*Quality Engineering*, 27:304–316, 2015

ISSN: 0898-2112 print / 1532-4222 online
DOI: 10.1080/08982112.2014.990033

Taylor & Francis
Taylor & Francis Group



# Improving Reliability Understanding Through Estimation and Prediction with Usage Information

**Lu Lu,[1] Christine M. Anderson-Cook[2]**
[1]Department of Mathematics and Statistics, University of South Florida, Tampa, Florida
[2]Statistical Sciences Group, Los Alamos National Laboratory, Los Alamos, New Mexico

**ABSTRACT** Using information about the usage or exposure of a complex system in addition to its age can provide additional understanding about mechanisms driving change in reliability as well as potentially improve the prediction. Both the individual reliability of particular units as well as population reliability can be improved with the inclusion of additional explanatory factors. In this article we consider an example based on a complex munition system. Using age alone to predict reliability can provide some information, but differences between units of the same age cannot be discerned. Subpopulations of the stockpile can be identified to help improve estimation, but the largest gains in understanding of the mechanisms driving change in reliability and prediction of future performance come from incorporating usage information.

**KEYWORDS** Bayesian analysis, individual reliability, population reliability, auxiliary information

## INTRODUCTION

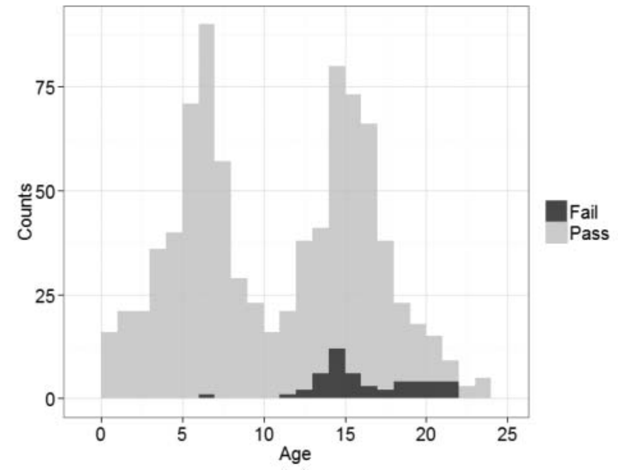
Understanding the reliability of units can lead to improved prediction of the health of the population or stockpile, as well as opportunities for their better management. If managers can anticipate when units will begin to fail, units can be removed from use before they are put into critical situations. In addition, having a dependable estimate of when failures are expected can allow healthy units to remain in the population for longer. Including information about the usage or exposure of a complex system in addition to its age can provide improved understanding about mechanisms driving change in reliability. Of interest is considering both the accuracy and precision of the estimated reliability. Missing a feature in the data can lead to erroneous decisions, and inflated uncertainty in the estimates can weaken the ability of the decision maker to draw conclusions. Similar to the reliability of cars being improved by using both the age and mileage of the vehicle compared to basing estimates on age alone, unit-by-unit management of the population becomes possible when all units of the same age are not necessarily treated as being the same. Of interest

Address correspondence to Christine M. Anderson-Cook, Statistical Sciences Group, Los Alamos National Laboratory, P.O. Box 1663, MS F600, Los Alamos, NM 87545. E-mail: c-and-cook@lanl.gov

may be a statistical parametric model to describe the reliability of individual units as a function of their age and usage history, as well as population reliability, which allows prediction of the probability of a randomly selected unit working at a given point in the future (Lu and Anderson-Cook 2011a, 2011b).

Consider an example based on a complex munition system, where fictitious data are shown in lieu of the proprietary data but with the key characteristics of the data structure preserved. The units are stored in warehouses or close to combat locations and are transferred between locations as required. It is thought that the handling of the units is stressful on several parts and may accelerate degradation. Because this is a single-use destructively tested system, it is not possible to monitor the health of the unit without a full system test. Data consist of a pass/fail measure at a given point in time, where the age and other summary information about the system are also recorded. Figure 1a shows the histogram of the age distribution for the test data, with the darker gray representing failures and the lighter gray for passes. We can see a bimodal age distribution for the test data, which suggests that the tested units are a mixture of some younger units (with age between 0 and 10 years and centered around 6 years) and some older units (with age between 10 and 25 years and centered around 15 years). The peaks in the age distribution with more testing coincide with the scheduled maintenance plan for the units. The goal is to do testing near 6 and 15 years but, due to logistical constraints, there is considerable variability in when the testing is actually performed. There appears to be a larger proportion of failures among older units than younger units indicating the systems may be aging over time.

At the point of purchase, the units are assigned to one of the two branches of the military, the Army or Navy, respectively. Due to the differences in the storage and handling history, it is possible for the units from the two subpopulations to have very different degradation mechanisms and reliability performance. Hence, it is natural to consider dividing the units into two distinct subpopulations. Figure 1b shows the age distribution of the tested units from the two subpopulations. We can see many more failures among the 190 Army units than among the 240 Navy units. In addition, failures happen primarily for units of age greater than 10 years.



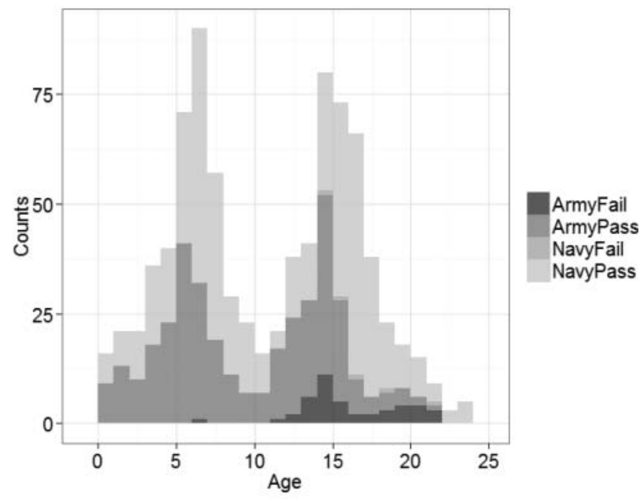




**FIGURE 1** **Histogram of the age (measured in years) for the test data: (a) the whole population and (b) distinguishing the Army and Navy subpopulations.**

In addition to the age and subpopulation membership information, we know the number of transfers the units have had during their lifetime, which can be considered as the usage history of the single-use missile systems. The usage rate, defined as the ratio of the usage and the age of a system, can be used for capturing the average usage level over a certain period of interest (lifetime in our case). Figure 2a shows the usage rate versus age for the Army (black) and Navy (gray) units. We observe more units with higher usage rate and larger variation for the Army units than for the Navy ones. The scatter plot of usage versus age is shown in Figure 2b, which indicates a positive association of the age and usage of the systems with generally heavier handling for the Army units.

