Available online at www.sciencedirect.com
ScienceDirect
Reliability Engineering and System Safety 92 (2007) 243–251
RELIABILITY ENGINEERING & SYSTEM SAFETY
www.elsevier.com/locate/ress

A non-homogeneous Poisson process predictive model for automobile warranty claims

Karl D. Majeske*

School of Business Administration, Oakland University, Rochester, MI 48309-4493, USA

Received 9 February 2005; received in revised form 18 November 2005; accepted 19 December 2005
Available online 10 February 2006

Abstract

Automobile warranties and thus lifetimes are characterized in the two-dimensional space of time and mileage. This paper presents a non-homogenous Poisson process (NHPP) predictive model for automobile warranty claims consisting of two components: a population size function and a failure or warranty claim rate. The population size function tracks the population in the time domain and accounts for mileage by removing vehicles from the population when they exceed the warranty mileage limitation. The model uses the intensity function of a NHPP—the instantaneous probability of failure—to model the occurrence of warranty claims. The approach was developed to support automobile manufacturers' process of using claims observed during the early portion (first 7 months) of vehicle life to predict claims for the remainder of coverage, typically between 3 and 5 years. This paper uses manufacturer provided warranty data from a luxury car to demonstrate the NHPP model by predicting claims for three vehicle subsystems. Warranty predictions are then compared with the actual observed values and predictions made with a standard forecasting technique.

Keywords: Automobile warranty data; Two-dimensional lifetime; Non-homogeneous Poisson process; Warranty claim predictions

1. Introduction

Current trends in automobile manufacturing and marketing have increased basic warranty coverage from 12 months/12,000 miles to as much as 5 years/50,000 miles on the complete vehicle [1,2], and 10 years/100,000 miles on the power train [3]. These longer coverage periods have increased manufacturer costs and spurred research in modeling warranty data. There are many nuances to warranties and the associated claim data [4]. Manufacturers offer either a renewing warranty in which coverage starts over after product failure or a non-renewing warranty with fixed duration of coverage at time of sale [5]. In the event of product failure, warranties compensate customers via free replacement, lump sum payment, or a pro-rata reimbursement [6]. For free replacement, manufacturers utilize a maintenance policy such as minimal repair (good as old), repair (good as new), or replace [7].

Manufacturers express the duration of warranty coverage using a quantitative definition of product lifetime. When one variable quantifies product life (e.g., time since customer purchase) the product carries a one-dimensional warranty [8]. One-dimensional renewal process cost models exist for both a renewing warranty [9] and a non-renewing free replacement warranty [10]. A one-dimensional renewal process-type model that relaxes the independence assumption on time between arrivals (claims) has been proposed [11]. When two variables quantify product lifetime, the product carries a two-dimensional warranty. Two-dimensional warranty cost models include a non-homogeneous Poisson process (NHPP) model that assumes the manufacturer repairs first failures, and subsequent failures are minimally repaired [12]. Using a NHPP cost model, it has been shown that a minimal repair strategy has lower cost than a replace strategy [13].

Manufacturers sell automobiles with a non-renewing two-dimensional warranty coverage characterized by time and mileage that offers free replacement. The two-dimensional automobile warranty provides customers a flexible coverage based on their personal usage pattern.

*Tel.: +1 248 370 4976.
E-mail address: majeske2@oakland.edu.

0951-8320/$ - see front matter
doi:10.1016/j.ress.2005.12.004

The two-dimensional automobile warranty also protects the manufacturer from replacing components on high mileage newer vehicles and limits long-term manufacturer liability on low usage vehicles. One aspect of automobile warranty that complicated modeling observational data is the lack of complete information on mileage. Defining automobile life purely in the time metric has a great deal of intuitive appeal and facilitates warranty data modeling. Some automobile warranty data models use only the time domain as if the coverage were one-dimensional [3,14].

Automobile manufacturers track warranty claims using a normalized cumulative value. This allows comparing claim rates for vehicles with different sales rates in a common metric. The statistic combines f_i, the observed number or frequency of claims in time period i, and $N(i)$, the number of vehicles at risk (or the population size) for warranty claims at time i. Manufacturers then track cumulative claims (per thousand vehicles) by month in service t using

$$R(t) = \sum_{i=0}^{t} \frac{f_i}{N(i)} 1000. \tag{1}$$

While the notation differs between manufacturers, this approach appears to be used throughout the industry. One method used to predict $R(t)$ values [15–17] utilizes transformed data by regressing $\log[R(t)]$ on $\log(t)$, starting in month 1, to fit the simple regression model

$$\log(R(t)) = \beta_0 + \beta_1 \log(t) + \varepsilon_t. \tag{2}$$

To predict a future $R(t)$ value, denoted $\hat{R}(t)$, simply extend the fit line and transform back to the original units of $R(t)$ using

$$\hat{R}(t) = \exp(\hat{\beta}_0 + \hat{\beta}_1 \log(t)). \tag{3}$$

The motivation for this research was the need to develop a technique that provided the manufacturer warranty predictions more accurate than those obtained with the log–log $R(t)$ model, especially in the later stages of warranty coverage. The manufacturer also desired a model that incorporated mileage, to take the place of the informal "correction factor" applied to time-based forecasts. This paper contains the result of that work, a NHPP predictive model for automobile warranty claims. This technique forecasts the number of warranty claims that a population of vehicles will experience based on the claims observed early in the vehicle lifetime. In addition to providing manufacturers a prediction of the total number of claims, the technique also forecasts the timing of claims during the vehicle lifetime.

