UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

NOVEL AND FASTER WAYS FOR SOLVING SEMI-MARKOV PROCESSES: MATHEMATICAL AND NUMERICAL ISSUES

TESE SUBMETIDA À UFPE

PARA OBTENÇÃO DE GRAU DE DOUTOR

POR

MÁRCIO JOSÉ DAS CHAGAS MOURA

Orientador: Enrique López Droguett, Ph.D.

RECIFE, Junho / 2009

M838n **Moura, Marcio José das Chagas.**

Novel and faster ways for solving semi-markov processes: mathematical and numerical issues / Marcio José das Chagas Moura. - Recife: O Autor, 2009.

x, 106 folhas, il : tabs. grafs., figs

Tese (Doutorado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia de Produção.

Inclui Referência e Apêndice.

1. Engenharia de Produção. 2.Processos Semi-Markovianos. 3.Densidade de Freqüência de Transição. 4.Quadratura Gaussiana. 5. Confiabilidade. 6. Avaliação da Disponibilidade. I. Título

UFPE

658.5 **BCTG/ 2009-110**

UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

PARECER DA COMISSÃO EXAMINADORA

DE DEFESA DE TESE DE

DOUTORADO DE

MÁRCIO JOSÉ DAS CHAGAS MOURA

"NOVEL AND FASTER WAYS FOR SOLVING SEMI-MARKOV PROCESSES: MATHEMATICAL AND NUMERAL ISSUES"

ÁREA DE CONCENTRAÇÃO: PESQUISA OPERACIONAL

A comissão examinadora, composta pelos professores abaixo, sob a presidência do(a) primeiro(a), considera o candidato **MÁRCIO JOSÉ DAS CHAGAS MOURA** APROVADO.

Recife, 29 de junho de 2009.

Prof. ENRIQUE ANDRÉS LÓPEZ DROGUETT, PhD (UFPE)

Prof.ª ANA PAULA CABRAL SEIXAS COSTA, Doutor (UFPE)

Prof. JOSÉ LAMARTINE TÁVORA JÚNIOR, Doutor (UFPE)

Prof. PAULI ADRIANO DE ALMADA GARCIA, Doutor (UFF)

Prof.ª MARTA AFONSO FREITAS, Doutor (UFMG)

PÁGINA DEDICATÓRIA

À minha avó, Iracema das Chagas (*in memorian*).

À minha mãe, Maria da Conceição.

À minha filha, Geovanna.

AGRADECIMENTOS

Esse sem dúvida é o tópico mais fácil de redigir desse trabalho. As palavras fluem naturalmente. A causa disso é o sentimento de alegria e alívio desse momento. Inúmeras pessoas contribuíram para a concretização desse sonho.

Em primeiro lugar, agradeço a Deus por te me concedido paz e saúde, pré-requisitos necessários e suficientes para se iniciar qualquer coisa nesta vida.

Agradeço à minha avó, Iracema das Chagas (*in memorian*), à minha mãe, Maria da Conceição e à minha filha, Geovanna, mulheres que fazem minha vida ter sentido diariamente. Agradeço à minha família, em geral, especialmente a meus tios (Marinalva, Mário, Marina, Zezinho e Carminha (*in memorian*)) e meu pai, Marcos Antônio. Agradeço à minha namorada, Milena Nascimento por ser fonte de afeto, compreensão e reflexão. Agradeço a meus amigos mais próximos que, indubitavelmente, são considerados como família: Flavinho, Reginha, Dona Lourdes, Selma, Dona Lindalva.

Agradeço ao professor Enrique López por ter creditado em mim confiança e me concedido a oportunidade de trabalhar no CEERMA – Centro de Estudos e Ensaios em Risco e Modelagem Ambiental. Agradeço também a todos que compõem tal grupo de pesquisa, especialmente a Paulo Firmino e Isis Didier pelas discussões científicas sempre proveitosas.

Agradeço aos professores membros da banca, Ana Paula Cabral, José Lamartine, Pauli Garcia e Marta Afonso, pelas sugestões valiosas e à Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo financiamento acadêmico concedido desde o meu mestrado.

Agradeço a todos os membros (professores, secretárias, coordenadores, colaboradores, colegas de classe, etc.) do DEP – UFPE (Departamento de Engenharia de Produção), especialmente Juliane, que sempre estiveram dispostos a ajudar desde a minha graduação.

Agradeço a todos meus amigos, especialmente a Mário Vieira, Vanessa Valentim, Hélder Diniz, Adiel Filho, Douglas Tomé, Marcio Mello e Felipe Fernandez por fazerem parte de minha estrutura.

Enfim, agradeço a todos que direta ou indiretamente me ajudaram no desenvolvimento do presente trabalho.

ABSTRACT

Continuous-time semi-Markov processes (SMP) are important stochastic tools for modeling reliability metrics over time for systems where the future behavior depends on the current and next states as well as on sojourn times. The classical approach for solving the interval transition probabilities of SMP consists of directly applying any general quadrature method to the integral equations. However, this approach has a considerable computational effort. Namely N^2 coupled integral equations must be solved, where N is the number of states. Therefore, this thesis proposes more efficient mathematical and numerical treatments for SMP. The first approach, which is called *2N*-method, is based on transition frequency densities and general quadrature methods. Basically, it consists of only solving N coupled integral equations and N straightforward integrations. Another proposed method, named *Lap*-method, is based on the application of Laplace transforms that are inverted by the Gauss quadrature method known as Gauss Legendre to obtain the state probabilities on the time domain. Mathematical formulation of these approaches as well as descriptions of their numerical treatment, including accurateness and time convergence issues, are developed and provided with details. The effectiveness of the novel *2N*- and *Lap*-developments will be compared against the results provided by the classical method by using examples in the context of reliability engineering. From these examples, it is showed that the *2N*- and the Laplace-based approach are significantly less time-consuming and have accuracy comparable to the classical method.

Keywords: Semi-Markov Process; Transition Frequency Densities; Quadrature Methods; Laplace Transforms; Gauss Quadrature; Reliability; Availability Assessment.

RESUMO

Processos semi-Markovianos (SMP) contínuos no tempo são importantes ferramentas estocásticas para modelagem de métricas de confiabilidade ao longo do tempo para sistemas para os quais o comportamento futuro depende dos estados presente e seguinte assim como do tempo de residência. O método clássico para resolver as probabilidades intervalares de transição de SMP consiste em aplicar diretamente um método geral de quadratura às equações integrais. Entretanto, esta técnica possui um esforço computacional considerável, isto é, N^2 equações integrais conjugadas devem ser resolvidas, onde N é o número de estados. Portanto, esta tese propõe tratamentos matemáticos e numéricos mais eficientes para SMP. O primeiro método, o qual é denominado *2N*-, é baseado em densidades de frequência de transição e métodos gerais de quadratura. Basicamente, o método *2N* consiste em resolver N equações integrais conjugadas e N integrais diretas. Outro método proposto, chamado *Lap*-, é baseado na aplicação de transformadas de Laplace as quais são invertidas por um método de quadratura Gaussiana, chamado *Gauss Legendre*, para obter as probabilidades de estado no domínio do tempo. Formulação matemática destes métodos assim como descrições de seus tratamentos numéricos, incluindo questões de exatidão e tempo para convergência, são desenvolvidas e fornecidas com detalhes. A efetividade dos novos desenvolvimentos *2N*- e *Lap*- serão comparados contra os resultados fornecidos pelo método clássico por meio de exemplos no contexto de engenharia de confiabilidade. A partir destes exemplos, é mostrado que os métodos *2N*- e *Lap*- são significantemente menos custosos e têm acurácia comparável ao método clássico.

Palavras-chave: Processos semi-Markovianos; Densidades de Frequência de Transição; Métodos de Quadratura; Transformadas de Laplace; Quadratura Gaussiana; Confiabilidade; Avaliação da Disponibilidade.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS[1]

ANN – Artificial Neural Network

BBN – Bayesian Belief Network

CBM – Condition-Based Maintenance

CDF – Cumulative Distribution Function

CPT – Conditional Probability Table

DBN – Dynamic Bayesian Network

EHMP – Embedded Homogeneous Markov Process

ENHMP – Embedded Non-Homogeneous Markov Process

FTS – Fault Tolerant Systems

GA - Genetic Algorithm

HSMP – Homogeneous Semi-Markov Process

HEP – Human Error Probability

HRA – Human Reliability Analysis

LT – Laplace Transform

MC – Monte Carlo

MP – Markov Processes

MTBF – Mean Time Between Failures

MTTF – Mean Time To Failure

MTTR – Mean Time To Repair

NHSMP – Non-Homogeneous Semi-Markov Process

OMS – Optical Monitoring Systems

PDF – Probability Density Function

ROCOF – Rate Of Occurrence Of Failures

ROI – Return of Investment

SMP – Semi-Markov Process

SMDP – Semi-Markov decision Process

SVM – Support Vector Machines

TDT – Tolerable Down-Time

[1] The singular and plural of an acronym are always spelled the same.

SUMMARY

1. INTRODUCTION

1.1 Overview

A homogeneous semi-Markov process (HSMP) can be understood as a probabilistic model whose future behavior is dependent on sojourn times which are random variables dependent on the current state i and on the state j to which the process will transit next. According to Ouhbi and Limnios (2003), HSMP are more flexible models than ordinary Markov processes as it is no longer required to assume that sojourn times are exponentially distributed.

Recent applications and theoretical developments on HSMP have been proposed in the context of reliability engineering. For example, Perman et al. (1997) apply a recursive procedure to approximate the interval transition probabilities, which are used to assess the future behavior of an HSMP over time. Limnios (1997) proposes a dependability analysis for HSMP in discrete time by using a method based on algebraic calculus. Ouhbi and Limnios (1997) estimate reliability and availability through HSMP of a turbo-generator rotor using a set of real data. Ouhbi and Limnios (2002) propose a statistical formula for assessing the rate of occurrence of failures (ROCOF) of HSMP. Through this result, ROCOF of the Markov and alternated renewal processes are given as special cases. Some other applications of HSMP may be encountered in related literature, mainly in the reliability field (as exemplified in Janssen and Manca (2007); Limnios and Oprisan (2001); Pievatolo and Valadè (2003)).

The future behavior of an HSMP is assessed through its interval transition probability equations which are comprised of a set of N^2 coupled convolution integral equations, where N is the number of states. The classical method for solving these equations is explained in Corradi et al. (2004), and consists of directly applying a general quadrature method to these N^2 coupled convolution integral equations. However, such an approach is quite burdensome with a computational cost sometimes greater than the Monte Carlo (MC) simulation.

In a non-homogeneous semi-Markov process (NHSMP), transitions between two states in turn may depend not only on such states and on the sojourn times (x), but also on both times of the last (τ) and next (t) transitions, with $x = t - \tau$. The time variable τ is also known as the most recent arrival time or last entry time, and the time variable t is the calendar or process time. Thus, NHSMP extend other stochastic processes such as HSMP. As a result, NHSMP are powerful modeling tools, mainly in the context of reliability engineering (as exemplified in Janssen and Manca (2007)).

In spite of that, there are two main reasons to explain the scarcity of NHSMP applications: (*i*) Janssen and Manca (2001) argue the non-homogeneity on the continuous time semi-Markov environment implies additional difficulties in treating NHSMP; (*ii*) in accordance with Nelson and Wang (2007), for practical applications, gathering of high level required data (transition probabilities and/or rates) is likely to be a significant challenge, mainly in the presence of censoring implied by preventive maintenance.

Specifically regarding the first claim, it gives rise to more intricate mathematical methods and numerical solutions. Indeed, as it will be discussed in upcoming sections, the future behavior of an NHSMP is assessed through its interval transition probability equations which are comprised of a system of N^2 coupled integral equations with two variables, where N is the number of states. The classical method to solve the non-homogeneous equations is explained in Janssen and Manca (2001), and also consists of directly applying a general quadrature method to these N^2 coupled integral equations, as for HSMP. However, such an approach is more complex than in the case of homogeneous counterpart, because the integrals involved are not of convolution type anymore, and also, since the interval transition probabilities to be determined depend on two parameters.

As it can be seen from this overview on homogeneous and non-homogeneous semi-Markov processes, the dynamic behavior analysis of both these models requires solving a set of N^2 integral equations which increase considerably the computational time and intricacy of the related solution. Therefore, this thesis proposes alternative methods for solving the probability equations of HSMP and NHSMP in continuous time as an attempt to reduce the complexity associated with these stochastic models and to foster their applicability, mainly of NHSMP.

Basically, one of these approaches consists of casting the N^2 coupled integral equations of either HSMP or NHSMP into an initial value problem involving transition frequency densities, and then solve N coupled integral equations with one variable and N straightforward integrations. As it will be seen in upcoming chapters, this approach considerably reduces the computational effort in relation to the abovementioned classical method and MC simulation since it is not needed solving N^2 integral equations anymore.

This proposed approach is partly based on the work of Becker et al. (2000) where it is presented the mathematical formulation for semi-Markov processes (SMP) described by transition rates $\lambda_{ij}(\cdot)$. Similarly to Becker et al. (2000), the proposed approach also involves transition frequency densities. However, from this point the method discussed throughout this thesis departs from the one presented in Becker et al. (2000). Firstly, in the proposed method the HSMP and NHSMP may be specified in terms of not only transition rates $\lambda_{ij}(\cdot)$, but also through transition probabilities $C_{ij}(\cdot)$. SMP described via transition probabilities represent important modeling tools, mainly in reliability applications such as in Janssen and Limnios (1999) and Droguett et al. (2007). Thus, this thesis proposes an extension to the work developed in Becker et al. (2000) so that SMP described through both $\lambda_{ij}(\cdot)$ and $C_{ij}(\cdot)$ may be handled in a more efficient and integrated form.

Secondly, and conversely to Becker et al. (2000), this thesis is also numerical-based, i.e., a numerical treatment for the proposed mathematical formulation will be discussed. This numerical solution is based on general quadrature methods and will have its effectiveness compared against the classical method and the MC simulation by means of some examples in the reliability context.

Thus, the proposed approach is two-fold, i.e., it addresses mathematical and numerical issues related for solving SMP in continuous time. For the sake of simplicity, henceforth the classical and proposed approaches will be distinguished through their computational efforts as N^2-method and *2N*-method, respectively.

Another method which will be proposed here to handle specifically the behavior of HSMP is based on the Laplace Transform (LT) apparatus. The use of LT is not novel on problems involving HSMP. There are some works treating this issue in the related literature (Perman et al. (1997); Janssen and Manca (2006); Moura (2006); Howard (2007)). Through these approaches, LT are applied to the N^2 coupled convolution integral equations and thus the solution on time domain is obtained through respective inversion. However, as stated by Bellman and Roth (1984) (pp. 149), "We cannot expect that any specific method for the inversion of the LT will work equally well in all cases". Moreover, Csenki (1994) (p.233-234) argues that "no single method can be devised which will perform numerical LT inversion to a given accuracy". In other words, a unique numerical method to invert LT is not able to solve any problem in a general way.

In spite of these statements a method of LT inversion, which was developed by Bellman et al. (1966), has been applied by Oliveira et al. (2005) for solving the partial differential equations for non-homogeneous Markov processes described using supplementary variables. Great results attained on this situation (for distribution functions widely used in reliability context, like Exponential and Weibull) have led to delve on the feasibility of application of that LT method for solving SMP as well, for which the dynamic behavior rise from a generalization of the Kolmogorov backward differential equations of the Markov environment (see Feller (1964)).

Therefore, besides the *2N*- and N^2- approaches, there will also be described a method based on LT for solving SMP. This approach will be drawn only for HSMP (due to reasons that will come up over the text) and at the best of our knowledge, as *2N*-approach, it has not been used elsewhere within the semi-Markov environment.

This approach, which will be named *Lap*-method, will also be developed so that it can handle HSMP described through either transition probabilities or transition rates. The effectiveness of the *Lap*-numerical procedure will be compared against the *2N*- and N^2-methods and the MC simulation in terms of computational effort (time) and accuracy by means of some examples in the context of reliability engineering.

Therefore, the main question behind this thesis is: "*How to solve (homogeneous and non-homogeneous) semi-Markov processes through a less intricate and more efficient way?*".

1.2 Motivation and Justifications

In this section, the main contributions and justifications, under which the present thesis is backed up, will be discussed. Basically, two examples that may be faced by reliability practitioners are presented in order to show which type of practical problems will be solved by the proposed mathematical and numerical approaches.

The first example addresses a case where an HSMP described by transition rates is used to handle a repairable pumping oil unit that pumps oil to a storage tank. Then it is discussed another example which consists of an NHSMP described by transition probabilities used to model a repairable pressure-temperature optical monitoring system for oil wells.

Basically, these examples will be treated by the proposed mathematical and numerical approaches which will be designed in upcoming chapters as an attempt to answer the aforementioned question.

1.2.1 Example 1: Pumping Oil Unit

Most probabilistic models for system availability, reliability and maintainability assessment assume that the failure of one component immediately causes system failure. In some systems, however, the failure of a component leads to a system failure only when repair time has exceeded some time *T*, known as tolerable downtime (TDT). According to Vaurio (1997), systems that have this feature are known as fault tolerant systems (FTS).

This concept is usually employed in the context of software-based systems reliability, for example, in Madan et al. (2004) who use SMP to model a possible security intrusion and corresponding response of the fault tolerant software system to this event. Other related works include Littlewood et al. (2002), Levitin (2004), Levitin (2005) and Levitin (2006).

In the context of fault tolerant safety systems, some reliability assessment models have been developed. For example, Camarinopoulos and Obrowski (1981) propose a model for reliability quantification that takes into account the frequency as well as the duration of failures. In that work, however, the TDT is considered constant, i.e., it does not have a stochastic behavior.

Becker et al. (1994) and Chandra and Kumar (1997) use Markov processes (MP) in order to model safety systems with stochastic TDT. An MP is defined as a probabilistic model that satisfies the memoryless Markov property. According to this assumption, the future behavior of a system depends only on its present state and therefore is independent on the sojourn time in this state. According to Ouhbi and Limnios (1997), however, such an assumption is not always appropriate, since it is required to assume that sojourn times are exponentially distributed.

Becker et al. (2000) model the reliability of FTS through SMP. SMP is an extension of Markov processes and as such they provide greater flexibility in terms of modeling complex dynamic systems. According to Howard (2007), SMP are not strictly Markovian anymore as the Markov property is not required at all instants. However, as they share enough characteristics in common with these processes, SMP receive that denomination. Moreover, when non-homogeneous semi-Markov processes are considered, it is also possible to model a system that might be under improvement or aging processes. In this type of SMP, the future behavior depends on two types of time variables: sojourn time and process time, being the latter also known as calendar or global time.

A common characteristic shared by the aforementioned reliability/availability assessment models is that the future behavior of a system is conditioned only on time variables, either process or sojourn times or both. In some situations, however, other factors not necessarily time can influence the system behavior. Examples of such external factors include environmental variables (e.g., temperature, humidity), operational variables (e.g., hydrate and H_2S concentration in oil flow), and physiological (e.g., fatigue) and/or psychological conditions (e.g., workload, stress).

In these cases, the system's future behavior might be influenced by sojourn time variable as well as by those external factors. To take it into account, it is possible to integrate continuous time homogeneous semi-Markov processes and Bayesian belief networks (BBN) (see Moura (2006) for greater details on the hybrid model: SMP and BBN).

As an example, assume that one is uncertain about the true value of the mean time to failure (MTTF) of a downhole pumping oil system, i.e., one is interested in assessing the uncertainty distribution of MTTF. The BBN topology in Figure 1-1 characterizes how the random variable MTTF of the downhole pumping system is influenced by the variables BWSOT: "Percentage of H_2O and solids", PARAF: "Level of paraffin", FILTER: "Classification of the filter installed", DEPTH_PUMP: "Depth of the pump unit".

As it can be seen in Figure 1-1, BBN is composed of nodes, which represent the variables of interest (discrete or continuous), and arcs that characterize the cause-effect relationships among these variables.

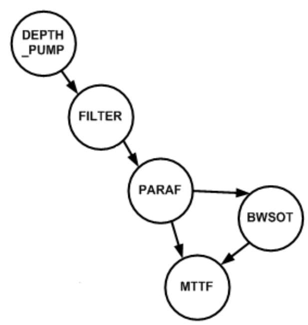


Figure 1-1 – BBN for MTTF of a pumping unit

The first step in setting up a BBN is the identification of random variables and their nature, i.e., whether they are discrete or continuous. Such values must be mutually exclusive. Next step is to designate the cause-effect relations among the relevant variables in order to construct the BBN topology.

In a BBN, a node is parent of a child node when there is an arc leaving the former in direction to the latter. In Figure 1-1, for instance, the variable "PARAF" is a parent of "BWSOT" and "MTTF". Any node with no parents is a root node, any node without children is a leaf node and any node that is neither a root nor leaf is an intermediary node. "DEPTH_PUMP" is a root node, "MTTF" is a leaf node and "PARAF" and "BWSOT" are intermediary nodes.

After the construction of the BBN topology, next step is to determine the strengths of the cause-effect relations among the connected variables. This is carried out by specifying a conditional probability distribution for each node. For discrete random variables, this consists of establishing conditional probabilities tables CPT for each node. These CPT can be generated from either data bases or engineering judgments, as in Langseth and Portinale (2007).

For the sake of simplicity, it is assumed that all variables in the BBN of Figure 1-1 are dichotomic unless MTTF that can assume the following values {100, 200, 1.000, 10.000} hours. The CPT given in Appendix A were obtained from a data base according to the methodology proposed in Barros Jr. (2006), where level 0 refers to an adequate condition and level 1 to an inadequate one. These CPT correspond to the prior distributions.

In this way, BBN is a graphic representation of a multivariate probability distribution where it is possible to represent cause-effect relations among random variables (Langseth and Portinale (2007)). Moreover, BBN provide flexibility in terms of knowledge updating through the Bayes theorem (see Bernardo and Smith (1994) for basic concepts on Bayesian inference) as discussed in Firmino (2004).

As an example of how to integrate a homogeneous SMP with BBN, consider a downhole pumping unit that pumps oil to a storage tank, which in turn is kept above a predetermined level L in order to be able to supply customers in case of a pumping unit failure. The tank level above L is set to a value such that a TDT holds before the oil level goes under L in case of a pumping unit failure. Therefore, upon the occurrence of this failure, it is assumed that repair starts immediately in order to not go under this predetermined level and consequently the TDT. Otherwise, the oil level in the storage tank goes under a low limit and the oil supply halts. When the pumping unit is under repair and the TDT has not expired yet, no damage to customers is inflicted as oil can still be supplied, i.e., although in a degraded state the system is still available. However, when the tolerable downtime is reached and repair has not been completed yet, the system fails and it is assumed to be unavailable.

It is clear that the elapsed time since the start of repair activities plays a relevant role with respect to system availability measure. Indeed, the system initially starts in state 1 (available) and upon failure (it is considered failure time follows an exponential distribution) of the pumping unit it transits to state 2 (failed, under repair and TDT not exceeded), as shown in Figure 1-2. When state 2 is reached, a local clock is started such that when the sojourn time in this state is greater than the TDT the system becomes unavailable, i.e., it transits to state 3 (failed, under repair and TDT exceeded).

In other words, the transition from state 2 to 3 depends on the elapsed time t since the pumping unit has failed. In both cases (either states 2 or 3), it is assumed repair rate μ is constant (see Figure 1-2). For the sake of simplicity, no failures are considered for pipelines, valves and the storage tank.

It is also assumed the TDT (in this case, time for the system transits from state 2 to 3) is distributed according to a Weibull distribution as follows:

$$f_{23}(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left(-\left({t}/{\alpha}\right)^{\beta}\right),$$

where α and β are scale and shape factors respectively.

Given that transitions outwards state 2 depend only on the sojourn time, it is considered a homogeneous semi-Markov process in order to address this FTS. Otherwise, an MP could have been chosen since in this case all transition rates would be constant.

Furthermore, suppose that, as it might happen in situations of practical interest, the MTTF characterizing transitions from state 1 to state 2 is influenced by some external factors. As discussed above, the causal relationships among external factors related to a transition rate can be characterized in terms of a BBN. As a result, availability measure of the pumping system could be estimated from the hybrid model based on HSMP and BBN (see Moura and Droguett (2008) in Attachment A).

In particular, for the system under consideration, assume that the MTTF of the exponentially distributed time up to pumping unit failures (i.e., sojourn time in state 1) is uncertain and influenced by the external factors shown in Figure 1-1.

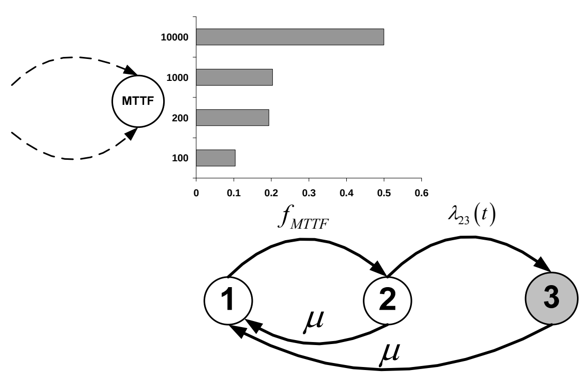


Figure 1-2 – HSMP for the downhole pumping oil unit

Figure 1-2 depicts an HSMP, which is described by transition rates $\lambda_{ij}(t)$, designed to model the oil pumping unit taking into account the influence of sojourn times and external factors on the future behavior of the system. In Figure 1-2, f_{MTTF} is marginal probability distribution of the MTTF and it is obtained from the BBN in Figure 1-1.

The requisite data needed for solving this sort of SMP are the parameters of the probability density functions (PDF) of the holding time in each state *i* given that system will go towards state *j*. In this example, it means to estimate of μ, α and β. Moreover, as MTTF of the transition from state 1 to 2 is uncertain, the CPT given in Appendix A are necessary to estimate the PDF on that parameter.

Through Figure 1-2, it is developed a model for a more realistic representation and quantification of availability measure for repairable FTS via the integration between continuous time HSMP and BBN. Such systems have a basic feature: the sojourn time in any state influences the transition probabilities. Moreover, external factors (e.g., environmental and operational conditions) not necessarily time variables also impact the future behavior of the system. Furthermore, as new evidence becomes available, the probability distributions of these parameters as well as the state of knowledge about the behavior of the system can be updated.

Thus, as the HSMP is described via transition rates then the integration between it and BBN is achieved through an interface represented by parameters of the intensity functions characterizing the transition rates. Such parameters are taken from BBN describing the cause and effect relationships among the relevant external factors and the corresponding parameters. The resulting uncertainty distribution about a particular parameter is then taken as input information for the HSMP.

In order to explicitly quantify the impact of the uncertainty in the transition rates on the state probabilities of the semi-Markov model, on the availability measure or on other relevant reliability metric, a numerical procedure for solving HSMP must be repeated for a considerable number of iterations.

Using the N^2-method (given in Corradi et al. (2004)), which is hardly time-consuming, the solution of the model in Figure 1-2 would become infeasible. Therefore, developing a faster and accurate way for solving HSMP is a must for practitioners who are used to face some sort of problems such as just described.

This example will be further discussed in chapter 4. In fact, it will be solved using *2N*-, *Lap*-, N^2- and MC approaches which will be compared in terms of computational cost and accuracy.

Next subsection presents another application of SMP. The problem characterization mainly draws from Droguett et al. (2007) and Droguett et al. (2008) which follow in attachments B and C respectively.

1.2.2 Example 2: Optical Monitoring Systems

Oil has been the most important source of energy since the early days of last century. The growing and continuous demand for energy associated with decreasing availability of this limited resource have led to a considerable increase in investment directed towards the development of renewable energy sources as well as to research efforts for optimizing technologies related to the exploration and production of oil.

Mostly because of the increasing oil price, a considerable attention has been given to the enhancement of production technologies that allow for anticipation of oil production volumes and an improved reservoir management and control. In line with such efforts, recent developments have led to the so called intelligent oil fields. The term 'intelligent' means: (*i*) data acquisition: sensors provide data on important well parameters in real time; (*ii*) flow remote control: it allows an operator to modify production or injection flow characteristics with no on-site intervention; (*iii*) data interpretation and optimization: it allows production and reservoir engineers feed simulation models and act on a particular well in real time. Therefore, intelligent oil field is a concept encompassing various technologies that allow for an integrated management of production and injection of one or several reservoirs.

Under these circumstances, availability is a key attribute: the higher availability the higher production volumes and therefore profit are. Moreover, in terms of intelligent oil fields, increased availability levels associated with the anticipation of production volumes in relation to what is currently attained by a conventional oil field might serve as evidence for justification of the considerable steep investment in new technology.

In this context, a research effort is underway for designing and implementing intelligent oil fields in mature wells located in the Northeast of Brazil. Part of this effort concerns the in-house development and installation of pressure-temperature optical monitoring systems (OMS).

At the current stage, only a few units of these systems have been deployed for field tests and, given the limited experience, availability assessment is usually performed under a considerable level of uncertainty. In spite of that scenario, this limited experience has suggested that an OMS might be comprised of components that are renewed after failures as well as components that are under deteriorating processes with failure intensity functions that are dependent on the total system age (process time).

Upon failure of the monitoring system, human performance during the reinstallation of an OMS (i.e., removing, repairing and then running an OMS in hole) is a relevant factor influencing its availability. Moreover, the time interval to accomplish the reinstallation plays an important role since it directly impacts the OMS availability as well as the human performance during the effort to recover the system. In fact, under real life oil production conditions in the Northeast of Brazil, there exists an available time to complete the reinstallation (tolerable downtime). Otherwise, the OMS reinstallation in the field is not longer feasible and, from the availability analysis perspective, it evolves to an unrecoverable state – this tolerable downtime is one of the factors that directly influence the human performance during the reinstallation and thus the OMS availability (see Droguett et al. (2008) for deeper details).

Therefore, there are three relevant aspects in estimating the OMS availability: (*i*) the available time to complete the reinstallation; (*ii*) the system deteriorating process and (*iii*) the maintenance crew's performance, which is influenced by tolerable downtime and other factors (e.g., experience, fatigue) in returning an OMS to its normal operational condition.

NHSMP may be used here in order to tackle the first two issues because: (*i*) the duration (sojourn time) in a state may influence the availability of an OMS and (*ii*) provided that some components might be under deteriorating processes, it should be considered time dependent transition intensity functions. In this context, the combined impact of these two time variables on the reliability of an OMS will be assessed through an NHSMP.

Indeed, OMS reinstallation process involves the repair of any possibly failed component as well as running the OMS system downhole. Thus, as it depicted in Figure 1-3, it is assumed that the system (OMS) starts at normal operation in state 1. Upon a system failure, the reinstallation process of the OMS starts, which is represented by state 2. If the reinstallation process cannot be completed, the system goes to state 3 where additional actions are taken to restore the system to its normal operating condition. If the operator is still not able to restore the system, all actions are halted as represented by state 4. Thus, the system is not functioning (unavailable) when in states 2, 3 or 4.

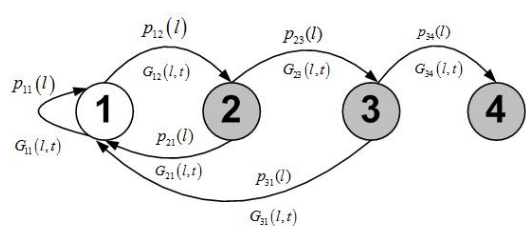


Figure 1-3 – Non-Homogeneous semi-Markov process for an OMS

OMS reinstallation procedures involve complex operations that require interactions between human elements and equipments. Thus, to take into account the third relevant issue (the crew's performance in recovering an OMS), BBN could again be used to address qualitatively and quantitatively the cause-effect relationships among factors that impact the Human Error Probability (HEP) during the reinstallation of an OMS.

