

Curve Fitting for Shot Peening Data Analysis

by David Kirk, Coventry University, UK

Introduction

Data abounds in shot peening. We have Almen arc heights, peening times, sieve analyses, shot weightings, image analysis values, air pressure variations, shot flow rates, residual stress data, etc. Most of this data is valuable – it cost time and money to produce – and it can all be stored in a readily-accessible format. Data that simply appears during a controlled process – such as MagnaValve readings during peening – would not normally need to be stored for subsequent analysis. Almen arc heights, on the other hand, should be stored together with associated variables such as peening time, shot size, shot type, air pressure, stand-off distance, feed rate, etc. Commercial spreadsheets (such as Microsoft's Excel) are useful for limited data storage but a proper database programme (such as Microsoft's Access) is needed for large data storage procedures. Most data analysis programs can input data straight from either a spreadsheet or a database.

Vast amounts of data can therefore be accumulated and stored, but it is all useless unless it can be analysed. This analysis must have specific objectives. These objectives may include:

- Presenting data in a palatable format,
- Determining required parameters – such as time to achieve a specified Almen arc height,
- Investigating unknown relationships between parameters and impressing customers!

Data is often presented in tabular form. Tabulation has specific advantages, the chief of which is that actual values are made available. On the other hand it is often difficult with tabulated data to visualise trends and deviations from those trends. Graphical representation is the main alternative to tabulation. Several formats are used including histograms, pie charts and X-Y plots. This paper will concentrate on the two-dimensional X-Y plotting of data and the consequent analysis of the data trends.

X-Y Plotting

We have two axes – “X” (abscissa), and “Y” (ordinate). The “X” values are generally referred to as the “**independent variables**” whereas the “Y” values are the measured “**dependent variables.**” For example, we can have a set of X values that represent specified peening times together with Y values that are Almen arc heights measured for each peening time. The magnitude of each Y value must depend in some way on the corresponding X value, hence the use of the term ‘dependent variable’. In most situations the independent variable is the one that we exercise control over and the dependent variable is the one that we subsequently measure. For example we may vary the shot flow rate by pre-determined amounts and measure the effect that has on Almen arc height (keeping all other parameters constant). Each point has an X and a Y value – its “coordi-

nates” - that specify where it will appear on a graph. If we plot a set of data points on a graph, we usually add some form of ‘data analysis’. This normally takes the form of **curve fitting**. We can employ either **interpolation** or **regression** curve fitting techniques.

Interpolation involves fitting a curve that must pass through every point. **Regression** uses some form of ‘model equation’ and attempts to minimise the differences between data points and the model equation.

Regression techniques are relatively intellectually demanding because the model chosen has to be justified. Whatever the objective is for curve fitting we should try to understand the mathematical basis of the particular data analysis procedure that is being used.

Consider the ‘fictitious’ data that has been presented in Fig.1. The data from a point whose coordinates are 0,0 and has a maximum at about 9,18 (the first value in each coordinate pair representing the X value and the second representing the Y value). The human brain is a very powerful computer. It will normally seize upon a “model” that represents the ‘behaviour’ of the data. In this case, most people would say that there appears to be a straight-line relationship between the X and the Y values. The general equation for a straight line is given by:

$$y = a + mx \quad (1)$$

where a is the value of y when x=0 and m is the slope of the straight line.

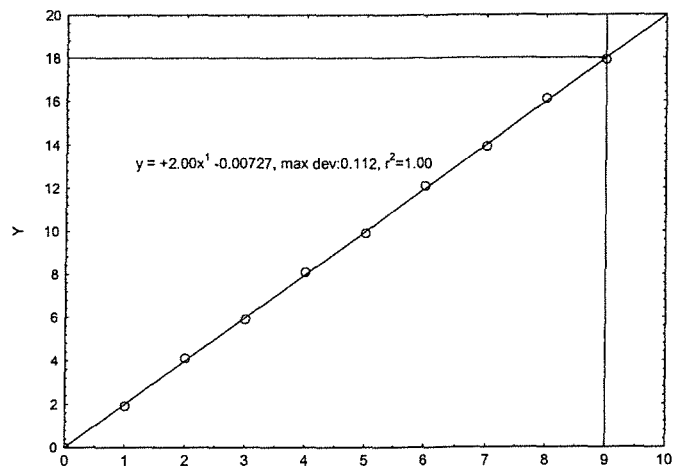


Fig.1. X-Y data plot with “least-squares” linear fit.

