Improving Risk Management

RIESGOS-CM

Análisis, Gestión y Aplicaciones

P2009/ESP-1685

UNION EUROPEA
FONDO SOCIAL EUROPEO

La Suma de Todos
Comunidad de Madrid
www.madrid.org

Technical Report 2010.03

Bayesian Reliability, Availability and Maintainability Analysis for Hardware Systems Described Through Continuous Time Markov Chains

Javier Cano, Javier Moguerza, David Rios Insua

http://www.analisisderiesgos.org

Universidad Carlos III de Madrid

Universidad Rey Juan Carlos

Bayesian Reliability, Availability and Maintainability Analysis for Hardware Systems Described Through Continuous Time Markov Chains

Javier Cano and **Javier M. Moguerza**

Rey Juan Carlos University
Department of Statistics and
Operations Research
Camino del Molino s/n
28943 Fuenlabrada (Madrid), Spain
(javier.cano@urjc.es)

David Ríos Insua

Royal Academy of Sciences
Valverde 22
28004 Madrid, Spain
(david.rios@urjc.es)

Abstract

Hardware systems are present in many fields of human activity. Markov models are sometimes used in hardware reliability, availability and maintainability (RAM) modeling. They are especially useful in situations in which the system we want to analyze may be modeled with several states through which the system evolves, some of them corresponding to ON states, the rest to OFF states. We provide here RAM analyses of such systems within a Bayesian framework, addressing both short-term and long-term performance.

Keywords: Bayesian analysis, Multinomial-Dirichlet model, Exponential-Gamma model, Markov Chain Monte Carlo, Phase-type distribution.

1 INTRODUCTION

There is a growing interest in Reliability, Availability and Maintainability (RAM) analyses of hardware (HW) systems, especially of safety critical ones, see Eusgeld et al. (2008); Xie, Dai, and Poh (2004); and Pukite and Pukite (1998) for extensive reviews. One of the ways to analyze such systems is through Continuous Time Markov Chains (CTMCs), which, in our context, are stochastic processes which evolve through a discrete set of states, some of them corresponding to ON configurations, the rest to OFF configurations, remaining at each state an exponential time. There are several ways to model a HW system through a CTMC, see Prowell and Poore (2004); or Xie et al. (2004) for various examples. Indeed, there are even several standard hardware configurations based on CTMC models, including the triple modular redundancy (TMR) system (Baier, Haverkort, Hermanns, and Katoen 2003), the dual-duplex system, or the all voting triple modular redundancy (AVTMR) system (Kim, Lee, and Lee 2005). These, and other, configurations are aimed at attaining very high system availability, say 99.999% of time, through transfer of workload when one or more system components fail, or through the inclusion of intermediate failure states with automated recovery.

The standard approach to RAM estimation of CTMC HW systems computes Maximum Likelihood Estimates (MLEs) for the involved parameters of the CTMC, substitutes parameters by the MLEs, computes the equilibrium distribution, and, consequently, estimates the long-term fraction of time that the system remains in ON and OFF configurations. However, this approach usually underrepresents uncertainty in the parameters, as discussed in Glynn (1986) or Berger and Ríos Insua (1998). Thus, we adopt here a Bayesian approach that fully acknowledges the uncertainty present and takes advantage of all information available. Moreover, we adopt a more short-term oriented approach.

To do so, in Section 2, we describe a general formulation of the problem, defining model parameters and specifying the data to proceed with inference and, more importantly, forecasting tasks. In Section 3, we briefly review how to estimate the parameters of a CTMC, with the aid of standard Dirichlet-Multinomial and Exponential-Gamma models and compute the posterior equilibrium distribution, which we shall use later on. In Sections 4 and 5, we show how to estimate, respectively, the reliability and maintainability of HW systems. We then perform inference for system availability. We conclude with an example in Section 7, followed by some discussion.

