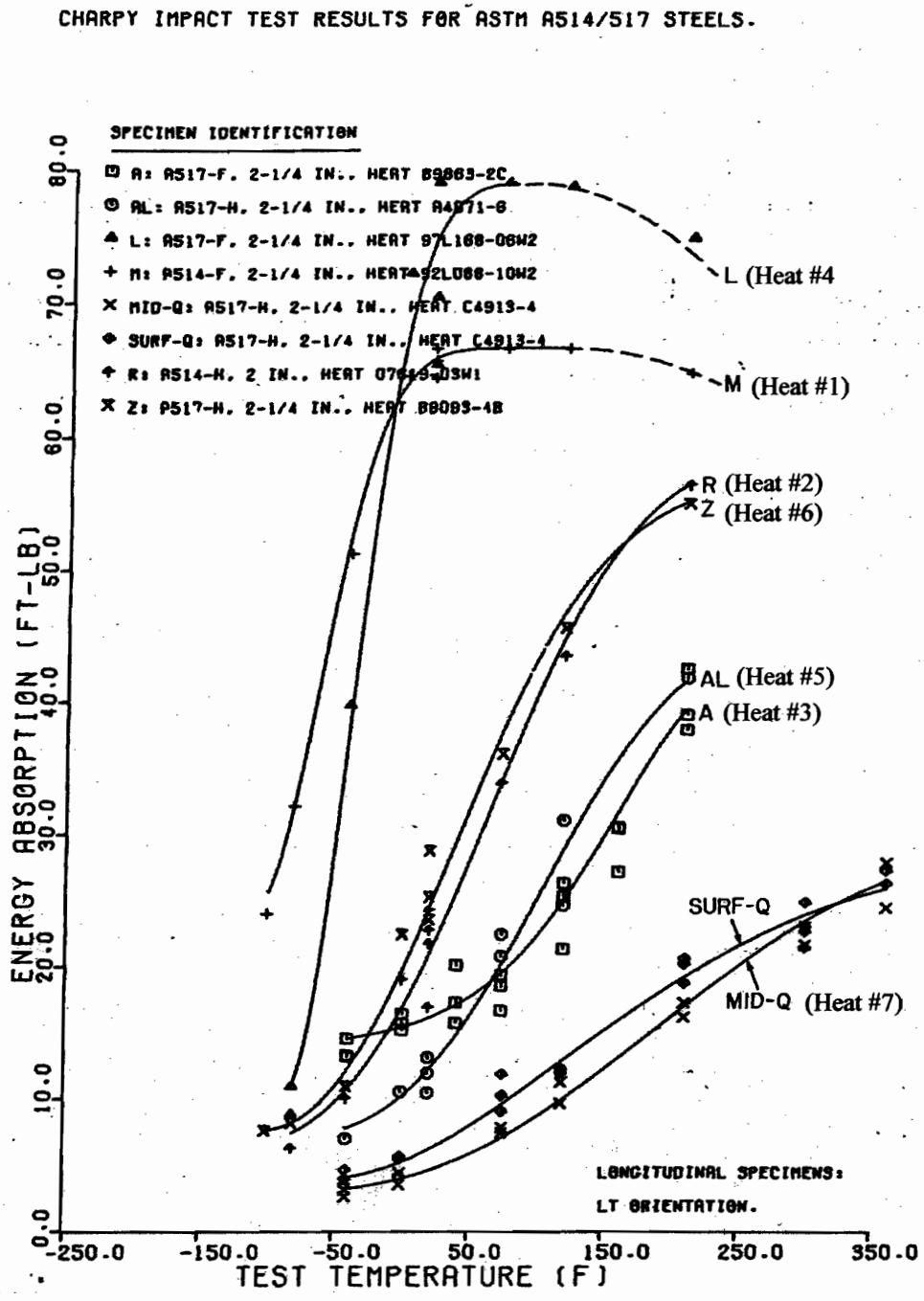

Fig. 7. Charpy V-notch impact test results for longitudinal specimens of one heat of ASTM A514 Grade F, one heat of A514 Grade H, two heats of A517 Grade F, and four heats of A517 Grade H steels (after Interrante and Hicho [10] and ASTM [11]). Curves are best fits of those eight sets of data using the least square method. Note the scatters within and between heats from -100 to 240 °C. The ASTM specification for A 517/A 517M - 93 steel [11] calls for a minimum tensile yield strength of 620 MPa (90 ksi) for plates of 65 to 150 mm (2.50 to 6.00 in) thick and higher yield strengths for thinner plates (690 MPa (100 ksi) for 65 mm (2.50 in) or less).

2. ***Problem*** *(Continued)*

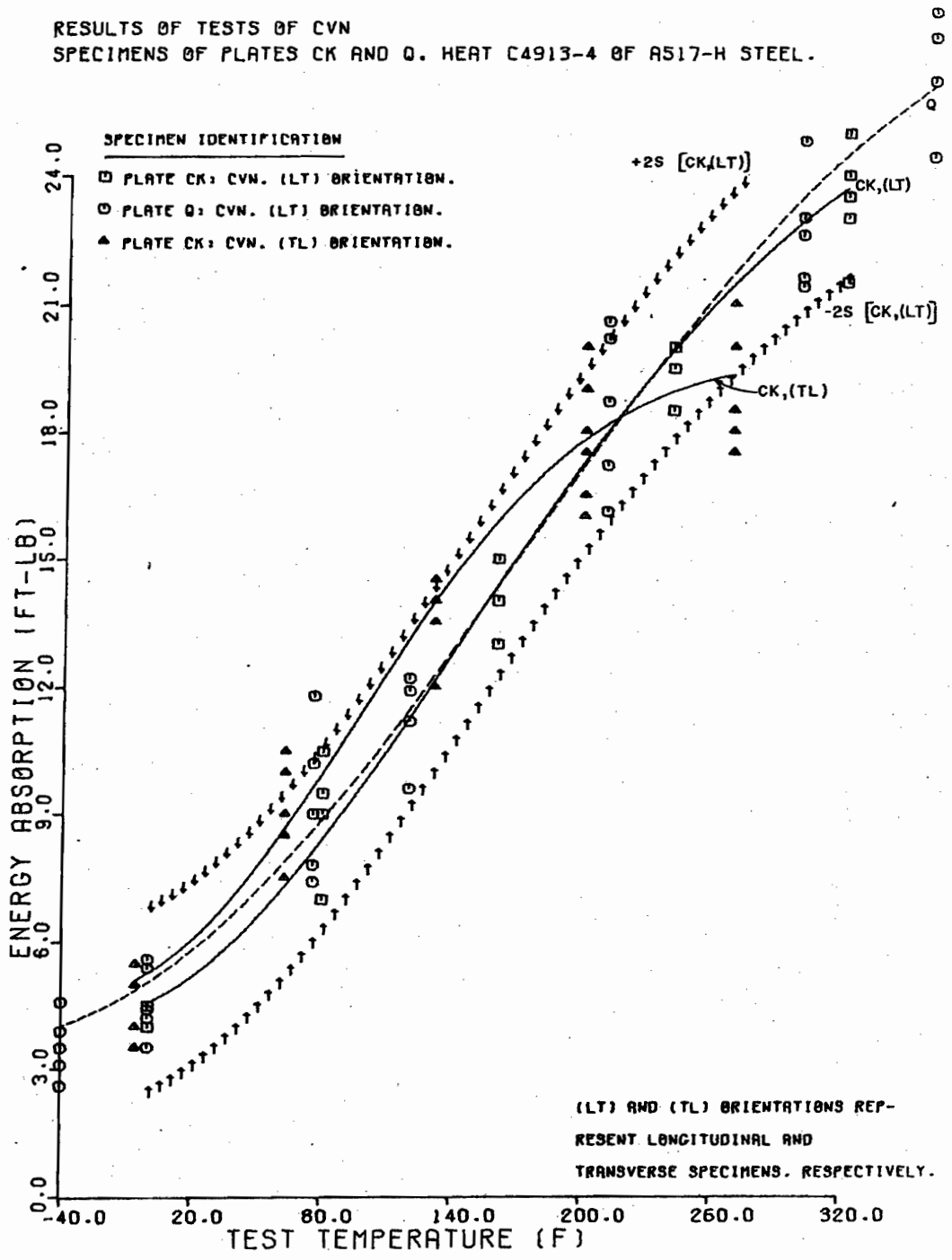

Fig. 8. Charpy V-notch impact test results for longitudinal (LT) and transverse (TL) specimens of a single heat of ASTM A517 Grade H steel (after Interrante and Hicho [10] and ASTM [11]). Curves are best fits of those eight sets of data using the least square method. Note the scatters within specimens and between specimens of different orientation from 0 to 360 °C.

3. THEORY OF PROPAGATION OF ERRORS

In 1939, Birge [18] wrote an expository paper on the theory of error propagation with a remark which is worth quoting below because it sounded as relevant almost 70 years ago as it is today:

> *"The question of what constitutes the most reliable value to be assigned as the uncertainty of any given measured quantity is one that has been discussed for many decades and, presumably, will continued to be discussed.*
>
> *"It is a question that involves many considerations and by its very nature has no unique answer.*
>
> *"The subject of the propagation of errors, on the contrary, is a purely mathematical matter, with very definite and easily ascertained conclusions.*
>
> *"Although the general subject of the present article is by no means new, many scientists (**and engineers**) still fail to avail themselves of the enlightening conclusions that may often thus be reached, while others frequently use the theory incorrectly and thus arrive at quite misleading conclusions."* [Note: Words in bold are added by this author to emphasize the applicability of Birge's remark to a broader audience.]

Inspired by the above, Ku [19] wrote a follow-up paper with a clear and easy-to-understand exposition of the same subject and included several tables with frequently-used formulas, some of which are reproduced in Table 1 of this paper.

In particular, two of the 11 formulas listed in Table 1, namely, the product of two variables and the square root of a single variable, are used in this paper to derive a variance formula for fracture toughness, K_{Ic}. For a fuller introduction of the subject of error propagation and the mathematical derivation of a large number of useful formulas, the reader is referred to either Ku's paper [19], or a more recent one by Nelson [20].

Of the four formulas for K_{Ic} listed in Section 2, the Wullaert-Server equation [16] is the simplest in having the yield strength as the second variable. Since our paper is to present a methodology and not to arrive at an answer to a specific problem, it does not matter which formula to use. So our starting point is the Wullaert-Server K_{Ic} -CVN relationship:

$$K_{Ic} = 2.1 \, (\sigma_y * CVN)^{1/2} \quad , \qquad (1)$$

where K_{Ic} is the static crack initiation toughness (ksi-in$^{1/2}$), σ_y, the tensile yield strength (ksi) corresponding to the approximate loading rate, and CVN, the Charpy V-notch energy (ft-lb).

