Inaccuracy and Variability of Shot Peening Measurements

Dr. David Kirk is a regular contributor to The Shot Peener. Since his retirement, Dr. Kirk has been an Honorary Research Fellow at Coventry University, U.K. and is now Visiting Professor in Materials, Faculty of Engineering and Computing at Coventry University.

INTRODUCTION

Every measurement ever made of a shot peening parameter has been inaccurate – to a greater or lesser extent! Every shot peening parameter varies – to a greater or lesser extent. Accuracy and variability have a powerful effect on the controllability of shot peening. They cannot, or at least should not, be ignored. Inaccuracy is the difference between a measured value and the true value. Variability is the extent to which a set of measurements deviates from its mean (average) value. Specified tolerance bands allow, however, for both inaccuracy and variability of parameters.

Three primary factors contribute to the inaccuracy and variability of shot peening measurements:

- Instrument Inaccuracy,
- Measurement Variability and
- Parameter Variability.

These three factors interact with one another – as shown schematically in Fig.1.

INACCURACY

Accuracy is often taken for granted. The emphasis in this article is therefore on inaccuracy. A simple equation connects inaccuracy with true and measured values:

True value - Measured value = Inaccuracy (1)

A measured value will differ from the true value in two respects: **Precision** and **Bias**. Precision is the last significant digit of the instrument's scale, e.g., this might be 1psi for an air pressure gage. Bias is the difference between the indicated value and the true value, e.g., if the pressure gage indicated 88psi when the true value was 91.00psi then the instrument bias would be 3psi.

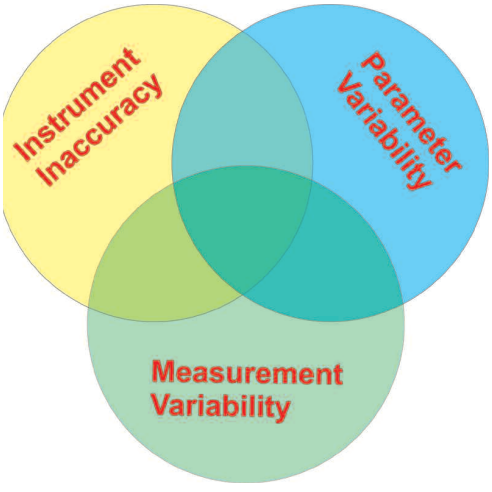

Fig.1 Interactions of Instrument Inaccuracy, Measurement Variability and Parameter Variability.

Precision

Precision is important because it determines how close the instrument's reading can possibly be to the true value for a parameter. Fig.2 is a schematic illustration of the effect of low precision on inaccuracy of Almen arc height measurements. Assume (a) that a given gage reads to the nearest thousandth of an inch (b) that the gage has zero bias and (c) that the true value for the arc height of a particular sample is 9.325×10^{-3} inch - to the nearest millionth of an inch. For this particular example there is a precision inaccuracy of 0.325 – the gage displaying 9×10^{-3} when the true value is 9.325×10^{-3}. The true value could, in fact, have been anywhere between 8.500 and 9.499 and this gage would still have displayed 9 as the arc height – so that the maximum precision inaccuracy is 0.5 (for other samples).

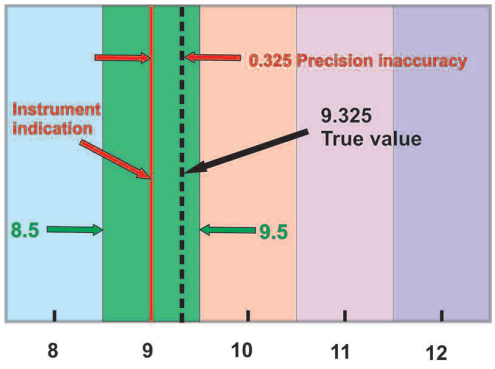

Fig.2 Inaccuracy of Arc Height Measurement caused by low instrument precision.

Modern digital Almen gages have a precision that is better than one thousandth of an inch. Fig.3 illustrates the reduction of precision inaccuracy, for the same specimen, because the gage is more precise - to one ten-thousandth of an inch. The precision inaccuracy is now, for this example, only 0.025 - as compared with 0.325 for the previous gage. True values could lie anywhere between 9.25 and 9.35 and this gage would still have displayed 9.3 as the arc height – so that the maximum precision inaccuracy is then 0.05 (for other samples).

