A Generic Approach to Model Complex System Reliability using Graphical Duration Models

Roland Donat	Laurent Bouillaut	Patrice Aknin	Philippe Leray	Didier Levy
INRETS - LTN	INRETS - LTN	INRETS - LTN	INSA Rouen - LITIS	RATP - EST Voie
Av. Malleret-Joinville	Av. Malleret-Joinville	Av. Malleret-Joinville	Av. de l'Université	50 rue R. Salengro
Arcueil	Arcueil	Arcueil	St-Étienne-du-Rouvray	Fontenay s/s Bois
94114 France	94114 France	94114 France	76801 France	94724 France
donat@inrets.fr	*bouillaut@inrets.fr*	*aknin@inrets.fr*	*pleray@insa-rouen.fr*	*didier.levy@ratp.fr*

Abstract

Nowadays, reliability analysis has become an integral part of system design and operating. This is especially true for systems performing critical tasks. Moreover, recent works in reliability involving the use of probabilistic graphical models, also known as bayesian networks, have been proved relevant.

This paper aims to describe a general methodology to model the stochastic degradation process of a complex system, allowing any kind of state sojourn distributions along with an accurate context description. We meet these objectives using a specific dynamic graphical model, namely a graphical duration model. In this article, we give qualitative and quantitative descriptions of the proposed model and describe a simple algorithm to estimate the system reliability and some of its related metrics.

Finally, we illustrate this approach by applying our methodology to a three-states system subjected to one context variable and with non exponential duration distributions.

1 Introduction

Nowadays, reliability analysis has become an integral part of system design and operating. This is especially true for systems performing critical applications. Typically, the results of such analysis are given as inputs to a decision support tool in order to optimise the maintenance operations. Unfortunately, in most of cases, the system state cannot be evaluated exactly. Indeed, it is uncommon to be able to deterministically describe the way a complex system reaches a failure state. This is one of the reasons which has led to the important development of probabilistic methods in reliability.

A wide range of works about reliability analysis is available in the literature. For instance in numerous applications, the aim is to model a multi-states system and therefore to capture how the system state changes over time. This problematic can be partially solved using the Markov framework (Aven and Jensen 1999). The major drawback of this approach comes from the constraint on state sojourn times which are necessarily exponentially distributed. This issue can be overcome by the use of semi-Markov models (Limnios and Oprisan 2001) which allow to consider any kind of sojourn time distributions. On the other hand, one can be interested in modeling the context impacting on the system degradation (Kalbfleisch and Prentice 2002). A classic manner to address such an issue consists in using a Cox model (Cox 1972) or a more general proportional hazard model (Kay 1977). Nevertheless, as far as we know, it is unusual to find works considering both approaches at the same time.

Moreover, recent works in reliability involving the use of Probabilistic Graphical Models (PGMs), also known as Bayesian Networks (BNs), have been proved relevant. For instance, Boudali and Dugan (2005) show how to model a complex system dependability by mean of PGMs. Langseth and Portinale (2006) explains hows fault trees can be represented by PGMs. Finally, Weber and Jouffe (2003) show how convenient Dynamic PGMs (DPGMs) are to study the reliability of a dynamic system represented by a Markov chain. Our work aims to describe a general methodology to model the stochastic degradation process of a system, allowing any kind of state sojourn distributions along with an accurate context description. We achieve to meet these objectives using a specific DPGM called Graphical Duration Model (GDM).

This paper is divided into five sections. Section 2 briefly describes the PGMs and DPGMs theory. Then, section 3 introduces the GDMs and explains how to model the reliability of complex systems using our proposed graphical approach. Section 4 depicts a simple recursive method aiming to estimate the reliability and some of its related metrics. To illustrate our methodology, we propose to study a three-states system subjected to one context variable and with non exponential duration distributions in section 5. Finally, some conclusions and perspectives are discussed in section 6.

2 Probabilistic Graphical Models

2.1 Definition

Probabilistic Graphical Models (PGMs), also known as Bayesian Networks (BNs) (Jensen 1996), are mathematical tools relying on the probability theory and the graph theory. They allow to qualitatively and quantitatively represent uncertain knowledge. Basically, PGMs are used to describe in a compact way the joint distribution of a set of random variables $\{X_1, \ldots, X_N\}$. Formally, a PGM, denoted by $\mathcal{M}$, is defined as a pair $\left(\mathcal{G}, \{p_i\}_{i=1}^N\right)$, where

- $\mathcal{G} = (X, E)$ is a Directed Acyclic Graph (DAG). $X = \{X_1, \ldots, X_N\}$ is a set of nodes representing random variables and E is a set of edges encoding the conditional independence relationship between the variables in the model. Thus, $\mathcal{G}$ is the qualitative description of $\mathcal{M}$.

- $\{p_i\}_{i=1}^N$ is a set of Conditional Probability Distributions (CPDs) aiming to describe the quantitative aspect of the model. It is worth noting that if the random variable X_i takes its values in a finite and countable set $\mathcal{X}_i$ (e.g. $\mathcal{X}_i = \{x_i^1, \ldots, x_i^K\}$), the CPD of X_i can be defined by a Conditional Probability Table (CPT). On the other hand, if $\mathcal{X}_i$ is an infinite set such as $\mathbb{R}$, then p_i is a conditional density function.

The underlying conditional independence assumptions introduced by this modelling allows to economically rewrite the joint probability distribution :

$$\mathbb{P}(X_1 = x_1, \ldots, X_N = x_N) = \prod_{i=1}^{N} \mathbb{P}(X_i = x_i | X_{\pi_i} = x_{\pi_i}) = \prod_{i=1}^{N} p_i(x_i, x_{\pi_i}), \quad (1)$$

where π_i denotes the subscripts of the i'th variable parents in the graph. Thereby, it is important to remark that X_{π_i} is no longer a single random variable but a set of random variables containing the parents of X_i in the graph $\mathcal{G}$. More formally, all conditional independence statements can be read off the graph structure by using the rule of d-separation (Pearl 1988).