Let us introduce an intermediate variable named Q, where Q is defined as the product of σ_y and CVN. Furthermore, let us abbreviate the variable, CVN, as simply, C. By using the 7th formula in Table 1, we derive the following variance formula for Q:

$$\begin{aligned} \mathrm{Var}(Q) &= \mathrm{Var}(\sigma_y * C) \\ &= (\sigma_y * C)^2 * (\mathrm{Var}(\sigma_y) / \sigma_y^2 + \mathrm{Var}(C) / C^2) \\ &= C^2 \, \mathrm{Var}(\sigma_y) + \sigma_y^2 \, \mathrm{Var}(C) \end{aligned} \qquad (2)$$

By using the 9th formula in Table 1, we derive the following variance formula for the square root of Q:

$$\mathrm{Var}(Q^{1/2}) = (1/4) \, \mathrm{Var}(Q) / Q$$

Substituting the result of eq. (2) in the above, we get

$$\begin{aligned} \mathrm{Var}(Q^{1/2}) &= (1/4) \, (C^2 \, \mathrm{Var}(\sigma_y) + \sigma_y^2 \, \mathrm{Var}(C)) / (\sigma_y * C) \\ &= (1/4) \, \{ (C/\sigma_y) \, \mathrm{Var}(\sigma_y) + (\sigma_y/C) \, \mathrm{Var}(C) \} \end{aligned} \qquad (3)$$

Substituting eq. (3) into eq. (1), we obtain the following variance formula for K_{Ic}:

$$\mathrm{Var}(K_{Ic}) = 1.103 \, \{ (C/\sigma_y) \, \mathrm{Var}(\sigma_y) + (\sigma_y/C) \, \mathrm{Var}(C) \} \qquad (4)$$

For different combinations of values of σ_y and C ($= CVN$), eq. (4) delivers a variety of correlations between the variance of K_{Ic} and the variances of σ_y and CVN. For example, the following lists some useful formulas for a numerical representation of eq. (4):

σ_y ksi	CVN ft-lb	Variance formulas for K_{Ic} in ksi-in$^{1/2}$ units (*) (Note: *Not valid for K_{Ic} in MPa-m$^{1/2}$ units*)	
60	30	$\mathrm{Var}(K_{Ic}) = 0.551 \, \mathrm{Var}(\sigma_y) + 2.205 \, \mathrm{Var}(CVN)$	(5)
90	45	$\mathrm{Var}(K_{Ic}) = 0.551 \, \mathrm{Var}(\sigma_y) + 2.205 \, \mathrm{Var}(CVN)$	(5)
120	60	$\mathrm{Var}(K_{Ic}) = 0.551 \, \mathrm{Var}(\sigma_y) + 2.205 \, \mathrm{Var}(CVN)$	(5)
45	15	$\mathrm{Var}(K_{Ic}) = 0.368 \, \mathrm{Var}(\sigma_y) + 3.308 \, \mathrm{Var}(CVN)$	(6)
90	30	$\mathrm{Var}(K_{Ic}) = 0.368 \, \mathrm{Var}(\sigma_y) + 3.308 \, \mathrm{Var}(CVN)$	(6)
80	20	$\mathrm{Var}(K_{Ic}) = 0.276 \, \mathrm{Var}(\sigma_y) + 4.410 \, \mathrm{Var}(CVN)$	(7)
100	20	$\mathrm{Var}(K_{Ic}) = 0.221 \, \mathrm{Var}(\sigma_y) + 5.513 \, \mathrm{Var}(CVN)$	(8)

Note the relative importance of $\mathrm{Var}(CVN)$ vs. $\mathrm{Var}(\sigma_y)$.

(*) These formulas are approximate, as noted in Table 1, and are valid if σ_y and CVN are statistically independent. Additional terms are needed if they are correlated (see Ku [19]).

Table 1. Propagation of error formulas for 11 commonly encountered functions as listed by Ku [19].

	Function form of $\hat{w}$ *	Approx. formula for $\hat{\text{var}}(\hat{w})$ (x and y are assumed to be statistically independent)	Term to be added if x and y are correlated, and a reliable estimate of $\sigma_{\overline{xy}}$, $s_{\overline{xy}}$, can be assumed
1.	$A\bar{x}+B\bar{y}$	$A^2 s_{\bar{x}}^2 + B^2 s_{\bar{y}}^2$	$2AB s_{\overline{xy}}$
2.	$\left(\frac{\bar{x}}{\sigma_{\bar{x}}^2}+\frac{\bar{y}}{\sigma_{\bar{y}}^2}\right)\Big/\left(\frac{1}{\sigma_{\bar{x}}^2}+\frac{1}{\sigma_{\bar{y}}^2}\right)$**	$1\Big/\left(\frac{1}{\sigma_{\bar{x}}^2}+\frac{1}{\sigma_{\bar{y}}^2}\right)$	
3.	$\frac{\bar{x}}{\bar{y}}$	$\left(\frac{\bar{x}}{\bar{y}}\right)^2\left(\frac{s_{\bar{x}}^2}{\bar{x}^2}+\frac{s_{\bar{y}}^2}{\bar{y}^2}\right)$	$\left(\frac{\bar{x}}{\bar{y}}\right)^2\left(-2\frac{s_{\overline{xy}}}{\bar{x}\bar{y}}\right)$
4.	$\frac{1}{\bar{y}}$	$\frac{s_{\bar{y}}^2}{\bar{y}^4}$	
5.	$\frac{\bar{x}}{\bar{x}+\bar{y}}$	$\left(\frac{\hat{w}}{\bar{x}}\right)^4(\bar{y}^2 s_{\bar{x}}^2+\bar{x}^2 s_{\bar{y}}^2)$	$\left(\frac{\hat{w}}{\bar{x}}\right)^4(-2\bar{x}\bar{y}s_{\overline{xy}})$
6.	$\frac{\bar{x}}{1+\bar{x}}$	$\frac{s_{\bar{x}}^2}{(1+\bar{x})^4}$	
7.	$\overline{xy}$	$(\overline{xy})^2\left(\frac{s_{\bar{x}}^2}{\bar{x}^2}+\frac{s_{\bar{y}}^2}{\bar{y}^2}\right)$	$(\overline{xy})^2\left(2\frac{s_{\overline{xy}}}{\bar{x}\bar{y}}\right)$
8.	$\bar{x}^2$	$4\bar{x}^2 s_{\bar{x}}^2$	
9.	$\sqrt{\bar{x}}$	$\frac{1}{4}\frac{s_{\bar{x}}^2}{\bar{x}}$, where, the numerator in Ku's notation, $s_{xbar}{}^2$, equals *Var(xbar)* .	
10.	$\ln \bar{x}$	$\frac{s_{\bar{x}}^2}{\bar{x}^2}$	
11.	$k\bar{x}^a\bar{y}^b$	$\hat{w}^2\left(a^2\frac{s_{\bar{x}}^2}{\bar{x}^2}+b^2\frac{s_{\bar{y}}^2}{\bar{y}^2}\right)$	$(\hat{w})^2\left(2ab\frac{s_{\overline{xy}}}{\bar{x}\bar{y}}\right)$

* It is assumed that the value of $\hat{w}$ is finite and real, e.g., $\bar{y}\neq 0$ for ratios with $\bar{y}$ as denominator, $\bar{x}>0$ for $\sqrt{\bar{x}}$ and $\ln \bar{x}$.
** Weighted mean as a special case of $A\bar{x}+B\bar{y}$, with σ_x and σ_y considered known.

4. ESTIMATE OF VARIANCE OF YIELD STRENGTH

The general formula, eq. (4), derived in the last section for the variance of K_{Ic} in terms of the variances of σ_y and *CVN*, requires us to determine the last two variances from experimental data of the same grade of steel.

Since our investigation in this paper to find the variance of *CVN* relies heavily on some Charpy energy data of the ASTM A514/517 steel [10], we shall present in this section some literature data (see, e.g., ASM [21], NRIM [22], etc.) on the tensile yield strength of a high-strength steel comparable to ASTM A514/517 as a source for estimating Var(σ_y).

In particular, we find the set of data in a 1994 report of the National Research Institute of Metals, Tokyo [22] so interesting and relevant as an example for statistical analysis that we simply reproduce that entire set in Table 2 for readers to pursue additional analyses of the same nature when a need arises.

In Fig. 9, we reproduce a plot from the NRIM report [22] showing the tensile yield strength data of 21 specimens per temperature for 10 temperatures from 20 °C to 600 °C. Note that the mean yield strength of the Class 590 MPa steel plate at 20 °C is 518.7 MPa (75.2 ksi), which is ~17% less than the minimum yield strength of ASTM A517 [11]. Since the effect of the variance of yield strength, as shown earlier in eqs. (5) through (8), is relatively less important than that of the Charpy energy, we believe we are justified in analyzing the data in Table 2 to get an estimate of Var(σ_y) for ASTM A517.

In Fig. 10, we plotted the 21 yield strength data of the Class 590 MPa steel at 20 °C in a histogram and used a normality test subroutine of the well-known Wilks-Shapiro test [23, 24] as implemented in a public-domain software package named DATAPLOT [25] to verify that the normality assumption for the data set was valid at the 95 % and 99 % confidence.

In Fig. 11, we did the same for the 21 yield strength data of the Class 590 MPa steel at 100 °C, and verified that the data set was normal.

The results of the analyses displayed in Figs. 10 and 11 are summarized below: (sd = standard deviation)

For 20 °C: Sample ave. = 518.67 MPa. sd(σ_y) = 47.15 MPa.

For 100 °C: Sample ave. = 500.24 MPa. sd(σ_y) = 52.22 MPa.

By a linear interpolation on the two sd's, we obtain:

For 48.9 °C (**120 °F**), sd(σ_y) = 48.98 MPa (**7.10 ksi**), which will be used in our subsequent analysis in determining the variance of the yield strength of ASTM A517 steel.

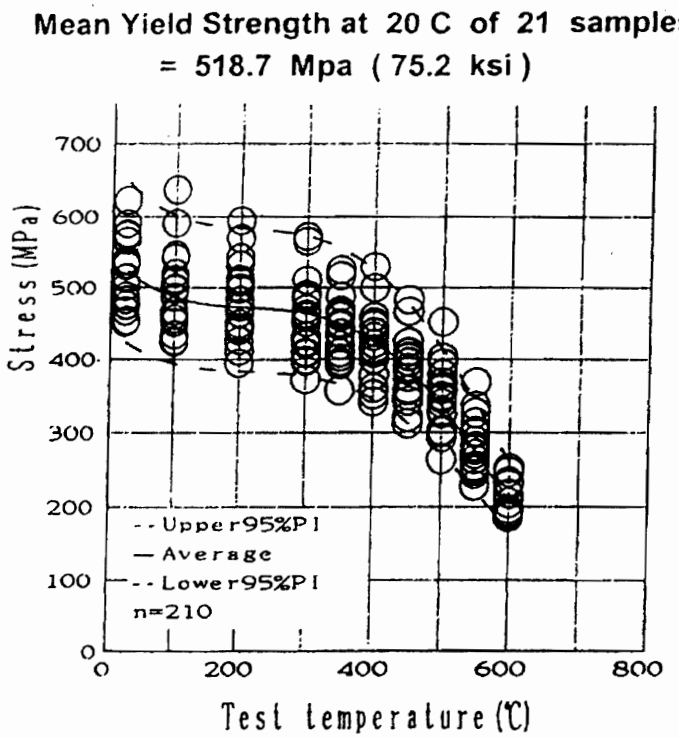

Fig. 9. Test data of 21 specimens per temperature from 20 °C to 600 °C for 0.2% proof stress or tensile yield strength of a Japanese high strength steel (Class 590 MPa) plate (see list of data in **Table 2** from NRIM[22]). The solid curve represents the average; the broken curves the upper and lower 95 % prediction confidence intervals. Note the scatter of the 21 data points per temperature.

Table 2. Test data of 21 specimens per temperature for 10 temperatures from 20 °C (RT) to 600 °C for tensile yield strength (0.2% proof stress), listed on the left, and ultimate strength, listed on the right of each data column as shown, for a Japanese high strength steel (Class 590 MPa) plate as reported by NRIM in a 1994 report [22]

Short-time tensile properties of high strength steel (Class 590 MPa) plates

0.2 % proof stress and tensile strength

NRIM reference code	Test temperature (℃) RT	100	200	300	350	400	450	500	550	600
SPV 490[1]	≥490 / 610–740	0.2 % proof stress (MPa)[3] / Tensile strength (MPa)								
SM 570[2]	≥450 / 570–720	0.2 % proof stress (MPa) / Tensile strength (MPa)								
CbA	468 / 592	434 / 545	439 / 557	421 / 596	403 / 566	410 / 513	350 / 457	349 / 426	271 / 344	194 / 276
CbB	483 / 625	513 / 645	450 / 623	452 / 634	436 / 614	415 / 569	375 / 490	303 / 410	264 / 363	189 / 288
CbC	590 / 692	591 / 682	570 / 670	566 / 701	526 / 659	530 / 600	467 / 543	403 / 476	288 / 387	186 / 316
CbD	574 / 698	512 / 637	513 / 641	486 / 649	461 / 630	431 / 567	424 / 535	380 / 466	326 / 412	252 / 338
CbE	515 / 629	500 / 592	473 / 589	459 / 624	446 / 595	399 / 530	395 / 487	368 / 430	298 / 367	218 / 300
CbF	451 / 578	424 / 534	411 / 533	396 / 571	397 / 551	351 / 490	342 / 432	300 / 380	248 / 321	189 / 269
CbG	476 / 604	456 / 572	443 / 578	432 / 602	409 / 582	422 / 548	403 / 495	357 / 433	299 / 374	230 / 310
CbH	514 / 612	497 / 593	511 / 599	451 / 618	459 / 602	415 / 544	388 / 491	328 / 409	249 / 333	190 / 282
CbJ	623 / 736	637 / 736	594 / 712	573 / 711	517 / 660	500 / 630	483 / 576	452 / 536	368 / 464	247 / 366
CbL	566 / 676	541 / 647	507 / 630	491 / 660	467 / 630	450 / 580	409 / 517	361 / 452	281 / 382	206 / 311
CbM	533 / 636	510 / 612	482 / 602	464 / 636	437 / 609	447 / 566	396 / 496	359 / 436	301 / 376	217 / 312
CbN	455 / 571	426 / 533	394 / 530	372 / 571	357 / 562	341 / 509	321 / 450	293 / 389	244 / 326	185 / 258
CbR	501 / 630	513 / 642	482 / 636	408 / 614	451 / 623	422 / 551	388 / 505	303 / 408	251 / 343	191 / 290
CbS	530 / 630	486 / 578	465 / 570	511 / 664	451 / 598	459 / 576	383 / 475	338 / 424	276 / 364	206 / 317
CbT	539 / 648	508 / 614	497 / 609	471 / 637	447 / 609	438 / 559	403 / 503	362 / 441	313 / 378	229 / 313
CbU	513 / 660	521 / 625	500 / 616	488 / 655	469 / 630	437 / 570	402 / 509	345 / 444	278 / 367	193 / 299
CbV	483 / 618	468 / 586	411 / 585	399 / 583	395 / 573	360 / 517	312 / 442	265 / 378	229 / 330	185 / 270
CbW	488 / 601	463 / 569	445 / 585	425 / 604	420 / 599	407 / 530	356 / 468	328 / 416	256 / 328	193 / 279
CbX	537 / 672	508 / 640	532 / 666	461 / 634	465 / 632	413 / 566	411 / 519	394 / 478	335 / 415	234 / 330
CbY	476 / 603	451 / 567	426 / 569	410 / 607	391 / 580	378 / 532	359 / 473	325 / 405	253 / 341	193 / 276
CbZ	577 / 693	546 / 663	542 / 679	483 / 664	487 / 655	452 / 583	404 / 513	360 / 450	279 / 375	198 / 313

