JOURNAL OF AIRCRAFT
Vol. 39, No. 4, July–August 2002

Bayesian Updating of Damage Size Probabilities for Aircraft Structural Life-Cycle Management

David T. Rusk*
Naval Air Systems Command, Patuxent River, Maryland 20670-1906
Kuen Y. Lin†
University of Washington, Seattle, Washington 98195-2400
David D. Swartz‡
Federal Aviation Administration, Renton, Washington 98055
and
Greg K. Ridgeway§
Rand Corporation, Santa Monica, California 90407-2138

A statistical approach to estimating the probabilistic distribution of composite damage sizes using aircraft service inspection data has been investigated. Bayesian updating methods were implemented to revise baseline composite damage size distributions using damage size data from the Federal Aviation Administration's Service Difficulty Reporting System. Updating was performed on the Boeing 757 and 767 wing composite trailing-edge devices, elevators and rudders, with the results demonstrating that the assumed baseline damage size estimates are conservative in nearly all cases. Component failure probabilities were recalculated using the updated damage size distributions, and these results show an overall improvement in reliability for the damage mechanisms analyzed. The results of the analysis demonstrate that an inspection and maintenance program that reports damage characteristics can be used to monitor the reliability of damage tolerant structures on a quantitative statistical basis. Recommendations are also made for improving current inspection data reporting systems, which would enhance the ability to gather detailed information on the characteristics of each structural damage event.

Nomenclature

A	=	random variable for damage size
a	=	sample damage size from domain A
a_c	=	critical damage size
a_{50}	=	median detection probability for probability of detection models
$\bar{a}$	=	sample mean of damage sizes
D	=	binary random variable for damage detection state (1 indicates damage is detected)
$E[\,]$	=	expected value of quantity in brackets
$f_A(a)$	=	probability density function of A
g	=	importance-sampled probability density function
k	=	shape parameter for log-odds probability of detection model
L	=	likelihood function
$\overline{\ln a}$	=	sample mean of the log of damage sizes
m	=	importance sample size
n	=	sample size of damages used for updating
$P_D(a)$	=	probability of detection for damage size a
$P(Y)$	=	probability of Y
$\bar{P}_D$	=	sample mean of the log of probabilities of detection
$p(a)$	=	probability density function of actual damage size
$p_0(a)$	=	probability density function of detected damage size
R	=	reliability
w	=	importance weight factor
α	=	shape parameter for prior distribution of gamma model parameter θ
β	=	shape parameter for Weibull distribution of damage sizes, or scale parameter for prior distribution of gamma model parameter θ
θ	=	scale parameter for actual damage size distributions
ξ	=	truncation value for detected damage size distribution
σ	=	shape parameter for lognormal probability of detection model
τ	=	shape parameter for gamma distribution of damage sizes

Subscripts

n	=	normalized values of importance weight factors
u	=	updated distribution of model parameters
0	=	earlier distribution of model parameters

Received 10 April 2001; presented as Paper 2001-1646 at the AIAA 42nd Structures, Structural Dynamics and Materials, Seattle, WA, 16–19 April 2001; revision received 11 October 2001; accepted for publication 17 December 2001.

*Aerospace Engineer, Structures Division, Code 4.3.3.1, Building 2187, Suite 2340A, 48110 Shaw Road Unit 5. Senior Member AIAA.

†Professor, Department of Aeronautics and Astronautics, Box 352400. Associate Fellow AIAA.

‡Metallurgy/Composites Technical Specialist, Seattle Manufacturing Inspection District Office, ANM 1085, 2500 East Valley Road, Suite C-2. Member AIAA.

§Associate Statistician, Statistics Group, P.O. Box 2138.

Introduction

THE nondeterministic approach to damage tolerance is beginning to gain acceptance as a means of quantifying safety and reliability in aircraft structures. Probabilistic methods applied to damage-tolerant designs enable the characterization of uncertainty associated with damage accumulation and growth, inspection reliability, and residual strength behavior of the structure. By the use of these methods, the safety and reliability of a structure can be assessed on a quantitative basis, allowing aircraft manufacturers, operators, and flight certification authorities to evaluate the risk associated with structural failures in an aircraft fleet. A simplified probabilistic approach for quantifying the reliability of damage-tolerant structures has been previously investigated by Lin et al.[1] Structural reliability for a single inspection opportunity is defined as the compliment of the probability that a single flaw size larger than the critical flaw size for residual strength of the structure exists and that the flaw will not be detected. The methodology derived from this definition is sufficient for use on composite structures designed for "no damage growth" certification criteria. One of the most challenging aspects of applying this or any other probabilistic methodology to a damage tolerance problem is the determination of the appropriate distribution of actual damage sizes for each damage mechanism the structure will see in service.

During the design phase and early operational life of the structure, little damage size data may be available because it is very difficult to simulate, in laboratory experiments, all of the conditions that cause damage to accumulate on an aircraft structure. Under the current philosophy of commercial and military aircraft operations, periodic scheduled and unscheduled airframe inspections are typically carried out by maintenance personnel to ensure the airworthiness of the fleet. These inspections provide a good opportunity to collect damage size information on all of the structural damage that accumulates in a fleet of aircraft. One of the benefits of utilizing a probabilistic approach to damage tolerance is that Bayesian statistical tools can be used to update the damage size probability distributions when new data become available. This technique was previously demonstrated by Harris for initial crack depths on a center-cracked panel.[2] In this research, Bayesian updating techniques are used to revise initial estimates of damage size distributions using composite damage size data from the Federal Aviation Administration's Service Difficulty Reporting System (SDRS) database.