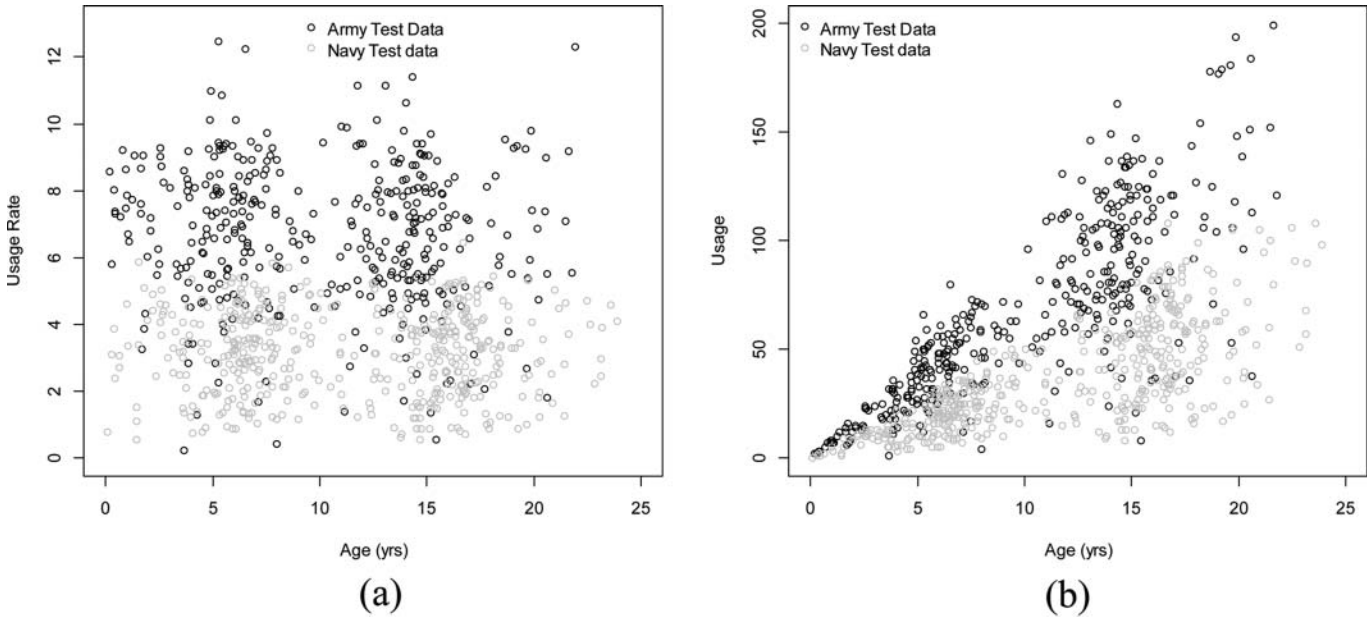



**FIGURE 2** **Scatter plots of (a) usage rate vs. age and (b) usage vs. age for the test data. The black and gray circles represent the Army and Navy subpopulations, respectively.**

A Probit regression model (Lu and Anderson-Cook 2011a, 2011b) of the form

$$\begin{aligned} Y_i &\sim Bernoulli(p_i) \\ p_i &= \Phi(\boldsymbol{\beta}' \boldsymbol{x_i}) \end{aligned} \quad [1]$$

is used to predict reliability as a function of the available information. In the above model, $p_i = P(Y_i = 1)$ is the probability of the $i$th system working at a given time point, $\Phi$ is the cumulative distribution function for a standard normal distribution, $\boldsymbol{x_i}$ denotes the vector of explanatory variables (which, depending on the selected model, may include age and/or usage) for the $i$th system at the given time, and $\boldsymbol{\beta}$ is the vector of the coefficient parameters including the intercept.

Given the available information for the test data, we explore four different models for assessing the system reliability for the individual Army and Navy units as well as summarizing the overall performance of all units from the two subpopulations and the combined whole population. The four models are as follows:

1. A single population with reliability estimated using only the age of the systems.
2. Separate models for reliability for each of the Army and Navy subpopulations using age only.
3. A single population with reliability estimated as a function of age and usage (number of transfers).
4. Separate Army/Navy models using age and usage.

Comparison between the models considers the possible benefits of using additional information for modeling reliability. The following section briefly describes the individual and population reliabilities as well as their relevance to quantities of interest for stockpile or population managers as well as differences between the two summaries. The next section compares the four different models and discusses how the choice of explanatory factors affects the estimated individual and population reliability summaries. Concluding remarks are provided in the last section.

## INDIVIDUAL RELIABILITY (INDREL) AND POPULATION RELIABILITY (POPREL)

Reliability analyses are typically conducted to estimate the probabilities of individual systems performing their required functions for a stated period of time under some given environmental and operating conditions. This reliability summary for a single system is named "individual reliability" or IndRel in Lu and Anderson-Cook (2011a). For single-use nonrepairable systems that age over time, reliability is often estimated as a changing function of the age of the system. If additional information on usage or exposure history that describes the life experience of the system is useful for predicting reliability, then it should be included in the model for improved precision of estimation. By using

Bayesian or frequentist methods for estimating the model parameters based on a representative sample of data obtained from a population of homogeneous systems, the reliability for an individual system (IndRel) with characteristics consistent with the tested population at a given time point can be estimated if we know its age and usage values. This summary is useful for understanding and managing individual units; for example, to remove units with lower reliability from the stockpile or to send them for maintenance before they fail in the field.

However, when managing a population of systems, interest may lie in understanding the reliability of the overall population at the current or some future time, rather than whether particular units will function or not. An overall summary of the health of the population can be measured by describing the reliability for the collection of units. This is helpful for scheduling and planning resources for maintenance services as well as ordering replacements. Lu and Anderson-Cook (2011a) proposed a population aggregate reliability (PopRel) summary that estimates the probability of a randomly selected unit from the population functioning properly at the current or some future time. To calculate PopRel, both test data and the age distribution of the population at the current time are needed. The test data are required to estimate the parameters of the reliability model and the age distribution of the current population is useful for predicting future reliability for individual units based on their current ages. An estimate of PopRel is summarized as a function of time from present and can be obtained by averaging the predicted individual reliabilities at the given time point over the entire population. Another summary that estimates the proportion of the population having reliability above a certain threshold level was also proposed to provide an alternative quantitative measure of the overall population performance. More details about the calculation of the PopRel and its alternative summary are available in Lu and Anderson-Cook (2011a), where the implementations under both the Bayesian and frequentist paradigms were discussed. Lu and Anderson-Cook (2011b) adapted the PopRel summary for cases where reliability changes as a function of both age and usage. In addition to the test data and the age distribution of the current population, we need the usage information for the current population, which will be used for predicting future usage at some future time of interest. Because usage does not advance in a deterministic pattern like age, Lu and Anderson-Cook (2011b) proposed a resampling-based method assuming that a similar usage pattern as observed from the current population will be preserved in the future, which allows for the estimation of PopRel taking into account the additional uncertainty of the unknown future usage. The method can also be adapted if there is an anticipated change in the future usage pattern (Lu and Anderson-Cook 2011b). Examples of using PopRel summary in stockpile management are explored in Lu, Anderson-Cook, and Wilson (2011).

## COMPARISON OF THE FOUR MODELS BASED ON THE INDIVIDUAL AND POPULATION RELIABILITIES

In this section, we consider four models for estimating and predicting future reliability. We use Bayesian analyses (see Lu and Anderson-Cook [2011a] for details) to estimate the posterior distribution of the model parameters. In a Bayesian analysis, prior distributions are specified to describe knowledge known a priori about the model parameters. The priors are combined with the observed data summarized through the likelihood function to obtain an improved understanding of the model, which is summarized in the posterior distribution of the model parameters. By manipulating the posterior distributions, we are able to obtain reliability estimates and predictions for different combinations of inputs. Informative priors can be used when there is prior knowledge or external information available prior to collecting the data; otherwise, noninformative or diffuse priors should be used, which usually lead to results that are consistent with using the frequentist approaches.