The remainder of this paper is organized as follows. Section 2 defines the warranty lifetime of an automobile in the two-dimensional space defined by time and mileage, and introduces the usage function. Section 3 develops the NHPP predictive model that consists of two components: a time-based population size that accounts for mileage by censoring vehicles from the population when they exceed the warranty mileage limitation; and the failure or the warranty claim rate, captured by the intensity function of a NHPP. The model predicts claims by integrating the product of the population size and the intensity function over the time domain of the warranty coverage. Section 4 presents an application of the approach to three subsystems of a luxury automobile using manufacturer supplied warranty data. Section 5 contains the conclusion.

2. Automobile lifetime under a two-dimensional warranty

Fig. 1 shows the warranty timeline for an automobile from final assembly until the warranty time limit. The date of final assembly establishes the physical existence of an automobile but the warranty lifetime does not begin until a customer purchases the vehicle. Let the random variable L represent the sales lag, the time from final vehicle assembly to vehicle sale. Sales lag is an unknown value at time of assembly; however, all vehicles at risk for warranty claims have an observed value because they have been sold to a customer. It has been shown that the larger the sales lag, the more likely a vehicle will observe warranty claims [15]. The cutoff date is the calendar date of a warranty prediction. Let the random variable T represent time in service—the elapsed time from product sale to the cutoff date—which captures product exposure or the amount of time the vehicle was eligible for warranty claims.

Fig. 2 shows the two-dimensional automobile warranty lifetime space which begins at time of sale. Defining the time limitation and the mileage limitation of the two-dimensional coverage as t_{WL} and m_{WL}, respectively, results in a rectangular coverage region. Fig. 2 also shows the usage function $U(t)$, the non-decreasing path through the lifetime space that relates time to mileage for two vehicles. Vehicles repaired while inside the basic coverage region, $U(t)$ such that $(t \leqslant t_{WL} \cap U(t) \leqslant m_{WL})$, are eligible for warranty claims. From Fig. 2, notice that vehicle 2 leaves

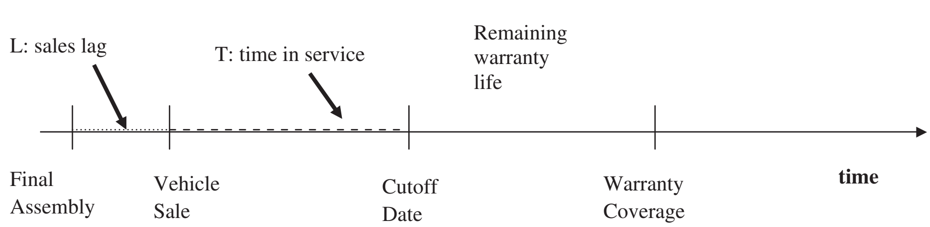

Fig. 1. Automobile warranty timeline.

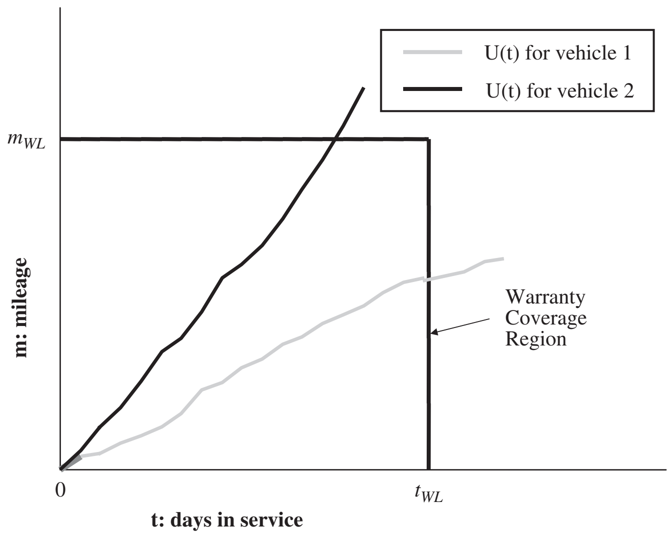

Fig. 2. Two-dimensional automobile warranty lifetime space.

coverage by exceeding the mileage limit when $U(t) > m_{\mathrm{WL}}$ for some $t < t_{\mathrm{WL}}$. Therefore, when making a prediction, the remaining coverage—in both time and mileage—is unknown for each vehicle.

3. Automobile warranty predictive model

This section presents a NHPP predictive model for automobile warranty claims that provides a great deal of flexibility when fitting observational data. The approach presented here makes predictions in the time domain but accounts for mileage by removing vehicles from the population when they exceed the mileage limit.

3.1. Time in service-based number at risk

This model forms vehicle populations based on assembly date. This allows relating warranty claims and costs to time-based events such as product design changes, component manufacturing process changes, assembly process changes, and vendor issues. Given the variation in sales lag, and the random nature of when vehicles leave warranty coverage, the number at risk varies as the population moves through calendar time. Specifically, the number at risk starts at zero, increases by one each time a vehicle is sold, and decreases by one each time a vehicle leaves warranty coverage. To place vehicles in a common metric for warranty predictions, automobile manufacturers characterize lifetime using time in service which begins at vehicle sale. Due to the random nature of sales lag, a population of vehicles assembled together will have varying values for time in service at a given cutoff date.

