Reliability-based regression model for complex systems considering environmental uncertainties

Amin Moniri-Morad[a], Mohammad Pourgol-Mohammad[a*], Hamid Aghababaei[a], Javad Sattarvand[b]

[a] Sahand University of Technology, Tabriz, Iran
[b] University of Nevada Reno, Reno, USA

Abstract: Raw material transportation equipment is considered as a key factor to meet production planning goals in mining industry. Mine dynamic circumstances lead to variations in equipment reliability, which has a direct effect on the attained production capacity during operating intervals. Prevalent reliability evaluation techniques are concentrated only on analyzing failure time data based on fully known probabilistic distribution functions. Nevertheless, time between failures may be influenced by various operational and environmental conditions. In this study, Cox proportional hazard regression model is proposed to estimate the reliability of complex systems considering significant covariates in dynamic situations. In this method, baseline hazard function is parametrically modeled for simulation and extrapolation purposes. A real case study was completed to investigate the impact of different scenarios on the mining equipment reliability. The collected covariates data consist of severity of failure incidents, type of failed sub-systems, environmental temperature and climate conditions. The research results revealed that identification and quantification of the covariates are taken into consideration as a main aspect in prediction of failure rate of mining equipment. This framework results in strong insights into critical sub-systems and environmental uncertainties, leading to high risk levels on the complex systems.

Keywords: mining equipment, reliability, failure rate, Cox proportional hazard regression model, dynamic environment, production plan

1. INTRODUCTION

Haulage equipment is considered as a critical element to meet the production planning targets. However, operational variations and dynamic environment considerably affect the haulage equipment reliability and thus the attained production capacity is declined. In mining industry, these kinds of situations can be classified into two main groups including operational and environmental factors. The operational factors refer to various working bench locations, rock types, conditions of haulage roads, skill of operators and maintenance crews, as well as different maintenance strategies and so on. The environmental factors are consisted of harsh climate, temperature, presence or absence of wind, dust, snow and rain fall and so on. As a result, the equipment reliability, maintainability and availability should be addressed by identifying and analyzing these factors.

The conventional methods are associated with fitting failure time data set to a suitable distribution function such as exponential, Weibull and Power law process [1-6]. Many researches have been accomplished for evaluation of mining equipment reliability using the standard parametric models. Some of the studies were comprised of reliability evaluation of crushing plant [7], dragline reliability estimation [8], reliability evaluation of off-highway trucks [9], modeling reliability block diagram for horizontal drilling equipment [10]. Although, the standard parametric models have the capability of simulating and extrapolation of the complex system reliability, nonetheless, the model is dependent only on failure time data and it doesn't the ability to quantify and address the impact of covariates on the system reliability and performance. As an alternative, Cox proportional hazard model (PHM) can be considered to predict the effect of operational and environmental covariates in engineering regions. The Cox proportional hazard regression model was

* Corresponding Author: pourgolmohammad@sut.ac.ir

developed by Cox [11] in 1972. This model is expressed by baseline hazard rate, and a multiplicative term comprising set of covariates. It is employed when the covariates have multiplicative influence on the observed hazard rate.

There are various studies about modeling equipment reliability and performance evaluation considering significant covariates. Kumar [12] utilized the proportional hazard model to consider the most influential covariates dealing with different maintenance strategies (perfect, imperfect and minimal repairs) for electric load-haul-dump machines. The author computed hazard ratios with respect to the maintenance strategies. Ghodrati et al. [13] employed the proportional hazard model to predict the quantity of required spare parts considering the environmental factors for hydraulic jacks on a LHD machine. Gasmi et al. [14] applied the proportional hazard model to investigate the effect of changes in failure intensity due to switching of operating circumstances. Barabadi et al. [15] used a proportional repair model (PRM) [16] to analyze the effect of climate conditions on the quality of maintenance activities in harsh environment.