For the model using only age to predict reliability for the entire population (Model I), a diffuse prior centered at zero with a uniform distribution, $U(-1000, 1000)$, is used for both of the coefficient parameters $\beta_0$ and $\beta_1$. The Markov chain Monte Carlo approach was implemented through JAGS (Plummer 2003) and R (through the rjags package) to obtain an approximation of the posterior distributions of the model parameters. The estimated regression model based on using the posterior means of the model parameters is given by $p_i = \Phi(3.15 - 0.17A_i)$, where $A_i$ is the age of the $i$th system at the given time. The 95 percent credible intervals for $\beta_0$ and $\beta_1$ given by (2.71,

**TABLE 1** Summary of the Posterior Distributions of Model Coefficient Parameters for the Four Models

| Mean (5th, 50th, 95th percentiles) | Age only | | Age and usage | | |
|---|---|---|---|---|---|
| | $\beta_0$ | $\beta_1$ | $\beta_0$ | $\beta_1$ | $\beta_2$ |
| | Model I (DIC = 322.4) | | Model III (DIC = 200.4) | | |
| Combined | 3.15 | −0.17 | 4.08 | −0.01 | −0.03 |
| | [2.71, 3.14, 3.64] | [−0.15, −0.12, −0.09] | [3.39, 4.05, 4.89] | [−0.06, −0.01, 0.04] | [−0.04, −0.03, −0.02] |
| | Model II (DIC = 254.8) | | Model IV (DIC = 202.0) | | |
| Army | 3.72 | −0.19 | 4.44 | −0.01 | −0.03 |
| | [3.03, 3.71, 4.49] | [−0.24, −0.19, −0.15] | [3.57, 4.39, 5.43] | [−0.09, −0.01, 0.07] | [−0.04, −0.03, −0.02] |
| Navy | 4.05 | −0.11 | 4.05 | −0.07 | −0.01 |
| | [2.83, 3.93, 5.72] | [−0.21, −0.11, −0.04] | [2.75, 3.94, 5.70] | [−0.18, −0.07, 0.04] | [−0.03, −0.01, 0.00] |



3.64) and (−0.15, −0.09), respectively. Note that because the entire credible interval for $\beta_1$ is negative, this suggests that the units are degrading over time with diminishing reliability. More detailed summary for the estimated model parameters is available in Table 1.

Based on the differences observed in the usage patterns between the Army and Navy subpopulations, there is a reason to anticipate that the reliabilities of the two subpopulations may differ. A simple approach for taking this into account suggests separate reliability models for each subpopulation. Hence, the second model (Model II) considers using this additional information to divide the overall stockpile into two subpopulations based on whether the Army or Navy manages the particular units. This information is included into a model of the form

$$p_i = \Phi\left(\beta_0^{(j)} + \beta_1^{(j)} A_i\right), j \in \{\text{Army}, \text{Navy}\},$$

where the superscript denotes the membership of the Army and Navy subpopulations. Then we are able to discern differences in the predicted reliability of the two subpopulations. By distinguishing the units from the two subpopulations, we can obtain the estimated regression models based on using the posterior means of the model parameters as $p_i = \Phi(3.72 - 0.19A_i)$ and $p_i = \Phi(4.05 - 0.11A_i)$ for the Army and Navy subpopulations, respectively. The estimate based on using the posterior median is very close to using the posterior mean, which is also provided in Table 1. The 95 percent credible intervals for $\beta_0$ and $\beta_1$ of the Probit model are (3.03, 4.49) and (−0.24, −0.15) for the Army units and (2.85, 5.72) and (−0.21, −0.04) for the Navy units. The Navy units are estimated to have larger intercept and smaller magnitude of slope for the linear link function and hence have higher estimated reliability than the Army units with the same age. In addition, due to the small number of failures observed for the Navy units, the estimated model parameters are associated with larger variances than the Army units.

To compare the estimation obtained using Models I and II, Figure 3a shows the posterior distributions of the coefficient of the age variable, $\beta_1$, using the different models. The $\beta_1$ estimate has a smaller center value as well as a smaller dispersion for the Army units than the Navy units for Model II. The $\beta_1$ estimated using the combined data from Model I has the posterior mean in between of the values obtained for the two subpopulations from Model II. Because it is based on a larger sample size, the combined $\beta_1$ estimate has smaller variance than the estimate obtained from either one of the two subpopulations.

Figure 4 shows the estimated IndRels as a function of the age of a system using the first two models. We can see that the Navy units have generally higher estimated reliability than the Army units of the same age. The younger Army units (with age under 23 years) have wider credible intervals for the estimated reliability than the Navy units, whereas the older units (of age above 23 years) have smaller variance for the estimated reliability. The estimated reliability and its associated uncertainty using the model with the whole population data (Model I) are in between of the estimates for the two subpopulations (Model II). The distinction

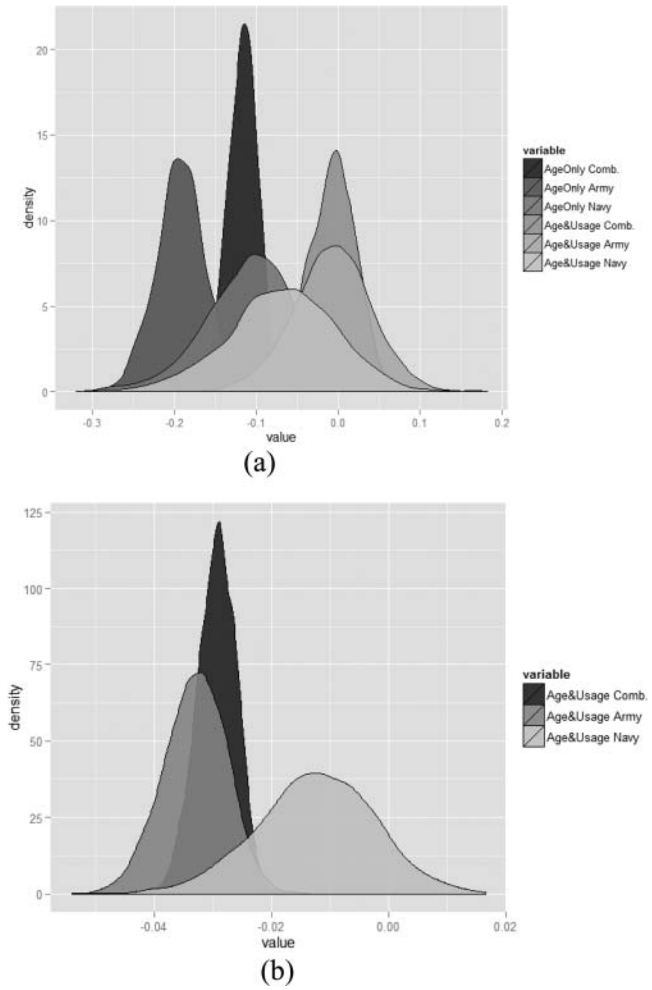



**FIGURE 3** Posterior distributions of model coefficients for using the four different models: age only or age and usage crossed with combined data or separate subpopulation data: (a) posterior distributions of $\beta_1$ and (b) posterior distributions of $\beta_2$.