In the context of the OMS reinstallation, the variables "available time to complete the reinstallation" and "the capacity to accomplish the task" directly influence the probability of the repairman to properly accomplish the procedure, as it can be seen in Figure 1-4.

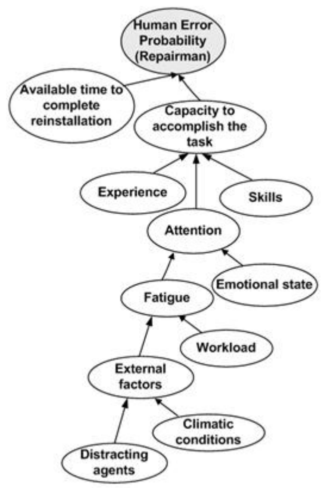


Figure 1-4 – BBN model for the OMS repairman

Furthermore, according to Figure 1-4 the repairman is influenced by external factors. Two of these factors are considered here: climatic conditions (e.g., temperature and humidity) and distracting agents (e.g., informal parallel chats in work environment, noise, glare, movement, flicker and color). These external factors associated with workload can cause fatigue (i.e., physical and/or mental fatigues). Fatigue associated with emotional state can influence the attention of the repairman to the current task. It is possible, for example, that due to fatigue and an unfavorable emotional state, the attention level is negatively impacted. It is considered that three factors can influence the repairman capacity to carry out his activities: attention, skills and experience. Attention refers to whether sufficient cognitive and physical resources are put at the "right" places. Skills are the ability to understand situations and perform needed actions without much cognitive activity. Deficiency of skills can manifest itself in reduced job quality and time delay. Experience is the accumulation of information and knowledge acquired through direct or indirect interactions with the system (see Chang and Mosleh (2007)). The repairman performance measured by the HEP is directly influenced by his capacity to carry out the task and the available time to complete the reinstallation. Both factors are considered to have major impact on the HEP, thus they are parents of the "human error" node. For a detailed discussion on how a Human Reliability Analysis (HRA) is performed, see Menêzes (2005). Table 1-1 summarizes the BBN nodes and the levels which they can assume.

Table 1-1 – Variables and their levels

Variable	**Levels**	
	0	1
Human Error	Yes	No
Available time to complete reinstallation	Adequate	Inadequate
Capacity to accomplish the task	Adequate	Inadequate
Experience	Average	High
Attention	Adequate	Inadequate
Skills	Adequate	Inadequate
Emotional State	Adequate	Inadequate
Fatigue	Adequate	Inadequate
Workload	Adequate	Inadequate
External Factors	Adequate	Inadequate
Distracting Agents	Yes	No
Climatic Conditions	Adequate	Inadequate

Thus, when the system is in state 2, it is assumed that the operator has an appropriate available time to complete the reinstallation tasks. Under this situation, the probability p_{23} corresponds to the HEP under a condition of "adequate available time to complete the repair (evidence 0)". If the operator does not complete the reinstallation in the allotted time frame, the system transits to state 3. In this state, the operator takes additional actions to restore the system but now under a time pressure situation, i.e., it is considered that a substantially reduced time frame is available to restore the system to its normal operating condition. Correspondingly, the HEP p_{34} reflects the situation of "inadequate available time to complete the repair (evidence 1)". In both cases (states 2 and 3), if the operator ends the reinstallation within the available time, the system returns to its normal operating condition (state 1). Otherwise, the system transits to state 4.

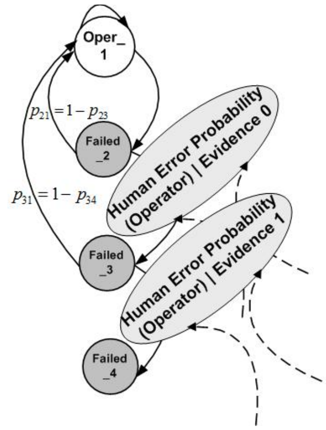


Figure 1-5 – A semi-Markov model with Bayesian belief network based human error probability for availability assessment of downhole optical monitoring systems

Therefore, when the system enters states 2 or 3, the BBN node "human error" is directly connected to the probabilities (parameters) of the NHSMP of Figure 1-3 what can be seen in Figure 1-5. Thus, the parameters p_{23} and p_{34} are the HEP in the BBN of Figure 1-4 given the evidences 0 and 1, respectively. The parameters p_{21} and p_{31} are the respective complements of p_{23} and p_{34}.

In this way, the conditioning factors influencing the error probability of an OMS repairman as well as the cause-effect relationships among them are taken into account for the availability assessment of an OMS via the continuous-time NHSMP.

As it may be noticed in the preceding description, the requisite data needed for solving this NHSMP are different from those for example 1. Indeed, Figure 1-3 illustrates an NHSMP described by transition probabilities. The required data to estimate the system (un)availability over time via this type of NHSMP are the transition probabilities $p_{ij}(\cdot)$ and the conditional Cumulative Distribution Function (CDF) $G_{ij}(\cdot,\cdot)$. These terms will be further described in detail in next chapter.

Due to the lack of a robust and efficient method to solve the example just described, Droguett et al. (2008) have solved it by using MC. Another possibility is to resort to the N^2-method drawn for NHSMP in Janssen and Manca (2001). However, due to computational time reasons this approach becomes impracticable.

Therefore, the *2N*-method for NHSMP will be developed in chapter 5. Then in chapter 6, the example described in the present section will be widely solved by using the *2N*-method.

1.2.3 Contributions

Actually, examples 1 and 2 address the availability assessment problem with somewhat no use of simplistic assumptions on the system's behavior. However, trying to approach as much as possible towards reality requires a price to be paid. In these cases, the penalty corresponds to the intricacy and complexity of mathematical and numerical formulations involved with SMP what also implies impracticable computational times.

Indeed, traditionally examples 1 and 2 could be solved by using N^2-method given in Corradi et al. (2004) (HSMP) and Janssen and Manca (2001) (NHSMP), respectively. However, both of them are rather cumbersome with a computational cost greater than MC.

This situation motivates the development of a novel and more efficient (faster) mathematical and numerical formulation for SMP that has less computational effort, but keeps the accuracy in relation to the available methods in the related literature, that is, MC simulation and the N^2-approach. In fact, the $2N$-mathematical formulation and numerical treatment consists of casting the N^2 coupled integral equations into an initial value problem involving transition frequency densities, and then solve N coupled integral equations with one variable and N straightforward integrations. As it will be proved in next chapters, this approach possesses both abovementioned features: it is significantly less time-consuming and has roughly accuracy equals to the N^2-method.

Specifically regarding NHSMP, although they are powerful modeling tools, the mathematical and computational difficulties of the non-homogeneous environment are usually blamed as accountable for the scarcity of applications of this type of stochastic process. Thus, this thesis plays an important role as an attempt to increase the feasibility of application of this kind of stochastic model.

Moreover, this thesis describes another alternative method (called *Lap*-approach) for solving the state probability equations of an HSMP on continuous time. This numerical procedure is based on the application of LT. As there will be seen in detail, the main advantage of this approach is that it is not required adjusting the number of steps in order to obtain the desired convergence. There will be a pre-set number of steps, which is independent on the problem to be solved and thus, this method is likely to have a considerable reduced computational effort in relation to the abovementioned $2N$- and N^2-methods and MC as well.

Finally, the $2N$-mathematical formulation and numerical approach will also be illustrated by means of some examples of application in the context of reliability assessment (including those which have been described in the two previous sections), where the effectiveness and the required computational effort of the *2N*-method will be compared against the MC and the N^2- and *Lap*-approaches.

1.3 Objectives

1.3.1 General Objective

Developing a novel mathematical formulation and a faster numerical treatment for solving on continuous time (homogeneous and non-homogeneous) semi-Markov processes described through either transition rates or transition probabilities.

1.3.2 Specific Objectives

- Surveying at the theoretical background of homogenous and non-homogeneous semi-Markov processes;
- Surveying at the N^2-method: numerical perspective;
- Developing the *2N*-mathematical formulation and numerical treatment for HSMP;
- Developing the *Lap*-numerical treatment for HSMP;
- Developing the *2N*-mathematical formulation and numerical treatment for NHSMP;
- Implementing numerically on C++ platform the solutions for *2N*-, *Lap-*, N^2- approaches and MC simulation in order to make comparisons among them in terms of computational time and accuracy;
- Applying the abovementioned methods to solve some examples in the context of reliability engineering in order to evaluate their effectiveness in terms of computational cost and accuracy.

1.3 Thesis Layout

The remainder of this thesis is organized as follows. Next chapter presents the theoretical background related to SMP: HSMP and NHSMP described through either transition probabilities or transition rates. Chapter 3 (section 3.1) develops the *2N*-method for homogeneous semi-Markov processes. In this chapter, the mathematical formulation involving transition frequency densities and the description of the numerical method (including the analysis of the discretization error) will be described. Chapter 3 (section 3.2) also describes the *Lap*-numerical method for HSMP. This method is composed of the application of LT and its corresponding inversion. Both issues will be discussed in that chapter. In chapter 4, comparisons among *2N*-, N^2-, *Lap*- and MC approaches will be made through some examples of application. Chapter 5 will show the description of the *2N*-method for NHSMP. As for HSMP, in this chapter, the mathematical formulation and the numerical treatment (including also the analysis of the discretization error) will be discussed taking into account non-homogeneity issues. Chapter 6 will challenge the effectiveness of the *2N*-method drawn for NHSMP against the results from N^2- and MC approaches. In this chapter, the example 2 described in section 1.2.2 will be solved using the *2N*-method. Next, chapter 7 presents two further examples. The first one will show how the *2N*-method may be used for determining a maintenance optimization policy so that to maximize the mean availability measure. The second example is also inserted inside the optimization context. Basically, it is designed for determining which maintenance decisions should be made so that the mean availability and expected costs are jointly optimized over the system's age. Thus, the *Lap*-method will be used to estimate the mean availability in this framework. Finally, chapter 8 presents some conclusions, discussing final remarks and challenges for ongoing and future research.

2. THEORETICAL BACKGROUND: SEMI-MARKOV PROCESSES

2.1. Applications and terminology

According to Howard (2007), an SMP can be understood as a probabilistic model in which the successive occupation of states is governed by the transition probabilities of an MP, known as embedded MP, but the sojourn times in each state is described by a random variable that depends on the current state and on the state to which the next transition will be done.

In an SMP, the Markov property is required only at the transition times between states and, therefore, it is not strictly Markovian. Thus, the sojourn time distribution can be arbitrary, following any probability density function not necessarily exponential.

Some recent scientific developments on SMP may be quoted. Grabski (2003) presents the properties of the reliability function of a component under a random load process with failure rate modeled according to an SMP. The reliability functions were obtained through application of Laplace-*Stieltjes* transforms to transition probability equations and, by using a commercial computational software, the analytical solution of the inverse transform were obtained.

Ouhbi and Limnios (2003) introduce non-parametric estimators for the reliability and availability of SMP by assessing the asymptotical properties of these types of metrics. A method to compute confidence intervals for such estimators is proposed and an example of application is given for a three state SMP. Limnios and Oprisan (2001) demonstrate some results and applications of SMP in the context of reliability.

Pievatolo and Valadè (2003) assess the reliability of electrical systems in situations of continuous operation. An analytical model is developed which allows for non-exponential distributions of failure and repair times. SMP are used to compute the mean time between failures (MTBF) and mean time to repair (MTTR) of a compensator output voltage.

El-Gohary (2004) presents maximum likelihood and Bayesian estimators for reliability parameters of semi-Markovian models. Other recent works that have SMP as main issue are Afchain (2004), Chen and Trivedi (2005), Limnios and Ouhbi (2006), Xie et al. (2005), Soszynska (2006) and Jenab and Dhillon (2006).

A common characteristic of the aforementioned works is that defining an SMP requires the specification of N^2 probabilities of the embedded MP and N^2 conditional probability density functions of the sojourn times in each state given the next state. This is the usual definition of SMP which is presented in most of related literature, for example in Ross (1997) and Limnios and Oprisan (2001).

However, in the context of reliability engineering, transition rates rather than transition probabilities are also usually employed to define continuous time MP and, therefore, transition rates should be attractive for defining SMP as well. Indeed, Becker et al. (2000) develop the mathematical formulation of SMP described through transition rates. Such transition rates are different from those of MP which are either constant (homogeneous Markov processes) or dependent on process time (non-homogeneous Markov processes).

In fact, the transition rates of an SMP may only depend on sojourn time in a state for the case of an HSMP, or both sojourn and process times for an NHSMP. In both cases, the transition rates can be used to represent failure and repair rates as for MP.

Both ways (transition probabilities and transition rates) will be used in next two sections to define SMP. In this way, the mathematical and numerical developments which will be proposed in chapters 3 and 5 could address SMP described through either transition probabilities or transition rates in the same fashion.

2.2. Homogeneous semi-Markov processes

HSMP in continuous time are introduced in this section using a similar nomenclature to the one given in Corradi et al. (2004). Let $S=\{1,...,N\}$ represent the finite state space and define the following random variables:

$$Z_n:\Omega\to S, T_n:\Omega\to[0,\infty[,$$

where Z_n and T_n are, respectively, the state and the time in the nth transition.

The process (Z_n, T_n) is called homogeneous Markov renewal process if

$$\begin{aligned}&\Pr[Z_{n+1}=j, T_{n+1}-T_n\le t \mid Z_n, T_n, Z_{n-1}, T_{n-1}, ..., Z_0, T_0]\\ &\qquad = \Pr[Z_{n+1}=j, T_{n+1}-T_n\le t \mid Z_n=i]\end{aligned}.$$

The kernel $C_{ij}(t)$ of an HSMP is defined as:

$$C_{ij}(t)=\Pr[Z_{n+1}=j, T_{n+1}-T_n\le t \mid Z_n=i]. \quad (2\text{-}1)$$

Eq. (2-1) is the probability of the HSMP to reach state j at time T_{n+1} given that it has remained in state i for $T_{n+1}-T_n\le t$. According to Howard (2007), the kernel $C_{ij}(t)$ is the fundamental describer of an HSMP as its elements determine the transitions between states as well as the sojourn time (t) both conditioned on the current state (i).

It follows that:

$$p_{ij} = \Pr[Z_n = j \mid Z_{n-1} = i] = \lim_{t\to\infty} C_{ij}(t),\ i, j \in S\,,$$

where $P = [p_{ij}]$ is the matrix of transition probabilities of the continuous-time embedded homogeneous Markov process (EHMP), which is the homogeneous Markov process relevant to the HSMP.

HSMP will leave state i after it has stayed there for t with probability given by:

$$F_i(t) = \Pr\left[T_{n+1} - T_n \le t \mid Z_n = i\right], \qquad (2\text{-}2)$$

which represents the CDF of the waiting time in state i.

Eq. (2-1) and (2-2) are related as follows:

$$F_i(t) = \sum\nolimits_{j=1}^{N} C_{ij}(t)\,.$$

In fact, $F_i(t)$ means the probability that the HSMP leaves state i when its successor state j is unknown.

The conditional CDF of the sojourn time given the current *(i)* and next states (j) to be occupied by the process is given as:

$$G_{ij}(t) = \Pr\left[T_{n+1} - T_n \le t \mid Z_n = i, Z_{n+1} = j\right],$$

which corresponds to CDF of the holding time given i and j.

The probabilities are related as follows:

$$G_{ij}(t) = \begin{cases} \dfrac{C_{ij}(t)}{p_{ij}} & , \textit{if } p_{ij} \ne 0, \\ 1 & , \textit{otherwise.} \end{cases}$$

Basically, an HSMP works in the following way: when state i is reached, the next state j to be occupied by the process is immediately drawn from the transition probabilities p_{ij} of the EHMP. Given the current (i) and next (j) states, the sojourn time (t) in state (i) is sampled from the CDF $G_{ij}(t)$. Thus, the next transition time (t_{n+1}) is determined as $t_{n+1} = t_n + t$.

The future behavior of an HSMP over time may be assessed through its interval transition probabilities $\phi_{ij}(t)=Pr[Z_t=j \mid Z_0=i]$, $Z = (Z_t,\ t \in R_0^+)$, which are given as follows (see Corradi et al. (2004)):

$$\phi_{ij}(t) = \delta_{ij}\left(1 - F_i(t)\right) + \sum_{k=1}^{N} \int_0^t \dot{C}_{ik}(\tau) \cdot \phi_{kj}(t-\tau)\, d\tau \qquad (2\text{-}3)$$

where $\dot{C}_{ij}(t) = p_{ij} \cdot \dfrac{d\left[G_{ij}(t)\right]}{dt}$ is the derivative of the kernel of the HSMP in relation to the sojourn time t, and δ_{ij} is the Kronecker's delta for which holds $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$, otherwise. Eq. (2-3) assumes that the kernel $C_{ij}(t)$ is absolutely continuous with respect to the sojourn time.

Eq. (2-3), which is a set of convolution integral equations, is interpreted as follows: the first part represents the probability of the process to remain in state i from 0 to t, with no state change in this time interval. The second part represents the probability of the process to stay in state i during the sojourn time τ, transiting to the intermediary state k at this time and from this state to j at time t, remaining $(t - \tau)$ in state k with $t > \tau$.

When an HSMP is defined in this way, it is said that this process is described through transition probabilities. However, in the context of reliability engineering, transition rates $\lambda_{ij}(t)$ rather than transition probabilities could also be attractive to define HSMP. Indeed, Becker et al. (2000) and Ouhbi and Limnios (1999) have modified eq. (2-3) in order to handle HSMP described through transition rates as follows:

$$\begin{aligned}\phi_{ij}(t) &= \delta_{ij}\exp(-\int_0^t \lambda_i(x)dx) + \\ &+\sum_{k=1}^{N}\int_0^t \lambda_{ik}(\tau)\exp(-\int_0^\tau \lambda_i(x)dx)\phi_{kj}(t-\tau)d\tau\end{aligned}, \quad (2\text{-}4)$$

where $\lambda_{ij}(t)$ is the transition rate of an HSMP defined as:

$$\lambda_{ij}(t)dt = P\{T_n - T_{n-1} \in (t, t+dt) \cap Z(T_n) = j \mid Z(T_{n-1}) = i \cap T_n - T_{n-1} > t\}, \quad (2\text{-}5)$$

Eq. (2-5) indicates that a transition to state j occurs in an infinitesimal time interval after the process has remained in state i for duration t, given that no transition leaving this state has occurred. Moreover, $\lambda_i(\cdot)$ is the transition rate leaving state i and is given by the following equation:

$$\lambda_i(\cdot) = \sum_{k=1}^{N}\lambda_{ik}(\cdot).$$

The interpretation of eq. (2-4) is the same as the one provided for (2-3). However, the kernel $C_{ij}(\cdot)$ and the CDF $F_i(\cdot)$ are defined in a different way as follows:

$$C_{ij}(t) = \int_0^t \lambda_{ij}(z)\exp\left(-\int_0^z \lambda_i(x)dx\right)dz \quad (2\text{-}6)$$

and

$$F_i(t) = 1 - \exp(-\int_0^t \lambda_i(x)dx). \quad (2\text{-}7)$$

Corradi et al. (2004) have developed a numerical method for HSMP directly applying a general numerical quadrature method to equation (2-3) (i.e., only for NHSMP described by transition probabilities). However, the computational cost of this numerical solution is considerably high mainly because it involves solving N^2 coupled integral equations with one variable in the time domain, t.

Although the numerical method proposed by Corradi et al. (2004) has been originally developed to handle HSMP described through transition probabilities (eq. (2-3)), it is likely to be extended to address HSMP described through transition rates for which the interval transition probabilities are given in equation (2-4).

Thus, in chapter 3 (section 3.1) it is presented a novel mathematical formulation for HSMP (described through either transition probabilities (eqs. (2-1) and (2-2)) or transition rates (eqs. (2-6) and (2-7))) as an initial value problem involving transition frequency densities. Moreover, in the same chapter a numerical and straightforward treatment for this new mathematical is drawn as an attempt to reduce the inherent computational cost that is present in the solution of HSMP through the N^2-method. As it said, this approach is called *2N*-due to its complexity.

Moreover, in chapter 3 (section 3.2) there will be proposed an alternative method to the *2N*- and N^2-methods for solving the interval convolution transition probability equations of an HSMP on continuous time. Taking advantage of the convolution feature present in homogenous environment, this numerical procedure is based on the application of LT which will be inverted by using the Gauss quadrature method known as Gauss Legendre. Basically, LT plays an important role since they will change the integral domain by an algebraic environment which is likely to reduce the computational time of the solution.

Comparisons in terms of computational time and accuracy among the N^2-, *2N*- and *Lap*-methods and Monte Carlo simulation will be accomplished in chapter 4 in order to validate the effectiveness of the proposed models for solving HSMP.

2.3. Non-Homogeneous semi-Markov processes

NHSMP are introduced here using a similar nomenclature to the one given in Janssen and Manca (2001). Thus, let define the following random variables:

$Z_n : \Omega \to S, T_n : \Omega \to [0, \infty[,$

where Z_n, T_n and $X_n = T_n - T_{n-1}$ are the state, process time, and sojourn time in the n^{th} transition, respectively.

The process (Z_n, T_n) is called non-homogeneous Markov renewal process if

$$\Pr[Z_{n+1}=j,T_{n+1}\le t\,|\,Z_n=i,T_n=l,Z_{n-1},T_{n-1},...,Z_0,T_0]$$
$$=\Pr[Z_{n+1}=j,T_{n+1}\le t\,|\,Z_n=i,T_n=l].$$

The kernel $C_{ij}(\cdot,\cdot)$ of an NHSMP is defined as:

$$C_{ij}(l,t)=\Pr[Z_{n+1}=j,T_{n+1}\le t\,|\,Z_n=i,T_n=l]. \tag{2-8}$$

Eq. (2-8) is the probability of the process to reach state *j* at the time $T_{n+1}\le t$ given that it has reached state *i* at the time *l*, and remained there for $X_{n+1}\le x$. The kernel $C_{ij}(\cdot,\cdot)$ is the fundamental describer of an NHSMP as its elements determine the transitions between states, the time of the next transition (*t*) and then sojourn time (*x*) conditioned on the current state (*i*) and the last transition time (*l*).

It follows that:

$$p_{ij}(l)=\Pr[Z_n=j\,|\,Z_{n-1}=i,T_{n-1}=l]=\lim_{t\to\infty}C_{ij}(l,t),\ i,j\in S,$$

where $P(l)=[p_{ij}(l)]$ is the matrix of transition probabilities of the continuous-time embedded non-homogeneous Markov process (ENHMP), which is the non-homogeneous Markov process associated to the NHSMP.

NHSMP will leave state *i* within the time interval from *l* to *t* with probability given by:

$$F_i(l,t)=\Pr\left[T_{n+1}\le t\,|\,Z_n=i,T_n=l\right], \tag{2-9}$$

which represents the CDF of the waiting time in state *i*.

Eqs. (2-8) and (2-9) are related as follows:

$$F_i(l,t)=\sum_{j=1}^{N}C_{ij}(l,t).$$

In fact, $F_i(\cdot,\cdot)$ means the probability that the NHSMP leaves state *i* when its successor state *j* is unknown.

The conditional CDF of the sojourn time in the current state (*i*) given the next state (*j*) to be occupied by the process and the last transition time (*l*) is given as:

$$G_{ij}(l,t)=\Pr\left[T_{n+1}\le t\,|\,Z_n=i,Z_{n+1}=j,T_n=l\right].$$

The probabilities are related as follows:

$$G_{ij}(l,t)=\begin{cases}\dfrac{C_{ij}(l,t)}{p_{ij}(l)}, & if\ p_{ij}(l)\neq 0,\\ 1 & , otherwise.\end{cases}$$

According to D'amico et al. (2005), the main difference between a non-homogeneous Markov process and an NHSMP is on the CDF $G_{ij}(l,t)$. In a Markovian environment, such functions must be exponential negative whereas in a semi-Markov context $G_{ij}(l,t)$ may be arbitrary and not necessarily exponential.

Basically, an NHSMP works in the following way: when state *i* is reached at the time *l*, the next state *j* to be occupied by the process is immediately drawn from the transition probabilities $p_{ij}(l)$ of the ENHMP. Given the current (*i*) and next (*j*) states and the last transition time (*l*), the sojourn time (*x*) in state (*i*) is sampled from the CDF $G_{ij}(\cdot,\cdot)$. Thus, the next transition time (*t*) is determined as $t = l + x$.

The future behavior of NHSMP over time is assessed through its interval transition probabilities $\phi_{ij}(l,t)=\Pr[Z_t=j \mid Z_l=i]$, $Z = (Z_t,\ t \in R_0^+)$, which are given as follows (see Janssen and Manca (2001)):

$$\phi_{ij}(l,t) = \delta_{ij}\left(1 - F_i(l,t)\right) + \sum_{k=1}^{N}\int_{l}^{t}\dot{C}_{ik}(l,\tau)\cdot\phi_{kj}(\tau,t)\,d\tau \tag{2-10}$$

where $\dot{C}_{ij}(l,t) = \dfrac{d\left[p_{ij}(l)\cdot G_{ij}(l,t)\right]}{dl}$ is the derivative of the kernel of the NHSMP in relation to *l*, and δ_{ij} is the Kronecker's delta for which holds $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$, otherwise. Eq. (2-10) assumes that the kernel $C_{ij}(\cdot,\cdot)$ is absolutely continuous with respect to the process time.

Eq. (2-10) is interpreted as follows. The first part represents the probability of the process to remain in state *i* from *l* to *t*, with no state change in this time interval. The second part represents the probability of the process to reach state *i* at the time *l*, and the intermediary state *k* at the time τ, and to transit from this state to *j* at the time *t*, remaining ($x = \tau - l$) in the state *i* and ($x = t - \tau$) in the state *k* before reaching the state *j* at *t*, with $t > \tau > l$.

Transition rates that depend on both types of time variables (sojourn and process) may occur in problems of practical interest, for example, in FTS which are under deteriorating processes (see Moura (2006), for example). Therefore, it is necessary to develop a non-homogeneous semi-Markovian model described by transition rates which depend on both sojourn and process times.

Becker et al. (2000) defined the transition rate of an NHSMP as follows:

$$\begin{aligned}\lambda_{ij}(l,t)dx = \Pr\{&t - l \in (x, x+dx) \cap Z(t) = j \mid Z(l) \\ &= i \cap t - l > x\}.\end{aligned} \tag{2-11}$$

Eq. (2-11) represents a transition to state *j* which occurs in an infinitesimal time interval after the process has reached state *i* at time *l*, given that no transition leaving such a state occurred before. Note that both *t* and *l* represent process times, but *l* is the time of the last transition so that $t=x+l$ as occurs for NHSMP described by transition probabilities.

To obtain the interval transition probabilities of an NHSMP defined by transition rates, one has to consider that the process is non-homogeneous in relation to the process time. Thus, $\phi(l,t)$ are given as follows:

$$\phi_{ij}(l,t) = \delta_{ij} \exp\left(-\int_l^t \lambda_i(l,t)\,dx\right) + \sum_{k=1}^{N} \int_l^t \lambda_{ik}(l,\tau) \exp\left(-\int_l^\tau \lambda_i(l,x)\,dx\right) \phi_{kj}(\tau,t)\,d\tau. \tag{2-12}$$

Eq. (2-12) means the same of eq. (2-10). The only idiosyncrasy regards the kernel $C_{ij}(\cdot)$ and the CDF $F_i(\cdot)$ since they are defined in a different way as follows:

$$C_{ij}(l,t) = \int_l^t \lambda_{ik}(l,\tau) \exp\left(-\int_l^\tau \lambda_i(l,\tau)\,dx\right) dz \tag{2-13}$$

and

$$F_i(l,t) = 1 - \exp\left(-\int_l^t \lambda_i(l,t)\,dx\right). \tag{2-14}$$

Note that eqs. (2-10) and (2-12) are not of convolution type as is for the case of homogeneous semi-Markov processes (see previous section). In this way, the LT technique that is widely applied for HSMP could not be used in the non-homogeneous environment.

Therefore, Janssen and Manca (2001) have developed the N^2-numerical method for NHSMP directly applying a general numerical quadrature method to eq. (2-10) (i.e., only for NHSMP described by transition probabilities). However, the computational cost of this numerical solution is considerably high mainly because it involves the solution of N^2 non-convolution coupled integral equations and with two variables in the time domain, *l* and *t*.

Thus, in chapter 5 it is developed a novel mathematical formulation for NHSMP (described through either transition probabilities (eqs. (2-8) and (2-9)) or transition rates (eqs. (2-13) and (2-14)) as an initial value problem involving transition frequency densities. Moreover, in the same chapter a numerical and straightforward treatment for this new mathematical is drawn as an attempt to reduce the inherent computational cost that plagues the solution of NHSMP through the N^2-method. This approach is called *2N-* due to its complexity.

3. SOLVING HOMOGENEOUS SEMI-MARKOV PROCESSES: 2N- AND LAP- APPROACHES

3.1. *2N*-method: Mathematical Formulation and Numerical Treatment

The mathematical and numerical treatments, which is called *2N*-method and will be developed in this chapter, is put forward as an attempt for untangling the inherent computational cost that plagues the solution of HSMP via the classical method given in Corradi et al. (2004).

Basically, by changing N^2-effort by *2N*-complexity, the *2N*-approach tends to reduce considerably the computational effort in relation to the N^2-method and MC simulation as well, keeping the accuracy of the abovementioned approaches. The main findings of this section can be encountered in Moura and Droguett (2009d) which follows in attachment D.

3.1.1. An initial value problem involving transition frequency densities

Depending on how an HSMP is described, the kernel $C_{ij}(\cdot)$ and the CDF $F_i(\cdot)$ are given by equations (2-1) and (2-2) in case of transition probabilities or by equations (2-6) and (2-7) in case of transition rates. The *2N*-numerical approach will be developed in a general way in order to handle both situations.

Thus, an initial value problem consists of computing the probabilities $\phi_j(t)$ given the initial conditions $\phi_j(0)$. By using a similar nomenclature to the one given in Becker et al. (2000), let $N_j(t)$ be the number of times that state *j* of an HSMP is visited from any state in the interval $[0,t]$. Let also $H_j(t) = E[N_j(t)]$ be its expected value. If $H_j(t)$ is continuously differentiable, then $h_j(t)dt = dH_j(t)$ is its corresponding density function.

As the stochastic process under consideration is regular, i.e., no more than one transition can occur in any interval $(t, t+dt)$, then $h_j(t)$ can be assumed as the probability that a transition occurs to state *j* in an infinitesimal time interval *dt* as follows:

$h_j(t)dt = Pr\{to\ reach\ state\ j\ in\ (t,\ t+dt)\}$.

Thus, it follows that:

$$h_j(t) = \sum_{i=1}^{N} \phi_i(0) \cdot \dot{C}_{ij}(t) + \sum_{i=1}^{N} \int_0^t h_i(\tau) \cdot \dot{C}_{ij}(t-\tau)\, d\tau. \quad (3\text{-}1)$$

According to the description of eq. (2-1), (3-1) means state *j* can be reached either if the process was initially in state *i* and remains there until time *t*, when a transition to state *j* occurs; or if the process reached state *i* at time τ, remaining there for $x = t - \tau$, then a transition to state j occurs. The summation over the state number N is for all possible intermediary states, where HSMP can transit. The integral term in turn means that the transition to state i may occur in any time $\tau \in [0,t]$.

Eq. (3-1) corresponds to a system of N integral equations with unknowns $h_j(t)$, $j = 1,...,N$. The probabilities $\phi_j(t)=\Pr[Z_t=j \mid Z_0]$ can be obtained from the initial conditions $\phi_j(0)$ as follows:

$$\phi_j(t) = \phi_j(0)\cdot[1-F_j(t)] + \int_0^t h_j(\tau)\cdot[1-F_j(t-\tau)]d\tau \,. \tag{3-2}$$

Eq. (3-2) argues that a process can be in state j in time t either if it was initially in the state j and remained there at least up to time t; or if it visited state j at any time $\tau \in [0,t]$ with probability $h_j(t)$ and stayed there for $(t - \tau)$. Eq. (3-2) corresponds to N straightforward integrations that can be solved independently by using the solution of equation (3-1).