This paper is aimed to develop an integrated methodology based on proportional hazard model for analyzing off-highway trucks reliability and performance under dynamic environment of open pit mines. In this case, various covariates have been developed and examined to recognize the most significant covariates leading to variations in truck reliability and performance under heterogeneous environment. The covariates are composed of truck sub-systems, severity of failure incidents and climate conditions, which are investigated as set of covariates and Weibull distribution function is employed for modeling baseline hazard rate function. The final results provides a valuable insight into the truck crucial sub-systems and type of truck failure incidents. Accordingly, the complex system reliability and performance are characterized more efficiently for developing maintenance plan strategies.

The paper is organized as follows. The research theory is described in section 2. Section 3 presents a case study to address the off-highway trucks reliability and performance with respect to the collected environmental and operational factors at Sungun copper mine. In section 4, conclusion and remarks are characterized for the proposed reliability evaluation method under dynamic environment.

2. RESEARCH THEORY AND FRAMEWORK

The research methodology is presented in Figure 1. According to the figure, the study is initially started by data collection process. The data collection is dealt with complex system familiarization and specifying boundaries for reliability evaluation (step 1). The required data is gathered based on operational and environmental evidences. The operational document is associated with recognizing equipment sub-systems, collection of failure and maintenance time data. The environmental document is gathered based on geology and climate conditions. The step 2 is coincided with determination of set of important covariates for the regression model. To do this, it is essential to follow a covariate selection structure. In the first stage, the potential covariates are divided into qualitative and quantitative parameters. In the second stage, the covariates are numerically quantified. The qualitative covariates are addressed by binary or categorical values. There are some examples for the qualitative covariates including presence or absence of snow or icing. The quantitative covariates are characterized by continuous values (e.g. system age and temperature). In the third stage, the outliers are found and eliminated from the collected data set. In the fourth stage, time dependency of the covariates are tested. From time dependency point of view, a proper hazard model is selected for the analysis process (i.e. PHM, Stratified or extended PHM) [17, 18]. The fifth stage is related to elimination of unimportant covariates, leading to some bias and errors on the regression model. Finally, there are various approaches for covariate selection comprising polynomials models [19], Akaike Information Criterion (AIC) [20], Bayes Information Criterion (BIC) [21], stepwise method (Backward Elimination (BE), Forward Selection (FS) and Stepwise Regression (SR)) [22].

