Degradation modelling using a Phase Type Distribution (PHD)

Ewa Laskowska

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Norway. E-mail: ewa.m.laskowska@ntnu.no

Jørn Vatn

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Norway. E-mail: jorn.vatn@ntnu.no

The aim of this study is to use Phase Type Distribution (PHD) for modelling degradation and maintenance of safety valves used in the oil and gas industry. The model is built as a phase expansion of the existing degradation model in form of multiphase CTMC. PHD approximation is performed in order to reflect the aging impact on the system degradation, so that sojourn times in degradation states of Markov model become Weibull distributed. In this work the practical application of PHD approximation is presented, including model assumptions, choice of approximation algorithm, PHD structure and merging of PHDs arisen by phase expansion of main states.

Keywords: phase type distribution, Markov model, degradation modelling, Weibull distribution

1. Introduction

The purpose of predictive maintenance is to provide safety and cost-efficient operation of considered systems. Quantification of these characteristics is often performed by use of stochastic models.

In the field of reliability, the most widely used class of stochastic models are Markov models. Their popularity comes from the fact that they are mathematically tractable and can describe the probabilistic behaviour of a target system as a state-based stochastic model. (Okamura and Dohi, 2016a). However, most of technical equipment is subjected to continuous deterioration due to aging, fatigue or other degradation mechanisms. In such a situation modelling of lifetime by Markov model is unrealistic due to its memoryless property. The memoryless means that the item has a constant probability to fail no matter what the current usage and level of deterioration is. Typically, lifetimes of degrading items are described by Weibull, gamma or log-normal distributions. However, for mentioned distributions deriving important system performance indices becomes very troublesome in many cases. (Okamura and Dohi, 2016b).

This challenge can be overcomed by approximation of general distributions or general stochastic point processes by use of Phase-Type Distributions (PHDs), (Okamura and Dohi, 2016a). PHD were introduced by Neuts (1975) as generalization of Erlang distribution and are defined as distributions of absorption times in Markov processes with *n* transient states and one absorbing state (Asmussen, Nerman et al., 1996). Any lifetime distribution can be approximated arbitrarily close by a phase-type distribution defined as a Continuous Time Markov Chain (CTMC) (Lindqvist, 2016), This fact makes the class of PHD very attractive for use as a parametric model for failure data with degradation trend, enabling computation of various quantities for homogeneous Markov chains (Aalen, 1995).

Although PHDs based models are widely used in queuing theory, medical studies or lifetime analysis, Buchholz, Kriege et al. (2014) noticed that PHDs have been rarely used to model failure data.

Some of phase-type applications in systems' lifetimes modelling are presented by Faddy (1995), Pérez-Ocón and Ruiz Castro (2004), Buchholz and Kriege (2014). In mentioned works transient phases can be treated as degradation stages, but they don't have any specific physical meaning, the deterioration is just a general interpretation of phases. More explicit modelling of degradation phases is proposed in work of Montoro-Cazorla and Pérez-Ocón (2006) where the considered system is continuously monitored. Although authors didn't connect phases with specific degradation level, all states before absorption are split into good and middling phases.

In this work the phase-type expansion of the existing degradation model considered by (Laskowska and Vatn, 2019) is proposed. The model presented in the previous work was a CTMC including maintenance strategy. The model was based on the empirical data where the

Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference
Edited by Piero Baraldi, Francesco Di Maio and Enrico Zio

ISBN: 978-981-14-8593-0; doi:10.3850/978-981-14-8593-0

condition of the system was defined by five deteriorating states. The fact of condition description by discrete states implied the use of the Markov model. Another argument for application of the Markov framework was the easiness of incorporation of maintenance model. As usual in Markov model, sojourn times in model states were exponentially distributed. In this paper sojourn times in main states are assumed to be Weibull distributed in order to reflect aging impact on the system degradation. This is done by introduction of additional, intermediate states between main condition states.

Similar solution of the phase expansion of an existing deterioration model was presented in paper by Yeh (1997). However, the article does not provide any information about the approach for approximating general distributions between main states, neither it explains the branching of created PHDs. Finally, authors do not provide any information about the estimation algorithm they applied, neither whether the estimation was based on the failure data, empirical or theoretical moments.