1) SPV 490, JIS G 3115-1990
2) SM 570, JIS G 3106-1992
3) 1 MPa = 1 N/mm² = 0.101972 kgf/mm²

*4. **Yield Strength Variance** (Continued)*

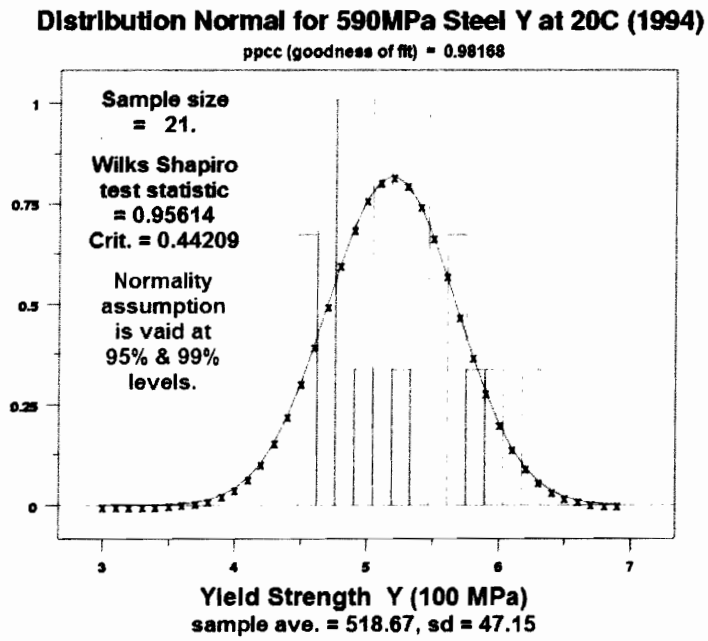

Fig. 10. A histogram plot of 21 yield strength data of a Japanese high strength steel (Class 590 MPa) plate at 20 °C (after NRIM [22]). Using the Wilks-Shapiro normality test [23, 24] as implemented in a public domain statistical data analysis software package named DATAPLOT [25], we verified that the normality assumption for the data set is valid at the 95 % and 99 % levels.

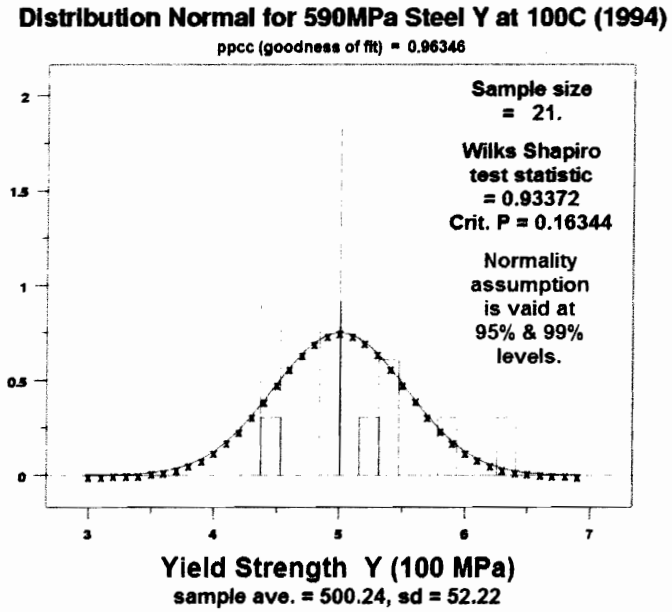

Fig. 11. A histogram plot of 21 yield strength data of a Japanese high strength steel (Class 590 MPa) plate at 100 °C (after NRIM [22]). Using the Wilks-Shapiro normality test [23, 24] as implemented in a public domain statistical data analysis software package named DATAPLOT [25], we verified that the normality assumption for the data set is valid at the 95 % and 99 % levels.

5. A DILEMMA IN ESTIMATING HEAT-TO-HEAT VARIABILITY OF CHARPY V-NOTCH IMPACT ENERGY

Having determined the standard deviation of the yield strength at 48.9 °C (120 °F) for a high-strength steel comparable to ASTM A517, we shall return to the task of sorting out the Charpy energy data reported by Interrante and Hicho [10] as shown in Fig. 7, where we found a huge heat-to-heat variation (from 10 to 80 ft-lb, or, 13.6 to 108 N-m) at 48.9 °C (120 °F).

Before we begin, it may be of interest to some readers to have a review of the terminology in statistics on two types of confidence intervals we shall use in this paper. [Note: For a reader familiar with those concepts, one can skip Notes 5.1 through 5.6 and go directly to the analysis of Charpy data following the review notes.]

Following Nelson, Coffin, and Copeland [27, pp. 165-181]), we introduce the concepts in a sequence of review notes:

Note 5.1. For a given sample data of a large population, we define the sample average, *ybar*, as a number that estimates the unknown mean, μ, of that population. Here, *ybar* is called a *point estimate.*

Note 5.2. For that same population, we then derive some formulas to add and subtract an amount, say, d, from *ybar* to get an interval, $(ybar-d,\ ybar+d)$, and call it an *interval estimate*, or a confidence interval for μ. The amount d depends on what level of confidence we are interested in, say, 90 %, 95 %, or 99 %. For our purposes here, let us just define two types of 95 % confidence intervals as follows:

Type 1. 95 % confidence interval for μ, i.e., $(ybar-d_1,\ ybar+d_1)$

This makes a statement about where the unknown mean of the population, μ, is likely to be within that interval 19 times out of 20. A formula for calculating d_1 is given below:

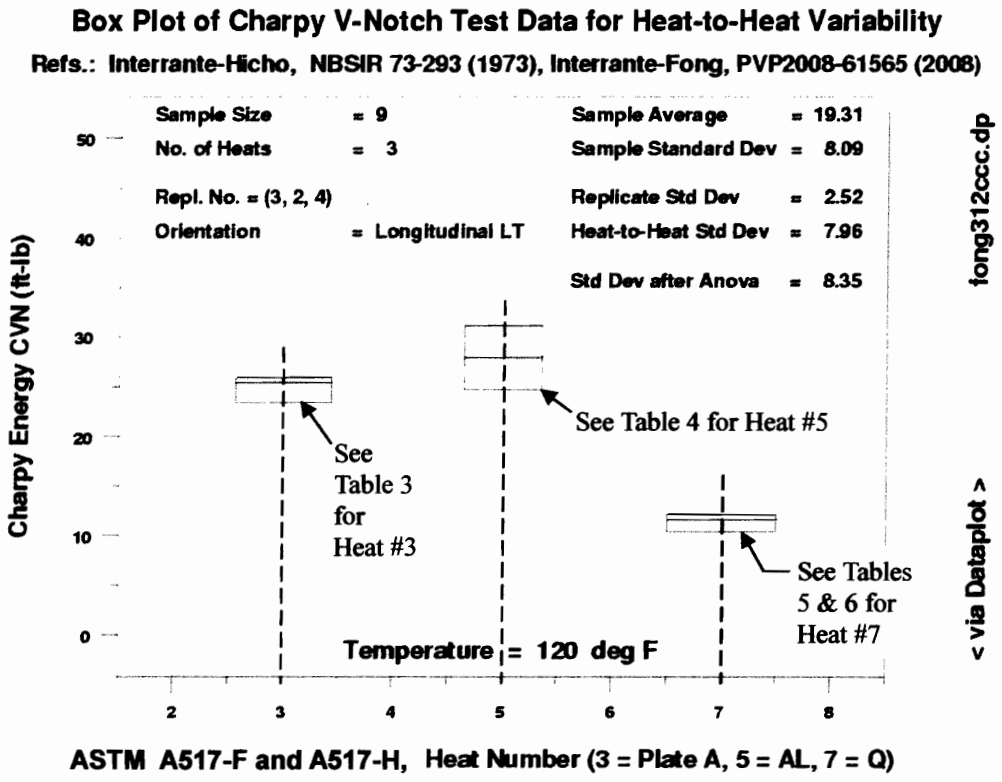

Fig. 12. Box plot and results of an analysis of variance (ANOVA) of three heats of Charpy V-notch data at 120 °F (48.9 °C) for a high strength ASTM A517 steel (one heat of Grade F and two heats of Grade H, as listed in **Tables 3-6** from Interrante & Hicho [10]). All 9 specimens, of which 3 were for A517-F Heat 89863-2C, 2 for A517-H Heat A4071-6, and 4 for A517-H Heat C4913-4, were orientated in the longitudinal (LT) direction. Note the presence of both within-heat (2.52) and between-heat (7.96) effects in the composite data set with the between-heat or heat-to-heat standard deviation (sd) about 3 times larger than the within-heat sd.

Table 3. 21 test data of 3 **surface** longitudinal specimens per temperature for 7 temperatures from -40 °F (-40 °C) to 210 °F (98.9 °C) for Charpy V-notch impact test of ASTM A517 Grade F **Heat #3** (89863-2C) steel plate named **Plate A** (after Interrante & Hicho [10])

IMPACT TEST RESULTS FOR PLATE A.

CALCULATIONS FOR ENERGY ABSORPTION DATA OF CHARPY V-NOTCH TESTS OF SURFACE.
LONGITUDINAL SPECIMENS: LT ORIENTATION.
A: A517-F, 2-1/4 IN., HEAT B9863-2C.

SPECIMEN	TEMPERATURE(F)	OBSERVED ENERGY ABSORPTION(FT-LB)
6A1	-40.0	13.20
6A2	-40.0	14.55
6A3	-40.0	13.28
6A4	.0	15.20
6A5	.0	15.73
6A6	.0	16.42
6A7	40.0	17.29
6A8	40.0	20.19
6A9	40.0	15.73
6A10	74.0	19.40
6A11	74.0	18.58
6A12	74.0	16.68
6A13	120.0	21.40
6A14	120.0	25.36
6A15	120.0	26.35
6A16	160.0	30.57
6A17	160.0	30.48
6A18	160.0	27.19
6A19	210.0	37.93
6A20	210.0	39.06
6A21	210.0	42.53

Table 4. 10 test data of one or more **surface** longitudinal specimens per temperature for 6 temperatures from -40 °F (- 40 °C) to 210 °F (98.9 °C) for Charpy V-notch impact test of ASTM A517 Grade H **Heat #5** (A4071-6) steel plate named **Plate AL** (after Interrante & Hicho [10])

IMPACT TEST RESULTS FOR PLATE AL.

CALCULATIONS FOR ENERGY ABSORPTION DATA OF CHARPY V-NOTCH TESTS OF SURFACE.
LONGITUDINAL SPECIMENS: LT ORIENTATION.
AL: A517-H, 2-1/4 IN., HEAT A4071-6.

