Robust empirical predictions of residual performance of damaged composites with quantified uncertainties

W. J. R. Christian*, E. A. Patterson and F. A. DiazDelaO

School of Engineering, University of Liverpool, L69 3GH, UK

*Email: W.Christian@liverpool.ac.uk

Abstract

Robust predictions with estimated uncertainties were made for the residual strength of impact-damaged composite laminates based on simple non-destructive measurements of the size of the damage from ultrasound C-scans. Experimental data was acquired for two sets of composite coupons, one with a crossply and the other with a quasi-isotropic layup. The laminates were subject to drop-weight impacts, non-destructively evaluated using ultrasound and then loaded to failure in bending. An empirical model of the residual strength of each laminate layup, as a function of the ultrasound measurements, was generated by fitting a Bayesian linear regression model to the normalised measured data. Bayesian linear regression was demonstrated to provide conservative estimates when only minimal data is available. Unlike classical regression, this technique provides a robust treatment of outliers, which avoids underestimation of residual strength. The Leave-One-Out-Cross-Validation (LOOCV) metric was used to assess the performance of models allowing for the quantitative comparison of the predictive power of regression models as well as being consistent in the presence of outliers in the data. The LOOCV metric indicated that predictions of residual strength are up to 25% more accurate when based on damage area than when using measurements of the damage width or length. The proposed approach provides a robust methodology for performing damage assessments in safety critical composite components based on reliable predictions with quantified uncertainties.

Keywords

Residual strength prediction, Robust Bayesian regression, Composite materials, Impact damage, Quantified uncertainty, Ultrasonic NDE

Acknowledgements

The authors would like to thank Kathryn Atherton and Eszter Szigeti of Airbus for helpful discussions. WJRC is studying on an EPSRC CASE Award PhD sponsored by Airbus.

1. Introduction

Composite materials are used in aircraft structures to assist in reducing weight without compromising strength or durability. Structural design analyses ensure that operators can be confident of the aircraft's safety and durability when only minor damage is present. However, defects may be introduced during manufacture or impacts may occur during service that cause substantial reductions in the strength or fatigue life of a composite structure. Therefore, it is important to be able to locate and size defects in composites using non-destructive evaluation (NDE). A number of techniques are used to make such measurements with perhaps ultrasound [1] and thermography [2] being amongst the most popular. Once damage has been located and its size estimated, the next step can take a number of forms: the simplest is to repair or replace the component as soon as a damage indication occurs. This is a common approach in the aerospace industry because predictions of residual strength and life of composites have substantial levels of uncertainty [3]. At the opposite end of the complexity scale, a finite element model of the damaged component can be created with the available data [4] and used to simulate damage propagation under service conditions. It is difficult to establish the extent to which a finite element model provides an accurate representation of the damaged component and thus it is usually costly in time and resources to obtain high-fidelity predictions of residual performance with a known level of uncertainty. Alternative approaches utilise empirical models relating measurements of damage to experimental measurements of component residual performance when the same type of damage is present [5]. Once the database has been established to support the empirical model, this type of approach can be applied quickly and without in-depth expert analysis, thus reducing revenue lost during structural assessments and the costs of performing such assessments. It is also more straightforward to establish the uncertainty associated with predictions from empirical models and this allows for conservative estimations of residual performance to be made that ensure the safety of the aircraft whilst limiting the number of unnecessary repairs.

Empirical models have been created for a wide range of damage types [6-9] including impact damage, for which it has been shown that the compressive residual strength of an impacted composite correlates linearly with ultrasound measurements of the width of damage perpendicular to the direction of external loading. This correlation has been used to make predictions of residual strength using the classical technique of least-squares regression [7].

In classical regression two assumptions are made regarding the data [10]: first, there is no prior knowledge of the potential values of the regression line's intercept and slope. These parameters are treated as unknown constants that are inferred directly from data. This means that, when a small number of specimens has been tested, the calculated values of these parameters will have no indication of their uncertainty. The second assumption is that the deviation of the measured data from the predictions has a normal distribution. The normal distribution can be a poor choice to represent the deviation of measurements from predictions when there is a high probability of very large deviations, which corresponds to a heavy-tailed probability distribution [11]. Situations where this can occur are: when environmental conditions have a significant effect on damage measurements [12], the mechanics that generate the internal damage exhibit substantial variation during the impact [13] or substantial amounts of immeasurable microscopic damage is created [14]. Measurements that substantially deviate from the trend are often considered outliers and can cause problems when fitting a linear model. This is due to the outliers influencing the parameters of the line, i.e. the intercept and slope [12, 15], as well as the estimated uncertainty of the residual performance predictions made by the model [16]. As a consequence, outliers are often removed from engineering data using methods, such as the maximum normed residual [17] or the least trimmed squares [12]. These methods assume outliers are erroneous measurements; however in reality, the outliers may be a valid aspect of the behaviour of a structure and discarding such residual performance measurements will result in a regression model that under-estimates the uncertainty of residual performance predictions and could classify damage as safe that has a significant probability of causing failure. Thus, there is a need for a more robust approach to handling measured data with outliers and establishing the uncertainty of resultant predictions in the field of non-destructive evaluation of composites.

In this paper, Bayesian linear regression has been applied to simple ultrasound measurements of damage. Leave-One-Out-Cross-Validation (LOOCV) has been used to determine the best measurement parameter, i.e. damage area, length or width, for reliable strength prediction. Bayesian linear regression allows the uncertainty in residual strength predictions to be estimated whilst accounting for the presence of outliers and ensures prediction uncertainty is appropriately high when small quantities of data are available. The approach is illustrated using C-scan data from quasi-isotropic and crossply coupons manufactured from carbon fibre prepregs.

The paper is organised as follows. In Section 2 an overview of classical and Bayesian linear regression is provided. The experimental method is described in Section 3, detailing how specimens were manufactured and assessed. In Section 4 the specific Bayesian linear regression model used in this paper is described. Section 5 presents the results of the experimental work and the predictive models that have been fitted to the measurements. In Section 6 the results are discussed and some of the capabilities of Bayesian regression are explored. Concluding remarks are presented in Section 7.

2. Regression Techniques

2.1. Classical Linear Regression

The purpose of regression analysis is to predict a quantitative response based on measured data, where the relationship between the two is assumed to be linear. In classical regression analysis, the method of least-squares is used to determine the parameters of a linear model, assuming that the difference between the model and the data is normally distributed. This method is widely known and documented [7, 18] and thus the equations to perform classical regression are not included in this paper. The aim of least-squares regression is to fit a line, $y = mx + c$, to a set of n points (x_i, y_i), where y is the residual strength to be predicted and x the value of the damage metric measured using ultrasound. The parameters m and c are directly estimated from the residual strength data. By assuming the measured residual strengths are normally distributed around the regression line, a region can be defined that in 95% of cases is expected to contain the residual strength measurements. This region is known as a confidence interval. The lower bound of the confidence interval is then used to define the allowable damage size for a safe residual strength such that only 2.5% of components with the allowable damage size would be expected to fail at loads below the required strength. A potential problem with the classical approach arises when the data is scarce and there is strong prior information that is not being used. The sources of this prior information can be a combination of expert knowledge, engineering judgement, physical constraints, amongst many others. When prior information is ignored, the fitted regression model may generate inaccurate predictions. This problem is circumvented by adopting a Bayesian framework. By doing so, uncertainty is quantified when predicting the residual strength for any unobserved damage value.