Besides, both the qualitative and quantitative parts of a PGM can be automatically learnt (Neapolitan 2003) if some data or experts' opinions are available. The latter problem can be boiled down to a probability distribution estimation. We consider that the CPDs of the model are the $\{p_{\theta_i}\}_{i=1}^N$ where $\theta_i \in \Theta_i$ represents the parameters of the i'th CPD. Thereby, the objective is to estimate the set $\{\theta_i\}_{i=1}^N$ using the available knowledge. A classic manner to tackle this problem is to use the Maximum Likelihood (ML) method to exploit information in databases. If some a priori information is also available (e.g. experts knowledge), one can use bayesian methods (Gelman et al. 2003) and compute Maximum A Posteriori (MAP) estimates of the parameters. To that end, it is possible to use the factorisation property of PGMs, so that each θ_i can be locally estimated using only the i'th CPD.

Using PGMs is also particularly interesting because of the possibility to propagate knowledge through the network. Indeed, various inference algorithms can be used to compute marginal probability of the system variables. The most classical one relies on the use of a junction tree. Inference in PGMs (Huang and Darwiche 1996) allows to take into account any variable observations (also called evidence) so as to update the marginal distribution of the other variables. Without any evidence, the computation is based on a priori distributions. When evidence is given, this knowledge is integrated into the network and all the marginal distributions are updated accordingly.

Finally, it is important to notice that such modelling are unable to model the dynamic of a non stationary system. For instance, in reliability analysis, one can be interested in modelling how a system changes from an "up" state to a "down" state over time. For this kind of problem a possible solution consists in using the dynamic extension of PGMs which are presented in the next section.

2.2 Dynamic probabilistic graphical models

Dynamic Probabilistic Graphical Models (DPGMs), also known as Dynamic Bayesian Networks (DBNs) are convenient tools to represent complex dynamic systems. The term "dynamic system" makes reference to a system of which state can change over time but with a fixed graph structure. To that end, DPGMs allow variables to have temporal (or sequential) dependencies.

Strictly speaking, a DPGM is a way to extend PGM to model probability distributions over a collection of random variables $(X_{1,t}, \ldots, X_{N,t})_{t \in \mathbb{N}^*}$. A DPGM $\mathcal{M}_D$ is defined (Murphy 2002) to be a pair $(\mathcal{M}_1, \mathcal{M}_\rightarrow)$ where

- $\mathcal{M}_1$ is a PGM which defines the prior distribution $\mathbb{P}(X_{1,1}, \ldots, X_{N,1})$ as in equation (1).
- $\mathcal{M}_\rightarrow$ is a s-slice Sequential Probabilistic Graphical Model (s-SPGM), also named s-slice Temporal Bayes Net (s-TBN) in the literature. s makes reference to the temporal dependence order of the model. In this paper, we will limit ourselves to the case of a 2-SPGM which means that the present (slice t) is only dependent on the one step past (slice $t-1$). It is worth noting that we would rather use the more generic term "sequential" instead of "temporal". Indeed, in many applications (e.g. genetic or reliability), the dynamic of the studied system does not necessarily rely on time (e.g. nucleotides sequence in DNA, number of mechanical solicitations for a device).

  Basically, $\mathcal{M}_\rightarrow$ is also a PGM used to define the transition model which describes the dependencies between variables in slice $t-1$ and variables in slice t. It aims to specify the CPD $\mathbb{P}(X_{1,t}, \ldots, X_{N,t} | X_{1,t-1}, \ldots, X_{N,t-1})$ taking advantages of the factorization property in PGMs :

  $$\mathbb{P}(X_{1,t}, \ldots, X_{N,t} | X_{1,t-1}, \ldots, X_{N,t-1}) = \prod_{i=1}^{N} \mathbb{P}(X_{i,t} | X_{\pi_i,t}),$$

  where $X_{\pi_i,t}$ is a set of the parents of $X_{i,t}$ which could contain variables in the slices t and $t-1$.

  Note that if the first slice of $\mathcal{M}_\rightarrow$ is identical to $\mathcal{M}_1$, the latter can be omitted and the DPGM is strictly equivalent to a 2-SPGM.

Then, it is possible to deduce the distribution $\mathbb{P}\left((X_{1,t}, \ldots, X_{N,t})_{1 \leq t \leq T}\right)$ by "unrolling" the 2-SPGM until we have a sequence of length T :

$$\begin{aligned}
\mathbb{P}\left((X_{1,t}, \ldots, X_{N,t})_{1 \leq t \leq T}\right) &= \mathbb{P}(X_{1,1}, \ldots, X_{N,1}) \prod_{t=2}^{T} \mathbb{P}(X_{1,t}, \ldots, X_{N,t} | X_{1,t-1}, \ldots, X_{N,t-1}) \\
&= \prod_{i=1}^{N} \mathbb{P}(X_{i,1} | X_{\pi_i,1}) \prod_{t=2}^{T} \prod_{i=1}^{N} \mathbb{P}(X_{i,t} | X_{\pi_i,t}).
\end{aligned}$$

The fact that t is an integer means we only consider discrete stochastic processes. This restriction is not very penalizing in survival analysis because the duration variable is often expressed as an integer (e.g. number of hours, days, years, solicitations... before the failure hitting moment).

Finally, as it is possible to consider a DPGM as a big unrolled PGM, it is clear DPGM inherits the convenient properties of classic PGM. Indeed, learning methods do not change when using DPGM. Concerning the inference problem, most of the methods are based on static PGM inference algorithms (Murphy 2002).

3 Introducing the Graphical Duration Models

Recent works in reliability involving the use of PGMs have been proved relevant since they are particularly suitable to model the dynamic of a complex system (Boudali and Dugan 2005). Nevertheless in the field of reliability analysis and as far as we know, no author seems to have gone far the use of DPGM to represent more complex models than a Markov process with exogenous constraints (Weber and Jouffe 2003).