One method of ‘analysing’ the data in Fig.1 would be to draw a line using a pen and ruler. The brain would then employ a natural ‘minimal error’ technique so that the straight line would be placed such that the data points were at some perceived minimum distance from the straight line. That technique has the disadvantage that it is ‘subjective’ - each person will draw a slightly different line. Long before the universal employment of computerised curve-fitting techniques, it was possible to remove this subjective element by employ-

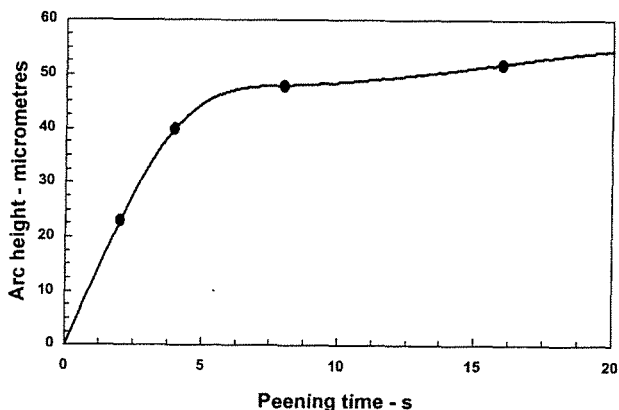


Fig.3 Cubic spline fitted by CurveExpert to data used for Fig.2.

Regression techniques

Regression techniques seek to minimise the disagreements between data points and a chosen mathematical equation (or model).

In applying regression techniques, we have to be very clear about our objectives. A simple objective would be to say, "I must have a curve that gives a perfect fit". Such an approach is wrong because that objective is only satisfied by applying interpolation – using, for example, a polynomial of one less order than the number of data points. Then the polynomial will be a 'perfect fit' since it must pass through all of the data points. The question of realistic objectives is examined later - by means of case studies.

In essence our problem is to select a 'model' that is 'satisfactory' when applied to a particular situation. This selection of a satisfactory model depends, to a greater or lesser extent, on our prior knowledge of what the shape of the resulting curve should be. At one extreme, we may know for certain what the shape of curve should be. For example, a curve of measured velocities against known time intervals for a falling steel ball should be a straight line (velocity = acceleration x time). At the other extreme, we may have no idea as to what shape a curve should be, but seek to fit a curve that at least represents some realistic physical concepts.

Having applied a particular mathematical model, it is usual to determine the corresponding 'goodness of fit'. Probably the most commonly-used test for 'goodness of fit' is the 'chi-squared' factor, χ^2 . χ^2 is computed as $(1.0 - se/sa)$ where 'se' is the sum of the errors squared (calculated Y-coordinates minus the data Y-coordinates) and 'sa' is the sum of the differences squared between the data Y-coordinates and the mean Y value of the original data points. The better the fit the smaller the value of se/sa so that r^2 approaches unity as se/sa approaches zero. χ^2 is presented as "r²" in the 'Easyplot' graphs that follow – presumably to avoid confusion between the Greek letter χ and x.

The following three case studies are used to illustrate how different objectives can be realised by correct selection of curve fitting technique.

Case Study 1: Residual Surface Stresses Induced by Peening.

A set of Almen A strips had been peened for different times in order to produce a saturation curve. The surface residual stress was measured for each strip using X-ray diffractometry. Three sets of parameters were therefore available – Almen arc height, residual

stress level and peening time, see Table 1. Plotting surface residual stress level against peening time gives Fig.4 where a simple linear interpolation has been applied. Linear interpolation is, however, unsatisfactory in this situation. That is because abrupt changes in stress at each point are impossible. A satisfactory model equation is needed which, of necessity, must involve regression analysis. One approach is to use a curve-fitting programme to solve the problem for us. Simply asking 'CurveExpert' for the best fit to the nine data points comes up with the fourth-order polynomial shown in Fig.5. The shape is quite unlike that which the linear interpolation (Fig.4) might lead most observers to expect! No criticism of CurveExpert is implied – it is an excellent program – the point being made is that subjective judgment is always needed.

Table 1 Arc heights and surface residual stress values for peened A strips.

Peening time - s	Almen arc height - μm	Surface Residual Stress - MPa
0	0	54
2	21	-89
4	38	-120
8	70	-250
16	113	-450
32	162	-613
60	209	-620
120	242	-598
240	265	-476

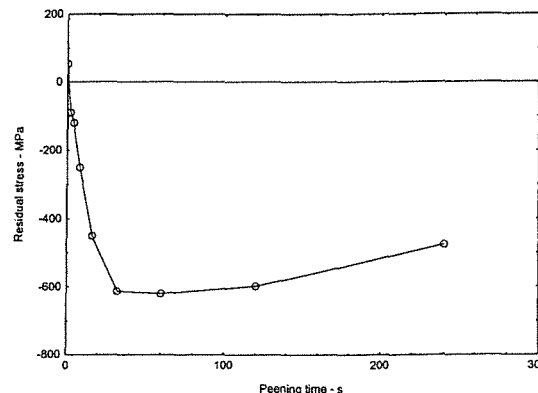


Fig.4 Surface residual stresses in peened A strips, S170 shot.

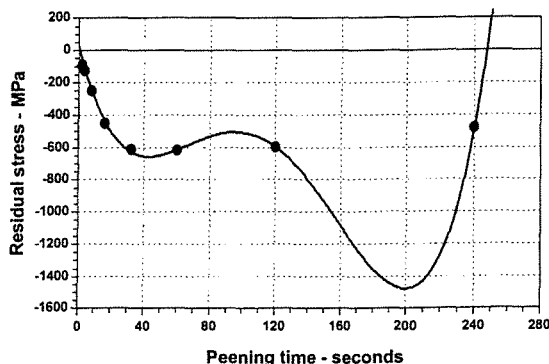


Fig.5 Fourth-order polynomial fit to peening data.