SPECIMEN	TEMPERATURE(F)	OBSERVED ENERGY ABSORPTION(FT-LB)
4AL1	-40.0	6.95
4AL2	.0	10.50
4AL3	20.0	11.90
4AL4	20.0	10.40
4AL5	20.0	13.10
4AL6	74.0	20.85
4AL10	74.0	22.50
4AL7	120.0	24.70
4AL9	120.0	31.10
4AL8	210.0	41.80

Table 5. 21 test data of two or more **surface** longitudinal specimens per temperature for 7 temperatures from -40 °F (-40 °C) to 360 °F (182.2 °C) for Charpy V-notch impact test of ASTM A517 Grade H **Heat #7** (C4913-4) steel plate named **Plate Q** (after Interrante & Hicho [10])

IMPACT TEST RESULTS FOR SURFACE OF PLATE Q.

CALCULATIONS FOR ENERGY ABSORPTION DATA OF CHARPY V-NOTCH TESTS OF SURFACE.
LONGITUDINAL SPECIMENS: LT ORIENTATION.
SURF-Q: A517-H, 2-1/4 IN., HEAT C4913-4.

SPECIMEN	TEMPERATURE(F)	OBSERVED ENERGY ABSORPTION(FT-LB)
2A16	-40.0	3.90
2B18	-40.0	3.50
1A15	-40.0	4.60
1B17	-40.0	3.50
1B2	.0	5.60
2B2	.0	4.20
3B2	.0	5.40
1B1	76.0	10.20
2B1	76.0	9.00
3B1	76.0	11.80
1B5	120.0	12.20
2B5	120.0	11.90
1B6	210.0	18.70
2B6	210.0	20.20
3B8	210.0	20.60
1A6	300.0	23.00
2A3	300.0	24.80
1B11	300.0	22.60
2B9	300.0	21.40
1A14	360.0	26.20
1A18	360.0	27.20

Table 6. 14 test data of two **midthickness** longitudinal specimens per temperature for 7 temperatures from -40 °F (-40 °C) to 360 °F (182.2 °C) for Charpy V-notch impact test of ASTM A517 Grade H **Heat #7** (C4913-4) steel plate named **Plate Q** (after Interrante & Hicho [10])

IMPACT TEST RESULTS FOR MIDTHICKNESS OF PLATE Q

CALCULATIONS FOR ENERGY ABSORPTION DATA OF CHARPY V-NOTCH TESTS OF MIDTHICKNESS.
LONGITUDINAL SPECIMENS: LT ORIENTATION.
MID-Q: A517-H, 2-1/4 IN., HEAT C4913-4.

SPECIMEN	TEMPERATURE(F)	OBSERVED ENERGY ABSORPTION(FT-LB)
1C13	-40.0	2.60
2C19	-40.0	3.10
1C2	.0	4.40
2C2	.0	3.50
1C1	76.0	7.40
2C1	76.0	7.80
1C6	120.0	11.20
2C5	120.0	9.60
1C7	210.0	17.20
2C6	210.0	16.10
1C12	300.0	21.60
2C9	300.0	23.00
1C14	360.0	24.40
2C17	360.0	27.80

5. A Dilemma in Estimating Heat-to-Heat Variability (Cont'd)

$$d_1 = t(0.025;\ n-1) * s / (n)^{1/2} \quad , \qquad (9)$$

where $t(\alpha,\nu)$ is the t-distribution for a normal population with values tabulated in most statistics textbooks (see, e.g., Nelson, et al. [27, p. 444]).

Type 2. 95% prediction interval for μ, i.e., $(ybar-d_2,\ ybar+d_2)$

This makes a statement about where a *single* future observation is likely to be within that interval 19 times out of 20. Because a prediction interval is for an individual observation, i.e., a "moving target," it is necessarily wider than the first type ($d_2 > d_1$).

The formula for calculating d_2 in order to construct the 95% prediction interval of the mean, μ, of a population, for which the sample average, $ybar$, the sample size, n, and the sample standard deviation, s, are known, is given below:

$$d_2 = t(0.025;\ n-1) * s * (1 + 1/n)^{1/2}, \qquad (10)$$

where t is the t-distribution for a normal population with values tabulated in most statistics textbooks (see, e.g., Nelson, et al [14, p. 444]).

Note 5.3. From eq. (9), we note that d_1 approaches zero as n approaches infinity. In theory, we can make a confidence interval (Type 1) as narrow as we like simply by increasing the sample size, n. (In practice, this is often difficult to accomplish.)

Note 5.4. From eq. (10), we note that d_2 approaches $t * s$ as n approaches infinity. This means the width of a prediction interval depends mostly on the amount of spread (e.g., standard deviation) in the population. If the members of the population vary widely, prediction intervals will necessarily be wide, no matter how large a sample is taken.

Note 5.5. When constructing a confidence interval (Type 1), we need not worry too much about whether the underlying population is normally distributed. The same is not true for prediction intervals (Type 2). Since we are trying to pinpoint where individual observation might lie, the shape of the distribution is very important. The simplest situation for constructing prediction intervals is when the underlying population is (at least approximately) normally distributed.

Note 5.6. In this paper, all of our discussion on confidence intervals will be for the second type (prediction intervals) based on eq. (10).

(End of Review Notes on Confidence Intervals.)

Analysis of Charpy Data of 7 Heats given in Ref. [10]

Given the set of Charpy data of 7 heats in Ref. [10], we are interested in finding two kinds of variances, namly, the replicate (within) and the heat-to-heat (between) variances.

The analysis tool we use is the classical statistical technique known as the analysis of variance, or, ANOVA (see, e.g., Draper and Smith [26], and Nelson, Coffin, and Copeland [27]). The data we need should have at least two specimens (replicates) from the same heat for each temperature we are interested in.

In Tables 3 through 6, we list Charpy energy data for Plates A, AL, and Q from 3 heats that satisfy the replication criterion for 48.9 °C (120 °F). Those data were plotted in Fig. 12, where we showed the results of an ANOVA as implemented in DATAPLOT [25], with, indeed, a strong heat-to-heat effect (sd = 7.96) vs. a small replicate effect (sd = 2.52).

For n = 9, the sample average, $ybar$, for 9 specimens from 3 heats was found to be 19.31 ft-lb, and the standard deviation, s, after ANOVA was found to be 8.35 ft-lb.

Using a table of t distributions (see, e.g., Nelson, et al [27, p. 444]), where we found the critical value of $t(0.025, 9)$ to be 2.262, we calculate d_2, the 95 % confidence half-interval, to be $2.262 * 8.35 * (1 + 1/9)^{1/2} = 19.91$ ft-lb, and the sample mean with predicted intervals to be (-0.60 ft-lb, 39.22 ft-lb), an answer we are not willing to accept, because the lower bound is negative.

Another reason why the analysis results for the 3 heats with replicates is not acceptable is that we have ignored the single-specimen data of the other four heats, and those four data values are very high (see below) as compared to the sample average ($ybar$ = 19.31) of the 9 replicated specimens:

Plate	L	M	R	Z	Unit
Heat # (See Fig.7)	4 97L168-06W2	1 92L088-10W2	2 07619-03W1	6 B9093-4B	
CVN =	78.80	66.70	43.50	45.60	ft-lb.

If we close our eyes, add those 4 single-specimen data to the 9 replicated ones to make a sample of 13, calculate the sample ave ($ybar$ = 31.42), standard deviation (s = 21.76), the critical $t(0.025,13) = 2.16$, and the half-interval (d_2 = 48.77), we end up with a non-sensical prediction interval of (-17.35, 80.19) ft-lb, because its lower bound is again negative.

In the next section, we shall present a way out of the above dilemma in solving a heat-to-heat variability problem.

6. A METHODOOGY FOR ESTIMATING VARIANCE OF CHARPY V-NOTCH IMPACT ENERGY

The analysis methodology described in the previous section depended on two classical statistical data analysis techniques, namely, the ANOVA and the prediction confidence intervals. The methodology is capable of delivering an answer to a user on the question whether a specific set of data is normally distributed and homogeneous, and if it is not homogeneous, one can estimate the replicate (within) and between-treatment variability to obtain an improved estimate of variance. But the methodology does not go far enough to prescribe an efficient and rigorous follow-up plan of how to do more tests, if and when the analysis yields non-sensical results as it did in Section 5.

In this section, we introduce a third statistical data analysis technique, known as the design of experiments (DOE) as described in Box, Hunter, and Hunter [28], and Montgomery [29], that will rectify the shortcomings of the previous methodology. For a detailed exposition of this DOE technique, the reader may consult Croarkin, et al [30], or a recent tutorial paper with two examples by Fong, et al [31].

To acquaint the reader with the fundamentals of DOE, we present below a series of 8 questions and answers as a background for applying the concept to estimating the variance of Charpy V-notch energy, Var(*CVN*). (Note: For a reader familiar with the DOE technique, one can skip the parts through DOE 1.8 and go directly to the *CVN* energy example that follows.]:

DOE 1.1 What is design of experiments (DOE)?

Ans. In an experiment, we change one or more process variables (factors) in order to observe the effect the changes have on one or more response variables. DOE is an efficient procedure for planning experiments so that the data obtained can be analyzed to yield valid and objective conclusions.

DOE begins with determining the *objectives* of an experiment and selecting the *process factors* for the study. An Experimental Design is the laying out of a detailed experimental plan in advance of doing the experiment. Well chosen experimental designs maximize the amount of "information" that can be obtained for a given amount of experimental effort.

DOE 1.2 What is the first step in applying the DOE method?

Ans. The statistical theory underlying DOE begins with the concept of *process models*. A process *model* of the 'black box' type is formulated with several discrete or continuous input *factors* that can be controlled, and one or more measured output *responses*. The output responses are assumed continuous. Both real (results of physical experiments) and virtual (values conceived from experience and judgment of the experimentalist) data are used to derive an empirical (approximate) model linking the outputs and inputs. These empirical models generally contain *first-order (linear) and second-order (quadratic) terms.*

DOE 1.3 What is a first order model?

Ans. A first-order model with two factors, X_1 and X_2, can be written as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + errors \qquad (11)$$

Here, Y is the response for given levels of the main effects X_1 and X_2, and the X_1X_2 term is included to account for a possible interaction effect between X_1 and X_2. The constant β_0 is the response of Y when both main effects are 0.

In the example that follows for an application in NDE, we use a linear model with five factors and one response variable, and the total number of terms on the right hand side of eq. (11) is 2^5, or 32, as shown in the following eq. (12):

$$\begin{aligned} Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \\ \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{14} X_1 X_4 + \beta_{15} X_1 X_5 + \\ \beta_{23} X_2 X_3 + \beta_{24} X_2 X_4 + \beta_{25} X_2 X_5 + \\ \beta_{34} X_3 X_4 + \beta_{35} X_3 X_5 + \\ \beta_{45} X_4 X_5 + \\ \beta_{123} X_1 X_2 X_3 + \beta_{124} X_1 X_2 X_4 + \ldots + errors \end{aligned} \qquad (12)$$

A second-order model (typically used in response surface DOE's with suspected curvature as described in Ref. [28], pp. 510-539, or Ref. [30], Chap. 5, Sect. 5.1, pp. 9-20) does not include the three-way interaction terms such as $\beta_{123} X_1 X_2 X_3$, $\beta_{124} X_1 X_2 X_4$, etc., as shown above, but adds five more terms to the first order model (12), namely,

$$\beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2 + \beta_{55} X_5^2 . \qquad (13)$$

Note: Clearly, a full model could include many cross-product (or interaction) terms involving squared *X's*. However, in general these terms are not needed and most DOE software defaults to leaving them out of the model.

DOE 1.4 How does one select factors and responses?

Ans. Process variables of an experiment include both inputs (factors) and outputs (responses). The selection criteria are:

(a) Include all important factors (based on judgment).

(b) Be bold (consider worst cases from experience) in choosing

the low and high factor levels.

(c) Check the factor levels/settings for impractical or impossible com-binations, such as very low pressure or very high gas flows.

(d) Include all relevant responses.

(e) Avoid using only the combined responses of two or more measurements of the process. For example, if interested in the ratio of two factors, measure both factors, not just the ratios.

In choosing the range of levels/settings, it is wise to give this some thought beforehand rather than just try extreme values.

DOE 1.5 How does one select an experimental design?

Ans. The most popular experimental designs are *two-level designs*. Why only two levels? There are a number of good reasons why two is the most common choice; one reason is that it is ideal for screening designs, simple and economical; it also gives most of the information required to go to a multilevel response surface experiment if one is needed.