In this article, we propose to extend the variable duration model introduced by Murphy (2002) to build a comprehensive model for complex survival distributions.

3.1 Qualitative definition

The proposed model, which we denote by Graphical Duration Model (GDM), is depicted in figure 1 as a 2-SPGM. This model allows to describe, in a flexible and accurate way, the behaviour of a complex system given its context. Indeed, three different parts are considered :

1. The system state (X_t) over the sequence.
2. The covariates $(Z_{1,t}, \ldots, Z_{P,t})$ used to describe the system context.
3. The duration variable (X_t^D) describing how long the system remains in a specific state.

Moreover, a transition variable (J_t) is added to explicitly characterise when the system jumps into another state. Indeed, if $J_t = 1$, it means that the system state will change at time $t + 1$. On the other hand, while $J_t = 0$, the system remains in the same state. This variable is not necessary but appears to be convenient for further generalization and CPDs definition. In addition for the sake of readability, we denote by $\mathbf{Z}_t$ the random variable vector $(Z_{1,t}, \ldots, Z_{P,t})$ and by $\mathbf{z}_t$ an observation of Z_t, in other words $\mathbf{z}_t = (z_{1,t}, \ldots, z_{P,t})$ with $z_{p,t}$ an observation of the variable $Z_{p,t}$.

Then, it is possible to factorise the joint probability distribution of the variables $X_t, X_t^D, J_t, \mathbf{Z}_t$ over a sequence of length T as it follows

$$\begin{aligned}\mathbb{P}\left((\mathbf{Z}_t, X_t, X_t^D, J_t)_{1\leq t\leq T}\right) &= \left[\prod_{p=1}^{P} \mathbb{P}(Z_{p,1})\right] \mathbb{P}(X_1|\mathbf{Z}_1)\mathbb{P}(X_1^D|X_1, \mathbf{Z}_1)\mathbb{P}(J_1|X_1^D) \\ &\times \prod_{t=2}^{T}\left[\prod_{p=1}^{P} \mathbb{P}(Z_{p,t})\right] \mathbb{P}(X_t|X_{t-1}, J_{t-1}, \mathbf{Z}_t)\mathbb{P}(X_t^D|X_{t-1}^D, J_{t-1}, \mathbf{Z}_t, X_t)\mathbb{P}(J_t|X_t^D).\end{aligned} \quad (2)$$

Besides, in this model the system state transition depends on the duration spent in the current system state and the current state itself. Thus, we are in a discrete semi-markovian approach (Barbu et al. (2004); Barbu and Limnios (2006)). Indeed, we can specify any kind of state sojourn time distribution by contrast with a classic markovian approach in which all durations have to be exponentially distributed.

This modelling is particularly interesting since it allows to take into account complex degradation distributions and context effects at the same time.

3.2 CPDs definition

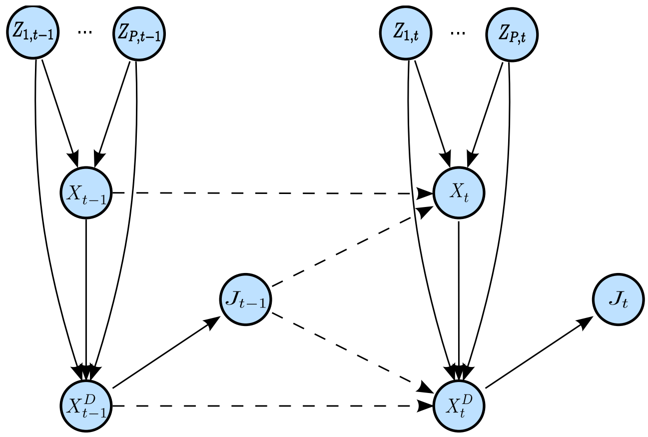
The following paragraph addresses the specification of each CPD involved in equation (2) characterising the joint probability distribution in a Graphical Duration Model (GDM).

3.2.1 Covariates PDFs

We suppose that each covariate $Z_{p,t}$ is identically distributed over time and takes its values in the set $\mathcal{Z}_p$, so that $\mathbf{Z}_t$ is defined over the set $\mathcal{Z} = \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_P$. As the covariates do not have any parent in the model, their Probability Distribution Function (PDF) is not conditional and for each p we have :

$$\mathbb{P}(Z_{p,1} = z_p) = \ldots = \mathbb{P}(Z_{p,t} = z_p) = \ldots = \mathbb{P}(Z_{p,T} = z_p) = p_{Z_p}(z_p).$$

Figure 1: Representation of a GDM. The $Z_{p,t}$'s represent the system covariates. X_t is the system state and X_t^D is the duration variable in the current state. J_t is the explicit transition variable of the system. A dotted arrow characterises a dependency between two slices.



3.2.2 System state CPDs

We make the assumption that the number of system states is finite and let $\mathcal{S} = \{1, \ldots, K\}$ be the set of the K different states. The first CPD concerns the distribution of the initial system state according to its context :

$$\mathbb{P}(\underbrace{X_1 = i}_{\text{initial state}} | \underbrace{\mathbf{Z}_1 = \mathbf{z}}_{\text{context}}) = p_{X_1}(\mathbf{z}, i),$$

where $p_{X_1}(\cdot, \mathbf{z})$ is a vector of K elements depending on the vector of values $\mathbf{z} = (z_1, \ldots, z_P)$.

The second state CPD concerns the dynamic transition from one state to another. As a transition occurs at time t if and only if the variable $J_{t-1} = 1$, we have

$$\mathbb{P}(\underbrace{X_t = j}_{\text{current state}} | \underbrace{X_{t-1} = i}_{\text{previous state}}, \underbrace{J_{t-1} = 1}_{\text{transtion at time } t}, \mathbf{Z}_t = \mathbf{z}) = A(i, \mathbf{z}, j), \tag{3}$$

where $A(\cdot, \mathbf{z}, \cdot)$ is a stochastic matrix of size $K \times K$ depending on the covariate values $\mathbf{z}$. In general, $A(\cdot, \mathbf{z}, \cdot)$ is called the transition matrix of the system.