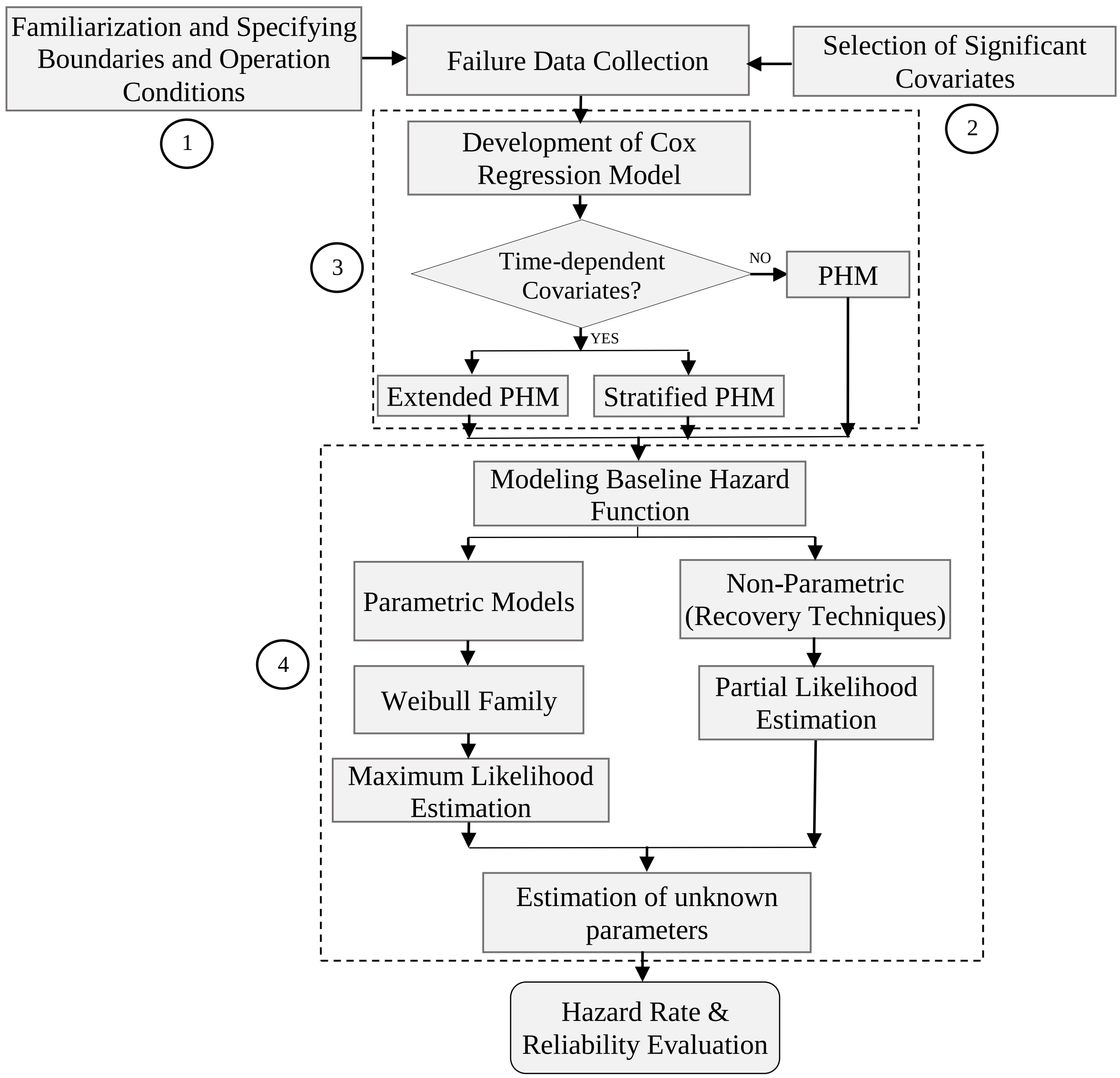

Figure 1: Step-By-Step Study Flowchart for Modeling Cox Regression Model

According to the analysis process, the step 3 is associated with employing Cox regression model to evaluate the complex system reliability. In general, Cox model contains two key elements encompassing baseline hazard function, dependent only on failure time data and an exponential term for set of covariates. It is introduced as follows [23]:

$$h(t|Z) = h_0(t)\, exp(\beta^T Z) \tag{1}$$

where, $h(t|z)$ denotes the observed hazard function, Z is a $p \times 1$ vector containing set of covariates, β^T is a $1 \times p$ vector of regression coefficients, entailing the effect of covariates, $h_0(t)$ presents the baseline hazard function, occurring when the covariates have no influence on the failure rate ($Z = 0$ or $\exp(\beta^T Z) = 1$). Proportionality of hazard is considered as a key assumption for Cox proportional hazard model. According to this assumption, the hazard ratio remains constant for any two observations over time. If the proportionality assumption is violated, non-proportional hazard models are utilized to compute the complex system failure rate and reliability considering time-dependent covariates (e.g. extended and stratified PHM

[17, 18]). Otherwise, the PHM is employed for analyzing process. The proportionality assumption is derived by:

$$HR = \frac{h_1(t|Z_1)}{h_2(t|Z_2)} = \frac{h_0(t)\, exp(\beta^T Z_1)}{h_0(t)\, exp(\beta^T Z_2)} = exp[\beta^T (Z_1 - Z_2)] \tag{2}$$

where, Z_1 and Z_2 are introduced as sets of covariates for two different observations, the term $HR = exp[\beta^T (Z_1 - Z_2)]$ is denoted by hazard ratio and gets a constant value over time.