At this stage, it is important to consider any clues coming from a consideration of the original data. In this case we are starting with strips that each originally contain a small level of tensile surface residual stress (+54MPa). Using the principle of 'superposition of stresses', it is reasonable to consider an equation of the form:

$$\text{Residual stress} = 54\text{MPa} + f(\text{peening time}) \quad (3)$$

where $f(\text{peening time})$ means 'some mathematical function of peening time'.

We can then try subtracting 54MPa from each of the measured residual stress values - regarding 54MPa as a 'fixed component'. Trying to find an equation to fit ' f ' (peening time) to the residual data still doesn't work well so we have to look further at the original data. It seems possible, examining the data represented in Fig.4 that there may be a **linear component** - making the compressive stress level fall after reaching a peak value. A reasonable physical explanation of that would be the well-known phenomenon of 'work softening'. Inspection of the data in Table 1 shows that the stress level has fallen by about 120MPa in the 120s between peening for 120s and peening for 240s. That would imply a 'linear component' equivalent to a stress increase equal to t MPa - where t is the peening time in seconds. If again we use a spreadsheet to organise the data we get now three sets of 'component' data as shown in Table 2.

Peening time - s	Raw Data - MPa	Fixed component - MPa	Linear component - MPa	Residual component - MPa
0	54	54	0	0
2	-89	54	2	-145
4	-120	54	4	-178
8	-250	54	8	-312
16	-450	54	16	-520
32	-613	54	32	-699
60	-620	54	60	-734
120	-598	54	120	-772
240	-476	54	240	-770

Table 2 Splitting of Raw Data into Components.

The components of the data are plotted/represented in Fig.6. One problem remains - to find an equation that is a reasonably good fit to the 'Residual' component. Standard computer curve fitting programs did not provide an acceptable answer for the author. That seems to be odd since the residual component has quite a simple shape. In desperation, the author thought that he could certainly fit an equation to the residual data if the stress values were positive rather than negative. That was because the shape would then be very similar to a 'coverage curve' - for which several solutions were available. This 'reversed sign' residual data is plotted in Fig.7. The automatic best fit equation provided by CurveExpert was the familiar Avrami equation that is so appropriate to shot peening saturation curves.

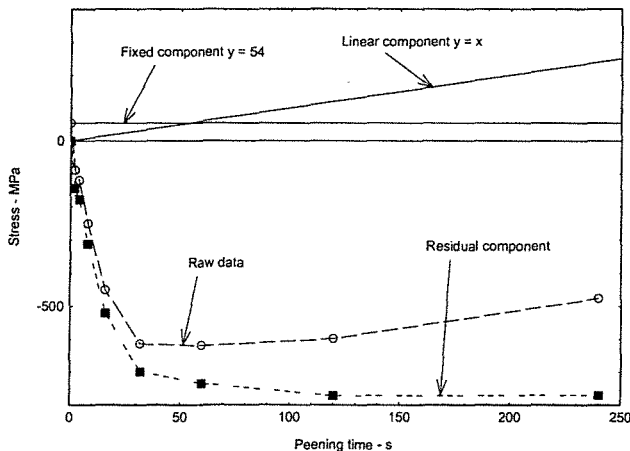


Fig.6 Components of raw stress/peening time data.

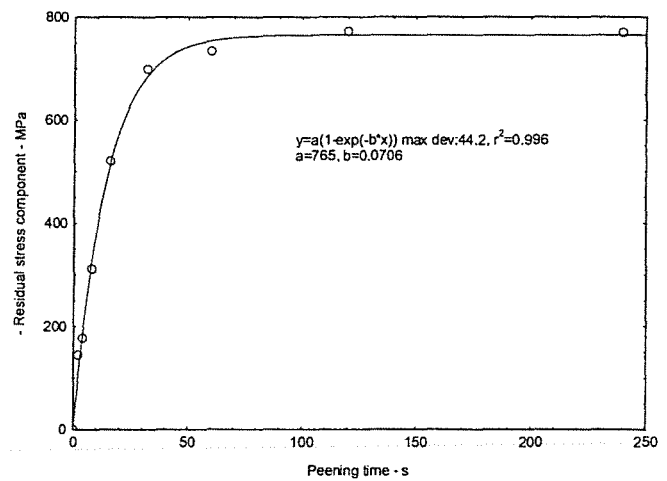


Fig.7 Reversed sign plot of residual component from Fig.6.