The standard layout for a 2-level design uses +1 and -1 notation to denote the "high level" and the "low level" respectively, for each factor. For example, the matrix below

	Factor 1 (*X1*)	Factor 2 (*X2*)
Trial 1	-1	-1
Trial 2	+1	-1
Trial 3	-1	+1
Trial 4	+1	+1

describes an experiment in which 4 trials (or runs) were conducted with each factor set to high or low during a run according to whether the matrix had a +1 or -1 set for the factor during that trial. If the experiment had more than 2 factors, there would be an additional column in the matrix for each additional factor. For example, a 3-factor full factorial design is represented by the following matrix:

Order of Run	*X1*	*X2*	*X3*
1	- 1	- 1	- 1
2	+1	- 1	- 1
3	- 1	+1	- 1
4	+1	+1	- 1
5	- 1	- 1	+1
6	+1	- 1	+1
7	- 1	+1	+1
8	+1	+1	+1

DOE 1.6 What is a 2-level full factorial DOE?

Ans. A common experimental design is one with all input factors set at two levels each. These levels are called 'high' and 'low', or '+1' and '-1', respectively. A design with all possible high/low combinations of all the input factors is called a full factorial design of experiments in two levels.

If there are k factors, each at 2 levels, a full factorial DOE has 2^k runs. Fig. 13 is a graphical representation of a 2-level, 3-factor, 2^3 or 8-run full factorial DOE. This implies eight runs (not counting replications or center point runs). The arrows show the direction of increase of the factors. The numbers '1' through '8' at the corners of the design box reference the "Standard Order" of runs (also referred to as the "Yates Order", see Ref. [30]). When the number of factors is 5 or greater, a full factorial DOE requires a large number of runs and is not very efficient. This is where and why there is a need for a fractional factorial DOE.

DOE 1.7 What is a Center Point in a 2-level design?

Ans. To introduce the concept of a center point, we again refer to Fig. 13, a graphical representation of a two-level, full factorial design for three factors, namely, the 2^3 design. As mentioned earlier, we adopt the convention of +1 and -1 for the factor settings of a two-level design. When we include a center point during the experiment, we mean a point located in the middle of the design cube, and the convention is to denote a center point by the value "0".

DOE 1.8 What is a 2-level fractional factorial DOE?

Ans. A fractional factorial DOE is a factorial experiment in which only an adequately chosen fraction of the treatment combinations required for the complete factorial experiment is selected to be run.

In general, we pick a fraction such as ½, ¼, etc. of the runs called for by the full factorial. We use various strategies that ensure an appropriate choice of runs. Properly chosen fractional factorial designs for 2-level experiments have the desirable properties of being both *balanced* and *orthogonal*. For example, the following matrix represents a 3-factor half-factorial design:

Order of Run	*X1*	*X2*	*X3 (X1*X2)*
1 (new), 5 (old)	- 1	- 1	+1
2 (new), 2 (old)	+1	- 1	- 1
3 (new), 3 (old)	- 1	+1	- 1
4 (new), 8 (old)	+1	+1	+1

A comparison of the half-fractional factorial design (new) with that of the full (old, see DOE 1.5) with the help of Fig. 13 shows the balanced and orthogonal nature of the DOE concept.

6. *Design of Experiments Approach* (Continued)

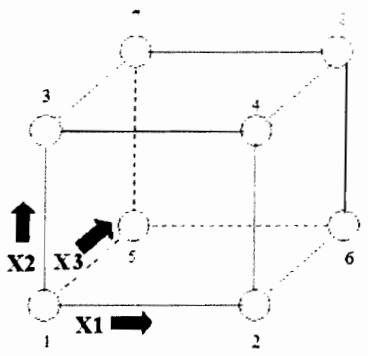

Fig. 13. A 2^3 2-level, full factorial design; factors *X1, X2, X3*.

A Charpy V-notch Energy (*CVN*) Example of DOE

Let us consider a. fictitious scenario where a Charpy V-notch impact test team, Team A, reported the result (=19.31 ft-lb) of a single-specimen test at 48.9 °C (120 °F). The question is:

> *What additional tests should Team A make in order to report the result with an expression of uncertainty in the form of a number with a 2-sided prediction 95% level of confidence, i.e., 19.31 (?) ft-lb?*

To answer this question, we asked Team A to name factors that they would consider to be highly important in getting a good response variable (output *Y1*), which, in this case, had already been selected, i.e., the Charpy energy, *CVN*. We also asked Team A to give us numerical values of those five factors when they did the single-specimen test. Within minutes, Team A reported back with one real factor (*X1*) and four fictitious factors (*X2, ... X5*) estimated from judgment:

X1	Temperature (deg. °F)	120 °F (48.9 °C)
X2	Manganese Sulfide (%)	0.10 %
X3	Initial Strain (%)	2.0 %
X4	Mis-orientation (degree)	8.0 deg.
X5	Notch radius (mm)	0.25 mm

According to our approach, a 5-factor experimental design requires 2^5, or, 32 specimens for a two-level full factorial DOE. If it is a 3-level one, the number goes up to 3^5, or 243, which makes the cost and time of testing prohibitively high. In both cases, the alternative is a fractional factorial design. As shown in Box, Hunter and Hunter [12, p. 410], the minimum number of specimens, n , for a two-level k-factor fractional factorial design is below: (Note: n is always equal or greater than k+1.)

k =	3	4	5	6	7	8	9 ... 15	16 ... 31
Min. no. of specimens, n =	4	8	8	8	8	16	16 ... 16	32 ... 32

Thus we advised Team A that with five factors, the minimum number of additional tests we needed to get credible estimate of the uncertainty was 8 , and Team A agreed. At that point, we were ready to construct a 5-factor, 9-run fractional factorial DOE with the experimental data (*Y1* = 19.31 ft-lb) as the center point and 8 additional tests yet to be performed as the corners of a 5-dimensional hypercube.

The last question we asked would be the hardest. We asked Team A to commit to the numerical values of the high and low settings of each factor with a request that they produce 8 specimens to the precise specifications of those ten settings. For example, if they thought the manganese sulfide content could vary between 0.05 % and 0.15 % around the center point of 0.10 %, then they had to work hard to make specimens with those contents. Let us imagine that, for this example of a fictitious experiment, Team A were most cooperative in responding to our "tough" request. The settings they considered to be feasible and were willing to commit to were as follows:

Factor	*Title (Unit)*	*Low*	*Center*	*High*
X1	Temperature (deg. °F)	80.0	120.0	160.0
X2	Manganese Sulfide (%)	0.05	0.10	0.15
X3	Initial Strain (%)	1.0	2.0	3.0
X4	Mis-orientation (degree)	1.0	8.0	15.0
X5	Notch radius (mm)	0.23	0.25	0.27

We immediately told Team A to get busy in order to run those 8 tests and we gave Team A the design matrix as shown in the upper left corner of Fig. 14. For example, the first row of the matrix, (-1, -1, -1, +1, +1), corresponded to one of the 8 tests Team A had to do with the following specifications:

Temp.	*Mn Sulfide*	*Init. Strain*	*Mis-orientation*	*Notch Radius*
80 F	0.05 %	1.0 %	15.0 deg.	0.27 mm

After several days, Team A reported results of their 8 additional tests as shown in the leftmost column of the matrix table in Fig. 14. Note that we have placed the center point values in red at the bottom of that matrix.

Using a 10-step exploratory-data-analysis routine of DATAPLOT [25], we obtained results in the form of plots and tables as shown in Figs. 14 through 19. In particular, Fig. 14 is a display of the virtual data (consisting of 8 values of *Y1*) as an ordered set.

In Fig. 15, we display the results of fitting the empirical first order model, eq. (12), with the coefficients, β_0 , β_1, β_2, β_3 , β_4, β_5, β_{12} , β_{13} , etc., given in the upper left (for β_0) and right (for the rest) cornrs of the plot. This plot shows a ranking of the absolute values of all one-way and two-way effects of the virtual data we envisioned for a 5-factor, 8-run DOE. Note that the two two-term interactions, *X2*X3* and *X2*X5* are extremely small in comparison with the other one-

6. (Continued)

term effects. Moreover, we observe the following confounding phenomena:

Main Effect	Confounding Factors
X4	*X1*X2*
X1	*X2*X4, X3*X5*
X2	*X1*X4*
X5	*X1*X3*

This means we had to be careful in choosing which two factors among the five as "key parameters" for constructing a simpler model. Let us examine our results with some more plots. Fig. 16 is a main effects plot where *X1* and *X4* are clearly shown as dominant, with *X2* a close third. Fig. 17 is an interaction effects plot where we verify that 5 of the 10 two-term interactions, *X1*X2, X1*X3, X1*X4, X2*X4,* and *X3*X5,* appear to be confounding. By observing that *X1* appears in the confounding terms more often than *X2*, and the fact that the main effects of *X3* and *X5* are much smaller than the three dominant ones, we conclude it is reasonable for us to choose *X4* and *X1* as candidates for simplification.

Fig. 18 is a 2-parameter least square fit for *X4* and *X1*, and Fig. 19 is an uncertainty analysis, where the predicted mean and the 95 % confidence intervals are reported. Note that the analysis predicted an estimated mean (18.996) different from the initial value (19.13) of the center point.

Conclusion: The real-plus-virtual 9-run DOE yielded an estimate for the *CVN* at 48.9 °C (120 °F) to be **18.996 (12.787)**, or 18.996 ± 12.787 ft-lb, with 95 % confidence.

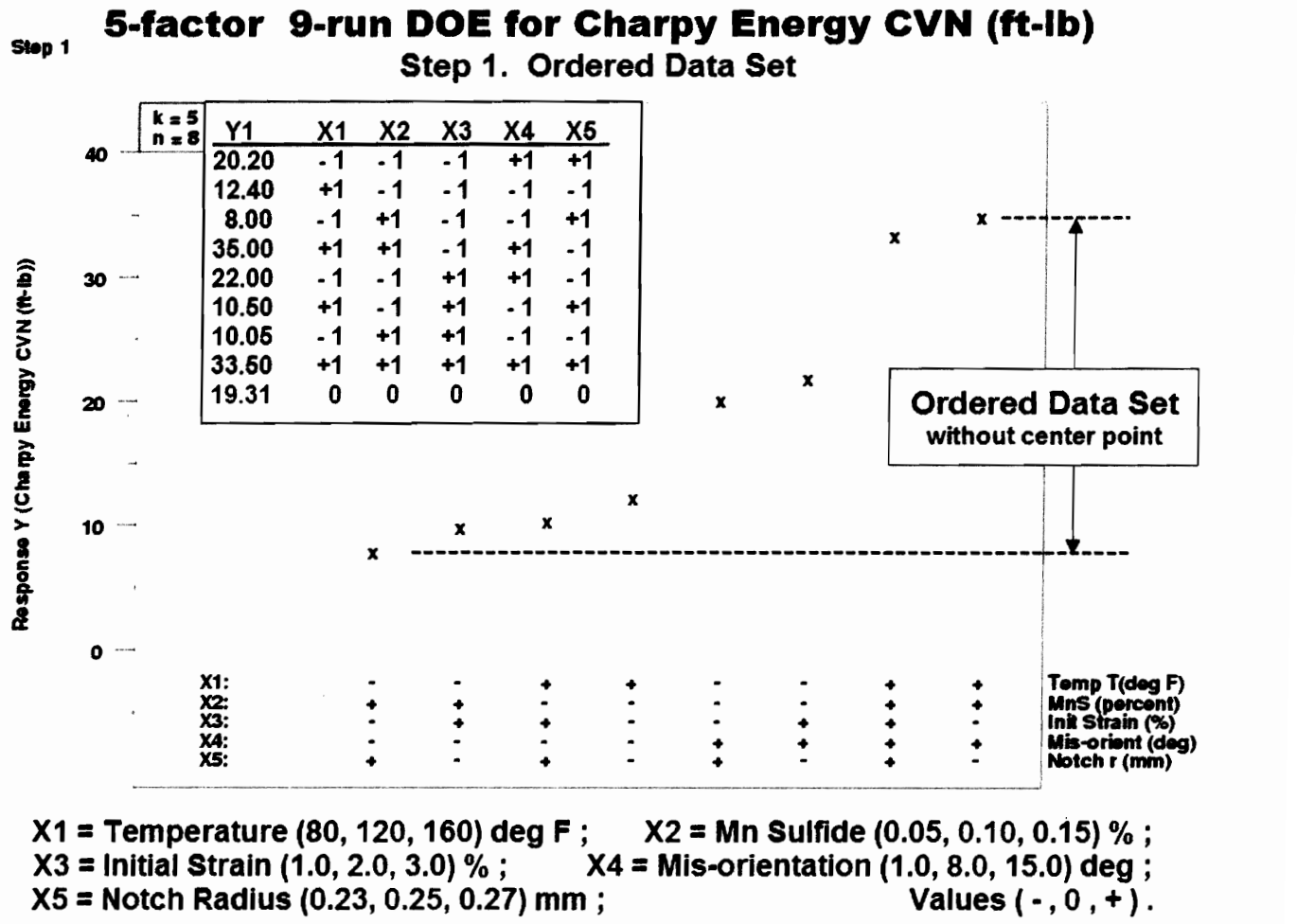

Fig. 14. First of 10 plots by DATAPLOT [25] showing a fictitious Charpy V-notch impact test data set in the upper left corner as an ordered set for a fractional factorial orthogonal design of experiments (DOE, 5-factor, 8-run-plus-center-point). Note the table at the bottom of the plot being the transposed DOE matrix with re-ordered columns.