The step 4 is involved in modeling and computation of baseline hazard function. As shown in the flowchart, the baseline hazard function is derived by non-parametric and parametric models. In the case of non-parametric model, value of the baseline hazard rate is obtained through recovery techniques [24]. The outcomes of the Cox regression model (e.g. hazard ratio and regression coefficients) are estimated without realizing the functional form of the baseline hazard rate function. The regression coefficients are estimated by maximizing partial likelihood function [25]. The obstacle of non-parametric model is concerned with inaccessibility of reliability function for simulation and extrapolation purposes. According to the flowchart, the parametric baseline hazard function is characterized by standard parametric distribution functions. In this case, the baseline hazard function is modeled through some of the probabilistic distribution functions. There are few probabilistic distributions (i.e. Weibull and exponential) to fulfill the proportionality assumption for Cox proportional hazard model [26]. According to the definition, combination of multiplicative term into the Weibull and exponential baseline hazard functions lead to same shape for the observed hazard function. The Weibull baseline hazard function is expressed according to Equation (3) [13].

$$h_0(t) = \lambda_0 . \gamma_0 . t^{\gamma_0 - 1} \tag{3}$$

Assumed that $w = exp(\beta^T Z)$ and $g = \frac{1}{w^{1/\gamma_0}}$. Then,

$$h(t|Z) = \frac{\gamma_0}{(\alpha_0 \times g)^{\gamma_0}} t^{\gamma_0 - 1} \tag{4}$$

Finally, the observed hazard function ($h(t|Z)$) is derived as a form of Weibull baseline hazard function:

$$h(t|Z) = \frac{\gamma_0}{(\alpha^*)^{\gamma_0}} t^{\gamma_0 - 1} = \lambda^* . \gamma^* . t^{\gamma^* - 1} \tag{5}$$

where, λ^* is equal to $\frac{\gamma^*}{(\alpha^*)^{\gamma^*}}$; and $\alpha^* = \alpha_0 (exp(\beta^T Z))^{(-\frac{1}{\gamma_0})}$, $\gamma^* = \gamma_0$ are scale and shape parameters of the observed hazard function, respectively; α_0, γ_0 are the shape and scale parameters for the Weibull baseline hazard function, respectively.

The Equation (5) demonstrates that the covariates effect changes the scale parameter of the Weibull baseline hazard function, whereas the shape parameter remains constant for the Weibull baseline hazard function. Therefore, the Weibull family fulfills the proportionality assumption. Finally, maximum likelihood estimation is utilized to compute the unknown regression coefficients and Weibull parameters for the parametric Cox regression model.

The step by step analysis process leads to establish a comprehensive formulation considering significant covariates for reliability evaluation in dynamic and heterogeneous environment.

3. CASE STUDY

The off-highway trucks play a vital role in the planned production capacity at Sungun copper mine. The mine dynamic environment causes to major variations in failure rate and performance of the trucks. An accurate estimation of the truck reliability is viable to develop a realistic production plan for the mining operation.

The required database was collected based on failure time data and covariate information including operational and environmental parameters. In this study, one type of covariate was concerned with five main sub-systems (engine, transmission, gearbox, hydraulics, body and chassis). Another type of covariate was associated with severity of failure incidents (mild or severe failure incidents). This kind of covariate plays an important role in repair times and maintenance quality. As a matter of fact, mild failure incidents are frequent, whereas they have short repair times and low repair costs. In contrast, severe failure incidents occasionally occur on the truck system, whereas they spend long repair times and high repair costs, leading to considerable mining operation interruption. Furthermore, the environmental conditions were considered as the environmental covariates such as climate conditions (i.e. temperature, rain fall). The gathered failure data set is expressed in Table 1. In Table 2, the binary codes are introduced for each covariate.

Table 1: The Historical Data Set with Respect to Covariates

No. Failures	TBFs (Hr)	Engine Failure Incidents	Transmission Failure Incidents	Hydraulics Failure Incidents	Body Failure Incidents	Gearbox Failure Incidents	Climate conditions	Severity	Temperature (◦c)
1	50	0	0	0	1	0	1	0	-0.1
2	73	1	1	0	1	0	1	1	7
3	45	1	0	0	0	0	1	0	7
4	17	1	1	1	0	0	1	0	7
5	70	0	1	0	0	0	1	0	3.3
6	30	0	0	1	0	0	1	1	3.3
7	18	0	0	1	0	0	1	1	3.3
8	37	1	0	0	1	0	0	0	9.1
9	185	1	0	0	0	0	1	0	7.6
10	30	1	0	0	0	0	0	1	11.6
11	18	0	1	0	0	0	0	1	11.6
12	25	0	1	0	0	0	0	1	11.6
13	92	1	0	0	0	0	1	1	12.6
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

