Springer

KSME

Journal of Mechanical Science and Technology 31 (6) (2017) 2747~2757
www.springerlink.com/content/1738-494x(Print)/1976-3824(Online)
DOI 10.1007/s12206-017-0518-1

Data-driven modeling of truck engine exhaust valve failures: A case study †

Yusen He[1,*], Andrew Kusiak[1], Tinghui Ouyang[2] and Wei Teng[3]

[1]*Department of Mechanical and Industrial Engineering, The University of Iowa, Iowa City, IA, 52242, USA*
[2]*School of Electrical Engineering, Wuhan University, Wuhan 430072, China*
[3]*School of Energy Power and Mechanical Engineering, North China Electric Power University, Beijing 102206, China*

(Manuscript Received September 26, 2016; Revised January 9, 2017; Accepted February 9, 2017)

Abstract

Exhaust valve is an essential part of truck engine. Dynamic and unpredictable thermal and mechanical stress cause valves to wear prematurely, leading to increased maintenance costs. In this paper, a data-driven approach is presented to predict failures of exhaust valves of truck engines. The failure datasets of exhaust valves recorded from 13 truck engines are divided into three groups: First failure, second failure, and third or more failures. The Kaplan-Meier estimator is selected to express the distribution of survival probability of the three groups of failures. In order to find the hazard indicator, two data-mining algorithms, a wrapper and a boosting tree are applied to select parameters highly relevant to the hazard rate. A Cox proportional hazard model is used to conduct regression analysis on each selected parameter. Based on the derived hazard ratio, the time-dependent baseline hazard rate is computed. Five parametric reliability models are selected to capture the baseline hazard rate for the three groups. The value-at-risk for each group of failures is computed to express the risk at different confidence levels. Life circle of truck engine exhaust valves can be estimated.

Keywords: Exhaust valve failure; Multi-dimensional imputation; Kaplan-Meier estimate; Cox proportional regression; Reliability model fitting; Value-at-risk; Kolmogorov-Smirnov two sample test

1. Introduction

Failures of truck engines may have hazardous consequences, e.g., a sudden stop of a truck carrying a heavy load. Depending on its severity, a failure may take long hours to fix, thus causing economic losses.

One of the most frequent failed components of a truck engine is the exhaust valve. Truck engine manufacturers usually use 21-4N heat-resistant steel to produce the plate material of exhaust valves due to its outstanding heat-resistance and PbO-corrosion resistance. However, being exposed to thermal and mechanical overstress after long period, the material structure of exhaust valve would experience decomposition which decreases the toughness, strength, loading capacity and corrosion resistance. For this reason, exhaust valves are more vulnerable to erosion and combustion than other truck engine components. Therefore, aged exhaust valves usually fail as a results of wearing, fatigue and corrosion.

A detection ahead of exhaust valve failure occurrence is essential for manufacturers to produce engines with better quality and reduced failure rates. Statistical analysis of failure distribution is critical for failure detection and prediction. For a preliminary analysis of the survival probability distribution of incomplete time-to-event datasets, the Kaplan-Meier non-parametric method can be used [1]. However, the Kaplan-Meier curve may only illustrate the general survival curve. A specific type of partial hazard is unlikely to be detected. The Cox partial likelihood regression can overcome this deficiency to detect a potential hazard or latent hazard factor among all variables in a dataset [2]. The hazard ratio is calculated and a baseline hazard rate is evaluated.

The reliability literature includes various models for health condition monitoring of components, assemblies, and products. Parametric models applied to fit the failure data and hazard rate has been widely discussed. The Weibull model has served as the basis of reliability models or their ensembles [3]. Mudholkar [4] developed a Data-transformation model (DTM) to describe the relationship between reliability and time. Jiang [5] proposed a 3-parameter Finite support model (FSM) based on the Weibull model to estimate the reliability of failure data in a finite time interval. A new exponential-type mixed Weibull model (ENH) was proposed by Lemonte [6] for fitting bathtub shaped failures. Wang [7] developed a generalized 4-parameter Finite interval lifetime distribution model for reliability engineering (FIRE). The models published in the literature considered the failure time only without attention to the partial hazard.

*Corresponding author. Tel.: +1 319 333 5337, Fax.: +1 319 335 5669
E-mail address: yusen-he@uiowa.edu
† Recommended by Editor Chongdu Cho

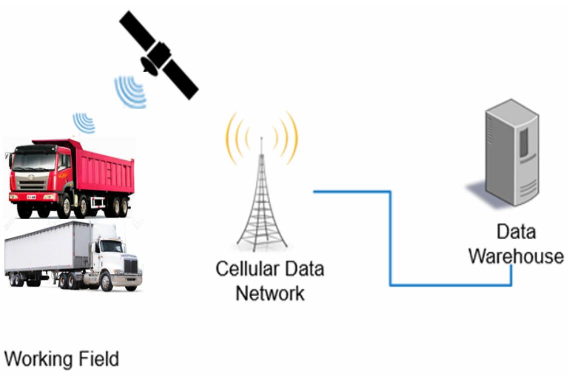


Fig. 1. Data collection process.

The detection of exhaust valve failures is challenging as their occurrence is difficult to predict. The need for dynamic evaluation of risks of a failure is apparent. Survival analysis allows estimating the time-dependent hazard rate. Cox regression is essential to determine the baseline hazard. For discrete hazard data, the performance of different parametric models varies. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) are commonly used for model evaluation. Based on the fitted parametric models, the value-at-risk approach measures the risk of an exhaust valve failure at different confidence levels and hence it informs a user ahead of the event occurrence.

2. Case study

2.1 Experiment design and data collection

The study discussed in this paper was conducted on the data collected from engines of 13 different trucks illustrated in Fig. 1. The dataset includes 26 variables, e.g., temperature, pressure, fuel consumption, recorded every 30 minutes. All truck engines were monitored with these 26 variables for 320 days. Performance data was recorded only when the truck was functioning. Hence, the number of records of each truck engine varies due to different functioning time lengths. Meanwhile, missing values were also generated during the functioning times. Alerts were also included in the dataset.