Thus, eqs. (3-1) and (3-2) extend the formulation presented in Becker et al. (2000) in order to address HSMP described in terms of both transition probabilities and transition rates.

The computational effort to solve eqs. (3-1) and (3-2) is less demanding than in the case of the N^2-method described in Corradi et al. (2004). In fact, the *2N*-method consists of solving N coupled convolution integral equations with one variable (eq. (3-1)) and N straightforward integrations (eq. (3-2)) whereas the N^2-method consists of solving N^2 coupled convolution integral equations. Comparisons among the *2N*-approach, the N^2-method and the Monte Carlo-based solution will be presented in chapter 4. Before that, a general quadrature based method for simultaneously solving equations (3-1) and (3-2) is presented in next section.

3.1.2. Numerical formulation

3.1.2.1. Description of the numerical solution

A numerical integration or quadrature method can be written as follows (see Press et al. (2002)):

$$\int_0^{kh} \varphi(\tau)d\tau \cong \sum_{\tau=0}^{k} w_{\tau,k}\cdot\varphi(\tau h)\,, \tag{3-3}$$

where h is the step measure and $w_{\tau,k}$ are the weights related to the quadrature formula (3-3). Note that such weights do not depend on the integrand function $\varphi(\cdot)$; they are function of the start point (0), of the end (kh) and the intermediary (τh) points at which the function value is computed. Moreover, one has M such that $Mh = T$, with $0 \le kh \le T$, $k \le M, k, M \in \mathbb{N}$ where $[0,T]$ is the integration interval, M is the number of steps and h is the step interval.

Using (3-3), a solution for (3-1) is given as:

$$\tilde{h}_j(kh) = a_j(kh) + \sum_{i=1}^{N}\left(\sum_{\tau=0}^{k} w_{\tau,k}\cdot\tilde{h}_i(\tau h)\cdot\dot{C}_{ij}((k-\tau)h)\right) \tag{3-4}$$

where

$$a_j(kh) = \sum_{i=1}^{N}\phi_i(0)\cdot\dot{C}_{ij}(kh), \tag{3-5}$$

where the notation $\sim$ means an approximation, i.e., $\tilde{h}_j(kh)$ is a numerical approximation for $h_j(kh)$.

The system of eqs. (3-4) can be written as follows:

$$\tilde{h}_j(kh) - \sum_{i=1}^{N}\left(\sum_{\tau=0}^{k} w_{\tau,k}\cdot\tilde{h}_i(\tau h)\cdot\dot{C}_{ij}((k-\tau)h)\right) = a_j(kh) \tag{3-6}$$

or in matrix form:

$$\widetilde{H}^T_{(kh)} - \sum_{\tau=0}^{k}\Psi^T_{((k-\tau)h)}\cdot\widetilde{H}^T_{(\tau h)} = A^T_{(kh)}, \tag{3-7}$$

where the symbol ^{T} represents the transpose matrix, $\widetilde{H}^T_{(\cdot)}$ and $A^T_{(\cdot)}$ are N-order matrices, and $\Psi^T_{(\cdot)}$ is N^2-order matrix, where:

$$\psi_{ij}(\cdot) = w_{\tau,k}\,\dot{C}_{ij}(\cdot). \tag{3-8}$$

Alternatively, eq. (3-7) can be rewritten as follows:

$$\mathbf{U}^T\cdot\widetilde{\mathbf{H}}^T = \mathbf{A}^T, \tag{3-9}$$

where

$$\boldsymbol{U^T} = \begin{bmatrix} I-\Psi^T_{(0h)} & 0 & \cdots & 0 & \cdots & 0 \\ -\Psi^T_{(1h)} & I-\Psi^T_{(0h)} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\Psi^T_{(\tau h)} & -\Psi^T_{((\tau-1)h)} & \cdots & I-\Psi^T_{(0h)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\Psi^T_{(Mh)} & -\Psi^T_{((M-1)h)} & \cdots & -\Psi^T_{((M-\tau)h)} & \cdots & I-\Psi^T_{(0h)} \end{bmatrix},$$

$$\widetilde{\boldsymbol{H}}^T = \begin{bmatrix} \widetilde{H}^T_{(0h)} \\ \widetilde{H}^T_{(1h)} \\ \ddots \\ \widetilde{H}^T_{(\tau h)} \\ \ddots \\ \widetilde{H}^T_{(Mh)} \end{bmatrix} \quad \text{and} \quad \boldsymbol{A^T} = \begin{bmatrix} A^T_{(0h)} \\ A^T_{(1h)} \\ \ddots \\ A^T_{(\tau h)} \\ \ddots \\ A^T_{(Mh)} \end{bmatrix}. \tag{3-10}$$

A bold face notation is used for a matrix of matrices as in $\mathbf{U}^T, \widetilde{\mathbf{H}}^T$, and $\mathbf{A}^T$. Eq. (3-9) is used to compute the *N*-order matrices $\widetilde{H}^T_{(\tau h)}$, whose elements $\tilde{h}_j(\tau h)$ are the density functions of the number of times that state *j* of an HSMP is visited from any state in the interval [0,*τh*], with $j = 1,...,N$ and $\tau = 0,...,M$.

Having estimated the matrix $\widetilde{\mathbf{H}}^T$, the next step consists in computing the *N* state probabilities $\tilde{\phi}_j(t)$. In fact, a numerical solution for the state probabilities $\tilde{\phi}_j(t)$ can be computed as follows:

$$\tilde{\phi}_j(kh) = \phi_j(0)\cdot[1-F_j(kh)] + \sum_{\tau=0}^{k} w_{\tau,k}\cdot\tilde{h}_j(\tau h)\cdot[1-F_j((k-\tau)h)], \tag{3-11}$$

where $\tilde{h}_j(\tau h)$ (*j = 1,…,N* and *τ = 0,…,M*) are step-solutions obtained from eq. (3-8). In this way, $\widetilde{\mathbf{\Phi}} = \left[\widetilde{\Phi}_{(0h)},...,\widetilde{\Phi}_{(\tau h)},...,\widetilde{\Phi}_{(Mh)}\right]$ is comprised of *M* matrices $\widetilde{\Phi}_{(\tau h)}$ each one of order *N*.

Note that, rather than solving the N^2 coupled integral equations through the direct application of any general quadrature method to either equation (2-10) or (2-12), the approach just described estimates the state probabilities of an HSMP by solving *N* coupled integral equations (eq. (3-6)) and *N* straightforward integrations (eq. (3-11)).

For example, by using the extended Simpson's rule given in Press et al. (2002) (p. 138), eqs. (3-6) and (3-11) can be written in the following way:

$$\begin{aligned}\tilde{h}_j(kh) = a_j(kh) &+ \frac{h}{3}\sum_{i=1}^{N}\tilde{h}_i(0h)\cdot\dot{C}_{ij}(kh) \\ &+\frac{4h}{3}\sum_{i=1}^{N}\sum_{\tau=1}^{\frac{k}{2}}\tilde{h}_i((2\tau-1)h)\cdot\dot{C}_{ij}((k-2\tau+1)h \\ &+\frac{2h}{3}\sum_{i=1}^{N}\sum_{\tau=1}^{\frac{k}{2}-1}\tilde{h}_i(2\tau h)\cdot\dot{C}_{ij}((k-2\tau)h) \\ &+\frac{h}{3}\sum_{i=1}^{N}\tilde{h}_i(kh)\cdot\dot{C}_{ij}(0h)\end{aligned}$$

and

$$\tilde{\phi}_j(kh) = \phi_j(0)\cdot[1-F_j(kh)] + \frac{h}{3}\cdot\tilde{h}_j(0h)\cdot[1-F_j(kh)] + \frac{4h}{3}\cdot\sum_{\tau=1}^{\frac{k}{2}}\tilde{h}_j((2\tau-1)h)\cdot[1-F_j((k-2\tau+1)h)] + \frac{2h}{3}\cdot\sum_{\tau=1}^{\frac{k}{2}-1}\tilde{h}_j(2\tau h)\cdot[1-F_j((k-2\tau)h)] + \frac{h}{3}\cdot\tilde{h}_j(kh)\cdot[1-F_j(0h)]. \quad (3\text{-}12)$$

3.1.2.2. Solution conditions and upper limit estimate of the discretization error

From eq. (3-4) follows that:

$$\tilde{h}_j(kh) - \sum_{i=1}^{N} w_{k,k}\cdot\tilde{h}_i(kh)\cdot\dot{C}_{ij}(0h) = a_j(kh) + \sum_{i=1}^{N}\left(\sum_{\tau=0}^{k-1} w_{\tau,k}\cdot\tilde{h}_i(\tau h)\cdot\dot{C}_{ij}((k-\tau)h)\right). \quad (3\text{-}13)$$

By writing eq. (3-13) in terms of $j = 1,\ldots,N$, it follows that:

$$\tilde{h}_1(kh) - \sum_{i=1}^{N} w_{k,k}\cdot\tilde{h}_i(kh)\cdot\dot{C}_{i1}(0h) = d_1$$

$$\tilde{h}_2(kh) - \sum_{i=1}^{N} w_{k,k}\cdot\tilde{h}_i(kh)\cdot\dot{C}_{i2}(0h) = d_2$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$\tilde{h}_N(kh) - \sum_{i=1}^{N} w_{k,k}\cdot\tilde{h}_i(kh)\cdot\dot{C}_{iN}(0h) = d_N, \quad (3\text{-}14)$$

where d_j, $j = 1,\ldots,N$ is the right-hand side term of equation (3-13).

According to Press et al. (2002), it is guaranteed that the set of linear algebraic equations given in eq. (3-14) has solution if the matrix of known coefficients

$$\begin{bmatrix} 1-\psi_{11}(0h) & -\psi_{12}(0h) & \cdots & -\psi_{1j}(0h) & \cdots & -\psi_{1N}(0h) \\ -\psi_{21}(0h) & 1-\psi_{22}(0h) & \cdots & -\psi_{2j}(0h) & \cdots & -\psi_{2N}(0h) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\psi_{i1}(0h) & -\psi_{j2}(0h) & \cdots & 1-\psi_{ij}(0h) & \cdots & -\psi_{iN}(0h) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\psi_{N1}(0h) & -\psi_{N2}(0h) & \cdots & -\psi_{Nj}(0h) & \cdots & 1-\psi_{NN}(0h) \end{bmatrix},$$

is not degenerate, i.e., is not singular, where $\psi_{ij}(\cdot)$ is given in eq. (3-8). The non-singularity happens if none of the N eqs. in (3-14) is a linear combination of the others. If this condition is satisfied, the eq. (3-11) of unknowns $\tilde{\phi}_j$ also admits solution given the solution of (3-4).

In the subsequent development, we make use of the following Lemmas 1 and 2 and Remark 1 from Baker (1977), and described in Corradi et al. (2004), that also hold for the *2N*-mathematical formulation:

Lemma 1 (Baker (1977), p. 925). If $|\xi_v| \le A\sum_{i=0}^{v-1}|\xi_i| + B, v = q, q+1,..., q \ge 1$, where $A > 0, B > 0$ and $\sum_{i=0}^{q-1}|\xi_i| \le \xi$, then:

$|\xi_v| \le (A\xi + B)(1+A)^{v-q}, v = q, q+1,....$

Lemma 2 (Baker (1977), p.926). Suppose that: $A = h\overline{L} \le 0$ and $ph = x \ge 0$. Then

$(1+A)^{p-q} \le \exp(\overline{L}x)$ if $p \ge q$.

Remark 1 from Lemmas 1 and 2 results:

$|\xi_v| \le (A\xi + B)(1+A)^{v-q} \le (A\xi + B)\exp(\overline{L}x) = (h\overline{L}\xi + B)\exp(\overline{L}rh)$.

Theorem 1: Let

$h_j(t): [0, Y] \to \mathbb{R}$

Let $q \in \{0,..., M-1\}$, $M \in \mathbb{N}$, such that $0 \le qh \le M$.

Let

$$\xi_j^k(h) = \tilde{h}_j(kh) - h_j(kh) \qquad k = 0,1,..., M, \tag{3-15}$$

where $h_j(kh)$ is solution of equation (3-1) and $\tilde{h}_j(kh)$ is solution of equation (3-4).

Define also:

$$\eta^k(h) = \sum_{i=1}^{N}\left|\xi_i^k(h)\right|.$$

Moreover, let

$$w^l = w_M^l = \max_{0\le\tau\le k\le M}\frac{|w_{\tau,k}|}{h} < \infty \tag{3-16}$$

$$t_j^k(h) = \sum_{i=1}^{N}\int_0^{kh} h_i(\vartheta)\cdot\dot{C}_{ij}(kh-\vartheta)\,d\vartheta - \sum_{i=1}^{N}\left[\sum_{\tau=0}^{k} w_{\tau,k}\cdot h_i(\tau h)\cdot\dot{C}_{ij}((k-\tau)h)\right] \tag{3-17}$$

$$\sigma^k(h) = \sum_{i=1}^{N}\left|t_i^k(h)\right|,$$

$$\rho^l(h) = \rho_M^l(h) = \max_{q\le k\le M}\sigma^k(h) \tag{3-18}$$

$$\xi^l(h) = \sum_{\tau=0}^{q-1}\eta^\tau(h).$$

Suppose that $\left|\dot{C}_{ij}(t)\right| \le d_1$ for $t \in [0, T]$.

Then

$$\eta^k(h) \le \left(\frac{\rho^l(h) + Nhw_M^l d_1 \xi^l(h)}{1 - Nhw_M^l d_1} \right) \exp\left(\frac{Nd_1 w_M^l kh}{1 - Nd_1 w_M^l h} \right), \tag{3-19}$$

$k = q, q+1, \ldots, M$

given that $Nhw_M^l d_1 < 1$.

Proof: From eq. (3-4) it follows that for $k \ge q$:

$$\begin{aligned} & h_j(kh) - \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}((k-\tau)h) \right) = a_j(kh) \\ & \quad + \sum_{i=1}^{N} \int_0^{kh} h_i(\vartheta) \cdot \dot{C}_{ij}(kh - \vartheta) d\vartheta - \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}((k-\tau)h) \right) \end{aligned} . \tag{3-20}$$

From eqs. (3-4) and (3-20), it follows that:

$$\begin{aligned} & \tilde{h}_j(kh) - h_j(kh) + \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}((k-\tau)h) \right) \\ & \quad = \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot \tilde{h}_i(\tau h) \cdot \dot{C}_{ij}((k-\tau)h) \right) - \sum_{i=1}^{N} \int_0^{kh} h_i(\vartheta) \cdot \dot{C}_{ij}(kh - \vartheta) d\vartheta \\ & \quad + \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}((k-\tau)h) \right) \end{aligned}$$

and from eqs. (3-15) and (3-17):

$$\begin{aligned} & \xi_j^k(h) - \sum_{i=1}^{N} \left[\sum_{\tau=0}^{k} w_{\tau,k} \cdot \tilde{h}_i(\tau h) \cdot \dot{C}_{ij}((k-\tau)h) - \sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}((k-\tau)h) \right] \\ & \quad = -t_j^k(h). \end{aligned}$$

Hence,

$$\xi_j^k(h) - \sum_{i=1}^{N} \left\{ \sum_{\tau=0}^{k} w_{\tau,k} \cdot \dot{C}_{ij}((k-\tau)h) \cdot \left[\tilde{h}_i(\tau h) - h_i(\tau h) \right] \right\} = -t_j^k(h)$$

and

$$\xi_j^k(h) - \sum_{i=1}^{N} \left[\sum_{\tau=0}^{k} w_{\tau,k} \cdot \dot{C}_{ij}((k-\tau)h) \cdot \xi_i^\tau(h) \right] = -t_j^k(h).$$

It follows from eq. (3-16) that:

$$\begin{aligned} \left| \xi_j^k(h) \right| & \le \left| t_j^k(h) \right| + \sum_{i=1}^{N} \sum_{\tau=0}^{k} \left| w_{\tau,k} \right| \cdot \left| \dot{C}_{ij}((k-\tau)h) \right| \cdot \left| \xi_i^\tau(h) \right| \\ & \le \left| t_j^k(h) \right| + d_1 w_M h \sum_{i=1}^{N} \sum_{\tau=0}^{k} \left| \xi_i^\tau(h) \right|. \end{aligned} .$$

Therefore,

$$\left| \xi_j^k(h) \right| \le \left| t_j^k(h) \right| + d_1 w_M h \sum_{\tau=0}^{k} \sum_{i=1}^{N} \left| \xi_i^\tau(h) \right|.$$

Performing the summation in relation to the index *j = 1,…,N*, one obtains that:

$$\sum_{j=1}^{N}\left|\xi_j^k(h)\right| \le \sum_{j=1}^{N}\left|t_j^k(h)\right| + Nd_1 w_M h \sum_{\tau=0}^{k}\sum_{i=1}^{N}\left|\xi_i^\tau(h)\right|. \qquad (3\text{-}21)$$

Eq. (3-21) implies that:

$$\eta^k(h) \le \sigma^k(h) + Nd_1 w_M h \sum_{\tau=0}^{k}\eta^\tau(h)$$

and

$$(1 - Nd_1 w_M h)\eta^k(h) \le \sigma^k(h) + Nd_1 w_M h \sum_{\tau=0}^{k-1}\eta^\tau(h).$$

From eq. (3-18), it follows that:

$$(1 - Nd_1 w_M h)\eta^k(h) \le \rho_M(h) + Nd_1 w_M h \sum_{\tau=0}^{k-1}\eta^\tau(h).$$

Now, from Lemmas 1 and 2, it follows that:

$$A = \frac{Nd_1 w_M h}{1 - Nd_1 w_M h},$$

$$B = \frac{\rho_M(h)}{1 - Nd_1 w_M h}$$

From Remark 1, it follows that:

$$\overline{L} = \frac{A}{h} = \frac{Nd_1 w_M}{1 - Nd_1 w_M h}.$$

Therefore, one has that:

$$\eta^k(h) \le (A\xi(h) + B)\exp\left(\frac{Nd_1 w_M kh}{1 - Nd_1 w_M h}\right)$$

what implies that:

$$\eta^k(h) \le \left(\frac{\rho_M(h) + Nhw_M d_1 \xi(h)}{1 - Nhw_M d_1}\right)\exp\left(\frac{Nd_1 w_M kh}{1 - Nd_1 w_M h}\right). \qquad \bullet$$

This proves the result in eq. (3-19). The same conclusion was reached in Corradi et al. (2004) for the N^2-method. In other words, the same upper limit of the discretization error in solving the interval transition probabilities via N^2-method is found when $h_j(t)$ is approximated via eq. (3-4).

However, the upper limit of the *2N*-discretization error function should also take into account the error made in eq. (3-11). Actually, the error estimate $\varsigma_j^k(h) = \tilde{\phi}_j(kh) - \phi_j(kh)$ in computing $\tilde{\phi}_j(t)$ through eq. (3-11) given the solution of (3-4) depends only on the choice of the quadrature method. In eq. (3-12), for instance, the error term is equal to $\varsigma_j^k(h) = O_j\left(1/k^4\right)$.

This means that the true answer $\phi_j(kh)$ differs from the estimate $\tilde{\phi}_j(kh)$ by an amount equals to $1/k^4$ (see Press et al. (2002) for further details) which tends to zero when *k* increases.

Disregarding the inequality in eq. (3-19), the most important finding encountered in the preceding developments is that even though *2N*-method has a greater discretization error than N^2-approach their accuracy will approximately be equal when the number *M* of steps increases.

In other words, for the same *M*, the *2N*-approach tends to be significantly less time-consuming and has rough accuracy to the N^2-method. In chapter 5, when the *2N*-method drawn for NHSMP will be discussed the inequality in eq. (3-19) will again be considered.

3.2. *Lap*-method: A Laplace-based numerical procedure to solve the state probability equations of HSMP

In the present section the *Lap*-procedure is discussed. This numerical method is also drawn for HSMP described through either transition probabilities or transition rates. Basically, it is based on the application of Laplace transforms which will be inverted by using the Gauss quadrature method known as Gauss Legendre.

As it will be seen, conversely to the *2N*- and N^2-approaches, the main advantage of this approach is that it is not required adjusting the number *M* of steps in order to obtain desired convergence. There will be a pre-set number of steps, which is independent on the problem to be solved and thus, this method is likely to have a considerable reduced computational effort in relation to the *2N*- and N^2-methods and MC as well. In other words, the features of this method will quit the need of previously specifying a number *M* of steps for each problem, thus reducing time for performing calculus. The main developments of this chapter may be found in Moura and Droguett (2008) and Moura and Droguett (2009b) which follow in attachments A and E respectively.

3.2.1. State probabilities for HSMP via *Lap*-procedure

Basically, eq. (3-1) can be written in matrix form as:

$$\left[h_1(t)\cdots h_N(t)\right]=\left[a_1(t)\cdots a_N(t)\right] + \int_0^t \left[h_1(\tau)\cdots h_N(\tau)\right]\cdot\begin{bmatrix}\dot{C}_{11}(t-\tau) & \cdots & \dot{C}_{1N}(t-\tau)\\ \vdots & \ddots & \vdots\\ \dot{C}_{N1}(t-\tau) & \cdots & \dot{C}_{NN}(t-\tau)\end{bmatrix}d\tau, \quad (3\text{-}22)$$

where $\alpha_j(t)$ is given by eq. (3-5) and $\dot{C}_{ij}(t)$ is the derivative of the HSMP's kernel.

Eq. (3-22) can be rewritten as:

$$\begin{bmatrix} h_1(t) \\ \vdots \\ h_N(t) \end{bmatrix}^T = \begin{bmatrix} a_1(t) \\ \vdots \\ a_N(t) \end{bmatrix}^T + \int_0^t \begin{bmatrix} \dot{C}_{11}(t-\tau) & \cdots & \dot{C}_{1N}(t-\tau) \\ \vdots & \ddots & \vdots \\ \dot{C}_{N1}(t-\tau) & \cdots & \dot{C}_{NN}(t-\tau) \end{bmatrix}^T \cdot \begin{bmatrix} h_1(\tau) \\ \vdots \\ h_N(\tau) \end{bmatrix}^T d\tau \ .$$

where symbol ^{T} represents transpose matrix. Hence, in a more compact way:

$$\left[H(t)\right]^T = \left[A(t)\right]^T + \int_0^t \left[\dot{C}(t-\tau)\right]^T \cdot \left[H(\tau)\right]^T d\tau \ . \tag{3-23}$$

Notice this formulation is general since it can address HSMP described by transition probabilities or transition rates. The only difference is how kernel is defined: either by eq. (2-1) for transition probabilities or (2-6) for transition rates.

In order to compute the state probabilities of HSMP, the *Lap*-procedure is based on the application of the LT to these equations and the corresponding inversion to obtain the solution on time domain t.

Indeed, by applying LT to eq. (3-23), taking into account that the LT of the convolution of two independent functions ($\dot{C}(\cdot)$ and $H(\cdot)$) is equal to the product between their individual LT and using $\widetilde{f}(s)$ as the LT of a function $f(t)$, it follows that:

$$\left[\widetilde{H}(s)\right]^T = \left[\widetilde{A}(s)\right]^T + \left[\widetilde{K}(s)\right]^T * \left[\widetilde{H}(s)\right]^T , \tag{3-24}$$

where s is the transformed variable and $\widetilde{K}(s)$ is the matrix of the LT of kernel derivative $\dot{C}(\cdot)$.

By solving (3-24) for $\widetilde{H}(s)$, it follows that:

$$\left[I - \left[\widetilde{K}(s)\right]^T\right] \cdot \left[\widetilde{H}(s)\right]^T = \left[\widetilde{A}(s)\right]^T . \tag{3-25}$$

Eq. (3-25) corresponds to a system of linear algebraic equations that can be simultaneously solved by using any numerical solution method. The unknowns of (3-25) are the values of $\tilde{h}_j(s)$ (with $j = 1,\ldots,N$) and will be used for computing $\tilde{\phi}_j(s)$ which comes up from applying LT to convolution integral equations given in (3-2) as follows:

$$\tilde{\phi}_j(s) = \phi_j(0) \cdot \tilde{v}_j(s) + \tilde{h}_j(s) * \tilde{v}_j(s) , \tag{3-26}$$

where $\tilde{v}_j(s)$ is the LT of the term $\exp\left(-\int_0^t \lambda_j(x)\,dx\right)$. The values $\tilde{\phi}_j(s)$ represent the solution of (3-26) which can be independently solved for *j=1,…,N* using the values $\tilde{h}_j(s)$ obtained from (3-25).

Given the solution of (3-26) (LT of the state probabilities for HSMP), the problem now consists of inverting the LT to obtain the state probabilities on the time variable *t*. The inversion method on which the *Lap*-method is backed up will be described in next subsection.

3.2.2. Numerical Inversion of Laplace Transforms: Gauss-Legendre based method

The numerical inversion of LT consists of obtaining estimates for *f(t)* given numerical values of the transform function $\tilde{f}(s)$:

$$\tilde{f}(s) = \int_0^\infty e^{-st} f(t)dt, \tag{3-27}$$

where *s* is the transformed variable.

Some methods have been proposed in the literature to solve this problem such as Valkó and Abate (2004), Abate and Valkó (2004), Kryzhniy (2004), Milovanovic and Cvetkovic (2005) and Cuomo et al. (2007).

The numerical inversion method of LT presented here to compute the interval transition probabilities of an HSMP is based on a Gaussian quadrature method known as Gauss Legendre (Bellman et al. (1966) and Abramowitz and Stegun (1972)). Recently, Oliveira et al. (2005) has applied a similar procedure to compute the state probabilities of non-homogeneous MP with supplementary variables. Great results attained on this situation have led to delve on the feasibility of its application for solving HSMP as well, for which the dynamic behavior rise from a generalization of the Kolmogorov backward differential equations of the Markov environment.

Thus, making the change of variables *z=exp(-t)*, eq. (3-27) reduces to a finite Mellin transform (see Haidar (1997)):

$$\tilde{f}(s) = \int_0^1 z^{s-1} f(-\ln(z))dz. \tag{3-28}$$

The integral of the right hand side of (3-28) can be approximated by a Gaussian Quadrature which involves the weighted sum of function $f(\cdot)$ in the natural log of the abscissas z_k provided, in this case, by the Gauss Legendre integration method. Thus,

$$\tilde{f}(s) = \sum_{k=1}^{M} w_k z_k^{s-1} f(-\ln(z_k)), \tag{3-29}$$

where w_k and z_k are the weights and abscissas, respectively, provided by the Gauss Legendre method. Note that w_k and z_k do not depend on the function $f(\cdot)$, but only on the number *M* of quadrature points and on the integration interval. See Press et al. (2002) for further details on how to obtain w_k and z_k by the Gauss Legendre method.

According to Press et al. (2002), the idea of Gaussian quadrature is to provide the freedom to choose not only the weighting coefficients, but also the location of the abscissas at which the function is to be evaluated: they are no longer equally spaced as occurs, for example, with trapezoidal rule and Simpson method.

Representing (3-29) in matrix form, it follows that:

$$\Psi \cdot \mathbf{\Theta} = \widetilde{\mathbf{\Theta}}, \qquad (3\text{-}30)$$

where Ψ is M^2-order matrix with $\psi_{vk} = w_k z_k^{v-1}$ and $v, k = 1, \ldots, M$; $\mathbf{\Theta}$ is M-order matrix of the state probabilities $\phi_j(-\ln z_k)$; $\widetilde{\mathbf{\Theta}}$ is M-order matrix of the LT of the state probabilities $\tilde{\phi}_j(s)$, with $s = 1, \ldots, M$ and j fixed. Given the transformed solution, eq. (3-30) is solved N times in order to obtain the state probabilities $\phi_j(-\ln z_k)$ for $j = 1 \ldots N$ by using any method of solution of linear algebraic equations such as Lower-Upper decomposition (see Press et al. (2002)).

Before solving eq. (3-30), (3-26) ($\tilde{\phi}_j(s) = \phi_j(0) \cdot \tilde{v}_j(s) + \tilde{h}_j(s) * \tilde{v}_j(s)$) is solved M runs in order to obtain the LT of the state probabilities $\tilde{\phi}_j(s)$, with $s = 1, \ldots, M$.

Theoretically, the bigger the number M of discretization points the greater results will be obtained. However, as we are using Gaussian Quadratures (rather than an equally spaced general quadrature methods, such as Newton-Cotes or Simpson formulas, as holds for the *2N*- and N^2- approaches) one has more freedom in choosing the coefficients and abscissas at which the function f will be evaluated thus, achieving integrations formulas of higher and higher accuracy with a smaller number of function evaluations than Newton-Cotes formula requires, for example (see Press et al. (2002) for more details).

Indeed, in accordance with a sensitivity analysis performed by Oliveira et al. (1997) one may obtain reasonable great accuracies with a number M of discretization points equals to *16*. Although these sensitivity tests have been taken on the Markov environment, one has considerable chances to reach the same findings here provided the semi-Markov processes are an extension of Markov models.

At this point, an important advantage of the *Lap*-numerical method comes up: whereas for the *2N*- and N^2-approaches the number of discretization points should be increased to obtain improved accuracies, the *Lap*-procedure with only *16* points will be able to provide valuable results with less computational cost than the *2N*- and N^2-methods, as it will be showed through the examples in chapter 4.

Before that, notice that the numerical method as just described is only able to compute the state probabilities $\phi_j(\cdot)$ at the points $t = -\ln z_k$. However, by using the result $\left(\int_0^\infty e^{-st} f(at)dt = \frac{1}{a} \tilde{f}\left(\frac{s}{a}\right) \right)$, the *Lap*-procedure may be used to obtain $\phi_j\left(-a \ln t_k\right)$, where $a > 0$ works as a scale factor and it will be defined as $a = {-T}/{\ln z_1}$, where z_1 is the minimum value of the abscissa provided by the Gauss Legendre method.

4. ASSESSING THE EFFECTIVENESS OF THE 2N- AND LAP-METHODS FOR SOLVING HSMP

The present chapter focuses on scrutinizing the main findings of chapter 3, comparing the effectiveness in terms of computational time and accuracy of *2N*- and *Lap*- approaches against the results obtained through the N^2- and MC methods by means of three examples in the context of availability assessment.

At first, *2N*- and N^2-approaches will be required for solving the homogeneous version of the availability assessment problem of optical monitoring systems described in section 1.2.2. Next, *2N*- and *Lap*-methods will solve the pumping oil unit problem given in section 1.2.1. Finally, the *Lap*-procedure will be also used in order to perform uncertainty analysis on the availability measure of the same pumping oil unit.

4.1. Optical Monitoring System Case: A Comparison Between *2N*- and N^2-approaches

In section 1.2.2, it was mentioned that the OMS dynamics should be treated by NHSMP due to the its peculiar characteristics. However, in this section let disregard the non-homogeneous idiosyncrasy of that system in order to compare performances of the *2N*- and N^2-approaches designed for HSMP in treating the problem of predicting the OMS dynamic behavior over time.

Given that the problem description is provided in 1.2.2, let's move on directly towards the required data for modeling. In fact, data needed to estimate the state probabilities and system unavailability for HSMP described through transition probabilities, such as OMS, are the transition probabilities p_{ij} and the CDF $G_{ij}(\cdot)$ as indicated in Figure 1-3. The probabilities p_{23} and p_{34} correspond to HEP that are assessed taking into consideration whether the "available time to complete the repair" is adequate or not (see section 1.2.2).