The goal of this paper is to present how PHD can be practically used for degradation modelling, including model assumptions, choice of approximation algorithm, phase type distribution structure and branching of PHDs arisen by phase expansion of main states. The model is applied to real use case and results are compared with those derived from the previously used classical CTMC (multiphase Markov model when considering maintenance).

2. Phase type distribution

Consider a time-homogeneous Markov process: $X(t_{t\geq 0}^{\propto})$ which is an absorbing CTMC on the finite state space S = 1, 2,…,n, n+1, where n states are transient and n+1 is an absorbing state. The infinitesimal generator matrix Q of considered CTMC is following:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{T} & \boldsymbol{\tau} \\ \mathbf{0} & 0 \end{bmatrix} \qquad (1)$$

where **T** is an infinitesimal generator between transient states and $\boldsymbol{\tau}$ is a column vector of transition rates from transient states to the absorbing state. For **1** being a column vector whose all elements are 1, the vector $\boldsymbol{\tau}$= -**T1.**

A PH random variable is defined by the first passage time to the absorbing state (state n+1) on the phase process $X(t_{t\geq 0}^{\propto})$. The initial state of X(t) is given by a probability vector $\boldsymbol{\pi}$ over transient states: 1,…, n. (Buchholz et al., 2014). The cumulative distribution function of a phase-type distributed variable with representation ($\boldsymbol{\pi}$, **T**) is given by:

$$F(x) = 1 - \boldsymbol{\pi} e^{\boldsymbol{T}x}\mathbf{1}, \qquad \text{for } x \geq 0 \qquad (2)$$

and its density function is given by

$$f(x) = \boldsymbol{\pi} e^{\boldsymbol{T}x}\boldsymbol{\tau}, \qquad \text{for } x \geq 0 \qquad (3)$$

The PHD is fully determined by the representation ($\boldsymbol{\pi}$, **T**), and the idea of PH fit is to find such parameters ($\boldsymbol{\pi}$, **T**) which can approximate the original distribution by the estimated PHD (Okamura and Dohi, 2016a).

The efficiency of phase-type (PH) approximation depends on the number of phases, phase structure and PH estimation algorithm. Among many possible structures of PH the most widely used are following: mixture of exponential/Erlang distributions, Acyclic Phase-type Distributions (APHD) and general PHDs (Okamura and Dohi, 2016a).When it comes to number of used phases, the rule of thumb is to specify the number of phases based on the moments of fitted distribution. The idea is to keep the minimal number of phases so that moments of approximated PHD fit moments of the original distribution (empirical or theoretical moments). The literature proposes also other methods for determining minimal number of phases, but this topic is out of the scope of this work. Considering approximation methods, the two most widely used approaches are Method of Moments (MM) and Maximum Likelihood Estimation (MLE), which is often performed by Expectation Maximization (EM) algorithm. The idea of MM is to search for such PH parameters which fit original moments. When one works on statistical data and empirical moments the MM estimation can be not so accurate, especially when a sample is of small size. In addition, in situation when some data are missing, which is very common in real case scenarios, it is not possible to calculate even first moment of original distribution (Okamura and Dohi, 2016a). The MLE is a commonly used technique to estimate model parameters by maximizing the likelihood with regard to unobserved data when observed data are given. The MLE is especially efficient when one has to work on the observed failure data. (Okamura and Dohi, 2016a)

3. Model

3.1 Use case & existing degradation model

In the previous work (Laskowska et al. 2019) a multi-phase Markov model was proposed for degradation modelling of Emergency Shutdown Valves (ESVs) used in the oil and gas industry. The model is built based on the maintenance data for 92 ESVs being periodically subjected to functional tests and other types of inspections. Just before and just after each inspection the condition of the valve was described by one of five states: 1=*new*, 2=*good*, 3=*slightly degraded*, 4=*significantly degraded*, 5=*failed*. Because the degradation is assumed to be gradual only transitions between adjacent states are considered. Transition rates between two states follow the exponential distribution and are estimated as a ratio of the sum of transitions to another state to the cumulative number of hours spent in a given state.

The degradation process follows CTMC until the time of inspection, when a decision about maintenance action is undertaken. A decision about repair is instantaneous, according to a probability matrix of Deterministic Time Markov Chain, and allows the transition of the system to one of logical states showed in Fig. 1.