Let $N(t)$ represent the number of vehicles at risk (or the population size) for warranty claims at time in service t. For a one-dimensional coverage in the time domain, the manufacturer would have complete information on $N(t)$. When making a prediction of total claims for a population of size N

$$N(t) = \begin{cases} N, & t \leqslant t_{\mathrm{WL}}, \\ 0, & t > t_{\mathrm{WL}}. \end{cases} \tag{4}$$

When performing warranty data analysis, many automobile manufacturers use a population definition similar to Eq. (4) due to the limited information on vehicle mileage. Using Eq. (4) fails to remove vehicles when they exceed the mileage limit which, in one analysis, was more than half the population [18]. Ignoring mileage results in overestimating the number of at risk vehicles for higher time in service values. Adjustments for vehicles exceeding the mileage limit have been accounted for in forecasting models [2].

To include the mileage domain, define u as the rate a vehicle accumulates mileage, the derivative or slope of the usage function $U(t)$ shown in Fig. 2. The constant usage rate assumption is common in two-dimensional warranty models [12,19–21] and has also appeared in automobile warranty data models [1]. The constant usage rate assumption allows predicting the time duration of the warranty coverage. Specifically, a vehicle is no longer in the warranty coverage region if the mileage exceeds the mileage limit or

$$ut > m_{\mathrm{WL}}. \tag{5}$$

By letting u_i represent the usage rate of the ith vehicle, the variable

$$\delta_i = \begin{cases} 1 & \text{if } u_i t \leqslant m_{\mathrm{WL}}, \\ 0 & \text{else}, \end{cases} \tag{6}$$

indicates if a vehicle remains at risk for warranty claims at time t. To determine the number at risk at time t, sum the indicator variables of Eq. (6) to obtain

$$N_m(t) = \begin{cases} \sum_{i=1}^{N} \delta_i, & t \leqslant t_{\mathrm{WL}}, \\ 0, & t > t_{\mathrm{WL}}. \end{cases} \tag{7}$$

Eq. (7) defines $N_m(t)$ a population size that quantifies the number at risk in the time domain and takes into account the effects of mileage by removing vehicles when they exceed the mileage limit.

Automobile manufacturers do not have complete information on the population size $N_m(t)$ as they do for $N(t)$, due to the lack of information on vehicle usage for vehicles without warranty claims. Vehicles with a warranty claim have an observed value on the lifetime path that can be used to calculate a vehicle specific value for u. To estimate usage for vehicles without a claim, the manufacturer can either use parametric or non-parametric techniques. A two-dimensional simulation model, motivated by warranty for an automobile starter, uses three rates of mileage accumulation [22]. Authors have suggested log-normal [1] and Gamma distributions [20,23] as parametric models for U, the distribution of usage rates for the vehicles in the population. When using a parametric model for U, estimate the distribution parameters using the observed data and then randomly assign a usage value to vehicles without claims following the parametric model. Alternatively, one could assign usage rates without using a parametric model by randomly selecting values from a Uniform distribution on the $[0, 1]$ interval and assigning the corresponding value from the empirical cumulative distribution function. Once each vehicle has a usage value, either observed or estimated, determine $N_m(t)$ using Eq. (7).

3.2. Time-based warranty claim rate

The second component of the predictive technique models the warranty claims associated with a vehicle. A NHPP provides a flexible model to capture the varied warranty claim patterns observed in automobile warranty data [12,13]. The NHPP is characterized by the intensity function

$$v(t) = \lim_{\Delta t \to 0} \frac{\Pr(\text{observe a failure in interval } t, t + \Delta t)}{\Delta t} \tag{8}$$

that represents the instantaneous probability of a failure. To use the predictive approach in this paper, one must identify a functional form for the intensity function and then use sample data to estimate the parameters. Maximum likelihood approaches exist for estimating NHPP parameters to multiple observations from a single system [24] and for fitting a power-law process to a population where each system has at least one observed failure [25]. However, when modeling automobile warranty data, not all vehicles have experienced a claim (first failure) that complicates NHPP parameter estimation.

For any NHPP, the intensity function and the hazard function of T_1, the time to first failure, have the same functional form [26]

$$v(t) = h(t_1). \tag{9}$$

This provides an alternate approach to characterizing the intensity function. First, estimate the distribution parameters for time to first warranty claim (failure). Then, use the time to first failure coefficient values to characterize a NHPP intensity function. Including non-failed vehicles as censored observations allows incorporating vehicles without a warranty claim when fitting the stochastic process. Standard methods exist for fitting parametric models to right-censored lifetime data [24,27,28] that have been used to estimate time to first failure distributions from automobile warranty data [29].

Identifying a functional form for the time to first warranty claim (or the intensity function) that properly captures the failure process will ultimately determine aptness of the NHPP predictive model presented in this paper. When using the model to fit data ex-ante (at the end of warranty coverage) the modeler will have observed values for a substantial portion of the products useful life. From this complete data one can construct empirical plots (cumulative distribution, survivor, or hazard functions) that provide input to parametric model selection [30]. Hazard plots specifically have been applied to automobile data to as graphical technique for identifying an appropriate parametric model [31]. These graphical techniques also apply when using the NHPP model to fit data early in the coverage period to make forecasts for the remaining coverage; e.g., to set aside reserves to pay for future claims. However, the observational data may not span enough of the useful life to provide an assessment of distributional form. In many applications, manufacturers already use a parametric model for product lifetime, which could serve as a starting point for modeling warranty claims. By first fitting this accepted parametric model to the limited data, the modeler has a baseline model to use when evaluating other possible selections.