2.2. Bayesian Linear Regression

For Bayesian regression, initial probability distributions, called prior distributions, are defined for the parameters of the regression line (e.g. line gradient, m, and intercept, c). These prior distributions are expressed in mathematical notation as $p(\theta)$ where θ is a column vector of the parameters of the regression line. The prior distributions can be specific and well-defined if previous experimental data is available or vague if no such data exists. A vague prior distribution tends to be very wide in the sense that a large number of potential values of the parameter have a high probability of occurring, indicating the lack of pre-existing knowledge. When residual strength measurements are taken the data can be used to update the prior distributions using Baye's rule,

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{\int p(D|\theta)p(\theta)d\theta} \quad (1)$$

After the parameter distributions have been updated they are called 'posterior distributions' and have the notation $p(\theta|D)$ where D is the residual strength data. The term $p(D|\theta)$ is called the likelihood function and is the probability that the residual strength measurements come from a system with the regression parameters in θ. The denominator of equation (1) is used to normalise the posterior distribution such that the area under the probability density functions for the posterior distributions is equal to unity. For simple models, the posterior distributions can be calculated analytically. However for regression models such as the one described in section 4 this is not possible. Instead, random samples from the posterior distribution can be generated using Markov Chain Monte Carlo (MCMC) techniques. For this paper, the MCMC algorithm known as Gibbs sampling, described extensively in [10], has been used to generate a large number of samples of the possible parameter values for the regression model. The histogram of these random samples will have the same shape as the posterior distribution for each parameter and thus can be used to make residual strength predictions.

3. Experimental Method

Two sets of composite specimens, with in-plane dimensions of 250 mm by 90 mm, were manufactured using two different material systems and layups. Since laminates consisting of unidirectional plies are particularly sensitive to impacts [14], unidirectional prepregs were used to manufacture the specimens for this study. Twelve crossply carbon fibre laminates were produced from M10RUD150 unidirectional prepreg (Hexcel, USA) with a $[0_2/90_2/0_2/90_2/0_2/90]_S$ layup where the 0° plies were parallel to the longest edge. The prepreg laminate was cured in a hot press (APV-3530, Meyer, Germany) at a temperature of 130 °C and a pressure of 2.5 bar for 45 minutes as per the manufacturer's recommendations. The press was heated up to temperature at a rate of 10 °C/min and after curing was left to cool naturally with the pressure maintained. A second batch of specimens with a different layup was produced to demonstrate the rigour of the damage assessment. The different layup affects the size and shape of impact damage in the laminate [14]. Twenty-five quasi-isotropic carbon fibre laminates were produced from RP507 prepreg (PRF, UK) with a $[0_2/90_2/45_2/-45_2]_S$ layup using an identical process. The nominal thickness of the crossply and quasi-isotropic specimens were 2.90 mm and 3.02 mm respectively. Whilst an autoclave would typically be used for curing load-bearing aerospace components, a hot press has been used for this study as only flat laminates were required.

A drop weight impactor, identical to the one employed by Yang and Cantwell [19], was used to produce barely visible impact damage in the laminates with a 20mm diameter hemispherical tup of mass 2.67kg. The specimens were clamped onto a metal support with a 125 mm by 75 mm opening underneath them. The eleven crossply specimens were impacted with a range of energies between 4 and 14 J leaving a single virgin specimen, and 18 quasi-isotropic specimens were impacted with a range of energies between 4 and 15J leaving seven virgin specimens.

The specimens were then cut down to 240 mm by 40 mm coupons using a wet diamond saw (Versatile 103450, Vitrex, USA). The use of the wet diamond saw resulted in clean edges on the specimens and did not produce edge delaminations. The impact location was used as the centre of the specimens.

After each impact the delaminations in the composite were evaluated using pulse-echo ultrasound. Ultrasound was the chosen inspection technique as it is commonly utilised in the aerospace industry for the sizing of delaminations, such as those produced by impacts. Delaminations represent significant discontinuities in composites, which reflect a large portion of the ultrasound energy back towards the probe [1], resulting in clear images of subsurface damage. A multi-axis scanner (Midas NDT, UK) was used to produce C-scan time-of-flight images of the internal damage. The scanner used a focused probe with a natural frequency of 10 MHz. The probe and coupon were immersed in water for ultrasonic coupling with a standoff distance equal to the focal length (50 mm) of the probe. The probe was attached to an ultrasonic flaw detector (Epoch 4+, Olympus, Japan) that produced an A-scan of the laminate at the location of the probe. Using the A-scan data the vertical position of damage within the coupon was identified. The flaw detector's time-variable-gain feature was used to correct for attenuation of the ultrasound signal through the depth of the coupon. Delaminations stop the ultrasound energy so that only the topmost damage is visible in each C-scan. The output voltage from the flaw detector was sampled at 100 µm increments along a series of lines at a spacing of 200 µm. The flaw detector was calibrated by scanning virgin specimens of known thickness so that the output voltage could be converted into a measurement of the depth of the detected delamination from the top surface of the laminate. The width, length and area of the damage zone were evaluated from the C-scans.

After the damage evaluation using ultrasound, the coupons were loaded to failure using a servo-hydraulic load frame (8501, Instron, USA) in a four-point bend setup with a support span of 160 mm and load span of 80 mm. Each coupon was placed in the loading rig such that the impact damage was centred in the load span and the impacted surface was in tension. The specimens were loaded under displacement control at a rate of 0.8 mm/min and failure was defined as the point at which the force dropped below 50% of its peak value. The maximum bending moment was recorded and used as the measure of the residual strength of the coupon.

The ultrasound damage, x_i and residual strength, y_i measurements were normalised using:

$$x_i^* = \frac{x_i - \bar{x}}{s_x} \qquad (2)$$

$$y_i^* = \frac{y_i - \bar{y}}{s_y} \quad (3)$$

where $\bar{x}$ and $\bar{y}$ are the mean of the ultrasound measurements of damage size and residual strength measurements respectively, and s_x and s_y are their standard deviations. The normalisation allowed for the use of the same prior distributions for Bayesian regression regardless of the units used in the ultrasound and residual strength measurements. The normalisation also increased the efficiency of the Gibbs sampler, by reducing the number of iterations required to represent the predictive distribution, because it ensured that the potential values for the gradient and intercept were close to zero and of approximately the same scale so that large perturbations were not introduced during sampling.