2 PROBLEM FORMULATION

We are concerned with hardware systems, which we assume can be modeled through Continuous Time Markov Chains (CTMC). We consider that the chain evolves in a state space $E = \{1, 2, \ldots, m\}$. States $\{1, 2, \ldots, l\}$ correspond to operational (ON) configurations, whereas states $\{l + 1, \ldots, m\}$ correspond to OFF configurations. A transition from an ON to an OFF state describes a failure, whereas a transition from an OFF to an ON state represents a repair. We shall denote by X_t the state of the system at time t.

The behavior of a CTMC is characterized by the transition probabilities and the permanence rates. The transition probability matrix is an $m \times m$ matrix $\boldsymbol{P} = (p_{ij})$, where p_{ij} is the probability that a transition out of state i leads to state j. Hence, $\sum_j p_{ij} = 1$, $\forall i$, with $p_{ii} = 0$. Clearly, for physical or logical reasons, some additional p_{ij} matrix entries could be 0. The permanence rates are ν_i, $i = 1, \ldots, m$, where $1/\nu_i$ is the mean of the exponential random variable which models the time spent by the system in state i before leaving it. See Ross (2007) for a full description.

For a given initial ON state i, assume we may compute its conditional reliability $R_i(t|\boldsymbol{\nu}, \boldsymbol{P}) = \Pr\{T \geq t|\boldsymbol{\nu}, \boldsymbol{P}, X_0 = i\}$, i.e., the probability that the system will remain ON, for a time longer than t, where $\boldsymbol{\nu}$ is the vector of $\nu_i's$, and T

is the random variable which represents the time passed until system failure. Then, once we have computed the posterior $\pi(\boldsymbol{\nu}, \boldsymbol{P}|data)$, we shall be able to compute the posterior predictive reliability

$$R_i(t|data) \quad = \quad \iint R_i(t|\boldsymbol{\nu}, \boldsymbol{P})\pi(\boldsymbol{\nu}, \boldsymbol{P}|data)\, d\boldsymbol{\nu}\, d\boldsymbol{P}, \tag{1}$$

which will be one of our key computational objectives, dealt with in Section 4. Computation of maintainability will follow a similar scheme, interchanging the roles of ON and OFF states with respect to the strategy we shall adopt to estimate system reliability. Regarding the calculation of availability, we shall adopt here a short-term approach, computing the so-called interval availability, which is the expected proportion of time that the system will be ON within a given interval, provided it was ON initially.

3 INFERENCE FOR CTMC PARAMETERS

We shall briefly discuss first inference for the transition probabilities, permanence rates and equilibrium distribution of CTMC. We shall use these later on in our RAM computations. We assume that the data available are the times until state i is left, $t_{i1}, t_{i2}, \cdots, t_{in_i}$, for various states $i = 1, \ldots, m$, and the counts for the transitions, n_{ij}, from state i to state j, that is, $(n_{i1}, \ldots, n_{i,i-1}, n_{i,i+1}, \ldots, n_{im})$, with $i, j = 1, \ldots, m$. Clearly, n_{ij} will be zero if transitions from state i to state j are logically or physically impossible.

3.1 Estimating transition probabilities

We start with inference for the transition probabilities. Unless based on a specific model, in general we shall proceed as follows. For the $i-$th row, we assume that, a priori:

$$\begin{aligned}&(p_{i1}, \ldots, p_{i,i-1}, p_{i,i+1}, \ldots, p_{im})\\ &\qquad\qquad\qquad\qquad \sim \mathcal{D}(\delta_{i1}, \ldots, \delta_{i,i-1}, \delta_{i,i+1}, \ldots, \delta_{im}),\end{aligned}$$

where δ_{ij} are the parameters of a Dirichlet distribution, which will be zero if the corresponding p_{ij} are known to be zero. Then, the posterior distribution is

$$\begin{aligned}&(p_{i1}, \ldots,\ p_{i,i-1},\ p_{i,i+1}, \ldots,\ p_{im})|data\\ &\qquad \sim \mathcal{D}(\delta_{i1} + n_{i1}, \ldots,\ \delta_{i,i-1} + n_{i,i-1},\ \delta_{i,i+1} + n_{i,i+1}, \ldots,\ \delta_{im} + n_{im}),\end{aligned}$$

(French and Ríos Insua 2000). More sophisticated models which take into account possible row dependence may be seen in, e.g., Diaconis and Rolles (2006).