3.3. Predicting warranty claims with the model

Let the function $C(t)$ represent the cumulative number of warranty claims observed by the population through t time in service. Let $\hat{C}(t)$ represent the predicted number of claims through time t as the expected value of the cumulative warranty claim function

$$\hat{C}(t) = E[C(t)]. \tag{10}$$

To predict claims for a population of vehicles requires combining the warranty claim rate (intensity function) and the number of vehicles at risk. First, assume a homogenous population or assume that all vehicles in the population have the same intensity function. Predict the expected number of warranty claims by integrating the product of $N_m(t)$, the number at risk function of Eq. (7), and the intensity function

$$\hat{C}(t_{\mathrm{WL}}) = E[C(t_{\mathrm{WL}})] = \int_0^{t_{\mathrm{WL}}} N_m(t) v(t)\, \mathrm{d}t. \tag{11}$$

In some cases, manufacturers may want to relax the homogeneity assumption. Allowing the model parameters to be a function of $\vec{X}$ (covariates such as vehicle option content) would allow the manufacturer to make inference on how these variables affect warranty claim rates. Including covariates makes predicting claims somewhat more difficult because vehicles are at risk for different lengths of time. Letting t_i represent the duration of time the ith vehicle is at risk for warranty claims, the expected number of claims for the population of vehicles is

$$\hat{C}(t_{\mathrm{WL}}) = E[C(t_{\mathrm{WL}})] = \sum_{i=1}^{N} \int_0^{t_i} v(t, \vec{X}_i)\, \mathrm{d}t. \tag{12}$$

The complicating factor with this approach is that the t_i values depend on the vehicle mileage, a random variable with limited or no information.

An automobile is a collection of mechanical, electrical, cosmetic and personal comfort subsystems driven by very diverse failure mechanisms. The resultant claims represent a mixture of these failure distributions. To account for these various warranty claim patterns, this research suggests predicting warranty claims by vehicle subsystem (e.g., powertrain, wiring, fit and finish, etc.) and then aggregating subsystem predictions to obtain a prediction for the population of vehicles. To account for differences in option content (e.g., V6 versus V8 engine) the manufacturer can stratify the population for each subsystem and make a prediction with Eq. (11) for each strata within each subsystem. Then, aggregate the predictions across strata (option content differences) and subsystem (portions of the vehicle) to arrive at a prediction for the population of vehicles.

4. Application of predictive technique

This section assesses the predictive validity of the NHPP model using a set of 9168 luxury cars. Manufacturers use warranty claim data available at a cutoff date to predict warranty claims through some future date (a prediction period). To emulate the manufacturer's prediction process, this section predicts future warranty claims using only the data available to the manufacturer at a cutoff date approximately 7 months after vehicle assembly. Predictions are made for a time period of 45 months that allows comparing predictions to actual claims and thus assess the predictive validity of the NHPP technique.

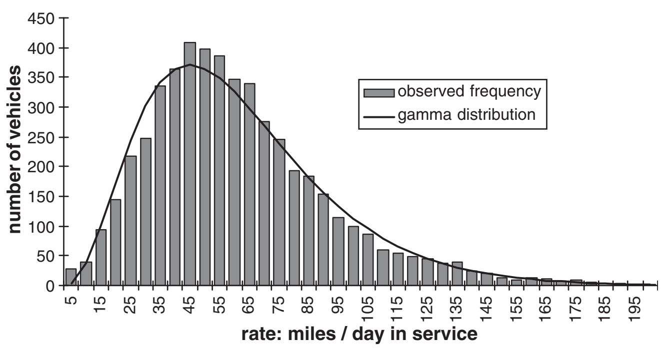

Fig. 3. Histogram of observed usage rates assuming a Gamma distribution.

4.1. Estimating the number at risk

To estimate the number at risk function $N_m(t)$ for the prediction period, a usage value u was calculated for each of the 5127 vehicles with an observed mileage value prior to the cutoff date. A histogram of the usage values appears as Fig. 3 and suggests using a Gamma distribution with probability density function

$$f(u) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} u^{\alpha-1} \mathrm{e}^{-u/\beta}, \quad u > 0. \tag{13}$$

The sample average $\bar{u} = 58.99$ and sample variance $S_u^2 = 980.98$ were set equal to the mean and variance of the Gamma. Solving these two equations yielded the parameter estimates $\hat{\alpha} = 3.55$ and $\hat{\beta} = 16.63$, and the fit distribution also appears in Fig. 3. To generate usage values for vehicles without a claim, usage rates were then simulated for the 4041 vehicles without an observed value from a gamma distribution with $\alpha = 3.55$ and $\beta = 16.63$ to match the observed data.