Reliability Formulation

A definition for damage-tolerant reliability was previously derived in Ref. 1 and will be used here in modified form for the subsequent analyses. The reliability equations rely on a probabilistic characterization of actual structural damage sizes and damage detection capability for the inspection technique being used:

$$R = 1 - P(A \geq a_c, D = 0) = 1 - \mathrm{PF} \tag{1}$$

$$\mathrm{PF} = \int_{a_c}^{\infty} p(a)[1 - P_D(a)]\,\mathrm{d}a \tag{2}$$

This definition, depressed in terms probability of failure (PF), assumes that only a single flaw is present in the structure at a single inspection opportunity and that the flaw is not growing with time. The definition should, thus, be sufficient for characterizing composite structures designed under "no damage growth" certification criteria. An additional assumption is that a single characteristic dimension can describe the damage mechanism being modeled.

The reliability equation itself is independent of the particular damage mechanism being modeled because all of the configuration-specific information in the problem is contained within the parameters of the probability distributions. Therefore, the choice of appropriate probability models is important to accurately describe the nature of uncertainty for the specific problem of interest. Berens and Hovey[3] and Berens[4] have conducted significant research to characterize probability of detection (POD) models for cracks in metal aircraft structures. The results of these studies show that a cumulative lognormal distribution [Eq. (3)] can be used to model the mean hit/miss response data from crack detection experiments:

$$P_D(a) = \int_0^a \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}\,\ell n^2\left(\frac{x}{a_{50}}\right)\right]\mathrm{d}x \tag{3}$$

Efforts to extend this research to the determination of POD characteristics for composite damage inspection techniques have so far been minimal. Thus, for the purposes of this analysis, the cumulative lognormal POD model is assumed to apply equally well to composite damage types. The form of the POD model is such that the probability of damage detection goes to zero as the damage size approaches either zero or some minimum detection threshold. This means that the frequency of occurrence for structural damage sizes is not completely observable over the range of possible damage sizes, and so the exact shape of $p(a)$ can never be completely characterized. A choice of probability models for $p(a)$ that can accommodate this type of uncertainty is the gamma [Eq. (4)] or Weibull [Eq. (5)] probability density functions (PDFs). Both of these models can assume either an exponential form, or a form that is bounded near zero, depending on the value of the shape factor term:

$$\mathrm{Gam}(a; \tau, \theta) = [1/\theta^{\tau}\Gamma(\tau)]a^{\tau-1}\exp(-a/\theta) \tag{4}$$

$$\mathrm{Wei}(a; \beta, \theta) = (\beta/\theta^{\beta})a^{\beta-1}\exp[-(a/\theta)^{\beta}] \tag{5}$$

Any damage size data collected from structural inspections represent a random sample, not from the actual damage size distribution, but from the detected damage size distribution, which is a product of the actual damage size distribution and the detection probability of the particular inspection technique used:

$$p_0(a) = \frac{p(a)P_D(a)}{\int_0^{\infty} p(a)P_D(a)\,\mathrm{d}a} \tag{6}$$

The analytic detected damage size models are shown in Eq. (7) for a gamma actual damage size distribution and in Eq. (8) for a Weibull actual damage size distribution:

$$p_0(a) = \frac{\mathrm{Gam}(a; \tau, \theta)P_D(a)}{E[P_D(a)]} \tag{7}$$

$$p_0(a) = \frac{\mathrm{Wei}(a; \beta, \theta)P_D(a)}{E\,[P_D(a)]} \tag{8}$$

Baseline Damage Size Data

One of the most difficult aspects of applying a probabilistic approach to damage-tolerant structural analyses is in determining the appropriate distribution of actual damage sizes that will accumulate on a structure in service. At present, little quantitative data exist on the damage size characteristics of various composite structural applications. One of the few published examples of such data was compiled by Gray and Riskalla.[5] An excerpt of these data is reprinted in Table 1 in modified form and was used to derive damage size distributions for the composite sandwich reliability analysis of Ref. 1. The data will also be used here to provide a baseline estimate of damage size distributions for existing commercial aircraft composite structures.

Before fitting the baseline damage size data to Eqs. (7) and (8), the POD model parameter values for each damage inspection technique must be known. The log-odds POD model parameter values previously assumed in Ref. 1 will be used here, after being transformed to the cumulative lognormal POD parameters by the transformation relation outlined by Berens in Ref. 4:

$$k = \pi\big/\sqrt{3}\sigma \tag{9}$$

The transformed POD parameter values are listed in Table 2 and the resulting POD curves are shown graphically in Fig. 1. Parameter values were chosen to represent detection probabilities that can

Table 1 Composite damage size data from Ref. 5

	Damage size, in.(mm)		
Damage type	<1.5, %	1.5–3.0, %	>3.0, %
Hole damage	51.4	34.3	14.3
Delaminations	11.1	31.1	57.8
Cracks	30	30	40

Table 2 Parameters for cumulative lognormal detection probability distributions

Damage type	a_{50}	σ
Hole damage	0.5	0.726
Delaminations	2.0	0.698
Delaminations (NDE)	0.5	0.698
Cracks	0.8	1.01

Table 3 Model parameters for actual damage size distributions

	Gamma $P(a)$		Weibull $P(a)$	
Damage type	τ	θ	β	θ
Hole damage	1.26	1.09	1.10	1.40
Delaminations	0.834	2.63	0.919	2.16
Cracks	0.752	2.85	0.869	2.07

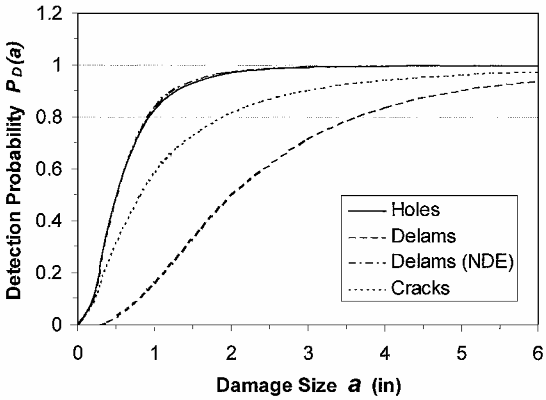

Fig. 1 Cumulative lognormal detection probability functions.