6. (Continued)

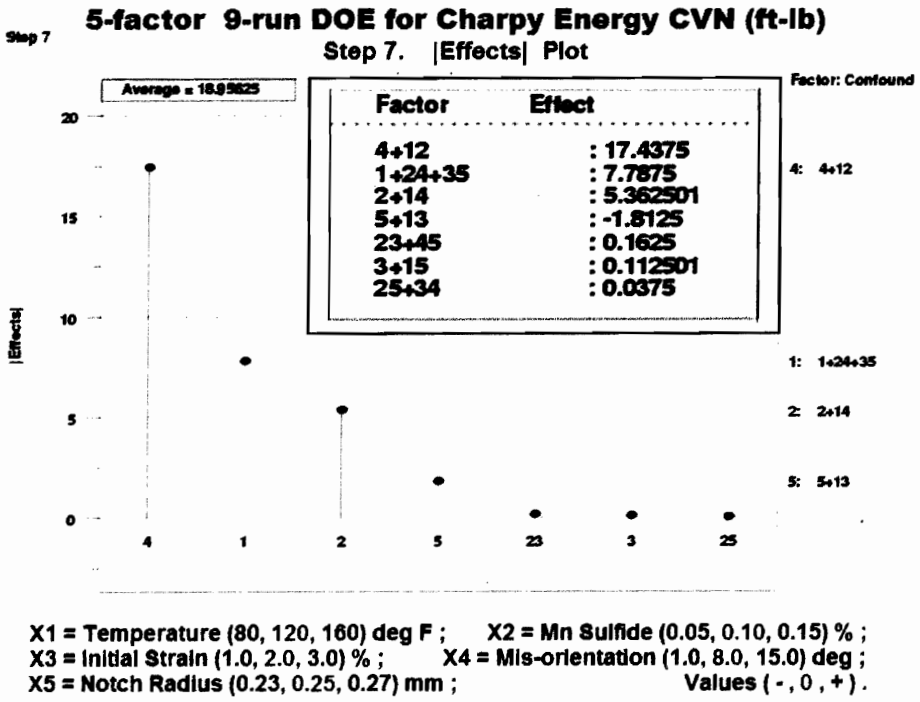

Fig. 15. Seventh of 10 plots by DATAPLOT [25] showing a ranking of the absolute values of all one-way and two-way effects of a fictitious Charpy V-notch impact test data set for a fractional factorial orthogonal design of experiments (DOE, 5-factor, 8-run-plus-center-point). In particular, we observe the confounding of a two-way interaction, *X1-X2*, in the dominant factor, *X4* , and two two-way interactions, *X2-X4* and *X3-X5*, in the second dominant factor, *X1*. A competing third dominant factor, *X2* (MnS content), is a surprise as discussed in the text.

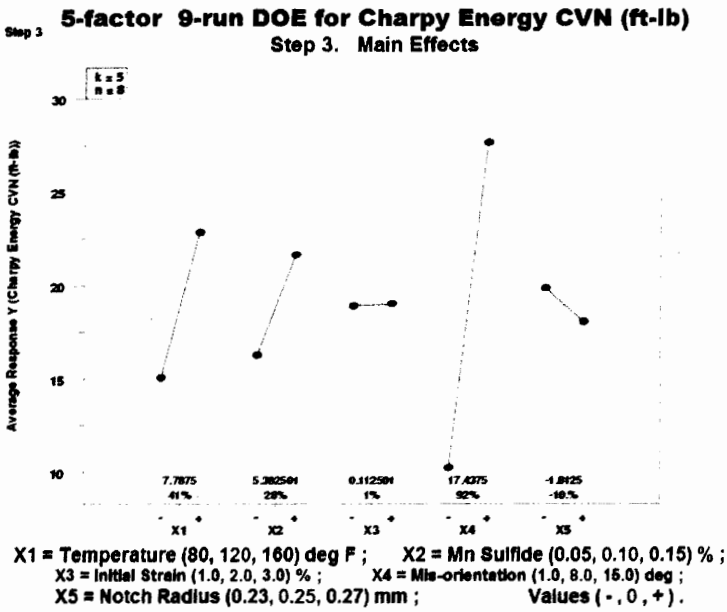

Fig. 16. Third of 10 plots by DATAPLOT [25] showing the main effects of a fictitious Charpy V-notch impact test data set for a fractional factorial orthogonal design of experiments (DOE, 5-factor, 8-run-plus-center-point). Note the dominance of *X1* and *X4*.

6. (Continued)

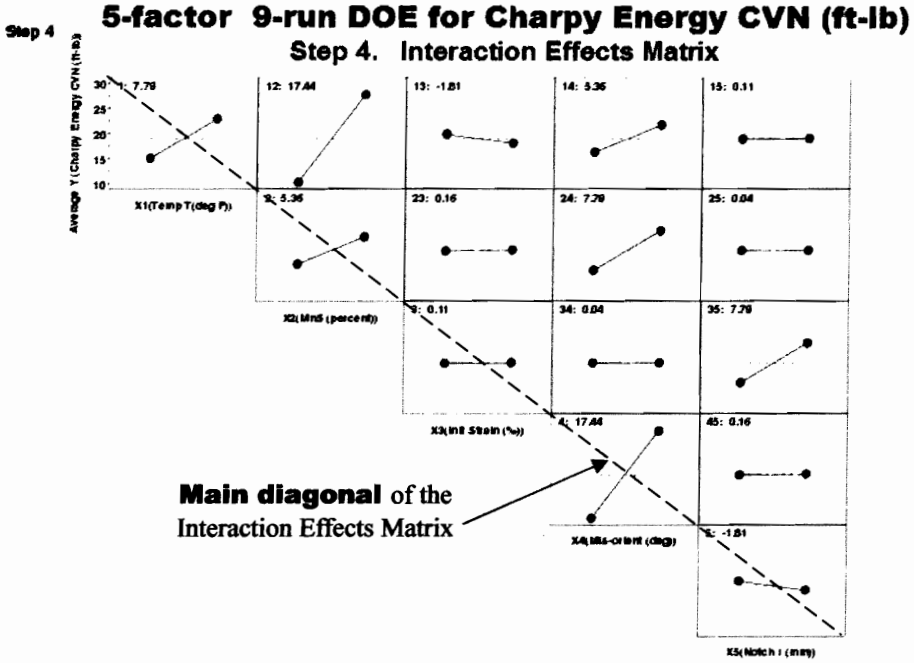

Fig. 17. Fourth of 10 plots by DATAPLOT [25] showing the interaction matrix of a fictitious Charpy V-notch impact test data set for a fractional factorial orthogonal design of experiments (DOE, 5-factor, 8-run-plus-center-point). The boxes on the main diagonal are re-plots of the results of **Fig. 16**. Note the presence of several two-way interactions in this plot, namely, *X1-X2, X1-X3, X1-X4, X2-X4,* and *X3-X5*.

```
                    LEAST SQUARES MULTILINEAR FIT

SAMPLE SIZE N       =        9
NUMBER OF VARIABLES =        2

REPLICATION CASE
REPLICATION STANDARD DEVIATION =          1.289501
REPLICATION DEGREES OF FREEDOM =             4
NUMBER OF DISTINCT SUBSETS     =             5

        PARAMETER ESTIMATES              (APPROX. ST. DEV.)     T VALUE
 1  A0                      18.9956       ( 1.091    )         17.41
 2  A1       X1             3.89375       ( 1.157    )         3.365
 3  A2       X4             8.71875       ( 1.157    )         7.534

RESIDUAL    STANDARD DEVIATION =          3.273003
RESIDUAL    DEGREES OF FREEDOM =             6

REPLICATION STANDARD DEVIATION =          1.2895007133
REPLICATION DEGREES OF FREEDOM =             4

LACK OF FIT F RATIO =      17.3273 = THE  98.9292% POINT OF THE
     F DISTRIBUTION WITH       2 AND       4 DEGREES OF FREEDOM
```

Fig. 18. Sensitivity analysis of the 5-factor, 8-run-plus-center-point fictitious Charpy V-notch impact test data set using a two-parameter (*X1, X4*) least square fit routine of DATAPLOT [25].

6. (Continued)

```
                    Uncertainty Analysis

A0                                        = 18.99556
A1                                        = 3.89375
A2                                        = 8.71875

Resdiual SD                               = 3.273003
Resdiual DF                               = 6

Variance(Y)                               = 24.5799
SD(Y)                                     = 4.957812

Upper 95% Confidence Bound for Y          = 31.7831
Lower 95% Confidence Bound for Y          = 6.208013
```

Fig. 19. Results of an uncertainty analysis generated by DATAPLOT 10-step code [25] showing that the predicted mean (*A0*) of the fictitious Charpy V-notch data set is 18.996 ft-lb and the estimated variance is **24.58,** which will be used in computing variance of K_{Ic}.

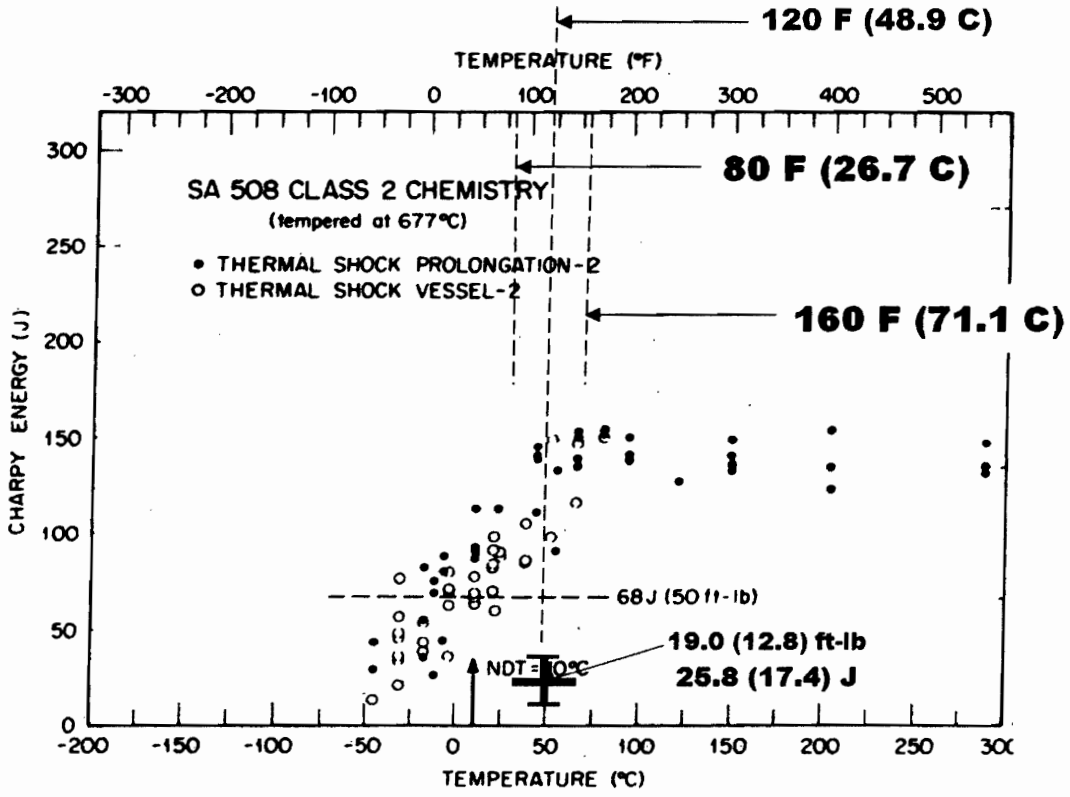

Fig. 20. Comparison of the estimated mean and standard deviation of Charpy V-notch energy at 120 °F (48.9 °C) based on a 5-factor, 9-run DOE-generated *fictitious* test data for ASTM A517 Grade H steel plate (620 MPa min. room temperature yield strength), with Charpy energy data of a thermal shock experiment test cylinder, TSE-5A, of comparable yield strength (see **Fig. 1** based on Cheverton et al [4]). Note that none of the TSE-5A data (black dots and circles) was plotted with an expression of uncertainty as the fictitious DOE-generated result, 19.0 (12.8) ft-lb, or, 25.8 (17.4) J, plotted in red.