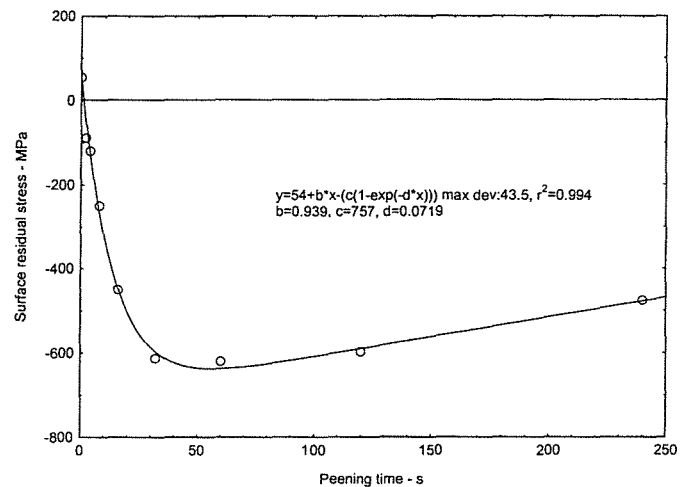


Fig.8 Curve fit to surface residual stress/peening time data for Almen A strips.

We are now in a position to curve fit an equation that contains all three components - Fixed, Linear and Residual. The mathematical form of that equation is:

$$y = a + b*x - (c(1 - \exp(-d*x))) \quad (4)$$

where a is the 'fixed component', $b*x$ is the 'linear component' and $-(c(1 - \exp(-d*x)))$ is the 'residual component'. The resulting curve fit to the original raw data is shown as Fig.8 for which the value of ' a ' has been deliberately set as 54.

The curve shown in Fig.8 does appear to be a reasonably good fit to the data. Note that the curve-fitting programme has decided that the linear component parameter b should be 0.939 - which is close to the original guess of 1. The physical model for the data is one of a compressive stress being induced by peening which obeys a coverage-type equation, which is superimposed on the original tensile stress and is accompanied by a linear stress-relaxation factor. From a scientific point of view this model should be (a) tested against any other published data, (b) examined for reproducibility (using the same peening conditions) and (c) applied to data sets obtained using different peening conditions. If general agreement is then found it may be concluded that this model is generally satisfactory.

The preceding case study was an example of the extreme type of case for which no model of curve shape was known to exist - one had to be established. There are many intermediate cases

where one has some idea of what the shape of curve should be. A classic example is that of Almen saturation curves for which the general shape is well-known.

Case Study 2: Shape of Almen Saturation Curves.

This is, without doubt, the most frequently-encountered curve shape in shot peening – but what is that shape? Fig.9 is a copy of the type of curve shown in an American Military Specification, where there is a sharp 'knee' at the peening time T. The existence of such a sharp 'knee' is, in the author's opinion, erroneous and gives a false impression of the actual shape of real saturation curves. There is no apparent scientific reason for the existence of a sharp 'knee'. Curiously, different copies of the same Military Specification, ref.3, have different Fig.9's. Some show a sharp knee and others don't.

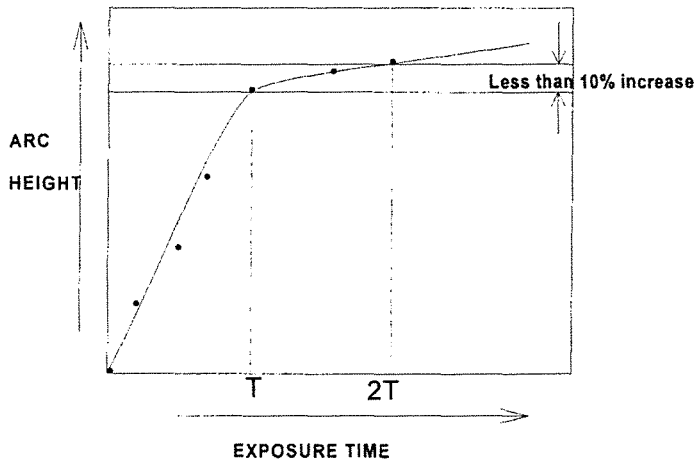


Fig.9 Almen Saturation Curve.

In order to investigate the true shape of Almen saturation curves it is necessary to look at a large collection of data that has been obtained using carefully controlled production variables. One such collection is that presented by Wieland, ref. 4, where 388 Almen strips were shot peened and a number of average values obtained. Fig.10 shows Wieland's data without a curve having been fitted. In terms of perceived shape this data has some similarities to a 'coverage curve' for which one is plotting the percentage area covered by random impacts as a function of time, ref.5. Coverage curves have the form:

$$\text{Percentage Coverage} = 100(1 - \exp(-R*t)) \quad (5)$$

where R is a constant determined by the rate of impacting.

The main difference between coverage and saturation curves is that the Y scale cannot be a percentage but is a real dimension. That can readily be accommodated by applying an equation of the type:

$$\text{Almen arc height} = a(1 - \exp(-b*t)) \quad (6)$$

where a and b are constants and t is peening time.

If we now use a curve-fitting program to fit a curve having the form given in equation (6) to the data of Fig.10 we get Fig.11.