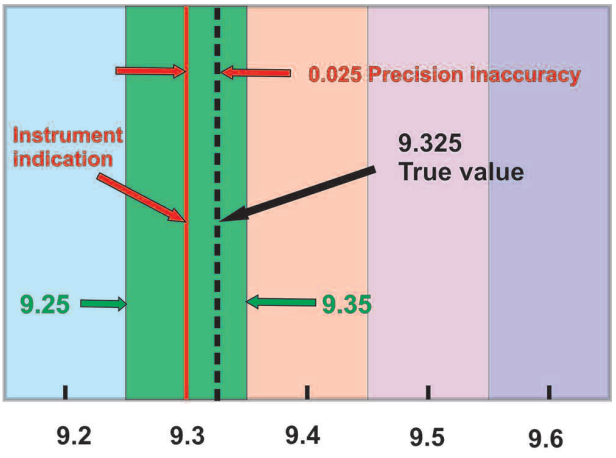

Fig.3 Reduction of inaccuracy of Arc Height Measurement by using improved instrument precision.

There is generally an optimum level of instrument precision for any given application. For example, it would be ludicrous to use scales precise to the nearest milligram when weighing shot to fill bags to nominally 50kg. Scale precision is related to maximum capacity so that typical milligram scales would have a maximum capacity of just over 100g. Such scales would have to be used about 500 times just to fill a single 50kg bag – multiplying the cost of the shot to the customer. Single weighings on a scale precise to the nearest gram would offer more than adequate accuracy.

Bias

Bias is the difference between an instrument's indicated value and a true value. This can only be detected if the bias is greater than the instrument's precision. If the bias is greater than the instrument's precision then it will have a significant effect on accuracy. The degree of bias generally changes over the range of a given instrument. Reference specimens, i.e., specimens with known true values, are needed in order to detect and determine the amount of bias.

Consider the following test question: "An Almen gage reads 10.3 (thousandths of an inch) every time a reference specimen having an arc height of 10.000 is placed on the gage. What is the bias of the Almen gage?". A quick, inaccurate, answer would be "0.3". The correct answer would be "At the moment it is somewhere between 0.25 and 0.35". "At the moment" is appropriate because the difference of 0.3 might change with time – instrument instability could be a factor. Fig.4 shows why the bias, for this hypothetical example, lies somewhere between 0.25 and 0.35 and is not precisely 0.3. The gage would 'round' any value

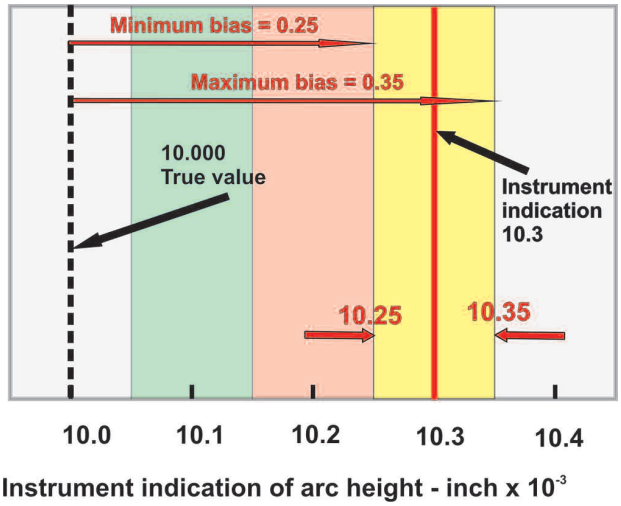

Fig.4 Example of Almen Gage bias lying between 0.25 and 0.35 (thousandths of an inch).

between 10.25 and 10.35 to its nearest precision value – 10.3.

If a bias of, for example, 0.25 to 0.35 was left uncorrected then it would have a significant effect on the accuracy of indicated arc heights for peened strips.