In Droguett et al. (2008), these probabilities and their causal relationships with performance shaping factors are handled through BBN (see Pearl (2000) and Korb and Nicholson (2003) for further details on BBN). Due to the scarcity of failure data on the OMS performance, the probabilities p_{23} and p_{34} (and their complements p_{21} and p_{31}, respectively) were obtained from engineering judgments (see Langseth and Portinale (2007) for further details on BBN in the reliability context).

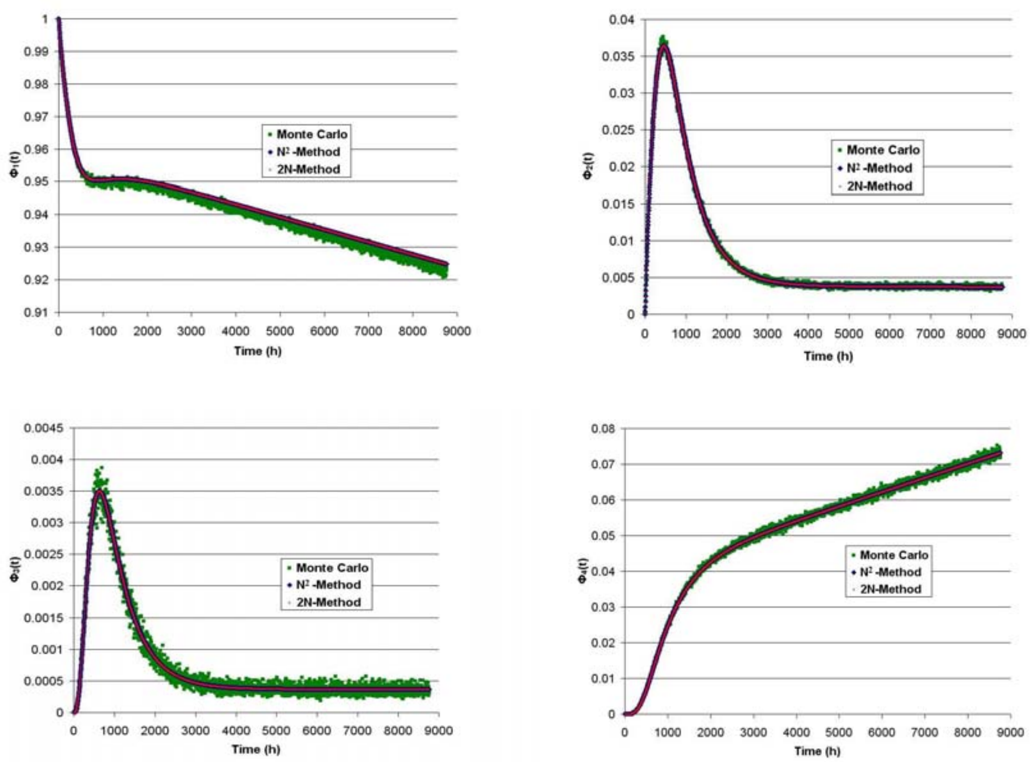
For the sake of simplicity, the integration between HSMP and BBN is not considered here. Therefore, in this example it is assumed that the operator HEP corresponds to 0.18 when under a situation of adequate available time to complete the repair, i.e., when the system reaches state 2. When the system reaches state 3, the time window to complete the repair becomes tighter, and it is assumed an HEP equals to 0.38 (these values are from Droguett et al. (2008) and may also be obtained through the CPT given in Appendix B).

The probability values p_{ij} as well as the CDF $G_{ij}(\cdot)$ of the HSMP are summarized in Table 4-1. Note that none of the parameters given in Table 4-1 depends on the process time what implies the non-homogeneity has not actually been considered.

Table 4-1 – Parameter estimates of the HSMP: p_{ij} and $G_{ij}(t)$ for the OMS example

Transition	**Transition probabilities of the EHMP**	**Conditional CDF**
$i \to j$	p_{ij}	$G_{ij}(t)$
$1 \to 1$	0.60	Exponential (1E-04h^{-1})
$1 \to 2$	0.40	Weibull (500.0h, 1.35)
$2 \to 1$	0.82	Exponential (0.05h^{-1})
$2 \to 3$	0.18	Weibull (300.0h, 1.75)
$3 \to 1$	0.62	Exponential (0.05h^{-1})
$3 \to 4$	0.38	Lognormal (4.0h, $(0.40)^2$h^2)

It is now possible to estimate the state probabilities (eq. (3-2)) and the system unavailability by solving the *2N*-numerical procedure described in chapter 3 (section 3.1), as shown in Figure 4-1 for a mission time of 1 year (8760 hours) and M = 2,500 steps for the *2N*-method, as well as according to the N^2-method described in Corradi et al. (2004) (with M = 2,500) and to the MC simulation (with M = 2,500 and 100,000 samples).



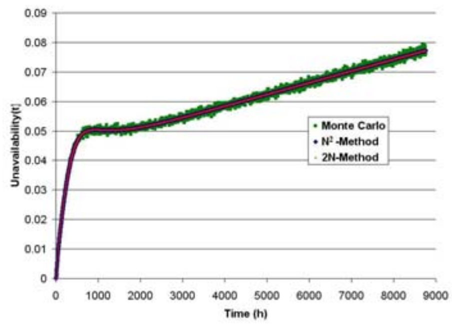


Figure 4-1 – State probabilities and unavailability curves for example 2 - *2N*-method x N^2 and Monte Carlo approaches: (a) $\phi_1(t)$; (b) $\phi_2(t)$; (c) $\phi_3(t)$; (d) $\phi_4(t)$; (e) unavailability

Note that the measures estimated by the *2N*-method match the computed values obtained by the application of MC simulation as well as by the N^2-solution. These results provide a validation of the accuracy of *2N*-numerical treatment then.

Furthermore, even though the simulation times depend on the computer settings, a considerable difference in terms of computational effort is verified in this example. Indeed, in an Intel® Core Dual Core Processor 32-bit Operating System, 2.00 GHz, 250.0 Gb and 2.00 GB of RAM[1], the *2N*-method required 27.56 seconds, whereas the MC and the N^2 approaches spent 1,163.25 seconds and 40,551.05 seconds, respectively.

The difference in CPU time required by the N^2 and MC methods and the *2N*-approach underlines the efficiency of the latter in quickly achieving meaningful results with accuracy comparable to the N^2-method.

4.2. Pumping Oil Unit Case: A Comparison Between *2N-* and *Lap*-approaches

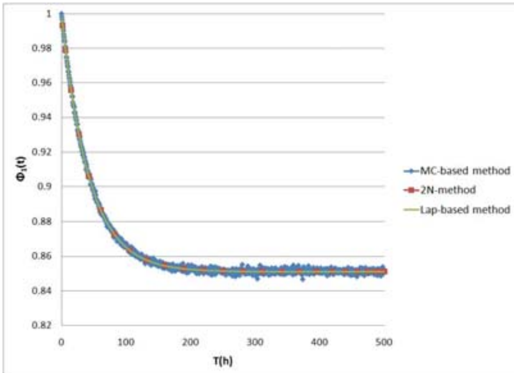
Since the *2N*-method has obtained noteworthy results, this section aims comparing it against the *Lap*-procedure through the pumping oil unit case, which was described in section 1.2.1 and may be modeled as an HSMP described by transition rates.

Then, assume that system is initially in state 1, the failure (λ) and repair (μ) rates are constant and equal to 3.5e-3 failure/h and 0.020 repair/h, respectively, and α = 100h and β = 2.08. Conversely to section 1.2.1, notice that the uncertainty on MTTF value has not been considered here. The focus on uncertainty analysis will be switched back in next section.

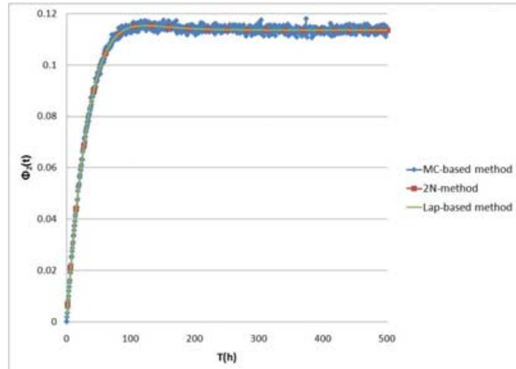
By solving the state probabilities for this HSMP described by transition rates given in eq. (3-2), Figure 4-2 shows the outcomes provided by the *Lap*-procedure, *2N*- and Monte Carlo approaches for a mission time of T = 500.0h. For sake of illustration, Figure 4-2 does not

[1] The same computer setting will be used for the next examples.

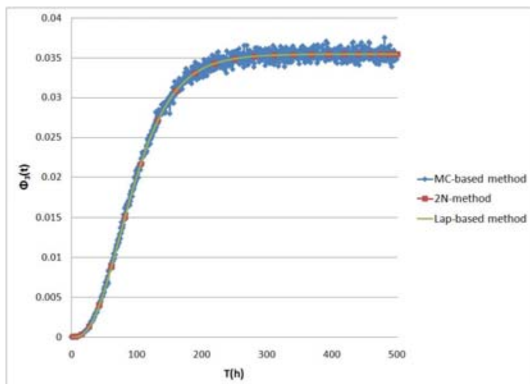
depict the results attained by N^2- since previous section already showed it and the *2N*-approach show close agreement.



(a)



(b)



(c)

Figure 4-2 – State probabilities – *Lap*-method x *2N*- and MC approaches: (a) $\phi_1(t)$; (b) $\phi_2(t)$; (c) $\phi_3(t)$

In this example of application, the MC algorithm ran with 100,000 iterations and M = 500 steps. For the *2N*- and N^2- methods, M = 500 steps were used. Note that in Figure 4-2, the state probabilities computed by the *Lap*- and *2N*- numerical procedures and by simulation show close agreement, providing a validation of the *Lap*-technique in relation to its numerical approximation.

Furthermore, the *Lap*-numerical technique has computed the state probabilities for this HSMP described by transition rates considerably faster than the *2N*- and N^2- and MC approaches. Indeed, the *Lap*-method spent less than one second (0.06 seconds), while the *2N*-, N^2- and MC approaches took 1.97, 1,317.12 and 64.10 seconds, respectively.

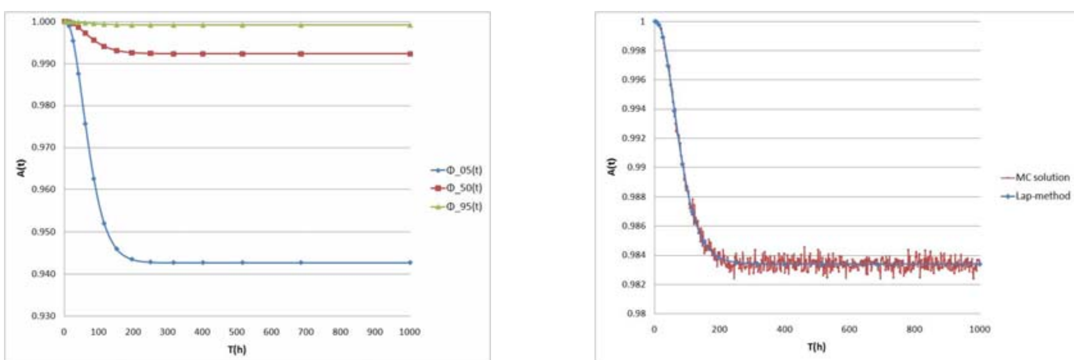
4.3. Pumping Oil Unit: Availability Uncertainty Analysis Through *Lap*-method

In section 1.2.1, the pumping oil unit problem had been described considering the uncertainty on MTTF value which was influenced by external factors (see Figure 1-1). Basically, it is required that an HSMP-based numerical method runs several times in order to catch the impact of MTTF uncertainty on the (un)availability measure. Since the *Lap*-procedure has showed meaningful outcomes in terms of computational time and accuracy as well, it will be used to perform the availability uncertainty analysis in this section. Furthermore, it will be showed how to update probabilistic beliefs on availability curve as new evidence becomes available.

4.3.1. Availability Measure Estimation

Considering that the system starts its operation in state 1 (available), the availability is assessed for a mission time equals to T=1,000.0h. Considering also the relation $\lambda = 1/\text{MTTF}$, λ: failure rate, the *Lap*-algorithm described in section 3.2 is replicated for 100,000 iterations in order to explicitly quantify the uncertainty on the availability measure given uncertainty in the MTTF characterized in terms of the BBN in Figure 1-1.

Indeed, Figure 4-3(a) shows the 5th, 50th, and 95th percentile curves computed for availability by the *Lap*-numerical procedure. Each percentile corresponds to the probability that the availability measure value is smaller than the computed one at a specific point in the mission time, thus explicitly quantifying the impact of the uncertainty about the parameter MTTF on availability measure. Figure 4-3(b) in turn illustrates the availability curves computed by both *Lap*-numerical method and MC simulation considering the prior[1] mean value for $\overline{MTTF}_{prior} = 5246.98$ h.



(a) (b)

Figure 4-3 – Availability Measure Curve: (a) uncertainty on availability measure; (b) Proposed numerical procedure *x* Monte Carlo simulation

In this example, the MC algorithm for HSMP described in Moura and Droguett (2007) (which follows in attachment F) has run with M = 1,000 steps and 100,000 iterations for each step. For the *Lap*-method, M = 16 steps was used as discussed in section 3.2.

[1] The term prior means CPT in Appendix A is directly used to estimate MTTF distribution.

Note that in Figure 4-3(b), the availability measure computed by both the *Lap*- and MC procedures show close agreement, again providing a validation of the *Lap*-technique in relation to its numerical approximation.

Moreover, the *Lap*-method spent roughly 0.02 seconds per replication, while the MC took 47.81 seconds to compute the results showed in Figure 4-3(b).

4.3.2. Updating Probabilistic Beliefs

The hybrid (HSMP-BBN) model allows for uncertainty updating regarding the availability measure as new evidence about any of the external factors influencing MTTF (Figure 1-1) becomes available at any point in the mission time. In fact, suppose it is known that the level of paraffin is inadequate for the oil to be handled by the pumping unit. This new evidence does not imply any changes in the BBN topology (Figure 1-1) or for the state diagram of the HSMP (Figure 1-2). However, the CPT of the BBN, which are given in Appendix A, are modified as now it is known that P(Inadequate level of paraffin) = 1.

This new evidence impacts the future behavior of the system and consequently its availability measure. The uncertainty on system availability metric given the new evidence is characterized in terms of a posterior distribution whereas a prior distribution characterizes the uncertainty about availability metric before the evidence has become available as was done in previous subsection.

Updating the uncertainty distribution for the MTTF according to Firmino (2004), Figure 4-4 shows the comparison between marginal prior and posterior probability distributions for the MTTF. The results show that because of the inadequate paraffin level there is a shift of probability mass towards lower values of MTTF.

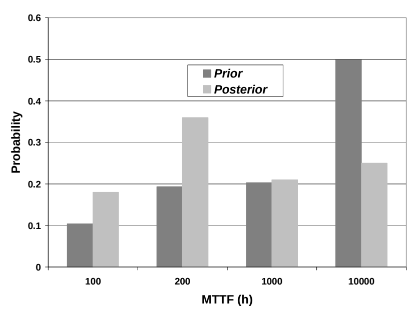
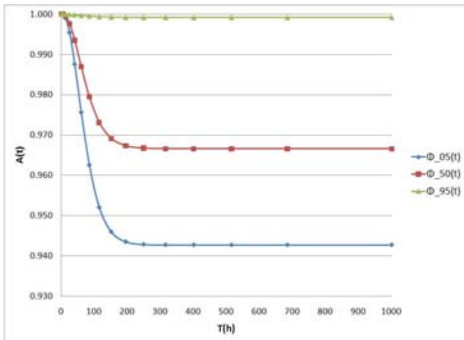


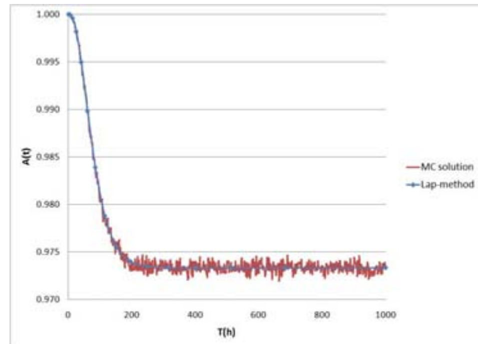
Figure 4-4 – Prior and posterior marginal probability distributions of MTTF

Considering the posterior mean value for MTTF, $\overline{MTTF}_{posterior} = 2800.0\,\text{h}$, Figure 4-5 illustrates the impact on the availability uncertainty given the updated MTTF probability distribution. More specifically, the marginal probability distribution of MTTF, called

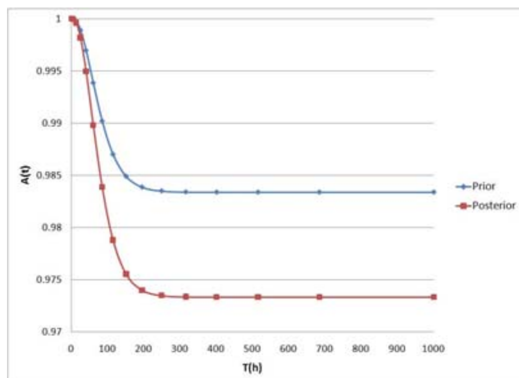
$MTTF_{prior}$ in previous subsection, is updated given the evidence regarding the level of paraffin which in turn affects the availability metric.



(a)



(b)



(c)

Figure 4-5 – Prior and posterior availability measure: (a) *Lap*-method: uncertainty on posterior availability measure; (b) *Lap*-numerical procedure *x* Monte Carlo simulation; (c) *Lap*-method: Prior *x* posterior availability measure

This analysis was solely based on evidence about the variable “level of paraffin”. Nevertheless, similar studies can be performed provided that evidence becomes available for other variables, for example, “percentage of H_2O and solids - BWSOT”. Moreover, uncertainty on other reliability measures such as reliability and maintainability could be assessed from the proposed hybrid model (see Moura (2006) and Moura and Droguett (2008)).

4.4. Comments

Continuous-time homogeneous semi-Markov processes are important probabilistic tools to model reliability measures for systems whose future behavior is dependent on the current and next states of the process as well as on sojourn times, as for fault tolerant systems where the failure of a component leads to a system failure only when repair time has exceeded some tolerable downtime, i.e., component failure does not immediately cause a system failure.

HSMP were usually computed via the N^2-method described in Corradi et al. (2004) where the future behavior of the system is assessed via interval transition probability equations comprised of a set of N^2 coupled integral equations with one time variable *t*. However, this approach is rather time-consuming with a computational cost greater than the MC simulation.

This situation motivated the search for more efficient numerical treatments of HSMP with less computational effort and with a comparable accuracy in relation to the available methods in the related literature (MC simulation and the N^2-approach).

In fact, the *2N*-mathematical formulation and numerical treatment consists of casting the N^2 coupled convolution integral equations into an initial value problem involving transition frequency densities, and then solve *N* coupled integral equations with one variable and *N* straightforward integrations.

Moreover, this thesis has proposed the alternative *Lap*-method which makes use of LT. Although using Laplace apparatus on HSMP field is not novel, the *Lap*-procedure showed some noteworthy advantages: (*i*) it used a pre-set number of steps, which is independent on the problem to be solved. Thus, it is not required anymore adjusting (through either trial-error tests or dynamically) the number *M* of steps in order to attain the desired convergence. (*ii*) thus, it reduced considerably the computational effort in relation to the abovementioned *2N*- and N^2-methods and MC as well. The *Lap*-approach was also validated by comparison with the results from both the *2N*- and N^2-solutions and MC method.

One drawback that deserves attention is since *Lap*-method is backed up on Gaussian Quadratures theory there is not a quite simple way to obtain an estimate of the absolute error made by the approach (see Press et al. (2002) for more details). Moreover, unfortunately, the application of the *Lap*-procedure within the non-homogeneous environment has not catched notable outcomes, as showed in Moura and Droguett (2007).

5. 2N-METHOD FOR NON-HOMOGENEOUS SEMI-MARKOV PROCESSES

Non-homogeneity implies higher difficulties on the continuous-time semi-Markov processes (NHSMP) environment. This gives rise as more intricate mathematical methods and related numerical solutions and is one of the main reasons behind the scarcity of NHSMP applications.

Indeed, the N^2-method for solving NHSMP is rather burdensome (as it has already been seen for HSMP case in chapter 3), consisting of directly applying a general quadrature method to N^2 coupled integral equations with two variables, where N is the number of states. Therefore, the next two chapters focus on developing a new and faster numerical treatment, which is also called *2N*-method, for NHSMP and scrutinizing its effectiveness comparing against the results provided by the N^2- method and MC. Rather than computing N^2 integral equations, this approach consists of solving only N coupled integral equations with one variable and N straightforward integrations so that the high and inherent computational cost that plagues the solution of NHSMP is likely to be reduced.

The main findings within this chapter can be found in Moura and Droguett (2009c), which follows in attachment G. As the homogenous counterpart may be considered a special case of NHSMP, the conclusions made over the present chapter may be specialized for the developments showed in chapter 3 (section 3.1). In spite of that, it was quite meaningful developing a particular *2N*-mathematical and numerical formulations for HSMP since in this way its peculiar characteristics have been taken into account.

5.1. An initial value problem for NHSMP involving transition frequency densities

Depending on how an NHSMP is described, the kernel $C_{ij}(\cdot)$ and the CDF $F_i(\cdot)$ are given by equations (2-8) and (2-9) in case of transition probabilities or by equations (2-13) and (2-14) in case of transition rates. The *2N*-numerical approach will be developed in a general way in order to handle both situations.

By using a similar nomenclature to the one in Becker et al. (2000), let recall $\mathrm{N}(t)$ to be the number of times that state j of an NHSMP is visited from any state in the interval $[0,t]$. Let also $\mathrm{H_j}(t) = \mathrm{E}[\mathrm{N}(t)]$ be its expected value. If $\mathrm{H_j}(t)$ is continuously differentiable, then $\mathrm{h_j}(t)\mathrm{dt} = \mathrm{dH_j}(t)$ is its corresponding density function.

As the stochastic process under consideration is regular, i.e., no more than one transition can occur in any interval ($t, t+dt$), then $h_j(t)$ can be assumed as the probability that a transition occurs to state j in an infinitesimal time interval dt as follows:

$h_j(t)dt = Pr\{to\ reach\ state\ j\ in\ (t,\ t+dt)\}$.

Thus, it follows that:

$$h_j(t) = \sum_{i=1}^{N} \phi_i(0) \cdot \dot{C}_{ij}(0,t) + \sum_{i=1}^{N} \int_0^t h_i(\tau) \cdot \dot{C}_{ij}(\tau,t)\, d\tau. \qquad (5\text{-}1)$$

According to the description of eq. (2-8), (5-1) means state j can be reached either if the process was initially in state i and remains there until time t, when a transition to state j occurs; or if the process reached state i at time τ, remaining there for $x = t - \tau$, then a transition to state j occurs. The summation over the state number N is for all possible intermediary states, where NHSMP can transit. The integral term in turn means that the transition to state i may occur at any time $\tau \in [0,t]$.

Eq. (5-1) corresponds to a system of N integral equations with unknowns $h_j(t)$, $j = 1,\ldots,N$. The probabilities $\phi_j(t)=Pr[Z_t=j \mid Z_0]$ can be obtained from the initial conditions $\phi_j(0)$ as follows:

$$\phi_j(t) = \phi_j(0) \cdot [1 - F_j(0,t)] + \int_0^t h_j(\tau) \cdot [1 - F_j(\tau,t)] d\tau\ . \qquad (5\text{-}2)$$

Eq. (5-2) says that a process can be in state j at time t either if it was initially in the state j and remained there at least up to time t; or if it visited state j at any time $\tau \in [0,t]$ with probability $h_j(t)$ and stayed there for ($x = t - \tau$). Eq. (5-2) corresponds to N straightforward integrations that can be solved independently by using the solution of equation (5-1).

Eqs. (5-1) and (5-2) extend the formulation presented in Becker et al. (2000) in order to address NHSMP described in terms of both transition probabilities and transition rates. Moreover, the computational effort to solve eqs. (5-1) and (5-2) is less intricate than in the case of the N^2-method described in Janssen and Manca (2001). In fact, the *2N*-method consists of solving N coupled integral equations with one variable (eq. (5-1)) and N straightforward integrations (eq. (5-2)) whereas the N^2-method consists of solving N^2 coupled integral equations (eq. (2-10)). A comparison among the *2N*-approach, the N^2-method and the MC-based solution will be discussed in chapter 6.

Before that in next section, a general quadrature based method for simultaneously solving the eqs. (5-1) and (5-2) will be presented. Moreover, the convergence conditions and error analysis are also developed and demonstrated.

5.2. Numerical formulation

5.2.1. Description of the numerical solution

A numerical integration or quadrature method can be written as follows (see Press et al. (2002)):

$$\int_0^{kh} \varphi(\tau,t)d\tau \cong \sum_{\tau=0}^{k} w_{\tau,k} \cdot \varphi(\tau h, kh)\,, \tag{5-3}$$

where h is the step measure, and $w_{\tau,k}$ are the weights related to the quadrature formula (5-3). Note that such weights also do not depend on the integrand function φ(· , ·); they are function of the start point (0), of the end point (kh) and of the intermediary point (τh) at which the function value is computed. Moreover, one has M such that $Mh = T$, with $0 \le kh \le T$, $k \le M, k, M \in \mathbb{N}$ where M is the number of steps and T is the mission time.

Using eq. (5-3), a solution for (5-1) is given as:

$$\tilde{h}_j(kh) = a_j(0h,kh) + \sum_{i=1}^{N}\left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot \tilde{h}_i(\tau h) \cdot \dot{C}_{ij}(\tau h, kh)\right), \tag{5-4}$$

where

$$a_j(0h,kh) = \sum_{i=1}^{N} \phi_i(0) \cdot \dot{C}_{ij}(0h,kh)\,,$$

where the notation $\sim$ again means an approximation.

The system of eqs. (5-4) can be written as follows:

$$\tilde{h}_j(kh) - \sum_{i=1}^{N}\left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot \tilde{h}_i(\tau h) \cdot \dot{C}_{ij}(\tau h, kh)\right) = a_j(0h,kh) \tag{5-5}$$

or in matrix form:

$$\widetilde{H}^T_{(kh)} - \sum_{\tau=0}^{k} \Psi^T_{(\tau h,kh)} \cdot \widetilde{H}^T_{(\tau h)} = A^T_{(0h,kh)}\,, \tag{5-6}$$

where the symbol ^{T} represents the transpose matrix, $\widetilde{H}^T_{(\cdot)}$ and $A^T_{(\cdot,\cdot)}$ are N-order matrices, and $\Psi^T_{(\cdot,\cdot)}$ is N^2-order matrix, where:

$$\psi_{ij}(\cdot,\cdot) = w_{\tau,k}\, \dot{C}_{ij}(\cdot,\cdot)\,. \tag{5-7}$$

Alternatively, eq. (5-6) can be written as follows:

$$\mathbf{U}^T \cdot \widetilde{\mathbf{H}}^T = \mathbf{A}^T\,, \tag{5-8}$$

where

$$\boldsymbol{U^T} = \begin{bmatrix} I-\Psi^T_{(0h,0h)} & 0 & \cdots & 0 & \cdots & 0 \\ -\Psi^T_{(0h,1h)} & I-\Psi^T_{(1h,1h)} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\Psi^T_{(0h,\tau h)} & -\Psi^T_{(1h,\tau h)} & \cdots & I-\Psi^T_{(\tau h,\tau h)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\Psi^T_{(0h,Mh)} & -\Psi^T_{(1h,Mh)} & \cdots & -\Psi^T_{(\tau h,Mh)} & \cdots & I-\Psi^T_{(Mh,Mh)} \end{bmatrix},$$

$$\widetilde{\boldsymbol{H}}^T = \begin{bmatrix} \widetilde{H}^T_{(0h)} \\ \widetilde{H}^T_{(1h)} \\ \ddots \\ \widetilde{H}^T_{(\tau h)} \\ \ddots \\ \widetilde{H}^T_{(Mh)} \end{bmatrix} \quad \text{and} \quad \boldsymbol{A^T} = \begin{bmatrix} A^T_{(0h,0h)} \\ A^T_{(0h,1h)} \\ \ddots \\ A^T_{(0h,\tau h)} \\ \ddots \\ A^T_{(0h,Mh)} \end{bmatrix}. \tag{5-9}$$

A bold face notation is used for a matrix of matrices as in $\mathbf{U}^T, \widetilde{\mathbf{H}}^T$, and $\mathbf{A}^T$. Eq. (5-9) is used to compute the N-order matrices $\widetilde{H}^T_{(\tau h)}$, whose elements $\tilde{h}_j(\tau h)$ are the density functions of the number of times that state $j = 1,\ldots,N$ of an NHSMP is visited from any state in the interval $[0,\tau h]$.

Basically, eq. (5-9) represents the main difference between the *2N*- and N^2-approaches. While *2N*-method needs to solve eq. (5-9) (set of N linear algebraic equations) M times as a requirement to catch the system's dynamic behavior over time, N^2-method computes the interval transition probabilities by using a somewhat modified formula which possesses a main difference in relation to that equation: the solution of the N^2-counterpart for eq. (5-9) is composed of an M^2-order matrix of matrices which in turn are of order N^2. Thus, through N^2-method the number of each set of N^2 linear equations to be solved is equal to $(M^2+M)/2$. In this way, *2N*-method tends to reduce considerably the computational time in relation to N^2-method.

Having estimated the solutions matrix $\widetilde{\mathbf{H}}^T$ (eq. (5-9)), the next step consists in computing the N state probabilities $\tilde{\phi}_j(t)$. In fact, a numerical solution for the state probabilities $\tilde{\phi}_j(t)$ can be computed as follows:

$$\tilde{\phi}_j(kh) = \phi_j(0h)\cdot[1-F_j(0h,kh)] + \sum_{\tau=0}^{k} w_{\tau,k}\cdot\tilde{h}_j(\tau h)\cdot[1-F_j(\tau h,kh)], \tag{5-10}$$

where $\tilde{h}_j(\tau h)$ ($j = 1,\dots,N$ and $\tau = 0,\dots,M$) are step-solutions obtained from eq. (5-9). In this way, $\widetilde{\mathbf{\Phi}} = \left[\widetilde{\Phi}_{(0h)},\dots,\widetilde{\Phi}_{(\tau h)},\dots,\widetilde{\Phi}_{(Mh)}\right]$ is comprised of M matrices $\widetilde{\Phi}_{(\tau h)}$ each one of order N.

Note that, rather than solving the N^2 coupled integral equations with two variables through the direct application of any general quadrature method to eq. (2-10), the approach just described estimates the state probabilities of an NHSMP by solving N coupled integral equations with one variable (eq. (5-9)) and N straightforward integrations (eq. (5-10)).