We see from Fig.11 that the curve fit to the data is fairly good but does not, however, accurately define the shape. There are consistent deviations: from 75 to 125s the curve overestimates and from 125s onwards the curve underestimates. It is important to remember that the shape of curve should represent some physical model of the causes of X-Y variation. Equation (6) as a model is simply implying that Almen arc height is directly proportional to coverage based on

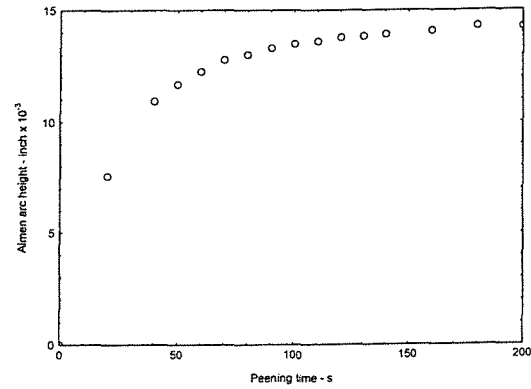


Fig.10 Representation of Almen arc heights means, from ref.4.

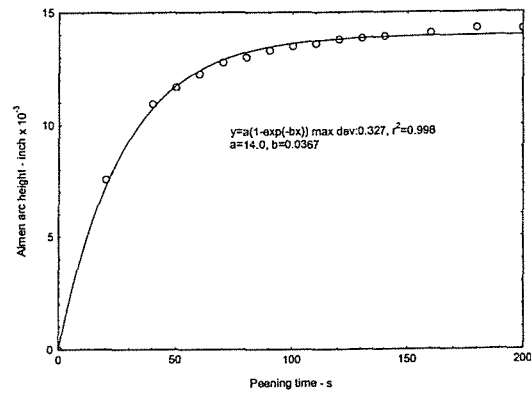


Fig.11 Curve fit of equation (6) to Wieland data.

identical spherical impacts. It is reasonable, however, to suppose that, as the surface is progressively work-hardened by peening, the impact areas will become correspondingly smaller. We can accommodate that observation by modifying equation (6) as follows:

$$y = a(1 - \exp(-b*x^c)) \quad (7)$$

where y is Almen arc height; a, b and c are constants and x is the peening time.

The constant c must be less than unity so that xc gets smaller as x increases. Applying equation (7) to the Wieland data gives Fig.12.

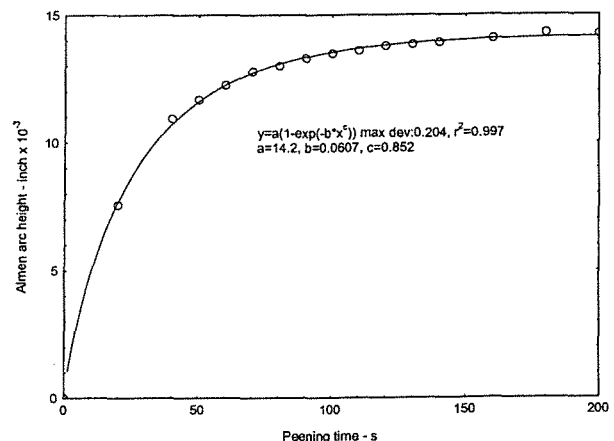


Fig.12 Curve fit of equation (7) to Wieland data.

The curve fit given in Fig.12 still shows the consistent deviations mentioned previously. A common observation of saturation data is that there appears to be a linear component – at long peening times the curve tends towards a straight line. We can, therefore, further refine our physical model by proposing that a linear component

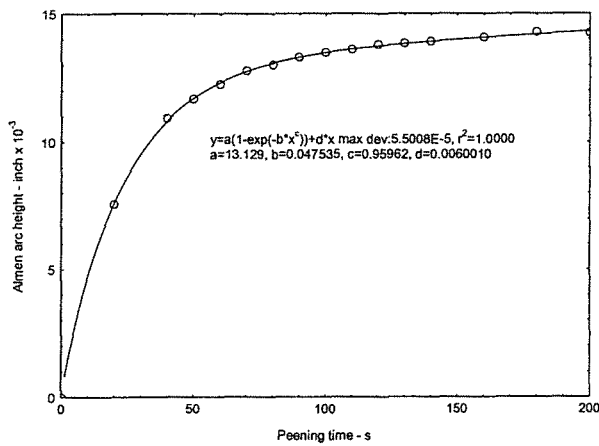


Fig.13 Curve fit of equation (8) to Wieland data.

should be included. We can accommodate that observation by modifying equation (7) as follows:

$$y = a(1 - \exp(-b*x^c)) + d*x \quad (8)$$

where $d*x$ is the linear element, d being a fourth constant.

With the modification incorporated in equation (8) we get the curve fit shown in Fig.13.

The curve fit given by equation (8) is excellent – as indicated by the 'r²' value of 1.0000. Before one can propose that equation (8) truly represents the shape that every saturation curve should have it has to be tested on other data sets. The author has carried out such tests on dozens of data sets from a variety of sources and it has proved to be extremely reliable. Furthermore, no evidence of a sharp 'knee' has been found in any of the data sets examined.