3.2 Estimating permanence rates

We proceed now with inference for the permanence rates, ν_i, $i = 1, \cdots, m$. We assume that the times $t_{i1}, t_{i2}, \cdots, t_{in_i}$ follow an exponential distribution, with parameter ν_i, that is,

$$t_{i1}, t_{i2}, \ldots, t_{in_i} \sim \mathcal{E}(\nu_i).$$

This is a reasonable assumption, as many hardware architectures around are based on CTMCs, see Xie et al. (2004) and references therein for abundant examples, but see as well our final discussion.

Assuming that the parameter ν_i follows a gamma prior distribution,

$$\nu_i \sim \mathcal{G}(\alpha_i, \beta_i),$$

then,

$$\nu_i | data \sim \mathcal{G}\left(\alpha_i + n_i, \beta_i + \sum_{j=1}^{n_i} t_{ij}\right),$$

see, e.g., French and Ríos Insua (2000).

3.3 Computing the posterior equilibrium distribution

We outline now how to compute the posterior equilibrium distribution of the CTMC, which we shall need later on. For fixed values of the p_{ij}'s and ν_i's, the equilibrium distribution $\{\pi_j\}_{j=1}^m$ is obtained, if it exists, by solving the system

$$\begin{aligned} \nu_j \pi_j = \sum_{i \neq j} r_{ij} \pi_i; \qquad & \forall j \in \{1, \ldots, m\}, \\ \sum_j \pi_j = 1; \qquad & \pi_j \geq 0, \end{aligned} \tag{2}$$

with $r_{ij} = \nu_i p_{ij}$. Sufficient conditions for the existence of the π_i's may be consulted, e.g., in Ross (2007, pg. 385). The π_j's may be interpreted as long-term time fractions that the system spends in various states.

When the posterior distributions of ν_i, p_{ij}, are, respectively, very peaked around certain values $\hat{\nu}_i, \hat{p}_{ij}$, say their posterior modes, we could substitute the parameters by their posterior modes,

$$\hat{\nu}_i = \frac{\alpha_i + n_i - 1}{\beta_i + \sum_{j=1}^{n_i} t_{ij}}; \ \hat{p}_{ij} = \frac{\delta_{ij} + n_{ij} - 1}{\sum_{l \neq i}(n_{il} + \delta_{il}) - m + 1}; \ \hat{r}_{ij} = \hat{\nu}_i \hat{p}_{ij},$$

and solve system (2), to obtain the appropriate solution $\{\hat{\pi}_i\}_{i=1}^m$. If this is not the case, we could obtain samples from the posteriors $\{\nu_i^\eta\}_{\eta=1}^N$, $\{p_{ij}^\eta\}_{\eta=1}^N$, and, consequently, obtain a sample $\{\pi_i^\eta\}_{\eta=1}^N$ through the repeated solution of (2). If needed, we could summarize such a sample by, e.g., its mean,

$$\hat{\pi}_i = \frac{1}{N} \sum_{\eta=1}^N \pi_i^\eta, \ i = 1, \ldots, m.$$

3.4 Estimating the intensity matrix

The r_{ij}'s in (2) are designated jumping intensities from state i to state j. In addition, we set $r_{ii} = -\sum_{j \neq i} r_{ij} = -\nu_i$, $i = 1, \ldots, m$, and place them together in the intensity matrix $\mathbf{\Lambda} = (r_{ij})$, also called the infinitesimal generator of the process, which shall have a key role in later computations.

Following a similar reasoning to that of Section 3.3, if the posterior distributions of ν_i, p_{ij} are very peaked around, e.g., their posterior modes $\hat{\nu}_i, \hat{p}_{ij}$, we could use $\hat{r}_{ij} = \hat{\nu}_i \hat{p}_{ij}$, $i \neq j$. For $i = j$, we set $\hat{r}_{ii} = -\hat{\nu}_i$, $i = 1, \ldots, m$. Otherwise, we may obtain samples $\{\nu_i^\eta\}_{\eta=1}^N$ and $\{p_{ij}^\eta\}_{\eta=1}^N$ from the posteriors above, and use the relationship $r_{ij} = \nu_i p_{ij}$ to obtain samples from the posterior $\{r_{ij}^\eta\}_{\eta=1}^N$, $i \neq j$. For $i = j$, we use the posterior sample $\{r_{ii}^\eta = -\nu_i^\eta\}_{\eta=1}^N$, $i = 1, \ldots, m$. If needed, we could summarize all samples appropriately, by, e.g., their means.