4. Bayesian Modelling

The model used to predict residual strength based on ultrasound measurements was linear, of the form $y = mx + c$, with the assumption that the data was distributed around the regression line as a Student's t-distribution. This distribution was employed because it allowed the deviation of residual strength measurements from the regression line to have a greater probability of outliers, compared to the normal distribution. Thus, when present, outliers were expected to have less effect on the intercept and gradient of the regression line. The form of the model is shown schematically in the bottom portion of Fig. 1, in a similar style as the one employed in [10]. The shape of the distribution about the line shown in the centre portion as a t-distribution defined by its normality, ν_S, spread, σ_S, and mean, μ_S, which in turn is defined by the gradient of the line, m_S, and its intercept, c_S. Each of these four parameters: normality, spread, gradient and intercept are described by prior distributions that are shown schematically in the top portion of Fig. 1. It was decided to represent the prior distribution for the normality parameter by an exponential distribution in order to allow it to vary from low values that causes the t-distribution to have long tails, to high values at which the t-distribution behaved like a normal distribution. The probability density function, $p(y)$, for the t-distribution illustrating this behaviour is shown in Fig. 2. The exponential distribution has a scale parameter, K, of 29. The prior distributions of the gradient and intercept of the line-of-best-fit have been defined as normal distributions, which is the common choice for simple regression models [20]. In both normal distributions the means, M_c and M_m, were set to zero as the normalisation process in equations (2) and (3) cause the values of both the intercept and gradient to be close to zero. The standard deviations, S_c and S_m, of the normal distributions are also defined with a high value of 100, which ensures that the prior distributions are vague and thus will not limit the potential values of the gradient and intercept. The probability distribution describing the prior distribution of the spread of the t-distribution, σ_S, was a uniform distribution that assigned an equal probability to values of the spread from $L = 1 \times 10^{-3}$ to $H = 1 \times 10^{3}$. All of the parameters used to define the initial or prior probability distributions are listed in Table 1.

The regression model was fitted to the residual strength data using the open-source software, JAGS [21] to determine the posterior distributions of the parameters, that is the normality, ν and spread, σ of the t-distribution together with the gradient, m and intercept, c of the regression line. Subsequently, these posterior parameters can be used to predict the residual strength of a damaged composite structure when a particular size of impact damage has been identified. Since the posterior parameters are each represented by a probability distribution, the prediction also results in a probability distribution that is called the posterior predictive distribution. This predictive distribution can be computed using a random number generator such that when $\nu_{(j)}$, $m_{(j)}$, $c_{(j)}$ and $\sigma_{(j)}$ are the jth samples from the posterior distributions, then the jth sample of the predicted strength is given by

$$y_{(j)} = t_{rnd}\left(m_{(j)}\left(\frac{x-\bar{x}}{s_x}\right) + c_{(j)}, \sigma_{(j)}, \nu_{(j)}\right) \times s_y + \bar{y} \qquad (4)$$

where $t_{rnd}(\,)$ is a generator of random numbers conforming to a t-distribution and, as in equations (2) and (3), $\bar{x}$ and $\bar{y}$ are the mean of the ultrasound damage and residual strength measurements respectively, and s_x and s_y are their corresponding standard deviations. The JAGS software package was used to generate 500,000 random samples at each x location and ranked by the magnitude of the predictive residual strength in order to allow the 2.5th, 50th and 97.5th percentiles to be identified. A typical result is shown in Fig. 3 with an interval defined using the 2.5th and 97.5th percentiles from the predictive distribution. The 50th percentile of y was used to plot the line-of-best-fit. The locations at which the percentiles were calculated are marked with dots. It is important to note that when using percentiles of the predictive distribution, the interval generated is commonly called a credible interval. The distinction between the classical confidence intervals and the Bayesian credible interval is subtle but important. In the classical (frequentist) view, parameters are fixed but unknown. Thus, a 95% confidence interval is expected to include the true value of the parameter 95% of the time in repeated sampling. The interval, not the true value of the unknown parameter, is random. From the Bayesian point of view, the value of the parameter is considered random and a 95% credible interval is expected to contain 95% of the probability distribution of the parameter. Under certain conditions, confidence and credible intervals may coincide. For a detailed account on the similarities and differences in methodology and philosophy of construction refer to [22].

The performance of the regression model was assessed using the Leave-One-Out-Cross-Validation (LOOCV) technique [20, 23] to alleviate the effect of double-counting when using the same measured data to fit the model and assess its performance with a traditional correlation coefficient such as R^2. Using the LOOCV technique, the regression model was fitted to the data n times, where n was the number of coupons tested and omitting or leaving out the data from one coupon each time, i.e. using $(n-1)$ points of data, and the fitted model was used to predict the left-out measured data. The numerical difference between the predicted and measured residual strength of the left-out specimen was recorded as the prediction error for that coupon, e_i, and the LOOCV performance metric was taken as:

$$LOOCV\ Performance\ Metric = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_i}^2} \qquad (5)$$

Since, the extent of the damage detected in the ultrasound scans can be characterised in a number of ways, including area, width and length, as shown in Fig. 4, the LOOCV metric was used to identify the damage characteristic that provided the best predictions in terms of the lowest value of root mean squared error in equation (5) and the results are listed in Table 2.

5. Results

A typical time-of-flight ultrasound C-scan from an impacted coupon is shown in Fig. 4. The area, width and length of the damage were used as damage metrics and their efficacy for predicting residual strength compared, as described in the previous section. The width was defined as the total width of the delaminations perpendicular to the loading direction when viewed from above, as defined by Pritchard and Hogg [7]. Damage length was measured similarly, but parallel to the loading direction. The damage area was taken as the total projected area of the delaminations when viewed from the impacted face using time-of-flight ultrasound.

Fig. 5 shows the Bayesian linear regression predictions of the residual strength of quasi-isotropic coupons using area, length and width of the ultrasound data (as shown in Fig. 4) as the damage metric. The values of LOOCV performance metric in Table 2 reveal that the area of the damage is the best metric for predicting the residual strength and so only this prediction is shown for the crossply coupons in Fig. 6.

The predictions from Bayesian linear regression form a straight line as expected while the corresponding uncertainties are a pair of curves above and below the regression line defining a credible interval such that 95% of results for all coupons evaluated with ultrasound (to provide a value of the damage metric, x^*) and then tested to failure (to provide a value of residual strength, y^*) would be expected to lie in. Residual strength predictions were made for a range of damage metric values from zero, indicating no damage found, to the point at which the damage causes the residual strength to be approximately half the undamaged strength.

6. Discussion

The graphs in Fig. 5 and Fig. 6 show the predictions of the residual strength as a function of size of damage and are directly comparable to the ones produced by Pritchard and Hogg [7] with the same correlation between damage width and residual strength demonstrated. There is substantial scatter above and below the regression line in the measured residual strengths, which is likely due to the complex nature of impact damage involving both fibre breakage and matrix cracking within the area of impact damage [14]. It is not possible to resolve these microscale features using ultrasound and a very much more sophisticated model would be required to incorporate such detail into the residual strength predictions (e.g. [24] where an FEM model was developed that predicted the propagation of delaminations produced by impact damage). However, the Bayesian approach employed here together with the use of the calculated credible interval allows all of the potential sources of uncertainty, including both the noise in the ultrasound measurements and the behaviour of the unmeasured microscale damage, to be accounted for in the predictions. The lower bound of the credible interval can be used to specify the maximum allowable size of damage for a specified minimum residual strength. For instance, if the minimum residual bending strength that can be tolerated is 32 Nm for the crossply coupons then the maximum allowable area of damage detected by ultrasound would be 341 mm^2, as shown by the dotted lines in Fig. 6, in order to ensure that coupons with a probability of 2.5% or more of failing at loads below the minimum residual strength are taken out of service for repair or disposal.