2.2 Data cleaning and multiple imputation

The data collected from 13 trucks was merged into two datasets for analysis. The first dataset contains exhaust valve failure data of ten trucks and is utilized to build models. The second dataset contains exhaust valve failure data of the rest three trucks and is used for model validation process. Numerous values of variables in the dataset were missing, including the average inlet air temperature, average Charge-air-cooler (CAC) temperature, average fuel temperature, time since last regeneration, last soot time, period fuel consumption, and total fuel consumed. The percentage of missing values and number of records are provided in Table 1. A Neural network (NN) imputation method was used to generate the missing values [8].

Table 1. Percentage of missing values.

Dataset	Number of records	Percentage of missing values
1	6709	2.70 %
2	16606	9.40 %
3	18579	5.30 %
4	15511	0.40 %
5	13118	7.80 %
6	12924	3.30 %
7	16429	3.70 %
8	16956	6.10 %
9	13962	10.60 %
10	14476	2.90 %
11	13037	4.40 %
12	14925	11.30 %
13	16411	0.90 %

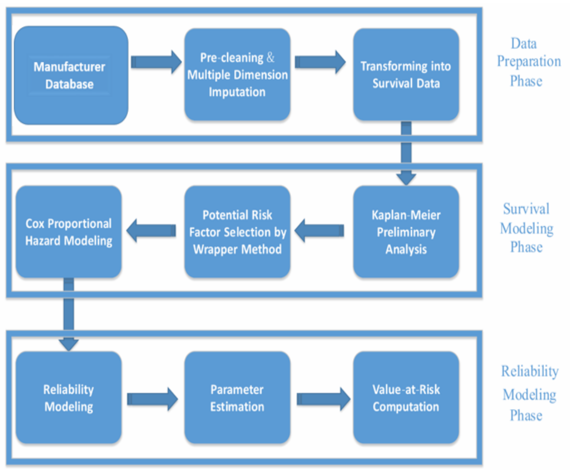


Fig. 2. Methodological framework.

To impute the missing values, the complete records were used to train a Neural network (NN) and the missing values were predicted. The survival analysis performed on the dataset is discussed in the next section.

2.3 Methodological framework

After data cleaning and imputation, survival analysis was conducted on the dataset containing failure occurrences. Reliability modeling was also implemented on the baseline hazard rates. The methodological framework of the data-driven approach is illustrated in Fig. 2.

3. Survival analysis of exhaust valve failures

3.1 Preliminary data description

Survival analysis has been widely applied, including medi-

Table 2. Failure records.

Number	First failure (days)	Second failure (days)
1	4.1	19.05
2	8.17	24.95
3	11.35	47.40
4	46.67	63.30
5	46.95	74.95
6	52.58	184.14
7	71.96	184.97
8	183.29	
9	226.74	
10	228.79	

cal research, e.g., human life prediction, clinical trials, and experiments. Survival data represents the duration between the starting and the end-point time of an event. For instance, the duration between two heart attacks of a patient in a clinical trial is a survival time [9].

In the case discussed in this paper, the target is to predict a truck engine exhaust valve failure. The exhaust valve failure is analogous to the heart attack of the clinical domain. During our experiment, once there was an exhaust valve failure, it was fixed immediately and then the truck became operational. Since the data was collected from 13 different trucks, each with multiple failure events, it became a multiple events dataset. Kelly [10] categorized multiple events data as recurrent or multiple events. In this paper, time is considered as a predictor variable and the warranty claim is considered as a dependent variable. The initial recurrence dataset is characterized.

In the dataset, there are a total of 46 exhaust valve failures recorded between the first 24 engine hours and the 7660th engine hour (320th day). The risk set is 10 for the first groups of failures, and it is decreasing in longer engine hours. In the second group, the risk set is 7. In the third group, every repaired engine is considered as one risk set and hence the number of risk set is 29. The detailed failure times are summarized in Table 2. There are 10 first failures and 7 second failures. The number of third failures is 29. Due to the limit of article space, only the first failures and second failures are illustrated.

3.2 Kaplan-Meier estimate of survival function

There are four commonly used probability distribution functions of survival analysis discussed next [10].

Survival function (1) that measures the cumulative survival probability at time t. It is a decreasing function that evaluates the probability that the survival time T is larger than t.

$$S(t) = \Pr(T > t) \ . \tag{1}$$

Cumulative failure function (2) that measures the probability that the target fails before time t or the survival time T of

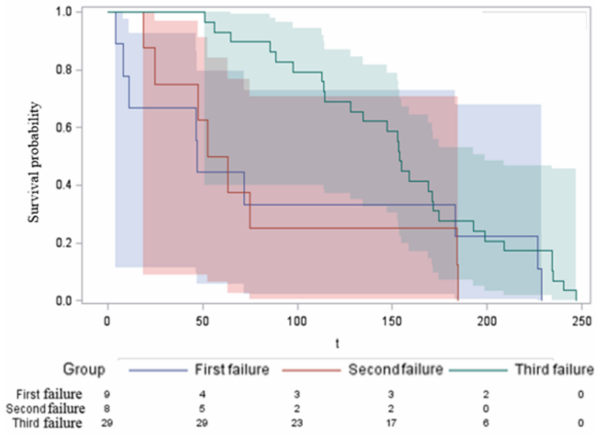


Fig. 3. Log-transformed 95 % confidence interval of the Kaplan-Meier estimates.

the target is smaller than t.

$$F(t) = \Pr(T \le t) = 1 - S(t) \ . \tag{2}$$

Failure density function (3) is the Probability density function (PDF) of a failure. It evaluates the mortality rate of the target at the time point t.

$$f(t) = \frac{dF(t)}{dt} = \frac{d}{dt}\{1 - S(t)\} = -S'(t) \ . \tag{3}$$

Hazard rate function (4) which is also known as a conditional failure rate function measuring the age-specific failure rate or the force of mortality [13]. It is the instantaneous failure rate of the truck engines.

$$h(t) = \lim \frac{\Pr(t + \Delta t > T > t \mid T > t)}{\Delta t} = \frac{f(t)}{S(t)} \ . \tag{4}$$

As the initial dataset contains risk sets and the time of each failure, non-parametric estimation of incomplete observations can be performed [1]. The Kaplan-Meier estimator (5) is selected to express the survival probability of each time point a claim is made.

$$S(t) = \prod_{i=1}^{j-1} \left(1 - \frac{s_i}{r_i}\right), \quad y_{j-1} \le t \le y_j \tag{5}$$

where i denotes the ith exhaust valve failure; j denotes the total number of failures prior to time t. The Kaplan-Meier estimates of the survival probability for three types of failures $S(t)$ are computed by SAS programming and are illustrated in Fig. 3.

