QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL, VOL. 11, 361-369 (1995)

PROPORTIONAL HAZARDS MODELLING OF REPAIRABLE SYSTEMS

DHANANJAY KUMAR
Department of Civil and Mining Engineering, Luleå University of Technology, Luleå, S-97187, Sweden

SUMMARY

The purpose of this paper is to illustrate some situations under which the proportional hazards model (PHM) and its extensions can be used for identification of the most important covariates influencing a repairable system. First of all an overview of the application of the PHM in engineering is presented. Then the concepts of the PHM and its extensions, such as stratified PHM, PHM in the case of non-homogeneous Poisson processes and PHM in the case of jumps in the hazard rate or different intensity function at failures of a large number of copies of a repairable system, are presented. Selection of a suitable extension of the PHM for given data on the basis of residual plots is also discussed. Finally applications of the PHM and its extensions are illustrated with a suitable example. Only the semi-parametric method has been considered. The assumptions made in the PHM for the analysis of repairable systems have been explained graphically as far as possible. Perfect, minimal or imperfect repairs carried out on repairable systems can be taken into consideration for the reliability analysis using the PHM.

KEY WORDS: proportional hazards model; covariates; repairable system; reliability; hazard rate; intensity function; perfect repair; minimal repair; good-as-new condition; bad-as-old condition

1. INTRODUCTION

This paper deals with the application of the proportional hazards model (PHM) proposed by Cox[1] and its extensions for repairable systems. A repairable system is regarded as one which is repaired after failure but not replaced. However, some of its components may be replaced. The failure time of a system may be dependent on covariates, explanatory or concomitant variables. A covariate may, for instance, include the operating environment (e.g. temperature, pressure, humidity or dust), the operating history of a machine (e.g. overhauls, effects of repair, preventive or opportunistic maintenance), or the type of design or material. A covariate may be constant over time (e.g. design or material) or may vary over time (e.g. number of repairs done on a machine or accumulated stress factor in a component).

The PHM can be used to model the influence of covariates on the observed failure process, of repairable as well as non-repairable systems. Perfect, minimal or imperfect repairs carried out on a repairable systems can be taken into consideration for the reliability analysis using the PHM. The PHM can also be used to decide optimum maintenance time intervals of repairable systems taking covariates into consideration.[2] One major advantage of the PHM is that one need not make any specific assumptions about the functional form of the hazard rate.

Estimating the magnitudes of effects of a covariate can assist management in deciding which critical parameters are more important to be controlled or improved for avoiding failures of a system. From a manufacturer's point of view, some of these parameters can be considered during the design stage of a system. Such a study can help in analysing the effects of complex and varying operating conditions imposed on a system in order to improve the system reliability. It can also help in planning of maintenance activities as well as in condition monitoring of critical parameters of a system.

As we are applying the model to a repairable system, some changes in notation and terminology are necessary. We denote the times between failures by x and the total operating time by t. The term hazard rate is denoted by $h(x)$, and $h(x)\mathrm{d}x$ is defined as the conditional probability that a system will fail in a small time interval $(x,x+\Delta x]$ provided it has not failed upto time x. The term intensity function is denoted by $\nu(t)$, and $\nu(t)\mathrm{d}t$ is interpreted as the probability that a failure, not necessarily the first, occurs in $(t,t+\Delta t]$. More details can be found in Reference 3.

An overview of the problem areas where the PHM has been applied will be given in Section 2. The basic model, its extension and the selection of a suitable form of the PHM for a given data set are discussed in Sections 3, 4 and 5, respectively. The applications of the extension of the PHM, residual plots and results for data on electrical cable failures are discussed in Sections 6, 7 and 8, respectively.

2. AN OVERVIEW OF THE APPLICATIONS OF THE PHM AND ITS EXTENSIONS

PHM has been extensively used in the medical field. Its application in reliability analysis is diverse.

CCC 0748-8017/95/050361-09

Received 14 March 1995
Revised 5 August 1995

Booker *et al.*[4,5] have applied the PHM to compare the hazard rates of various types of valves operating under different conditions in nuclear power plants. The problem they encountered was that if the data were compartmented into homogenous groups, there were too few failures available for meaningful analysis. The PHM facilitates a precise estimate of the reliability by borrowing the strength from the total data set. Ascher[6] used the PHM for drawing conclusions about the reliability characteristics for highly censored failure time data of ship sonars used under different operating conditions. In the case of large numbers of items with repeated failure and repair, an extension of the PHM has been used to identify important covariates influencing the life length of a marine gas turbine.[6] Dale[7] has illustrated the application of the PHM as an alternative to the accelerated failure time model in estimating the hazard rate of motorettes operating at various temperatures. Jardine and Anderson[8] and Jardine *et al.*[9] apply the PHM for more precise reliability prediction using oil analysis for aircraft engines. A comparison of precision in estimation of hazard rates with and without considering covariates indicates that the PHM gives better results.[9] The PHM helps in gaining a better understanding of the effects of operating variables and their interactions. The PHM was used to explain the reason for higher hazard rate of brake discs used for high-speed trains when a metallurgically improved material was used.[10] Ansell and Ansell[11] use the PHM in modelling reliability of the sodium sulphur cells. Lindqvist *et al.*[12] apply the PHM and its extension in identifying the critical covariates influencing the hazard rate of surface controlled subsurface safety valves used in the North Sea for oil or gas well safety systems. Mazzuchi and Soyer[13] used the PHM in modelling tool life considering the effect of tool wear and the machining environment, such as tool feed rate and cutting speed. Previously, the effect of either tool wear or the machining environment was considered in describing the stochastic nature of the tool life. But, in the PHM both could be considered together. An application of the PHM in predicting failures on the basis of condition monitoring can be found in Reference 14. The effects of different failure modes and the maintenance policies of aircraft cargo doors have been analysed to quantify the hazard of a delay related to a cause and to predict future failures using the PHM.[15] Kumar *et al.*,[16] Kumar[17] and Kumar and Klefsjö[18] have used the PHM for identifying the most important mining operating conditions influencing the life length of power supply cables of electric load-haul-dump machines. In almost all the cases it has been tacitly assumed that after every failure, the items return to the good-as-new condition through repair or the item is replaced with a new one. It can be said that the application of the PHM is still mostly restricted to the case of good-as-new after repair in the engineering applications. Recently, the extension of the PHM has been applied to estimate the optimum time interval between two preventive maintenance actions considering the bad-as-old as well as good-as-new conditions.[2,19] Reviews of the PHM can be found in References 20 and 21.