On the other hand, if $J_{t-1} = 0$, the system remains in the same state. Therefore, the CPD is deterministic and the transition matrix has to be equal to identity :

$$\mathbb{P}(X_t = j | X_{t-1} = i, \underbrace{J_{t-1} = 0}_{\text{no transition}}, \mathbf{Z}_t = \mathbf{z}) = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

3.2.3 Duration CPDs

The first duration CPD describes the sojourn time distribution for each state and each combination of covariates. We denote by p_{X^D} this conditional duration distribution and we have :

$$\mathbb{P}(\underbrace{X_1^D = d}_{\text{sojourn time in } i} | \mathbf{Z}_t = \mathbf{z}, X_1 = i) = p_{X^D}(\mathbf{z}, i, d), \ d \in \mathbb{N}^*, \tag{4}$$

where $p_{X^D}(\mathbf{z}, i, d)$ represents the probability to stay during d time units in the state i given the context z. Besides, as we are studying discrete-time models, the duration unit d is a non negative integer. $p_{X^D}(\cdot, \mathbf{z}, \cdot)$ can be seen as an infinite matrix (along its first dimension) depending of covariates values z. In practice, it is possible to set an upper time bound D so that the time scale becomes finite. Indeed in this case, $d \in \{1, \dots, D\}$ and $p_{X^D}(\cdot, \mathbf{z}, \cdot)$ is a finite matrix of $K \times D$ elements.

The dynamic duration CPD aims to memorize the time spent in the current state. Indeed, if the previous remaining duration is greater than one, we update the remaining time by deterministically decreasing it by one unit.

$$\mathbb{P}(X_t^D = d | X_{t-1}^D = d', \underbrace{J_{t-1} = 0}_{\text{no transition}}, \mathbf{Z}_t = \mathbf{z}, X_t = i) = \begin{cases} 1 & \text{if } d = d' - 1 \text{ and } d' \geq 2, \\ 0 & \text{otherwise.} \end{cases}$$

If the previous remaining time reach the value one, a transition occurs at time t and the duration in the new current state is drawn according to the CPD p_{X^D}. In other words,

$$\mathbb{P}(X_t^D = d | \underbrace{X_{t-1}^D = 1, J_{t-1} = 1}_{\text{transition occurred}}, \mathbf{Z}_t = \mathbf{z}, X_t = i) = p_{X^D}(\mathbf{z}, i, d).$$

Let us note that the case $X_{t-1}^D = 1$ and $J_{t-1} \neq 1$ is not consistent. As a consequence the previous CPD is undefined for these values.

Finally, it is worth noting that the discrete-time assumption laid on duration distributions by the modelling can be easily overcome. Indeed, Bracquemond and Gaudoin (2003) present a survey of discrete lifetime distributions and describe some of them which derive from usual continuous ones (e.g. exponential, Weibull).

3.2.4 Transition CPD

J_t is a random variable characterising the occurrence of a system state change at time t. More precisely, when $J_t = 1$, a transition is triggered at time $t+1$ and the state remains unchanged while $J_t = 0$. Besides, a transition occurs at time $t+1$ if and only if the remaining duration in the current state at time t equals one. As a consequence, the CPD of variable J_t is deterministic and defined by

$$\mathbb{P}(J_t = 1 | X_t^D = d) = \begin{cases} 1 & \text{if } d = 1, \\ 0 & \text{otherwise.} \end{cases}$$

4 Reliability analysis using GDM

4.1 Basic definitions

In this section, we suppose that the set of system states $\mathcal{S}$ is partitioned into two sets $\mathcal{U}$ and $\mathcal{D}$ (i.e. $\mathcal{S} = \mathcal{U} \cup \mathcal{D}$ with $\mathcal{U} \cap \mathcal{D} = \varnothing$), respectively for "up" states and for "down" states (i.e. OK and failure situations). The system transition matrix from equation (3) can be decomposed as follows :

$$A(\cdot, \mathbf{z}, \cdot) = \begin{pmatrix} A_{\mathcal{U}\to\mathcal{U}}(\cdot, \mathbf{z}, \cdot) & A_{\mathcal{U}\to\mathcal{D}}(\cdot, \mathbf{z}, \cdot) \\ A_{\mathcal{D}\to\mathcal{U}}(\cdot, \mathbf{z}, \cdot) & A_{\mathcal{D}\to\mathcal{D}}(\cdot, \mathbf{z}, \cdot) \end{pmatrix}. \quad (5)$$

The four submatrices introduced in (5) allow to specifically describe the transition rates between up and down states. Typically, in a reliability study without maintenance action, it is impossible to go back into an up state if the system reached a down state (except for self reparable system). In this paper we assume that the matrix $A_{\mathcal{D}\to\mathcal{U}}(\cdot, \mathbf{z}, \cdot)$ is equal to zero.

4.1.1 Reliability

Let $R : \mathbb{N}^* \mapsto [0, 1]$ denote the reliability of the system. $R(t)$ represents the probability that the system is always stayed in an up state until moment t. In other words,

$$R(t) = \mathbb{P}(X_1 \in \mathcal{U}, \ldots, X_t \in \mathcal{U}),\ t \in \mathbb{N}^*.$$

Similarly, let $T_{\mathcal{D}}$ denote the random variable describing the first hitting time of the subset $\mathcal{D}$, i.e. $T_{\mathcal{D}} = \inf \{t \in \mathbb{N}^* | X_t \in \mathcal{D}\}$ and then the reliability is given by

$$R(t) = \mathbb{P}(T_{\mathcal{D}} > t),\ t \in \mathbb{N}^*.$$

Although the reliability is supposed to be undefined for $t = 0$, we set that $R(0) = 1$ by convention which is useful for the definitions given in the sequel.