As shown in the Table 1, time between failures (TBFs), value of the continuous covariates, as well as the status of the binary covariates are characterized for each failure event. The table rows demonstrate time between failures of the truck with respect to the status of the covariates. The number of 103 failure incidents were collected for one year of truck operation. The status of each covariate was also gathered to start the reliability analysis process. As an example, the first failure occurred after 50 operation hour. In addition, the failure event was due to failure in body & chassis sub-system. The failure incident was mild event, the weather was rainy during the day (bad climate condition) and the average temperature was -0.1 for a week.

Table 2: Definition of Binary Codes for the Collected Covariates

Severity		Engine		Transmission		Hydraulics		Body		Gearbox		Climate condition	
Severe	Mild	F	S	F	S	F	S	F	S	F	S	Bad	Good
1	0	1	0	1	0	1	0	1	0	1	0	1	0

F: Failure, S: Success

3.1. Covariate Analysis Process

The time dependency (time-dependent or time-independent) of the covariates must be examined to make decision regarding an appropriate Cox regression model for evaluation of system reliability. Graphical technique was completed to test the proportionality assumption. The plot of *log[-log(R(t))]* against *log(time)* was drawn for each covariate. As an example, Figure 2 illustrates the proportionality test for the engine covariate. If the curve is nearly parallel, the proportionality assumption is fulfilled and the PHM is employed for modeling hazard rate. Otherwise, the stratified or extended PHM are considered for modeling of the hazard rate function. As shown in the graph, the engine covariate is taken into account as a time-independent covariate.

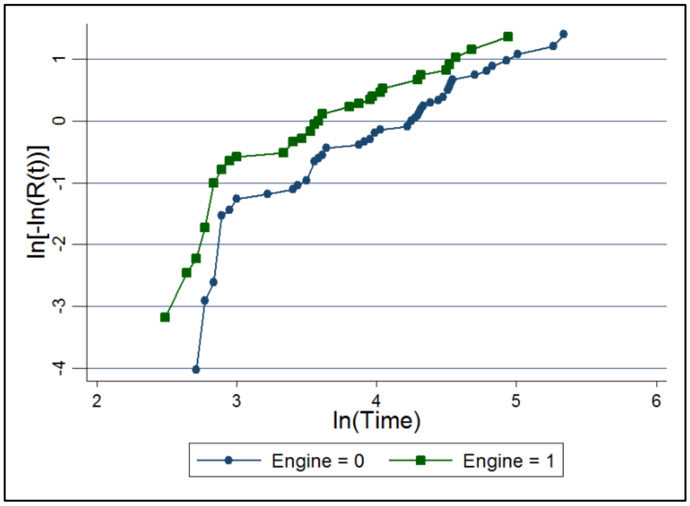

Figure 2: Evaluation of Proportionality Hazard Assumption for Engine Covariate

The next step is concentrated on employing stepwise method (i.e. BE, FS and SR) to choose a set of important covariates for Cox regression model (Table 3Table 4). To select the set of significant covariates, the probability value (p-value) is utilized as a statistical hypothesis testing. In this approach, the analysis procedure is implemented through specifying a null hypothesis (H) and significance level (α_{crit}). The null hypothesis is rejected where the computed p-value is less than the pre-defined significance level. As a result, the BE method is started by full model consisting of all the potential covariates. Afterwards, one of the covariates is eliminated with highest p-value greater than the significance level (α_{crit}). Then, the regression model is refitted and the p-value of the covariates are assessed to remove another candidate covariate. The stopping criterion happens when all the p-values are less than α_{crit}. The FS technique is accomplished in reverse direction of the BE. First, there is no covariate in the regression model. Then, the covariates are subsequently imposed on the regression model with respect to their p-values (the covariate with the greatest improvement in the fitted model is added at each step). The SR is a combination of BE and FS methods. This process is started by FS technique. After adding second covariate, this method examines if an existing covariate can be removed from the model without a significant effect on the fitted model. In the Table 3, the BE technique was considered to recognize the final set of covariates for Cox regression model.