Bias can vary over the available range of any given instrument. Weighing scales are perhaps the easiest for detecting bias over a scale's range. Table 1 gives the measurements obtained by using a set of calibrated applied masses on a 50g capacity "Digital Pocket Scale". The scale was advertised as having an "Accuracy: ± 0.01g" and as having "Auto Calibration".

Table 1 Applied versus Indicated Masses for "Digital Pocket Scale"

Applied Mass - g	Indicated Mass - g	% Bias
1.000	1.01	1.00
2.000	2.00	0.00
5.000	5.03	0.60
10.000	10.03	0.30
20.000	20.05	0.25
50.000	50.12	0.24

The values given in Table 1 (a) illustrate the fact that manufacturers often confuse "accuracy" with "precision" and (b) reveal that the scale has a small bias that varies with the magnitude of applied mass.

Some instruments, such as Almen gages, are notoriously difficult to calibrate accurately. Almen gages support strips on four balls that are subject to wear. The author's calibration solution is to employ a carefully-preserved set of eight stress-relieved, peened, 'A' strips. These are a set that had been peened to produce a saturation curve and therefore had different arc heights. Stress-relieving involved heating for four hours at 500°C – which reduced the arc heights by only about 10%. Polishing the stress-

relieved set 'face-up' on fine emery paper induced tiny flats on each of the four corners. A precision surface grinder was then used to produce a small central flat on the convex surface of each stress-relieved strip. Placing each such strip on an 'Engineer's stand' equipped with a calibrated digital gage allowed the 'height' (ground flat over base) of each strip to be measured. This was to determine the curvature stability of the stress-relieved strips. In practice no detectable change occurred over a ten-year period for any of this set of calibration strips.

Checking for bias, and changes of bias, is easier for some instruments than it is for others. An additional consideration is that checking takes time and therefore costs money. For some instruments e.g., air pressure gages, it is tempting to assume that the instrument does not have a bias. Complete reliance is then being placed on the inbuilt accuracy of the instrument. Critical measurements, such as arc heights, require regular checking for bias. An important guiding principle is that: "Calibration specimens should have values near to those of the objects to be measured."

VARIABILITY

Every instrument normally indicates different values when it is being used. Variability can be quantified in terms of "Variance". Variance, **V**, is the square of the measured standard deviation, **σ**, of a set of measurements. Hence:

$$\textbf{Variance, } V = \sigma^2$$

The key to understanding and using variances is to appreciate three of its features:

1 - Constituent variances are additive,

2 - Contributing variances must be identified and

3 - Contributing variances with small standard deviations can be ignored.

1 - Constituent variances are additive. Assume, for example, that single measurements of mass made on each of 50 Almen A strips indicated a variance of 11 (in arbitrary units). 50 repeat measurements made on just one of the 50 strips indicated a variance of 1. The observed variance is therefore 11 and the measurement variance is 1. Now:

Observed variance =
Measurement variance + Mass variance

so that, for this example:

11 = 1 + Mass variance

Hence we can deduce that the mass variance, for this example, is **10**, (11 – 1).

2 - Contributing variances must be identified. For example: the variances that contribute to the mass (weight) of an Almen strip can be identified as being length, width, thickness and steel density. No other properties of an Almen strip (such as hardness) contribute to its mass. If, for example, it was established that the variances of length, width and steel density for the strips were all equal to 1 then:

10 = 1 + 1 + 1 + Thickness variance

from which we can deduce that the thickness variance must be **7**, (10 - 1 -1 -1).

3 - Contributing variances with small standard deviations can be ignored. This is a very important point that is rarely highlighted. Imagine, for example, that the observed standard deviations (not variances) for length, width and steel density for a given batch of Almen strips all had a magnitude of 1 and that the observed standard deviation for mass was 10. Converting these into variances gives that:

100 = 1 + 1 + 1 + 97 (thickness variation)

That means that 97% of the observed variability can be attributed to thickness variation so that variations of length, width and steel density can effectively be ignored (as being insignificant).

Measurement Variance

Measurement variance arises when an instrument indicates different values for repeat measurements made on the same specimen. For example, a high-precision Almen gage may well indicate slightly different values for arc height when the same peened strip is measured several times. The causes of measurement variance are normally identifiable and involve a combination of operator and instrument factors. Reputable instrument manufacturers usually try to offset measurement variance. Every case is, however, different making it difficult to generalize.