Finally, we can remark that in the case $A_{\mathcal{D}\to\mathcal{U}}(\mathbf{z}) = 0$ (i.e. $\mathcal{D}$ is a set of absorbing states), we have

$$R(t) = \mathbb{P}(X_t \in \mathcal{U}),\ t \in \mathbb{N}^*.$$

4.1.2 Failure rate

The failure rate $h : \mathbb{N}^* \mapsto [0, 1]$ is defined as the conditional probability that the failure of the system occurs at the moment t given that it has worked until moment $t - 1$. In other terms,

$$h(t) = \mathbb{P}(T_{\mathcal{D}} = t | T_{\mathcal{D}} \geq t),\ t \in \mathbb{N}^*.$$

Besides, the failure rate can be expressed using the reliability as follows :

$$h(t) = \begin{cases} 1 - \frac{R(t)}{R(t-1)}, & R(t-1) \neq 0 \\ 0, & \text{otherwise} \end{cases}, t \in \mathbb{N}^*.$$

4.1.3 Mean Time To Failure (MTTF)

The Mean Time To Failure (MTTF) is defined as the expectation of the lifetime (i.e. the expectation of the hitting time to "down" states $\mathcal{D}$) :

$$MTTF = \mathbb{E}\left[T_{\mathcal{D}}\right] = \sum_{t=1}^{+\infty} t\mathbb{P}(T_{\mathcal{D}} = t).$$

Once again, it is possible to express the MTTF with the reliability :

$$MTTF = 1 + \sum_{t=1}^{+\infty} R(t).$$

As the failure rate and the MTTF can be expressed using the reliability, the objective is to build an algorithm able to compute $R(t)$. The simplest approach consists in using a generic graphical model inference method like those based on junction trees. Unfortunately, this kind of methods are not optimized for the problem of reliability estimation since they involve extra-calculations (e.g. junction tree building, backward probability propagation) which are not necessary for our problematic. That is why in the next section, we propose an efficient and simple algorithm aimed to estimate the probability of any sequence of state subsets $\mathcal{E}$ defined as

$$\mathcal{E} = \{\mathcal{E}_1, \ldots, \mathcal{E}_T\},$$

where each $\mathcal{E}_t \subseteq \mathcal{S}$ with $\mathcal{S}$, the set of the system states. Thus, by setting $\mathcal{E} = \{\underbrace{\mathcal{U}, \ldots, \mathcal{U}}_{t \text{ times}}\}$, the method will provide $R(t)$.

Besides, let us note that exact calculations can be carried out provided all the CPDs in the model are finite and discrete (i.e. represented as CPTs). Otherwise, one can resort to use approximate inference algorithms which allow in general to work with any kind of CPDs.

4.2 Estimation method

By analogy with the Markov property, it is straightforward to verify that in a GDM, the pair (X_t, X_t^D) d-separates the future (slices $\tau \geq t+1$) from the past (slices $1 \leq \tau \leq t-1$). In other words, the future is independent from the past given (X_t, X_t^D) which is noted

$$\underbrace{(\mathbf{Z}_\tau, X_\tau, X_\tau^D, J_\tau)_{1\leq\tau\leq t-1}}_{\text{past}}, \mathbf{Z}_t \perp\!\!\!\perp \underbrace{(\mathbf{Z}_\tau, X_\tau, X_\tau^D, J_\tau)_{\tau\geq t+1}}_{\text{future}}, J_t \mid \underbrace{X_t, X_t^D}_{\text{present}}. \tag{6}$$

As it is shown in the next paragraphs, this property is the key to make tractable calculations in GDMs.

The aim of the following paragraph is to build an algorithm able to compute the probability

$$\mathbb{P}(X_1 \in \mathcal{E}_1, \ldots, X_T \in \mathcal{E}_T),\ \forall \mathcal{E}_t \subseteq \mathcal{S},\ t \in \{1, \ldots, T\}.$$

To begin, let $\alpha_t(i,d) = \mathbb{P}(X_1 \in \mathcal{E}_1, \ldots, X_{t-1} \in \mathcal{E}_{t-1}, X_t = i, X_t^D = d)$, then using the GDM joint distribution factorization given in (2), we obtain

$$\begin{aligned}
\alpha_1(i,d) &= \mathbb{P}(X_1 = i, X_1^D = d) = \int_{\mathcal{Z}_1}\int_{\mathcal{Z}_2}\cdots\int_{\mathcal{Z}_P} \mathbb{P}(\mathbf{Z}_1 = \mathbf{z}, X_1 = i, X_1^D = d) dz_P \ldots dz_2 dz_1 \\
&= \int_{\mathcal{Z}_1}\int_{\mathcal{Z}_2}\cdots\int_{\mathcal{Z}_P} \left[\prod_{p=1}^{P} \mathbb{P}(Z_{p,1} = z_p)\right] \mathbb{P}(X_1 = i|\mathbf{Z}_1 = \mathbf{z})\mathbb{P}(X_1^D = d|\mathbf{Z}_1 = \mathbf{z}, X_1 = i) dz_P \ldots dz_2 dz_1,
\end{aligned} \tag{7}$$

where $\mathbf{z} = (z_1, \ldots, z_P)$. As a consequence, from (7), we deduce the initial probability to begin in the subset of states $\mathcal{E}_1$:

$$\mathbb{P}(X_1 \in \mathcal{E}_1) = \sum_{i\in\mathcal{E}_1}\sum_{d\in\mathbb{N}^*} \alpha_1(i,d).$$