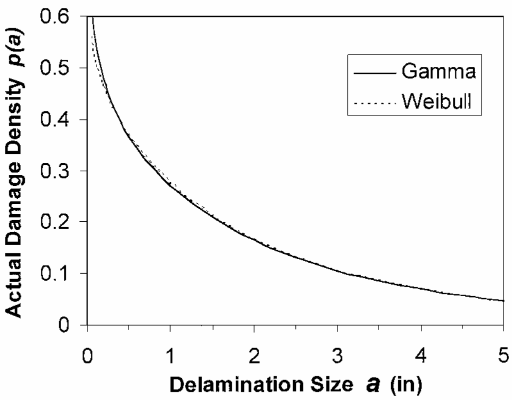

Fig. 2 Actual delamination size distributions for gamma and Weibull models.

be reasonably obtained for operational inspections and that would likely result in the distribution of damage sizes observed in Ref. 5. The POD curves represent visual inspection capability for all hole and crack damage and a combination of visual and tap testing capability for delaminations. In addition, a POD curve representing an automated nondestructive evaluation (NDE) method for detecting delaminations is also assumed.

To calculate the parameter values of $p(a)$ from the baseline damage data, Eqs. (7) and (8) must be integrated numerically over the damage sizes corresponding to the cumulative probability of occurrence data in Table 1. A double-precision FORTRAN program was written to solve for the $p(a)$ model parameters using the secant method for sets of nonlinear algebraic equations.[6] The function integrations were solved using the SLATEC subroutines DQAG and DQAGI.[7] Integrations were carried out to a relative precision of 10^{-12}. Parameter values converged to a cumulative absolute error of 10^{-6} and were solved for each damage type, for both gamma and Weibull actual damage size models. Results are listed in Table 3 and are shown graphically for the delamination case in Fig. 2. Figure 2 demonstrates that the choice of either a gamma or Weibull PDF for $p(a)$ yields similar results for the damage type and the range of damage sizes shown.

SDRS Data

The damage size distributions derived from the data of Ref. 5 represent the aggregate response of a large number of composite structures to in-service damage. To determine the response of specific structural components to damage, more detailed data need to be collected for the individual structures of interest. Use of the damage-tolerant design philosophy requires periodic inspection to detect and repair structural damage, and the results from those inspections can be used to revise the baseline damage size distributions as new data are accumulated. The Federal Aviation Administration (FAA) requires commercial aircraft operators in the United States to submit their aircraft to periodic inspection and to report any failure, malfunction, or defect that threatens flight safety or exceeds allowable limits. This information is submitted to the FAA in the form of a malfunction or defect report, and the individual reports are collected in SDRS. Although not expressly designed for the purpose, SDRS can be used in some cases to obtain damage size data for individual structural components. Previous research efforts by Brewer utilized SDRS crack size data to estimate POD curves for inspection of metallic fuselage lap splice joints.[8] In this study, damage size data available from SDRS will be used to demonstrate Bayesian updating of prior detected damage size distributions. These data will also be used as a preliminary validation of the baseline composite damage densities derived from Ref. 5, for specific airframe structural components. A third purpose in using these data will be to demonstrate how the existing commercial aircraft maintenance infrastructure can be adapted to gather statistically useful data on the damage threat environment of in-service aircraft.

An archive of all SDRs covering the period from January 1990 to April 1999 was obtained from the FAA. Only records pertaining to structural problems on large commercial transports with a significant number of high-performance composite components were retained. All other records were deleted. The remaining records were searched by aircraft type for damage occurring on any major composite structures. Damage sizes and methods of detection are not required by the FAA to be submitted on structural damage reports. However, many inspectors choose to report this information anyway. The largest sample set of reported damage sizes found in the database was for the Boeing 757 and 767 airframes. The breakdown of damage events, reported damage sizes, and components affected are listed in Table 4 for the 757 and Table 5 for the 767. Analysis of the records with damage sizes indicated that detected disbond and delamination damage events were not reported consistently. Often, disbond damage is reported as delamination in the records. Dents, gouges, and general damage have no apparent delamination associated with them, and so are treated as a separate case. The detection method is usually not reported, and so is assumed to be visual unless otherwise stated.

Before using any of the SDR damage data in a statistical analysis, the limitations of the reported information must be addressed. Structural damage classified as a major repair is often handled through the FAA's Designated Engineering Representative, or directly by the airframe manufacturer. In those cases, an SDR may not be filed by the maintenance activity. This means that the SDRS database does not contain all incidences of major structural damage that occur on

Table 4 SDR damage data for 757 composite structure

	SDR Records	
Damage type	Damage events	Damage sizes
Ailerons, flaps, and spoilers		
Dents, gouges, and General damage	26	9
Cracks	16	5
Delaminations	39	9
Holes	76	9
Lightning strike	2	0
Elevators and rudders		
Dents, gouges, and general damage	3	0
Cracks	1	1
Delaminations	7	1
Holes	26	2
Lightning strike	5	0

Table 5 SDR damage data for 767 composite structure

Damage type	SDR records: Damage events	SDR records: Damage sizes
Ailerons, flaps, and spoilers		
Dents, gouges, and General damage	27	7
Cracks	11	6
Disbonds	3	0
Delaminations	32	9
Holes	8	1
Lightning strike	3	0
Wing trailing-edge skin panels		
Cracks	1	1
Delaminations	2	1
Elevators and rudders		
Dents, gouges, and general damage	14	3
Cracks	6	1
Delaminations	15	7
Holes	27	17
Lightning strike	2	1

a component in service. Also, only damage that is beyond the maximum acceptable limits is required to be reported. These limits are usually set by the airframe manufacturer's structural repair manual (SRM) and are specific to the damage type and location on each component. Damage sizes below these limits are not required to be reported and usually are not. As a result, damage sizes derived from the SDRS database do not represent a random sample from the overall damage size distribution of a component. These issues must be taken into account in the subsequent analysis, or the results may be significantly biased.