The LOOCV performance metric, values given in Table 2, was used to demonstrate that damage area was the best choice for residual strength predictions for both the crossply and quasi-isotropic specimens as predictions were up to 25% more accurate than when using the other damage metrics. The R^2 statistics, also shown in Table 2, also demonstrate that damage area is the best empirical model however this statistic does not estimate the average uncertainty of future predictions. The use of the best available damage metric for residual strength predictions reduces the size of the credible intervals, as can be seen in Fig. 5, which raises the maximum allowable size of damage permitted for the same probability of failure for a given required residual strength. This is likely to reduce the number of components that are removed unnecessarily from service for repair or disposal and hence reduce operating costs. Table 2 contains LOOCV performance metrics for both Bayesian and classical regression models showing that for data that does not contain outliers the two regression techniques have similar behaviour.

The influence of outliers on Bayesian regression was explored using an artificial set of data generated using the following linear function:

$$y_i = -3x_i + 60 + N_{rnd}(0, 0.3^2) \tag{6}$$

where $x_i = \{3,4,5,\cdots,12\}$ and $N_{rnd}(0, 0.3^2)$ is a random number generator that produces numbers with a normal-distribution of mean value of 0 and standard deviation of 0.3. The classical and Bayesian linear regressions and the corresponding intervals calculated from this data were identical, as expected, and thus only the Bayesian linear regression is shown in Fig. 7. However, when an outlier was introduced by changing a single value, $y_2 = 30$, then the behaviour of the two techniques was substantially different, as shown in Fig. 8. The gradient and intercept of the Bayesian regression is unchanged with the line-of-best-fit equation being almost identical to the linear function in equation (6). For classical linear regression the gradient has increased to -2.2, an error of 26% of the value used in equation (6) to generate the data. The credible intervals are larger in the presence of the outlier by a factor of 3.3 and 12.2 respectively for the Bayesian and classical regression models. This suggests that the Bayesian regression model is more robust in the presence of an outlier, which is expected, because in these circumstances Gibbs sampling will assign higher probabilities to low values of the normality parameter of the t-distribution so that the outlier has little effect on the slope and gradient distributions for the regression line. Consequently, when using the Bayesian model, there is no need to identify and remove outliers from measured data as is common practice when using other regression techniques [12, 15]. The LOOCV performance metric and R^2 were calculated for both regression models when an outlier was present and are shown in Table 3. When choosing the best performing regression model, the R^2 statistic incorrectly identified the classical regression model as having marginally higher performance despite the regression line not following the data trend. The LOOCV performance metric is 13% lower for the Bayesian regression model and thus correctly identified the Bayesian regression model as the best choice for future predictions. This is because the LOOCV performance metric actually tests the regression models ability to predict new data whilst R^2 only indicates how well the regression line fits the existing data.

The relative performance of the classical and Bayesian linear regression models were also compared for a small dataset by using a subset of data from just four of the crossply coupons. The two techniques produce the same regression line but substantially different intervals as shown in Fig. 9. When calculating the allowable damage size for a coupon with a residual strength of 32 Nm the classical regression model over predicts the allowable damage size by almost 40%. In contrast as the data set is small the uncertainty on the Bayesian regression predictions are high and thus a conservative allowable damage size is determined that is 32% smaller than when the damage size is calculated using the full dataset. This implies that classical regression tends to under-estimate the prediction uncertainty in the presence of limited or small datasets while the Bayesian analysis is conservative. This occurs because the normalisation, described in equations (2) and (3), ensures that the spread, σ_S of the measured residual strengths around the regression line are always towards the lower end of the corresponding prior distribution, shown in Fig. 1 and Table 1, and thus as more coupons are tested the credible interval will collapse towards the measurements. These conservative estimates of the allowable damage size ensure that unexpected failures are unlikely.

The Gibbs sampling algorithm used to perform the Bayesian regression took on average 33 s to fit the regression model using a PC with an Intel Core i7-960 processor. The classical regression was performed using the same data in approximately 0.02 s on the same computer. The disparity in computational speed is due to classical linear regression using analytical equations whilst Bayesian linear regression uses iterative Gibbs sampling. However, the application of the Bayesian regression model is simpler than a computational mechanics model and allows damage prognoses to be made quickly whilst incorporating all of the uncertainties. In this study ultrasound C-scans were correlated to residual bending strength but the approach is generic and any measurement from a non-destructive evaluation technique could be correlated to an important performance metric, such as remnant fatigue life.

7. Conclusions

Bayesian linear regression has been used to create a model of the residual strength in bending after impact of carbon fibre composite coupons, based on ultrasound measurements in both crossply and quasi-isotropic laminates. The performance of the model was found to be more robust than classical linear regression. In more detail, C-scan data was generated for twenty-five quasi-isotropic coupons and twelve crossply coupons that had been subject to drop-weight impacts in the range 0 to 15J using a 20mm diameter tup. The coupons were tested to failure in four-point bending in order to determine their residual strength. A Bayesian linear regression model was fitted to the measured ultrasound and strength data using Gibbs sampling and the performance of the model evaluated using the Leave-One-Out-Cross-Validation (LOOCV) performance metric.

The use of the area of the damaged zone observable in the C-scan was found to yield a smaller uncertainty in the predictions from the model than using either the length or width of the damage zone. The Leave-One-Out-Cross-Validation (LOOCV) performance metric was used to estimate the average uncertainty of future predictions and found that for both the crossply and quasi-isotropic specimens using damage area results in prediction uncertainties that are approximately 25% smaller than using either damage length or damage width.

It was found that a classical linear regression model tended to underestimate the uncertainty in predictions of residual strength and thus overestimate the allowable damage size when only a limited amount of experimental measurements was available. In the example given in this paper, this overestimation was 40%. In these circumstances, the Bayesian linear regression model provided conservative estimates of the prediction uncertainty thus underestimating the allowable damage size. This implies that the use of the Bayesian model might lead to the rejection of some serviceable components but it is unlikely to compromise safety. An additional advantage of the Bayesian approach is that as more experimental data becomes available the model can be updated using the original posterior distributions of the model's parameters as updated prior distributions and generating updated posterior distributions using the new data. In the presence of outliers in the measured data, Bayesian regression was shown to provide accurate estimations of prediction uncertainties whereas the classical linear regression provided very conservative estimations with a 95% confidence interval that was almost four times larger than the interval calculated using Bayesian regression. Bayesian regression was still able to accurately estimate the parameters of the function used to generate the data containing outliers whilst classical regression estimates of gradient were different by 26%. The R^2 statistic in the presence of data outliers wrongly identified the classical regression model as being more adequate, unlike LOOCV.

Normalisation of the measured data was used to provide a generic approach to implementing the Bayesian regression model using Gibbs sampling. However, the main drawback of both the Bayesian regression model and the LOOCV performance metric is the high computation time required for the calculations. Once the calculations have been completed the damage assessments can be performed rapidly using residual strength predictions based on non-destructive evaluation (NDE), without the need for expert interpretation or analysis. It is proposed that the Bayesian model has the potential to permit improved decision-making on the serviceability of damaged composite components as a result of more robust and reliable predictions of residual behaviour and accurate estimates of the uncertainty in the prediction, which is likely to reduce unnecessary repairs and replacements and thus decrease maintenance costs and down-time.