Table 1 contains the predicted number at risk $\hat{N}_m(t)$ values through the future cutoff date calculated using Eq. (7). Table 1 also contains the $N(t)$ values, the known population size ignoring the effect of mileage, calculated with Eq. (4). Fig. 4 plots the data of Table 1 to provide a graphical comparison of the estimated number at risk function $\hat{N}_m(t)$ with the time only population size $N(t)$. Notice that by 3 years in service, the $\hat{N}_m(t)$ estimates more than half the vehicles have left warranty coverage by exceeding the mileage limit.

Table 1
Population size by time in service (months)

Time in service (Months)	$\hat{N}_m(t)$	$N(t)$	Time in service (Months)	$\hat{N}_m(t)$	$N(t)$	Time in service (Months)	$\hat{N}_m(t)$	$N(t)$
0	9168	9168	16	8480	9168	32	4561	9168
1	9168	9168	17	8296	9168	33	4328	9168
2	9168	9168	18	8082	9168	34	4098	9168
3	9168	9168	19	7875	9168	35	3903	9168
4	9168	9168	20	7620	9168	36	3705	9168
5	9168	9168	21	7360	9168	37	3538	9168
6	9166	9168	22	7131	9168	38	3340	9168
7	9166	9168	23	6849	9168	39	3173	9168
8	9158	9168	24	6601	9168	40	2979	9052
9	9141	9168	25	6341	9168	41	2810	8895
10	9109	9168	26	6059	9168	42	2586	8691
11	9061	9168	27	5794	9168	43	2383	8450
12	8992	9168	28	5523	9168	44	2156	8165
13	8881	9168	29	5290	9168	45	1963	7813
14	8754	9168	30	5054	9168			
15	8620	9168	31	4806	9168			

4.2. Estimating the NHPP intensity function

This section characterizes the NHPP intensity function for three subsystems of the luxury cars. To do this, a parametric model must be selected for time to first warranty claim. Empirical hazard plots, constructed from the warranty data available 7 months after vehicle assembly, give input to the model selection. Then model parameters are then estimated for the model from the observational data.

Subsystem 1:

The Weibull-uniform mixture model with cumulative distribution function

$$F(t) = \begin{cases} p\left(\dfrac{t}{A}\right) + q(1 - \mathrm{e}^{-(\alpha t)^{\beta}}), & 0 \leqslant t < A, \\ p + q(1 - \mathrm{e}^{-(\alpha t)^{\beta}}), & t \geqslant A \end{cases} \tag{14}$$

was derived to fit warranty claims from subsystem 1 of these same luxury cars [29]. To apply the prediction method, the mixture model parameters were estimated using only the data available 7 months after vehicle assembly. Table 2 contains the mixture distribution parameter estimates obtained via maximum likelihood estimation and the standard errors estimated with the observed Fisher information matrix. Fig. 5 shows the empirical hazard function $\hat{h}(t)$ and the mixture hazard function

$$h(t) = \begin{cases} \dfrac{\frac{p(1-\theta)}{A} + (1-p)[\alpha\beta(\alpha t)^{\beta-1}\mathrm{e}^{-(\alpha t)^{\beta}}}{1 - p\left(\theta + \frac{(1-\theta)t}{A}\right) - (1-p)(1 - \mathrm{e}^{-(\alpha t)^{\beta}})}, & 0 \leqslant t < A, \\ \alpha\beta(\alpha t)^{\beta-1}, & t \geqslant A \end{cases} \tag{15}$$

with the parameter estimates from Table 2 and using $A = 17$ weeks.

Subsystem 2:

Reliability engineers often use the Weibull distribution to model field failure and bench test data for subsystem 2. Modeling time to first failure with a Weibull distribution results in the power-law process [32], a special case of the NHPP. The Weibull distribution was fit to the time to first subsystem 2 claim data available at the current cutoff. The parameter estimates and associated standard errors appear in Table 2. Fig. 6 plots the Weibull hazard with the observed values for subsystem 2.

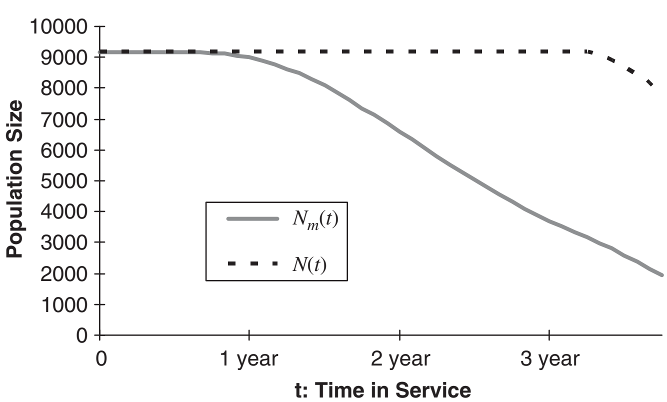

Fig. 4. Estimated population size with and without adjusting for mileage.

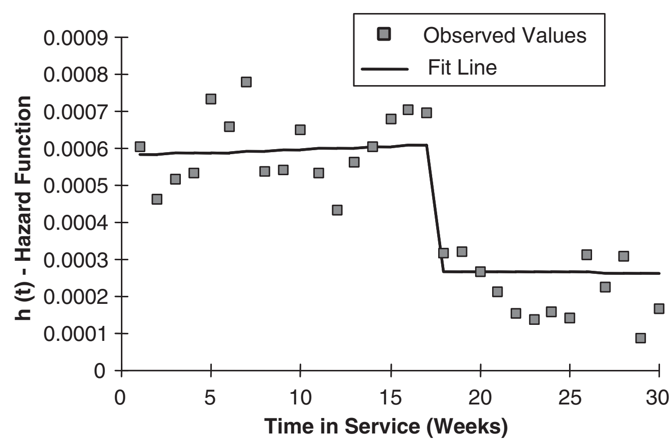

Fig. 5. Subsystem 1—empirical and mixture model hazard functions.