4 RELIABILITY FORECASTING WITH CTMCs

We describe now how to estimate reliabilities in our system, distinguishing the case in which we know the initial ON state, and that in which we only know that the system is initially ON.

For each ON state $i \in \{1, 2, \ldots, l\}$, we may compute its unconditional posterior reliability (1), which will be typically approximated through simulation

$$R_i(t|data) \simeq \frac{1}{N} \sum_{\eta=1}^{N} R_i(t|\boldsymbol{\nu}^\eta, \boldsymbol{P}^\eta),$$

for a sample $\{\boldsymbol{\nu}^\eta, \boldsymbol{P}^\eta\}_{\eta=1}^N$ from the above posteriors. If we only know that the system is initially ON, but we do not know which specific ON state the system is initially at, we could follow this approach. Let π_i be the posterior equilibrium probability for the i-th ON state, $i = 1, \ldots, l$, as computed from (2). Then, we could compute

$$R(t|\boldsymbol{\nu}, \boldsymbol{P}) = \frac{1}{\bar{\pi}} \sum_{i=1}^{l} \pi_i \; R_i(t|\boldsymbol{\nu}, \boldsymbol{P}),$$

where $\bar{\pi} = \sum_{j=1}^{l} \pi_j$, assuming the values of π_i, $\boldsymbol{\nu}$ and $\boldsymbol{P}$ are known. As we do not know them, we could approximate the posterior predictive reliability through Monte Carlo simulation

$$R(t|data) \simeq \frac{1}{N} \sum_{\eta=1}^{N} \sum_{i=1}^{l} \frac{\pi_i^\eta}{\bar{\pi}^\eta} \; R_i(t|\boldsymbol{\nu}^\eta, \boldsymbol{P}^\eta),$$

where $\bar{\pi}^\eta = \sum_{j=1}^{l} \pi_j^\eta$, and $\{\pi_i^\eta\}_{\eta=1}^N$ is a sample from the posterior calculated in Section 3.3, associated with $\{\boldsymbol{\nu}^\eta, \boldsymbol{P}^\eta\}_{\eta=1}^N$.

The key issue is then how to compute the reliabilities $R_i(t|\boldsymbol{\nu}, \boldsymbol{P})$, $i = 1, \ldots, l$. To do so, we adopt the strategy of subsuming all OFF states into an absorbing state which we designate a, conveniently redefining the transition probabilities. In Figure 1, we sketch such process for an arbitrary Markov chain with m states, the first l ON; the rest, OFF. Indeed, for our general CTMC, we divide the transition probability matrix into four blocks as follows:

$$\boldsymbol{P} = \left(\begin{array}{ccccc|cccc} 0 & p_{12} & p_{13} & \cdots & p_{1l} & p_{1,l+1} & \cdots & \cdots & p_{1m} \\ p_{21} & 0 & p_{23} & \cdots & p_{2l} & p_{2,l+1} & \cdots & \cdots & p_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ p_{l1} & p_{l2} & p_{l3} & \cdots & 0 & p_{l,l+1} & \cdots & \cdots & p_{lm} \\ \hline p_{l+1,1} & p_{l+1,2} & p_{l+1,3} & \cdots & p_{l+1,l} & 0 & \cdots & \cdots & p_{l+1,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ p_{m1} & p_{m2} & p_{m3} & \cdots & p_{ml} & p_{m,l+1} & \cdots & p_{m,m-1} & 0 \end{array} \right), \tag{3}$$

leading to a chain with probability matrix

$$\boldsymbol{P}_1 = \left(\begin{array}{ccccc|c} 0 & p_{12} & p_{13} & \cdots & p_{1l} & p_{1a} = \sum_{m>l} p_{1m} \\ p_{21} & 0 & p_{23} & \cdots & p_{2l} & p_{2a} = \sum_{m>l} p_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{l1} & p_{l2} & p_{l3} & \cdots & 0 & p_{la} = \sum_{m>l} p_{lm} \\ \hline 0 & 0 & 0 & \cdots & 0 & 1 \end{array}\right).$$