Besides, using the property (6) and equation (2), we can write

$$\begin{aligned}
&\mathbb{P}\left((\mathbf{Z}_\tau, X_\tau \in \mathcal{E}_\tau, X_\tau^D, J_\tau)_{1\leq\tau\leq t-1}, \mathbf{Z}_t, X_t, X_t^D\right) \\
=\ &\mathbb{P}\left(J_{t-1}, \mathbf{Z}_t, X_t, X_t^D|(\mathbf{Z}_\tau, X_\tau \in \mathcal{E}_\tau, X_\tau^D, J_{\tau-1})_{1\leq\tau\leq t-1}\right)\mathbb{P}\left((\mathbf{Z}_\tau, X_\tau \in \mathcal{E}_\tau, X_\tau^D, J_{\tau-1})_{1\leq\tau\leq t-1}\right) \\
=\ &\underbrace{\mathbb{P}\left(J_{t-1}, \mathbf{Z}_t, X_t, X_t^D|X_{t-1} \in \mathcal{E}_{t-1}, X_{t-1}^D\right)\mathbb{P}\left((X_\tau \in \mathcal{E}_\tau)_{1\leq\tau\leq t-1}, X_{t-1}^D\right)}_{\text{term } C_1} \\
&\times\underbrace{\mathbb{P}\left((\mathbf{Z}_\tau, X_{\tau-1}^D, J_{\tau-1})_{1\leq\tau\leq t-1}|(X_\tau \in \mathcal{E}_\tau)_{1\leq\tau\leq t-1}, X_{t-1}^D\right)}_{\text{term } C_2}
\end{aligned} \tag{8}$$

Observing that marginalizing (8) onto variables $(\mathbf{Z}_\tau, X_{\tau-1}^D, J_{\tau-1})_{1\leq\tau\leq t-1}$ lets the term C_1 unchanged and sums to one the term C_2, it follows :

$$\begin{aligned}
&\mathbb{P}\left((X_\tau \in \mathcal{E}_\tau))_{1\leq\tau\leq t-1}, J_{t-1}, \mathbf{Z}_t, X_t, X_t^D\right) \\
=\ &\mathbb{P}\left(J_{t-1}, \mathbf{Z}_t, X_t, X_t^D|X_{t-1} \in \mathcal{E}_{t-1}, X_{t-1}^D\right)\mathbb{P}\left((X_\tau \in \mathcal{E}_\tau)_{1\leq\tau\leq t-1}, X_{t-1}^D\right) \\
=\ &\sum_{j\in\mathcal{E}_{t-1}} \mathbb{P}(J_{t-1}, \mathbf{Z}_t, X_t, X_t^D|X_{t-1} = j, X_{t-1}^D)\alpha_{t-1}(j,\cdot) \\
=\ &\sum_{j\in\mathcal{E}_{t-1}} \mathbb{P}\left(J_{t-1}|X_{t-1}^D\right)\left[\prod_{p=1}^{P}\mathbb{P}(Z_{t,p})\right]\mathbb{P}(X_t|X_{t-1} = j, J_{t-1}, \mathbf{Z}_t)\mathbb{P}(X_t^D|X_{t-1}^D, J_{t-1}, \mathbf{Z}_t, X_t)\alpha_{t-1}(j,\cdot)
\end{aligned} \tag{9}$$

Hence using (9), we can express α_t as a function of α_{t-1} since we have

$$\begin{aligned}
\alpha_t(i,d) &= \sum_{j\in\mathcal{E}_{t-1}} \sum_{d'\in\mathbb{N}^*} \sum_{\zeta\in\{0,1\}} \int_{\mathcal{Z}_1} \cdots \int_{\mathcal{Z}_P} \mathbb{P}\left(J_{t-1}=\zeta | X_{t-1}^D = d'\right) \left[\prod_{p=1}^{P} \mathbb{P}(Z_{t,p}=z_p)\right] \\
&\quad \times \mathbb{P}(X_t = i | X_{t-1}=j, J_{t-1}=\zeta, \mathbf{Z}_t = \mathbf{z}) \\
&\quad \times \mathbb{P}(X_t^D = d | X_{t-1}^D = d', J_{t-1}=\zeta, \mathbf{Z}_t=\mathbf{z}, X_t = i)\alpha_{t-1}(j,d') dz_P \ldots dz_1 \\
&= \sum_{j\in\mathcal{E}_{t-1}} \sum_{d'\in\mathbb{N}^*} \alpha_{t-1}(j,d') \sum_{\zeta\in\{0,1\}} \mathbb{P}\left(J_{t-1}=\zeta | X_{t-1}^D = d'\right) \int_{\mathcal{Z}_1} \cdots \int_{\mathcal{Z}_P} \left[\prod_{p=1}^{P} \mathbb{P}(Z_{t,p}=z_p)\right] \\
&\quad \times \mathbb{P}(X_t = i | X_{t-1}=j, J_{t-1}=\zeta, \mathbf{Z}_t = \mathbf{z}) \mathbb{P}(X_t^D = d | X_{t-1}^D = d', J_{t-1}=\zeta, \mathbf{Z}_t=\mathbf{z}, X_t = i) dz_P \ldots dz_1.
\end{aligned} \tag{10}$$

In addition for the sake of readability, let $m_t(j,d',i,d)$ denote the quantity

$$\begin{aligned}
m_t(j,d',i,d) &= \sum_{\zeta\in\{0,1\}} \mathbb{P}\left(J_{t-1}=\zeta | X_{t-1}^D = d'\right) \int_{\mathcal{Z}_1} \cdots \int_{\mathcal{Z}_P} \left[\prod_{p=1}^{P} \mathbb{P}(Z_{t,p}=z_p)\right] \\
&\quad \times \mathbb{P}(X_t = i | X_{t-1}=j, J_{t-1}=\zeta, \mathbf{Z}_t = \mathbf{z}) \mathbb{P}(X_t^D = d | X_{t-1}^D = d', J_{t-1}=\zeta, \mathbf{Z}_t=\mathbf{z}, X_t = i) dz_P \ldots dz_1,
\end{aligned} \tag{11}$$

such that equation (10) can be rewrite as follows

$$\alpha_t(i,d) = \sum_{j\in\mathcal{E}_{t-1}} \sum_{d'\in\mathbb{N}^*} \alpha_{t-1}(j,d') m_t(j,d',i,d).$$