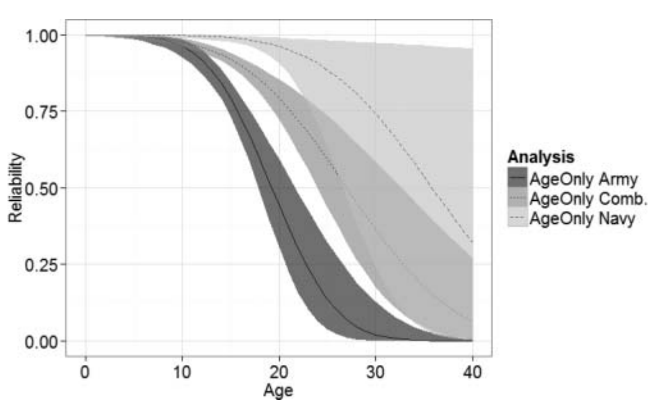



**FIGURE 4** Estimated IndRel with its associated uncertainty for Models I and II. The credible regions for different models and different subpopulations are shown in different shades of gray.

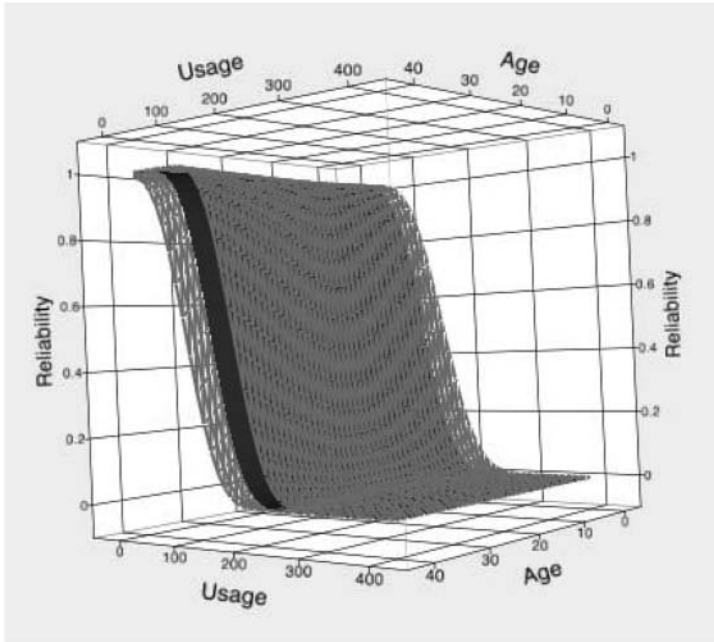
between the two subpopulations suggests some benefits for managers to be able to discern the different predicted reliabilities. Understanding the differences in performance for the two subpopulations would allow managers to adapt their decision making to take this information into account.

Another approach to take into account the significant observed difference in usage history for the two subpopulations is to incorporate the usage directly into the model. If the usage is a main factor driving change in reliability between the two subpopulations, then including the usage for modeling the reliability will improve estimation and prediction. In addition, this approach potentially allows for more precise information to be used to distinguish reliabilities within a subpopulation, because units with different numbers of transfers within a subpopulation could have distinct predicted reliabilities. When the usage information is included, the third model for reliability is given in the form of

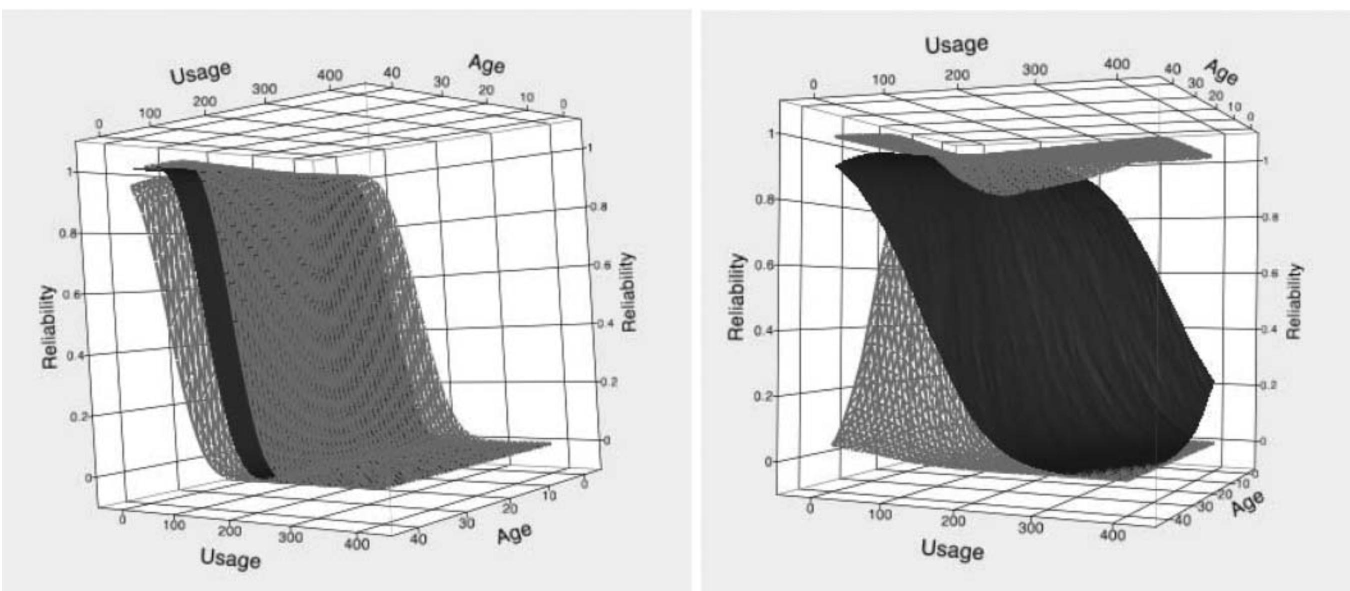
$$p_i = \Phi(\beta_0 + \beta_1 A_i + \beta_2 U_i), \qquad [2]$$

where $U_i$ denotes the usage (the number of transfers) of the $i$th system at the given time, and $\beta_2$ is the model coefficient parameter for the usage variable. The approximated posterior distributions of $\beta_1$ and $\beta_2$ are shown in Figure 3, with more detailed numerical summaries available in Table 1. The 95 percent credible intervals for $\beta_1$ and $\beta_2$ are (−0.06, 0.04) and (−0.04, −0.02), respectively. Note that when the usage is included in the model, the posterior distribution for $\beta_1$ includes the value zero, suggesting that there is an even stronger usage effect on estimating reliability compared to considering age effect only and that it might be reasonable to remove the age term. The estimated reliability as a function of both age and usage is displayed in Figure 5a, where the reliability is seen to drop quickly as the usage increases but is less affected by increasing age.