To construct the log-transformed confidence interval, the variance of the Kaplan-Meier estimate needs to be obtained. Greenwood approximation is a common approach to approximate the variance of Kaplan-Meier estimator [12]. The Greenwood expression for the variance is presented in Eq. (6).

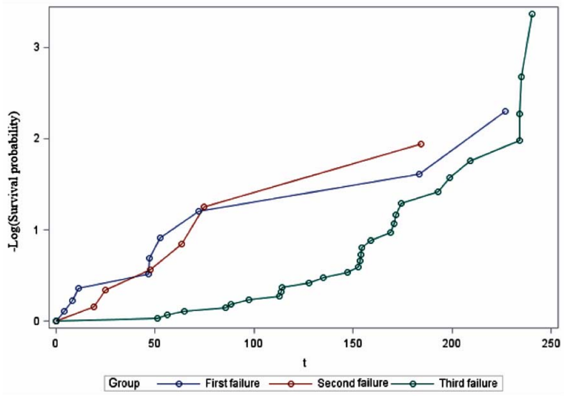
$$Var(S(t)) = S(t)^2 \sum_{y_j \le t} \frac{s_j}{r_j(r_j - s_j)} \tag{6}$$

where j is the j^{th} observation of failure; T is the time of occurrence of the j^{th} failure; $S(t)$ is the survival probability at time t; and s_j is the number of failures at the j^{th} failure. Hence, the log-transformed confidence interval is $(S_j(t)^{1/U}, S_j(t)^U)$ with U defined in Eq. (7).

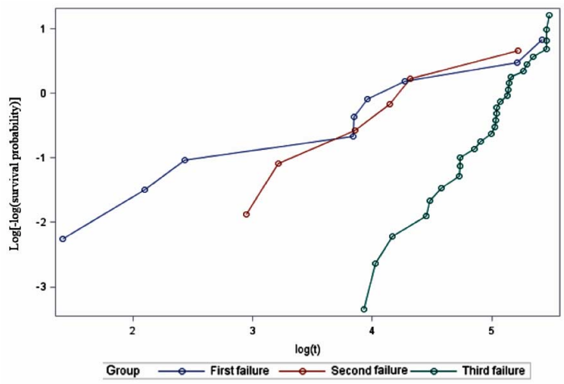
$$U = \exp\left(\frac{Z_{(1+p)/2}\sqrt{Var(S_j(t))}}{S_j(t)\ln(S_j(t))}\right) \tag{7}$$

where p is the confidence level of this interval [11]. For p = 95 %, there is a 95 % probability that the survival probability would fall within the range. The Kaplan-Meier curve of the three groups of failures are derived along with the log-transformed confidence interval at every failure event time illustrated in Fig. 3. The computed negative log-transformed survival probabilities and the log of negative log-transformed survival probabilities are illustrated in Fig. 4. In Fig. 3, the third failure refers to the third or more failures.

By merging Eqs. (3) and (4), the time-dependent hazard rate can be computed by the formula $h(t) = f(t)/S(t)$. And the pdf of failure rate can be computed by $f(t) = -S'(t)$. Hence, the hazard rate can calculated from $h(t) = -(S'(t)/S(t))$.



(a)



(b)

Fig. 4. Negative log$S(t)$ and log of negative log$S(t)$ of survival probabilities.

4. Dimension reduction and cox regression analysis

4.1 Variable selection

In the combined dataset, there are in total 26 different variables indicating the performance of the truck engines. However, not every variable is relevant to the failure of exhaust valves. A variable selection process is needed to eliminate the less important variables. The frequently used variable selection approaches include: Wrapper method, Random forest, and Boosting tree [13].

Boosting tree involves training a number of classifiers. The importance of variables determined by the boosting tree algorithm is provided in Fig. 5. The most relevant variables (with importance above 80 %) are: Maximum fuel temperature, Lifetime aborted regeneration, Lifetime complete regeneration, Present aborted regeneration, Periodic fuel consumption, Number of alerts, Survival time length, Average barometric pressure, Average CAC temperature, Maximum coolant temperature, Parked regeneration time, Total fuel used, and Beginning engine hours. Based on the computational results from Boosting tree algorithms, they have the high relevancy to the hazard rate of exhaust valve failures.

The Wrapper method considers all possible subsets of variables and examines their importance using different algorithms [7]. In this paper, the following algorithms have been used within the Wrapper method: Genetic search, Best first search, Linear forward selection, Decision tree, Logistic regression, Multi-layer perceptron (MLP). The results of the Wrapper method are summarized in Table 3.

Illustrated in Table 3, the Beginning engine hours, Average barometric pressure, Maximum coolant pressure, Cum ETM time, Maximum DPR state, Maximum fuel temperature and Survival time length have the highest number of votes (above 5) and hence they are considered as relevant variables to the hazard rate. Combining results in Fig. 5, four variables as Beginning engine hours, Average barometric pressure, Maximum coolant temperature, and Maximum fuel temperature are selected as potential risk factors for the Cox regression analysis discussed in the next section.

4.2 Cox regression analysis

Cox [2] proposed a semi-parametric proportional model to evaluate multiple hazardous factors in biomedical applications. It has been widely used in modeling time-to-event data with censoring and covariates [14].

The semi-parametric Cox proportional hazard model performs well for datasets with missing values, censored or truncated data, and noisy data [15]. It also considers potential risk

Table 3. Voting-based Wrapper's method.