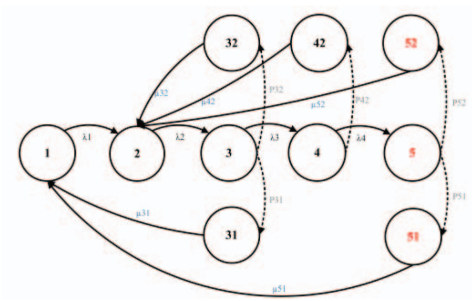

Fig. 1. Multiphase Markov model including maintenance. Dashed lines indicate repair decisions. Logical states are described by double digits, for instance in state 32, the first digit denotes that the physical degradation is at level 3, and the second digit says that the valve condition will be improved to state 2.

Logical states represent a repair decision, i.e. to which state the valve condition will be restored. Repair states are defined as logical because from the physical point of view the level of valve degradation is identical as in the state before decision about repair. After an instantaneous repair decision, valve being in the logical state can be repaired according to assigned repair rate, but it can also deteriorate further according to the same transition rate as the one assigned to its parent state.

3.2 *Phase-type expansion of the existing model*

In this work, in order to reflect deterioration of the system, sojourn times in main states (1 to 5) are assumed to follow Weibull distributions. The Weibull distribution is known for its flexibility, and therefore the capacity to model many types of failure rate behaviours (Rausand, 2014).

It is worth to mention that the application of PHD doesn't require specification of any distribution. In such a case a PHD estimation is based directly on the failure data or empirical moments. However, for the considered use case, the data is interval censored due to periodic condition monitoring. It means that exact interarrival times remain unknown. In addition, due to short time in observation there are very few observed transitions between model states, what also decreases the applicability of data-based estimation methods.

For five main states of the proposed degradation model, there will be four Weibull distributions to consider. The choice of parameters for these distributions was made partially based on observed data and partially on the assumption related to deterioration processes typical for ESVs. Weibull distribution is the extension of the exponential distribution where the shape parameters accounts for the trend in the lifetime. In the exponential case, there is no aging and shape parameter is equal 1. When the system experiences aging, the shape parameter is larger than 1. The faster proceeds the aging, the larger is shape parameter. In this model the shape parameter α is assumed to be the same for all distributions, α=3 for medium aging. Such a value is assumed for the situation where 2-3 failure mechanisms can lead to failure. (Vatn, 2007). Then having assumed value of the shape parameter and knowing average sojourn time in each state, estimated from the data, the scale parameter of considered Weibull, β is derived from the following formula:

$$E[X] = \frac{1}{\beta}\Gamma\left(\frac{1}{\alpha}+1\right) \quad (4)$$

where E[X] is the expected value of Weibull distribution.

Estimated parameters of Weibull distributions describing sojourn times in main states of the degradation Markov model are presented in Table 1. The table contains also parameters for exponential transition rates between main states applied in the previously proposed model.

Table 1 Parameters of distributions of transition times between main degradation states in two considered models: "classical" CTMC and phase espanded Markov model.

	Parameters	
State	Exponential transition rates	Theoretical Weibull
1-2	λ_1 =0 0.001	[β, α] = [11198, 3]
2-3	λ_2 =0 0.001	[β, α] = [11198, 3]
3-4	λ_3 =0 0.007	[β, α] = [1600, 3]

4-5	λ_4 =0 0.005	[β, α] = [2240, 3]

As it was mentioned already, main issues related to phase-type modelling relies on the choice of model structure, number of phases and the estimation number. Acyclic phase-type distribution has been chosen for approximation of Weibull distribution due to the simplicity of parameters estimation. Applied APHD are of order two, what is denoted by: APHD (2) and means that given Weibull distribution is approximated by a CTMC that consists of two transient and one absorbing state. From the perspective of existing model, it means that there is only one intermediate state added between every two main states. Therefore, the transition rate between state 1 and 2 is modelled by APHD (2) where one additional state 1b is added, so 1 (now 1a) and 1b are transient states and state 2 is the absorbing state. Of course, the increase of the number of phases by adding more intermediate states would provide better approximation of assumed Weibull distribution. However, from the practical point of view, the application of APHD of order two allows to gain a significant reduction of variance compare to classical exponential distribution, while keeping the model structure relatively simple.