3. THE BASIC MODEL

In the PHM, the hazard rate of a system is decomposed into two components, one as a function of time only, and the other as a function of covariates only. It is assumed that the hazard rate of a system is the product of an unspecified baseline ROCOF $h_0(t)$, dependent on time only, and a positive functional term, generally the exponential function, incorporating the effects of a number of covariates such as temperature, pressure and changes in design. Thus

$$h(x;\mathbf{z}) = h_0(x)\exp(\mathbf{z}\boldsymbol{\beta}) \tag{1}$$

where $\mathbf{z}$ is a row vector consisting of the covariates and $\boldsymbol{\beta}$ is a column vector consisting of the corresponding regression parameters. The regression vector $\boldsymbol{\beta}$ is the unknown parameter of the model, defining the effects of the covariates. The baseline hazard rate represents the hazard rate that a piece of equipment would experience if the effects of all the covariates are equal to zero. This may correspond, depending on how a covariate has been defined in the model, to either a natural zero, or an arbitrarily assigned zero value. Roughly, we can say that the baseline hazard rate is the hazard rate when the covariates have no influence on the failure pattern. For discussions on methods of estimation of the baseline hazard rate, see References 9, 19, 21 and 22.

Generally, the regression vector $\boldsymbol{\beta}$ is estimated by maximizing a likelihood function which is obtained by considering the contribution to the hazard rate by the times between failures. Let $x_{(1)} < x_{(2)} < \ldots < x_{(k)}$, $k \le n$, be uncensored and independent and identically distributed times between failures of m repairable systems with repeated events; here n is the total number of observed failure times and $m \le n$. Let $F(x_i)$ be the risk set of the failure events and let there be l failure events (censored or uncensored) which have not occurred just prior to the failure event at time x_i. The product of the conditional probabilities of occurrences of a failure event at time x_i, over all such failure events, is defined as a conditional or partial likelihood,[1,23] i.e.

Prob {occurrence of a failure at $x_i | l$ failure events in the risk set and exactly one failure at x_i}

$$= \frac{h(x_i;\mathbf{z})}{\sum_{l \in F(x_i)} h(x_i;\mathbf{z}_l)} = \frac{\exp(\mathbf{z}_i\boldsymbol{\beta})}{\sum_{l \in F(x_i)} \exp(\mathbf{z}_l\boldsymbol{\beta})}$$

or

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{k} \mathrm{L}_i(\boldsymbol{\beta}) = \prod_{i=1}^{k} \frac{\exp(\mathbf{z}_i\boldsymbol{\beta})}{\sum_{l \in F(x_i)} \exp(\mathbf{z}_l\boldsymbol{\beta})} \tag{2a}$$

$L_i(\beta)$ is the conditional probability that a failure occurred at the time x_i.

If the number of tied failures, denoted by d_i, at each failure point is small compared to l, an approximation of the above partial likelihood function is given by[22]

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{k} \frac{\exp(\mathbf{z}_i\boldsymbol{\beta})}{\left[\sum_{l \in F(x_i)} \exp(\mathbf{z}_l\boldsymbol{\beta})\right]^{d_i}} \tag{2b}$$

The value of $\boldsymbol{\beta}$ that maximizes (2a) or (2b) may be obtained by using numerical methods. The estimated value is then tested for its significance so that it can be verified whether the particular covariate has any effect on the failure behaviour of the system. For this purpose, analytical[22,24] or graphical[22,25–27] methods may be used. (A review on the goodness-of-fit tests of the PHM can be found in Reference 28). The significant covariates are kept in the model.

The assumption of the multiplicative effect of covariates on the baseline hazard rate implies that the ratio of the hazard rates of any two items observed at any time x associated with covariate sets $\mathbf{z}_1$ and $\mathbf{z}_2$, respectively, will be a constant with respect to time and proportional to each other, i.e. $h(x;\mathbf{z}_1) \propto h(x;\mathbf{z}_2)$. This explains why the model is called the proportional hazards model. However, in a real situation it may happen that a covariate has different amounts of influence on the hazard rate depending on the operating time of the system. Such a covariate is called a time-dependent covariate. A different approach[29,30] needs to be adopted while modelling a time dependent covariate in the PHM.

4. EXTENSIONS OF THE PROPORTIONAL HAZARDS MODEL

Assume that a plant operates at three different temperature levels, say low, medium and high, and other covariates such as dust, pressure, etc., are also measured and also have different levels associated with the operation of the plant. The three different levels of temperature will define three strata of the failure time data. If the assumption of the multiplicative effect of a covariate is violated, i.e. if the hazard rates corresponding to these strata are not proportional, stratified PHM may be used to model the hazard rate. Similarly, if the assumption of proportionality of the hazard rate is not true for a set of one or more covariates for their certain levels, this set of covariates may also be used to model the hazard rate. The system is allowed to have different hazard rates in different strata. Within the same stratum the system is assumed to have hazard rates which are proportional, and this is not necessarily the case for the system among different strata.

A system is said to be in j different strata if it can be grouped into j groups based on one or more covariates. The above situation is explained graphically in Figure 1, where j represents the number of strata and i is the number of failures occurring within stratum j, $j = 1, 2, \ldots, r$.