Variable name	Decision tree + genetic search	Logistic regression + best first search	MLP + linear forward selection	Total votes
Beginning engine hours	9	8	9	26
Average barometric pressures	10	3	3	16
Maximum coolant temperature	3	3	5	11
Maximum fuel temperature	4	3	3	10
Cumulative etm time	1	6	2	9
Max diesel particulate filters state	4	1	3	8
Survival time length	1	1	6	8
Lifetime aborted regens	0	2	3	5
Lifetime complete regens	2	2	1	5
Number of alerts	0	2	3	5
Inhibit switch time	0	2	2	4

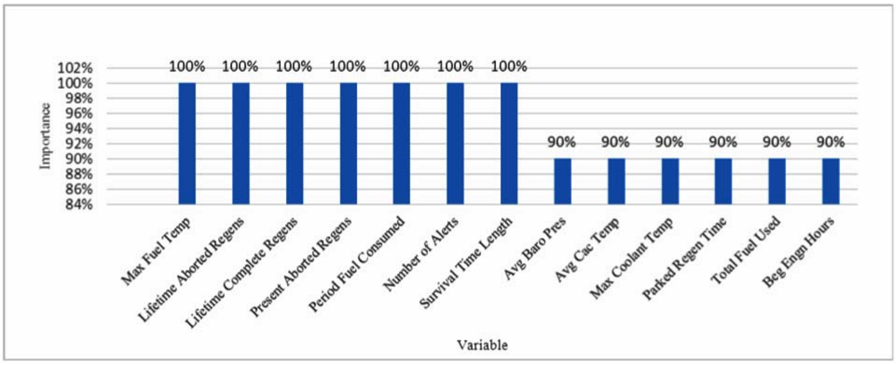


Fig. 5. Boosting tree variable selection results.

factors [16]. Hence, it is selected to evaluate proportional hazard rate of each hazard factor selected in Table 3 and Fig. 5.

The Cox proportional model is expressed in Eq. (8). $h_i(t)$ is defined as the hazard rate of i^{th} hazard factor that impacts the survival time. In other words, it is the instantaneous failure rate at time t. $h_0(t)$ is defined as the basic hazard rate without any impact from all risk factors at time t. Note that the $h_0(t)$ values have been derived in Sec. III. Hence, the general formula of Cox proportional hazard regression model is expressed in Ref. [17].

$$h_j(t) = h_0(t)\exp(\sum_{i=1}^{m} \beta_i X_{ij}) \tag{8}$$

where m denotes the total number of potential risk factors that impact the hazard rate; β_i is the weight assigned to each potential risk factor. To estimate the parameter β_i, the maximum likelihood estimation is widely applied [15]. If the j^{th} observation that is alive at time t_j but fails at time t_{j+1}, it has the instantaneous proportional hazard rate presented in Eq. (9) [17].

$$h(t_j, x_j) = h_0(t_j)\exp(\sum_{i=1}^{4} \beta_i X_{ij}) . \tag{9}$$

At time t_j, the j^{th} observation has the failure probability expressed in Eq. (10):

$$p_j = \frac{h(t_j, x_j)}{\sum^{n} h(t_j, x_j)} = \frac{h_0(t)\exp(\sum_{i=1}^{4} \beta_i X_{ij})}{\sum^{n} h_0(t)\exp(\sum_{i=1}^{4} \beta_i X_{ij})} . \tag{10}$$

For discrete failure occurrence, the likelihood function becomes the expression as Eq. (11) [18]:

$$L = \prod_{j=1}^{n} p_j = \prod_{j=1}^{n} (\frac{\exp(\sum_{i=1}^{4} \beta_i X_{ij})}{[\sum \exp(\sum_{i=1}^{4} \beta_i X_{ij})]^{d_j}})^{\delta^j} \tag{11}$$

where δ^j is 0 for truncated data and 1 for complete data, d_j is the number of failures at time t_j.

Based on variable selection results presented in Sec. 4.1, 4 potential risk factors that have high relevancy to the hazard rate. The MLE (SAS programming) estimates the weights of

Table 4. Analysis of maximum likelihood estimates.

Risk factor for group 1	DF	Parameter estimate	Chi-square	Pr > Chisq	Hazard ratio
Beginning engine hours	1	-1.8301	0	0.9993	0.1600
Average barometric pressure	1	-2.0386	0	0.9999	0.1300
Max coolant temp	1	1.5909	0	0.9998	4.5250
Max fuel temp	1	-0.6625	0	0.9999	0.5160
Risk factor for group 2	DF	Parameter estimate	Chi-square	Pr > Chisq	Hazard ratio
Beginning engine hours	1	-0.0047	0.2111	0.6459	0.995
Average barometric pressure	1	-0.1050	0.8101	0.3681	0.900
Max coolant temp	1	0.0584	0.3801	0.5375	1.060
Max fuel temp	1	0.0093	0.0295	0.8721	1.009
Risk factor for group 3	DF	Parameter estimate	Chi-square	Pr > Chisq	Hazard ratio
Beginning engine hours	1	0.0012	0.1175	0.7318	1.0010
Average barometric pressure	1	0.0549	2.1291	0.1445	1.0560
Max coolant temp	1	0.0585	1.5863	0.2079	1.0600
Max fuel temp	1	-0.0009	0.0010	0.9742	0.9990

the four proportional hazards for three groups as illustrated in Table 4.

The results in Table 4 summarize the estimated parameters of the Cox regression equations. Also, the hazard ratio is calculated. Since all P values are larger than 0.05, no statistically significant differences exist among risk factors.

Based on expression Eq. (5), the Cox regression equations for the three groups are presented in Eqs. (12)-(14), where X_1 denotes the beginning engine hours which is an indicator of engine age, X_2 represents the average barometric pressure, X_3 is the maximum coolant temperature, and X_4 is the maximum fuel temperature. Since the hazard rate can be derived from the Kaplan-Meier survival probability, it considers all related hazard risk factors. Hence from Eqs. (12)-(14), the baseline hazard rate is derived as $h_0(t)$.

The overall distribution of the baseline hazard rate at each occurrence time for the three groups is presented in Fig. 5.

The values of baseline hazard rate of first failures (diamonds) in Fig. 6 are higher at two sides of the timeline and lower in the middle area which resembles a bathtub curve. The baseline hazard rates of third failures (dots) are similar to the ones of the first failures. The baseline hazard rate for the third or higher failures (triangles) has an increasing trend in engine days.

$$\text{Group 1: } h(t,x) = h_0(t)\exp(-1.83015 * X_1 - 2.03816 * X_2 + 1.590955 * X_3 - 0.6625 * X_4) \tag{12}$$

$$\text{Group 2: } h(t,x) = h_0(t)\exp(-0.0047 * X_1 - 0.10501 * X_2 + 0.05843 * X_3 - 0.00937 * X_4) \tag{13}$$

$$\text{Group 3: } h(t,x) = h_0(t)\exp(-0.00118 * X_1 - 0.05847 * X_2 + 0.05849 * X_3 - 0.0008749 * X_4)\,. \tag{14}$$

5. Model fitting and risk measurement

5.1 Parametric model fitting

The basic additive Weibull model has been extended in the reliability engineering literature. Wang [7] proposed the finite interval distribution model for reliability engineering (FIRE) and a three parameter 3-FIRE model. Mudholkar [16] presented the Data-transformation model (DTM) and Jiang [5] developed the Finite support model (FSM) for performance assessment in finite intervals. The extension-type models faced challenges in modeling rapidly increasing failure rates on finite intervals [7]. For the infinite interval cases, Almalki's New modified Weibull (NMW) model [19] and Wang's additive Burr XII model (ABXII) [20] were constructed by embedding the additive Weibull or additive Burr models. Five typical bathtub models are selected for each dataset and performances of the five models are assessed.