It is worth noting that a simpler form of equation (8) gives a very good fit to some data sets. That form is:

$$y = a(1 - \exp(-b*x)) + c*x \quad (9)$$

The power term x^c has been replaced by x so that there are now only three constants. It is not advocated that equation (9) should be used generally as a test of the shape of a saturation data set. Equation (8) is the preferred option.

Case Study 3: Residual Stress/Depth Profiles.

The majority of residual stress/depth profiles for peened components have a similar shape. That shape is illustrated in Fig.14. We see that there is a surface layer of compressed material with the highest level of compression occurring just below the extreme surface. The compressive residual stress then continues to a much greater depth before it has to give way to balancing tensile residual stress.

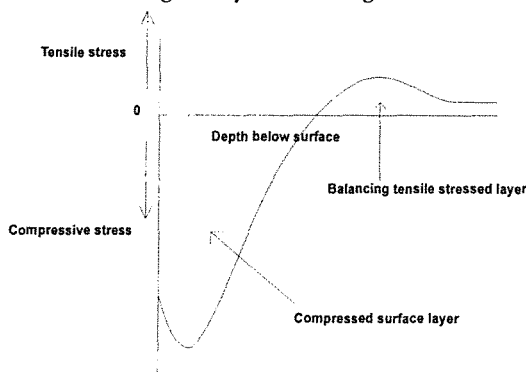


Fig.14 Classic residual stress profile for shot-peened component.

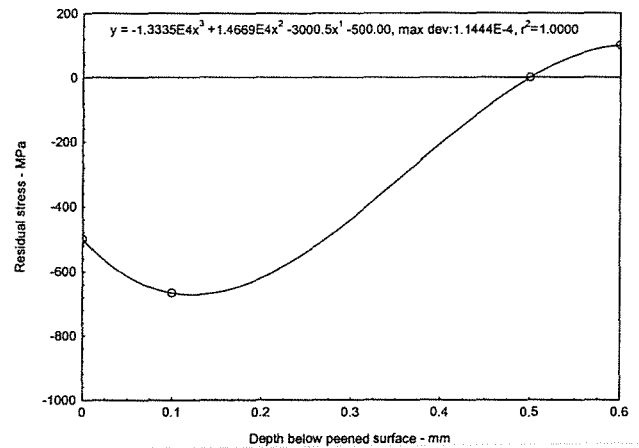


Fig.15 Model of residual stress/depth profile.

For this case study the problem to be tackled is to produce a model that will predict residual stress profile curves. All such models have to be based on a set of assumptions.

For this particular model the following assumptions will be made:

1. The level of surface compressive stress is half of the yield strength of the as-peened material, Y .
2. The maximum level of compressive stress is two-thirds of Y and occurs at 20% of the depth of compressed material, D .
3. A balancing tensile stress of 10% of Y is reached at $1.2D$.
4. A cubic polynomial interpolation will be appropriate.

For a material whose Y value is 1000MPa, peened to give a compression depth, D , of 0.5mm, then the model produces the curve as shown in Fig.15. The curve appears to have a reasonable shape relative to our general experience of residual stress/depth profiles.

The parameters of the cubic equation fit are given in Fig.15. Differentiating that cubic and putting it equal to zero gives us the value of depth (x) for maximum compressive stress. Substituting that value back into the cubic equation gives us the value of maximum compressive stress. Applying those procedures predicts that a depth of 0.123mm gives the maximum compressive stress of -672MPa. The computer program has, correctly, yielded an r^2 value of 1.0000 (perfect fit). That is because we are using interpolation not regression.

The model may be extended in several ways to make it more generally applicable:

1. Because we rarely know the yield strength of the as-peened material we can use some other measure. It is suggested that the ultimate tensile strength (U.T.S.) of the unpeened material is a good indication of the yield strength of as-peened material. That is because the U.T.S., as measured in a tensile test, is the strength of material deformed to the point of plastic instability. After the U.T.S. is reached further strengthening (true strength) occurs up to the point of fracture. During peening the material is subjected to multiple impacts that strengthen the material to about the U.T.S. level without any chance of plastic instability occurring.
2. Because we cannot know the depth of the compressed layer in advance we can make an assumption that it is equal to the dimple diameter. Dimple diameter can either be measured for a given peening situation or can be predicted. The following is a useful formula that allows us to predict dimple diameter, D :

Continued on page 30

$$D = k \cdot S \cdot \rho^{0.25} \cdot v^{0.5} \cdot Y^{0.25} \quad (10)$$

where k is a constant (equal to 1.278 when using SI units), S is mean shot diameter, ρ is the shot density, v is shot velocity and Y is the yield strength of the material being impacted.

As an example of using equation (10) if we require a dimple diameter of 0.5mm (as for deriving Fig.15) then with steel shot of density 7868kgm^{-3} fired with a velocity of 20.32ms^{-1} at a surface having a Y value of 1.381GPa then we would need to be using shot with a diameter of 1.78mm.

This case study is an example of developing a model that will predict a given type of curve without having to produce any actual experimental data. The reliability of the predicted curve is only as good as the assumptions that have been made. Hence extreme care has to be taken before any reliance can be put on the predictions.