The *R mapfit* package written by Okamura and Dohi (2015) is used for model estimation. The applied estimation approach is MLE baed on a theoretical density function defined on the positive domain [0, ∞).

The final step to derive phase expanded degradation model is to branch four CTMCs resulting from the performed estimation.

Each of obtained APHDs consists of two parameters: transition rates in a matrix form and initial probability vector $\boldsymbol{\pi}_i$ for *i*th APHD. When merging CTMCs one must remember to preserve those initial probabilities. This is done by allowing transitions between intermediate states. The idea of branching is explained in Fig 2. where two APHD (2) are merged. Because estimated transition rates must remain unchanged during merging, the exit rate of the "preceding" APHD (λ_{1b}) is multiplied by the initial probability vector of the "next" APHD ($\boldsymbol{\pi}_2$). This vector consists of two values: p_2 – probability of staring in state 2*a* and 1- p_2 – probability of starting in state 2*b*. This operation doesn't change anything for the "preceding" model, because from its perspective next states (2*a* and 2*b*) can be seen as one absorbing state (2). From the perspective of the "next" APHD, initial probabilities for state 2*a* and 2b are equal to values estimated by the *Rmapfit* algorithm, so

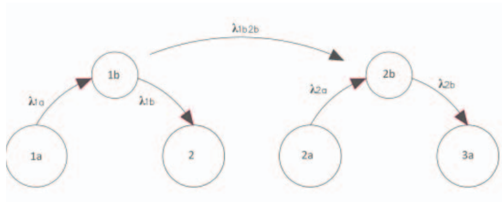

Fig. 2 Branching of two CTMCs resulting from APHD approximation of Weibull distributed transition times between main states of the "classical" Markov model. States 1a and 2a correspond to main degradation states 1 and 2 from the previous model. States 1b and 2b are artificial intermediate states added in the result of phase-type approximation.

the estimated distribution is maintained. See appendix A for the proof and details about branching of processes. The whole phase type degradation model derived in the described way is presented by the graph in Fig.3

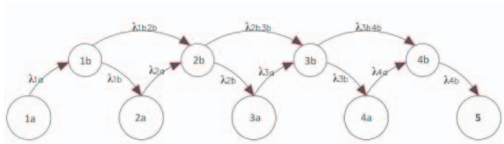

Fig.3 The phase-type degradation model resulting from phase expansion of the existing degradation model by application of APHD distributions to model Weibull distributed transition times between main degraded states (bottom states).

3.3 *Maintenance in the phase-type model*

The inclusion of maintenance policy to this model is performed in a similar way as it has been done in the previous work. The graph for phase-type degradation model including maintenance is presented in Fig. 4. Because it is assumed that the system can experience further degradation from logical states, it is necessary to include decision states for both main and intermediate states (for instance states 32*a* and 32*b* on the graph in Fig 4.). This is equal to the assumption that intermediate states are likely to be discovered during an inspection. Assuming that during an inspection a small degradation of an ESV is discovered imagine that a maintenance expert can distinct whether the level of this degradation is smaller or larger.

Finally, the application of intermediate states enables modelling of the repair efficiency. Each logical state provides the information to which level the condition of the valve will be improved. Because now each main state, except of the absorbing one, has assigned one intermediate

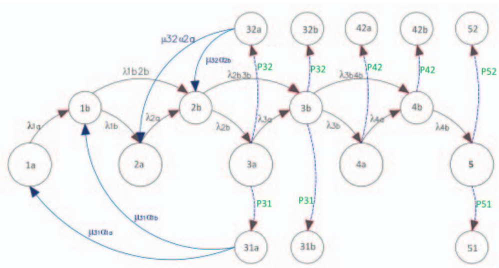

Fig.4 Phase-expanded degradation model including maintenance strategy.

state, it is possible that a repair is imperfect, and the valve is restored to an intermediate state, instead of the desirable main state. This is reflected by multiplying considered repair rate by the initial probability of suitable phases (blue continuous lines on the graph in Fig 4.). In order to keep the model legible only few repair transitions are depicted on the graph. The model resulting from inclusion of maintenance is quite complex and has a total of 17 states (see Fig 4.). Therefore, modelled repair rates remain exponentially distributed. Nevertheless, the application of phase expansion for modelling of repair transitions could reflect their deterministic nature.