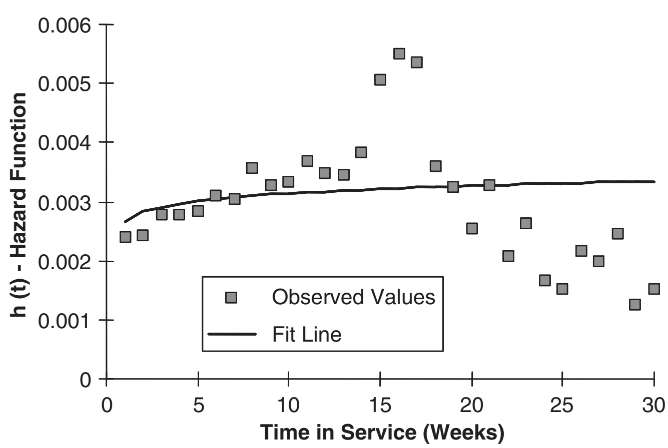

Fig. 6. Subsystem 2—empirical and Weibull hazard functions.

Table 2
Parameter estimates for time to first warranty claim

Subsystem 1 mixture distribution			Subsystem 2 Weibull distribution			Subsystem 3 linear hazard		
Parameter	Estimate	Std. error	Parameter	Estimate	Std. error	Parameter	Estimate	Std. error
α	0.00027	0.000021	α	0.00324	0.00006	β_0	0.0001558	0.0000379
β	1.006	0.0339	β	1.057	0.0154	β_1	0.0000174	0.0000022
p	0.047	0.0048						
θ	0.232	0.0295						

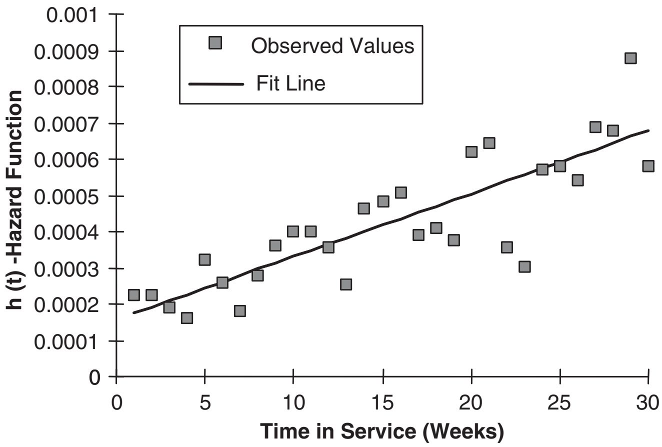

Fig. 7. Subsystem 3—empirical and linear hazard functions.

Subsystem 3:

The manufacturer did not use a parametric technique for modeling the lifetime distribution of subsystem 3. After evaluating the empirical hazard plot shown in Fig. 7, the linear hazard function

$$h(t) = \beta_0 + \beta_1 t, \tag{16}$$

was identified for the subsystem 3 data. The linear hazard of Eq. (16) defines a NHPP with intensity function

$$v(t) = \beta_0 + \beta_1 t. \tag{17}$$

The parameters of the linear hazard model were estimated using ordinary least squares by fitting a line to the empirical hazard function. The parameter estimates and associated standard errors appear in Table 2. Fig. 7 shows the empirical hazard along with the linear hazard, which appears to provide a good fit to the data.

4.3. Predicting warranty claims

This section predicts warranty claims through 45 months in service using the estimated number at risk function $\hat{N}_m(t)$ and NHPP intensify functions developed in Sections 4.1 and 4.2, respectively. To make predictions, Eq. (11) was evaluated for each 30 day (monthly) interval. The predictive validity of the NHPP process model is evaluated in two ways. First, predictions are compared with observed values to give an indication of the technique performs. Secondly, the NHPP predictions are compared to predictions made using Eq. (3) with the log–log $R(t)$ method. These forecasted cumulative claims can be converted into monthly counts by taking the difference in success $R(t)$ values and adjusting for the number at risk

$$\frac{[\hat{R}(t) - \hat{R}(t-1)]N(t)}{1000}. \tag{18}$$

Comparing predictions made with the NHPP model to the log–log $R(t)$ model provides insight into the potential improvement in warranty forecasts associated with the NHPP model.

Fig. 8 shows the observed and predicted claims by month in service (MIS) for subsystem 1. Evaluating Fig. 8, the Weibull portion of the mixture ($t > 4$ MIS) overpredicts in the period 9–22 MIS and underpredicts in the period 28–45 MIS. While the model has some correlation structure in the residuals, it does provide a good fit to the subsystem 1 warranty claims. For the 45 month period, the NHPP process underpredicts the total number of claims by less than 1% of the actual. From Fig. 8, notice the log–log $R(t)$ model consistently overpredicts the number of claims after 4 MIS due to the high rates of claims early in product life. For the 45 month period, the $R(t)$ model overpredicts total claims by 73%.