Conversely it can be a very good guide as to the peening parameters that may lead to a required residual stress depth profile. Measured residual stress profiles can be used to confirm the applicability of the model.

Discussion

This paper has shown that there is a considerable range of curve-fitting procedures that can usefully be applied to different aspects of shot peening data analysis. The analysis of any given data set should start with deciding the most appropriate curve-fitting procedure. Having then obtained a curve, together with its mathematical equation, it is often appropriate to use that equation to determine specific characteristics. One approach, in the absence of standardized routines, is to employ a mathematically-orientated computer program – such as Mathcad, ref.7. Routines can be written to determine, for example, the ‘saturation point’, T , on a saturation curve. Once written, such routines greatly simplify the problem of deriving critical points objectively. Getting good results from curve fitting is not, however, as simple as having good data and the right model. When performing non-linear fitting, it is critical to take notice of all of the information that is provided and check that the fit is really a good one that makes physical sense. The algorithm being used by a given computer program knows nothing about shot peening! It is therefore up to the individual to accept or reject its results based on one’s knowledge and experience.

Perhaps the most important application of curve fitting is in connection with saturation curves. There is, for example, considerable interest in producing computer-generated saturation curves based on a limited number of data points. We must, however, be very careful about our objectives and situation before relying on such curves. Consider the situation of a fully-automated, computer-controlled, shot peening facility where we hope to guarantee that a pre-determined regime of shot type, shot velocity, gun-to-workpiece distance, impact angle, etc. will always give us the same Almen arc height for any given peening time. It can then be argued that there is no need to produce a confirmatory saturation curve. Such an ideal situation probably cannot exist due to unavoidable process variation. At least two Almen strip tests at different peening times would be required at reasonable intervals in order to have confidence that the ‘ideal’ regime was being maintained. If we need to establish the saturation curve that arises from a new set of process variables we must, of course,peen at least four strips.

It has been argued that four strips are sufficient to determine the shape of a saturation curve. That argument relies on a belief that (a) each reading contains no error and (b) an equation with only three coefficients is perfectly satisfactory for shape determination. It is contended here that at least five strips should be used so that a four-coefficient equation (equation (8)) can be derived as a ‘fit’ to the six data points that then arise (including 0,0). That contention is based on the physical model that supports equation (8) and on a perceived error when applying a three-coefficient equation. That error is illustrated in Fig.16 where the slope of the linear component is almost fifty times greater ($0.75588/0.015989$) for the three-coefficient curve than it is for the four-coefficient curve. The peening time axis has been deliberately extended in order to illustrate the gross difference in linear components.

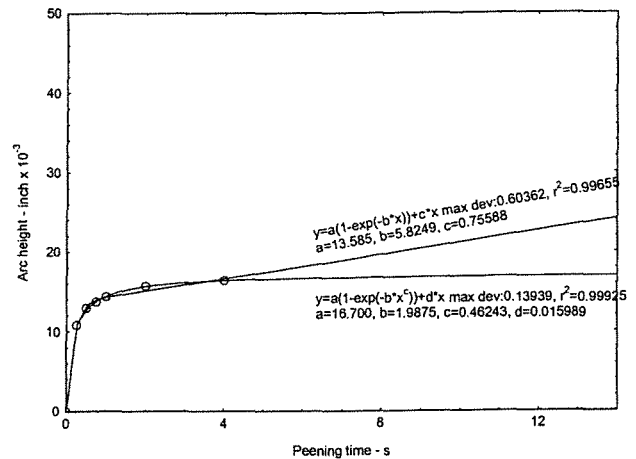


Fig.16 Comparison of three- and four-coefficient curve fitting.

The relatively small linear component predicted by the four-coefficient curve is consistent with the author’s experience of analysing a wide range of saturation curves.

Rather than being constrained by a limited number of Almen strips there is a more satisfactory alternative. Fig.17 (next page) shows a representation for a single Almen A strip peened with eight passes whilst remaining clamped. The displacement is monitored by an LVDT during peening and continuously recorded, ref.6. Eight peening passes were made giving total peening times of 2, 4, 8, 16, 32, 60, 120 and 240 seconds. The corresponding saturation curve is plotted as Fig.18. Hence we have a complete saturation curve produced during peening using only one strip. This ‘interactive’ technique is also useful for monitoring any variation in peening parameters during the actual process. It should be noted, however, that the displacements for as-clamped Almen strips are much less than for ‘released’ strips of the same thickness. Clamped Almen N strips, on the other hand, give displacements similar to those of Almen A strips after their release for measurement (for equivalent degrees of peening). LVDT displacements are easily calibrated against Arc heights achieved on release from clamping.

A small, but significant, feature of computerised curve fitting is that most programs require that “initial guesses” of coefficient values be made. Those programs often include built-in default values that, if not over-riden, commonly result in a “Math Error” message. With practice, initial guesses become easier to predict. For example, with saturation curves the ‘a’ value is similar to the maximum arc height value.