A final option is to allow for completely separate relationships between reliability and the input factors, age and usage, for each of the subpopulations. Here a model of the form in Eq. [2] would be fit separately to the Army and Navy subpopulations. The approximated posterior distributions of coefficient parameters are shown in Figure 3, with the numerical summaries available in Table 1. The posterior distribution of $\beta_1$ for the Army units is centered around a similar value as




(a)




(b)

**FIGURE 5** **Estimated IndRel with its associated uncertainty for Models (a) III and (b) IV. The dark gray surface represents the estimated response surface based on the posterior median. The lighter gray surfaces represent the 5th and 95th percentiles in the posterior distribution.**

for the whole population using Model III but has larger dispersion due to the smaller amount of available data. The posterior distribution of $\beta_1$ for the Navy units has even more dispersion with smaller posterior mean value. However, the two subpopulation posterior distributions still have substantial amount of overlap. The posterior distributions of $\beta_2$ are more different between the two subpopulations, with larger posterior mean and variance for the Navy units due to the smaller number of failures observed. Figure 5b shows the estimated IndRels for the Army and Navy units, where we see not so different estimated response surfaces for the two subpopulations based on the posterior median. However, the Navy units have much larger uncertainty associated with the estimated reliability, due to the smaller number of observed failures.

Noted that in Table 1 there is a comparison between the four models based on the deviance information criterion (DIC). The DIC (Gelman et al. 2004) uses the estimated expected deviance, which is a measure of the average predicting error with additional penalty of model complexity, to evaluate the overall model fit. A smaller DIC value suggests better predictive powers and hence better model fit. Model I has the largest DIC value (322.4) and hence the worst model fit among the four models. Model II has a smaller DIC (254.8) than Model I, which indicates improvement in the predictive power by taking into account the two subpopulations. Models III and IV have similar DIC values (200.4 and 202.0), which are considerably smaller than those for Models I and II. Hence, including the usage information in modeling the reliability is more effective for improving the precision of predicting reliability than simply discerning the two subpopulations into different homogeneous groups. However, after including the usage information, which naturally adjusts the reliability for the units from different subpopulations with different usage patterns, using separate models for the different subpopulations yields no extra benefit (with a slightly worse DIC value). This also suggests that usage is the internal main factor driving change in reliability and a single overall degradation mechanism applies to both subpopulations.

Therefore, the simpler model (Model III), which does not distinguishing between the two groups, has several substantial advantages over Model IV:

1. A single model would suggest that the same degradation mechanism applies to both the Army and Navy subpopulations. The simplicity of having a single relationship for reliability across both subpopulations may more closely match with the subject-matter expert's understanding of the drivers of change in reliability. By knowing the age and number of transfers for a given unit, it would be possible to predict its reliability without the knowledge of which subpopulation it came from.
2. By using the entire population to estimate the model parameters, we are able to obtain more precise estimates of reliability for all combinations of the explanatory factors.

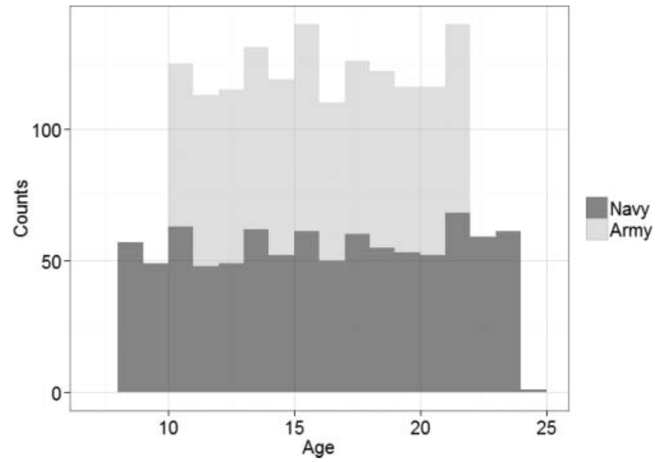



**FIGURE 6** Age distribution for the current population (1,700 units) by distinguishing the Army (800) and Navy (900) units.

When we compare Model III to Model II, there are also some advantages:

1. If the usage pattern changed within the Army or Navy subpopulations, we should have greatly reduced confidence in the ability of Model II to predict future reliability. Model III allows the change in pattern of usage to be directly incorporated into the model.
2. Model III allows for specific information for each unit to be entered into the model through the inputs, allowing for individualized estimates of reliability specific to each unit and its history. With Model II, differences in usage within the subpopulation are not leveraged and we obtain only a summary based on the overall characteristics of the subpopulation.

Based on the estimated IndRels with different models, next we examine the estimated PopRel for the combined population as well as the Army and Navy subpopulations. Figure 6 shows the age distribution for the 800 Army and 900 Navy units in the current population. Figure 7 shows the usage rate (as calculated by dividing the current usage by the current age) and usage versus age for the current population. Similar to the patterns observed in Figure 2, the Army usage rate is consistently higher than the Navy subpopulation. Figures 8 and 9 display the estimated PopRel (as estimated by the posterior median) and its associated uncertainty (credible interval based on the fifth and ninety-fifth percentiles of the posterior distribution) for the next 12 years for the Army and Navy

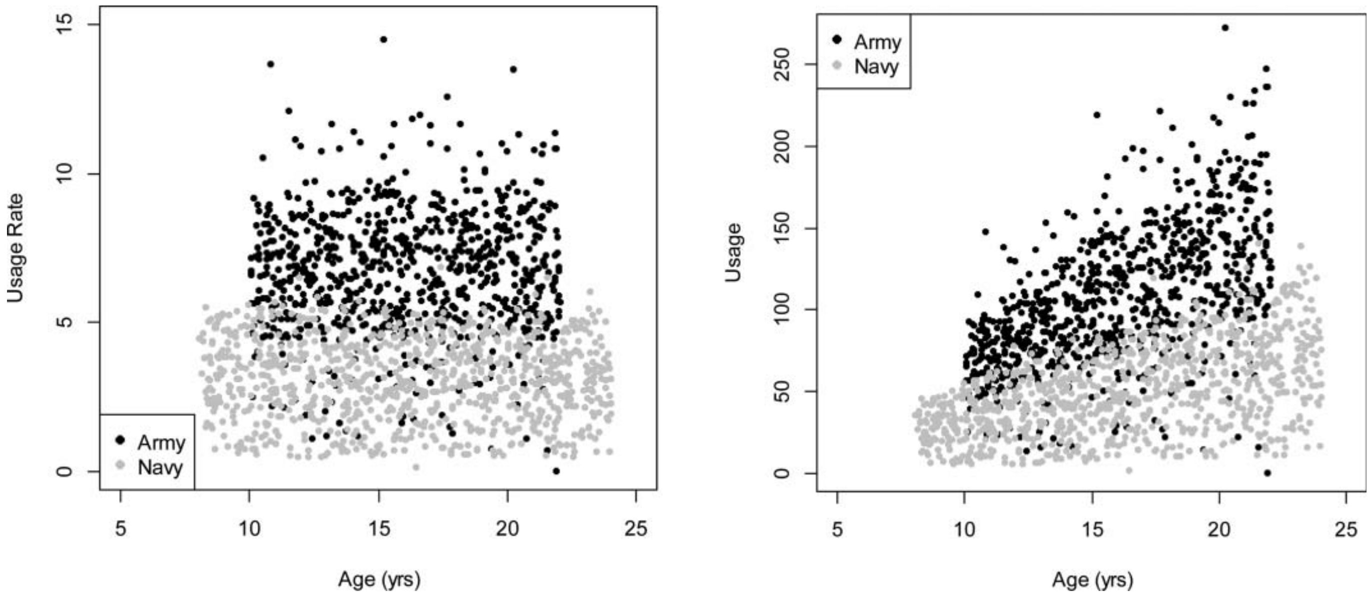



**FIGURE 7** Scatter plots of (a) age vs. usage rate and (b) age vs. usage for the current population of 1,700 units. The black and gray circles represent the Army and Navy subpopulations, respectively.

subpopulations, respectively. Note that the posterior distributions of the PopRel are nearly symmetric for all four models, leading to the median being close to the center in the credible intervals. In addition, the estimates obtained using the posterior mean (not shown) and median are very similar.