Table 3: The Results of Backward Elimination Method for Covariate Selection Procedure

Covariates	p-values	Status in Cox Model	Estimated Cox Regression Coefficients
Engine	0.000	Keep	0.834
Transmission	0.002	Keep	0.703
Hydraulics	0.000	Keep	1.400
Gearbox	0.000	Keep	1.058
Body and Chassis	0.059	Keep	-0.385
Severity	0.049	Keep	0.401
Climate Conditions	0.92	Remove	N/A
Temperature	0.260	Remove	N/A

In this table, the covariate selection was carried out based on the significance level of 10 percent. Based on the Table 3, the major covariates were identified as engine, transmission, severity and body and chassis. The eliminated covariates were comprised of temperature, hydraulics, climate conditions.

3.2. Fitting Parametric PHM

The hazard rate function has been formulated by employing Weibull distribution function as the baseline hazard rate function for Cox regression model. The Weibull distribution was used because of its ability to describe the service lifetime of mechanical elements. The maximum likelihood estimation was applied to estimate the unknown parameters for the Cox regression model. The unknown parameters were composed of set of covariates and the Weibull distribution parameters. The log-likelihood function is formulated for PHM with respect to Weibull baseline hazard function as Equation (6). In this equation, it is assumed that d_i =1, if each event is a failure incident, otherwise d_i =0.

$$\ln(L_i) = d_i\left[\ln(\gamma_i) - \ln(\alpha_i) + (\gamma_i - 1)\left(ln(t_i) - ln(\alpha_i)\right) + \beta Z\right] - \left(\frac{t_i}{\alpha_i}\right)^{\gamma_i} \exp(\beta Z) \tag{6}$$

Table 4 indicates the results of estimation of parameters for parametric Cox regression proportional hazard model under dynamic environment.

Table 4: Estimation of Unknown Parameters for Parametric Cox Proportional Hazard Model

Parameters	Estimated Unknown Parameters	p-values
Engine	0.795	0.000
Transmission	0.614	0.006
Body & Chassis	-0.37	0.062
Severity	0.406	0.048
Hydraulics	1.22	0.000
Gearbox	0.941	0.000
λ_0	0.0004	0.000
γ	1.66	0.000

In this table, there are six significant covariates encompassing engine, transmission, chassis & body, hydraulics, gearbox, severity. In addition, there are two parameters for Weibull distribution function. These unknown parameters were estimated by using maximum likelihood estimation. The p-values were employed to characterize the statistical significance level for the estimated parameters in the model. The significance level was defined at the level of 10%. The output results reveal that the most important covariate is associated with hydraulics sub-system so that it has the greatest effect on the baseline hazard rate function. The lowest effect is related into the body & chassis sub-system. The estimated coefficients were 1.22 and -0.37 for hydraulics and body & chassis sub-systems, respectively. This process gives a good guideline to investigate the effect of failure of each sub-system, environmental and operational conditions on the of hazard rate value. Therefore, a comprehensive maintenance strategy can be developed based on the outputs of Cox regression model.

3.3. Estimation Results

This section is dealt with prediction of reliability and hazard rate values for the developed Cox regression model. These graphs make a proper validation for the proposed Cox PHM. The reliability function is introduced as probability of implementation of component mission without occurrence of failure incident at a certain time interval. Figure 3 indicates the estimated reliability curve by using Cox regression model.

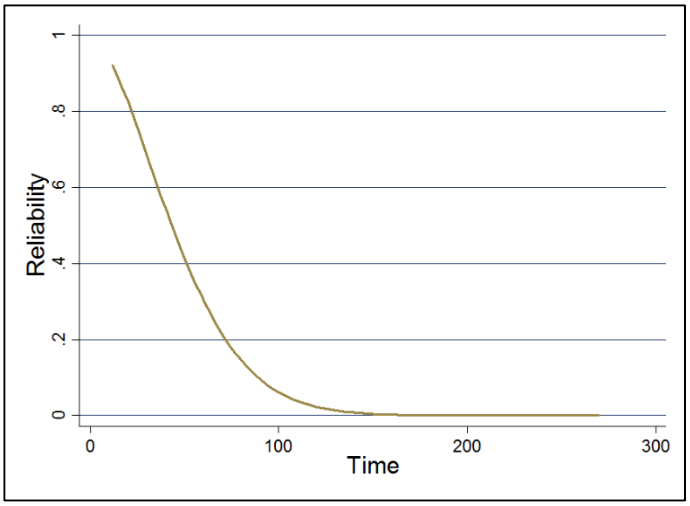

Figure 3: Reliability Graph for Off-Highway Truck by using Cox Regression Proportional Hazard Model