Bayesian Updating Formulation

For each damage type, damage size data from the SDRS database represent a sample from the detected damage size distribution. With the form of Eqs. (7) or (8) for $p_0(a)$, the model parameters of the actual damage size distribution can be updated. The damage size data do not represent a random sample from $p_0(a)$, however, because only damage sizes larger than the repair size limits are reported. If the size threshold for reporting damage is known a priori, then a particular damage size data point will represent a random sample only from the area of the detected damage PDF [Eqs. (7) or (8)] that is above the size threshold. This can be accounted for in probabilistic terms by the use of a truncated PDF, where the detected damage PDF can be expressed strictly as a function of damage sizes larger than the threshold value. The entire actual damage size distribution model can then be updated with data sampled strictly from the region beyond the threshold limit. Details of the required derivations, the resulting equations, and the solution methods utilized are listed in the Appendix. The gamma damage size model was chosen for this particular application to illustrate the methodology; however, the Bayesian updating method can be applied equally well to the Weibull damage size model, or any other model one wishes to choose.

Updated Damage Sizes

It is clear from Tables 4 and 5 that the small sample of damage sizes available from the SDR database precludes the ability to derive, with any significant degree of confidence, initial component damage size estimates from these data alone. However, with the Bayesian updating formulations just derived, SDR damage size data can be used to revise baseline probability distributions for each damage type. The baseline distributions used here were previously derived from the data of Ref. 5. For all of the damage types reported in Tables 4 and 5, only disbonds, delaminations, holes, and cracks were used to perform updating. Disbonds were grouped with delaminations because of the difficulty of sorting out the specific damage mechanisms from the individual records. The lightning strike damage on the 767 elevator was treated as a delamination. Four of the delaminations on the 767 were detected using (NDE)/ultrasound techniques, and these data points were accounted for in the updating calculation by using a likelihood function that incorporated multiple inspection techniques for a given damage type. SDR damage records that report damage sizes usually do not report damage shapes. For disbond, delamination, and hole records that only report a single dimension, the damage was assumed to be circular. If two dimensions were given, the damage was assumed to be elliptical. The damage size was then recorded as the diameter of a circular area equivalent to the area of the ellipse. Crack dimensions were assumed to be the overall crack length.

Each reported damage size has a maximum repair size limit associated with it, and the limits are typically set by the manufacturer's SRM. All of the reported damage sizes were cross checked with the appropriate Boeing SRM to determine the corresponding repair size limits. Without detailed dimensions of damage locations from the SDR data, and drawings of the affected part, it was difficult to ascertain which repair limit criteria caused the damage event to be reported. By the use of information in the SRM, criteria for choosing repair limits were established to provide a systematic approach to setting damage size thresholds based on the repair information reported in each SDR data record. For each reported damage size, the repair limit criteria used were as follows:

1) If the damage size is smaller than all SRM repair limits, the threshold is set to zero.

2) If the damage size is smaller than the permanent repair limits, and is repaired with an unspecified permanent repair, then the threshold is set to the interim repair limit.

3) For damage sizes above all SRM repair limits, if the repair type is not specified, then the threshold is set to the interim repair limit.

4) For large damage with no size limit specified on the repair type, the threshold is set to the largest repair limit for other repair types that is smaller than the reported damage size.

Applying the repair limit criteria gives a threshold value associated with the reported damage size. The threshold value will be in terms of the largest damage dimension and must be corrected when the damage area is noncircular. This was accomplished by calculating the aspect ratio of the reported damage area, then shrinking the damage area until the major axis is equal to the threshold limit, while keeping the aspect ratio constant. An equivalent circular diameter for the threshold limit is then calculated from the reduced major and minor axes dimensions. This technique reduces the effect of damage shape variation on the damage size results because the threshold values are calculated by holding the damage shape constant.

A double-precision FORTRAN program was written to carry out the updating using the damage size and threshold data. Integrations were carried out using the SLATEC subroutines DQAG and DQAGI, to a relative precision of 10^{-8}. Importance samples were drawn from a gamma distribution of the τ parameter and an inverse-gamma distribution of the θ parameter. The uniform random number generator used was the function DUNI from the NMS software package.[7] The gamma random number generator function used was a variation of the program listed in Ripley's Appendix for sampling from a standard gamma distribution.[9] The parameter values of the importance-sampled gamma distributions were manually optimized by trial and error to maximize the effective sample size (ESS) for a solution based on $m = 1000$. The final computational runs were performed using an importance sample size of $m = 100{,}000$. This large sample size was used so that histograms of the marginal posterior parameter distributions could be created. Generally, an importance sample size of only a few hundred gives sufficient solution accuracy for ESS values greater than 50%.

Analysis cases were divided into three categories. The first category is 767 damage, which is subdivided into two component groups of wing trailing-edge surfaces and empennage control surfaces. The second category is 757 damage, which consists of a single component group of wing trailing-edge surfaces. There were insufficient data points to perform a separate analysis of the 757 empennage control surfaces. The third category combines data for the 757 and 767 structures and is subdivided into wing trailing-edge and

Table 6 Bayesian updated composite damage size parameters for 757 and 767 component groups

	Baseline		25% COV			50% COV		
Component damage	τ	θ	τ	θ	ESS, %	τ	θ	ESS, %
767 flap delamination	0.834	2.63	0.735	2.52	97.0	0.567	2.57	93.0
767 flap crack	0.752	2.85	0.726	2.92	98.4	0.658	3.08	92.8
767 tail hole	1.26	1.09	1.11	0.815	85.0	1.25	0.666	59.8
767 tail delamination	0.834	2.63	0.730	2.46	93.7	0.579	2.44	90.9
757 flap hole	1.26	1.09	1.14	0.939	92.1	1.10	0.857	73.9
757 flap delamination	0.834	2.63	0.528	2.48	95.3	0.269	2.71	93.3
All flap hole	1.26	1.09	1.06	0.912	92.1	0.933	0.841	76.5
All flap delamination	0.834	2.63	0.496	2.52	94.3	0.239	2.81	92.7
All flap crack	0.752	2.85	0.686	2.73	96.3	0.580	2.79	87.7
All tail hole	1.26	1.09	1.02	0.789	84.4	1.03	0.668	62.7
All tail delamination	0.834	2.63	0.746	2.75	95.5	0.580	3.08	89.5