For the Army subpopulation with 800 units, Model I consistently estimates the reliability to be higher than Model II because the IndRel estimated using the combined data does not recognize the more severe usage of the Army subpopulation and hence is consistently higher than using the Army subpopulation data as shown in Figure 4. In addition, the PopRel estimated with Model I has larger variance after around 5 years from the current time because the IndRel estimated using the combined data is associated with larger variance than using the Army subpopulation data for old units (older than 25 years). Comparing the estimates

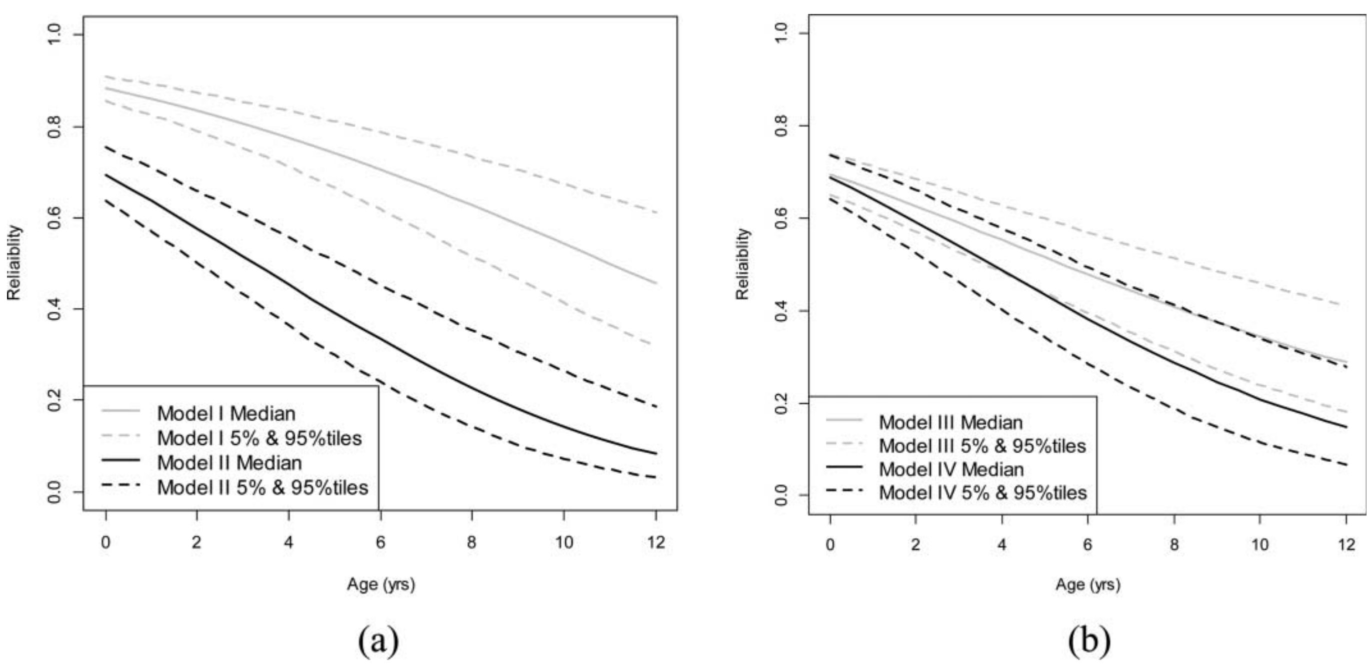



**FIGURE 8** PopRel summary for the Army subpopulation of 800 units for all four models: (a) Models I and II and (b) Models III and IV.

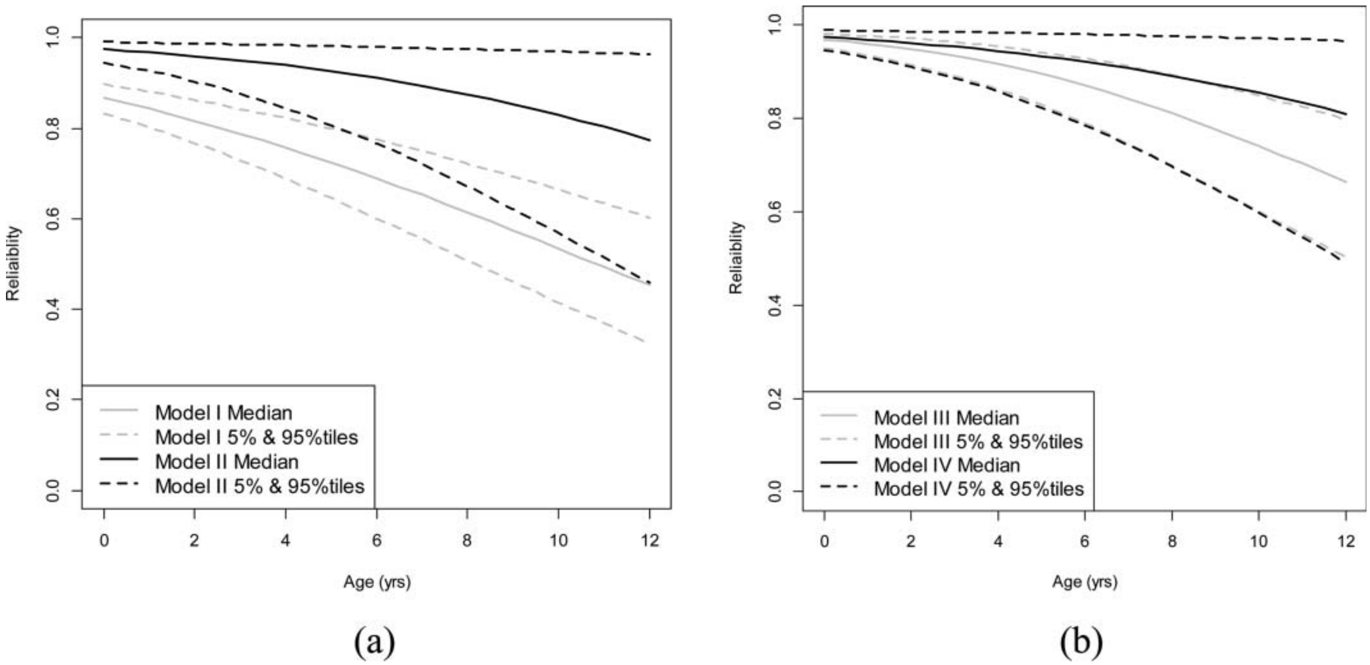



**FIGURE 9** PopRel summary for the Navy subpopuation of 900 units for all four models: (a) Models I and II and (b) Models III and IV.

obtained using the models incorporating usage (III and IV), Model III has a consistently larger estimated reliability, whereas Model IV tends to produce a lower estimate of reliability with larger differences as the prediction is made further into the future. When we compare the predictions from all four models, we see that Models III and IV have estimates that are between the extremes of Models I and II, and the widths of the uncertainty intervals are generally similar.