(1) *FIRE Model*: The FIRE (Finite interval lifetime distribution for reliability engineering) model [7] has the reliability function Eq. (15) and hazard rate function Eq. (16).

$$R(t) = \exp[-\frac{\omega(\frac{t}{\theta})^{\alpha}}{(1-\frac{t}{\theta})^{\beta}}] \tag{15}$$

$$h(t) = \frac{\omega(\frac{t}{\theta})^{\alpha-1}}{\theta(1-\frac{t}{\theta})^{\beta+1}}[(\beta-\alpha)\frac{t}{\theta}+\alpha] \tag{16}$$

where $\alpha > 0, \beta > 0, \theta > 0$ and $\omega > 0$.

(2) *3-FIRE Model*: The 3-FIRE model [7], is a revised FIRE model to fit reliability data in an infinite interval. When the multiplication factor $\omega = 1$, the FIRE model is transformed into a 3-FIRE model [7]. Hence, the reliability function Eq. (17) and the hazard rate function Eq. (18) are obtained.

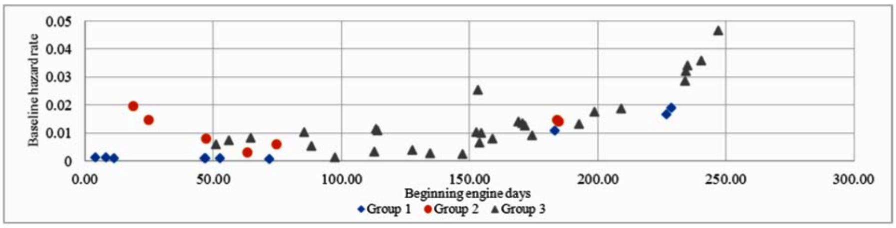


Fig. 6. The baseline hazard rate of three groups of failures.

$$R(t) = \exp[-\frac{(\frac{t}{\theta})^{\alpha}}{(1-\frac{t}{\theta})^{\beta}}] \quad (17)$$

$$h(t) = \frac{(\frac{t}{\theta})^{\alpha-1}}{\theta(1-\frac{t}{\theta})^{\beta+1}}[(\beta-\alpha)\frac{t}{\theta}+\alpha] \quad (18)$$

where α > 0, β > 0, θ > 0.

(3) *DTM Model*: Mudholkar [4] proposed the DTM (data-transformation) model to analyze the failure rate data. The DTM model is a three parameter model defined for a finite interval. It belongs to the exponential-Weibull family of parametric models. It follows the CDF expressed in Eq. (19), PDF in Eq. (20), and hazard rate function in Eq. (21).

$$F(t) = \exp[1-\exp(-(\frac{t}{\sigma}))^{\alpha}]^{\theta}, 0 \le t \le \infty \quad (19)$$

$$f(t) = (\frac{\alpha\theta}{\sigma})[1-\exp(-(\frac{t}{\sigma}))^{\alpha}]^{\theta-1} * \exp(-(\frac{t}{\sigma}))^{\alpha} * (\frac{t}{\sigma})^{\alpha-1} \quad (20)$$

$$h(t) = \frac{f(t)}{[1-F(t)]}. \quad (21)$$

(4) *ENH Model*: The exponential distribution type (ENH) model involves a three-parameter family of exponential-type distributions on the infinite interval [6]. It has the survival function Eq. (22) and the hazard rate function Eq. (23).

$$S(t) = 1-[1-\exp\{1-(1+\omega t)^{\alpha}\}]^{\beta}, t > 0 \quad (22)$$

$$h(t) = \alpha\beta\omega\frac{(1+\omega t)^{\alpha-1}\exp\{1-(1+\omega t)^{\alpha}\}[1-\exp\{1-(1+\omega t)^{\alpha}\}]^{\beta-1}}{1-[1-\exp\{1-(1+\omega t)^{\alpha}\}]^{\beta}}. \quad (23)$$

(5) *Additive Weibull Model*: The additive Weibull model combines two Weibull distributions. Among the two Weibull distributions, one has a decreasing failure rate and the other one has an increasing failure rate. The reliability function is presented in Eq. (24) and the hazard rate function in Eq. (25)

$$R(t) = \exp(-(at)^{b}-(ct)^{d}), t \ge 0 \quad (24)$$

$$h(t) = ab(at)^{b-1}+cd(ct)^{d-1}, t \ge 0. \quad (25)$$

For the parameter models above except the additive Weibull model, ω represents a multiplication factor of the failure rate function; θ is the scale parameter that refers to the maximum life of the truck engine exhaust valve; α and β are the shape parameters.

Step 1: Evaluate the interval and scale parameter θ using the formula $\theta = t_n+[(t_n-t_{n-k})/(n*k)]$ where n is the sample size; t_n is the time of the n^{th} failure; and k is the number of failures. For DTM model, the scale parameter is $1/\theta$.

Step 2: Estimate the shape parameters α and β plus the multiplication factor by maximum likelihood estimation.

Step 3: Compute AIC (Akaike information criterion), BIC (Bayesian information criterion), and the negative log-likelihood value -log$L$.

Application of the above 3-step procedure on the failure records has produced the results in Table 5.

For model selection for the three groups of failures, AIC and BIC evaluation matric are used. The AIC value of each model is derived from AIC = $2k$-2ln(L) and the BIC value is derived from $BIC = -2\ln(L)+k\ln(n)$. The best performing (data fitting) model is one with the minimum AIC or BIC value. Based on the results in Table 5, the 3-FIRE model performs best for the first failures. For the second failures, the Additive Weibull model performs best and the FIRE model performs best for third or more failures. The reliability and hazard rate curves are plotted in Figs. 7-9.