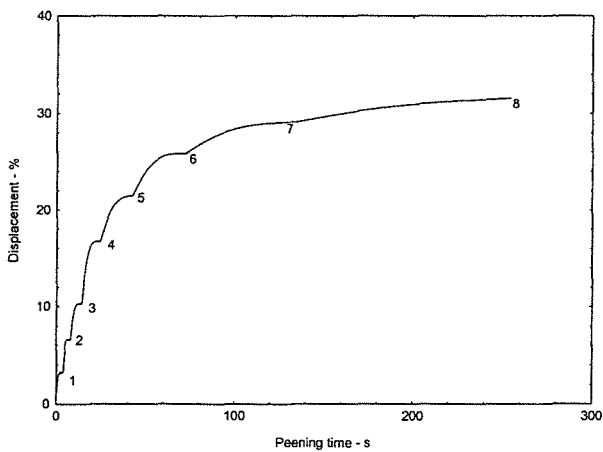


Fig. 17. Interactive monitoring of peening intensity.

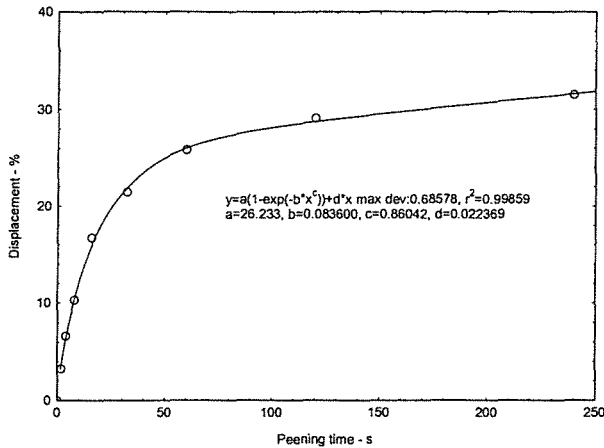


Fig. 18. Saturation curve plotted with data from Fig. 17.

Computer-generated residual stress profiles can give a useful guide as to the peening parameters that will give a required profile. They are, however, only predictions and suffer from the number and detail of the assumptions that must be made. One major problem is that the behaviour of the material during peening cannot be accurately predicted. For example, stainless steel may develop transformations that have a profound effect on the residual stresses. At least one major peening company uses its own program for predicting stress profiles.

In conclusion it may be argued that the application of computerised curve fitting routines should become a standard feature of peening control. Specific computer programs are mentioned in this paper simply because the author is familiar with them.

'CurveExpert' has the advantage that it is 'shareware' so that newcomers to curve fitting can try out a fully-functional program without incurring any expense. 'Easyplot' can display several curves on the same graph. It is, however, primarily a plotting (rather than curve-fitting) program and is quite expensive. In some situations CurveExpert's algorithm provides fits that defeat Easyplot when using the same data - and vice versa!

Finally, the best way to learn about curve-fitting is to do it. Do not be put off by mathematics!

References

- 1 Easyplot, version 4.03, www.spiralsoftware.com.
- 2 CurveExpert, version 1.37, www.ebicom.net/~dhyams/cvxpt.htm.
- 3 MIL-S-13165C, 1989, "Shot Peening of Metal Parts".
- 4 Wieland, R. C., "A Statistical Analysis of the Shot Peening Intensity Measurement", pp27-38, Proceedings I.C.S.P.5, Oxford, 1993.
- 5 Kirk, D. and Abyaneh, M. Y., "Theoretical Basis of Shot Peening Coverage Control", *ibid*, pp183-190.
- 6 Kirk, D., "Interactive Shot Peening Control", *ibid*, pp9-14.
- 7 Mathcad, www.mathsoft.com.

**"HAVE A
BLAST
WITH US"**

Vulkan Blast Shot Technology
P.O. Box 2052, 10 Plant Farm Blvd, Unit 2
Brantford, Ontario N3T 5W5

~ Call **1-800-263-7674** (Canada and U.S.) ~

Tel: 519-753-2226 ~ Fax: 519-759-8472
E mail: vulkan@vulkanshot.com

Supplier of:

VULKAN CHRONITAL

A high quality cast stainless steel shot media used for cleaning, deburring, descaling and surface finishing of a wide variety of aluminum and other non ferrous alloy castings. Ferrous contamination is avoided with a single step blast cleaning process.

Sizes from .005 to 0.1 inch are readily available.

CARBON & STAINLESS STEEL CUT WIRE

A high hardness, peening quality steel shot used for shot peening a wide variety of machine parts and structural components, primarily used in the automotive and aircraft manufacturing industries, meeting their respective specifications.

Standard particle sizes from .016 to .035 are readily available.
Available as cut or conditioned, graded and certified to meet Mil specifications

With Vulkan you can count on:

- HIGH QUALITY
- STRONG TECHNICAL SUPPORT
- EFFICIENT SERVICE

All at a Competitive Price.

VULKAN
Blast Shot Technology

Division of VULKAN HAREX Steelfiber (North America) Inc.