The hazard rate for the system in stratum j can be expressed as

$$h_j(x;\mathbf{z}) = h_{0j}(x)\exp(\mathbf{z}\boldsymbol{\beta}) \tag{3}$$

The baseline hazard rate, $h_{0j}(x)$, may remain completely unrelated in the different strata. A likelihood function is obtained for each one of j strata as given in (2), and $\boldsymbol{\beta}$ is estimated by maximizing the overall likelihood

$$L(\boldsymbol{\beta}) = \prod_{j=1}^{r}\prod_{i=1}^{k} L_j(\boldsymbol{\beta}) = \prod_{j=1}^{r}\prod_{i=1}^{k} \frac{\exp(\mathbf{z}_{ij}\boldsymbol{\beta})}{\left[\sum_{l,j \in F(x_{ij})} \exp(\mathbf{z}_{lj}\boldsymbol{\beta})\right]^{d_i}}. \tag{4}$$

The above approach to the stratified PHM has been suggested with some or almost no modifications for the reliability analysis of a possibly small number of failure times on a fairly large number of a repairable system.[2,31,32] Prentice *et al.*[31] suggested the following two models based on the time variable t, the total time since initial start-up of a system:

$$\nu(t|N(t),\mathbf{z}) = \nu_{0j}(t)\exp(\mathbf{z}\boldsymbol{\beta}_j) \tag{5}$$

$$\nu(t|N(t),\mathbf{z}) = h_{0j}(t - t_{n(t)})\exp(\mathbf{z}\boldsymbol{\beta}_j) \tag{6}$$

The symbol $N(t), t \geq 0$, is a counting process of failures and $N(t)$ is the number of failures of a system prior to time t, j represents the stratum which a system occupies at any instant t. The regression vector $\boldsymbol{\beta}_j$ is the regression coefficient for the jth stratum. The advantages of the models (5) and (6) are that depending on situations, one may assume the same baseline intensity function, baseline hazard rate or $\boldsymbol{\beta}_j = \boldsymbol{\beta}$ for all $j \geq 1$. In the case of repairable systems, it can be assumed that a system enters stratum j at the occurrence of the $(j - 1)$th failure, $j = 1, 2, \ldots, r$ and it enters stratum 1 at $t = 0$.

The approximation of the likelihood functions for estimating $\boldsymbol{\beta}$ in models (5) and (6), can be written as given in (7) and (8), respectively:[31]

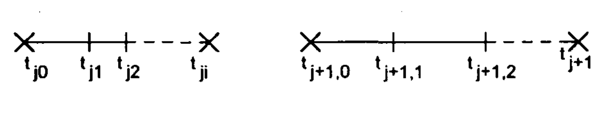

Figure 1. A graphical representation of the strata for a data set

$$L(\boldsymbol{\beta}) = \prod_{j=1}^{r} \prod_{i=1}^{k} \frac{\exp(\mathbf{z}_{ij}\boldsymbol{\beta}_j)}{\left[\sum_{l,j \in F(t_{ij})} \exp(\mathbf{z}_{lj}\boldsymbol{\beta}_j)\right]^{d_i}} \quad (7)$$

$$L(\boldsymbol{\beta}) = \prod_{j=1}^{r} \prod_{i=1}^{k} \frac{\exp(\mathbf{z}_{ij}\boldsymbol{\beta}_j)}{\left[\sum_{l,j \in F(x_{ij})} \exp(\mathbf{z}_{lj}\boldsymbol{\beta}_j)\right]^{d_i}} \quad (8)$$

Equation (8) is similar to equation (4) if $\boldsymbol{\beta}_j = \boldsymbol{\beta}$, $j \geq 1$. Lawless[32] and Love and Guo[2] considered that each one of m systems, i.e. $j = 1, 2, \ldots, m$ should form a stratum, if the failures of a system occur according to NHPP. The Lawless[32] approach was based on fixed covariates during the total operating time. This approach was modified by Love and Guo[2] so that the covariate value can vary at different failure times.

Recently, some modification in the stratified PHM has been suggested by Guo and Love[33] to model imperfectly repaired systems. This modification is based on a time transformation of a non-homogeneous Poisson process to a homogeneous Poisson process. After a repair, the system is assumed to be younger than its chronological age. The system does not follow its intensity function in chronological order. The effect of a repair is modelled via an age reduction factor, which may be assumed to be a constant for all repairs or may be assumed to take any value between 0 and 1 depending on the effectiveness of a repair. The details of the likelihood functions, the corresponding algorithm to estimate parameters and examples can be found in Reference 33.

The approach of the stratified PHM is one of the simplest and most useful extensions of the PHM for its application in different situations, e.g. goodness-of-fit test for the PHM, cause-specific hazards problems and competing risk problems. Other possible extensions of the PHM are discussed by Kalbfleisch and Prentice.[22]

5. SELECTION OF THE SUITABLE FORM OF THE MODEL FOR A GIVEN FAILURE DATA SET

For selecting the suitable form of the PHM for a given failure time data set, one may use the different methods of goodness-of-fit tests (see Reference 28 for details). The easiest way is to compare the residuals after fitting the various extensions of the PHM. The residuals should follow an exponential distribution with unit mean if the PHM assumption is suitable.[21,24,25] The residuals are estimated by using the estimated values of the cumulative baseline hazard rate, $H_0(x)$, and the regression vector $\boldsymbol{\beta}$, i.e.

$$\hat{e}_i = \hat{H}_0(x_i) \exp(\hat{\boldsymbol{\beta}}\mathbf{z}_i) \quad (9)$$

The observed residuals plotted against the expected order statistic (or the cumulative hazard rate of the residual) should lie on a straight line of unit slope. The cumulative baseline hazard rate, $H_0(x)$, may be estimated using the methods suggested by Breslow,[34] i.e.

$$\hat{H}_0(x_i) = \sum_{i:x_i \leq x} \frac{d_i}{\sum_{l \in F(x_i)} \exp(\mathbf{z}_l\hat{\boldsymbol{\beta}})} \quad (10)$$

In the case of stratified PHM, we use the following equivalent for estimating cumulative baseline intensity function:

$$\hat{V}_0(t_{ij}) = \sum_{i,j:t_{ij} \leq t} \frac{d_i}{\sum_{jl \in F(t_{ij})} \exp(\mathbf{z}_{jl}\hat{\boldsymbol{\beta}})} \quad (11)$$

where $F(t_{ij})$ is the risk set at time t_{ij} in stratum j. We use time between failures x_{ij} in place of total operating time t_{ij} in equation (11) for estimating cumulative baseline hazard rate, $H_0(x_{ij})$, if the stratified PHM given in (3) is used for the reliability analysis.