Table 7 Mean and standard deviation of updated actual damage size distributions for 757 and 767 composite components

	Baseline		25% COV		50% COV	
Component damage	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
767 flap delamination	2.195	2.404	1.857	2.166	1.414	1.859
767 flap crack	2.139	2.467	2.121	2.488	2.027	2.499
767 tail hole	1.382	1.230	0.9078	0.8600	0.8329	0.7447
767 tail delamination	2.195	2.404	1.795	2.101	1.414	1.859
757 flap hole	1.382	1.230	1.071	1.002	0.9383	0.8965
757 flap delamination	2.195	2.404	1.307	1.799	0.7311	1.409
All flap hole	1.382	1.230	0.9705	0.9410	0.7850	0.8125
All flap delamination	2.195	2.404	1.250	1.775	0.6697	1.371
All flap crack	2.139	2.467	1.875	2.264	1.617	2.123
All tail hole	1.382	1.230	0.8061	0.7974	0.6859	0.6770
All tail delamination	2.195	2.404	2.052	2.376	1.790	2.349

empennage control surface component groups. The earlier distributions of the actual damage size model parameters are characterized in terms of two levels of uncertainty. The low uncertainty level sets the coefficient of variation (COV) of the earlier parameter distributions to 25%. The high uncertainty level sets the COV to 50% for the earlier parameter distributions. These COV percentages were chosen arbitrarily because information on the qualitative uncertainty of the damage size data of Ref. 5 is unknown. Earlier mean values were taken from the baseline damage size distributions for each damage type. The updated mean parameter values, along with the ESSs for each analysis run, are shown in Table 6 for all categories analyzed. The results show that ESSs of 90% or greater were achieved for most of the analysis runs, with some exceptions. The effects of Bayesian updating on the baseline damage size distributions can be quantified by comparing the means and standard deviations of each damage size distribution. These results are tabulated for the damage cases analyzed and are listed in Table 7. All values have dimensions of inches.

In nearly every case examined, damage size updating using SDR data resulted in a significant reduction in the mean and standard deviation of the damage size distributions compared to their baseline values. The only exception to this is the updated standard deviation for 767 flap crack damage, which increases slightly as the earlier uncertainty level increases. Both flap delamination cases show a significant reduction in the updated mean and standard deviation values over the baseline values. The rate of decrease is faster in the 757 flap delamination case, mainly due to the smaller amount of scatter compared to the 767 flap delamination data set.

The effect that the Bayesian updated damage size distributions have on the reliability for each damage mechanism can be demonstrated by recalculating the PF [Eq. (2)] using the posterior parameter values for the actual damage size distributions. These integrations were carried out numerically to a relative precision of 10^{-12}. Selected results are plotted as a function of design critical damage size in Figs. 3–8. Bayesian updating of the damage sizes reduces the failure probability in nearly all damage cases studied. The only exception is the 767 flap crack case, where there was no significant

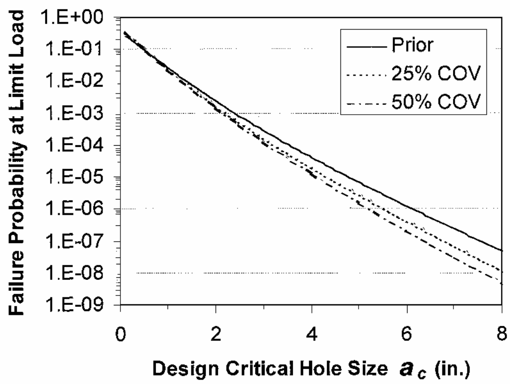

Fig. 3 Bayesian updated failure probabilities for 757 flap holes.

change in the failure probability curve due to updating. These results serve as a preliminary validation of the baseline composite damage size assumptions as applied to the structural components analyzed here. This validation is limited in scope, however, due to the small sample sizes of damage utilized, the data reporting uncertainties, and the assumptions of inspection POD characteristics.

Another noteworthy aspect of the analysis results is how Bayesian updating reduces the uncertainty in the parameter distributions of the damage size model. Histograms of the posterior marginal parameter distributions can be plotted by summing the normalized importance weights over intervals of the sampled parameter values. Examples of this type of plot are shown in Figs. 9 and 10 for the two preceding levels of uncertainty assumed. These charts show the relative reduction in variance of the parameter values over the earlier variances and also illustrate the shift in the distribution mean values due to updating.

The overall results of the Bayesian updating analysis demonstrate that damage size data from scheduled and unscheduled aircraft structural inspections can be effectively utilized to refine damage size

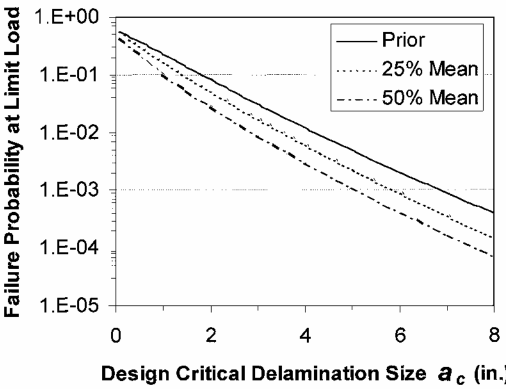

Fig. 4 Bayesian updated failure probabilities for 757 flap delaminations.

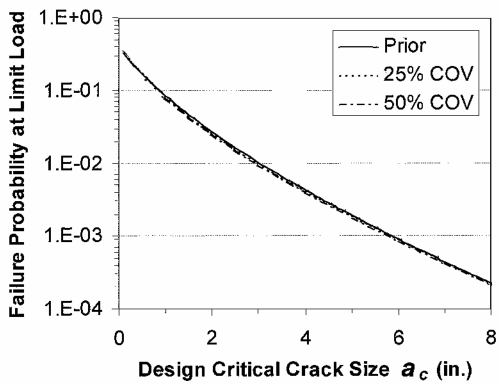

Fig. 5 Bayesian updated failure probabilities for 767 flap cracks.