Fig. 9 shows the observed and predicted claims by MIS for subsystem 2. From Fig. 9, the power law process underpredicts the subsystem 2 infant mortality failures. By comparing Fig. 9 to Fig. 6, notice that the relative magnitude of claims in the $0 > 4$ MIS region has increased.

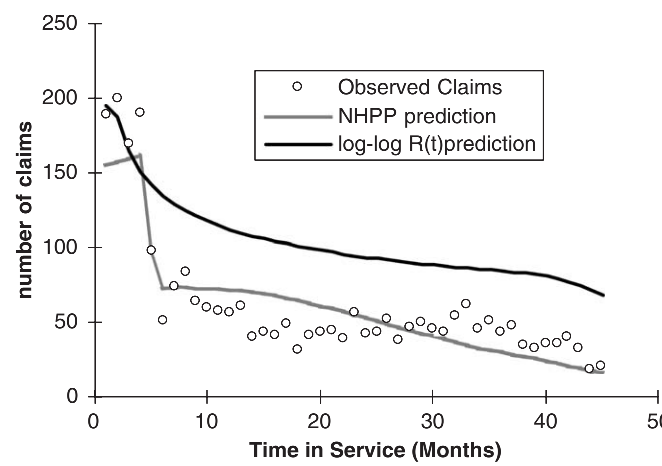

Fig. 8. Subsystem 1—mixture model prediction with observed claims.

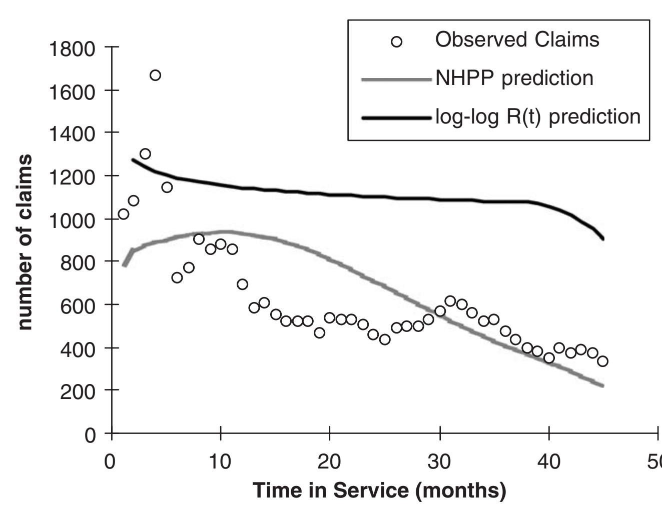

Fig. 9. Subsystem 2—power-law process prediction with observed claims.

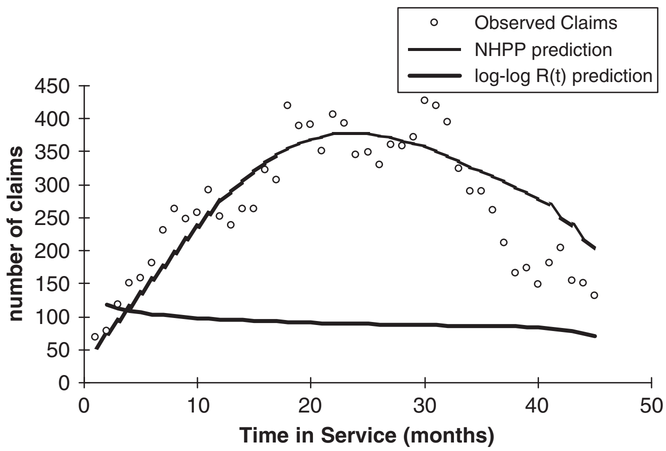

Fig. 10. Subsystem 3—linear intensity function model prediction with observed claims.

This suggests that vehicles with higher sales lag experienced higher infant mortality claim rates. The NHPP model also does not predict the downward trend at 10 MIS; however, the model very accurately predicts claims from 30 to 45 MIS. For the entire prediction period, the NHPP model overpredicts total claims by 7%. As with subsystem 1, the log–log $R(t)$ model consistently overpredicts the number of warranty claims with an over prediction of 62% for the 45 month period.

Fig. 10 shows the observed and predicted claims by month in service for subsystem 3. Without removing vehicles that exceed the mileage limitation from the number at risk for claims the NHPP prediction would continue the linear path established in the 1–10 MIS interval. This suggests that the leveling off of claim rates in the 20–30 MIS range, and subsequent drop in rates over 30 MIS are purely due to vehicles leaving warranty coverage. In other words, vehicles continue to experience these failures at the same (constant) rate but more and more customers bares the financial burden of repair. Fig. 10 shows the linear intensity model lends a good fit to the subsystem 3 data, yet slightly over predicts claims after 30 MIS. For the entire coverage period the NHPP model overpredicts total claims by 5%. Due to the relatively low claim rates early in vehicle life, the log–log $R(t)$ model dramatically underpredicts total claims. For the 45 month prediction period, the log–log $R(t)$ model predicts less than one-third the observed claims.