For the Navy subpopulation with 900 units, Model I has consistently lower estimates of the reliability compared to Model II. Model III produces a smaller estimate than Model IV with greater differences as the prediction is made further into the future. However, it offers substantial reduction in the estimated variance by using the combined data with larger size. Model II gives estimate of reliability similar to that of Model IV with generally larger uncertainty. This indicates that the additional uncertainty introduced by predicting usage is offset by the substantial reduction in the residual variance by including the usage variable. Model III provides predicted reliability for the subpopulation that is more moderate with generally less variability than either Model II or IV, which has some provision for taking into account of subpopulation differences.

When we compare the results between the Army and Navy subpopulations, the overestimated reliability for the Army and the underestimated reliability for the Navy stem from not distinguishing the differences in the usage with Model I are potentially costly. Units from the Navy could be pulled from service prematurely, whereas less reliable units in the Army subpopulation could remain in the field too long.

Figure 10 shows the PopRel estimate for the entire population with 1,700 units (800 Army and 900 Navy units) that are mainly middle-aged or old units. By averaging over the subpopulation summaries proportional to their population sizes, Model I has a generally larger estimate of the overall population reliability than Model II with larger differences as prediction is made further into the future. Models III and IV have similar predicted PopRel, with the former associated with slightly smaller predicted reliability and slightly smaller variance. In addition, Models I and II are generally associated with larger variance than using the models incorporating usage (III and IV). From the PopRel summaries, we can see that including the usage information is helpful for producing improved prediction of reliability with less uncertainty. When the usage is one of the main factors driving change in reliability and the reliability degradation mechanisms between the two subpopulations are believed to be the same, using separate models for the different subpopulations in addition to the inclusion of the usage variable can introduce more variation into the prediction.

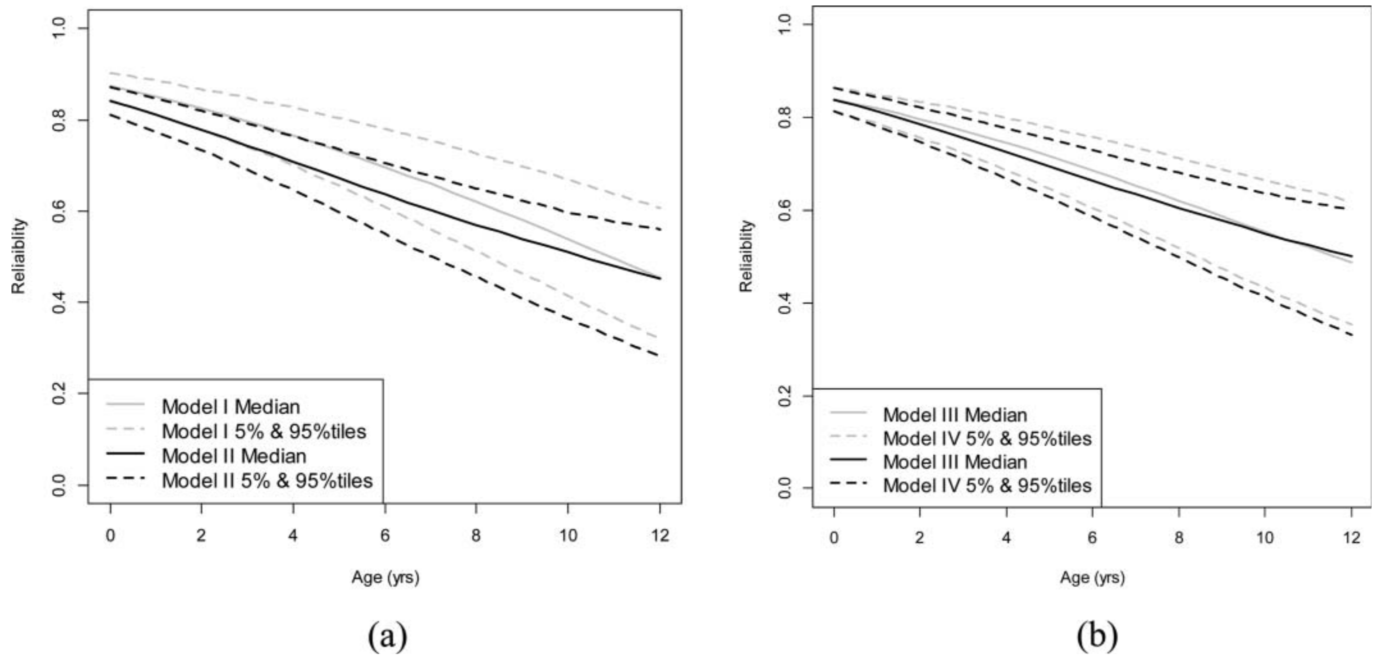



**FIGURE 10** **PopRel summary for the entire current population of 1,700 units for all four models: (a) Models I and II and (b) Models III and IV.**

## DISCUSSION AND CONCLUSIONS

A resampling-based method (Lu and Anderson-Cook 2011b) was used for predicting future usage when calculating the PopRel using models with the usage included (Models III and IV). This method does not assume a parametric model for the usage or usage rate distribution and instead draws repeated samples from the collection of usage rates from the observed data. Given an assumption of similar historical and future usage, future usage is determined based on the sampled usage rate given the current usage value. This method is easy to implement and flexible to use because it does not rely on the validity of an assumed model. However, like all nonparametric methods, it relies on large data to get reliable estimation. An alternative is to use a parametric approach. More specifically for our case study, we can fit truncated normal distributions to the usage rate for the two subpopulations. Truncated normal distributions make sense because a negative number of transfers (the usage measure) does not make sense. The fitted distributions can then be used for predicting future usage in the calculation of the PopRel.

Figures 11a and b show the histogram of the observed usage rates overlaid with the estimated truncated normal distribution for the two subpopulations. For the example, the Army units have larger average usage rates, which is more concentrated around seven transfers per year, whereas the Navy units distribution is centered between three and four but with a spread between zero and seven. Figure 11c illustrates how these different usage distributions translate into the variability in the predicted future usage. The dotted lines and the dashed lines define the range of the predicted usage rates determined by the estimated distributions for the Army and Navy subpopulations. The gray lines that radiate from the point denoting the current age and usage represent random samples drawn from the estimated usage rate distribution. Because the truncated normal distributions are fitted to both subpopulations, the gray lines that radiate from the current age and usage combination are more dense around the center of the estimated distribution but more widely spread near the tail ends of the distribution. This method relies on specifying an appropriate parametric distribution (or model) for the usage rate (or usage). If the model or distribution is properly specified, then it could lead to improved precision compared to using the resampling-based approach. An additional advantage of the parametric approach might be that it allows flexibility for a different future usage distribution to be specified if management of the units and the anticipated number of transfers were expected to change.