The standard method for countering measurement variance is to take the average of repeat measurements on the same specimen. If two successive measurements are identical then it is generally assumed that there is no significant variance and the average is self-calculated. If, on the other hand, two successive measurements are different then further action is necessary. If the difference is only one instrument unit one can either take the average or take a third measurement. For three measurements with two the same and one differing by only one measurable digit then the value of the two identical measurements is generally accepted.

Parameter Variance

Every shot peening parameter varies. For example, Fig.5 illustrates the variability of indent size. Different parameters vary, however, in different ways. For example the variability of cut wire shot diameter is quite different from that of cast steel shot. The type of variation affects how it can be measured and controlled.

Fig.5 Variability of indent size.

APPLICATION OF VARIANCE TECHNIQUES

Management and control of variability requires that it is can be measured quantitatively. Standard deviation and variance can then be calculated automatically, for example by using Excel.

Studies of parameter variability involve several other defined terms. These include:

Population – this is the total number of identifiable objects that could be measured. A 50kg bag of 110 size steel shot would contain about two hundred and fifty million particles. The population size would then be two hundred and fifty million. Taking ten seconds per particle to measure just one parameter would take eighty years for the whole population. This leads to the need for selecting a representative sample.

Sample Size – this is the number of identifiable objects properly selected as being adequately representative of the whole population. An "adequate number" will depend on the variability of the object and the ease of making individual measurements. The greater the variability the greater is the sample size needed to be representative.

Parameter Distribution – the measured parameter values for a particular sample may have different 'distributions'. A frequently-encountered distribution is the "Normal Distribution" which has a bell shape.

Range and Average – range is the difference between the largest and smallest measurements made on a sample. Average (or Mean) is the total of the measurements divided by the number of measurements.

The following Case Study is an example of how variability techniques can be applied and analyzed.

Case Study One:
Variability of Almen 'A' Strips for Two Boxes of 50

For this study. two unopened boxes of 'A' strips, Box A and Box B, were available. The defined objectives were to (a) determine the types of size distribution, (b) calculate and compare the variability of the strips and (c) to determine the most important factor contributing to any observed size variation.

Readily-available instruments were micrometers, digital dial gages and digital weighing scales.

The easiest measurements to make were those of mass - using digital weighing scales. Complete box content weighings - on a 1000g capacity scale having a precision of 1g - gave identical values of 725g for Boxes A and B. This indicated that each strip would weigh about 14.5g (725g/50). Each strip from Box A was then weighed once – on a calibrated 50g capacity scale with a precision of 0.01g – and each strip from Box B weighed twice, (W1 and W2), once on each of successive days.

Excel provides a powerful range of analysis tools. Each of the three sets of 50 mass measurements can readily be sorted in, for example, descending order. This reveals the smallest and largest values in each batch together with the range. The average values and total mass for each batch are also indicated. Highlighting each batch of 50, then 'Formulas', 'More Functions', 'Statistical' and selecting 'STDEV' yields the standard deviation for each batch. Table 2 summarizes the application of these analysis tools. Only ten measurements from each fifty (five lowest and five highest) are shown in Table 2.

Table 2 Analyzed Measurements of Almen Strip Masses

Strip No.	Box A	Box B -W1	Box B - W2
1	14.39	14.39	14.39
2	14.43	14.39	14.40
3	14.43	14.40	14.40
4	14.43	14.42	14.41
5	14.44	14.43	14.43
etc	etc	etc	etc
46	14.50	14.50	14.50
47	14.50	14.50	14.50
48	14.50	14.51	14.51
49	14.50	14.51	14.51
50	14.51	14.51	14.52
RANGE	**14.39-14.51**	**14.39-14.51**	**14.39-14.52**
AVERAGE	**14.469**	**14.463**	**14.462**
STDEV	**0.0246**	**0.0277**	**0.0277**
SUM	**723.47**	**723.16**	**723.12**

Size distribution was assessed by constructing histograms (using Excel) for all three sets of 50 measurements – the histogram for Box A measurements being shown as fig.6. The shape of the histograms for Box B measurements had the same shape as that shown by Box A.