For example, by using the extended Simpson's rule given in Press et al. (2002) (p. 138), eqs. (5-5) and (5-10) can be written in the following way:

$$\begin{aligned}\tilde{h}_j(kh) &= a_j(0h,kh) + \frac{h}{3}\sum_{i=1}^{N}\tilde{h}_i(0h)\cdot\dot{C}_{ij}(0h,kh)\\&+\frac{4h}{3}\sum_{i=1}^{N}\sum_{\tau=1}^{\frac{k}{2}}\tilde{h}_i((2\tau-1)h)\cdot\dot{C}_{ij}((2\tau-1)h,kh)\\&+\frac{2h}{3}\sum_{i=1}^{N}\sum_{\tau=1}^{\frac{k}{2}-1}\tilde{h}_i(2\tau h)\cdot\dot{C}_{ij}(2\tau h,kh)\\&+\frac{h}{3}\sum_{i=1}^{N}\tilde{h}_i(kh)\cdot\dot{C}_{ij}(kh,kh)\end{aligned}$$

and

$$\begin{aligned}\tilde{\phi}_j(kh) &= \phi_j(0)\cdot[1-F_j(0h,kh)] + \frac{h}{3}\cdot\tilde{h}_j(0h)\cdot[1-F_j(0h,kh)]\\&+\frac{4h}{3}\cdot\sum_{\tau=1}^{\frac{k}{2}}\tilde{h}_j((2\tau-1)h)\cdot[1-F_j((2\tau-1)h,kh)]\\&+\frac{2h}{3}\cdot\sum_{\tau=1}^{\frac{k}{2}-1}\tilde{h}_j(2\tau h)\cdot[1-F_j(2\tau h,kh)]\\&+\frac{h}{3}\cdot\tilde{h}_j(kh)\cdot[1-F_j(kh,kh)].\end{aligned} \tag{5-11}$$

5.2.2. Solution conditions and upper limit estimate of the discretization error

From eq. (5-4) follows that:

$$\begin{aligned}&\tilde{h}_j(kh) - \sum_{i=1}^{N} w_{k,k}\cdot\tilde{h}_i(kh)\cdot\dot{C}_{ij}(kh,kh)\\&= a_j(0h,kh) + \sum_{i=1}^{N}\left(\sum_{\tau=0}^{k-1} w_{\tau,k}\cdot\tilde{h}_i(\tau h)\cdot\dot{C}_{ij}(\tau h,kh)\right).\end{aligned} \tag{5-12}$$

By writing eq. (5-12) in terms of $j = 1,\dots,N$, it follows that:

$$\tilde{h}_1(kh) - \sum_{i=1}^{N} w_{k,k} \cdot \tilde{h}_i(kh) \cdot \dot{C}_{i1}(kh,kh) = d_1$$

$$\tilde{h}_2(kh) - \sum_{i=1}^{N} w_{k,k} \cdot \tilde{h}_i(kh) \cdot \dot{C}_{i2}(kh,kh) = d_2$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$

$$\tilde{h}_N(kh) - \sum_{i=1}^{N} w_{k,k} \cdot \tilde{h}_i(kh) \cdot \dot{C}_{iN}(kh,kh) = d_N , \tag{5-13}$$

where d_j, $j = 1,...,N$ is the right-hand side term of eq. (5-12).

According to Press et al. (2002), it is guaranteed that the set of linear algebraic equations given in eq. (5-13) has solution if the matrix of known coefficients

$$\begin{bmatrix} 1-\psi_{11}(kh,kh) & -\psi_{12}(kh,kh) & \cdots & -\psi_{1j}(kh,kh) & \cdots & -\psi_{1N}(kh,kh) \\ -\psi_{21}(kh,kh) & 1-\psi_{22}(kh,kh) & \cdots & -\psi_{2j}(kh,kh) & \cdots & -\psi_{2N}(kh,kh) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\psi_{i1}(kh,kh) & -\psi_{i2}(kh,kh) & \cdots & 1-\psi_{ij}(kh,kh) & \cdots & -\psi_{iN}(kh,kh) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\psi_{N1}(kh,kh) & -\psi_{N2}(kh,kh) & \cdots & -\psi_{Nj}(kh,kh) & \cdots & 1-\psi_{NN}(kh,kh) \end{bmatrix},$$

is not degenerate, i.e., is not singular, where $\psi_{ij}(\cdot,\cdot)$ is given in eq. (5-7). The non-singularity happens if none of the N eqs. in (5-13) is a linear combination of the others. If this condition is satisfied, eq. (5-10) of unknowns $\tilde{\phi}_j$ also admits solution given the solution of eq. (5-4). Janssen and Manca (2001) reached an equivalent solution condition for the case of the N^2-method, i.e., by directly applying a general quadrature method to the N^2 coupled integral equations.

In the subsequent development, we again make use of the Lemmas 1 and 2 and Remark 1 from section 3.1, and described in Janssen and Manca (2001), that also hold for the *2N*-mathematical formulation proposed here for NHSMP:

Theorem 1: Let

$h_j(l) : [l, Y] \to \mathbb{R} \qquad l \in [0, T]$.

Let $q \in \{0,...,M-1\}$, $M \in \mathbb{N}$, such that $l \le qh \le T$.

Let

$$\xi_j^k(h) = \tilde{h}_j(kh) - h_j(kh) \qquad k = 0,1,...,M , \tag{5-14}$$

where $h_j(kh)$ is solution of eq. (5-1) and $\tilde{h}_j(kh)$ is solution of (5-4).

Define also:

$$\eta^k(h) = \sum_{i=1}^{N} \left| \xi_i^k(h) \right| .$$

Moreover, let

$$w^l = w_M^l = \max_{0 \le \tau \le k \le M} \frac{\left| w_{\tau,k} \right|}{h} < \infty \tag{5-15}$$

$$t_j^k(h) = \sum_{i=1}^{N} \int_0^{kh} h_i(\vartheta) \cdot \dot{C}_{ij}(\vartheta, kh)\, d\vartheta - \sum_{i=1}^{N} \left[\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}(\tau h, kh) \right] \tag{5-16}$$

$$\sigma^k(h) = \sum_{i=1}^{N} \left| t_i^k(h) \right| ,$$

$$\rho^l(h) = \rho_M^l(h) = \max_{q \le k \le M} \sigma^k(h) \tag{5-17}$$

$$\xi^l(h) = \sum_{\tau=0}^{q-1} \eta^\tau(h) .$$

Suppose that $\left| \dot{C}_{ij}(l,t) \right| \le d_1$ for $t \in [l, Y]$ and $x \in [0, t-l]$.

Then

$$\eta^k(h) \le \left(\frac{\rho^l(h) + Nhw_M^l d_1 \xi^l(h)}{1 - Nhw_M^l d_1} \right) \exp\left(\frac{Nd_1 w_M^l kh}{1 - Nd_1 w_M^l h} \right), \tag{5-18}$$

$k = q, q+1, ..., M$

given that $Nhw_M^l d_1 < 1$.

Proof: From eq. (5-1) it follows that for $k \ge q$:

$$\begin{aligned} h_j(kh) - \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}(\tau h, kh) \right) = a_j(0h, kh) \\ + \sum_{i=1}^{N} \int_0^{kh} h_i(\vartheta) \cdot \dot{C}_{ij}(\vartheta, kh)\, d\vartheta - \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}(\tau h, kh) \right) \end{aligned} . \tag{5-19}$$

From eqs. (5-4) and (5-19), it follows that:

$$\begin{aligned} \tilde{h}_j(kh) - h_j(kh) + \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}(\tau h, kh) \right) \\ = \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot \tilde{h}_i(\tau h) \cdot \dot{C}_{ij}(\tau h, kh) \right) - \sum_{i=1}^{N} \int_0^{kh} h_i(\vartheta) \cdot \dot{C}_{ij}(\vartheta, kh)\, d\vartheta \\ + \sum_{i=1}^{N} \left(\sum_{\tau=0}^{k} w_{\tau,k} \cdot h_i(\tau h) \cdot \dot{C}_{ij}(\tau h, kh) \right) \end{aligned}$$

and from eqs. (5-14) and (5-16):

$$\xi_j^k(h)-\sum_{i=1}^{N}\left[\sum_{\tau=0}^{k}w_{\tau,k}\cdot\tilde{h}_i(\tau h)\cdot\dot{C}_{ij}(\tau h,kh)-\sum_{\tau=0}^{k}w_{\tau,k}\cdot h_i(\tau h)\cdot\dot{C}_{ij}(\tau h,kh)\right]$$
$$=-t_j^k(h).$$

Hence,

$$\xi_j^k(h)-\sum_{i=1}^{N}\left\{\sum_{\tau=0}^{k}w_{\tau,k}\cdot\dot{C}_{ij}(\tau h,kh)\cdot\left[\tilde{h}_i(\tau h)-h_i(\tau h)\right]\right\}=-t_j^k(h)$$

and

$$\xi_j^k(h)-\sum_{i=1}^{N}\left[\sum_{\tau=0}^{k}w_{\tau,k}\cdot\dot{C}_{ij}(\tau h,kh)\cdot\xi_i^\tau(h)\right]=-t_j^k(h)\ .$$

It follows from eq. (5-15) that:

$$\begin{aligned}\left|\xi_j^k(h)\right| &\le \left|t_j^k(h)\right|+\sum_{i=1}^{N}\sum_{\tau=0}^{k}\left|w_{\tau,k}\right|\cdot\left|\dot{C}_{ij}(\tau h,kh)\right|\cdot\left|\xi_i^\tau(h)\right| \\ &\le \left|t_j^k(h)\right|+d_1w_Mh\sum_{i=1}^{N}\sum_{\tau=0}^{k}\left|\xi_i^\tau(h)\right|.\end{aligned}$$

Therefore,

$$\left|\xi_j^k(h)\right|\le\left|t_j^k(h)\right|+d_1w_Mh\sum_{\tau=0}^{k}\sum_{i=1}^{N}\left|\xi_i^\tau(h)\right|.$$

Performing the summation in relation to the index *j = 1,…,N*, one obtains that:

$$\sum_{j=1}^{N}\left|\xi_j^k(h)\right|\le\sum_{j=1}^{N}\left|t_j^k(h)\right|+Nd_1w_Mh\sum_{\tau=0}^{k}\sum_{i=1}^{N}\left|\xi_i^\tau(h)\right|. \quad (5\text{-}20)$$

Eq. (3-21) implies that:

$$\eta^k(h)\le\sigma^k(h)+Nd_1w_Mh\sum_{\tau=0}^{k}\eta^\tau(h)$$

and

$$(1-Nd_1w_Mh)\eta^k(h)\le\sigma^k(h)+Nd_1w_Mh\sum_{\tau=0}^{k-1}\eta^\tau(h)\ .$$

From eq. (5-17), it follows that:

$$(1-Nd_1w_Mh)\eta^k(h)\le\rho_M(h)+Nd_1w_Mh\sum_{\tau=0}^{k-1}\eta^\tau(h).$$

Now, from Lemmas 1 and 2, it follows that:

$$A = \frac{Nd_1 w_M h}{1 - Nd_1 w_M h},$$

$$B = \frac{\rho_M(h)}{1 - Nd_1 w_M h}$$

From Remark 1, it follows that:

$$\overline{L} = \frac{A}{h} = \frac{Nd_1 w_M}{1 - Nd_1 w_M h}.$$

Therefore, one has that:

$$\eta^k(h) \le (A\xi(h) + B)\exp\left(\frac{Nd_1 w_M kh}{1 - Nd_1 w_M h}\right)$$

what implies that:

$$\eta^k(h) \le \left(\frac{\rho_M(h) + Nhw_M d_1 \xi(h)}{1 - Nhw_M d_1}\right)\exp\left(\frac{Nd_1 w_M kh}{1 - Nd_1 w_M h}\right). \qquad \bullet$$

Eq. (5-18) computes the upper limit of the discretization error estimate whenever $\tilde{h}_j(kh)$ is used to approximate $h_j(kh)$. The same conclusion was reached in Janssen and Manca (2001) for the N^2-method. In other words, the upper limit of the numerical error computed by estimating the interval transition probabilities by using the N^2-method is the same given by eq. (5-18).

Given the solution of eq. (5-4), the error estimate $\varsigma_j^k(h) = \tilde{\phi}_j(kh) - \phi_j(kh)$ in computing $\phi_j(t)$ through eq. (5-10) in turn depends only on the choice of the quadrature method since each unknown may be solved independtly. In eq. (5-11), for example, the error term is equal to $\varsigma_j^k(h) = O_j\left(1/k^4\right)$. This means the true answer $\phi_j(kh)$ differs from the estimate $\tilde{\phi}_j(kh)$ by an amount equals to $1/k^4$ (see Press et al. (2002) for further details).

Therefore, the upper limit of the total error estimate of the *2N*-method is given by the sum of the discretization errors estimated from eqs. (5-1) and (5-2). In spite of that, as eq. (5-18) is an inequality it is even possible that the *2N*-method catches smaller errors than N^2-method.

In this way, it is expected that the *2N*-method be less time-consuming since it involves solving *N* coupled integral equations and *N* straightforward integrations. Moreover, generally speaking it has accuracy roughly equals to the N^2-method. However, it is also possible that the *2N*-method obtains minor errors than N^2-method. Research on these findings will be done in next chapter.

6. ASSESSING THE EFFECTIVENESS OF THE 2N-METHOD FOR SOLVING NHSMP

Two examples in the context of reliability are presented in this chapter. The first one addresses a case where a semi-analytical solution is available. Then it is discussed an example of application concerning the pressure-temperature optical monitoring systems for oil wells previously described in section 1.2.2. In both cases, the *2N*-approach is validated via the comparison against the results obtained from the semi-analytical solution (for the first example) as well as from both the N^2- and the MC methods.

6.1. A semi-analytical example

This section handles a simple three-state reliability semi-Markov example for a system comprised of a single component as illustrated in Figure 6-1. This system starts its operation in state 1. From there, it moves to state 2 if it reaches a determined non-critical degradation stage. If degradation level of the component attains a critical threshold, the system fails (state 3).

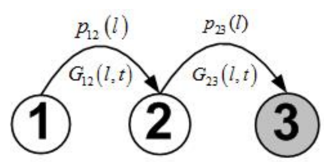


Figure 6-1 – HSMP for the semi-analytical example

This semi-Markov process is homogeneous in relation to the process time and, for this situation, a semi-analytical solution is possible as follows:

$$\phi_1(t) = 1 - F_1(0,t)\,,$$

$$\phi_2(t) = \int_0^t \dot{C}_{12}(0,\tau)\cdot[1 - F_2(\tau,t)]d\tau\,,$$

$$\phi_3(t) = 1 - [\phi_1(t) + \phi_2(t)]\,.$$

The solution is considered semi-analytical as the state probability $\phi_1(t)$ can be analytically estimated, while the state probability $\phi_2(t)$ is computed via numerical integration.

The required data to estimate the system reliability is given in Table 6-1. For the sake of validation, the results provided by the semi-analytical solution are compared against the ones from the *2N*-method, the N^2-approach, and the MC simulation drawn for NHSMP.

Table 6-1 – Parameters of the NHSMP: $p_{ij}(t)$ and $G_{ij}(l,t)$ for the semi-analytical example.

Transition	Transition probabilities of the ENHMP	Conditional CDF

$i \to j$	p_{ij}	$G_{ij}(l,t)$
$1 \to 2$	1.0	Exponential (1e-3h^{-1})
$2 \to 3$	1.0	Weibull (250.0h, 1.5)

Supposing that the system is functioning when in states 1 and 2, the system reliability is shown in Figure 6-2 for a mission time equals to 4,500.0 hours, M = 1,500 steps for the N^2- and *2N*-approaches, and 100,000 samples for the Monte Carlo simulation.

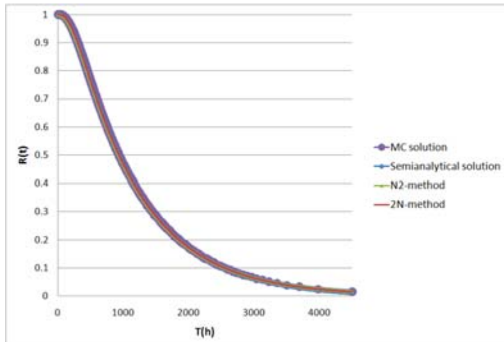


Figure 6-2 – Reliability for the semi-analytical example: mission time of 4,500 hours and M = 1500 steps.

Note that the reliability estimated by the *2N*-mathematical formulation and numerical treatment matches the computed values via the semi-analytical solution, as it also does for the MC and the N^2-solutions, thus providing a validation of the *2N*-model, with its numerical approximation.

Another important aspect to be considered is the computational effort involved in the three solution methods. Actually, the *2N*-approach with its numerical treatment computes the system reliability considerably faster than both MC simulation and N^2-method. Indeed, the *2N*-method spent 2.44 seconds, while the Monte Carlo took 133.56 seconds, and the N^2-required 2,442.19 seconds.

In the next section, it is discussed a more complex example of application in the context of temperature-pressure optical monitoring systems in oil industry.

6.2. Example of application: availability of downhole optical monitoring systems

The present section focuses on scrutinizing the main findings on the *2N*-method drawn for NHSMP on continuous-time.

For validation purposes, the solution provided by the *2N*-mathematical and numerical approaches will again be compared to the N^2-method given in Janssen and Manca (2001) and the MC simulation by means of an example in the context of reliability assessment of temperature-pressure optical monitoring systems for downhole applications in the oil industry.

The explanation of this application comes from the example 2 given in subsection 1.2.2. Provided that description, the focus will switch towards the data required for analyzing over time the availability of the optical monitoring system.

6.2.1. Required data

Optical Monitoring Systems are installed as part of intelligent completions in onshore oil wells. Each well is comprised of three production zones. Thus, a monitoring system is comprised of four pressure-temperature sensors, one for each zone plus one sensor for monitoring the artificial elevation system (oil pumping). All sensors are on the same optical cable.

Provided the lack of operational experience on these systems some simplifying assumptions are made. Indeed, although an OMS is comprised of several components (e.g., optical cable, sensor unit, jumpers, cable-cable and cable-sensor connections) as shown in Figure 6-3, the availability modeling is developed at the system level. Moreover, the reinstallation typically involves the repairman who runs the OMS downhole as well as a supervisor. The HEP is, however, modeled and quantified only for the repairman.

Therefore, and in light of these limitations, it aims to develop a model based on the combination of continuous-time SMP and BBN that is able to handle over time the joint impact of the tolerable downtime and process time as well as the human performance on the OMS availability during the execution of reinstallation activities.

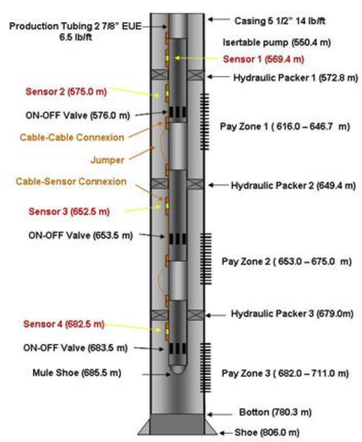


Figure 6-3 - Intelligent well with a pressure-temperature OMS

Spectral analysis of the limited OMS units that have been deployed has been used for gaining knowledge regarding OMS time to failure pattern. Figure 6-4 and Figure 6-5 show the spectral analysis for an OMS when it was installed and 20 months later, respectively.

Each sensor in the OMS is represented by a double optical power peak, where the left and right ones correspond to pressure and temperature, respectively. Thus, Figure 6-4 and Figure 6-5 have four pairs of peak corresponding to the four sensors in the OMS. If the optical power of one peak in relation to the others in the spectrum is small, then peaks are interpreted as noise, resulting in the loss of the monitoring capability.

As seen in Figure 6-4 and Figure 6-5, the spectral analysis has indicated a gradual attenuation of the optical power from the sensors, and it is more significant for the bottom sensor (indicated by a red contour).

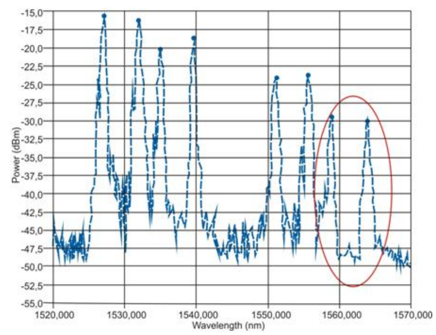


Figure 6-4 – Spectral analysis for OMS at installation

These results have led to the conclusion that the OMS units are under a deterioration process what, along with the tolerable downtime characteristics, justifies the use of an NHSMP. This process eventually leads to the complete loss of the signal, as shown in Figure 6-6 for the first two sensors from the well bottom after just 22 months from the installation date.

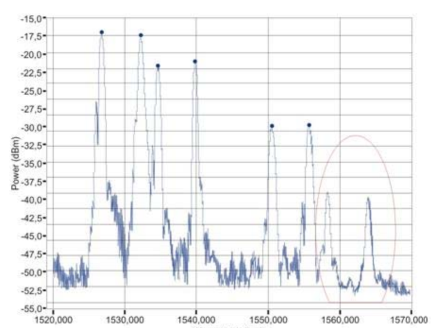


Figure 6-5 – Spectral analysis for OMS after 20 months of installation

The semi-Markov process is therefore considered as non-homogeneous so that this deterioration process may be adequately addressed. Therefore, the required data to estimate the system availability via this NHSMP model are the parameters $p_{ij}(\cdot)$ and $G_{ij}(\cdot,\cdot)$ of the NHSMP in Figure 1-3, through which the kernel $C_{ij}(\cdot,\cdot)$ can also be found.

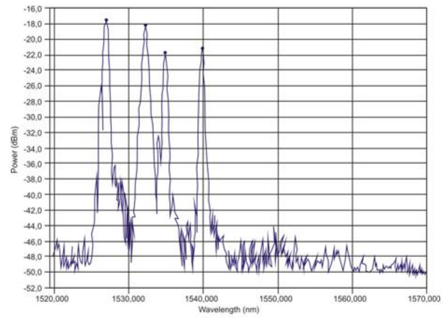


Figure 6-6 – Spectral analysis for OMS after 22 months of installation

Table 6-2 summarizes the requisite data for solving the NHSMP given in Figure 1-3. Looking at that, it can be seen probabilities $p_{11}(t)$ and $p_{12}(t)$ depend on the process time t, reflecting the deteriorating process to which the OMS is subjected. The probability $p_{12}(t)$ of leaving state 1 to state 2 is set to value 0.4 at $t_0 = 0$, increasing continuously in the interval of 15,000h. The probability $p_{11}(t)$ in turn decreases from 0.6 during the same interval.

Table 6-2 – Estimates of parameters of the NHSMP.

$i \rightarrow j$	$p_{ij}(t)$	$G_{ij}(l,t)$
1 → 1	(-0.00004 t) + 0.6	Exponential(1E-04h^{-1})
1 → 2	(0.00004 t) + 0.4	Weibull(270.0h, 1.86)
2 → 1	0.82	Exponential(0.0208h^{-1})
2 → 3	0.18	LogNormal(4.0h, (0.4h)^2)
3 → 1	0.62	Exponential(0.0416 h^{-1})
3 → 4	0.38	LogNormal(2.5h,(0.25h)^2)

Indeed, *2N*-method for NHSMP described in chapter 4 will be used for solving the NHSMP of Figure 1-3. In next subsection, the main findings on *2N*-method will be analyzed comparing its effectiveness, in terms of accuracy and computational time, with the results provided by N^2-method and MC simulation.

6.2.2. Results

In fact, for different number M of steps, from Figure 6-7 to Figure 6-10 there will be showed the estimated OMS availability ($A(t) = \phi_1(t)$) for a mission time of T = 1 year (8,760 hours) according to the *2N*- and N^2-methods and to MC simulation (100,000 samples). The main findings of this subsection are provided in Moura and Droguett (2009a), which follows in attachment H.

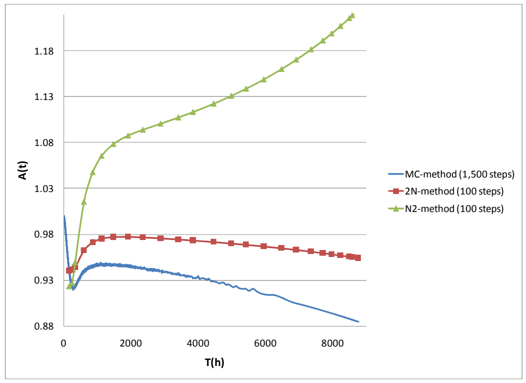


Figure 6-7 – OMS availability: mission time of 8,760 hours and $M = 100$ steps

In Figure 6-7, neither *2N*- nor N^2- show close agreement when compared to the MC solution. In this case, the distance (discretization error) between the *2N*-method and MC-simulation is shorter than that of the N^2-approach.

This figure also depicts the N^2-approach attains probability values greater than 1.0 for this number of steps, which does not look like to be as big as enough to reach the converged solution.

As the number of steps increases by 400 steps ($M = 500$) we can notice the *2N*-method already matches MC-simulation, whereas N^2-method keeps showing a noteworthy distance (Figure 6-8) from that.

Now setting up $M = 1{,}500$, Figure 6-9 depicts that the availability curve estimated from *2N*-method matches the computed values gathered from N^2- and MC approaches.

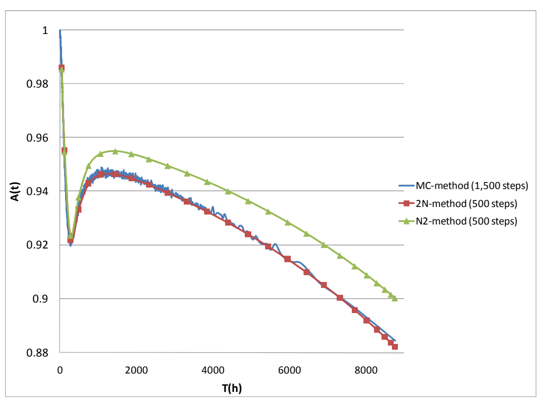


Figure 6-8 – OMS availability: mission time of 8,760 hours and $M = 500$ steps

Figure 6-10 summarizes how the *2N*-method reaches MC results as M increases. These results provide a illustrated validation, in terms of accuracy, of the mathematical formulation and numerical treatment given and developed by Moura and Droguett (2009c).

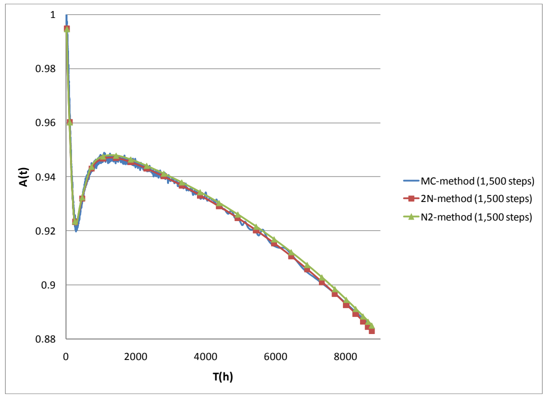


Figure 6-9 – OMS availability: mission time of 8,760 hours and M = 1,500 steps

Although the *2N*-method approximately reaches the desired convergence with only 500 steps (Figure 6-8), a similar analysis should be made for the other state probabilities ($\phi_i(t)$, i = 2, 3, 4) in order to find the number M_i necessary to converge. The number M of steps needed to achieve the process's convergence as a whole is the maximum of M_i, i = 1,…,4. In other words, this corresponds to find M_i which minimizes the upper limit of the discretization error of the *2N*-method which in turn is the sum of eq. (5-18) and the error estimated in approximating $\phi_j(kh)$ by $\tilde{\phi}_j(t)$.

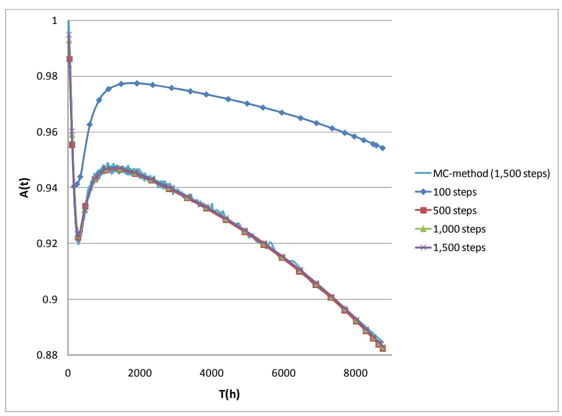


Figure 6-10 – OMS availability: *2N*-method *x* MC

It is worthwhile making two important considerations on the Figure 6-6 to Figure 6-10: (*i*) as in this case the NHSMP of Figure 1-3 is not an ergodic system (due to the absorbing state 4), availability equals reliability and the curve must go to zero. Because of this, the y-axis could have be quoted as reliability rather than availability, not reaching a steady state; (*ii*) it is observed a local minimum at the beginning of the mission. This behavior is somewhat a result of the deteriorating process under which the system is subjected when it occupies state 1. Then, it may be seen an increase in the system reliability as the failure/reinstallation cycles start (process reaches states 2 and 3).

Letting the illustrated analysis aside, we may make use of the cross-entropy measure (see Kullback and Leibler (1951) for more details) for analyzing how the discrepancy among the *2N*-, N^2- and MC-methods varies as a function of the number *M* of steps. Basically, the cross-entropy is given by

$$D(f,g)=\int f(x)\log\left(\frac{f(x)}{g(x)}\right)dx\ ,$$

where the function *f* corresponds to the results from MC-simulation and *g* could be the outcomes from the *2N*- or N^2-methods.

Figure 6-11 illustrates the cross-entropy measured for the availability metric by both the *2N*- and N^2-approaches. Generally speaking, it underpins the discussion made in section 5.2.2 on the upper limit of the discretization error estimate for NHSMP. In fact, even though the *2N*-method is faster it presents an error estimate smaller than N^2-method over the number *M* of steps and that tends to zero as *M* increases. The latter characteristic holds for the N^2-method as well.

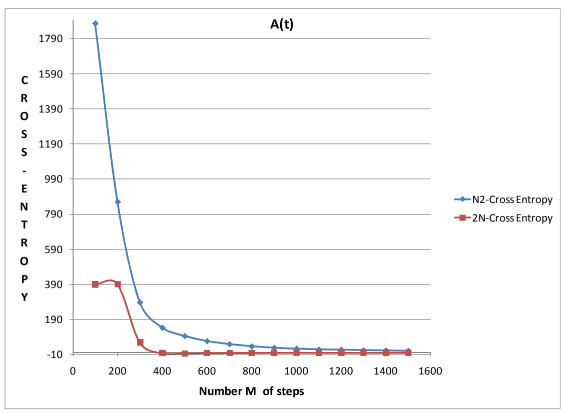


Figure 6-11 – Cross-entropy over number of steps: *2N*- x N^2-

Figure 6-11 also depicts that the findings given in section 3.1.2.2 on the accuracy of the *2N*-method for HSMP are somewhat conservative. In that occasion, it has been concluded the *2N*-discretization error would be greater or equal to N^2-error. However, in section 3.1.2.2 the inequality in eq. (3-19) was disregarded what implies the results from Figure 6-11 also holds for HSMP.

Besides validating the *2N*-method accuracy in comparison with the results of the N^2- and MC approaches, other criteria to contrast these different solutions for NHSMP is the time to converge, i.e., computational cost.

Indeed, by using the *2N*-approach it is only needed solving *N* coupled integral equations with one variable and *N* straightforward integrations, rather than computing N^2 integral equations as through the N^2-method. Indeed for *M* = 1,500 steps, the *2N*-method spent 7.10

seconds, whereas the MC took 246.43 seconds, and the N^2-approach required 8,370.24 seconds. Figure 6-12 shows the time (in seconds) took by both *2N*- and N^2-approaches as *M* varies.

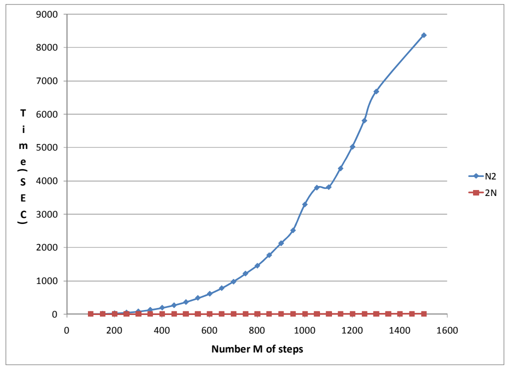


Figure 6-12 – Computational time over the number of steps: *2N*- x N^2-

This analysis provides a validation in terms of accuracy and computational effort of the *2N*-approach. Basically, it has showed that by the *2N*-method the state probabilities are obtained considerably faster than through the N^2-approach.