Finally by definition of α_t, we find that

$$P(X_1 \in \mathcal{E}_1, \ldots, X_t \in \mathcal{E}_t) = \sum_{i\in\mathcal{E}_t} \sum_{d\in\mathbb{N}^*} \alpha_t(i,d).$$

To sum up, theses results show that given any state sequence $\mathcal{E} = \{\mathcal{E}_1, \ldots, \mathcal{E}_T\}$, the following forward recursion allows to compute the probability of the sequence $\mathcal{E}$:

$$\begin{cases}
\text{Initialize } \alpha_1 & \text{according to equation (7)} \\
\mathbb{P}(X_1 \in \mathcal{E}_1) = \sum_{i\in\mathcal{E}_1} \sum_{d\in\mathbb{N}^*} \alpha_1(i,d) & \\
 & \\
\forall t \in \{1, \ldots, T\}, & \\
\text{Compute } m_t(j,d',i,d) & \text{according to equation (11)} \\
\alpha_t(i,d) = \sum_{j\in\mathcal{E}_{t-1}} \sum_{d'\in\mathbb{N}^*} \alpha_{t-1}(j,d') m_t(j,d',i,d) & \\
P(X_1 \in \mathcal{E}_1, \ldots, X_t \in \mathcal{E}_t) = \sum_{i\in\mathcal{E}_t} \sum_{d\in\mathbb{N}^*} \alpha_t(i,d) &
\end{cases} \tag{12}$$

Some aspects about the behaviour of the previous recursion are discussed in the three following remarks :

1. Under the classic assumption according to which all the CPDs are identically distributed over time as soon as $t \geq 2$, then m_t does not change during the recursion. Thus, it is possible to compute m_t once and for all which decreases the algorithm computing time.

2. If all the CPDs are finite and discrete (i.e. represented as CPTs), then integrals over the $\mathcal{Z}_p$'s and infinite sums over $\mathbb{N}^*$ become finite sums which allows to perform exact calculations.

3. For instance, if we set $\mathcal{E}^1 = \{\underbrace{\mathcal{U}, \ldots, \mathcal{U}}_{t \text{ times}}\}$, and $\mathcal{E}^2 = \{\underbrace{\mathcal{S}, \ldots, \mathcal{S}}_{t-1 \text{ times}}, \mathcal{E}_t\}$ respectively, then the previous method will compute $R(t)$ and $\mathbb{P}(X_t \in \mathcal{E}_t)$ respectively.

5 Application

To illustrate our approach, we use a GDM to model the behaviour of a 3-states system representing a production machine. In addition, we consider that the machine is subjected to one covariate, namely its production speed. The resulting GDM contains the following variables :

1. $Z_{1,t}$ represents the speed of the studied machine. Besides, we assume that two adjustments are available : "low speed" and "high speed" and then $\mathcal{Z}_1 = \{\text{low}, \text{high}\}$. As a consequence, the PDF of $Z_{1,t}$ gives the proportion of both adjustments during one time unit. Naturally, we consider that the longer the system works at high speed, the shorter its lifetime will be.

2. X_t is the state of our system. We consider $K = 3$ states of degradation : "no defect" (N), "minor defect" (M), "critical failure" (F) and then $\mathcal{S} = \{\text{N}, \text{M}, \text{F}\}$. The first two states do not bring on production stop which is the case of the last one. That is why we set $\mathcal{U} = \{\text{N}, \text{M}\}$ and $\mathcal{D} = \{\text{F}\}$. The transition rates between system states are given in table 2.

3. X_t^D represents the sojourn-time distribution for each state and each speed adjustments. The duration are expressed in months. Therefore, $|\mathcal{S}| \times |\mathcal{Z}_1| = 3 \times 2 = 6$ duration distributions have to specified. In fact, only 4 distributions are necessary if $\mathcal{D}$ is a set of absorbing states which is the case here since when the system reach $\mathcal{D}$, the associated sojourn-time is infinite. Besides, we assume the sojourn-times have truncated discrete Weibull distribution ($\mathcal{W}^D(\mu, \gamma)$), so that p_{X^D} (cf. equation (4)) is defined as (Bracquemond and Gaudoin 2003) :

$$p_{X^D}(z_1, i, d) = \begin{cases} \exp\left[-\left(\frac{d-1}{\mu(z_1,i)}\right)^{\gamma(z_1,i)}\right] - \exp\left[-\left(\frac{d}{\mu(z_1,i)}\right)^{\gamma(z_1,i)}\right] & \text{if } 1 \leq d < D, \\ \exp\left[-\left(\frac{d}{\mu(z_1,i)}\right)^{\gamma(z_1,i)}\right] & \text{if } d = D, \\ 0 & \text{otherwise,} \end{cases}$$

where $D \geq 2$ is the duration bound in such a way that we obtain discrete and finite duration distributions. In order to get relevant results, D has to take the highest value as possible. On the other hand, as by construction the size of the duration CPT is in $\mathcal{O}(D^2 K |\mathcal{Z}|)$, D has to take a reasonable value to avoid intractable structures. In the sequel, we set $D = 150$ months which appears to be a convenient value since the reliability analysis will be performed over only 100 months. Finally, table 1 gives a quantitative description of p_{X^D} for the system.

Table 1: Sojourn-time distribution for each state and each production speed.

	state	
production speed	N	M
low	$\mathcal{W}^D(30, 1)$	$\mathcal{W}^D(20, 3)$
high	$\mathcal{W}^D(20, 1)$	$\mathcal{W}^D(10, 3)$

The GDM used in this example has been implemented in `MATLAB`® environment, completed by the free Bayes Net Toolbox (BNT) written by Murphy (2001). The method presented in (12) has been used to compute reliability estimations presented in figure 3(a). Then, failure rate and MTTF have been deduced based on reliability estimations and depicted in figures 3(b) and 3(c) respectively.

These figures allow to characterise the behaviour of the studied system for different functioning types (i.e. proportions of production speed). As a consequence, useful information about the covariate effects on the system can be deduced from such analysis. In addition, these kind of results are essential when one wants to set up and optimize reliability-based maintenance policies.