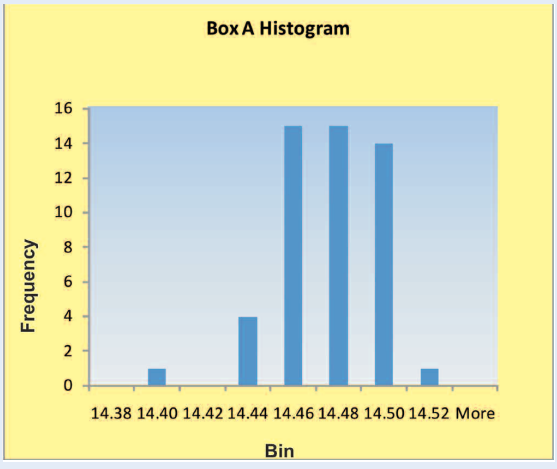

Fig.6 Histogram of mass measurements for Box A containing fifty Almen A strips.

The type of mass distribution shown in fig.6 is very similar to that of a "Normal Distribution". A Normal Distribution is very common and has an equation:

$$p = \exp[-(x - \mu)^2/(2\sigma^2)]/[(2\pi\sigma^2)^{0.5}] \qquad (2)$$

where **p** is probability, **x** is parameter value, **μ** is the average value and **σ** is the standard deviation (note that the variance, σ^2, is directly involved). Fig.7 on page 32 shows the Normal Probability Distribution for the Box A values (given in Table 2) of **μ = 14.469** and **σ = 0.0246**.

The mass of an individual Almen strip is its volume multiplied by its density. Volume of a rectangular strip is its length times its width times its thickness. This means that there are only four factors (length, width, thinness

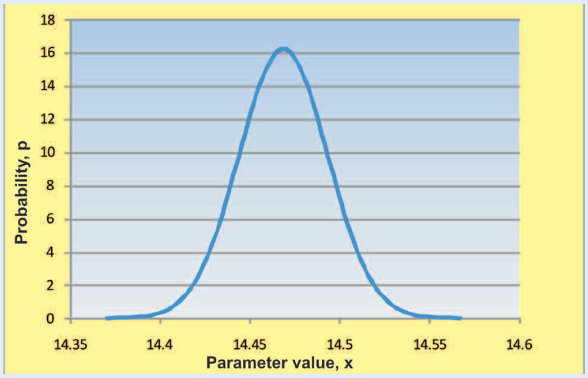

Fig.7 Normal Distribution curve for Box A parameters.

and steel density) that can possibly be responsible for the observed mass variability. Three of these factors (length, width and steel density) had such a small standard deviation that they can be ignored. This was established by selecting the lightest and heaviest strips and carefully measuring their length, width and thickness. For the Box A strips the lightest and heaviest strips were both 76.17mm long by 18.97mm wide (averages of seven measurements). The lightest strip, 14.39g, had a thickness of 1.281mm as compared with 1.295mm for the heaviest strip, 14.51g, in the batch (again averages of seven measurements). Dividing mass by volume gives a value of 7.76 for the density of both strips. Hence the only remaining significant variable is the Almen strip thickness.

The observed maximum difference in thickness for the Almen strips was 1.01%. Thickness difference will affect the magnitude of arc height induced by a given amount of shot peening. It has been established that the induced arc height is inversely proportional to the square of the strip thickness. Hence a 1.01% increase in thickness will reduce the induced arc height by 1.02% (1.01^2) e.g., from 9.76 to 9.57. Such a maximum effect is not likely to have a measurable effect on deduced peening intensity – because strips chosen for a saturation curve set would rarely include the thickest and thinnest from a box of 50.

MANAGEMENT OF INACCURACY AND VARIABILITY

Four independent factors are involved that require separate attention: **Instrument Precision, Instrument Bias, Measurement Variability and Parameter Variability**.

Instrument Precision. This the simplest factor to manage because the level of precision is pre-ordained by the instrument(s) being used. Initial purchase ensures that an appropriate level of precision is provided. Precision is, however, only part of accurate measurement.