Although the Markov graph is rather complex, once transitions between systems' states are in a matrix form, the derivation of states probabilities is straightforward as it consists on the integration of numerical equations: Eq. (5) in case of the continuous degradation:

$$\mathbf{P}(t+dt) = \mathbf{P}(t)\cdot(\mathbf{A}dt + \boldsymbol{I}) \qquad (5)$$

and Eq. (6) at inspection times τ:

$$\mathbf{P}(t^{+}) = \mathbf{P}(t)\cdot(\mathbf{A}\mathrm{t} + \mathrm{I})\cdot \boldsymbol{A_2} \qquad (6)$$

where $\mathbf{P}$ is the probability vector for considered states, $\mathbf{A}$ is the transition rate matrix and $\boldsymbol{A_2}$ is a matrix with probabilities for repair decisions.

4. Results

4.1 *Phase type approximation – simple case*

In this section the phase-type modelling of a Weibull distributed sojourn time for one state is investigated. Under the assumption of keeping the expected interarrival time unchanged with regard to the previous ("classical") model, the changes in variances for all proposed distributions are discussed.

Fig 5. shows the approximation of Weibull distribution by an acyclic phase type distribution of different orders. APHD (5) provides quite good approximation, especially on the left side of the distribution. The left tail of the distribution is the most important area, because the first passage time is the value of interest when it comes to the modelling of lifetime. When applying an acyclic phase type distribution of order 2, the quality of fit declines a lot compare to the fit by APHD (5). Although the left tail approximated by APHD (2) takes much larger values than the original Weibull distribution, it is still much better than the exponential distribution, what is visible in a significant reduction of variance in case of the APHD(2) comparing with the exponential distribution. When applying the acyclic phase type distribution for a single sojourn time, the number of intermediate states depends on APHD order, but transition rates for all phases are constant. In the previously used model, the transition rate from state 1 was equal 0.0001 (see Table 1). The transition rates of APHD (2) are almost identical being equal around 0.0002. In case of APHD (5) transition rates are also similar to each other and the obtained values are equal 0.0005. Such results of estimation look like scaling of the "original" rate according to the number of phases in the APHD approximation.

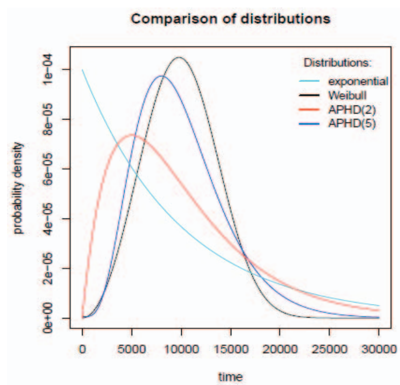

Fig. 5. Probability density functions of Weibull and exponential distributions with the same mean value. The Weibull distribution is approximated by two acyclic phase type distributions of order 2 and 5 accordingly.

4.2 *Phase type approximation – MTTF*

Two degradation models: the "classical" CTMC (Laskowska and Vatn, 2019) and phase type model are compared with regard to Mean Time To Failure (MTTF). Because of the assumption that expected interarrival times are the same in both models, the comparison of expected failure times allows to examine the accuracy of the model proposed in this study. MTTF is expressed by eq. (7), where $P_5(t)$ denotes the probability of being in the failed state at time t.

$$\int_{t=0}^{\infty} \left(1 - F\,(t)\right) dt = \int_{t=0}^{\infty} \left(1 - P_5\,(t)\right) dt \qquad (7)$$

Fig. 6 presents a comparison of expected failure times modelled by a classical CTMC and by phase expanded model. It is visible that both MTTFs are almost identical, what means that the phase type model is correct. As expected, the variance of failure time is smaller in the phase type model.

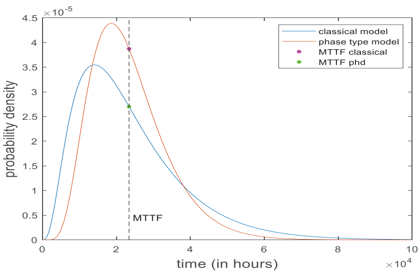

Fig. 6. Probability distributions of time to failure of an ESV, modelled by 5 state CTMC (classical model) and its phase expansion with 9states (phase type model).