References

1. US Department of Defense: MIL-HDBK-787 Nondestructive Testing Methods of Composite Materials – Ultrasonics (1988)

2. Ball, R.J., Almond, D.P.: The detection and measurement of impact damage in thick carbon fibre reinforced laminates by transient thermography. NDT&E Int. **31**(3), 165-173 (1998)

3. Patterson, E.A., Feligiotti, M., Hack, E.: On the integration of validation, quality assurance and non-destructive evaluation. J. Strain. Anal. Eng. Des. **48**(1), 48-58 (2013)

4. Smith, R.A., Nelson, L.J., Xie, N., Fraij, C., Hallett, S.R.: Progress in 3D characterisation and modelling of monolithic carbon-fibre composites. Insight Non-Destr. Test. Cond. Monit. **57**(3), 131-139 (2015)

5. ASTM International: ASTM D7137/D7137M-12 Standard Test Method for Compressive Residual Strength Properties of Damaged Polymer Matrix Composite Plates. ASTM International, West Conshohocken (2012)

6. Koo, J., Choi, J., Seok, C.: Evaluation for residual strength and fatigue characteristics after impact in CFRP composites. Compos. Struct. **105**, 58-65 (2013)

7. Prichard, J.C., Hogg P.J.: The role of impact damage in post-impact compression testing. Composites **21**(6), 503-511 (1990)

8. Caprino, G.: On the prediction of residual strength for notched laminates. J. Mater. Sci. **18**, 2269-2273 (1983)

9. Stone, D.E.W., Clarke, B.: Ultrasonic attenuation as a measure of void content in carbon-fibre reinforced plastics. Non-Destruct. Test. **8**(3), 137-145 (1975)

10. Kruschke, J.: Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. Elsevier Science, Amsterdam (2014)

11. Hawkins, D.M.: Identification of Outliers. Chapman and Hall, London (1980)

12. Dervilis, N., Worden, K., Cross, E.J.: On robust regression analysis as a means of exploring environmental and operational conditions for SHM data. J. Sound Vib. **347**, 279-296 (2015)

13. Schoeppner, G.A., Abrate, S.: Delamination threshold loads for low velocity impact on composite laminates. Compos. Part A-Appl. Sci. Technol. **31**(9), 903-915 (2000)

14. Abrate, S.: Impact on composite structures. Cambridge University Press, Cambridge (2005)

15. Rousseeuw, P.J.: Least Median of Squares Regression. J. Am. Stat. Assoc. **79**, 871-880 (1984)

16. Lange, K.L., Roderick, J.A.L., Jeremy, M.G.T.: Robust Statistical Modeling Using the t Distribution. J. Am. Stat. Assoc. **84**, 881-896 (1989)

17. US Department of Defense: MIL-HDBK-17-1F Composite Materials Handbook, Volume 1 - Polymer Matrix Composites Guidelines for Characterization of Structural Materials (2002)

18. Ang, A.H.S., Tang, W.H.: Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering. Wiley, New Jersey (2007)

19. Yang, F.J., Cantwell, W.J.: Impact damage initiation in composite materials. Compos. Sci. Technol. **70**(2), 336-342 (2010)

20. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. MIT Press, Boston (2006)

21. Plummer, M.: JAGS: Bayesian Graphical Models using MCMC. Ver 3.4.0. http://mcmc-jags.sourceforge.net/ (2015)

22. Lu, D., Ye, M., Hill, M.: Analysis of regression confidence intervals and Bayesian credible intervals for uncertainty quantification. Water Resour. Res. **48**(9) (2012)

23. Witten, I.H., Frank, E., Hall, M.A.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, Boston (2011)

24. Gong, W., Chen, J., Patterson, E.A.: Buckling and delamination growth behaviour of delaminated composite panels subject to four-point bending. Compos. Struct. **138**, 122-133 (2016)

Table 1 The values used to define the vague prior probability distributions for the regression model parameters, these values are suggested in [10].

Parameter	Prior Value
K	29
M_c	0
S_c	100
M_m	0
S_m	100
L	1×10^{-3}
H	1×10^{3}

Table 2 LOOCV performance metrics indicating the average uncertainties of residual strength predictions for crossply and quasi-isotropic specimens using three different damage metrics to make predictions. The $\boldsymbol{R^2}$ statistics for the same regression models are also shown.

	Bayesian Regression				**Classical Regression**			
	Crossply		**Quasi-isotropic**		**Crossply**		**Quasi-isotropic**	
Damage Metric	**LOOCV (Nm)**	$\boldsymbol{R^2}$ **(-)**	**LOOCV (Nm)**	$\boldsymbol{R^2}$ **(-)**	**LOOCV (Nm)**	$\boldsymbol{R^2}$ **(-)**	**LOOCV (Nm)**	$\boldsymbol{R^2}$ **(-)**
Damage Area	2.517	0.830	5.325	0.763	2.500	0.830	5.341	0.763
Damage Length	3.388	0.697	6.230	0.678	3.385	0.697	6.232	0.678
Damage Width	3.290	0.705	6.923	0.603	3.288	0.705	6.923	0.603

Table 3 LOOCV and $\boldsymbol{R^2}$ performance metrics for Bayesian and classical regression applied to data containing an outlier.

Performance Metric	**Bayesian Regression**	**Classical Regression**
LOOCV (Nm)	5.72	6.58
R^2 (-)	0.791	0.790