The expected value of the rth largest of n independent random variables following an exponential distribution with unit mean is[35]

$$E_{r,n} = \frac{1}{n} + \frac{1}{n-1} + \ldots + \frac{1}{n-r+1}, \quad r = 1, 2, \ldots, n \quad (12)$$

The applications of the PHM and its modifications under different assumptions and the selection of the most suitable model for a repairable system are discussed in the following sections.

6. APPLICATIONS OF THE PROPORTIONAL HAZARDS MODEL UNDER DIFFERENT ASSUMPTIONS

Electric load–haul–dump (LHD) machines are used for transport of the fragmented ore from a production face to an ore-pass at the LKAB Kiruna Iron Ore Mine, Kiruna, Sweden. The length of the cable that is used to supply electrical power to the 1000V, 50 Hz, AC motor of the machine varies from 275–325 m and weighs 2·90 kg/m. These cables are wrapped and unwrapped very frequently on a cable reel mounted on the machine, as the machine moves away from the electrical supply source for loading or dumping of ore. Generally, overheating and breaking due either to overrunning by the machine or to weak welded joints, are the main reasons for occurrences of faults in cables. The cable is repaired after failure and used again. Generally after a cable has been repaired 7–8 times, it is discarded.

The covariates that were considered included four fault types, the number of times repairs were carried out on the cable, and the age and identity number of the twelve machines on which these cables are used. These four fault types were selected after combining faults of similar nature into four groups. They correspond to faults due to breakage of a cable (fault type 1), electrical faults (fault type 2), problem at the welded joint (fault type 3) and the remaining type of faults (fault type 4). See References 17 and 18 for details. The data consisted of 600 times to failures of 68 cables. Each cable was assigned an identity number. This information was useful in identifying the failure times of each individual cable.

6.1. *Assuming perfect repair or good-as-new condition*

Perfect repair restores a system to 'good-as-new' condition (see Figure 2). It is assumed that the hazard rate is reset to that of a new system after repair. If the times between failures are independent and identically distributed, it can be assumed that the system undergoes perfect repair after failure. It is most easily thought of as a replacement of the failed system with a new one. This situation is treated in the same way as failures of non-repairable systems. In such cases, the original form of the PHM as introduced by Cox[1] and given in (1) can be used. The likelihood function given in (2) is used for estimating $\boldsymbol{\beta}$.

The failure times of all the cables irrespective of their identity numbers were considered together. All the covariates were considered together and the corresponding estimates of $\boldsymbol{\beta}$ were obtained by using equation (2). The statistical software S-PLUS[36] was used for the analysis. The estimates of $\boldsymbol{\beta}$ were tested for their significance at 10 per cent p-value on the basis of the t-statistic. The t-statistics were calculated by taking the ratio of the estimate of $\boldsymbol{\beta}$ to the standard deviation of the estimate. A p-value of the t-statistic corresponding to the estimate of $\boldsymbol{\beta}$ may be interpreted as the probability of obtaining such an extreme value for the estimate of $\boldsymbol{\beta}$, if it is equal to zero. The covariates whose p-values were less than 10 per cent were discarded from the model. The estimates of $\boldsymbol{\beta}$ were found significant only for

Table I. List of the covariates whose effects were found significant at 10 per cent p-value (assumption of good-as-new condition). The estimates of β and σ were obtained by using function (2)

Covariate	$\hat{\boldsymbol{\beta}}$	$\hat{\sigma}$	p-value
1. Fault type 1	0·2166	0·1039	$3{\cdot}72 \times 10^{-2}$
2. Fault type 2	0·2044	0·1107	$6{\cdot}47 \times 10^{-2}$
3. Fault type 3	0·3757	0·1277	$3{\cdot}25 \times 10^{-3}$
4. Machine age	0·5339	0·1177	$5{\cdot}73 \times 10^{-6}$
5. Machine identity number 5	0·3686	0·1734	$3{\cdot}36 \times 10^{-2}$
6. Machine identity number 9	0·4143	0·1930	$3{\cdot}19 \times 10^{-2}$
7. Machine identity number 10	0·6559	0·1787	$2{\cdot}41 \times 10^{-4}$
8. Machine identity number 11	0·5970	0·1961	$2{\cdot}33 \times 10^{-3}$
9. Machine identity number 12	0·6158	0·1820	$7{\cdot}17 \times 10^{-4}$
10. Failure number	0·0979	0·0103	0·00

the ten covariates listed in column 1 of Table I. The corresponding estimates of $\hat{\boldsymbol{\beta}}$ along with their standard errors, $\hat{\sigma}$, and the p-values are given in columns 2–4.

6.2. *Assuming minimal repair or bad-as-old condition*

In the case of minimal repair, a system has the same intensity function after repair as before the failure. Sometimes this condition is referred to as the 'bad-as-old' condition (Figure 3). Generally a minimal repair is carried out until the replacement costs of a system are more than its repair costs. Otherwise it is replaced with a new one. It may happen that the system undergoes emergency repair whenever a breakdown occurs. But after a certain time or depending on convenience, a thorough overhaul or preventive maintenance is carried out. In other words, a system after emergency breakdowns is repaired sufficiently to continue to operate it and a planned maintenance is carried out later on.