Instrument Bias. Management of instrument bias is based on the availability of reference standards and whether or not a proactive approach is in operation. Every instrument presents different problems, so that it is impossible to generalize on their solution. For example, reference standards are readily available for weighing scales and are very simple to use. Air pressure and Almen gages, on the other hand, present much more difficult problems. A Case Study is presented that illustrates how known problems with Almen arc height measurement can be overcome.

Measurement Variability. The standard method of overcoming measurement variability is to take the average of repeat measurements.

Parameter Variability. Parameter variability is unavoidable but can readily be quantified by taking enough measurements and applying procedures such as those described in Case Study 1.

Case Study Two: Reference Standards for Almen Gage Measurements

Check blocks are commonly used to zero the gage (using the flat side) and to check one gage reading (using the singly-curved side). This does not, however, provide a reference standard for the arc height of peened strips. These have a double curvature and contact the support balls at different points from those contacted by check blocks.

An appropriate reference standard for peened strips is a set of stabilized peened strips. It has been shown that approximately half of the curvature of a peened strip is caused by residual stress and half by plastic deformation. The residual stress contribution is unstable, in the sense that peened strips slowly 'self-anneal', whereas the plastic deformation contribution is permanent. Experimental studies have shown that peened strips lose only about one or two percent of their arc height after ten years at room temperature. Thereafter no further arc height reduction is detected. 'Stabilization' consists of low-temperature annealing, which is much more effective than even ten years at room temperature. Sets of peened strips that have been stabilized cannot change their arc heights and can therefore safely be used as reference standards. Fig.8 illustrates the principle of stabilization using a set of ten peened strips.

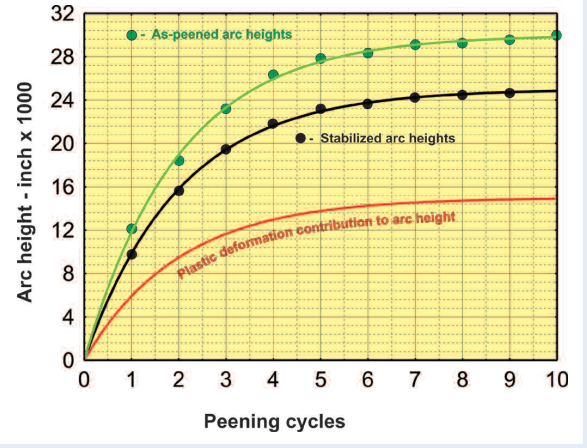

Fig.8 Arc heights and saturation curves for as-peened and stabilized strips.

Sets of stabilized strips should, ideally, have their arc heights measured several times using either a new or newly-calibrated Almen gage. A saturation curve is then produced for the set of strips and analyzed for the unique peening intensity, H, occurring at a determined fractional number of cycles, T. The individual arc heights, together with the deduced values of H and T, then act as the required reference standards.

DISCUSSION and CONCLUSIONS

Measurements inevitably involve some degree of inaccuracy and variability - in every branch of engineering. This is accommodated by having tolerance bands in specifica-

tions. Management of inaccuracy and variability costs time and money. A balance has to be achieved that is cost-effective. The optimum balance point depends upon the nature of the business involved. Some examination techniques have been included in this article but they are only intended to be illustrative of a very broad subject. Of particular note is the ease with which computer programs can be used to quantify the average, range and variability of a set of measurements. Determination of the type of measurement distribution is usually, however, of more theoretical than practical interest.

Regular use of reference standards is essential if inaccuracy is to be detected. It would be wrong to put blind faith in the accuracy of instrument readings. Measurement variability is readily countered by using repeat measurements.

It is possible to misinterpret the additive nature of variabilities. They are only additive if they are present. For example, it would be inconceivable that a saturation curve would be produced using six strips from six different batches, measured by six different operators on six different Almen gages. On the other hand, a set of stabilized strips can ensure that a given gage produces reliable measurements.

One silent enemy of accuracy is long-term drift. A relevant example is that of Almen gage ball wear. Periodic refurbishment and re-calibration is therefore necessary. Evolution of ball wear can be monitored via recorded checks using reference strips. That is facilitated if readings are fed into a computer program that can monitor progressive (and sudden) changes. ●