5.2 Value-at-risk measurement

Risk measurement models are constructed to evaluate exhaust valve failure hazard at different confidence levels. Value-at-risk (VaR) is a widely used measurement model to assess risks at extreme conditions. Initially proposed to measure the exposure of portfolio risks in the finance industry, VaR reflects what could cause default over a certain period of time [21]. Hendricks [22] presented three approaches to compute VaR values: Equally weighted variance-covariance approach; exponentially weighted variance-covariance approach; and the historical simulation approach. Due to the adequate number of historical failures, the third approach is employed here with the VaR defined in Eq. (26).

$$VaR_p(X) = \pi_p = F_X^{-1}(p) \quad (26)$$

Table 5. Model fitting for three groups of failures.

Model for group 1	Scale parameter	MLE for major parameters	AIC	BIC	-logL
FIRE	228.999	$\alpha = 0.666$ $\beta = 0.113$ $\omega = 1.469$	103.138	104.046	48.569
3-FIRE	228.999	$\alpha = 0.589$ $\beta = 0.155$	101.815	102.420	48.907
DTM	0.004	$\alpha = 0.265$ $\sigma = 0.649$	107.226	107.831	51.613
Additive weibull		$a = 0.004$ $b = 4.6$ $c = 5.2$ $d = 0.1$	113.300	113.600	52.63
ENH		$\alpha = 1.367$ $\beta = 0.771$ $\lambda = 0.006$	115.288	116.1957	54.644
Model for group 2	Scale parameter	MLE for major parameters	AIC	BIC	-logL
FIRE	185.089	$\alpha = 1.041$ $\beta = 0.016$ $\omega = 1.383$	80.333	80.171	37.167
3-FIRE	185.089	$\alpha = 0.941$ $\beta = 0.138$	78.697	78.589	37.349
DTM	0.005	$\alpha = 0.331$ $\sigma = 0.659$	71.324	71.929	33.662
Additive weibull		$a = 0.005$ $b = 5$ $c = 5$ $d = 0.09$	68.486	69.064	32.229
ENH		$\alpha = 0.3773$ $\beta = 3.176$ $\lambda = 0.128$	92.5026	92.3403	43.251
Model for group 3	Scale parameter	MLE for major parameters	AIC	BIC	-logL
FIRE	247.3876	$\alpha = 2.544$ $\beta = 0.115$ $\omega = 2.007$	313.000	317.103	153.515
3-FIRE	247.3876	$\alpha = 1.940$ $\beta = 0.236$	315.021	317.755	155.51
DTM	0.0040	$\alpha = 2.544$ $\sigma = 0.462$	341.323	344.057	168.66
Additive weibull		$a = 0.005$ $b = 5.2$ $c = 5.2$ $d = 0.07$	314.087	316.821	155.04
ENH		$\alpha = 3.01$ $\beta = 4.334$ $\lambda = 0.003$	322.124	326.225	158.06

where X denotes reliability which is a time-dependent variable; and P is the confidence level.

For the FIRE, the 3-FIRE, and the Additive Weibull model, the formula for VaR is not explicit, rather it is iteratively computed. For the three best performing models, the time at which a failure is likely to occur is computed at different confidence levels (see Table 6).

5.3 Model validation

A set of experiments are conducted to validate the correctness of the model output. With the establishment of the parametric reliability models for three different groups of failures, the main objective is to validate the statistical patterns of these failures.

In a typical statistical analysis, there are three major validation methods known as apparent validation, internal validation and external validation. In our case, we still have three engine datasets that are reserved for validation. The external validation method is more convenient to be applied with new datasets available. Hence, it is selected to validate the reliability for the three groups of failures.

Table 6. The value-at-risk estimates for three groups of failures.

P	0.5	0.8	0.9	0.95	0.99
Group 1	104.676	219.87	227.96	228.71	228.99
Group 2	0	39.26	172.1	204.5	243.6
Group 3	166.288	233.12	269.24	299.36	356.1

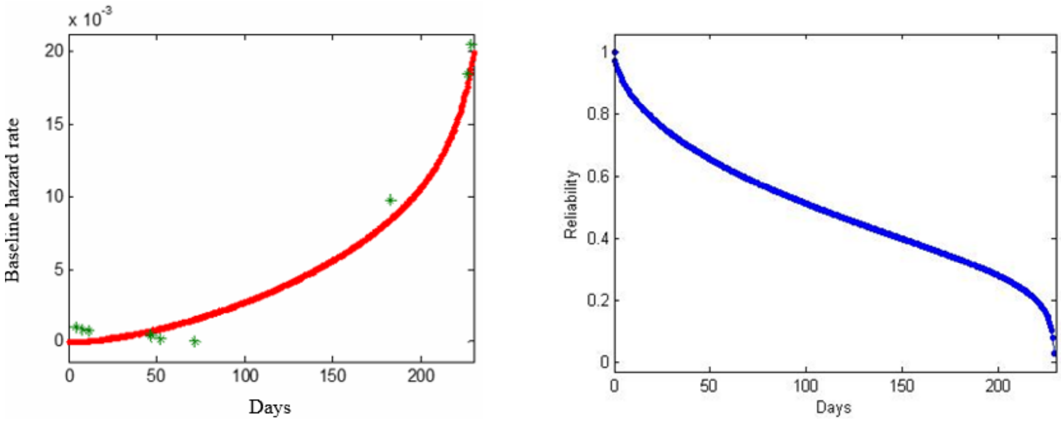


Fig. 7. Baseline hazard rate and reliability plot for group 1.