Moreover, it has been observed that even with a less intricate computational complexity, the *2N*-method reaches the converged solution with a truncation error smaller than the N^2-method. Obviously, one must scrutinize these outcomes in order to find whether they correspond to a general consequence or not.

6.3. Comments

6.3.1. *2N*-method

NHSMP were usually computed via the N^2-method described in Janssen and Manca (2001) where the future behavior of the system is assessed via interval transition probability equations comprised of a system of N^2 coupled integral equations with two variables, with *N* the number of states. However, this approach is rather cumbersome.

This situation motivated the development of a more efficient formulation for NHSMP that had less computational effort, but kept the accuracy in relation to the available methods in the related literature, that is, MC simulation and the N^2-approach. In fact, the proposed *2N*-mathematical formulation and numerical treatment consists of casting the N^2 coupled integral equations into an initial value problem involving transition frequency densities, and then solve *N* coupled integral equations with one variable and *N* straightforward integrations. This approach possesses the two aforementioned meaningful features: it is significantly less time-consuming and has accuracy equals to the N^2-method, as it was proved in section 5.2.2.

The *2N*-mathematical formulation and numerical approach were also illustrated by means of two examples of application in the context of reliability and availability assessment, where the effectiveness and the required computational effort of the *2N*-method were also compared against the MC and the N^2-approaches. From these examples, it was verified that the proposed approach is considerable faster than both the MC simulation and the N^2-method. Specifically on the second example, *2N*-approach reaches greater accuracy than N^2-method validating the findings in section 5.2.2 on the behavior of the upper limit of the discretization error.

6.3.2. OMS availability assessment

The availability analysis of an OMS is a complex subject. It is influenced by failure patterns of components that are usually under deteriorating processes. Maintenance also poses its own challenges, the most relevant ones being the human performance during reinstallations, and the impact of available time to complete the reinstallation on the human error probability. And the availability assessment must be carried out with high level of uncertainty as a result of the paucity of relevant empirical information.

In this context, section 6.2 has provided an availability model for pressure-temperature optical monitoring systems. The model is based on the integration between non-homogeneous continuous time semi-Markov processes and Bayesian belief networks. NHSMP portion is responsible for handling the OMS dynamics, and BBN are used to qualitatively and quantitatively model the cause-effect relationships among factors influencing the repairman error probability during reinstallations. The model has also been applied to the analysis of an OMS in an onshore oil well in the Northeast of Brazil.

With this model we only scratched the surface of the problem. Although we do not provide conclusive results from the application of the model to a real case situation, the availability model tackled the most relevant issues concerning the operation and maintenance characteristics of the system, providing the analysts with much needed flexibility for evaluating the availability of the system, as it was demonstrated in the example of application.

In the following we provide more specific comments on the limitations of the proposed availability model and discuss some alternative modeling approaches. (*i*) Given the lack of data on the failure processes, the availability analysis was done at the system level. Although the spectral analysis indicated that the optical monitoring system is under deterioration, the availability engineers and certainly the development process of the system will benefit from a disaggregated availability analysis (at component level). This might be done by including a Fault Tree sub-model which considers the combinations of component failure events leading the system to change its general state. This and failure data gathering and analysis from the deployed units (7 production and 1 injection wells have been recently equipped with the monitoring system) are part of the next stage of this research. (*ii*) The construction of the BBN model for the human error probability and the population of the table of conditional probabilities were based on expert opinion elicitation. Although all nodes were binaries, the expert had difficulties in providing quantitative assessments especially for the non-root nodes. As one of the next steps of this research, empirical data on at least some of the factors will be collected so to relief the cognitive burden on an expert. So from this and the previous step, one might get more reliable availability assessments. (*iii*) In the proposed model, the available time to complete a reinstallation is obtained from the sojourn time in an unavailable state. Although the semi-Markovian process is continuous in time, this information is passed as discrete (binary) evidence to the node "available time to complete the reinstallation" of the BBN model for the human error probability. In the next stage of this research, this node is treated as continuous. Thus, the use of hybrid BBN is a must and the iterative algorithm proposed by Neil et al. (2008) is currently being tested in the context of application of this work. (*iv*) Dynamic Bayesian networks (DBN) have been used in the area of dependability analysis, for example, Boudali and Dugan (2005), Weber et al. (2004) and Montani et al. (2008). However, at best of our knowledge, DBN-based approaches have been limited to deal with representations of homogeneous and non-homogeneous Markov processes mostly in discrete time ((Boudali and Dugan (2006) propose a continuous time based DBN framework for the analysis of dynamic fault trees), and with focus on non-repairable systems. However, DBN are an alternative approach to the availability assessment of the type of system analyzed in this research, and representation of semi-Markovian processes and repairable systems in a DBN framework are subjects of current research by the author.

7. FURTHER EXAMPLES AND COMMENTS

In this chapter two further examples will be discussed. The first one shows how the *2N*-method may be used for determining a maintenance optimization policy so that to maximize the mean availability measure. The second example is also inserted into the optimization context. Basically, it is designed for determining which maintenance decisions should be made so that the mean availability and expected costs are jointly optimized over the system's age. Thus, the *Lap*-method will be used to estimate the mean availability in this framework.

Basically, these examples will be described as an attempt to show how reliability problems, which would become infeasible due to the lack of an efficient method for solving NHSMP, may be modeled by using the proposed mathematical and numerical approaches.

7.1. System Availability Optimization

System availability optimization is one of the main issues to production managers: the greater the system availability the greater the production profits are. Provided that each preventive maintenance action promotes a rejuvenation impact on the availability measure, this section develops an approach to maximize the mean availability by identifying an optimal maintenance policy for a hypothetical system, which is modeled according to a non-homogeneous semi-Markov processes.

In order to solve the resulting optimization problem constrained by system performance costs, genetic algorithms (GA) operators will be used (see Marseguerra et al. (2006) for greater details on GA). The developments of this section are widely described in Moura et al. (2008), which follows in attachment I.

7.1.1. Description of the problem

Consider a system, which due to the same reasons discussed throughout the two last chapters, may be modeled through an NHSMP. Then, it is aimed establishing a preventive maintenance policy that maximizes the system's mean availability restricted to technological and cost constraints.

This optimal policy is comprised of operating times t_j up to the preventive action j^{th}, which has a rejuvenation impact q on the real age of the system. The parameter is incorporated into the state equations of an NHSMP so that the effectiveness of each preventive maintenance is taken into account in the optimization procedure of the mean system availability.

The mathematical programming problem relevant to the system is summarized as follows:

$$\text{Max } \bar{A}[T \mid (t_1, t_2, \dots, t_n); q] \tag{7-1}$$

$$\begin{aligned} s.t. \quad & C[T \mid (t_1, t_2, \dots, t_n); q; c_p; c_c] \leq K \\ & c_p, c_c, T, K > 0, t_i \in (0, T], N, n \in \mathbf{N}, \text{ and } q \in \Re \\ & n \leq N \end{aligned}$$

where T is the mission time under consideration; t_j is the operating time up to the j^{th} preventive maintenance action, with $t_0 = 0$; $\underline{t} = (t_1, t_2, \dots, t_n)$ composes a preventive maintenance policy; n is the number of preventive maintenance events in T and N is its upper bound; $\bar{A}[T \mid (t_1, t_2, \dots, t_n); q]$ is the system mean availability in T modeled in terms of an NHSMP and related to $(t_1, t_2, \dots, t_n)$ and q; $C[T \mid (t_1, t_2, \dots, t_n), q, c_p, c_c]$ is the cost related to the system performance in T given the maintenance policy $(t_1, t_2, \dots, t_n)$, q, the cost per time unit to perform preventive (c_p) and corrective (c_c) maintenances.

In order to compute $C[T \mid (t_1, t_2, \dots, t_n); q, c_p, c_c]$, the time spent by the system under preventive and corrective maintenances are estimated as a function of $\bar{A}[T \mid (t_1, t_2, \dots, t_n), q]$ and T_p (mean preventive maintenance time) [see eq. (7-2)]. Finally, K is a maximal cost constraint, i.e., the total cost incurred by performing corrective and preventive maintenance actions.

$$\begin{aligned} & C\left[T|\left(t_1,\dots,t_n\right);q;c_p;c_c\right] = \\ & n \cdot c_p \cdot T_p + \left[T - T \cdot \bar{A}\left[\left(t_1,\dots,t_n\right);q\right] - T_p\right] \cdot c_c \end{aligned} \tag{7-2}$$

To achieve the optimal maintenance policy, GA are introduced to the problem. Basically, GA consider a population of individuals, where each individual is a possible solution to the problem. In this context, genetic operators such as crossover and mutation are computationally mimicked in order to simulate the evolution process (see Michalewicz (1996) for more details).

7.1.2. Casting Maintenance Effectiveness into NHSMP

In order to take into account the effectiveness q of each maintenance action, let rewrite the eqs. (5-1) and (5-2). Thus, the future behavior of an NHSMP over time may now be assessed through its state probabilities $\phi_j(t)=\Pr[Z_t=j \mid Z_0]$ given by as follows:

$$\begin{aligned} \phi_j(t) = \phi_j(0) \cdot [1 - F_j(0, qt)] \\ + \int_0^t h_j(q\tau) \cdot [1 - F_j(q\tau, qt)] d(\tau) \end{aligned} \tag{7-3}$$

where $F_j(l,t)$ and $h_j(t)$ are defined in chapters 2 and 6 respectively. Thus, it follows that:

$$h_j(t) = \sum_{i=1}^{N} \phi_i(0) \cdot \dot{C}_{ij}(0, qt) + \sum_{i=1}^{N} \int_0^t h_i(q\tau) \cdot \dot{C}_{ij}(q\tau, qt) d\tau \tag{7-4}$$

where $\dot{C}_{ij}(\cdot,\cdot)$ is also defined in chapter 2.

Eq. (7-3) and (7-4) are modified versions of (5-1) and (5-2) respectively, keeping the same meaning as the latter though. In fact, backed up the virtual age model called General Renewal Process proposed by Kijima and Sumita (1986), the parameter q is introduced in order to measure the effectiveness of maintenance actions. In other words, the parameter q is used to handle the rejuvenation imposed to the system after the last maintenance event.

In accordance with q, a maintenance action can recover the system to some of the possible states: (*i*) $q = 0$ – as good as new (perfect repair); (*ii*) $q = 1$ – as bad as old (minimal repair); (*iii*) $0 < q < 1$ - better than old but worse than new (imperfect repair). The impact of these types of maintenance on the system's availability is illustrated in Figure 7-1.

According to Figure 7-1, it can be noticed that up to the first maintenance action there is no difference among the three types of repair analyzed. However, just after the first intervention the impact q of each maintenance action on system availability may be assessed. Moreover, Figure 7-1 illustrates that while the system is unavailable and under preventive maintenance the instantaneous availability is zero. For further details on the classical and Bayesian procedures for estimation of the parameter q see Yañez et al. (2002).

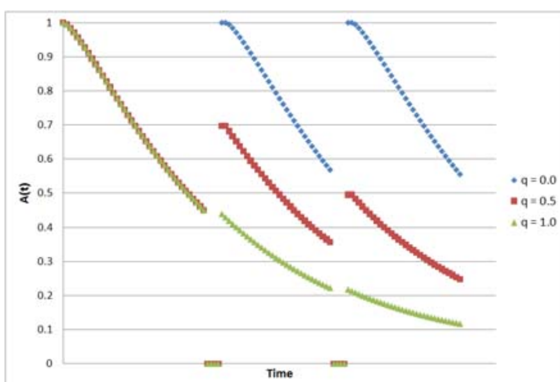


Figure 7-1 – Impact of different types of repair on the availability

The model developed here proposes a preventive maintenance policy that maximizes the mean availability of a system which is estimated according to an NHSMP by considering the impact q of each maintenance action on the system performance. Therefore, the objective function presented in (7-1) is the mean availability given by (7-5), where $\{A\}$ is the set of states in which the system is available. This mean availability as well as the system dynamics as a whole will be estimated by using the *2N*-method described in chapter 5 for NHSMP.

$$\overline{A} = \frac{1}{t} \sum_{k \in \{A\}} \int_0^t \phi_k(\tau) d\tau \tag{7-5}$$

7.1.3. Example

It is assumed that a hypothetical system starts in normal operation in state 1. Over time, due to operational and/or environmental conditions, the system may operate in a degraded state even though it is still available, which corresponds to the state 2. In this state, the corrective maintenance process, which consists of the installation or reinstallation of the system, starts. There is a tolerable downtime (TDT) inside which the system may operate in this degraded condition. If the repair process cannot be completed within this TDT, the system goes into state 3 where additional corrective actions are taken to restore it to its normal operating condition, but in this case the system is unavailable. It is assumed that all corrective actions recover the system to the same condition it had just before the failure. In other words, corrective actions are considered as minimal repairs (q = 1.0).

Besides the corrective maintenance actions, the system may also undergo preventive maintenance events (state 4) which possess an effectiveness q. It is assumed that all preventive maintenance occurrence times are known at the start of the mission (at $t=0$). This preventive maintenance policy corresponds to a particular individual in the GA optimization algorithm, i.e., the NHSMP model is evaluated for each potential solution, (t_1, t_2, …, t_n), during the execution of the optimization procedure.

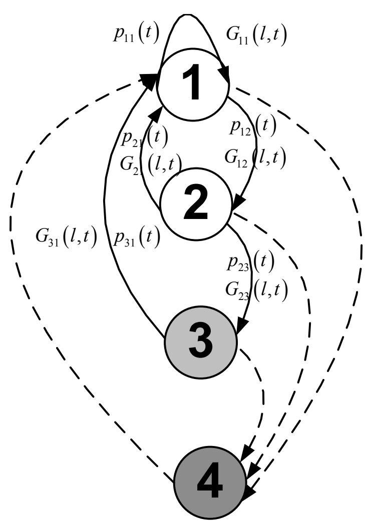


Figure 7-2 – Non-Homogeneous semi-Markov processes for a hypothetical system

It is also considered that a preventive maintenance action corresponds to an imperfect one, i.e., it recovers the system to a condition somewhere between perfect and minimal repairs. The state space diagram is illustrated in Figure 7-2, where the dotted lines represent the transitions due to the preventive maintenance events.

As it has been previously discussed, the goal is to establish a preventive maintenance policy that maximizes the mean availability of this system, which in turn is modeled via NHSMP whose required parameters are given in Table 7-1. Other requisite data are T = 150 days; T_p = 1 day; N = 5 maintenances; c_p = \$10.00; c_c = \$30.00; K = \$200.00. The GA-based required parameters are showed in Moura et al. (2008).

Table 7-1 – Parameters estimation for the NHSMP.

$i \rightarrow j$	$p_{ij}(t)$	$G_{ij}(l,t)$
$1 \rightarrow 1$	(-0.0034 t) + 0.5	Exponential (5E-01)
$1 \rightarrow 2$	(0. 0034 t) + 0.5	Weibull (30.0, 1.36)
$2 \rightarrow 1$	0.70	Exponential (1.0)
$2 \rightarrow 3$	0.30	LogNormal (2.5, 0.25)
$3 \rightarrow 1$	1.0	Exponential (1.0)

With the intention of evaluating the proposed approach, the preventive maintenance effectiveness parameter q is initially considered equals to 0.0, and the system performance cost is disregarded.

For q = 0.0, a preventive maintenance action restores the system to an "as good as new" condition, and therefore it is expected that t_j is approximately equally spaced over time. In relation to the system mean availability, it can be seen in Figure 7-3 that the optimal maintenance policy for the system is roughly given by $t_j \cong$ 24 days, for any j= 1, 2, …, 5. Call this t_j the target value. Under this condition, the resulting system mean availability is equal to 0.9975.

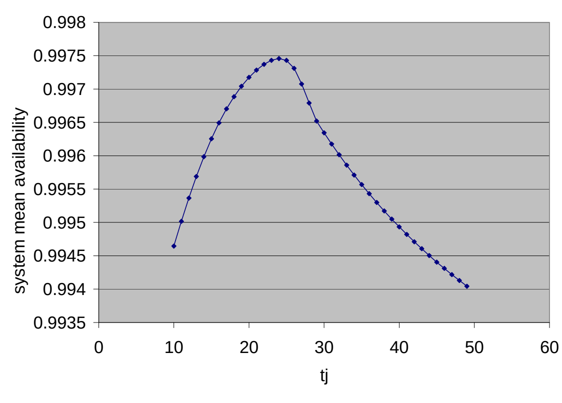


Figure 7-3 – System mean availability for different maintenance policies

By considering the input parameters, the best solution provided by the proposed NHSMP-GA approach corresponds to the preventive maintenance policy (in days) {21.10, 19.44, 24.91, 25.13, 29.16} for which the system mean availability is equal to 0.9974.

The maximum absolute error in comparison with the above target solution is about 5.16 days, generated by t_5=29.16 days. However, when considering a set of 50 replications of the algorithm the target t_j (24 days) lies inside the 95% confidence interval for each generated sample (see Table 7-2).

Table 7-2 – 95% Confidence interval (in days) for each preventive maintenance based on a GA sample of optimal solutions with q = 0.0.

Mean operating times up to maintenance	lower bound	upper bound
t_1	23.50	29.50
t_2	21.02	31.00
t_3	20.45	32.54
t_4	21.76	29.84
t_5	21.25	28.97

Now, considering q = 0.35, the best solution provided by the NHSMP-GA approach corresponds to the preventive maintenance policy (in days) {32.46, 17.46, 15.25, 15.03, 14.91} for which the system mean availability is equal to 0.9941. The resulting system performance cost is estimated as $50.00. Note that, on average, due to the high availability, the fraction of time the system spends under corrective maintenance is virtually zero, i.e., the preventive maintenance policy avoids the expenses with corrective maintenance actions.

Provided that no analytical solution is available for the system availability and for the optimal solution accordingly, the assessment of the uncertainty about the estimated optimal preventive maintenance policy is with no doubt relevant information for the decision maker. In fact, this uncertainty is characterized in terms of the 95% confidence intervals for the mean occurrence time of each maintenance event based on 50 replications of the algorithm (see Table 7-3).

Table 7-3 – 95% Confidence interval for each preventive maintenance based on a GA sample of optimal solutions with $q = 0.35$.

Mean operating times up to maintenance	lower bound	upper bound
t_1	21.56	36.41
t_2	17.08	23.87
t_3	14.98	21.15
t_4	14.34	20.07
t_5	13.83	19.98

7.1.4. Comments

This section has presented an approach for handling the maximization of system mean availability by determining an optimal preventive maintenance policy constrained to the system's performance costs.

In the approach, the system dynamics is modeled via non-homogeneous semi-Markov processes by using the *2N*-method given in chapter 5, where the idea behind Generalized Renewal Processes is employed to characterize the effectiveness of corrective and preventive maintenances on the system age.

7.2. Semi-Markov Processes for Decision-Making

As it has been seen previously in section 1.2.2 and chapter 6, pressure-temperature optical sensors have been developed to improve the management and control of oil reservoirs. One of their aims is to decrease the number and impact of intrusive maintenance interventions since the (re)installation procedures are human intensive and might influence the life of the monitored systems. Therefore, maintenance policies that jointly optimize mean availability and expected cost rate associated with maintenance interventions on monitored systems are a must in oil industries.

This section proposes a multiobjective optimization model based on semi-Markov decision processes (SMDP) to find a set of nondominated maintenance policies. Each obtained policy is of threshold type and it represents the optimal decision (do-nothing, minimal maintenance or replacement) whenever the system enters a new deterioration stage. An example of application is also discussed. This section is based on the findings of Moura et al. (2009), which follows in attachment J.

7.2.1. Description of the problem

Considered that data collected from OMS might be used in a pattern recognition technique (e.g., Support Vector Machines (SVM), see Burges (1998) and Shawe-Taylor (2000)) to indicate at which deterioration state the system is. Given that, adequate actions should be taken so that the number of interventions is minimized.

Due to the complexity of systems from oil industry, these interventions are intrusive, highly human-intensive and cost-consuming and thus minimizing them means decreasing the impact of human performance on the system and related costs as well.

Preventive actions, which set a periodic interval to perform planned maintenances, ignore the health status of a physical equipment/system. Therefore, they may not be adequate to oil industry systems since sometimes they would imply unnecessary actions, *i.e.*, as system has not crossed the critical deterioration line yet. On the other hand, pre-set times for preventive actions might also not pay enough attention on the system, even if a latent failure will take place next. Both situations are cost and time consuming and should be attenuated.

In this way, according to Jardine et al. (2006) more efficient maintenance approaches such as condition-based maintenance (CBM) may be implemented to handle this situation. CBM is a maintenance program that recommends maintenance actions based on the information collected through condition monitoring (OMS, for instance). CBM attempts to avoid systems being over or under maintained by taking maintenance actions only when there is evidence of abnormal behaviors of a physical asset.

If properly established and effectively implemented, a CBM program can significantly diminish maintenance costs by reducing the number of scheduled preventive maintenance operations.

One of the main key steps of a CBM program is to recommend efficient decision policies, which involves maintenance decision-making analysis that essentially depends on the system deterioration states.

Suppose a hypothetical system is monitored continuously, data on physical variables are collected from OMS, and then processed to find the system deterioration state. Moreover, assume there are three possible decisions which, generally speaking, will depend on the system state: do-nothing (N), minimal maintenance (M) or replacement (R).

Basically, this section is based on the work of Moustafa et al. (2004). Similarly to them, it is allowed one of three decisions $\delta_i = \{N, M, R\}$ at each deterioration state i. Moreover, SMDP will be used.

However, from this point the approach adopted here departs from the one presented in Moustafa et al. (2004). Firstly, it will be considered two objectives to optimize: the expected long-run cost rate and the expected availability, whereas in Moustafa et al. (2004) just the expected long-run cost rate is minimized. Secondly and conversely to Moustafa et al. (2004), in order to handle this multiobjective problem, a multiobjective genetic algorithm is applied (see Deb (1999)). Thirdly and finally, the *Lap*-method developed in this thesis will be used in order to estimate the mean availability measure.

Some optimization approaches have been presented in literature in order to attain optimal maintenance policies for the single objective problem. For instance, Castanier et al. (2003) investigate the problem of inspecting and maintaining a repairable system subject to continuous deterioration processes. They aim to find a policy, by means of Markov renewal approach, that optimizes system performance. Chen and Trivedi (2005) use an SMDP value iteration algorithm to find the optimal maintenance policy jointly with the optimal inspection rate. Kim and Makis (2009) apply SMDP with a modified policy iteration algorithm in order to find an optimal maintenance policy, such as in previous works, concerning only the minimization of the expected long-run cost rate.

Note that the mentioned works consider only a single objective optimization and, as asserted by Castanier et al. (2003), there is a necessity of optimization schemes adapted to the multiobjective nature of maintenance problems.

In a multiobjective optimization perspective, instead of finding a unique solution (an optimal maintenance policy), one may obtain a set of nondominated maintenance policies that present the compromise between the considered objectives (in this work, expected long run cost and mean availability).

Deb (1999) emphasizes that evolutionary algorithms such as GA are useful tools in handling multiobjective problems since they consider various potential solutions in a single run and several objectives can be treated separately.

Basically, given a solution provided by the multiobjective GA, the embedded SMDP calculates the associated values of both objectives, which are fed back to the multiobjective GA. Thus, solutions evolve throughout algorithm iterations by means of the genetic operators. An evaluation of the dominance-nondominance relation between every pair of potential solutions takes place and, at the end, a set of nondominated maintenance policies may be obtained.

Therefore, the main purpose of this section is to determine a way of how the decisions $\delta_i = \{N, M, R\}$ should be made in order to determine a set of nondominated steady state maintenance policies which minimize the expected long-run cost rate as well as maximizing the expected system availability via continuous time SMDP and multiobjective GA. In fact, the *Lap*-method described for continuous-time SMP will be adopted to compute the mean availability of the system. Regarding the multiobjective GA portion of the model, details may be found in Moura et al. (2009). The model SMDP-GA will be validated in subsection 7.2.3, comparing its results against an exhaustive multiobjective algorithm that assesses all possible maintenance policies.

7.2.2. Model Characteristics

SMDP will be used here to tackle the behavior of some systems in oil industry since it is assumed that the local time spent at each state influences the system dynamics. For the sake of simplicity, the analysis is accomplished at system level.

In this section, the SMDP's role is to determine which actions should be made at each decision epoch so that to optimize long-run cost rate and availability. Conversely to Love et al. (2000), who consider the time as the system goes down as a decision epoch, and similarly to Moustafa et al. (2004), it is assumed here a decision should be made at each time when the deterioration state changes.

At every decision epoch, an action must be taken, which implies some costs and elapsed times, for example, cost to replace the system or time to perform a minimal maintenance. Thus, the aim is to determine a sequence of decisions, which jointly optimize mean costs and availability, by using SMDP.

According to Makis and Jardine (1993), the optimal replacement policy for this type of system is of the control-limit form (threshold type). That is, for i^{th} state there is a decision $\delta_i = \{N, \text{ for } i < k_1; M, \text{ for } k_1 \leq i < k_2; R, \text{ for } i \geq k_2\}$, where k_1 and k_2 are the threshold indexes.

Let $S = \{1, \ldots, n+1\}$ represent the finite state space, where the state 1 represents the initial operational system and $n+1$ means the system reaches the most critical deterioration stage (see Figure 7-4). At these states, the decisions are do-nothing and replacement, respectively.

Between these extreme states, there are some intermediate ordered deterioration stages i, where it is needed to determine what is the more adequate action ($\delta_i = \{N; M; R\}$) that should be taken to optimize the expected long-run costs and availability.

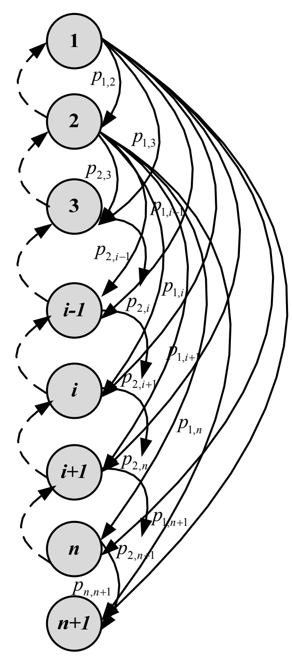


Figure 7-4 – State diagram for SMDP

Suppose, for example, the considered system might go through n deterioration stages before reach the highest deterioration level as can be seen in Figure 7-4. At each state, there are either rewards (availability) and/or losses (costs) depending on the decision to be made.

For decision "do-nothing", which may be made at any state $i \in S - \{n+1\}$, there is an operating cost per unit time of a_i. For decision "minimal maintenance", for each state $i \in S - \{1, n+1\}$, there are the mean maintenance cost b_i and time t_i. Just after this action, the system goes to state (i-1) and restart its operation from there. For decision "replacement", which may be taken at any state $i \in S - \{1\}$, there are the mean replacement cost c_i and time r_i. The system returns to "as good as new" condition at state 1, immediately after a replacement. For the two last decisions (M, R), there is also an idle cost m related to the elapsed time during which the system is not operational.

Thus, it is needed to choose the set of pairs (k_1, k_2) of thresholds that will point out the decisions to be made at each state so that the expected long-run cost rate G_∞^* and availability A_∞^* are minimized and maximized, respectively. G_∞^* and A_∞^* are given as follows:

$$G_\infty^* = \lim_{t \to \infty} \frac{E(C_p(t) + C_m(t) + C_r(t) + C_u(t))}{t} \tag{7-6}$$

$$A_\infty^* = \lim_{t \to \infty} \frac{E(D_o(t))}{t} \tag{7-7}$$

where $C_p(t)$, $C_m(t)$, $C_r(t)$ and $C_u(t)$ are the costs of production, minimal maintenance, replacement and interruption of the business over the time and $D_o(t)$ is the time portion during which the system is operational.

After a replacement, which might be either corrective (n+1-th state) or preventive (k^{th} state, where $k \in (2, n]$), the system is completely recovered to the first deterioration level. In accordance with Castanier et al. (2003), because of this regenerative property, and following a widely used approach in maintenance modeling based on the renewal theorem, the long-run study (*i.e.* on an infinite time span) of the deterioration process can be limited to the study of the system state evolution on a single renewal cycle defined by the time period between the instant when the system enters the first state and the moment when it undergoes a replacement.

In that way, one may calculate eqs. (7-6) and (7-7) by considering just a replacement cycle through eqs. (7-8) and (7-9):

$$G^*(k_1, k_2) = \frac{C(k_1, k_2)}{T(k_1, k_2)} \tag{7-8}$$

$$A^*(k_1,k_2)=\frac{T_o(k_1,k_2)}{T(k_1,k_2)} \quad (7\text{-}9)$$

where $T(k_1, k_2)$ is the expected long-run elapsed time required for the system goes from state 1 to state n+1 and is dependent on which decisions will be made at each state i. Associated with $T(k_1, k_2)$, follow $C(k_1, k_2)$ which is the expected long-run cost. Eq. (7-9) in turn is computed by settling the mission time at $T(k_1, k_2)$ and using the *Lap*-method described in section 3.2 since SMDP is considered homogeneous in relation to the process time.

7.2.3. Example

As previously discussed, the system under analysis might go through n states, before reaching the most critical deterioration state n+1. Hence, it is adopted here a multiobjective perspective integrated with SMDP, for handling the problem of characterizing the sequence of decisions along the n states. In this section, the results of the proposed model are validated by means of an exhaustive multiobjective example.

Decisions are made in accordance with the pairs (k_1, k_2) of thresholds that jointly optimize the expected long-run cost (eq. (7-8)) and availability (eq. (7-9)).

Table 7-3 presents cost data in monetary units and the distribution functions F_i of the waiting time in the state i needed to feed the SMDP portion. Apart from the first state, all F_i, $i \neq 1$, are exponential with parameter $\lambda_i = \lambda_{i-1}+0.0001$, i = 3, …, n, and $\lambda_2 = 0.0011$. Furthermore, it is considered that the cost m of the system loss per unit time is 18, $b_i = 0.03 \cdot c_i$ and $t_i = 0.03 \cdot r_i$.

Table 7-4 – Required data for SMDP portion.

State	F_i	a_i	c_i (10^3)	r_i
1	Wei(0.001, 1.36)	17	50	24
2	Exp(0.0011)	a_1+7	c_1+7.5	r_1+3
3	Exp(λ_2+0.0001)	a_2+7	c_2+7.5	r_2+3
⋮	⋮	⋮	⋮	⋮
i	Exp(λ_{i-1}+0.0001)	a_{i-1}+7	c_{i-1}+7.5	r_{i-1}+3
⋮	⋮	⋮	⋮	⋮
n	Exp(λ_{n-1}+0.0001)	a_{n-1}+7	c_{n-1}+7.5	r_{n-1}+3
n+1	–	–	c_n+7.5	r_n+3

The transition probabilities from deterioration state i to j are given by:

$$p_{ij}=\frac{n+2-j}{\sum_{j=i+1}^{n+1} n+2-j},\ i > j;\ p_{ij} = 0, \text{ otherwise}$$

The data acquired by continuous condition monitoring could have been used for directly estimating the failure rates and other parameters of the semi-Markov decision model. These data are also used for determining at which deterioration state the monitored system is likely to be.

The number of SMDP accesses by the exhaustive multiobjective algorithm in order to evaluate the considered objectives is exactly the quantity of potential maintenance policies (*nMP*). In the case of three possible decisions and hence two threshold indexes, this quantity is defined as:

$$nMP = \frac{n(n-1)}{2} + 1 \tag{7-10}$$

On the other hand, the multiobjective GA has an upper limit to the quantity of SMDP assessments, which is given by:

$$nEval \leq nInd \cdot (nGen+1) \tag{7-11}$$

where *nInd* and *nGen* are the number of individuals and generations respectively.