4.3 *Phase type approximation – PFD$_{avg}$*

In this section models including maintenance are compared. In case of the safety valves the Probability of Failure on Demand, PFD$_{avg}$, is used as a measure of the system reliability. PFD$_{avg}$ is the average probability that the system cannot respond on a demand and is expressed by t PFD$_{avg}$s were derived for both models for three different inspection intervals: τ = 2, 6 and 8 months. The choice of these intervals is discussed by Laskowska and Vatn (2019).

$$\mathrm{PFD_{avg}} = \frac{1}{\lim\limits_{t \to \propto} t} \int_{t=0}^{\propto} \left(P_5(t)\right) dt \qquad (8)$$

where P_5 (t) is the probability of being in the failed state. Because of periodic inspection the PFD$_{avg}$ does not converge to one value, but it has been checked that for the large t its value stabilizes oscillating around the values provided in Table 2 for different inspection intervals.

As it is shown in Fig. 7 the PFD$_{avg}$ derived from the phase type model is smaller than PFD$_{avg}$ provided by the “classical” model. This is due to the fact that the model with more phases becomes more deterministic. Therefore, for the same time values, the probability of arriving to failed state is smaller. It can be interpreted as a shift of the first hitting time in right direction. Of course, this decrease of the density on the left tail will be compensated by its higher value in the middle of the distribution, but before the density increases, the inspection of valve is already performed. The density plots for different inspection intervals depict that situation quite well. For the shortest inspection interval, the PFD$_{avg}$ obtained by the application of PH model is three times smaller than in case of the “classical” model. However, with the increase of the interval, the differences between calculated PFD$_{avg}$s are smaller. The exact values of PFD$_{avg}$ for both models are presented in Table 2.

Table 2 Comparison of PFDavg values for both models for different inspection intervals

	PFD$_{avg}$	
τ [days]	Classical model	PHD model
65	0.0039	0.0013
183	0.0345	0.0238
240	0.0552	0.0442

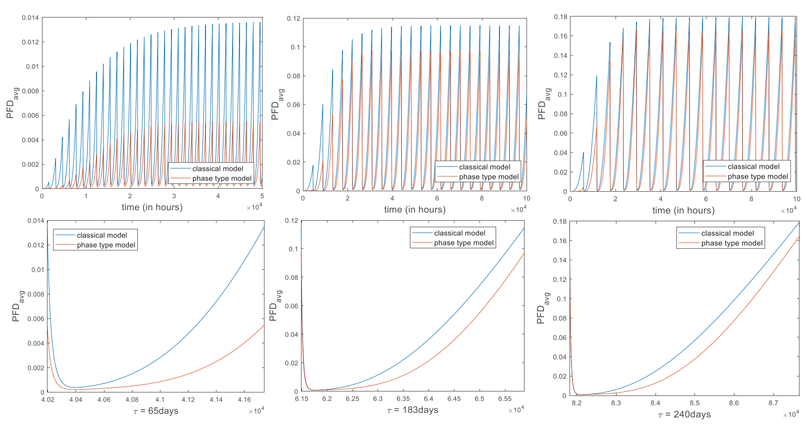

Fig. 7 PFD$_{avg}$ of an ESV for different inspection lengths: τ= 2, 6 and 8 months, shown in longer time perspective (top plots) and during one interval τ (bottom plots).

5. Conclusions

There exist several papers describing application of PHDs to model failure data. The contribution of this work relies on the enhancement of the degradation modelling of physical systems. The proposed model captures the impact of the aging on the deterioration of technical systems. This is done by the extension of existing homogeneous Markov model with intermediate states. The state expansion allows modelling of Weibull distributed sojourn times in degradation states. In that way the model becomes more realistic reflecting aging of the system. However, the PHD approximation which consists on adding intermediate states between main degradation states allows maintaining the memoryless property of the newly arisen Markov.

The phase expansion of the previously built degradation model allows a significant reduction of the variance in the distribution of time to failure of considered ESV. This in turn, by the reduction of the uncertainty leads to a decrease of derived PFD_{avg} what enables extending testing intervals, while still meeting safety requirements defined by IEC 61508. In addition, the reduction of uncertainty about valve state allows to capture the valve condition more precisely, and select a proper maintenance action for the investigated ESV. The improvement of the condition-based maintenance by the optimization of testing intervals and repair policy enables more cost-effective valve operation while providing required safety level.