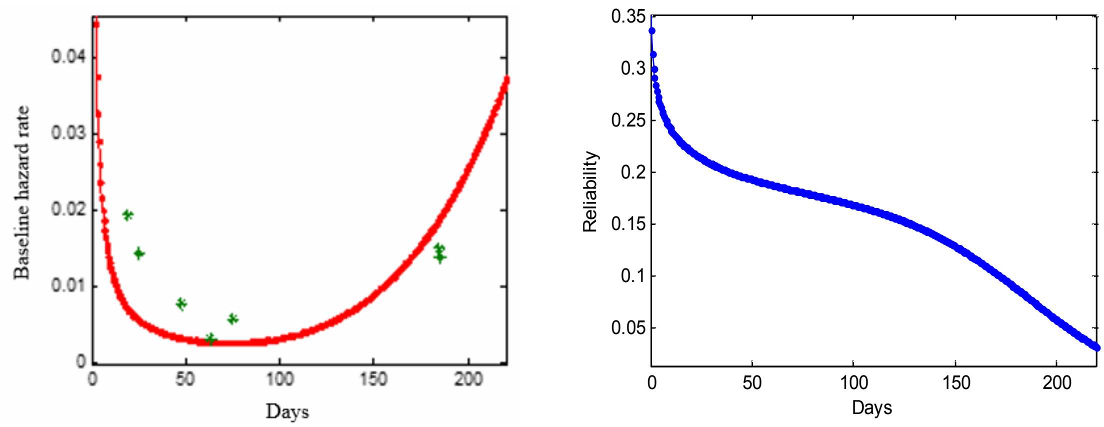


Fig. 8. Baseline hazard rate and reliability plot for group 2.

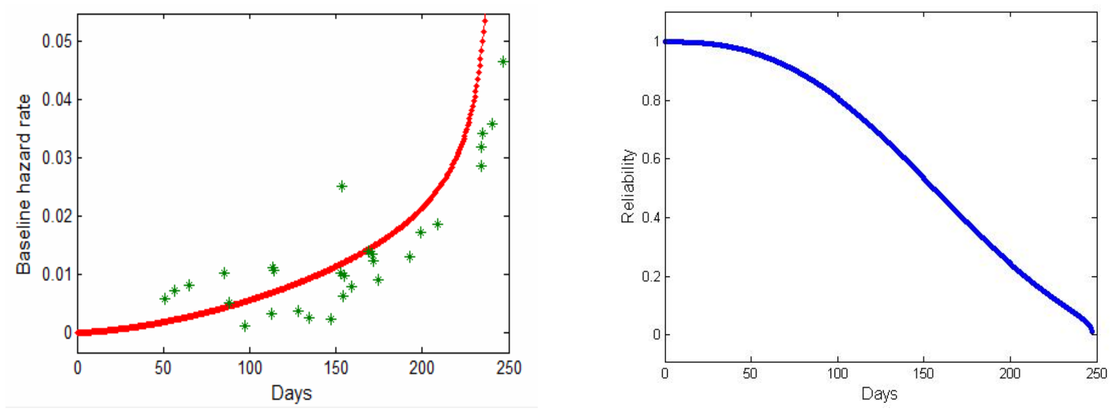


Fig. 9. Baseline hazard rate and reliability plot for group 3. The proposed FIRE and 3-FIRE models can be considered as a generalized Weibull model for fitting the bathtub curve. The FIRE model is equivalent to the Weibull model for $\beta = 0$. The FIRE model is equivalent to 3-FIRE model for $\omega = 0$. The FIRE model is equivalent to DTM model for $\alpha = \beta$ [15]. In order to evaluate performance of the proposed model, the following 3-step procedure is proposed [7].

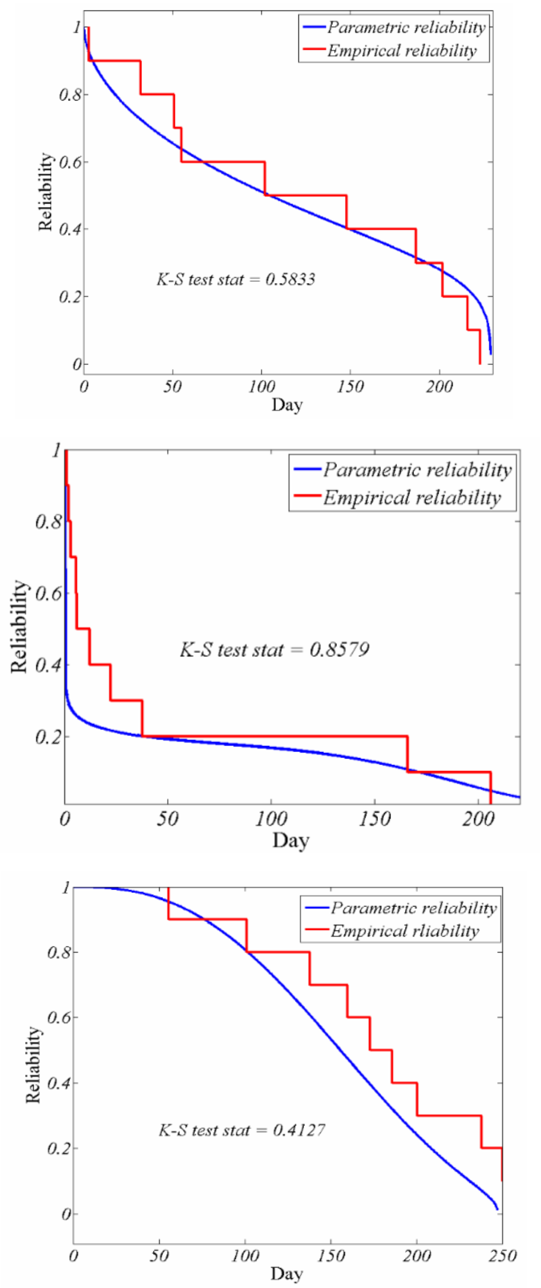


Fig. 10. Model validation results.

Due to the limited samples, a bootstrap method is applied to simulate the distribution of exhaust valve failures. In Fig. 10, the smoothed CDF of reliability curves of three groups of failures are plotted along with the empirical CDF derived based on real failures. A two sample Kolmogorov-Smirnov (K-S) test is conducted to compare the empirical distribution and derived parametric distribution. The K-S test statistic are computed and illustrated in each graph. At the level $\alpha = 0.05$, the threshold to reject the null hypothesis is 0.8601. Examination of the K-S test results indicates that three derived parametric models are all adequate in reproducing the actual distribution of exhaust valve failures.

6. Conclusion

This paper has focused on predicting exhaust valve failures of truck engines. The dataset used in the analysis was incomplete. Neural network algorithm was applied for imputing the missing values. The failures were categorized into three groups: the first failures, the second failures, and the third or more failures. The survival analysis was deployed to estimate the probability distribution of the hazard rate. The Kaplan-Meier survival curve was constructed and survival probabilities and discrete hazard rates were computed.