It may be assumed that a system is in bad-as-old conditions after emergency repairs and it is in good-as-new conditions after a preventive maintenance

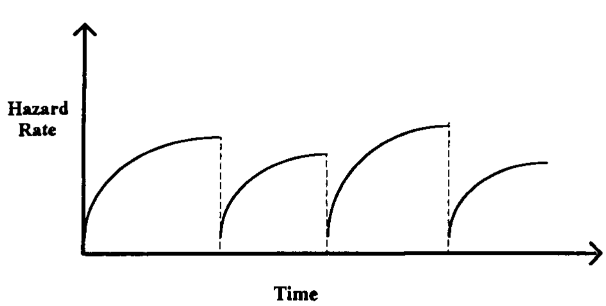

Figure 2. An illustration of the patterrn of the hazard rate in the case of perfect repair or good-as-new condition of a repairable system

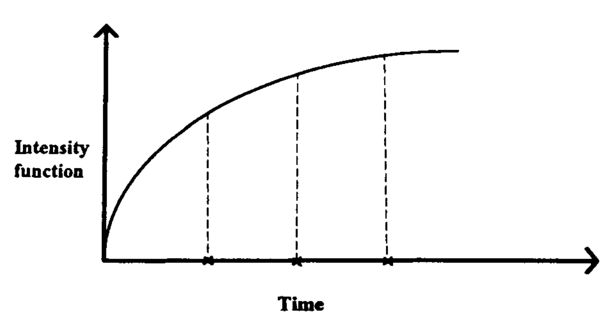

Figure 3. An illustration of the pattern of the intensity function in the case of minimal repair or 'bad-as-old' condition of a repairable system

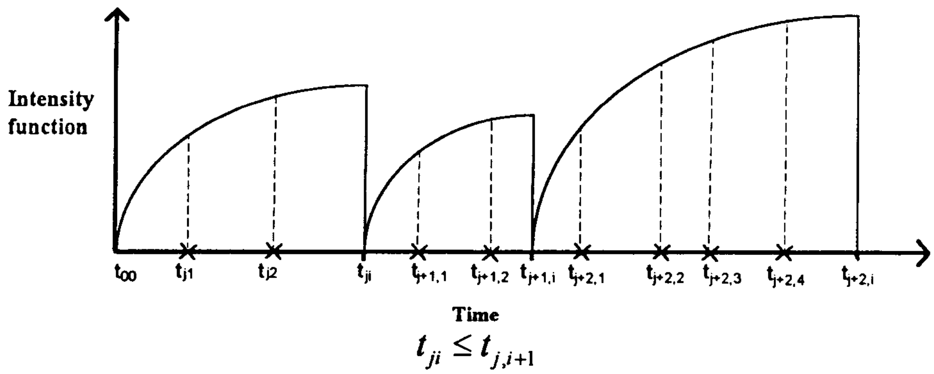

Figure 4. Graphical representation of a bad-as-old failure process modelled in the proportional hazards model. Here i is the number of failures of a cable and j is the cable identity number (variable defining the strata)

or a replacement. These situations are explained in Figure 4, where j represents the number of renewals through a thorough preventive maintenance, overhauls or replacements, and i is the total number of failures occurring in any fixed j, $j = 1, 2, \ldots, r$ and $i = 0, 1, 2, \ldots, k$. The failure times when minimal repair is carried out or the bad-as-old process can be thought as a non-homogeneous Poisson process (NHPP). It is assumed that the NHPP is cyclic and each cycle starts as a renewal process and within the cycle, failure times follow the NHPP.[2,32]

It was assumed that minimal repair is carried out on the cable whenever a failure occurs. The failure process can be graphically explained with the help of Figure 4, where j corresponds to the cable identity number which is the stratum variable. The renewal occurs whenever a new cable is put into operation. The failure times of an individual cable are assumed to follow an NHPP and the use of a new cable for operation is assumed to be the starting time point of a renewal process or the start of a cycle within which failure times are NHPP. The effects of the covariates, $\boldsymbol{\beta}$ were estimated using the likelihood function (7) and taking the cable identity number as strata. We consider the same value of the vector $\boldsymbol{\beta}$ across all strata. The results are listed in Table II.

Table II. List of the covariates whose effects were found significant at 10 per cent p-value (assumption of bad-as-old condition). The estimates of $\boldsymbol{\beta}$ and σ were obtained by using function (7)

Covariate	$\hat{\boldsymbol{\beta}}$	$\hat{\sigma}$	p-value
1. Fault type 3	0·3068	0·1625	$5{\cdot}91 \times 10^{-2}$
2. Failure number	0·0964	0·0144	$2{\cdot}45 \times 10^{-11}$
3. Machine identity number 9	0·5780	0·2418	$1{\cdot}68 \times 10^{-2}$
4. Machine identity number 10	0·8184	0·2257	$2{\cdot}87 \times 10^{-4}$
5. Machine identity number 11	0·8170	0·2372	$5{\cdot}71 \times 10^{-4}$
6. Machine identity number 12	0·6465	0·2204	$3{\cdot}35 \times 10^{-3}$
7. Machine age	0·5748	0·1608	$3{\cdot}51 \times 10^{-4}$

6.3. *Assuming jumps in the hazard rate after repair or different baseline hazard rate*

If the baseline hazard rate is not identical, stratified PHM as given in (3) can be used for analysing the effects of covariates. For example, the baseline hazard rate of all the copies of a repairable system can be assumed to be identical for all cables between repair numbers j and $j+1$ and the baseline hazard rates corresponding to repair numbers $j, j+1, \ldots, r$ are proportional to each other.

A simple analysis of the cable failure data indicates that as the number of failures increases the average failure time decreases.[18] Hence, we can assume that the baseline hazard rate for a particular failure number is different from the others. This has been graphically represented in Figure 5, where j refers to failure order numbers and i refers to all cable identity numbers. For example when $j=1$, the first failure times of all cables are assumed to be in stratum number one. The estimates of the effects of the covariates were roughly the same as listed in Table I, except that the tenth covariate (failure number) is not listed, because it was used to define strata in this model and hence no estimate is obtained. The results are listed in Table III.