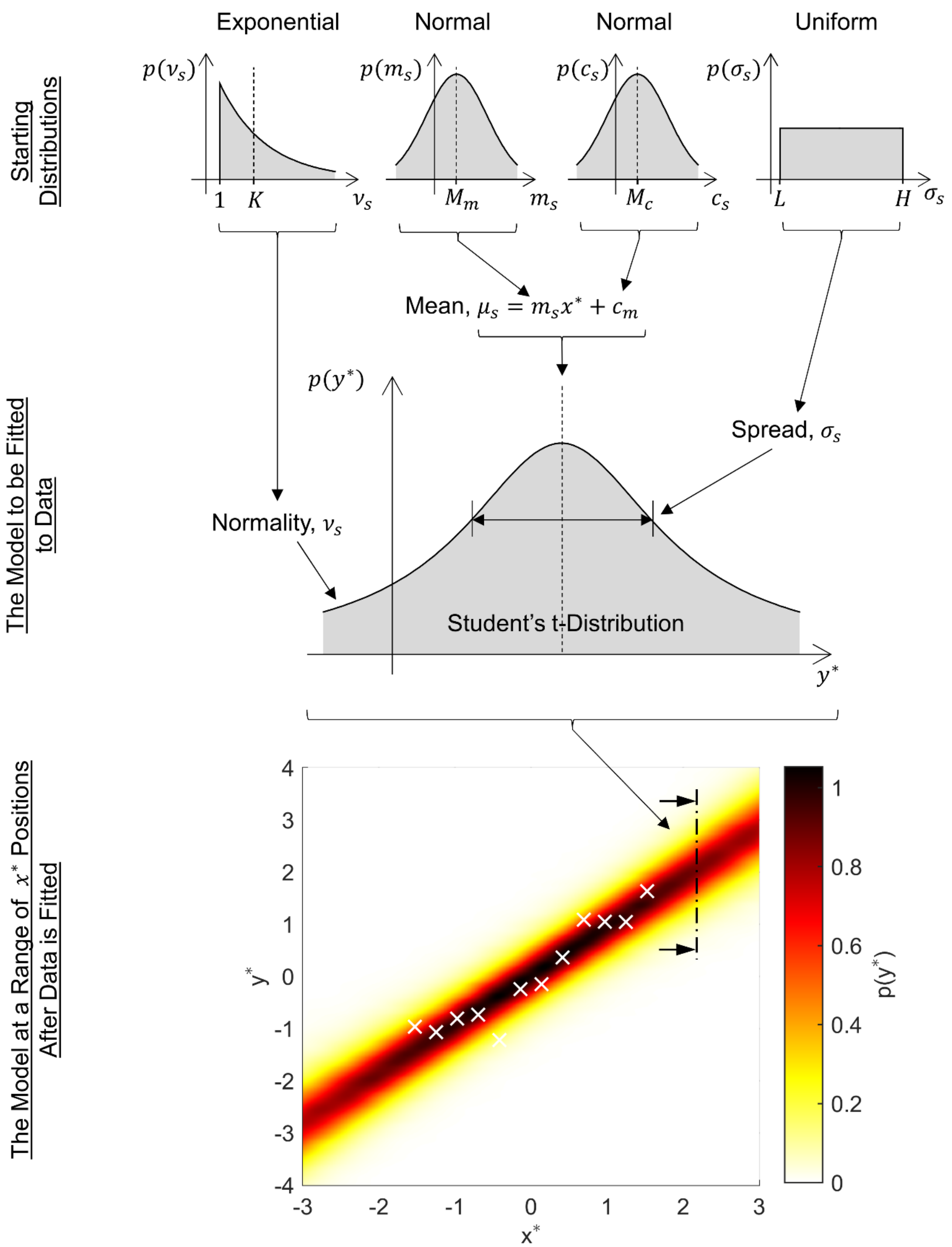

Fig. 1 Schematic diagram showing how the regression model (bottom) is formed from a linear regression line with the data distributed around it in the form of a t-distribution (middle). At the top are the initial or prior probability distributions for the values of the model parameters.

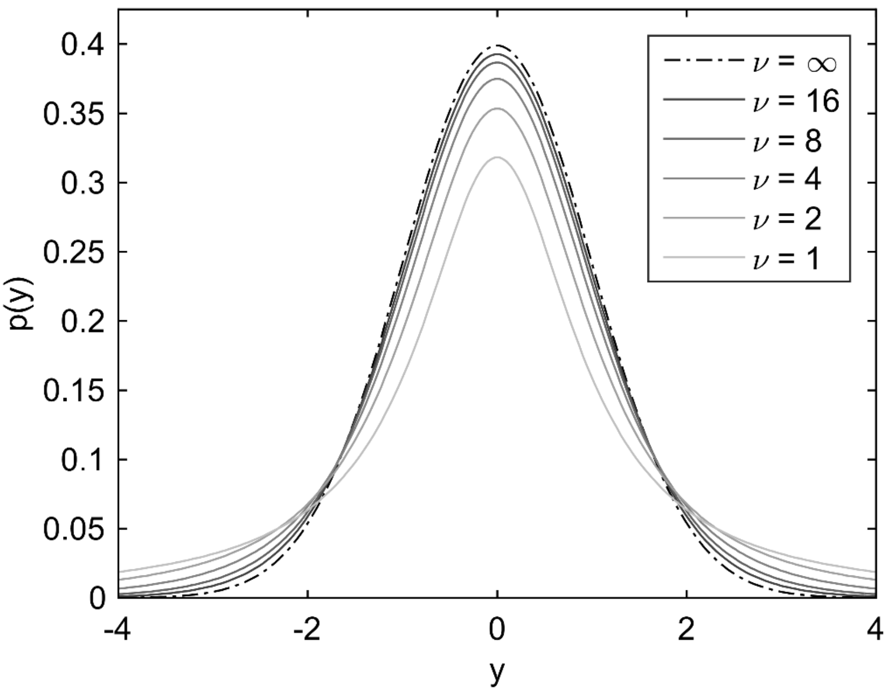

Fig. 2 Illustration of the probability density function for the t-distribution as a function of the normality parameter *v* for a mean, $\boldsymbol{\mu} = \mathbf{0}$ and spread, $\boldsymbol{\sigma} = \mathbf{1}$. As *v* tends to infinity, the t-distribution converges to a normal distribution.

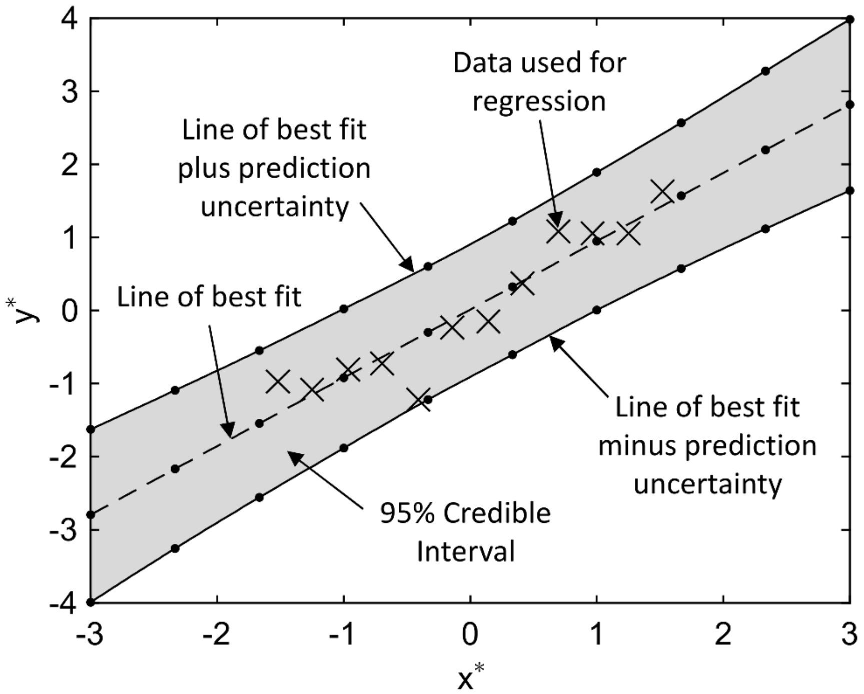

Fig. 3 A typical predictive distribution of residual strength, $y*$, as a function of ultrasound measurement, $x*$, based on a Bayesian linear regression model fitted to the measured data values (crosses) with prediction uncertainties and a 95% credible interval. The dots on the three lines indicate the locations at which percentiles of the predictive distribution were calculated. The lines are spline curves interpolated through the quantified points.