5. Conclusion

Automobile manufacturers rely on predictions for the number, timing, and cost of future warranty claims. Omitting the mileage data contained on automobile warranty claims, and making predictions purely in the time domain, ignores vehicles leaving coverage prior to the warranty time limit. Incorporating a two-dimensional aspect will result in a predictive model that better represents the process being modeled. This paper develops a non-homogenous Poisson process (NHPP) predictive model that has a parametric component (time to first failure) and provides a great deal of flexibility in application. The NHPP approach also allows incorporating past experience when identifying the failure process and bases predictions on early field performance of the products. Using the NHPP will result in more accurate predictions of warranty claims and support decision making when implementing engineering design and manufacturing process changes. More accurate predictions will also assist the manufacturer in allocating reserves to pay for repairs covered under warranty.

References

[1] Rai BK, Singh N. A modeling framework for assessing the impact of new time/mileage warranty limits on the number and cost of automotive warranty claims. Reliab Eng Syst Saf 2005;88:157–69.
[2] Kleyner A, Sandborn P. A warranty forecasting model based on piecewise statistical distributions and stochastic simulation. Reliab Eng Syst Saf 2005;88:207–14.
[3] Yang K, Cekecek E. Design vulnerability analysis and design improvement by using warranty data. Qual Reliab Eng Int 2004; 20:121–33.
[4] Karim MR, Suzuki K. Analysis of warranty claim data: a literature review. Int J Qual Reliab Manage 2005;22:667–86.
[5] Chukova S, Hayakawa Y. Warranty cost analysis: quasi-renewal inter-repair times. Int J Qual Reliab Manage 2005;22:687–98.
[6] Moskowitz H, Chun YH. A Poisson regression model for two-attribute warranty policies. Nav Res Log 1994;41:355–76.
[7] Chen T, Popova E. Maintenance policies with two-dimensional warranty. Reliab Eng Syst Safety 2002;77:61–9.
[8] Blischke WR, Murthy DNP. Warranty cost analysis. New York: Marcel Dekker; 1994.
[9] Yeh RH, Chen GC, Chen MY. Optimal age-replacement policy for nonrepairable products under renewing free-replacement warranty. IEEE Trans Reliab 2005;54:92–7.
[10] Chukova S, Hayakawa Y. Warranty cost analysis: non-renewing warranty with repair time. Appl Stochastic Models Bus Ind 2004; 20:59–71.
[11] Elkins DA, Wortman MA. An examination of correlation effects among warranty claims. IEEE Trans Reliab 2004;53:200–4.
[12] Iskandar BP, Murthy DNP, Jack N. A new repair-replace strategy for items sold with a two-dimensional warranty. Comput Oper Res 2005;32:669–82.
[13] Baik J, Murthy DNP, Jack N. Two-dimensional failure modeling with minimal repair. Nav Res Log 2004;51:345–62.
[14] Wu H, Meeker WQ. Early detection of reliability problems using information from warranty databases. Technometrics 2002;44(2): 120–33.
[15] Robinson JA, McDonald GC. In: Liepins GE, Uppuluri VRR, editors. Issues related to field reliability and warranty data. Data quality control: theory and pragmatics. New York: Marcel Dekker; 1991. p. 69–90.
[16] Majeske KD, Riches MD, Annandi HP. In: Blischke WR, Prabhakar Murthy DN, editors. Ford's reliability improvement process—a case study on automotive wheel bearings. Case studies in reliability and maintenance. New York: Wiley; 2003. p. 545–69.
[17] Wasserman GS. An application of dynamic linear models for predicting warranty claims. Comput Ind Eng 1992;22(1):37–47.
[18] Rai BK, Singh N. Modeling and analysis of automobile warranty data in presence of bias due to customer-rush near warranty expiration limit. Reliab Eng Syst Saf 2004;86:83–94.
[19] Frickenstein SG, Whitaker LR. Age replacement policies in two time scales. Nav Res Log 2003;50:592–613.

[20] Iskandar BP, Murthy DNP. Repair-replace strategies for two-dimensional warranty policies. Math Comput Model 2003;38: 1233–41.
[21] Pal S, Murthy GSR. An application of Gumbel's bivariate exponential distribution in estimation of warranty cost of motor cycles. Int J Qual Reliab Manage 2003;20:488–502.
[22] Manna DK, Pal S, Kulandaiyan A. Warranty cost estimation of a multi-module product. Int J Qual Reliab Manage 2004;21: 102–17.
[23] Phillips MJ. Bayesian estimation from censored data with incomplete information. Qual Reliab Eng Int 2004;20:237–45.
[24] Meeker WQ, Escobar LA. Statistical methods for reliability data. New York: Wiley; 1998.
[25] Rigdon SE, Basu AP. Statistical methods for the reliability of repairable systems. New York: Wiley; 2000.
[26] Thompson Jr. WA. On the foundations of reliability. Technometrics 1981;23:1–13.
[27] Blischke WR, Murthy DNP. Reliability modeling, prediction and optimization. New York: Wiley; 2000.
[28] Lawless JF. Statistical models and methods for lifetime data. New York: Wiley; 1982.
[29] Majeske KD. A mixture model for automobile warranty data. Reliab Eng Syst Saf 2003;81:71–7.
[30] Hoyland A, Rausand M, editors. System reliability theory, models and statistical methods. New York: Wiley; 2004.
[31] Majeske KD, Lynch-Caris T, Herrin GD. Evaluating product and process design changes with warranty data. Int J Prod Econ 1997;50:79–89.
[32] Rigdon SE, Basu AP. The power law process: a model for the reliability of repairable systems. J Quality Technol 1989;21:251–60.