According to the Figure 3, the truck reliability approaches to 0.8, 0.6 and 0.5, after nearly 22, 38 and 43 operation hours, respectively. In addition, the truck failure probability goes to one, after approximately 140 operation hours. Figure 4 demonstrates the hazard rate graph for the established Cox regression model. The graph illustrates that the truck failure rate is monotonically increasing during the operating intervals. Therefore, the truck age is in wear-out period, which requires holistic maintenance strategies like preventive maintenance, condition-based maintenance to improve the status of truck circumstance.

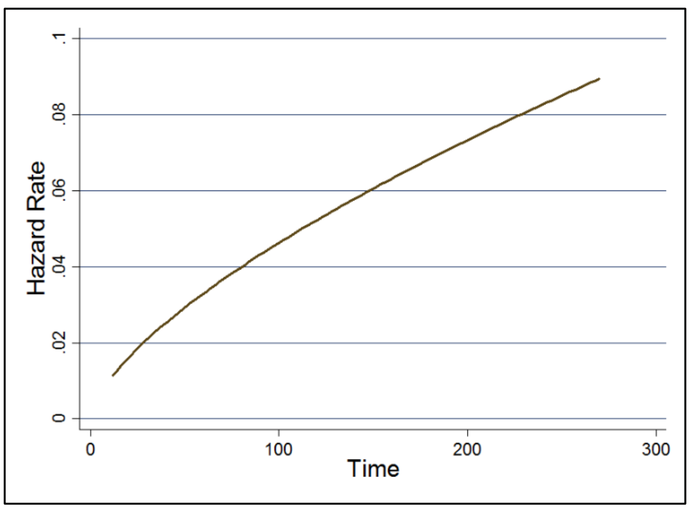

Figure 4: Hazard Rate Description for Off-Highway Truck

4. CONCLUSION

This research proposes an integrated methodology to analyze the complex system reliability and performance in dynamic conditions. The standard parametric model does not have the capability to consider various covariates in heterogonous environments. Cox regression proportional hazard model was developed to accurately compute the impact of different covariates on the complex mining systems. A case study was developed to validate the outcomes of the proposed Cox regression model. In this case, the off-highway truck system was considered for the reliability evaluation process. The selected covariates were associated with truck failure sub-systems, severity of failure events and climate conditions. The Weibull distribution function was considered to characterize the service lifetime of off-highway trucks for baseline hazard rate function. Maximum likelihood estimation was employed to compute the regression coefficients for Cox regression proportional hazard model. The results showed that Cox regression model has the ability to efficiently analyze the influence of different significant covariates on the truck hazard rate and reliability values. In addition, the truck sub-systems were ranked to be able to develop an appropriate maintenance policy for the truck system.

Acknowledgment

Authors would like to acknowledge the support of the maintenance management department at Sungun Copper mine.

References

[1] Modarres M, Kaminskiy MP, Krivtsov V. *"Reliability Engineering and Risk Analysis: A Practical Guide"*. CRC Press; 2016.

[2] Vagenas N, Runciman N, Clément SR. "*A methodology for maintenance analysis of mining equipment. International Journal of Surface Mining, Reclamation and Environment"*,11:33-40,(1997).

[3] Morad AM, Pourgol-Mohammad M, Sattarvand J. "*Application of reliability-centered maintenance for productivity improvement of open pit mining equipment: Case study of Sungun Copper Mine. Journal of Central South University"*,21:2372-82,(2014).

[4] Asghari J, Mohammad MP, Oskouyi FS. *"Improving Dynamic Fault Tree Method for Complex System Reliability Analysis: Case Study of a Wind Turbine"*. ASME 2015 International Mechanical Engineering Congress and Exposition: American Society of Mechanical Engineers. p. V014T08A7-VT08A7, (2015).