In this way, regarding the number of SMDP evaluations, it is surely worth using multiobjective GA instead of exhaustive multiobective algorithms since the equality in eq. (7-11) is hardly met in practice. This occurs since only different potential solutions may be evaluated by means of SMDP and, as the multiobjective GA evolves and converges towards the nondominated set, there is a reduction of the number of different solutions to be assessed. Taking this fact into consideration, it was set $n = 50$.

All experiments were executed in the same PC setting as in chapters 4 and 6. The exhaustive multiobjective algorithm found 38 nondominated solutions associated with the true Pareto front (see Table 7-5 and Figure 7-5). It required 6079.4 seconds to assess all of the 1226 possible pairs of thresholds (k_1, k_2).

Table 7-5 - True nondominated solutions

k_1	k_2	Exp. long-run cost rate	Mean availability
1	32	417.2220	0.5396
1	31	417.2625	0.5503
1	30	417.4126	0.5609
1	29	417.6861	0.5715
1	28	418.0987	0.5820
1	27	418.6685	0.5924
1	26	419.4156	0.6028
1	25	420.3632	0.6131
1	24	421.5382	0.6233
1	23	422.9713	0.6335
1	22	424.6987	0.6437
1	21	426.7623	0.6537
1	20	429.2119	0.6637
1	19	432.1065	0.6736
1	18	435.5165	0.6834
1	17	439.5271	0.6930
1	16	444.2420	0.7024
1	15	449.7887	0.7117
1	14	456.3262	0.7207
1	13	464.0547	0.7293
1	12	473.2302	0.7375
1	11	484.1855	0.7452
1	10	497.3612	0.7521
1	9	513.3531	0.7581

1	8	532.9875	0.7628
1	7	557.4443	0.7656
1	6	588.4696	0.7662
2	29	406.1186	0.4550
2	28	406.2036	0.4618
2	27	406.4278	0.4683
2	26	406.8091	0.4746
2	25	407.3677	0.4806
2	24	408.1267	0.4864
2	23	409.1130	0.4918
2	22	410.3578	0.4969
2	21	411.8974	0.5016
2	20	413.7748	0.5058
2	19	416.0406	0.5095

Then 10 trials of the multiobjective SMDP with GA were executed. Table 7-6 presents the number of obtained nondominated threshold pairs, the quantity of exact solutions of the true nondominated set, the number of SMDP evaluations and the execution time as well.

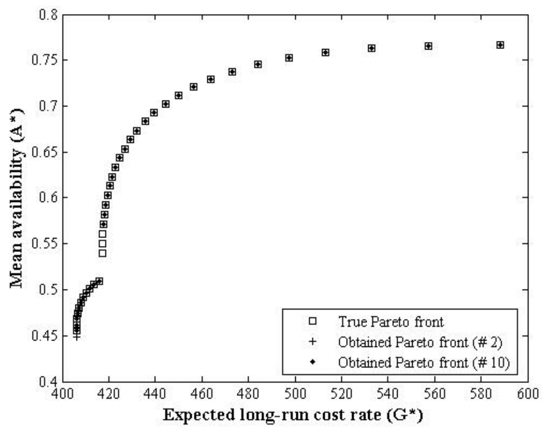


Figure 7-5 - True Pareto front and some obtained fronts from multiobjective SMDP + GA

Note that even the upper limit of *nEval* being relatively large (20100) for the present example application, the number of SMDP evaluations is, on average, about 37% of *nMP* = 1226. In addition, the multiobjective GA in junction with the SMDP was able to find approximately 91% of the true nondominated set in about 10% of the time required by the exhaustive multiobjective algorithm. Figure 7-5 depicts the true Pareto front and fronts #2 and #10 obtained by the SMDP with multiobjective GA. Note that all points from #2 and #10 fronts are on or very nearby the true solutions.

Table 7-6 - Results of multiobjective SMDP + GA

Trial #	Obtained (k_1, k_2)	Exact Pareto solutions	SMDP evaluations	Time (seconds)
1	35	35	443	616.5
2	35	34	460	640.8
3	35	35	445	604.7
4	35	35	460	644.8
5	29	29	412	626.6
6	34	34	441	578.3
7	36	36	470	647.1

8	36	36	468	657.8
9	36	36	463	626.7
10	35	35	448	623.8
Mean	34.6	34.5	451	626.7
Std. dev.	2.07	2.07	17.28	23.21

It can also be observed from Table 7-5 that all policies indicate it is interesting to perform minimal maintenance actions in early deterioration states. In addition, the sooner the replacement, the higher the mean availability and the expected long-run cost rate reached.

In this way, decision makers may evaluate how much they are disposal to spend in order to obtain a gain in mean availability. This can be done by means of a Return of Investment analysis (*ROI*) between two different maintenance policies from the solution set:

$ROI = (A_i^* - A_j^*) / (G_i^* - G_j^*),\ i \neq j$

For example, to change from policy (2, 23) to policy (1, 9), the *ROI* is equal to 0.00255.

7.2.4. Comments

This section proposed a multiobjective optimization model based on semi-Markov decision processes for the optimal replacement policy for monitored systems from oil industry. The proposed multiobjective approach was validated by means of an exhaustive algorithm and was able to find almost all solutions from the true nondominated set in a considerable reduced time frame.

The ongoing research is to integrate this multiobjective portion with a Gamma process which has been commonly used to address issues related to continuous degradation (see Noortwijk (2009) for more details). In this way, the work developed by Castanier et al. (2003) would be extended.

Moreover, following the findings in Kim and Makis (2009), in upcoming developments, it will also be considered imperfect maintenance actions which will allow recovering the system to an intermediate deterioration state between the "as good as new" and "as bad as old" conditions.

7.3. *Pattern Recognition Problem through Support Vector Machines*

Clustering the continuous data acquired via intelligent technologies into a set of discrete states, which is given in Figure 7-4, is the first step of the proposed model described in previous section. This requirement consists in collecting data, training them using a pattern recognition technique and analyzing its effectiveness on unseen data.

Basically, it corresponds to a multi-classification problem in which each state *i* represents a category where the system could be. The work presented in the previous section is supposed to start from the definition of this set of states and requires that a pattern recognition model determines periodically the condition state.

Support Vector Machines (SVM) has been successfully applied to pattern recognition and regression problems. SVM is able to obtain noteworthy results when a data set $\{(x_1,...,y_1),...,(x_l,...,y_l)\} \subset \chi \times \Re$ is available for training, where χ denotes the space of the input points (e.g., $\chi = \Re^d$, where d is the χ dimension). For instance, these might be time to failure (Hong and Pai (2006)), dissolved gases content in power transformer oil (Fei et al. (2009)) or reliability of software (Pai (2006)).

Experimental results have revealed SVM performs better than other techniques such as Artificial Neural Networks (ANN), Grey Model, Multi-Layer Perceptron network-based method, the Radial Basis Function network-based method, and autoregressive integrated moving average (see Pai (2006)).

There are two main reasons which explain SVM surpasses these techniques (mainly ANN, which are one of the commonest methods used in forecasting): (*i*) rather than backed up the empirical risk minimization (which minimizes the training errors) as ANN, SVM makes use of the structural risk minimization. Through this principle, SVM seeks to minimize an upper bound on the generalization error. This fact plays an important role since minimizing the number of training errors appears to be computationally demanding and it guarantees good generalization performance as well; (*ii*) solving a classification or regression problem via SVM corresponds to deal with a convex quadratic optimization problem. Karush-Kuhn-Tucker conditions state a necessary clause for a point $\epsilon\, \chi$ to be a global solution and also are sufficient conditions when the objective function is convex.

Therefore, SVM are not plagued with the problem of local minima as ANN are. For more details on these SVM characteristics see Shawe-Taylor (2000) and Burges (1998) for the classification problem and Smola and Scholkoff (2004) for the regression case.

In this way, SVM could have be used to address the pattern recognition (multi-classification) problem necessary as an intial step to tackle a condition monitoring problem such as the one discussed in previous section. This is issue of our ongoing research.

8. CONCLUSIONS

8.1. Final Remarks

Continuous-time homogeneous semi-Markov processes are important probabilistic tools to model reliability measures for systems whose future behavior is dependent on the current and next states of the process and on sojourn times, besides the process time in case of non-homogeneity.

SMP have been traditionally solved via the N^2-method described in Corradi et al. (2004) and Janssen and Manca (2001) for HSMP and NHSMP respectively, where the system dynamics are assessed via interval transition probability equations comprised of a set of N^2 coupled integral equations. However, as it can be seen in chapters 4 and 6, this approach has been rather burdensome and is not straightforward to implement.

This reason has motivated delving for a more efficient numerical treatment of SMP with less computational effort and with a comparable accuracy in relation to the available methods in the related literature (MC simulation and the N^2-approach).

Therefore, this research has given rise to the *2N*-mathematical formulation and numerical treatment which consists of casting the N^2 coupled integral equations into an initial value problem involving transition frequency densities, and then solve N coupled integral equations with one variable and N straightforward integrations. Through the analysis of some examples, it has been seen this approach possesses the two aforementioned features: the *2N*-method convergence speed is greater than the other approaches and has a discrepancy from the MC-results smaller than that of N^2-approach, corroborating the main findings provided in section 5.2 on the upper limit of the *2N*-discretization error.

Speaking specifically on NHSMP, the *2N*-method plays an important role to leverage the feasibility of application of this type of stochastic model. Although NHSMP are powerful modeling tools, the mathematical and computational difficulties inherent to the N^2-method on the non-homogeneous environment are usually blamed as accountable for the impracticability of this type of stochastic process.

Although *2N*-method has showed meaningful outcomes in terms of computational effort and accurateness as well, both approaches (*2N*- and N^2-) have an important drawback to be considered: they require increasing the algorithm's order (number M of steps) so that to attain greater accuracy. Hence, this reason increases the effort for solving semi-Markov processes, since the minimum number *M* of steps to reach a maximum discretization error should be known previously.

This situation has motivated the development of the *Lap*-numerical procedure which made use of Laplace Transforms for solving HSMP. Through a couple of examples of application in the context of reliability engineering, *Lap*-procedure has showed some noteworthy advantages: (*i*) it used a pre-set number of steps, which is independent on the problem to be solved. Thus, it is not required anymore adjusting (through either trial-error tests or dynamically) the number *M* of steps in order to attain the desired convergence. (*ii*) thus, it reduced considerably the computational effort in relation to the *2N*- and N^2-methods and MC as well. (*iii*) *Lap*-numerical procedure has been designed for treating HSMP specified in terms of either transition probabilities or transition rates (*iv*) it has possessed accurateness comparable to the *2N*- and N^2-method and MC solution. However, the same meaningful results have not been encountered for the non-homogeneous case, as can be seen in Moura and Droguett (2007).

Finally, this thesis has presented two further examples. In those examples, the numerical procedures developed in this work have been used in optimization and decision-making problems.

Indeed, the first example have developed an approach to maximize the mean availability by identifying an optimal maintenance policy for a hypothetical system, which is modeled according to a non-homogeneous semi-Markov processes. Hence, the *2N*-method, which has been drawn for NHSMP in chapter 6, has been used to estimate the availability measure. Genetic algorithms in turn have been adopted to perform the optimization task of the approach.

The aim of the second further example has been to establish a way of how the decisions do-nothing, minimal maintenance and replacement should be made in order to determine a set of nondominated steady state maintenance policies which jointly minimize the expected long-run cost rate as well as maximizing the expected system availability via continuous time SMDP and multiobjective GA. Thus, the *Lap*-method described for continuous-time SMP has been adopted to compute the mean availability of the system. The model SMDP-GA was validated comparing its results against an exhaustive multiobjective algorithm.

8.2. Limitations, Ongoing Research and Future Challenges

8.2.1. Semi-Markov Processes: Requisite data, *2N*- and *Lap*-methods

Regarding semi-Markov processes, three important limitations of this work deserve attention. Firstly discussing semi-Markov processes in general, we have the well-known and already quoted difficulty in obtaining the requisite data to analyze semi-Markov processes, mainly on the non-homogeneous environment. Regarding this issue, El-Gohary (2004) presents maximum likelihood and Bayesian estimates of the parameters included in a semi-Markov reliability model of three states.

The second limitation lies on how to find out a number M of steps to minimize the discretization error computed from the *2N*-method. Up to now, this variable is not calculated on simulation time what makes necessary to test several solutions of the *2N*-method (with different M) and check them out in comparison with the MC results. Sometimes, this is a quite tough task. This is the issue of our ongoing scientific researches.

Thirdly, one drawback that deserves attention on the *Lap*-method is since this approach is based upon Gaussian Quadratures theory there is not a quite simple way to obtain an estimate of the absolute error committed by the approach (see Press et al. (2002) for more details). This is also topic of our ongoing research.

As suggestion of future works, one could apply other numerical inversion Laplace transform methods such as Cuomo et al. (2007) in order to compare with the results provided in the present work.

8.2.2. Support Vector Machines

As it has been mentioned in section 7.3, the first steps in a condition-based maintenance problem correspond to clustering the continuous data acquired via intelligent technologies into a set of discrete states. This requirement would consist in gathering data, training them using a pattern recognition technique and analyzing its adequateness to test data.

Basically, it consists of a multi-classification problem in which each state i represents a pattern. Thus, the findings presented in section 7.2 are supposed to start from the definition of this set of states and requires that a pattern recognition model determines periodically the condition state.

One subject of our current research is to use support vector machines to address the former steps of a condition-based problem. Mainly, due to the reasons explained in section 7.3, SVM has been successfully applied not only to pattern recognition problems, but also to regression ones, surpassing techniques as Artificial Neural Networks, Grey Model, Multi-Layer Perceptron network-based method, the Radial Basis Function network-based method, and autoregressive integrated moving average (see Pai (2006)), what underpins its use for treating CBM related matters.

REFERENCES

Abate, J. and Valkó, P. P. Multi-precision Laplace transform inversion. *International Journal for Numerical Methods in Engineering*, v.60, p.979-993. 2004.

Abramowitz, M. and Stegun, J. A. *Handbook of Mathematical Functions*. New York: Dover. 1972.

Afchain, A. L. Non-parametric estimation of lifetime and repair time criteria for a semi-Markov process. *Comptes Rendus Mathematic*, v.339, p.137-140. 2004.

Baker, C. T. H. *The numerical treatment of integral equations*. New York: Clarendon Press. 1977.

Barros Jr., P. F. R. *A methodology for availability assessment of complex systems via hybridism between Bayesian Networks and Markov processes (in Portuguese)*. Department of Production Engineering, Technology and Geoscience Centre, Federal University of Pernambuco, Brazil, Recife, PE-Brazil, 2006.

Becker, G., Camarinopoulos, L. and Zioutas, G. A Markov Type model systems with tolerable downtimes. *Journal of the Operational Research Society*, v.45, p.1168-1178. 1994.

Becker, G., Camarinopoulos, L. and Zioutas, G. A semi-Markovian model allowing for inhomogenities with respect to process time. *Reliability Engineering & System Safety*, v.70, p.41-48. 2000.

Bellman, R., Kalaba, R. E. and Locket, J. A. *Numerical Inversion of the Laplace Transform*. New York: American Elsevier. 1966.

Bellman, R. E. and Roth, R. S. *The Laplace Transform*. Singapore: World Scientific Publishing Co Pte Ltd. 1984.

Bernardo, J. M. and Smith, A. F. M. *Bayesian Theory*. London, UK: John Wiley & Sons LTD. 1994.

Boudali, H. and Dugan, J. B. A discrete-time Bayesian network reliability modeling and analysis framework. *Reliability Engineering & System Safety*, v.87, p.337-349. 2005.

Boudali, H. and Dugan, J. B. A continuous-time Bayesian network reliability modeling, and analysis framework. *IEEE Transactions on Reliability*, v.55, n.1, p.86-97. 2006.

Burges, C. J. C. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, v.2, p.121-167. 1998.

Camarinopoulos, L. and Obrowski, W. Consideration of tolerable down times in the analysis of technical systems. *Nuclear Engineering and Design*, v.64, p.185-194. 1981.

Castanier, B., Bérenguer, C. and Grall, A. A sequential condition-based repair/replacement policy with non-periodic inspections for a system subject to continuous wear. *Applied Stochastic Models in Business and Industry*, v.19, n.4, p.327-247. 2003.

Chandra, V. and Kumar, K. V. Reliability and safety analysis of fault tolerant and fail safe node for use in a railway signalling system. *Reliability Engineering and System Safety*, v.57, p.177-183. 1997.

Chang, Y. H. J. and Mosleh, A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 2: IDAC performance influencing factors model. *Reliability Engineering & System Safety*, v.92, p.1014-1040. 2007.

Chen, D. and Trivedi, K. S. Optimization for condition-based maintenance with semi-Markov decision process. *Reliability Engineering & System Safety*, v.90, p.25-29. 2005.

Corradi, G., Janssen, J. and Manca, R. Numerical Treatment of Homogeneous semi Markov Processes in Transient Case - a Straightforward Approach. *Methodology and Computing in Applied Probability*, v.6, p.233-246. 2004.

Csenki, A. *Dependability for Systems with a Partitioned State Space (Markov and semi-Markov Theory and Computational Implementation)*. Berlin: Spring Verlag. 1994.

Cuomo, S., D'amore, L., Murli, A. and Rizzardi, M. Computation of the inverse Laplace transform based on a collocation method which uses only real values. *Journal of Computational and Applied Mathematics*, v.198, p.98-115. 2007.

D'amico, G., Janssen, J. and Manca, R. Homogeneous Semi-Markov reliability models for credit risk management. *Decisions in Economics and Finance*, v.28, p.79-93. 2005.

Deb, K. *Evolutionary algorithms for multicriterion optimization in engineering design*. Evolutionary Algorithms in Engineering and Computer Science (EUROGEN'99). University of Jyväskylä, Jyväskylä, Finland. May 30 - June 3, 1999. p.

Droguett, E. L., Moura, M. C., Jacinto, C. M. and Jr., M. F. S. A semi-Markov model with Bayesian belief network based human error probability for availability assessment of downhole optical monitoring systems. *Simulation Modelling Practice and Theory*, v.16, n.10, p.1713–1727. 2008.

Droguett, E. L., Moura, M. C., Jacinto, C. M. C., Jr., M. F. S. and Garcia, P. A. A. *A Semi-Markov Model with Bayesian Belief Network Based Human Reliability Modeling for Availability Assessment of Downhole Optical Monitoring Systems*. Esrel 2007-European Safety and Reliability Conference. Stavanger, Norway. 25-27 June, 2007. p.

El-Gohary, A. Bayesian estimations of parameters in a three state reliability semi-Markov models. *Applied Mathematics and Computation*, v.154, p.53-67. 2004.

Fei, S.-W., Wang, M.-J., Miao, Y.-B., Tu, J. and Liu, C.-L. Particle swarm optimization-based support vector machine for forecasting dissolved gases content in power transformer oil. *Energy Conversion and Management*, v.50, p.1604-1609. 2009.

Feller, W. On Semi-Markov Processes. In: (Ed.). *Mathematics*. Princeton: Princeton University, v.51, 1964. On Semi-Markov Processes

Firmino, P. R. A. *Bayesian Networks for parametrization of reliability in complex systems (in Portuguese)*. Department of Production Engineering, Center of Technology and Geoscience, Federal University of Pernambuco, Recife, PE-Brazil, 2004.

Grabski, F. The reliability of an object with semi-Markov failure rate. *Applied Mathematics and Computation*, v.135, p.1-16. 2003.

Haidar, N. H. S. Recursive Pseudo-Inversion of the Laplace Transform on the Real Line. *Applied Mathematics and Computation*, v.84, p.213-220. 1997.

Hong, W.-C. and Pai, P.-F. Predicting engine reliability by support vector machines. *Int J Adv Manuf Technol*, v.28, p.154-161. 2006.

Howard, R. A. *Dynamic Probabilistic Systems v.II: Semi-Markov and Decision Processes*. Mineola, New York: Dover Publications, INC. 2007.

Janssen, J. and Limnios, N. *Semi-Markov Models and Applications* Springer. 1999.

Janssen, J. and Manca, R. Numerical solution of Non Homogeneous semi Markov processes in Transient Case. *Methodology and Computing in Applied Probability*, v.3, p.271-293. 2001.

Janssen, J. and Manca, R. *Applied semi-Markov Processes*. New York: Springer Science+Business Media, LLC. 2006.

Janssen, J. and Manca, R. *Semi-Markov Risk Models for Finance, Insurance and Reliability*. New York: Springer Science+Business Media, LLC. 2007.

Jardine, A. K. S., Lin, D. and Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, v.20, n.7, p.1483-1510. 2006.

Jenab, K. and Dhillon, B. S. Assessment of reversible multi-state k-out-of-n:G/F/Load-Sharing systems with flow-graph models. *Reliability Engineering and System Safety*, v.91, p.765-771. 2006.

Kijima, M. and Sumita, N. A useful generalization of renewal theory: counting process governed by non-negative Markovian increments. *Journal of Applied Probability*, v.23, p.71-88. 1986.

Kim, M. J. and Makis, V. Optimal maintenance policy for a multi-state deteriorating system with two types of failures under general repair. *Computers & Industrial Engineering*, v.doi: 10.1016/j.cie.2008.11.023. 2009.

Korb, K. B. and Nicholson, A. E. *Bayesian artificial intelligence*. Florida: Chapman & Hall/CRC. 2003.

Kryzhniy, V. V. High-resolution exponential analysis via regularized numerical inversion of Laplace transforms. *Journal of Computational Physics*, v.199, p.618-630. 2004.

Kullback, S. and Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.*, v.22, p.79-86. 1951.

Langseth, H. and Portinale, L. Bayesian networks in reliability. *Reliability Engineering and System Safety*, v.92, p.92-108. 2007.

Levitin, G. Reliability and performance analysis for fault-tolerant programs consisting of versions with different characteristics. *Reliability Engineering & System Safety*, v.86, p.75-81. 2004.

Levitin, G. Optimal structure of fault-tolerant software systems. *Reliability Engineering and System Safety*, v.89, p.286-295. 2005.

Levitin, G. Reliability and performance analysis of hardware–software systems with fault-tolerant software components. *Reliability Engineering & System Safety*, v.91, p.570-579. 2006.

Limnios, N. Dependability analysis of semi-Markov systems. *Reliability Engineering & System Safety*, v.55, p.203-207. 1997.

Limnios, N. and Oprisan, G. *Semi-Markov processes and reliability*. Boston: Birkhauser. 2001.

Limnios, N. and Ouhbi, B. Nonparametric estimation of some important indicators in reliability for semi-Markov processes. *Statistical Methodology*, v.3, p.341-350. 2006.

Littlewood, B., Popov, P. and Strigini, L. Assessing the reliability of diverse fault-tolerant software-based systems. *Safety Science*, v.40, p.781-796. 2002.

Love, C. E., Zhang, Z. G., Zitron, M. A. and Guo, R. A discrete semi-Markov decision model to determine the optimal repair/replacement policy under general repairs. *European Journal of Operational Research*, v.125, n.2, p.398-409. 2000.

Madan, B. B., Goševa-Popstojanova, K., Vaidyanathan, K. and Trivedi, K. S. A method for modeling and quantifying the security attributes of intrusion tolerant systems. *Performance Evaluation*, v.56, p.167-186. 2004.

Makis, V. and Jardine, A. K. S. A note on optimal replacement policy under general repair. *European Journal of Operational Research*, v.69, n.1, p.75-82. 1993.

Marseguerra, M., Zio, E. and Martorell, S. Basics of genetic algorithms optimization for RAMS applications. *Reliability Engineering and System Safety*, v.91, p.977–991. 2006.

Menêzes, R. C. *Uma metodologia para a Avaliação da Confiabilidade Humana em atividades de substituição de cadeias de isoladores em linhas de transmissão*. Engenharia de Produção, Centro de Tecnologia e Geociências, UFPE, Recife, PE-Brasil, 2005.

Michalewicz, Z. *Genetic algorithms + data structures = evolution programs*. Berlin: Springer. 1996.

Milovanovic, G. V. and Cvetkovic, A. S. Numerical Inversion of the Laplace Transform. *ELEC. ENERG.*, v.18, n.3, p.515-530. 2005.

Montani, S., Portinale, L., Bobbio, A. and Codetta-Raiteri, D. RADYBAN: A tool for reliability analysis of dynamic fault trees through conversion into dynamic Bayesian Networks. *Reliability Engineering & System Safety*, v.93, n.7, p.922-932. 2008.

Moura, M. C. *Semi Markov Processes and Bayesian Networks for reliability performance indicators assessment of fault tolerant systems (in Portuguese)*. (Master). Production Engineering Department - CTG, Federal University of Pernambuco, Recife, PE, Brazil, 2006. 125 p.

Moura, M. C. and Droguett, E. L. *A numerical inversion of Laplace transforms based method to solve the interval transition probabilities of non-homogeneous semi-Markov processes*. ICORD VI - International Conference on Operational Research for Development. Fortaleza, CE, Brazil. 29 a 31 de August, 2007. 228-238 p.

Moura, M. C. and Droguett, E. L. A Continuous-Time semi-Markov Bayesian Belief Network Model for Availability Measure Estimation of Fault Tolerant Systems. *Pesquisa Operacional*, v.28, n.2, p.355-375. 2008.

Moura, M. C. and Droguett, E. L. *A faster numerical procedure for solving non-homogeneous semi-Markov processes*. Esrel 2009. Prague, Czech Republic: Taylor & Francis. 7-10 September, 2009a. p.

Moura, M. C. and Droguett, E. L. A Laplace Transform Numerical-Based Solution of Continuous-Time Homogeneous semi-Markov Processes for Dynamic Availability Assessment (under review). *Reliability Engineering & System Safety*. 2009b.

Moura, M. C. and Droguett, E. L. Mathematical Formulation and Numerical Treatment Based on Transition Frequency Densities and Quadrature Methods for non-Homogeneous semi-Markov Processes *Reliability Engineering & System Safety*, v.94, p.342-349. 2009c.

Moura, M. C. and Droguett, E. L. Numerical Approach for Assessing System Dynamic Availability Via Continuous Time Homogeneous Semi-Markov Processes (in press). *Methodol Comput Appl Probab*. 2009d.

Moura, M. C., Firmino, P. R. A., Droguett, E. A. L. and Jacinto, C. M. C. *Optical Monitoring System Availability Optimization via Semi-Markov Processes and Genetic Algorithms*. 54th IEEE Annual Reliability and Maintainability Symposium-RAMS, 2008. Las Vegas, USA, 2008. p.

Moura, M. C., Lins, I. D., Firmino, P. R., Droguett, E. L. and Jacinto, C. M. *Semi-Markov decision processes for determining optimal multiobjective maintenance policies*. Esrel 2009. Prague, Czech Republic: Taylor & Francis, 2009. 7-10 September p.

Moustafa, M. S., Maksoud, E. Y. A. and Sadek, S. Optimal major and minimal maintenance policies for deteriorating systems. *Reliability Engineering & System Safety*, v.83, n.3, p.363-368. 2004.

Neil, M., Tailor, M., Marquez, D., Fenton, N. and Hearty, P. Modelling dependable systems using hybrid Bayesian networks. *Reliability Engineering & System Safety*, v.93, n.7, p.933-939. 2008.

Nelson, P. and Wang, S. Dynamic reliability via computational solution of generalized state-transition equations for entry-processes. *Reliability Engineering & System Safety*, v.92, p.1281-1293. 2007.

Noortwijk, J. M. V. A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, v.94, n.1, p.2-21. 2009.

Oliveira, E. A., Alvim, A. C. M. and Frutuoso E Melo, P. F. *A Queueing Model for the Reliability Analysis of a System Considering its Age as a Supplementary Variable*. Annals of the XI Meeting of Reactors Physic and Termohydraulic. Poços de Caldas, 1997. 337-342 p.

Oliveira, E. A., Alvim, A. C. M. and Frutuoso E Melo, P. F. Unavailability analysis of safety systems under aging by supplementary variables with imperfect repair. *Annals Nuclear Energy*, v.32, p.241-252. 2005.

Ouhbi, B. and Limnios, N. Reliability estimation of semi-Markov systems: a case study. *Reliability Engineering & System Safety*, v.58, p.201-204. 1997.

Ouhbi, B. and Limnios, N. Non-parametric estimation for semi-Markov processes based on its hazard rate. *Statist. Infer. Stoch. Processes*, v.2, n.2, p.151-173. 1999.

Ouhbi, B. and Limnios, N. The rate of occurrence of failures for semi-Markov processes and estimation. *Statistics & Probability Letters*, v.59, p.245–255. 2002.

Ouhbi, B. and Limnios, N. Nonparametric reliability estimation of semi-Markov processes. *Journal of Statistical Planning and Inference*, v.109, p.155-165. 2003.

Pai, P.-F. System reliability forecasting by support vector machines with genetic algorithms. *Mathematical and Computer Modelling*, v.43, p.262-274. 2006.

Pearl, J. *Causality, Reasoning, and Inference*. New York: Cambridge University Press. 2000.

Perman, M., Senegacnik, A. and Tuma, M. Semi-Markov Models with an Application to Power-Plant Reliability Analysis. *IEEE Transactions on Reliability*, v.46(4), p.526-532. 1997.

Pievatolo, A. and Valadè, I. UPS reliability analysis with non-exponential duration distribution. *Reliability Engineering and System Safety*, v.81, p.183-189. 2003.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. *Numerical Recipes in C++*. Cambridge: Cambridge University Press. 2002.

Ross, S. M. *Introduction to Probability Models*. Berkeley, California: Academic Press. 1997.

Shawe-Taylor, N. C. A. J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge, UK: Cambridge University Press. 2000.

Smola, A. J. and Scholkoff, B. A tutorial on support vector regression. *Statistics and Computing*, v.14, p.199-222. 2004.

Soszynska, J. Reliability evaluation of a port oil transportation system in variable operation conditions. *International Journal of Pressure Vessels and Piping*, v.83, p.304-311. 2006.

Valkó, P. P. and Abate, J. Comparison of Sequence Accelerators for the Gaver Method of Numerical Laplace Transform Inversion. *An International Journal Computers & Mathematic with applications*, v.48, p.629-636. 2004.

Vaurio, J. K. Reliability characteristics of components and systems with tolerable repair times. *Reliability Engineering and System Safety*, v.56, p.43-52. 1997.

Weber, P., Munteanu, P. and Jouffe, L. *Dynamic Bayesian networks modelling the dependability of systems with degradations and exogenous constraints*. 11th IFAC Symposium on Information Control Problems in Manufacturing, INCOM'04. Salvador-Bahia, Brazil. April 5-7th, 2004. p.

Xie, W., Hong, Y. and Trivedi, K. Analysis of a two-level software rejuvenation policy. *Reliability Engineering & System Safety*, v.87, p.13-22. 2005.

Yañez, M., Joglar, F. and Modarres, M. Generalized renewal process for analysis of repairable systems with limited failure experience. *Reliability Engineering & System Safety*, v.77, p.167-180. 2002.

ATTACHMENTS

ATTACHMENT A[1]

versão impressa ISSN 0101-7438 / versão online ISSN 1678-5142

A CONTINUOUS-TIME SEMI-MARKOV BAYESIAN BELIEF NETWORK MODEL FOR AVAILABILITY MEASURE ESTIMATION OF FAULT TOLERANT SYSTEMS

Márcio das Chagas Moura*
Enrique López Droguett
Departamento de Engenharia de Produção
Universidade Federal de Pernambuco (UFPE)
Recife – PE
marciocmoura@gmail.com
ealopez@ufpe.br

* *Corresponding author / autor para quem as correspondências devem ser encaminhadas*

Recebido em 11/2007; aceito em 06/2008
Received November 2007; accepted June 2008

Abstract

In this work it is proposed a model for the assessment of availability measure of fault tolerant systems based on the integration of continuous time semi-Markov processes and Bayesian belief networks. This integration results in a hybrid stochastic model that is able to represent the dynamic characteristics of a system as well as to deal with cause-effect relationships among external factors such as environmental and operational conditions. The hybrid model also allows for uncertainty propagation on the system availability. It is also proposed a numerical procedure for the solution of the state probability equations of semi-Markov processes described in terms of transition rates. The numerical procedure is based on the application of Laplace transforms that are inverted by the Gauss quadrature method known as Gauss Legendre. The hybrid model and numerical procedure are illustrated by means of an example of application in the context of fault tolerant systems.