6.4. *Assuming the effect of operating and repair history*

To consider the effects of operating and repair history on the intensity function of cables, we need to take the failure order number as a stratum variable, the same as in Section 6.3. But the time variable should take the value of the age of each individual cable in each stratum instead of times between failure. A graphical representation of this situation is given in Figure 6.

The covariates for the cable age were excluded from the analysis. The cable age was considered as the basic time for estimating $\boldsymbol{\beta}$ in (7). Only one covariate was found significant (see Table IV). The result obtained using the stratified PHM as given in (5) is quite different compared to that in sections 6.1, 6.2, and 6.3 for the same cable data.

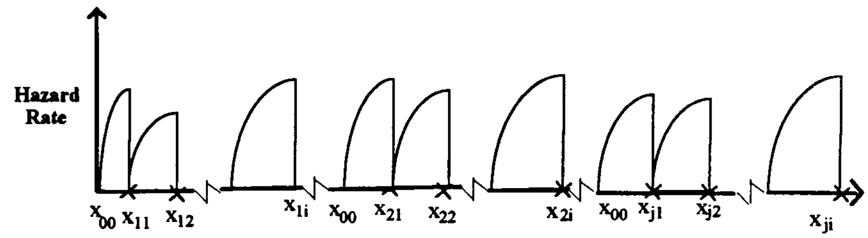

Figure 5. An illustration of the pattern of the hazard rate when assuming different baseline hazard rates after a repair has been carried out on the cable. Here i is the cable identity number and j is the failure order number of a cable (variable defining the strata)

Table III. List of the covariates whose effects were found significant at 10 per cent p-value (assumption of jumps in the hazard rate). The estimates of $\boldsymbol{\beta}$ and σ were obtained by using the function (4)

Covariate	$\hat{\boldsymbol{\beta}}$	$\hat{\sigma}$	p-value
1. Fault type 1	0·253	2·25	0·0246
2. Fault type 2	0·254	2·16	0·0309
3. Fault type 3	0·381	2·79	0·0051
4. Machine age	0·425	3·25	0·0011
5. Machine identity number 5	0·315	1·69	0·0914
6. Machine identity number 9	0·484	2·42	0·0154
7. Machine identity number 10	0·626	3·32	0·0008
8. Machine identity number 11	0·665	3·16	0·0015
9. Machine identity number 12	0·573	2·89	0·0037

7. RESIDUALS

The residuals were estimated using equations (9) and (10) or (11). In the case of censored failure times, residuals were estimated by adding one to the previously estimated value of the residual.[24] The estimated residuals were plotted against the expected order statistics (12) (see Figure 7). Plots in Figures 7(a), 7(b), 7(c) and 7(d) correspond to the assumptions made in Sections 6.1, 6.2, 6.3 and 6.4, respectively. The plotted points in Figure 6(d) lie very close to the straight line with unit slope. This implies that the corresponding assumptions and fitted model in section 6.4 is the most suitable one for the cable failure data.

8. EXPLANATION OF RESULTS

For the assumption of different baseline intensity function and the total operating time of a cable (i.e. cable age) as the basic time (see section 6.4), the

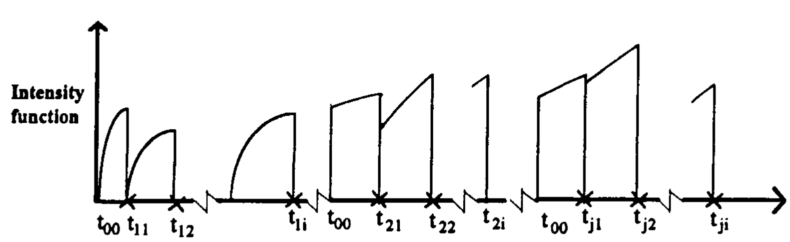

Figure 6. An illustration of the pattern of the intensity function when assuming different baseline intensity functions for different failure order numbers. Here i is the cable identity number and j is the number of failures of a cable (variable defining the strata)

intensity function of the cable is given by (see Table IV).

$$\nu_j(t;\mathbf{z}) = \nu_{0j}(t)\exp(0{\cdot}184z) \qquad (13)$$

where z represents the covariate fault type two and j is the failure order number which takes the value of one to the maximum number of failures observed on these cables. We should take into consideration the way covariates were formulated and the assump-

Table IV. List of the covariates whose effects were found significant at 10 per cent p-value (assumption of the effect of operating and repair history). The estimates of $\boldsymbol{\beta}$ and σ were obtained by using the function (7)

Covariate	$\hat{\boldsymbol{\beta}}$	$\hat{\sigma}$	p-value
1. Fault type 2	0·184	0·0977	0·0594