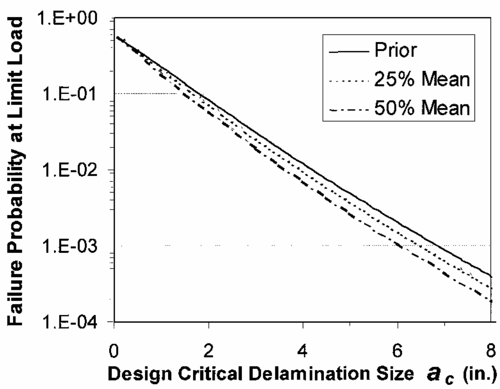

Fig. 6 Bayesian updated failure probabilities for 767 flap delaminations.

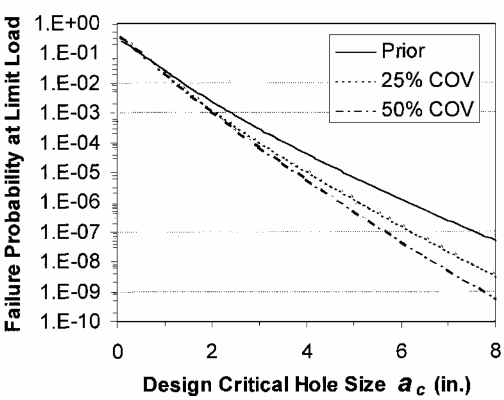

Fig. 7 Bayesian updated failure probabilities for 767 tail holes.

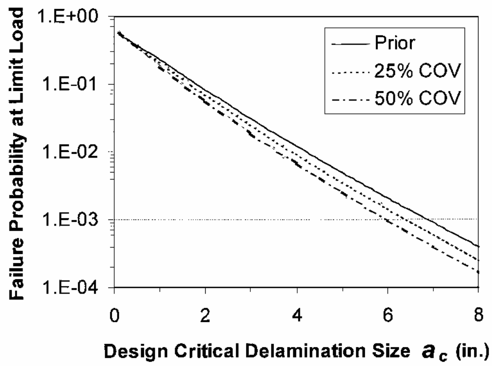

Fig. 8 Bayesian updated failure probabilities for 767 tail delaminations.

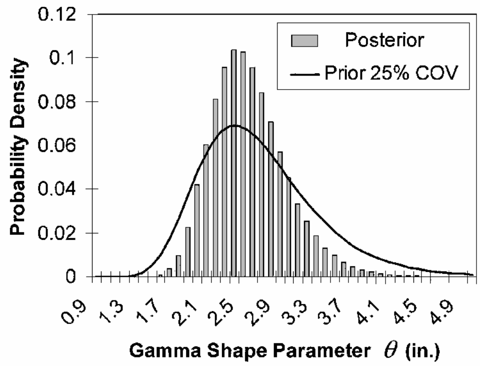

Fig. 9 Bayesian updated θ distribution for 757 and 767 flap delaminations (low uncertainty).

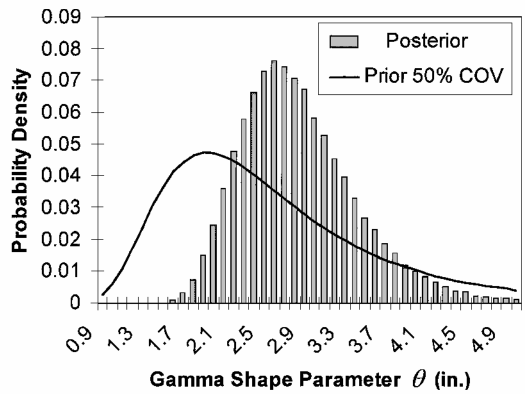

Fig. 10 Bayesian updated θ distribution for 757 and 767 flap delaminations (high uncertainty).

distributions on a quantitative basis. Even with the uncertainties associated with the damage size data reported in the SDR system, the results clearly demonstrate that significant reductions in damage size mean and scatter values are achievable, compared to conservative baseline values. These results translate into reduced failure probabilities for current designs and increase confidence that reliability-based methods can quantify uncertainty in damage-tolerant structural designs. This paper as a whole was derived from the work of Rusk in Ref. 10, in which further details of the analysis methods used, as well as a compilation of all relevant results, can be found.

Data Reporting Criteria

The results of the Bayesian updating analysis using SDR data demonstrate that an inspection and maintenance program that reports damage characteristics can be used to monitor the reliability of damage tolerant structures on a quantitative statistical basis. The criteria for reporting damage to SDRS poses some unique challenges when trying to use the data to update reliability predictions. Because only damage beyond the maximum repair size limit is reported, the volume of damage size data that accumulate are highly dependent on how large the repair size limits are in relation to the mean values of typical damage sizes. If the repair size limits cover nearly all damage sizes seen in service, only a few exceptional data points will be available for updating. It would, therefore, be difficult to make any meaningful statistical inferences from such a small data set. Ideally, one would like to have data on all damage events that are detected for every inspection opportunity. However, documenting all of this information could prove to be excessively burdensome for those doing the inspections. One compromise would be to report all damage sizes that are repaired. This should provide a much larger set of data to work with than only reporting events beyond the repair size limits.

Another issue related to the SDRS reporting requirements is whether or not all damage sizes larger than the repair size limits were actually reported. The size and effect that this missing data set has on the Bayesian updating results presented here is unknown. What is also unknown is the effect the records with unreported damage sizes have on the results. It is evident from the data in Tables 4 and 5 that the majority of the damage events reported to SDRS do not have damage sizes associated with them. The Bayesian updating analysis using these data assumes that the recorded damage sizes represent a random sample from all of the damage events reported. It may be such that only in the worst cases of damage were the damage sizes actually reported. This would skew the analysis results toward a larger damage size distribution than would otherwise be the case.