Keywords: semi-Markov processes; Bayesian belief networks; Laplace transforms; availability measure; fault tolerant systems.

Resumo

Neste trabalho, é proposto um modelo baseado na integração entre processos semi-Markovianos e redes Bayesianas para avaliação da disponibilidade de sistemas tolerantes à falha. Esta integração resulta em um modelo estocástico híbrido o qual é capaz de representar as características dinâmicas de um sistema assim como tratar as relações de causa e efeito entre fatores externos tais como condições ambientais e operacionais. Além disso, o modelo híbrido permite avaliar a propagação de incerteza sobre a disponibilidade do sistema. É também proposto um procedimento numérico para a solução das equações de probabilidade de estado de processos semi-Markovianos descritos por taxas de transição. Tal procedimento numérico é baseado na aplicação de transformadas de Laplace que são invertidas pelo método de quadratura Gaussiana conhecido como Gauss Legendre. O modelo híbrido e procedimento numérico são ilustrados por meio de um exemplo de aplicação no contexto de sistemas tolerantes à falha.

Palavras-chave: processos semi-Markovianos; redes Bayesianas; transformadas de Laplace; disponibilidade; sistemas tolerantes à falha.

Pesquisa Operacional, v.28, n.2, p.355-375, Maio a Agosto de 2008 355

[1] This paper is published at the journal *Pesquisa Operacional*.

ATTACHMENT B[1]

A Semi-Markov Model with Bayesian Belief Network Based Human Reliability Modeling for Availability Assessment of Downhole Optical Monitoring Systems

E.L. Droguett, M.C. Moura
Universidade Federal de Pernambuco - Brazil

C.M.C. Jacinto, M.F. Silva Jr. & P.A.A. Garcia
Petrobras - Brazil

ABSTRACT: In this article, it is proposed a hybrid model for the availability assessment of downhole optical monitoring systems where the system dynamics is modeled according to a non-homogeneous continuous time semi-Markov process and the human reliability for the installation and reinstallation procedures are given by Bayesian belief network models. Insights regarding the numerical solution of such hybrid model are also provided. The proposed modeling approach is illustrated by means of an application to a real case scenario in brown onshore fields in Brazil.

1 INTRODUCTION

1.1 *Optical monitoring system (OMS)*

Mostly because of the increasing oil price, a considerable attention has been given to the improvement of production technologies that allow for anticipation of oil production volumes and an improved reservoir management and control. In line with such efforts, recent developments in the context of reservoir management and control have led to the development and installation of temperature-pressure optical monitoring systems (OMS) for downhole applications.

Given the limited experience with this type of systems, the availability assessment is usually performed under a considerable level of uncertainty. Despite that scenario, this limited experience has suggested that an OMS is comprised of components that are renewed after failures as well as components that are under deteriorating processes with failure probabilities that are dependent on the total system age.

Furthermore, when brown (mature) fields are the focus of interest, cost is a relevant variable. In this context, the optical cable usually is responsible for approximately 80% of the total cost of an OMS for brown field applications. Upon a well failure, the OMS must be pull out of the whole even when no failures has been observed in any of the OMS components.

During these installation and reinstallation procedures, human performance is a relevant factor influencing the optical monitoring system life and the optical cable in particular.

Moreover, the time to accomplish a repair is determinant. It is considered that there is a tolerable maximum downtime for which the repair must be completed.

Therefore, there are three relevant aspects in estimating the system availability: (i) the system deteriorating process; (ii) the repair operator's capacity in returning an OMS to its normal operational condition; (iii) the available time to complete a repair action. In this context, this paper proposes a model for availability assessment that is able to handle the combined impact of failure process and the human error during the execution of repair activities.

1.2 *A hybrid model for availability analysis of OMS*

The proposed model combines the use of non-Homogeneous continuous time semi-Markov processes (NHSMP) and Bayesian Belief Networks (BBN) for human reliability estimation.

NHSMP will be used here in order to model the system dynamics because:

(i) Although the components of an OMS might be renewed upon a well failure, the duration (sojourn time) in a state influences the availability of an OMS and

(ii) Provided that some components might be under deteriorating processes, it should be considered time dependent transition probabilities.

The cause-effect relationships characterizing the human error probabilities (HEP) during installation and reinstallation procedures will be qualitatively and quantitatively modeled via BBNs according to a methodology proposed by (Droguett et al., 2006) and (Menêzes & Droguett, 2007).

[1] This article is published in the proceedings of Esrel (*Safety and Reliability Conference*) 2007, Stavanger, Norway.

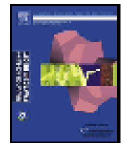
ATTACHMENT C[1]

Simulation Modelling Practice and Theory 16 (2008) 1713–1727

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat



A semi-Markov model with Bayesian belief network based human error probability for availability assessment of downhole optical monitoring systems

Enrique López Droguett [a], Márcio das Chagas Moura [a,*], Carlos Magno Jacinto [b], Manoel Feliciano Silva Jr. [b]

[a] *Department of Production Engineering, Federal University of Pernambuco Rua Acadêmico Hélio Ramos, S/N, Cidade Universitária, 50740-530, Recife-PE, Brazil*
[b] *Petrobras – Centro de Pesquisas e Desenvolvimento (CENPES), Ilha do Fundão, Cidade Universitária, 21941-598, Rio de Janeiro-RJ, Brazil*

ARTICLE INFO

Article history:
Received 23 April 2008
Received in revised form 25 July 2008
Accepted 27 August 2008
Available online 6 September 2008

Keywords:
Semi-Markov process
Bayesian belief network
Human error probability
Availability assessment

ABSTRACT

Compelled by increasing oil prices, a research effort is underway for designing and implementing intelligent oil fields in Brazil, with a first pilot directed towards mature wells in the Northeast. One of the major benefits of this technology is the anticipation of oil production volumes and an improved reservoir management and control. Given the considerable steep investment on the new technology, availability is a key attribute: higher availability means higher production volumes. An important part of this effort is the development of pressure–temperature optical monitoring systems (OMS) and their availability assessment. Availability analysis of an OMS impose some complexities, where the most relevant aspects are: (i) the system is under a deteriorating process; (ii) the available time to complete the maintenance; and (iii) human error probability (HEP) during maintenance that is influenced by the available time and other factors (e.g., experience, fatigue) in returning an OMS to its normal operational condition. In this paper we present a first attempt to solve this problem. It is developed an availability assessment model in which the system dynamics is described via a continuous-time semi-Markovian process specified in terms of probabilities. This model is integrated with a Bayesian belief network characterizing the cause-effect relationships among factors influencing the repairman error probability during maintenance. The model is applied to a real case concerning mature oil wells.

1. Introduction

Oil has been the most important source of energy since the early days of last century. The growing and continuous demand for energy associated with decreasing availability of this limited resource have led to a considerable increase in investment directed towards the development of alternative energy sources as well as to research efforts for optimizing technologies related to the exploration and production of oil.

Mostly because of the increasing oil price, a considerable attention has been given to the enhancement of production technologies that allow for anticipation of oil production volumes and an improved reservoir management and control. In line with such efforts, recent developments have led to the so called intelligent oil fields. The term 'intelligent' means: (i) data acquisition: sensors provide data on important well parameters in real time; (ii) flow remote control: it allows an operator to modify production or injection flow characteristics with no on-site intervention; (iii) data interpretation and

* Corresponding author. Tel.: +55 81 86477279; fax: +55 81 3271 8728/30.
E-mail address: marciocmoura@gmail.com (M. das Chagas Moura).

1569-190X/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.simpat.2008.08.011

[1] This paper is published at the journal *Simullation Modelling Practice and Theory*.

ATTACHMENT D[1]

Methodol Comput Appl Probab
DOI 10.1007/s11009-008-9114-2

Numerical Approach for Assessing System Dynamic Availability Via Continuous Time Homogeneous Semi-Markov Processes

Márcio das Chagas Moura · Enrique López Droguett

Received: 11 March 2008 / Revised: 5 September 2008 /
Accepted: 10 November 2008

Abstract Continuous-time homogeneous semi-Markov processes (CTHSMP) are important stochastic tools to model reliability measures for systems whose future behavior is dependent on the current and next states occupied by the process as well as on sojourn times in these states. A method to solve the interval transition probabilities of CTHSMP consists of directly applying any general quadrature method to the N^2 coupled integral equations which describe the future behavior of a CTHSMP, where N is the number of states. However, the major drawback of this approach is its considerable computational effort. In this work, it is proposed a new more efficient numerical approach for CTHSMPs described through either transition probabilities or transition rates. Rather than N^2 coupled integral equations, the approach consists of solving only N coupled integral equations and N straightforward integrations. Two examples in the context of availability assessment are presented in order to validate the effectiveness of this method against the comparison with the results provided by the classical and Monte Carlo approaches. From these examples, it is shown that the proposed approach is significantly less time-consuming and has accuracy comparable to the method of N^2 computational effort.

Keywords Homogeneous semi-markov processes · Integral equations · Quadrature methods · Availability assessment · Reliability

AMS 2000 Subject Classification 60K10 · 60K15

1 Introduction

A homogeneous semi-Markov process (HSMP) can be understood as a probabilistic model for which the future behavior is dependent on the sojourn times that in turn are random

M. d. C. Moura (✉) · E. L. Droguett
Department of Production Engineering, Federal University of Pernambuco,
Rua Acadêmico Hélio Ramos, s/n, Cidade Universitária, 50740-530 Recife, PE, Brazil
e-mail: marciocmoura@gmail.com

Published online: 13 December 2008 Springer

[1] This paper is published at the journal *Methodology Computing & Applied Probability*.

ATTACHMENT E[1]

A LAPLACE TRANSFORM NUMERICAL BASED SOLUTION OF CONTINUOUS TIME HOMOGENEOUS SEMI-MARKOV PROCESSES FOR DYNAMIC AVAILABILITY ASSESSMENT

Márcio das Chagas Moura*, Enrique López Droguett

Department of Production Engineering, Federal University of Pernambuco, Rua Acadêmico Hélio Ramos, s/n, Cidade Universitária, 50740-530 Recife, PE, Brazil

Abstract

Continuous-time homogeneous semi-Markov processes (CTHSMP) are important stochastic tools to model reliability measures for systems whose future behavior is dependent on the current and next states occupied by the process as well as on the sojourn times. A method (called *2N*-method) for solving the state probabilities of CTHSMP consists of directly applying a general quadrature method to *N* coupled integral equations and making *N* straightforward integrations, where *N* is the number of states. However, a drawback of this approach is the need to know previously the number *M* of steps enough to guarantee the algorithmic convergence. Therefore, it is proposed a numerical approach for CTHSMP described through either transition probabilities or transition rates. This numerical procedure is based on the application of Laplace transforms that are inverted by the Gauss quadrature method known as Gauss Legendre to obtain the state probabilities on the time domain. The main advantage of this approach is that it is not required adjusting dynamically the number of steps in order to obtain the desired convergence. There is a pre-set number of steps independent on the problem to be solved and thus, this method is likely to have a considerable reduced computational effort. Its effectiveness will be compared against the results provided by the *2N*- and Monte Carlo methods by using two examples in the context of reliability assessment. From these examples, it is showed that the Laplace-based approach is significantly less time-consuming and has accuracy comparable to the *2N*-method.

Keywords: Semi-Markov Process; Laplace Transforms; Gauss Quadrature; Availability Assessment; Reliability.

* Corresponding author. Fax: +55-81-2126-8728x30
E-mail address: marciocmoura@gmail.com (M. C. Moura)

[1] This article was submitted for possible publication at the journal RESS (*Reliability Engineering and System Safety*).

ATTACHMENT F[1]

ICORD VI
INTERNATIONAL CONFERENCE ON OPERATIONAL RESEARCH FOR DEVELOPMENT

29 a 31/08/07 Fortaleza, CE

074 - A numerical inversion of Laplace transforms based method to solve the interval transition probabilities of non-homogeneous semi-Markov processes

Márcio das Chagas Moura
Federal University of Pernambuco
Rua Acadêmico Hélio Ramos, s/n, Cidade Universitária, 50740-530 Recife, PE, Brazil
marciocmoura@gmail.com

Enrique López Droguett
Federal University of Pernambuco
Rua Acadêmico Hélio Ramos, s/n, Cidade Universitária, 50740-530 Recife, PE, Brazil
ealopez@uol.com.br

ABSTRACT

A non-homogeneous semi-Markov process (in short NHSMP) can be thought as a stochastic model in which the future behavior is dependent on the current and next states occupied by the process as well as on sojourn and process times. In this work, it is proposed a numerical approach to solve the interval transition probabilities of NHSMP described by transition rates. The numerical procedure is based on the application of Laplace transforms that are inverted by the Gauss Quadrature method known as Gauss Legendre. An example in the context of reliability assessment is presented in order to validate the effectiveness of the proposed method through the comparison with the results provided by Monte Carlo simulation.

KEYWORDS. Non-homogeneous semi-Markov processes. Numerical Inversion of Laplace transforms. Reliability.

[1] This article is published in the proceedings of ICORD (*International Conference of Operational Research and Development*) VI.

ATTACHMENT G[1]

Reliability Engineering and System Safety 94 (2009) 342–349

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Mathematical formulation and numerical treatment based on transition frequency densities and quadrature methods for non-homogeneous semi-Markov processes

Márcio das Chagas Moura *, Enrique López Droguett

Department of Production Engineering, Federal University of Pernambuco, Rua Acadêmico Hélio Ramos, S/N, Cidade Universitária, 50740-530 Recife, PE, Brazil

ARTICLE INFO

Article history:
Received 22 May 2007
Received in revised form
6 March 2008
Accepted 29 March 2008
Available online 11 April 2008

Keywords:
Non-homogeneous semi-Markov process
Transition frequency densities
Quadrature methods
Reliability

ABSTRACT

Non-homogeneous semi-Markov processes (NHSMP) are important stochastic tools for modeling reliability metrics over time for systems where the future behavior depends on the current and next states as well as on sojourn and process times. The classical method to solve the interval transition probabilities of NHSMPs consists of directly applying any general quadrature method to some non-convolution integral equations. However, this approach has a considerable computational effort. Namely, N^2-coupled integral equations with two variables must be solved, where N is the number of states. Therefore, this article proposes a more efficient mathematical formulation and numerical treatment, which are based on transition frequency densities and general quadrature methods respectively, for NHSMPs. The approach consists of only solving N-coupled integral equations with one variable and N straightforward integrations. Two examples in the context of reliability are also presented. The first one addresses a case where a semi-analytical solution is available. Then an example of application concerning pressure–temperature optical monitoring systems for oil wells is discussed. In both cases, the proposed approach is validated via the comparison against the results obtained from the semi-analytical solution (for the first example) as well as from both the classic and the Monte Carlo methods.

1. Introduction

A homogeneous semi-Markov process (HSMP) can be understood as a probabilistic model for which the future behavior is dependent on the sojourn times (x), which are random variables that depend on the current state and on the state to which the next transition will be done. According to Ouhbi and Limnios [1], HSMPs are more flexible models than ordinary Markov processes, as it is no longer required to assume that sojourn times are exponentially distributed.

Recent applications and theoretical developments on HSMPs have been proposed in the context of reliability engineering. For example, Perman et al. [2] applied a recursive procedure to approximate the interval transition probabilities, which are used to assess the future behavior of an HSMP over time. Closed formulas for reliability metrics are not available when the probability distributions of the sojourn time in a state are non-exponential. Limnios [3] proposes a discrete-time dependability analysis for HSMPs by using a method based on algebraic calculus. Ouhbi and Limnios [4] estimated reliability and availability through HSMPs of a turbo-generator rotor using a set of real data. Ouhbi and Limnios [5] proposed a statistical formula for assessing the rate of occurrence of failures (ROCOF) of HSMPs. Through this result, ROCOFs of the Markov and alternated renewal processes are also given as special cases.

In a non-homogeneous semi-Markov process (NHSMP), transitions between two states may depend not only on such states and on the sojourn times (x), but also on both times of the last (τ) and next (t) transitions, with $x = t - \tau$. The time variable τ is also known as the most recent arrival time or last entry time, and the time variable t is the calendar or process time. Thus, NHSMPs extend other stochastic processes such as HSMPs.

As a result, NHSMPs are powerful modeling tools, mainly in the reliability field (as exemplified in Janssen and Manca [6]). According to Becker et al. [7], NHSMPs are considered as approaches to model reliability characteristics of components or small systems with complex test and maintenance strategies. Janssen and Manca [8] argue, however, that the non-homogeneity in the continuous-time semi-Markov environment implies

* Corresponding author. Fax: +55 81 32718728x30

E-mail addresses: marciocmoura@gmail.com (M.C. Moura), ealopez@ufpe.br (E.L. Droguett).

0951-8320/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2008.03.032

[1] This article is published at the journal RESS (*Reliability Engineering and System Safety*).

ATTACHMENT H[1]

A faster numerical procedure for solving non-homogeneous semi-Markov processes

M. C. Moura & E. L. Droguett
Universidade Federal de Penambuco, Recife-PE, Brazil

ABSTRACT: Non-homogeneity implies higher difficulties on the continuous-time semi-Markov processes (CTNHSMP) environment. This gives rise as more intricate mathematical methods and related numerical solutions and is one of the main reasons behind the scarcity of CTNHSMP applications. Indeed, the classical method for solving CTNH

SMP is rather burdensome, consisting of directly applying a general quadrature method to N^2 coupled integral equations with two variables, where N is the number of states. Therefore, this article focuses on scrutinizing the effectiveness of a new and faster numerical treatment for CTNHSMP. Rather than computing N^2 integral equations, this approach consists of solving only N coupled integral equations with one variable and N straightforward integrations so that the high and inherent computational cost that plagues the solution of CTNHSMP is likely to be reduced. Comparisons against the results provided by the N^2- method and Monte Carlo will be performed.

1 INTRODUCTION

Continuous time non-homogeneous semi-Markov processes (CTNHSMP) are powerful modeling tools, especially in the reliability field (as exemplified in Janssen & Manca (2007)). According to Becker et al. (2000), CTNHSMP are considered as approaches to model reliability characteristics of components or small systems with complex test and maintenance strategies.

In a CTNHSMP, transitions between two states may depend not only on such states and on the sojourn times (x) (as it occurs with the homogeneous counterpart), but also on both times of the last (τ) and next (t) transitions, with $x = t - \tau$. The time variable τ is also known as the most recent arrival or last entry time, and t is the calendar or process time. Thus, CTNHSMP extend other models such as homogeneous semi-Markov, (non-) homogeneous ordinary Markov and other point stochastic processes.

In spite of that, there are two main reasons to explain the scarcity of CTNHSMP applications: (i) Janssen & Manca (2001) argue the non-homogeneity on the continuous time semi-Markov environment implies additional difficulties in treating CTNHSMP; (ii) in accordance with Nelson & Wang (2007), for practical applications, gathering of high level required data (transition probabilities and/or rates) is likely to be a significant challenge, mainly in the presence of censoring implied by preventive maintenance.

Specifically discussing the first cause, it gives rise as more intricate mathematical methods and numerical solutions. Indeed, the future behavior of a CTNHSMP is usually assessed through its interval transition probability equations which are comprised of a system of N^2 coupled integral equations with two variables, where N is the number of states.

The classical method for solving CTNHSMP is explained in Janssen & Manca (2001) and consists of directly applying a general quadrature method to these N^2 coupled integral equations, which are a generalization of the Kolmogorov backward differential equations of the Markov environment (see Feller (1964)). However, such an approach is rather cumbersome, with a computational cost greater than Monte Carlo (MC)-based algorithms.

Therefore, Moura & Droguett (2009) propose an alternative method for solving the probability equations of a CTNHSMP. The approach consists of casting the N^2 coupled integral equations into an initial value problem involving transition frequency densities, and then solve N coupled integral equations with one variable and N straightforward integrations.

The mathematical and numerical treatments developed by Moura & Droguett (2009) is put forward

[1] This article is accepted to be published in the proceedings of Esrel (*Safety and Reliability Conference*) 2009, Prague, Czech Republic.

ATTACHMENT I[1]

Optical Monitoring System Availability Optimization via Semi-Markov Processes and Genetic Algorithms

Márcio das Chagas Moura, M.Sc., Federal University of Pernambuco
Paulo Renato A. Firmino, M.Sc., Federal University of Pernambuco
Enrique López Droguett, Ph.D., Federal University of Pernambuco
Carlos Magno Jacinto, M.Sc., PETROBRAS-CENPES

Key Words: non-homogeneous semi-Markov processes, real-coded genetic algorithms, system mean availability, virtual age

SUMMARY & CONCLUSIONS

System availability optimization is one of the main issues to oil production managers: the greater the system availability the greater the production profits are. Provided that preventive maintenance actions promote rejuvenation impact on availability indicator, this paper proposes an approach to maximize the mean availability by identifying an optimal maintenance policy for downhole optical monitoring systems, which are modeled according to non-homogeneous semi-Markov processes. In order to solve the resulting optimization problem constrained by system performance costs, new real-coded GA operators are also presented. The proposed approach is exemplified by means of an application to a real scenario in onshore oil wells in Brazil.

1 INTRODUCTION

Mostly because of competitive reasons, a considerable attention is devoted to the development of optimal preventive maintenance policies that aim to maximize the system mean availability. In fact, these policies are influenced by system's stochastic failure and repair processes and have both cost and technological constraints.

In the context of oil production, recent breakthroughs for reservoir management and control have led to the development and installation of temperature-pressure optical monitoring systems (OMS) for downhole applications. Due to the huge variety and aggressiveness of the downhole operating environment, an OMS has components that upon a failure condition might either be completely renewed or have intensity functions dependent on the process (global) time.

Moreover, some of these components exhibit non-exponential times to failure. In this context, this article aims to develop an optimal preventive maintenance policy that maximizes the mean availability by using genetic algorithms for downhole OMSs based upon continuous time non-homogeneous semi-Markov processes (NHSMP).

In order to model the system dynamics, NHSMPs are employed because: (i) although the components of an OMS might be renewed upon a well failure, the duration (sojourn time) in a state influences the availability of an OMS and (ii) provided that some components might be under deteriorating processes, it should be considered time-dependent intensity functions.

Through genetic algorithms, it is established a preventive maintenance policy that maximizes the system's mean availability restricted to technological and cost constraints. This optimal policy is comprised of operating times t_j up to preventive actions, which have a rejuvenation impact q on the real age of the system. This parameter is incorporated into the state equations of the NHSMP so that the effectiveness of each preventive maintenance is taken into account in the optimization procedure of the OMS's availability.

The mathematical programming problem relevant to the system is summarized as follows:

$$\text{Maximize } \overline{A}\,[T \mid (t_1, t_2, \ldots, t_n); q] \text{ subject to} \quad (1)$$
$$n \le N \quad (2)$$
$$\mathrm{C}[T \mid (t_1, t_2, \ldots, t_n); q; c_p; c_c] \le K \quad (3)$$
$$c_p, c_c, T, K > 0,\ t_i \in (0, T],\ N, n \in \mathbb{N}, \text{ and } q \in \Re, \quad (4)$$

where T is the mission time under consideration; t_j is the operating time up to the j^{th} preventive maintenance action, with $t_0 = 0$; $(t_1, t_2, \ldots, t_n)$ composes a preventive maintenance policy; n is the number of preventive maintenance events in T and N is its upper bound; $\overline{A}\,[T \mid (t_1, t_2, \ldots, t_n); q]$ is the system mean availability in T modeled in terms of an NHSMP and related to $(t_1, t_2, \ldots, t_n)$ and q; $\mathrm{C}[T \mid (t_1, t_2, \ldots, t_n), q, c_p, c_c]$ is the cost related to the system performance in T given the maintenance policy $(t_1, t_2, \ldots, t_n)$, q, the cost per time unit to perform preventive (c_p) and corrective (c_c) maintenances.

In order to compute $\mathrm{C}[T \mid (t_1, t_2, \ldots, t_n); q, c_p, c_c]$, the time spent by the system under preventive and corrective maintenances are estimated as a function of $\overline{A}\,[T \mid (t_1, t_2, \ldots, t_n). q]$ and T_p (the average preventive maintenance time) [see eq. (5)]. Finally, K is a maximal cost constraint, i.e., the total cost incurred by performing corrective and preventive maintenance actions.

$$\mathrm{C}\left[\mathrm{T} \mid (t_1,\ldots,t_n); q; c_p; c_c\right] = n \cdot c_p \cdot T_p + \left[T - T \cdot \overline{A}\left[(t_1,\ldots,t_n); q\right] - T_p\right] \cdot c_c \quad (5)$$

[1] This article was published in the proceedings of IEEE RAMS 2008 (*Annual Reliability and Maintanability Symposium*), Las Vegas, USA.

ATTACHMENT J[1]

Semi-Markov decision processes for determining optimal multiobjective maintenance policies

M.C. Moura, I.D. Lins, P.R. Firmino & E.L. Droguett
Department of Production Engineering, Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil
C.M. Jacinto
PETROBRAS-CENPES, Rio de Janeiro, Brazil

ABSTRACT: Downhole pressure-temperature optical sensors have been developed to improve the management and control of oil reservoirs. One of their aims is to decrease the number and impact of intrusive maintenance interventions since the (re)installation procedures are human intensive and might influence the life of the monitored systems. Therefore, maintenance policies that jointly optimize mean availability and expected cost rate associated with maintenance interventions on monitored systems are a must in oil industries. This paper proposes a multiobjective optimization model based on semi-Markov decision processes to find a set of nondominated maintenance policies. Each obtained policy is of threshold type and it represents the optimal decision (do-nothing, minimal maintenance or replacement) whenever the system enters a new deterioration state. Insights of a multiobjective genetic algorithm and of an exhaustive method regarding numerical search for solutions are provided. The model is applied to a situation involving onshore brown fields in Brazil.

1 INTRODUCTION

Compelled by increasing oil prices, a research effort is underway for designing and implementing intelligent oil fields in Brazil, with a first pilot directed towards mature wells in the Northeast. One of the major benefits of this technology is the anticipation of oil production volumes and an improved reservoir management and control, including maintenance decisions. Given the considerable steep investment on new technologies, system availability is a key attribute: higher availability means higher production volumes.

The term “intelligent” means: (i) data acquisition: sensors provide data on important well parameters on real time; (ii) flow remote control: it allows an operator to modify production or injection flow characteristics with no on-site intervention; (iii) data interpretation and optimization: it allows production and reservoir engineers feed simulation models and act on a particular well on real time. Therefore, intelligent oil field is a concept encompassing various technologies that allow for an integrated management of production and injection of one or several reservoirs.

An important part of this technology is the pressure-temperature optical monitoring systems (OMS). They are responsible for acquiring physical data (pressure and temperature) from the downhole environment of oil industries. OMS availability analysis, mainly of the sensor component, is already done by Droguett et al. (2008). Moreover, Moura et al. (2008) developed optimal time maintenance policies for OMS systems.

Data collected from OMS might be used in a pattern recognition technique to indicate at which deterioration state the system is. Given that, adequate actions should be taken so that the number of interventions is minimized.

Due to the complexity of systems from oil industry, these interventions are intrusive, highly human-intensive and cost-consuming and thus minimizing them means decreasing the impact of human performance on the system and related costs as well.

Preventive actions, which set a periodic interval to perform planned maintenances, ignore the health status of a physical equipment/system. Therefore, they may not be adequate to oil industry systems since sometimes they would imply unnecessary actions, *i.e.*, as system did not cross the critical deterioration line yet. On the other hand, pre-set times for preventive actions might also not pay enough attention on the system, even if a latent failure will take place next. Both situations are cost and time consuming and should be attenuated.

In this way, according to Jardine et al. (2006) more efficient maintenance approaches such as condition-based maintenance (CBM) may be implemented to handle this situation. CBM is a maintenance program that recommends maintenance actions based on the information collected through condition monitoring (OMS, for instance). CBM attempts to

[1] This article is accepted to be published in the proceedings of Esrel (*Safety and Reliability Conference*) 2009, Prague, Czech Republic.

APPENDIX A

Table A - 1 – CPT of the MTTF given the level of paraffin (PARAF) and the percentage of H_2O and solids (BWSOT)

PARAF, BWSOT		P(MTTF (h) | PARAF, BWSOT)			
		100	200	1000.0	10000.0
0	0	0.05	0.10	0.15	0.70
0	1	0.15	0.15	0.30	0.40
1	0	0.15	0.20	0.40	0.25
1	1	0.20	0.50	0.15	0.15

Table A - 2 – CPT of the percentage of H_2O and solids (BWSOT) given the level of paraffin (PARAF)

PARAF	P(BWSOT | PARAF)	
	0	1
0	0.80	0.20
1	0.40	0.60

Table A - 3 – CPT of the level of paraffin (PARAF) given the classification of the filter installed (FILTER)

FILTER	P(PARAF | FILTER)	
	0	1
0	0.75	0.25
1	0.45	0.55

Table A - 4 – CPT of the classification of the filter installed (FILTER) given the depth of the pump (DEPTH_PUMP)

DEPTH_PUMP	P(FILTER | DEPTH_PUMP)	
	0	1
0	0.90	0.10
1	0.60	0.40

Table A - 5 – CPT of the depth of the pump (DEPTH_PUMP)

Variable	P(DEPTH_PUMP)	
	0	1
DEPTH_PUMP	0.70	0.30

APPENDIX B

Table B - 1 – CTP of the repairman's capacity given attention, experience and skill

Attention, Experience, Skill			P(Repairman's Capacity | Attention, Experience, Skill)	
			0	1
0	0	0	1.0	0.0
0	0	1	0.85	0.15
0	1	0	0.65	0.35
0	1	1	0.25	0.75
1	0	0	0.75	0.25
1	0	1	0.45	0.55
1	1	0	0.35	0.65
1	1	1	0.0	1.0

Table B - 2 – CTP of the repairman's attention given emotional state and fatigue

Emotional State, Fatigue		P(Repairman's Attention | Emotional State, Fatigue)	
		0	1
0	0	0.95	0.05
0	1	0.55	0.45
1	0	0.35	0.65
1	1	0.15	0.85

Table B - 3 – CTP of the repairman's fatigue given workload and external factors

Workload, External Factors		P(Repairman's Fatigue | Workload, External Factors)	
		0	1
0	0	0.95	0.05
0	1	0.75	0.25
1	0	0.55	0.45
1	1	0.15	0.85

Table B - 4 – CTP of the external factors given climatic conditions and distracter agent

Climatic conditions, distracter agents		P(External factors | Climatic conditions, Distracter agents)	
		0	1
0	0	0.85	0.15
0	1	0.55	0.45
1	0	0.75	0.25
1	1	0.45	0.55

Table B - 5 – CTP of the repairman given capacity and time available to complete reinstallation

Repairman's capacity, Available time		P(Repairman | Repairman's capacity, Time available)	
		0	1
0	0	0.90	0.10

0	1	0.75	0.25
1	0	0.55	0.45
1	1	0.20	0.80

Table B - 6 – CTP of the root nodes: Emotional state, Workload, Climatic Conditions, Distracter Agents, Experience, Skill, Available time

Node	**P(Root Nodes)**	
	0	1
Emotional State	0.45	0.55
Workload	0.85	0.15
Climatic Conditions	0.85	0.15
Distracting Agents	0.05	0.95
Experience	0.85	0.15
Skill	0.75	0.25
Available time to complete reinstallation	0.75	0.25