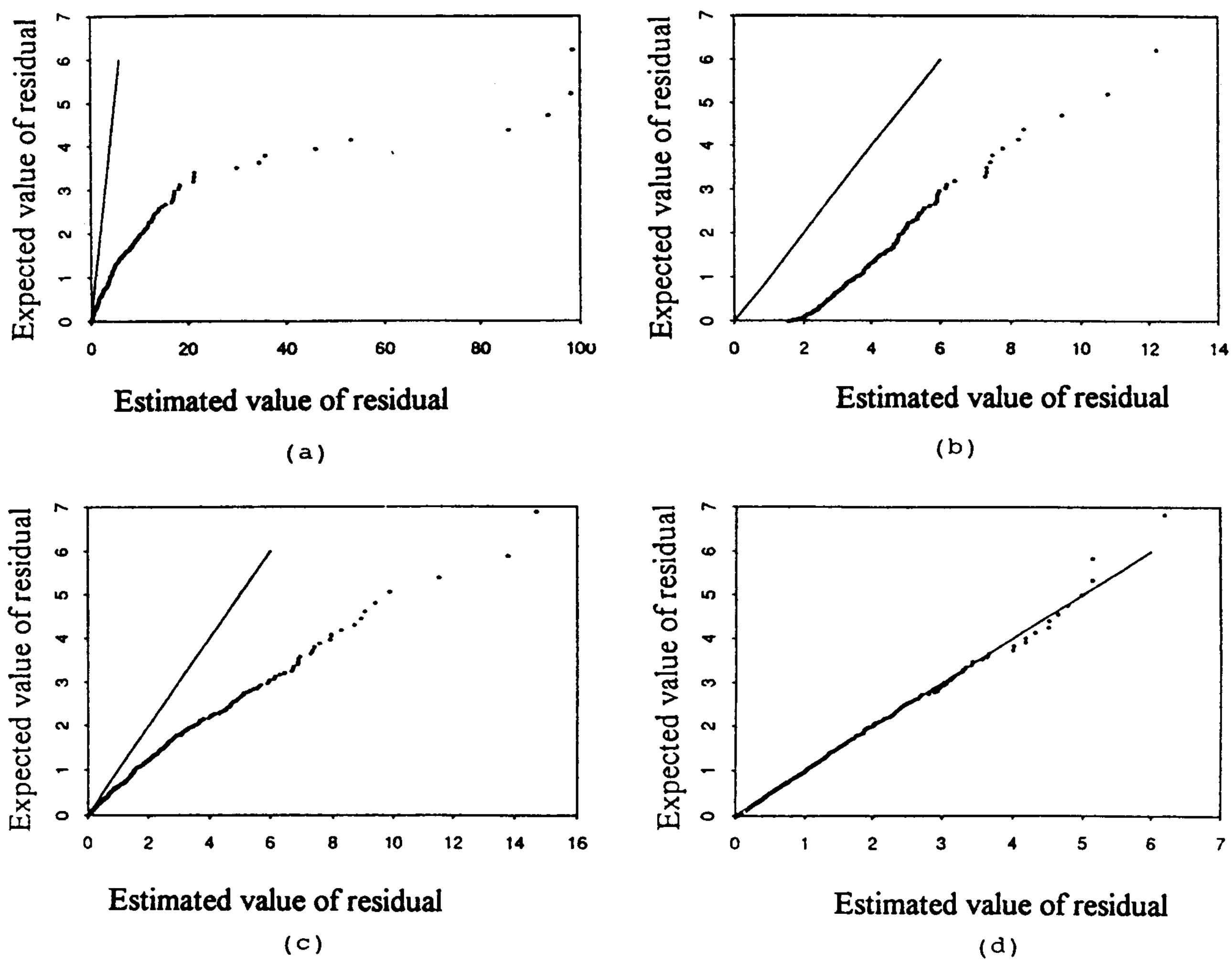

Figure 7. Plots of estimated and expected values of residuals to check whether they are exponentially distributed with parameter one. The plotted points should lie near to the straight line with unit slope. Figures 7(a) and 7(b) correspond to the assumptions of good-as-new and bad-as-old conditions, respectively, about the cable failure data. Figures 7(c) and 7(d) correspond to the assumptions of different baseline hazard rate and different baseline intensity function, respectively. The plotted points lie on the straight line with unit slope in Figure 7(d). This implies that the corresponding assumption about the cable failure data is a suitable one

tions that were made while interpreting their effects. The estimated effects are relative risks.

The fault type two is the most important covariate influencing the intensity function of the cable among the various covariates considered in the model. The occurrence of fault type two increases the intensity function of the cable by a multiplicative factor of 1·20 (exp(0·184) = 1·20) on average. Hence the elimination of causes of the fault type two can decrease the intensity function by a multiplicative factor of 1·20 on average.

Therefore to improve the reliability of the cable, we should try to eliminate the causes of fault type two. For further decisions about the reliability improvement, we may analyse the data without considering the covariate of fault type two in the model.

9. CONCLUSIONS

It is necessary to consider the covariates associated with the operation of a system for a better explanation of the failure characteristics. The extensions of the PHM facilitate the possibility of considering the different situations under which a repairable system is used. Perfect and minimal repair situations can be considered in the model. The concept of stratified PHM makes it a powerful tool for the reliability analysis of repairable systems.

The PHM analysis gives different results depending on the type of assumptions imposed on the failure data (e.g. strata selection). In reality, the repairs are neither good-as-new nor bad-as-old. Therefore it is important to check how closely the different assumptions are fulfilled. Selection of the suitable form of the PHM should be based on goodness-of-fit-tests, experience and the physical reality. It should preferably be used as tools to provide explanatory clues, rather than merely for fitting a model to a given data set.

ACKNOWLEDGEMENT

The author is grateful to Professor Bengt Klefsjö and Ulf Westberg, Division of Quality Technology and Statistics, Luleå University of Technology, Luleå, Sweden for their help and suggestions during the preparation of the paper. The financial support from the Swedish Research Council for Engineering

Sciences under the project number 95-55 is gratefully acknowledged.