6. *(Continued)*

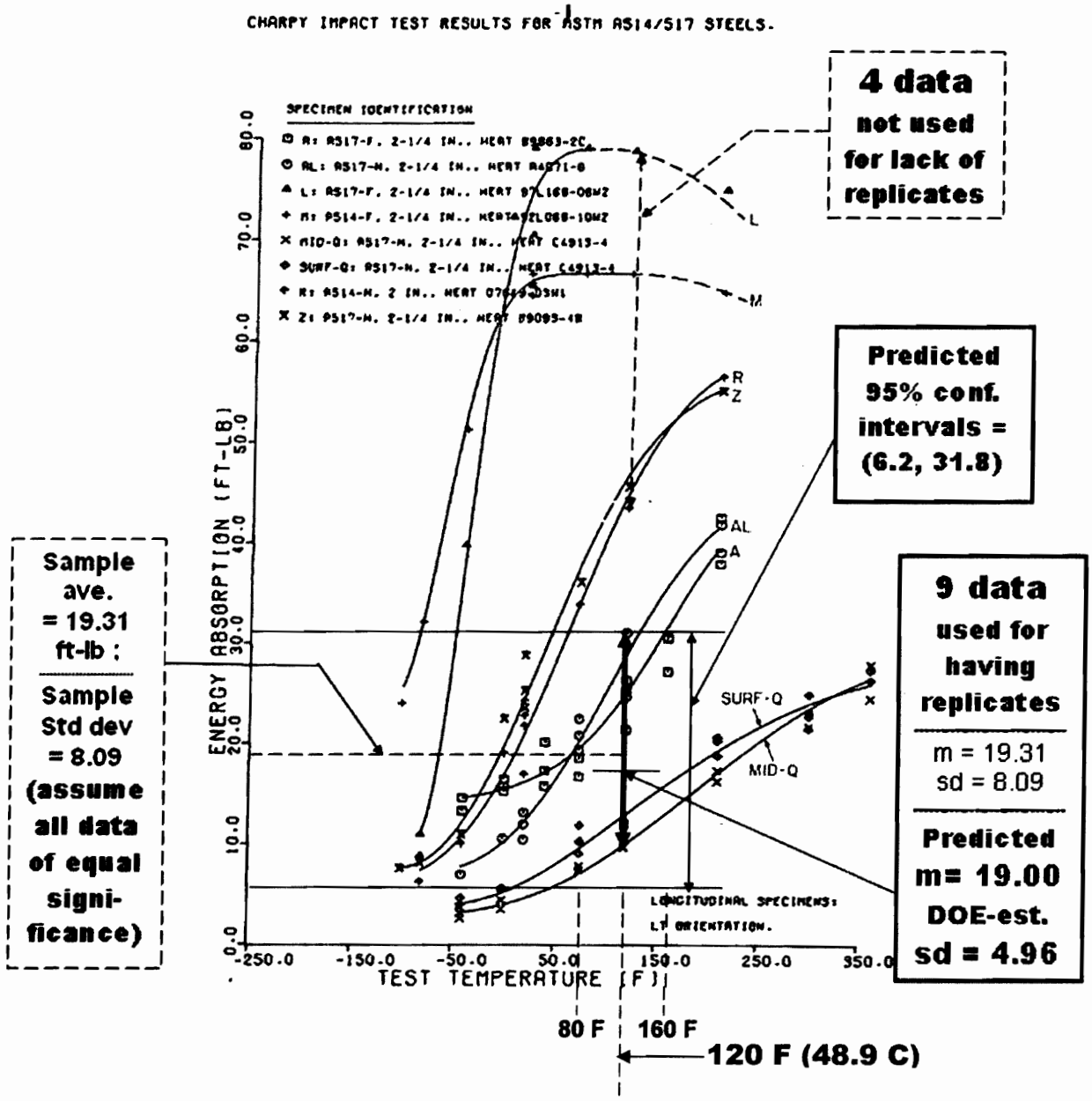

Fig. 21. Comparison of the estimated mean, standard deviation, and predicted 95 % confidence intervals of Charpy V-notch energy at 120 °F (48.9 °C) based on a 5-factor, 9-run DOE-generated *fictitious* test data for ASTM A517 Grade H steel plate (620 MPa min. room temperature yield strength), with the 3-heat, 9-specimen with replicates, Charpy V-notch energy data at 120 °F (48.9 °C) as reported by Interrante and Hicho in 1973 [10] and reproduced in Tables 3-6 and Fig. 7 of this paper. Note that if the 9 experimental data points of 1973 were treated with equal significance, the sample average and standard deviation, 19.31 and 8.09, respectively, differ from those estimated using the fictitious DOE-generated data.

7. A METHODOOGY FOR ESTIMATING VARIANCE OF FRACTURE TOUGHNESS (K_{IC})

The estimated Charpy energy at 48.9 °C (120 °F) of the last section, rounded to one decimal, equals 19.0 (12.8) ft-lb, or **25.8 (17.4)** J, and is plotted in Fig. 20 as a data point with an error bar to illustrate its uniqueness among a collection of experimental data reported by Cheverton et al [4]. In addition, the fictitious 5-factor, 9-run DOE result is plotted in Fig. 21 as an overlay on Fig. 7 to illustrate the value of a 9-replicate experiment involving three heats as represented by factor X2 (manganese sulfide). Instead of labelling each heat with the Charpy energy data as reported by Interrante and Hicho [10], Team A in this fictitious experiment went to the chemical composition of each heat. The resulting estimate is therefore more convincing in explaining the heat-to-heat variability with a physical basis.

Let us return to eq. (4) of Section 3 to estimate the variance of K_{Ic} from the estimated variances of σ_y and *CVN*. Since eq. (4) is strictly valid only for the units of σ_y and *CVN* to be ksi and ft-lb, respectively, we shall use the following estimates of σ_y and *CVN* and their variances to calculate the variance of K_{Ic} (ksi-in$^{1/2}$) at 120 °F (48.9 °C):

σ_y = 90.0 ksi , sd (σ_y) = 7.10 ksi , Var(σ_y) = 50.41 (ksi)^2 ,

CVN=19.0 ft-lb, sd(*CVN*)= 4.96 ft-lb, Var(*CVN*)= 24.60 (ft-lb)^2.

Substituting the above into eq. (4), we obtain Var(K_{Ic}) = 140.20 , which implies sd (K_{Ic}) = 11.84 ksi-in$^{1/2}$. Using the Wullaert-Server formula [16] to calculate K_{Ic} , we also obtain K_{Ic} = 86.84 ksi-in$^{1/2}$. Since the number of specimens of the yield strength data [22] is 21, and that of the fictitious Charpy energy data is 9 , we choose the smaller of the two sample sizes to find the critical value of the *t* distributions, i.e., t(0.025, 9) = 2.262. The half-interval for K_{Ic} equals 2.262 * 11.84 = 26.78, so the estimated value of K_{Ic} at 120 °F (48.9 °C) with an expression of uncertainty is 86.84 (26.78) ksi-in$^{1/2}$, or, **95.48 (29.43)** MPa-m$^{1/2}$. A plot of this estimate with an error bar is given in Fig. 22, which is an overlay of Fig. 3 (after Cheverton et al [4]).

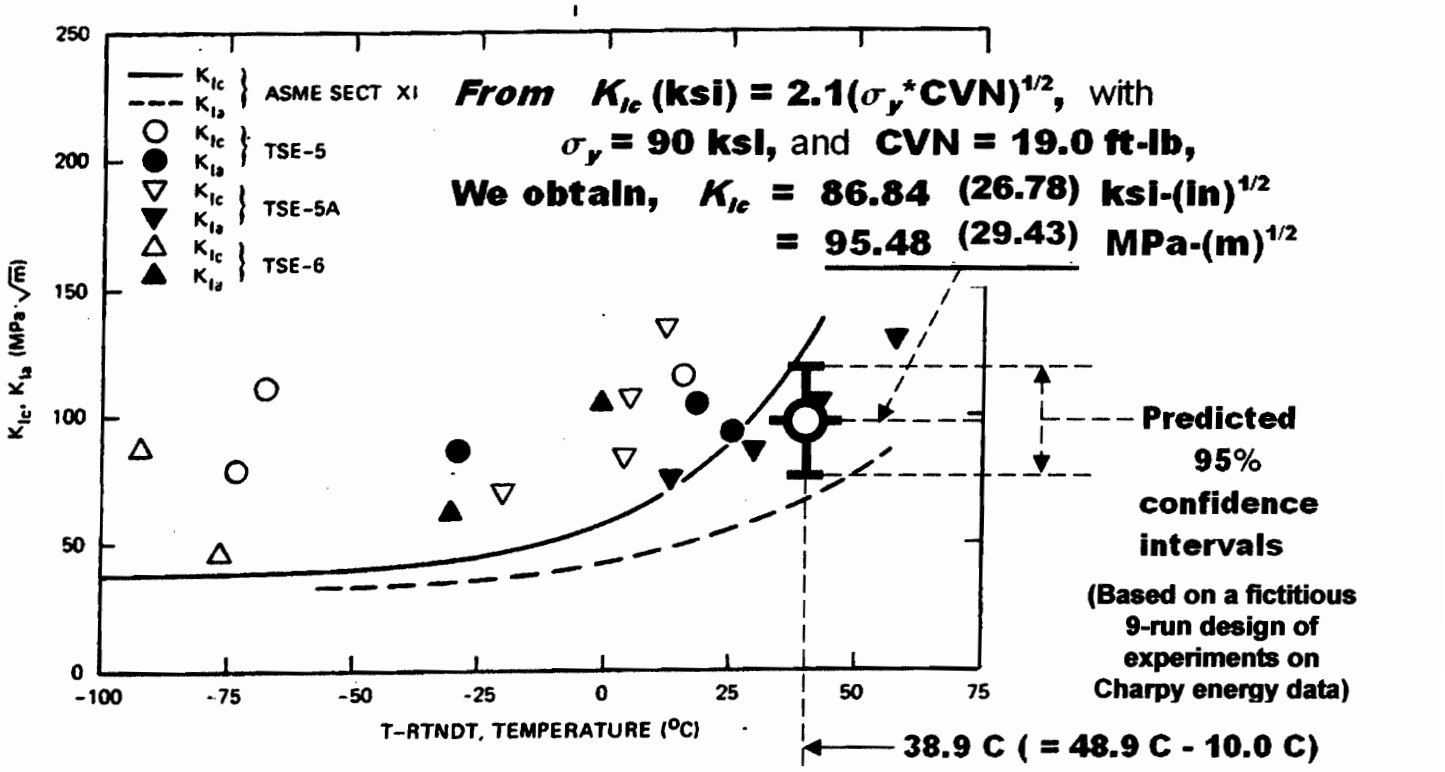

Fig. 22. Plot of an estimated static crack initiation toughness (K_{Ic}) value with an expression of uncertainty (error bar in red) based on *fictitious* design-of-experiments(**DOE**)-generated results at 120 °F (48.9 °C), in a *K* vs. (*T* - RTNDT) diagram where K_{Ic} and K_{Ia} data from three thermal shock experiment (TSE) test cylinders, TSE-5, 5A, and 6, and ASME Section XI K_{Ic} and K_{Ia} curves over a broad range of temperature shift, (*T* - RTNDT), were plotted by Cheverton et al [4] (see **Fig. 3** in this paper). Note that all experimental data or design curves are for comparable steels having an room temperature yield strength of about 90 ksi (620.6 MPa).

8. SIGNIFICANCE AND LIMITATIONS

This paper introduces two methodologies for estimating (a) the Charpy V-notch impact energy, *CVN*, with an expression of uncertainty, and (b) the static fracture initiation toughness, K_{Ic} , also with an expression of uncertainty if (a) is known together with the variance of either the yield strength, σ_y, or the Young's modulus E.