[5] Oskouei FS, Pourgol-Mohammad M. *"Optimal Sensor Placement for Efficient Fault Diagnosis in Condition Monitoring Process; A Case Study on Steam Turbine Monitoring"*. Current Trends in Reliability, Availability, Maintainability and Safety. Springer;2016,

[6] Kiad S, Pourgol-Mohammad M, Salimi H. *"Probabilistic reliability evaluation of space system considering physics of fatigue failure"*. Accelerated Stress Testing & Reliability Conference (ASTR), 2016 IEEE: IEEE. p. 1-8, (2016).

[7] Barabady J, Kumar U. "*Reliability analysis of mining equipment: a case study of a crushing plant at Jajarm Bauxite Mine in Iran. Reliability Engineering & System Safety"*,93:647-53,(2008).

[8] Uzgören N, Elevli S, Elevli B, Uysal Ö. "*Reliability analysis of draglines‘ mechanical failures. SCIENCE AND TECHNOLOGY*":23-8,(2010).

[9] Morad AM, Pourgol-Mohammad M, Sattarvand J. *"Reliability-centered maintenance for off-highway truck: case study of sungun copper mine operation equipment"*. Proceedings of the ASME International Mechanical Engineering Congress & Exposition, (2013).

[10] Soleimani M, Pourgol-Mohammad M. "*Design for reliability of complex system with limited failure data; case study of a horizontal drilling equipment. Proceedings of the Probabilistic Safety Assessment and Management PSAM"*,12,(2014).

[11] Cox DR. "*Regression models and life-tables. Journal of the Royal Statistical Society Series B (Methodological)"*:187-220,(1972).

[12] Kumar D. "*Proportional hazards modelling of repairable systems. Quality and reliability engineering international"*,11:361-9,(1995).

[13] Ghodrati B, Kumar U. "*Reliability and operating environment-based spare parts estimation approach: a case study in Kiruna Mine, Sweden. Journal of Quality in Maintenance Engineering"*,11:169-84,(2005).

[14] Gasmi S, Love CE, Kahle W. "*A general repair, proportional-hazards, framework to model complex repairable systems. IEEE Transactions on Reliability"*,52:26-32,(2003).

[15] Barabadi A, Gudmestad OT, Barabady J. "*RAMS data collection under Arctic conditions. Reliability Engineering & System Safety"*,135:92-9,(2015).

[16] Gao X, Barabady J, Markeset T. "*An approach for prediction of petroleum production facility performance considering Arctic influence factors. Reliability Engineering & System Safety"*,95:837-46,(2010).

[17] Kalbfleisch JD, Prentice RL. *"The statistical analysis of failure time data"*. John Wiley & Sons; 2011.

[18] Barabadi A, Barabady J, Markeset T. "*A methodology for throughput capacity analysis of a production facility considering environment condition. Reliability Engineering & System Safety"*,96:1637-46,(2011).
[19] Peixoto JL. "*Hierarchical variable selection in polynomial regression models. The American Statistician"*,41:311-3,(1987).
[20] Wang Z. "*Model selection using the Akaike information criterion. Stata Technical Bulletin"*,9,(2000).
[21] Konishi S, Kitagawa G. *"Information criteria and statistical modeling"*. Springer Science & Business Media; 2008.
[22] Rawlings JO, Pantula SG, Dickey DA. *"Applied regression analysis: a research tool"*. Springer Science & Business Media; 2001.
[23] Cox D. "*The statistical analysis of dependencies in point processes. Stochastic point processes"*:55-66,(1972).
[24] Singer JD, Willett JB. *"Applied longitudinal data analysis: Modeling change and event occurrence"*. Oxford university press; 2003.
[25] Kumar D, Klefsjö B. "*Proportional hazards model: a review. Reliability Engineering & System Safety"*,44:177-88,(1994).
[26] Bender R, Augustin T, Blettner M. "*Generating survival times to simulate Cox proportional hazards models. Statistics in medicine"*,24:1713-23,(2005).