As a summary, this article illustrates a case study where information on subpopulations with different user

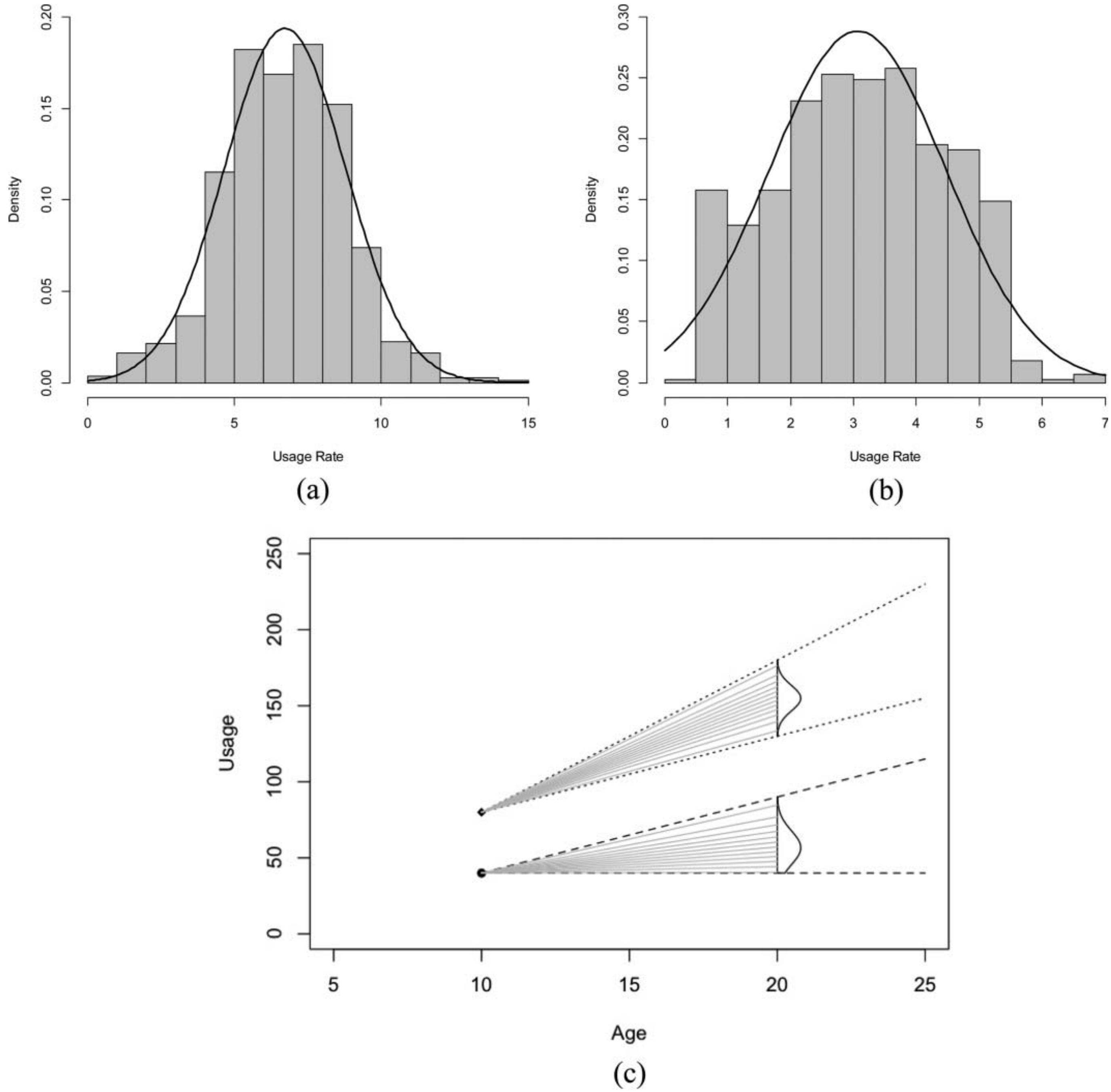



**FIGURE 11** **Histograms of the usage rate for the current (a) Army and (b) Navy subpopulations overlaid with the estimated density curves based on fitting truncated normal distributions and (c) an illustration for predicting future usage based on sampling from the estimated usage rate distributions.**

groups as well as the usage history was used to improve prediction of reliability for aging systems. Interest is in both good accuracy and precision for the estimates, because having bias in the reliability estimates can lead to errors in decisions about how to use the systems, and lack of precision comes with greater uncertainty about the patterns in the predicted reliability. Four different models considering whether or not discerning the Army and Navy subpopulations with separate models crossed with using age only or age and usage combined were used to estimate and predict individual and population reliabilities. It was shown that distinguishing between the two subpopulations (Model II) allows better estimation of reliability for different subpopulations than treating a heterogeneous population as a single group (Model I). However, directly using the usage in modeling the reliability is more effective for improving the reliability prediction because the differences between the individual units between and within the two subpopulations is better captured by the usage information than the more simplistic separation into two groups. This is true for this example because (a) the usage is a main factor driving change in reliability in addition to the age and (b) the same degradation mechanism driving changes in reliability from age and usage is suitable for both subpopulations. It was beneficial to discuss this hypothesis with the subject-matter experts and to verify this empirically by comparing the fitness of the four models based on the DIC or other model-checking criteria. Further, discerning the subpopulations in addition to including the usage in the reliability model may not be able to bring extra benefit when the subpopulations share the same reliability degradation mechanisms.

## ABOUT THE AUTHORS

Lu Lu is a visiting assistant professor in the Department of Mathematics and Statistics at the University of South Florida. She earned a doctorate in statistics from Iowa State University.

Christine M. Anderson-Cook is a research scientist in the Statistical Sciences Group at Los Alamos National Laboratory in Los Alamos, New Mexico. She earned a doctorate in statistics from the University of Waterloo in Ontario. She is a Fellow of both ASQ and the American Statistical Association.

## REFERENCES

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. *Bayesian data analysis*. 2nd ed. London, UK: Chapman & Hall/CRC.

Lu, L., and C. M. Anderson-Cook. 2011a. Prediction of reliability of an arbitrary system from a finite population. *Quality Engineering* 23: 71–83.

Lu, L., and C. M. Anderson-Cook. 2011b. Using age and usage for prediction of reliability of an arbitrary system from a finite population. *Quality and Reliability Engineering International* 27: 179–90. http://dx.doi.org/10.1002/qre.1109

Lu, L., C. Anderson-Cook, M., and A. G. Wilson. 2011. Choosing a consumption strategy for a population of systems based on reliability. *Proceedings of the Institution of Mechanical Engineers - Part O: Journal of Risk and Reliability* 225: 407–23. http://dx.doi.org/10.1177/1748006X11392287

Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Paper presented at the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), Vienna, Austria, March 20–22.