Appendix A. Merging two APHDs by applying the convolution theorem.

Consider a system depicted in Fig. 2 where state 3 in an absorbing failed state. Assume that sojourn times in state 1 and state 2 could be approximated by phase type distributions, i.e. $\tilde{T}_1 \sim F_1(t)$ and $\tilde{T}_2 \sim F_2(t)$. The sojourn times are assumed to be stochastically independent. For $\tilde{T}_1$ the phase type model comprises two transient states: 1a and 1b and one absorbing state: 2. The probability that the system starts in state 1a is p_1, and the probability that the system starts in state 1b is 1-p_1. All transition rates are constant. Given $F_1(t)$ and $F_2(t)$ the distribution of the time to failure for this system can be found by the convolution theorem. For the phase-type distribution approximating the first sojourn time there is a rate λ_{1b} from state 1b to an absorbing state outside the system. This transition is split into two transitions, one to state $2a$ and one to state $2b$. The corresponding rates are $\lambda_{1b2a} = p_2\lambda_{1b}$ and $\lambda_{1b2b} = (1-p_2)\lambda_{1b}$ respectively. Let $P_3(t)$ be the probability that the system is in state 3 at time t. $P_3(t)$ can easily be obtained by integration of the Markov equation for the compound system yielding the cumulative distribution function for time to system failure.

The idea behind the proof is to first find the cumulative distribution function for the time to system failure without making any assumption regarding how to link the two phase-type distributions. Then the procedure is repeated when the two phase-type distributions are linked in the proposed manner. If this gives the same result, the proposition is valid. First define $\tilde{T}_{2a}$ and $\tilde{T}_{2b}$ to be the sojourn times for states $2a$ and $2b$ respectively. It follows that:

$$\Pr(\tilde{T}_{2a} \le t) = p_2 \Pr(\tilde{T}_{2a} + \tilde{T}_{2b} \le t) + (1-p_2)\Pr(\tilde{T}_{2b} \le t) \qquad (A.1)$$

Let T be the time to failure. Now assume that the first sojourn time $\tilde{T}_1 = x$. It follows that:

$$\Pr(T = \tilde{T}_2 + x \le t | \tilde{T}_1 = x) = p_2 \Pr(\tilde{T}_{2a} + \tilde{T}_{2b} + x \le t) + (1-p_2)\Pr(\tilde{T}_{2b} + x \le t) \qquad (A.2)$$

Eq. (A.2) does not make any assumption regarding how the two phase-type models are linked. It just prescribes a way to obtain the cumulative distribution function (CDF) for the system failure time, given that the first sojourn time is x. To obtain the unconditional CDF we integrate over all possible x-values, $\Pr(T \le t)$:

$$\int_{t=0}^{\infty} \Pr(\tilde{T}_{2a} \le t - x)\, P_{1b}(x)\lambda_{1b}dx \qquad (A.3)$$

where $P_{1b}(x)$ is the time dependent probability that the first phase type system is in state $1b$ at time x. Using eq. (A.2) we get:

$$\Pr(T \le t) = \int_0^{\propto} p_2 \Pr(\tilde{T}_{2a} + \tilde{T}_{2b} \le t - x)\, P_{1b}(x)\lambda_{1b}dx + \int_0^{\propto} (1-p_2)\Pr(\tilde{T}_{2b} \le t - x)\, P_{1b}(x)\lambda_{1b}dx \quad (A.4)$$

Now, consider the compound model shown in Fig. 2. In the compound model there are two ways we may leave state $1b$, i.e., there is a transition from the "physical" state 1 to the "physical" state 2. These two possibilities are either going from state $1b$ to state $2a$ or going from state 1b to state $2b$.