A voting-based wrapper's variable selection method was deployed to reduce the number of variables. Four variables, Beginning engine hours, Average barometric pressure, Maximum coolant temperature, and Maximum fuel temperature were determined as most relevant to the time-dependent hazard rate of exhaust valve failures. The Cox partial likelihood regression model was used to estimate the baseline hazard rate. For each group, a Cox regression equation was established and the hazard ratio was computed. Based on the Cox equations, the time-dependent baseline hazard rate was derived.

Bathtub curves, commonly used in reliability engineering, were applied to analyze the time-dependent baseline hazard rate of each of the three groups. Modified parametric models and classical models such as the Finite interval distribution model for reliability engineering (FIRE), 3-parameter Finite interval distribution model for reliability engineering (3-FIRE), Data transformation model (DTM), additive Weibull, and exponential distribution type model (ENH) were selected to fit the baseline hazard data. The maximum likelihood estimates were derived for all parameters. The values of the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were computed to evaluate model performance. The FIRE, 3-FIRE, and the additive Weibull models were shown to be the best match for the failure data. The Value-at-risk (VaR) was calculated for each selected parametric model to predict failures of exhaust valves at different confidence levels. The three parametric models were validated by the two sample Kolmogorov-Smirnov (K-S) tests. All K-S test stats were under the threshold to reject null hypothesis at the 0.05 significance level. Hence, the parametric models evaluated the distribution of exhaust valve failures accurately.

Reference

[1] E. L. Kaplan and P. Meier, Nonparametric-Estimation from Incomplete Observations, *Journal of the American Statistical Association*, 53 (282) (1958) 457-481.

[2] D. R. Cox, Regression models and life-tables, Journal of the Royal Statistical Society, *Series B-Statistical Methodology*, 34 (2) (1972) 187-220.

[3] L. H. Crow, Tracking reliability growth, *Proceedings Annual Reliability and Maintainability Symposium*, 1 (28) (1975) 438-443.

[4] G. S. Mudholkar and D. K. Srivastava, Exponentiated

weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability*, 42 (2) (1993) 299-302.
[5] R. Jiang, A new bathtub curve model with a finite support, *Reliability Engineering & System Safety*, 119 (2013) 44-51.
[6] A. J. Lemonte, A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function, *Computational Statistics & Data Analysis*, 62 (2013) 149-170.
[7] X. H. Wang, C. Yu and Y. X. Li, A new finite interval lifetime distribution model for fitting bathtub-shaped failure rate curve, *Mathematical Problems in Engineering* (2015).
[8] S. E. Haupt, A. Pasini and C. Marzban, *Artificial Intelligence Methods in the Environmental Sciences*, Springer, New York, USA (2009) 424 .
[9] J. C. Houwelingen and H. Putter, Dynamic prediction in clinical survival analysis, *Monographs on Statistics and Applied Probability,* 6 (2012) 234 .
[10] P. J. Kelly and L. L. Y. Lim, Survival analysis for recurrent event data: An application to childhood infectious diseases, *Statistics in Medicine*, 19 (1) (2000) 13-33.
[11] W. J. Westlake, Symmetrical confidence-intervals for bioequivalence trials, *Biometrics*, 32 (4) (1976) 741-744.
[12] M. Greenwood, *A Report of the Natural Duration of Cancer,* H. M. Stationery. London, UK (1926).
[13] T. Ouyang, A. Kusiak and Y. He, Modeling wind-turbine power curve: A data partitioning and mining approach, *Renewable Energy*, 102 (2017) 1-8.
[14] D. R. Cox, Partial Likelihood, *Biometrika*, 62 (2) (1975) 269-276.
[15] G. S. Mudholkar and K. O. Asubonteng, Data-transformation approach to lifetimes data analysis: An overview, *Journal of Statistical Planning and Inference*, 140 (10) (2010) 2904-2917.
[16] P. Peduzzi, J. Concato and A. R. Feinstein, Importance of events per independent variable in proportional hazards analysis 1. background, goals, and general strategy, *Journal of Clinical Epidemiology*, 48 (12) (1995) 1495-1501.
[17] D. R. Cox and D. Oakes, Analysis of survival data, *Monographs on Statistics and Applied Probability,* Chapman and Hall, New York, USA (1984).
[18] D. G. Kleinbaum, Survival analysis: A self-learning text, *Statistics in the Health Sciences,* 12 (1996) 324.
[19] S. J. Almalki and J. S. Yuan, A new modified weibull distribution, *Reliability Engineering & System Safety*, 111 (2013) 164-170.
[20] F. K. Wang, A new model with bathtub-shaped failure rate using an additive burr XII distribution, *Reliability Engineering & System Safety*, 70 (3) (2000) 305-312.
[21] T. J. Linsmeier, Value at risk, *Financial Analysts Journal*, 56 (2) (2000).
[22] D. Hendricks, Evaluation of value-at-risk models using historical data, *Economic Policy Review Federal Reserve Bank of New York*, 27 (1) (1997) 39-67.
[23] T. Ouyang, A. Kusiak and Y. He, Predictive model of yaw error in a wind turbine, *Energy*, 123 (2017) 119-130.
[24] J. H. Kim, T. S. Kim and S. J. Moon, Development of a program for trasient behavior simulation of heavy-duty gas turbines, *Journal of Mechanical Science and Technology*, 30 (12) (2016) 5817-5828.
[25] L. Borello, G. Villero and M. D. Vedova, Flap failure and aircraft controllability: Developments in asymmetry monitoring techniques, *Journal of Mechanical Science and Technology*, 28 (11) (2014) 4593-4603.
[26] J. J. Han, H. W. Kim, Y. J. Kim, J. S. Kim, Y. J. Oh and H. B. Park, Failure assessment diagram analysis of high density polyethylene pipes, *Journal of Mechanical Science and Technology*, 28 (12) (2014) 4929-4938.
[27] G. M. Lim, D. M. Bae and J. H. Kim, Fault diagnosis of rorating machine by thermography method on support vector machine, *Journal of Mechanical Science and Technology,* 28 (8) (2014) 2947-2952.

Yusen He received his M.Sc. degree in Statistics from the University of Iowa, United States in 2013. He is currently persuing his Ph.D. degree in Industrial Engineering from the University of Iowa. His main research interests include application of data-mining on machinery failures and energy power systems.