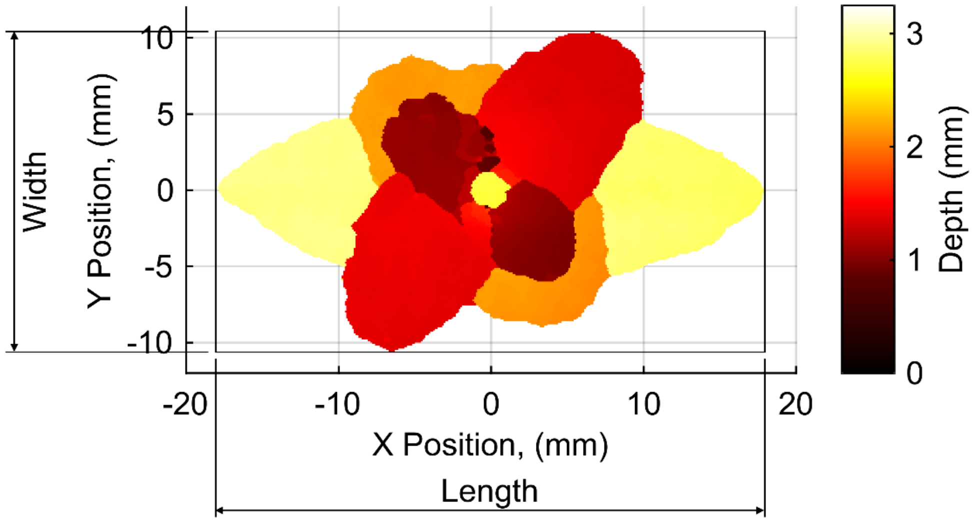

Fig. 4 A typical time-of-flight C-scan of impact damage in a quasi-isotropic composite laminate, showing the damage metrics used. The damage area was defined as the projected area of all the delaminations when viewed in the C-scan. Colour is used to indicate the depth of delaminations from the impacted surface.

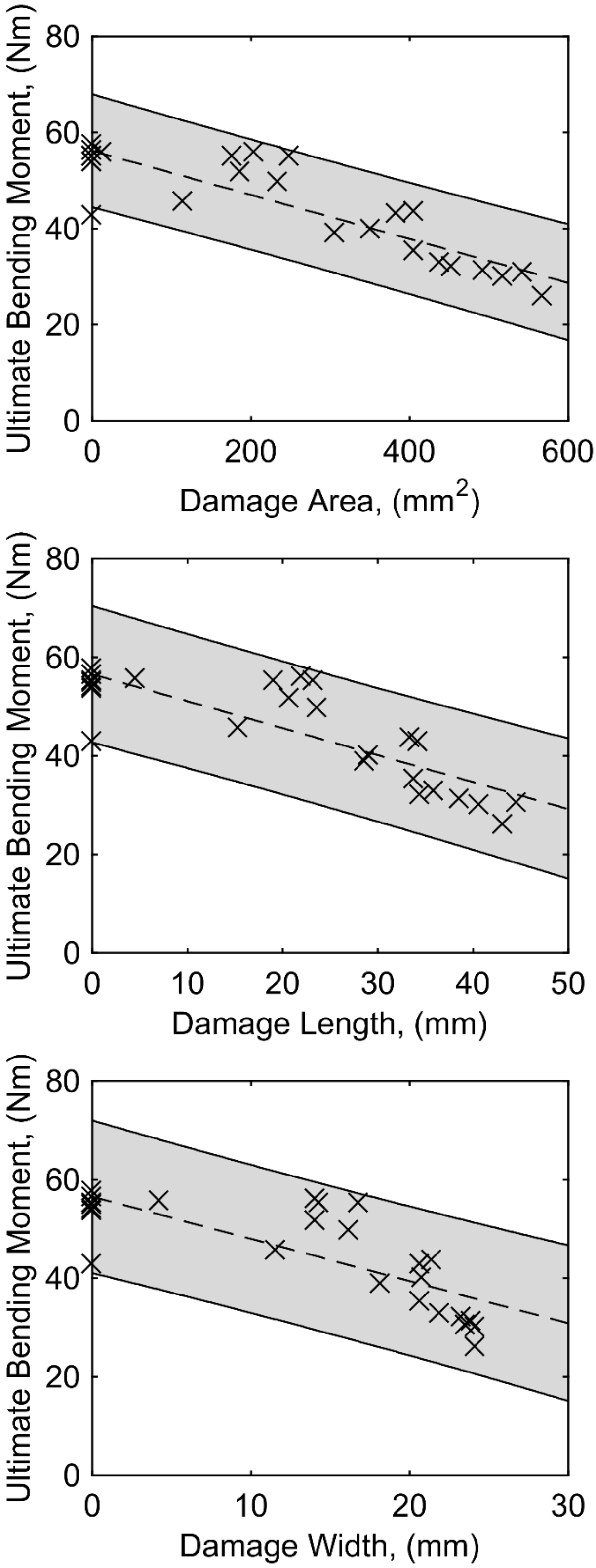

Fig. 5 Residual strength predictions made using Bayesian linear regression for impacted quasi-isotropic specimens using ultrasound measurements of damage area (top), length (middle) and width (bottom) as the damage metric. The size of the 95% credible interval (grey shading) indicates that the uncertainty is smallest when the area of damage was used as the damage metric which concurs with the LOOCV performance metric data in Table 2.

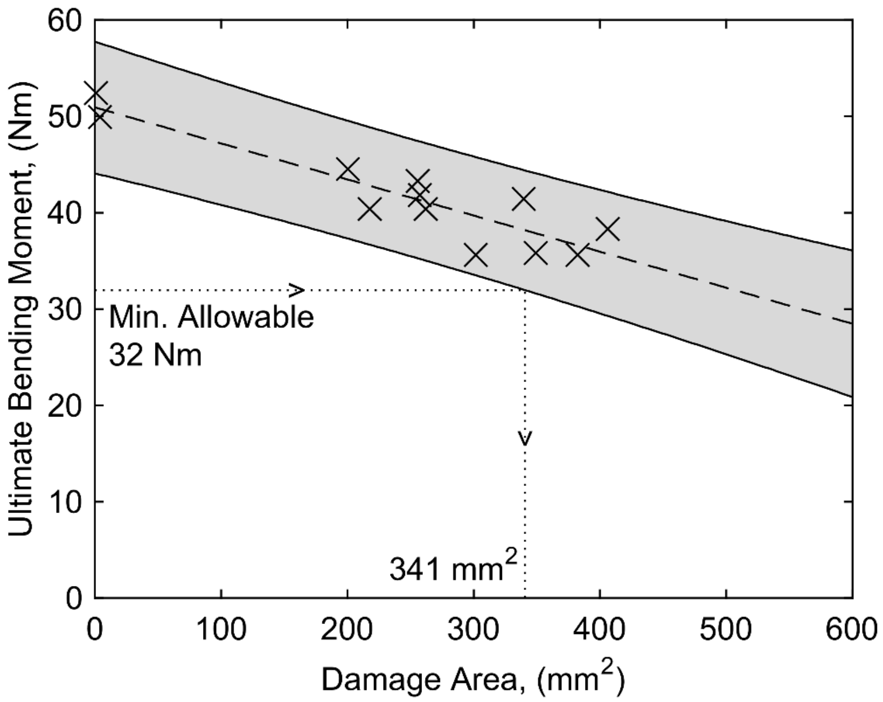

Fig. 6 Residual strength predictions made using Bayesian linear regression for impacted crossply specimens using the damage area from the ultrasound measurements at the damage metric together with the 95% credible interval (grey shading). The dotted lines indicate an exemplar minimum residual bending strength and the corresponding maximum allowable damage area for coupons with a probability of failure of less than 2.5%.