Many of the concerns about SDRS data arise primarily from trying to extract damage size information from a system that was not designed to record such data. If the updating methods were using data from an inspection and maintenance program that was specifically tailored to report damage characteristics, most of these problems would be eliminated, or at least significantly reduced. Modifications to the system's current data reporting format are suggested here that would enhance the ability to gather detailed information on the characteristics of each structural damage event. These suggestions are not unique to the SDR system and should be equally applicable to any other inspection and maintenance program that records damage information for use in structural reliability estimates:

1) Distinguish between disbond and delamination damage on composite sandwich structures.
2) Add a check box on the form for reported damage size beyond SRM limits.
3) Add database fields for recording number of damages, damage sizes, damage detection methods, and number of airframe hours or cycles for each report.
4) Add capability for characterizing damage events with multiple damage mechanisms present.
5) Include dimensional locations of damage site in report.
6) Specify repair method used to repair damage.

Conclusions

The results of this analysis demonstrate that Bayesian updating provides an efficient means to revise damage size probabilities of aircraft structural components when new damage data become available. Unfortunately, there is not enough relevant information contained in current FAA maintenance databases to characterize damage size distributions for individual composite components, at least with any reasonable degree of confidence. However, the Bayesian updating results using FAA data serve as a limited, component specific validation of the general composite damage size assumptions derived from earlier literature. In nearly every case, the results show that the baseline distributions are a conservative estimate of the range of damage sizes encountered on commercial composite structural applications. The only exceptions were the results for the 767 flap crack cases, where the damage size distributions did not noticeably change from earlier estimates.

In light of these results, changes to current inspection and maintenance reporting procedures are recommended that would allow the continuous collection of statistically useful structural damage data for application to reliability analyses. The increase in relevant data resulting from these changes would allow much more refined estimates of airframe component reliabilities compared to estimates derived from existing data only. Updating of damage size data would also allow fleet reliability estimates to be revised on an ongoing basis and enable the highlighting of adverse reliability trends before they lead to catastrophic failure.

Appendix: Bayesian Updating Equations

The basic form of truncated detected damage size PDF is

$$p_0(a|a>\xi) = \frac{p_0(a)}{1-\int_0^{\xi} p_0(x)\,\mathrm{d}x} \qquad \text{for} \qquad a>0 \tag{A1}$$

The truncated form of the gamma detected damage size model [Eq. (7)], where the damage size (a) is conditional on the truncation value and the model parameter values is

$$p_0(a|a>\xi,\tau,\theta) = \frac{\left\{1\big/\theta^{\tau}\Gamma(\tau)E[P_D(a)]\right\}a^{\tau-1}P_D(a)\exp(-a/\theta)}{1-\int_0^{\xi}\left\{1\big/\theta^{\tau}\Gamma(\tau)E[P_D(x)]\right\}x^{\tau-1}P_D(x)\exp(-x/\theta)\,\mathrm{d}x} \tag{A2}$$

For n new detected damage size data points, the likelihood of the points being randomly sampled from Eq. (A2) is

$$L(a_1,a_2,\ldots,a_n|a>\xi,\theta,\sigma) = \{\theta^{\tau}\Gamma(\tau)E[P_D(a)]\}^{-n} \times \prod_{i=1}^{n}\frac{a_i^{\tau-1}P_D(a_i)\exp(-a_i/\theta)}{1-\int_0^{\xi_i}\left\{x^{\tau-1}P_D(x)\exp(-x/\theta)\big/\theta^{\tau}\Gamma(\tau)E[P_D(x)]\right\}\mathrm{d}x} \tag{A3}$$

where the truncation point can be uniquely associated to each damage size data point. The likelihood can be rewritten in a simplified form by substituting damage size sample averages [Eqs. (A5–A7)]

$$L(a_1,a_2,\ldots,a_n|a>\xi,\tau,\theta) = \left\{\frac{\exp[-\bar{a}/\theta+(\tau-1)\overline{\ln a}+\bar{P}_D]}{\theta^{\tau}\Gamma(\tau)E[P_D(a)]}\right\}^{n} \times \prod_{i=1}^{n}\left\{1-\int_0^{\xi_i}\frac{x^{\tau-1}P_D(x)\exp(-x/\theta)}{\theta^{\tau}\Gamma(\tau)E[P_D(x)]}\,\mathrm{d}x\right\}^{-1} \tag{A4}$$

$$\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i \tag{A5}$$

$$\overline{\ln a} = \frac{1}{n}\sum_{i=1}^{n}\ln a_i \tag{A6}$$

$$\bar{P}_D = \frac{1}{n}\sum_{i=1}^{n}\ln[P_D(a_i)] \tag{A7}$$

When a joint earlier distribution is assumed for the model parameters τ and θ, Bayes theorem can provide an updated estimate of what the model parameter values should be in light of the new damage size data:

$$f_u(\tau,\theta|a_1,a_2,\ldots,a_n) \propto L(a_1,a_2,\ldots,a_n|a>\xi,\tau,\theta)f_0(\tau,\theta) \tag{A8}$$

For this application, the prior distributions of the model parameters τ and θ are assumed a priori to be independent. A two-parameter gamma PDF is used to model the τ parameter earlier distribution, and an inverse-gamma PDF is used to model the θ parameter earlier distribution. However, any continuous univariate PDF can be used to model the prior distributions:

$$f_0(\tau) = [1/\lambda^r \Gamma(r)]\tau^{r-1} \exp(-\tau/\lambda) \tag{A9}$$

$$f_0(\theta) = [1/\beta^\alpha \Gamma(\alpha)]\theta^{-(\alpha+1)} \exp(-1/\beta\theta) \tag{A10}$$

To obtain the joint posterior distribution of the model parameters, the Bayesian updating equation [Eq. (A8)] must be normalized and solved numerically because no closed-form solution exists. When attempting a numerical solution to this equation, considerable difficulty may be experienced in computing the normalizing constant using direct numerical integration. In such cases, Monte Carlo simulation is often used to construct probability surfaces from which the relevant statistics can be extracted. However, direct Monte Carlo simulation will not work here because of the need to generate random draws directly from the likelihood function [Eq. (A4)]. This problem can be circumvented by the use of importance sampling to generate the random draws.