REFERENCES

1. D. R. Cox, 'Regression models and life-tables', *Journal of the Royal Statistical Society,* **B34**, 187–220 (1972).
2. C. E. Love and R. Guo, 'Using proportional hazard modelling in plant maintenance', *Quality and Reliability Engineering International,* **7**, 7–17 (1991).
3. H. E. Ascher and H. Feingold, *Repairable Systems Reliability: Modeling, Inference, Misconceptions and their Causes*, Marcel Dekker, New York, 1984.
4. J. Booker, K. Campbell, A. G. Goldman, M. E. Johnson and M. C. Bryson, 'Application of Cox's proportional hazards model to light water reactor component failure data', *Report number: LA-8834-SR*, Los Alamos Scientific Laboratory, Los Alamos, New Mexico 87545, 1981.
5. J. B. Booker, M. E. Johnson and R. G. Easterling, 'Proportional hazards models for improving the analysis of light water reactor component failure data', *Report No. LA-UR-81-2803*, Los Alamos National Laboratory, Las Alamos, New Mexico 87545, 1981.
6. H. Ascher, 'Regression analysis of repairable systems reliability', in J. K. Skwirzinsky (ed.) *Electronic Systems Effectiveness and Life Cycle Costing*, 1983, pp. 119–133.
7. C. J. Dale, 'Application of proportional hazard model in the reliability field', *Proceedings of fourth National Reliability Conference*, Birmingham, U.K., 1983, pp. 5B/11/1–9.
8. A. K. S. Jardine and M. Anderson, 'Use of concomitant variables for reliability estimation', *Maintenance Management International,* **5**, 135–140 (1985).
9. A. K. S. Jardine, P. M. Anderson and D. S. Mann, 'Application of the Weibull proportional hazards model to aircraft and marine engine failure data', *Quality and Reliability Engineering International,* **3**, 77–82 (1987).
10. A. Bendell, M. Walley, D. W. Wightman and L. M. Wood, 'Proportional hazards modelling in reliability analysis—an application to brake discs on high speed trains', *Quality and Reliability Engineering International,* **2**, 45–52 (1986).
11. R. O. Ansell and J. Ansell, 'Modelling the reliability of sodium sulpher cells', *Reliability Engineering,* **17**, 127–137 (1987).
12. B. Lindqvist, E. Molnes and M. Rausand, 'Analysis of SCCV performance data', *Reliability Engineering and System Safety,* **20**, 3–17 (1989).
13. T. A. Mazzuchi and R. Soyer, 'Assessment of machine tool reliability using a proportional hazards model', *Naval Research Logistics,* **36**, 765–777 (1989).
14. A. K. S. Jardine, P. Ralston, N. Reid and J. Stafford, 'Proportional hazards analysis of diesel engine failure data', *Quality and Reliability Engineering International,* **5**, 207–216 (1989).
15. A. L. F. Leitao and D. W. Newton, 'Proportional hazards modelling of aircraft cargo door complaints', *Quality and Reliability Engineering International,* **5**, 229–238 (1989).
16. D. Kumar, B. Klefsjö and U. Kumar, 'Reliability analysis of power transmission cables of electric mine loaders using the proportional hazards model', *Reliability Engineering and System Safety,* **37**, (3), 217–222 (1992).
17. D. Kumar, 'Reliability analysis considering operating conditions in a mine', *Licentiate Thesis, 1993:17L*, Luleå University of Technology, Luleå, Sweden, 1993.
18. D. Kumar and B. Klefsjö, 'Proportional hazards model: an application to power supply cables of electrical mine loaders', *International Journal of Reliability, Quality and Safety Engineering,* **1**, (3), 337–352 (1994).
19. C. E. Love and R. Guo, 'Application of Weibull proportional hazards modelling to bad-as-old failure data', *Quality and Reliability Engineering International,* **7**, 149–157 (1991).
20. P. K. Andersen, 'The second decade of the proportional hazards regression model', *Research Report No. 91/1*, Statistical Research Unit, University of Copenhagen, Denmark 1991.
21. D. Kumar and B. Klefsjö, 'Proportional hazards model: a review', *International Journal of Reliability Engineering and System Safety,* **44**, (21), 217–222 (1994).
22. J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data,* Wiley, New York, 1980.
23. D. R. Cox, 'Partial likelihood', *Biometrika,* **62**, (2), 269–276 (1975).
24. J. F. Lawless, *Statistical Models and Methods for Lifetime Data,* Wiley, New York, 1982.
25. R. Kay, 'Proportional hazards regression models and the analysis of censored survival data', *Applied Statistics,* **26**, (3), 227–237 (1977).
26. D. Schoenfild, 'Partial residuals for the proportional hazards regression model', *Biometrika,* **69**, (1), 239–241 (1982).
27. E. Arjas, 'Graphical method for assessing goodness of fit in Cox's proportional hazards model', *Journal of American Statistical Association,* **83**, (401), 204–212 (1988).
28. R. Kay, 'Goodness of fit methods for the proportional hazards regression model: a review', *Revuè Êpidèmiologie et santè Publique,* **32**, 185–198 (1984).
29. J. A. Anderson and A. A. Senthilselvan, 'Two-step regression model for hazard functions', *Applied Statistics,* **31**, (1), 44–51 (1982).
30. D. Kumar and U. Westberg, 'Proportional hazards modelling of time-dependent covariates using a linear regression model: a case study', *Research report No. 1993:9*, Division of Quality Technology & Statistics, Luleå University, Luleå, Sweden. Submitted for publication.
31. R. L. Prentice, B. J. Williams and A. V. Peterson, 'On the regression analysis of multivariate failure data', *Biometrika,* **68**, (2), 373–379 (1981).
32. J. F. Lawless, 'Regression methods for Poisson process data', *Journal of American Statistical Association,* **82**, (399), 805–815 (1987).
33. R. Guo and C. E. Love, 'Statistical analysis of an age model for imperfectly repaired systems', *Quality and Reliability Engineering International,* **8**, 133–146 (1992).
34. N. Breslow, 'Covariance analysis of censored survival data', *Biometrics,* **30**, 89–99 (1974).
35. D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events*, Methuen & Co., London, 1966.
36. *S-PLUS*, Statistical Sciences Inc. Seattle, WA 98145-1625, U.S.A., 1991.

Author's biography:

Dhananjay Kumar obtained the degrees of B.Sc. (Mechanical Engineering) from Bihar Institute of Technology, Sindri, India, in 1987, Licentiate in Engineering from Luleå University of Technology, Luleå, Sweden in 1993 in the area of Reliability Engineering.

At present he is a research assistant at the Luleå University of Technology and working on projects related to reliability, quality and maintenance. Prior to obtaining his Licentiate degree, he worked from 1987 to 1990 in a hydroelectric power plant in India.

His research interests are in the areas of reliability, quality and maintenance engineering.