Table 2: System transition CPTs.

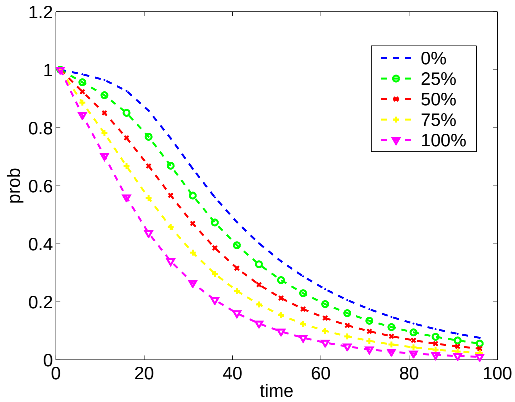
(a) $A(i, \text{low}, j)$

	N	M	F
N	0	9/10	1/10
M	0	0	1
F	0	0	1

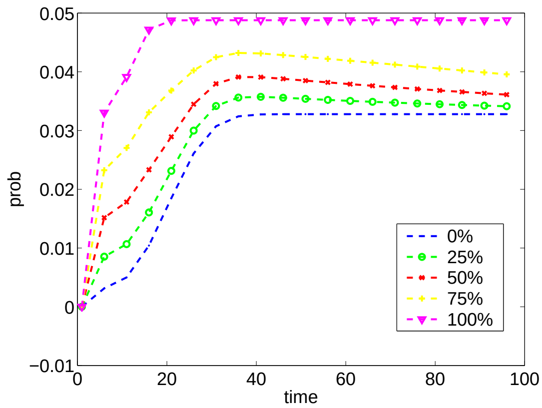
(b) $A(i, \text{high}, j)$

	N	M	F
N	0	3/10	7/10
M	0	0	1
F	0	0	1

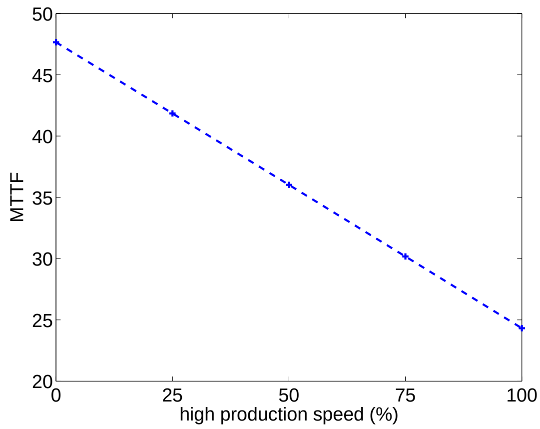
Figure 2: System reliability and related metrics over time (in months) for different proportions of high production speed.



(a) Reliability.



(b) Failure rate.



(c) MTTF.

6 Conclusion

The proposed method based on the GDMs aims to study the behaviour of a complex system. Our approach turns to be a satisfying and a comprehensive solution to model and estimate reliability and related metrics. Indeed, the proposed modelling is generic since it is possible to take into account the context of the system along with an accurate description of its survival distribution. In addition, as the method is based on graphical models, it makes it more intuitive and readable than more theoretical models.

The encouraging results presented in this paper confirm that GDMs are competitive reliability analysis tools for practical problems. In future work, we will address the problem of maintenance modelling. The aim is to build different reliability-based maintenance models which rely on GDMs. Finally, we will focus on the development of optimisation method in order to determine optimal maintenance policies.

References

Aven, T. and U. Jensen (1999). *Stochastic Models in reliability.* Number 41 in Stochastic Modelling and Applied Probability. Springer.

Barbu, V., M. Boussemart, and N. Limnios (2004). Discrete time semi-markov processes for reliability and survival analysis. *Communication in Statistics - Theory and Methods 33*(11), 2833–2868.

Barbu, V. and N. Limnios (2006). *Nonparametric Estimation for Failure Rate Functions of Discrete Time semi-Markov Processes*, pp. 53–72. Springer U.S.

Boudali, H. and J. B. Dugan (2005, March). A discrete-time bayesian network reliability modeling and analysis framework. *Reliability Engineering & System Safety 87*(3), 337–349.

Bracquemond, C. and O. Gaudoin (2003). A survey on discrete lifetime distributions. *International Journal on Reliability, Quality, and Safety Engineering 10*(1), 69–98.

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

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2003). *Bayesian Data Analysis* (Second ed.). Texts in statistical Science. Chapman & Hall/CRC.

Huang, C. and A. Darwiche (1996, October). Inference in belief networks: A procedural guide. *International Journal of Approximate Reasoning 15*(3), 225–263.

Jensen, F. V. (1996). *An introduction to Bayesian networks.* Press, U. C. L.

Kalbfleisch, J. D. and R. L. Prentice (2002). *The Statistical Analysis of Failure Time Data. Second Edition.* Wiley Series in Probability and Statistics. Wiley.

Kay, R. (1977). Proportional hazard regression models and the analysis of censored survival data. *Applied Statistics 26*(3), 227–237.

Langseth, H. and L. Portinale (2006). Bayesian networks in reliability. *Reliability Engineering & System Safety In Press, Corrected Proof.*

Limnios, N. and G. Oprisan (2001). *Semi-Markov Processes and Reliability.* Statistics for Industry & Technology. Springer.

Murphy, K. P. (2001). The bayes net toolbox for matlab.

Murphy, K. P. (2002). *Dynamic Bayesian Networks : Representation, Inference and Learning.* Ph. D. thesis, University of California, Berkeley.

Neapolitan, R. E. (2003, April). *Learning Bayesian Networks.* Prentice Hall.

Pearl, J. (1988, September). *Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference.* Morgan Kaufmann.

Weber, P. and L. Jouffe (2003, June). Reliability modelling with dynamic bayesian networks. In *5th IFAC Symposium on fault Detection, Supervision and Safety of Technical Processes*, Washington D.C., USA.