Gelman et al. outline a method for estimating the marginal distribution of the joint posterior model parameters using importance sampling.[11] The expected value of the updated marginal distribution for a model parameter can be expressed as a function of the joint posterior PDF:

$$E[\tau|a_1, a_2, \ldots, a_n] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tau f_u(\tau, \theta|a_1, a_2, \ldots, a_n)\, d\tau\, d\theta \tag{A11}$$

In importance sampling, a PDF, $g(\tau, \theta)$, is introduced that can be directly sampled from and that approximates the solution surface for the joint posterior distribution of the model parameters:

$$E[\tau|a_1, a_2, \ldots, a_n] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\tau f_u(\tau, \theta|a_1, a_2, \ldots, a_n) g(\tau, \theta)}{g(\tau, \theta)}\, d\tau\, d\theta \tag{A12}$$

The importance weight factor is defined as the ratio of the density to be simulated to the approximating density:

$$w(\tau, \theta) = \frac{f_u(\tau, \theta|a_1, a_2, \ldots, a_n)}{g(\tau, \theta)} \tag{A13}$$

Taking m samples from the approximating density, the expected values of the model parameters can be estimated as a weighted average of the sample values, if the joint posterior distribution has been normalized:

$$E[\tau|a_1, a_2, \ldots, a_n] \approx \frac{1}{m}\sum_{i=1}^{m} \tau_i w(\tau_i, \theta_i) \tag{A14}$$

For the unnormalized form of Eq. (A8) the expected values can be found by dividing the weighted average of the sample values by the average of the weights:

$$E[\tau|a_1, a_2, \ldots, a_n] \approx \frac{(1/m)\sum_{i=1}^{m} \tau_i w(\tau_i, \theta_i)}{(1/m)\sum_{i=1}^{m} w(\tau_i, \theta_i)} \tag{A15}$$

$$E[\theta|a_1, a_2, \ldots, a_n] \approx \frac{(1/m)\sum_{i=1}^{m} \theta_i w(\tau_i, \theta_i)}{(1/m)\sum_{i=1}^{m} w(\tau_i, \theta_i)} \tag{A16}$$

This approach eliminates the need to calculate normalizing constants for the Bayesian updating equation [Eq. (A8)].

The efficiency and accuracy of importance sampling depends on how well the approximating density matches the joint posterior distribution of the model parameters. Kong et al. define an ESS as a measure of importance sampling efficiency. ESS is interpreted as a ratio of importance sample size to joint posterior sample size, such that m draws from the importance sampling distribution offers the same estimation accuracy as ESS $\times\, m$ draws from the joint posterior distribution.[12] For m importance sampled draws, the ESS is expressed in terms of normalized values of the weight factors:

$$\mathrm{ESS} = \frac{\left[\sum_{i=1}^{m} w_n(\tau_i, \theta_i)\right]^2}{m\sum_{i=1}^{m} w_n^2(\tau_i, \theta_i)} \tag{A17}$$

Acknowledgments

This work was supported by the NASA Langley Research Center under Grant NAG-1-2055. The authors wish to thank W. Jefferson Stroud of NASA for his invaluable support in this research effort.

References

[1] Lin, K. Y., Rusk, D. T., and Du, J., "Equivalent Level of Safety Approach to Damage Tolerant Aircraft Structural Design," AIAA Paper 2000-1371, April 2000.

[2] Harris, D. O., "Probabilistic Crack Growth and Modeling," *Reliability-Based Mechanical Design*, edited by T. A. Cruse, Marcel Dekker, New York, 1997, pp. 300–318.

[3] Berens, A. P., and Hovey, P. W., "Statistical Methods for Estimating Crack Detection Probabilities," *Probabilistic Fracture Mechanics and Fatigue Methods: Applications for Structural Design and Maintenance*, STP 798, American Society for Testing and Materials, Philadelphia, 1983, pp. 79–94.

[4] Berens, A. P., "NDE Reliability Data Analysis," *Nondestructive Evaluation and Quality Control*, edited by S. R. Lampman and T. B. Zorc, 9th ed., Vol. 17, Metals Handbook, ASM International, Metals Park, OH, 1989, pp. 689–701.

[5] Gray, P. M., and Riskalla, M. G., "Development of Probabilistic Design Methodology for Composite Structures," U.S. Dept. of Transportation/Federal Aviation Administration, Rept. DOT/FAA/AR-95/17, Washington, DC, Aug. 1997.

[6] Hoffman, J. D., *Numerical Methods for Engineers and Scientists*, McGraw–Hill, New York, 1992, pp. 101–104.

[7] Boisvert, R. F., Howe, S. E., Kahaner, D. K., and Springman, J. L., "The Guide to Available Mathematical Software," National Inst. of Standards and Technology, Rept. NISTR 90-4237, Gaithersburg, MD, April 1990.

[8] Brewer, J. C., "Estimate of Probability of Crack Detection from Service Difficulty Report Data," U.S. Dept. of Transportation/Federal Aviation Administration, Rept. DOT/FAA/CT-94/90, Washington, DC, Sept. 1994.

[9] Ripley, B. D., *Stochastic Simulation*, Wiley, New York, 1987, pp. 230, 231.

[10] Rusk, D. T., "Reliability-Based Methods Applied to the Design of Damage Tolerant Aircraft Structures," M.S. Thesis, Dept. of Aeronautics and Astronautics, Univ. of Washington, Seattle, WA, June 2000.

[11] Gelman, A. B., Carlin, J. S., Stern, H. S., and Rubin, D.B., *Bayesian Data Analysis*, Chapman & Hall/CRC, New York, 1995, pp. 307–311.

[12] Kong, A., Liu, J. S., and Wong, W. H., "Sequential Imputations and Missing Data Problems," *Journal of the American Statistical Association*, Vol. 89, No. 425, 1994, pp. 278–288.