Assume there is a transition from state $1b$ to state $2a$ at time x. Let this event be denoted A_x, giving the conditional CDF:

$$\Pr(T \le t | A_x) = \Pr(\tilde{T}_{2a} + \tilde{T}_{2b} \le t - x) \quad (A.5)$$

Similarly, let B_x be the event that there is a transition from state 1b to state $2b$ at time x, giving the conditional CDF:

$$\Pr(T \leq t|B_x) = \Pr(\tilde{T}_{2b} \leq t - x) \qquad (A.6)$$

Since A_x and B_x are disjoint we integrate over all possible values to get the unconditional CDF:

$$\Pr(T \leq t) = \int_0^{\infty} \Pr(\tilde{T}_{2a} + \tilde{T}_{2b} \leq t - x)\, P_{1b}(x)\lambda_{1b2a}dx + \int_0^{\infty} \Pr(\tilde{T}_{2b} \leq t - x)\, P_{1b}(x)\lambda_{1b2a}dx \qquad (A.7)$$

which is the same as Eq. (4) provided: $\lambda_{1b2a} = p_2\lambda_{1b}$ and $\lambda_{1b2b} = (1 - p_2)\lambda_{1b}$. From the presented proof results following given two independent sojourn times approximated with two phase type-distributions having the structure presented in Fig. 2, the two models can be combined as illustrated in the figure, yielding the correct CDF for the sum of these two sojourn times. It is believed that a similar proof could be performed for the general case with more than two "physical" states. The proof is independent on how good the approximation of each sojourn time is.

References

Aalen, O. 1995. "Phase Type Distributions in Survival Analysis." *Scandinavian Journal of Statistics* 22 (4):447-463.

Asmussen, S., O. Nerman, and M. Olsson. 1996. "Fitting Phase-Type Distributions via the EM Algorithm." *Scandinavian Journal of Statistics* 23 (4):419-441.

Buchholz, P., and J. Kriege. 2014. "Markov Modeling of Availability and Unavailability Data." 2014 Tenth European Dependable Computing Conference, 13-16 May 2014.

Buchholz, P., J. Kriege, and I. Dohndorf. 2014. *Input modeling with phase-type distributions and Markov models. Theory and applications*

Faddy, M. J. 1995. "Phase-type distributions for failure times." *Mathematical and Computer Modelling* 22 (10):63-70.

International Electrotechnical Commission. 2005. "Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems (E/E/PE, or E/E/PES) "

Laskowska, E., and J. Vatn. 2019. "State Modelling and Prognostics of Safety Valves used in the Oil and Gas Industry." Proceedings of the 29th European Safety and Reliability Conference(ESREL), Hannover, Germany., 22-26 September 2019.

Lindqvist, B. H. 2016. "Phase-Type Models for Competing Risks." 2016 Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management (SMRLO), 15-18 Feb. 2016.

Montoro-Cazorla, D., and R. Pérez-Ocón. 2006. "A deteriorating two-system with two repair modes and sojourn times phase-type distributed." *Reliability Engineering & System Safety* 91 (1):1-9.

Neuts, Marcel F. 1975. "Computational uses of the method of phases in the theory of queues." *Computers & Mathematics with Applications* 1 (2):151-166.

Okamura, H., and T. Dohi. 2015. *mapfit: An R-Based Tool for PH/MAP Parameter Estimation*

Okamura, H., and T. Dohi. 2016a. "Fitting Phase-Type Distributions and Markovian Arrival Processes: Algorithms and Tools." In *Principles of Performance and Reliability Modeling and Evaluation: Essays in Honor of Kishor Trivedi on his 70th Birthday*, edited by Lance Fiondella and Antonio Puliafito, 49-75. Cham: Springer International Publishing

Okamura, H., and T. Dohi. 2016b. "Ph fitting algorithm and its application to reliability engineering." *Journal of the Operations Research Society of Japan* 59:72-109.

Pérez-Ocón, R., and J. E. Ruiz Castro. 2004. "Two models for a repairable two-system with phase-type sojourn time distributions." *Reliability Engineering & System Safety* 84 (3):253-260.

Rausand, M. 2014. *Reliability of safety-critical systems: theory and applications*: John Wiley & sons

Vatn, J. 2007. "Veien frem til "World Class Maintenance": Maintenance Optimisation." In: Norwegian University of Science and Technology (NTNU). http://folk.ntnu.no/jvatn/pdf/MaintenanceOptimisationWCM.pdf.

Yeh, R. H. 1997. "Optimal inspection and replacement policies for multi-state deteriorating systems." *European Journal of Operational Research* 96 (2):248-259.