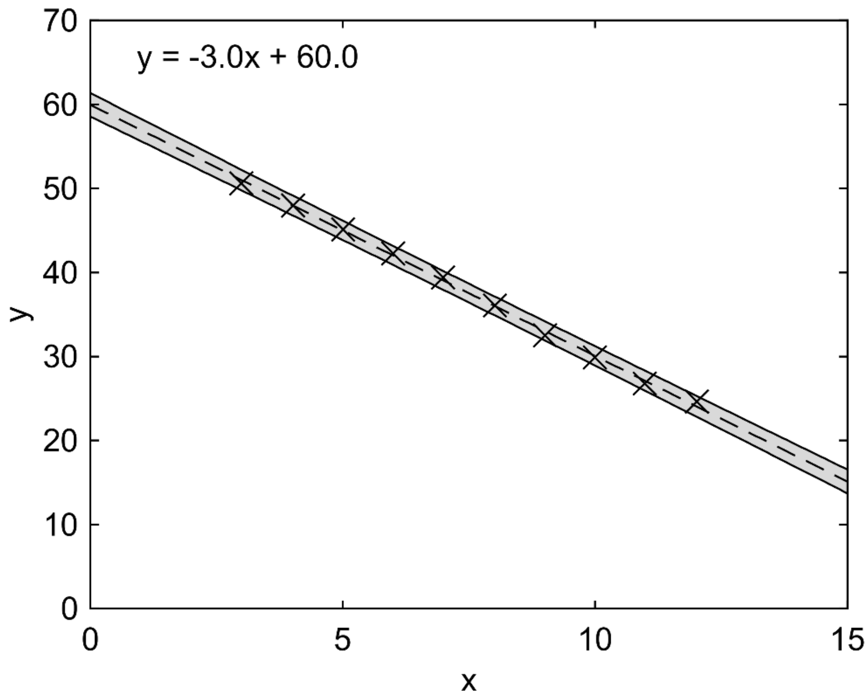

Fig. 7 Bayesian linear regression based on artificial data generated using the linear function in equation (6) with normally distributed measurement noise. The grey region shows the credible interval for the regression line. The equation for the line-of-best-fit (dashed line) is shown in the top-left corner of the figure.

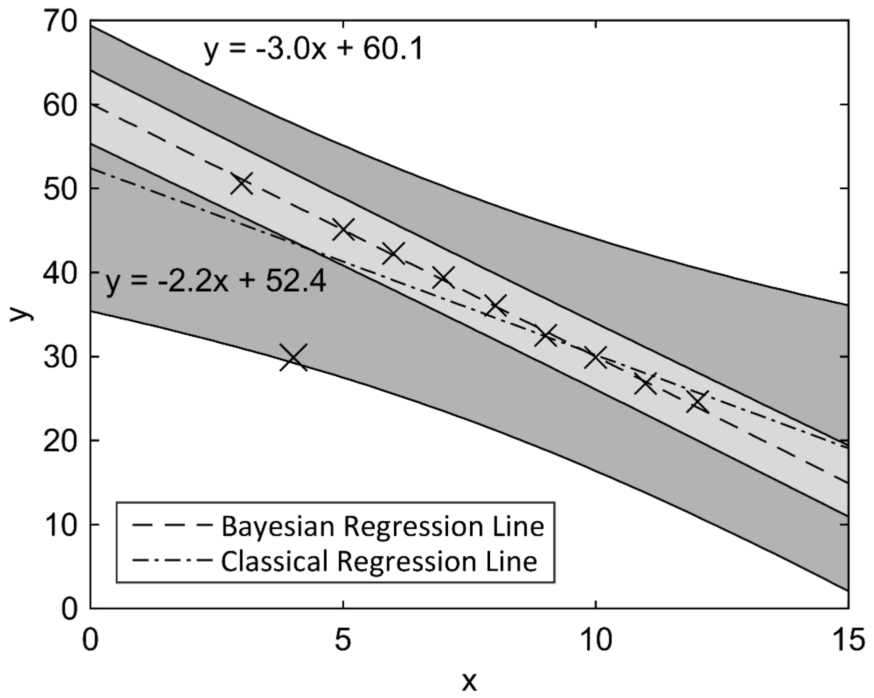

Fig. 8 Classical (dashed and dotted line with dark grey shading) and Bayesian (dashed line with light grey shading) linear regression and corresponding confidence and credible intervals based on the artificial data in Fig. 7 with the addition of an outlier at (4, 30). The equations for the Bayesian and classical lines of best fit are shown top-left and centre-left respectively.

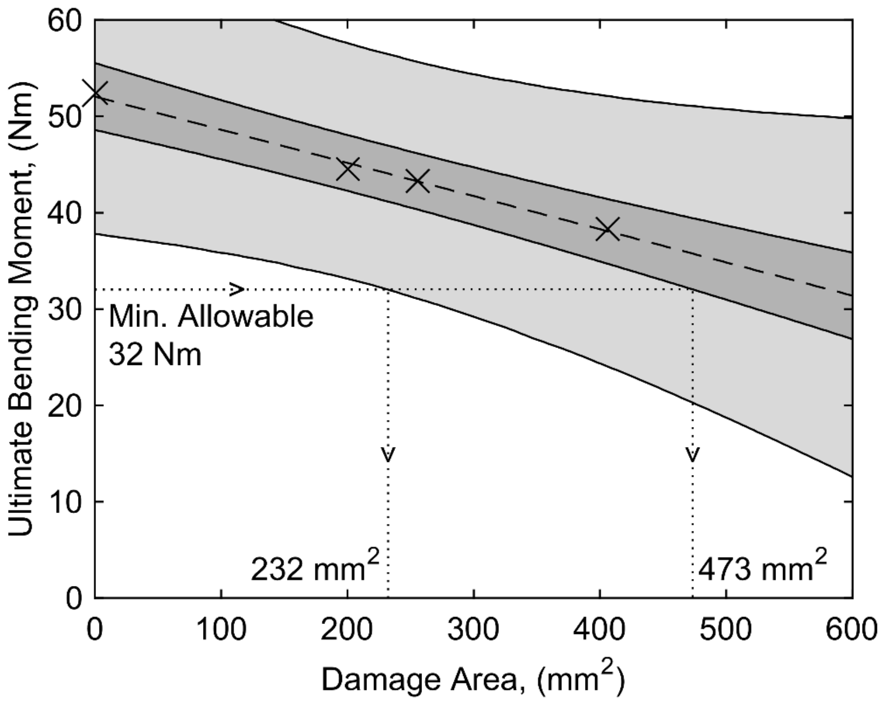

Fig. 9 Residual strength predictions made using Bayesian linear regression for a small set of four impacted quasi-isotropic specimens. Damage predictions were based on ultrasound measurements of damage area. The wide light grey region is the Bayesian regression credible interval. The narrow dark grey region is the classical regression confidence interval. The dotted lines indicate exemplar minimum residual bending strength and the corresponding maximum allowable damage area for coupons with a probability of failure of less than 2.5% calculated using Bayesian and classical regression.