The first methodology uses the concept of a statistical design of experiments (DOE) that calls for an engineer to use experience and judgment in planning a series of tests to overcome the problem of Charpy data scatter due to heat-to-heat variability. The second extends the results of the first by introducing error propagation formulas and a functional relationship between K_{Ic} and the combination of σ_y and *CVN*.

Both methodologies are significant in accomplishing two objectives in engineering and materials science, namely, to resolve the data scatter problem of a material characterization task such as the Charpy test, and to report the results of such a task with an error bar. As shown by Fong, et al in two recent articles on failure analysis [32, 33], the availability of an error bar in material property measurements such as yield strength, ultimate tensile strength, buckling strength, and fracture toughness provides a basis for using a non-deterministic and physically more realistic model to simulate progressive weakening and partial or complete failure of a structure or component in an abnormal or aging-related collapse scenario.

Clearly, the two methodologies as applied to Charpy energy and fracture toughness estimation are not without limitations. For the first methodology, a primary limitation stems from the difficulty of reducing a large list of test-related factors to a small and manageable number. The second and perhaps more serious limitation is the judgment-based requirement of high and low settings of each factor, making the experimental runs more difficult and the analysis outcome less rigorous when the settings were not achieved.

For the second methodology, the user needs to be aware that the validity of the approximate error propagation formulas requires the variables to be statistically independent. Furthermore, the functional form of the fracture toughness has a limited range of validity and is not unique, because the formulas given by ASM [12] came from correlations of experimental data of specific materials, specimen geometries and loading configurations. Nevertheless, as long as the user is aware of those limitations and interprets the results of the analysis accordingly, we believe the two methodologies are useful additions to the engineer's tool box, as illustrated by several other applications [34, 35] scheduled for presentation at this conference.

9. CONCLUDING REMARK

By introducing a statistical concept of the design of experiments (DOE), we develop a methodology to address the data scatter problem of the Charpy V-notch impact test, which is largely due to the heat-to-heat variability of the test specimens. We illustrate the DOE-base methodology by an example using a fictitious set of experimental data and succeded in obtaining an estimate of the Charpy energy with an expression of uncertainty (error bar).

By using error propagation formulas in the statistics literature, we successfully extend the results of the DOE-based methodology to an estimation of the variance of the static fracture initiation toughness, K_{Ic}. This second methodology also leads to an estimation of K_{Ic} with an expression of uncertainty (error bar).

10. ACKNOWLEDGMENTS

The authors wish to thank Roland deWit and Richard J. Fields of National Institute of Standards and Technology, Gaithersburg, MD, and Poh-Sang Lam and Karthik Subramanian of Savannah River National Laboratory, Aiken SC, for technical assistance and discussion during the course of preparing this paper.

11. REFERENCES

[1] Bush, S. H., and Hedden, O. F., 2007, "Flaw Detection, Location, and Sizing," Proc, Digest of ASME PVP Symposium on Engineering Safety, Applied Mechanics, and Non-destructive Evaluatioon (NDE), in honor of Dr. Spencer H. Bush (1920-2005), July 2007, San Antonio, TX, J. T. Fong and O. F. Hedden, Editors, pp. 153-156. Published by Stanford Mechanics Alumni Club, 104 King Farm Blvd, Suite C308, Rockville, MD 20850 (2007).

[2] Cheverton, P. D., Canonico, D. A., Iskander, S. K., Bolt, S. E., Holtz, P. P., Nanstad, R. K., and Stelzman, W. J., 1983, "Fracture Mechanics Data Deduced from Thermal Shock and Related Experiments with LWR Pressure Vessel Material," Journal of Pressure Vessel Technology, Vol. 105, pp. 102-110 (1983).

[3] Loss, F. J., Gray, R. A., Jr., and Hawthorne, J. R., 1979, "Investigation of Warm Prestress for the Case of Small ΔT During a Reactor Loss-of-Coolant Accident," Journal of Pressure Vessel Technology, Vol. 101, pp. 298-304 (1979).

[4] ASME, 1983, "Rules for Inservice Inspection of Nuclear Power Plant Components," ASME Boiler and Pressure Vessel Code, Section XI. The American Society of Mechanical Engineers, United Engineering Center, New York, NY (1983).

[5] deWit, R., Read, D. T., Low III, S. R., Harne, D. E., McColskey, D. J., Hicho, G. E., Smith, L. C., Danko, G., and Fields, R. J., 1988, "Wide Plate Crack Arrest Testing: A Description and Discussion of the First Two Wide Plate Tests and the Results of Six, Full Thickness, Bend Bar Tests," U. S. National Bureau of Standards Report NBSIR 87-3629, Feb. 1988. Available for purchase from Govt. Printing Office, P. O. Box 37082, Washington, DC 20013-7982, and National Technical Information Service, Springfield, VA 22161 (1988).

[6] deWit, R., Fields, R. J., and Irwin, G. R., 1993, "Use of Thickness Reduction to Estimate Fracture Toughness," Constraint Effects in Fracture, ASTM STP 1171, E. M. Hackett, K. H. Schwale, and R. H. Dodds, Eds., pp. 361-382. American Society for Testing and Materials, Philadelphia, PA (1993).

[7] ASTM Designation: E23-98, 1999, "Standard Test Methods for Notched Bar Impact Testing of Metallic Materials," Annual Book of ASTM Standards, Section 3, Metals Test Methods and Analytical Procedures, Vol. 03.01, Metals - Mechanical Testing; Elevated and Low-Temperature Tests; Metallography. ASTM, 100 Barr Harbor Dr., West Conshohocken, PA 19428-2959 (1999).

[8] Kanninen, M. F., and Popelar, C. H., 1985, Advanced Fracture Mechanics. Oxford University Press (1985).

[9] Dowling, N. E., 1999, Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue, 2nd ed., pp. 184-190. Prentice-Hall (1999).

[10] Interrante, C. G., and Hicho, G. E., 1973, "Fracture Toughness Evaluation of Quenched-and-Tempered Bridge Steels," U. S. National Bureau of Standards Report NBSIR 73-293, Oct. 1973. Available for purchase from Govt. Printing Office, P. O. Box 37082, Washington, DC 20013-7982, and National Technical Information Service, Springfield, VA 22161 (1973).

[11] ASTM Designation: A 517/A 517M-93, 1993, "Standard Specification for Pressure Vessel Plates, Alloy Steel, High-Strength, Quenched and Tempered," Annual Book of ASTM Standards, Vol. 01.04. ASTM, 100 Barr Harbor Dr., West Conshohocken, PA 19428-2959 (1993).

[12] ASM International, 1990, Mechanical Testing, Metals Handbook, Vol. 8, 10th ed., pp. 264-265. American Society for Metals (1990).

[13] Rolfe, S. T., and Novak, S. R., 1970, "Slow Bend KIc Testing of Medium-Strength High-Toughness Steels," in Review of Developments in Plane Strain Fracture Toughness Testing, ASTM STP 463, pp. 124-159. ASTM, Philadelphia, PA, (1970).

[14] Sailors, R. H., and Corten, H. T., 1972, "Relationship Between Material Fracture Toughness Using Fracture Mechanics and Transition Temperature Tests," in Fracture Toughness, Proc. 1971 National Symposium on Fracture Mechanics, STP 514, Part II, pp. 164-191. ASTM, Philadelphia, PA (1972).

[15] Marandet, B., and Sanz, G., 1977, "Evaluation of the Toughness of Thick Medium-Strength Steels by Using Linear Elastic Fracture Mechanics and Correlations Between K_{Ic} and Charpy V-Notch," in Flaw Growth and Fracture, STP 631, pp. 72-95. ASTM, Philadelphia, PA (1977).

[16] Wullaert, R. A., 1978, "Fracture Toughness Predictions from Charpy V-Notch Data," in What Does the Charpy Test Really Tell Us?, Proc. Amer. Inst. Mining, Metallurgical and Petroleum Engineers Conference, Denver, 1978. American Society for Metals (1978).

[17] Barsom, J. M., and Rolfe, S. T., 1979, "Correlations Between K_{Ic} and Charpy V-Notch Test Results in the Transition Temperature Range," in Impact Testing of Materials, STP 466, pp. 281-302. ASTM, Philadelphia, PA (1979).

[18] Birge, R. T., 1939, "The propagation of errors," The American Physics Teacher, Vol. 7, No. 6 (Dec. 1939).

[19] Ku, H. H., 1966, "Notes on the use of propagation of error formulas," NBS Journal of Research, Vol. 70C, No. 4, pp. 263-273 (1966).

[20] Nelson, L. S., 1992, "Propagation of Error," Journal of Quality Technology, Vol. 24, No. 4, pp. 232-235 (1992).

[21] ASM International, 1990, Properties and Selection: Irons, Steels, and High Performance Alloys, Metals Handbook, Vol. 10, 10th ed. American Society for Metals (1990).

[22] NRIM, 1994, "Data Sheets on the Elevated Temperature Properties of High Strength Steel (Class 590 MPa) Plates for Pressure Vessels," NRIM Creep Data Sheet, Second Revised (No. 25B), 30 September 1994. Tokyo: National Research Institute for Metals (1994).

[23] Shapiro, S. S., and Wilk, M. B., 1965, "An Analysis of Variance Test for Normality (Complete Samples)," Biometrika, Vol. 52, pp. 591-611 (1965).

[24] Anon., 1995, "Wilks Shapiro Normality Test Algorithm AS R94 (SWILK sub routine)," Applied Statistics Journal, Vol. 44, No. 4 (1995).

[25] Filliben, J. J., and Heckert, N. A., 2002, DATAPLOT: A Statistical Data Analysis Software System, a public domain software released by NIST, Gaith, MD 20899, http://www.itl.nist.gov/div898/software/dataplot.html (2002).

[26] Draper, N. R., and Smith, H., 1981, Applied Regression Analysis, 2nd ed. Wiley (1981).

[27] Nelson, P. R., Coffin, M., and Copeland, K. A. F., 2003, Introductory Statistics for Engineering Experimentation. Elsevier (2003)

[28] Box, G. E., Hunter, W. G., and Hunter, J. S., 1978, Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building. Wiley (1978).

[29] Montgomery, D. C., 2000, Design and Analysis of Experiments, 5th ed. Wiley (2000).

[30] Croarkin, C., Guthrie, W., Heckert, N. A., Filliben, J. J., Tobias, P., Prins, J., Zey, C., Hembree, B., and Trutna, eds., 2003, NIST/SEMATECH e-Handbook of Statistical Methods, Chap. 5 on Process Improvement, http://www.itl.nist.gov/div898/handbook/, first issued, June 1, 2003, and last updated July 18, 2006. Produced jointly by the Statistical Engineering Division of the National Institute of Standards & Technology, Gaithersburg, MD, and the Statistical Methods Group of SEMITECH, Austin, TX. Also

available as a NIST Interagency Report in a CD-ROM upon request to alan.heckert@nist.gov (2006).

[31] Fong, J. T., Filliben, J. J., Heckert, N. A., and deWit, R., 2008, "Design of Experiments Approach to Verification and Uncertainty Estimation of Simulations based on Finite Element Method," to appear in Proceedings of the Annual Conference of the American Society for Engineering Education, June 22-25, 2008, Pittsburgh, PA, Paper No. AC2008-2725 (2008).

[32] Fong, J. T., Filliben, J. J., deWit, R., Fields, R. J., Bernstein, B., and Marcal, P. V., 2006, "Uncertainty in Finite Element Modeling and Failure Analysis: A Metrology-Based Approach," ASME Trans., J. Press. Vess. Tech., Vol. 128, pp. 140-147 (2006).

[33] Fong, J. T., Filliben, J. J., and Heckert, N. A., deWit, R., and Bernstein, B., 2008, "Robust Engineering Design for Failure Prevention," Proc. ASME PVP Conference, July 27-31, 2008, Chicago, IL, Paper No. PVP2008-61602. New York, NY: ASME (2008).

[34] Fong, J. T., Hedden, O. F., Filliben, J. J., and Heckert, N. A., 2008, "A Web-based Data Analysis Methodology for Estimating Reliability of Weld Flaw Detection, Location, and Sizing," Proc. ASME PVP Conference, July 27-31, 2008, Chicago, IL, Paper No. PVP2008-61612. New York, NY: ASME (2008).

[35] Chao, Y., Lam, P. S., Fong, J. T., deWit, R., and Filliben, J. J., 2008, "A New Approach to Assessing the Reliability of Applying Laboratory Fracture Toughness Test Data to Full-Scale Structures," to appear in Proc. ASME PVP Conference, July 27-31, 2008, Chicago, IL, Paper No. PVP2008-61584. New York, NY: ASME (2008).