The intensity matrix of this chain will be designated

$$\boldsymbol{\Lambda}_1 = \left(\begin{array}{ccccc|c} -\nu_1 & r_{12} & r_{13} & \cdots & r_{1l} & \omega_1 \\ r_{21} & -\nu_2 & r_{23} & \cdots & r_{2l} & \omega_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{l1} & r_{l2} & r_{l3} & \cdots & -\nu_l & \omega_l \\ \hline 0 & 0 & 0 & \cdots & 0 & 0 \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{\Omega}_1 & \boldsymbol{\omega} \\ \hline \boldsymbol{0}^T & 0 \end{array}\right),$$

where $\omega_i = -\sum_{j=1}^{l} r_{ij}$, $i, j = 1, \ldots, l$, and $r_{ii} = -\nu_i$, $i = 1, \ldots, l$.

Note, now, that the time until failure coincides with the so-called time until absorption of our modified Markov chain, defined as

$$\tau = \inf\{t \geq 0 | X_t = a\}.$$

Consequently, τ has a phase-type distribution, given $\boldsymbol{\nu}, \boldsymbol{P}$,

$$\tau | \boldsymbol{\nu}, \boldsymbol{P} \sim PH(\boldsymbol{\pi}_{\text{ON}}^{(\mathbf{0})}, \boldsymbol{\Omega}_1 | \boldsymbol{\nu}, \boldsymbol{P}),$$

where $\boldsymbol{\pi}_{\text{ON}}^{(\mathbf{0})}$ is an initial state probability vector over ON states. From Bladt (2005), we may see that the distribution function F of τ, given $\boldsymbol{\nu}, \boldsymbol{P}$, is

$$F(t|\boldsymbol{\nu}, \boldsymbol{P}) = 1 - R(t|\boldsymbol{\nu}, \boldsymbol{P}) = 1 - (\boldsymbol{\pi}_{\text{ON}}^{(\mathbf{0})})^T \exp(\boldsymbol{\Omega}_1 t | \boldsymbol{\nu}, \boldsymbol{P})\boldsymbol{e}, \tag{4}$$

where $\boldsymbol{e} = (1, \ldots, 1)^T$ is the l-vector of 1's.

We use now these results for our purpose. Assume first that we know from which ON state we start, which, for simplicity, will be 1. Then, $\boldsymbol{\pi}_{\text{ON}}^{(\mathbf{0})} = (1\ 0\ \cdots\ 0)^T$ and, from (4), we have

$$\begin{aligned} R_1(t|\boldsymbol{\nu}, \boldsymbol{P}) &= (1\ 0\ \cdots\ 0) \exp(\boldsymbol{\Omega}_1 t | \boldsymbol{\nu}, \boldsymbol{P})\boldsymbol{e} \\ &= (1\ 0\ \cdots\ 0) \begin{pmatrix} \epsilon_{11}|\boldsymbol{\nu}, \boldsymbol{P} & \cdots & \epsilon_{1l}|\boldsymbol{\nu}, \boldsymbol{P} \\ \vdots & \ddots & \vdots \\ \epsilon_{l1}|\boldsymbol{\nu}, \boldsymbol{P} & \cdots & \epsilon_{ll}|\boldsymbol{\nu}, \boldsymbol{P} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ &= \sum_{j=1}^{l} \epsilon_{1j} | \boldsymbol{\nu}, \boldsymbol{P}. \end{aligned} \tag{5}$$

We need to compute $\exp(\boldsymbol{\Omega}_1 t)$ (for convenience, we omit the dependence on $\boldsymbol{\nu}, \boldsymbol{P}$ when the context is clear), a problem reviewed in Moler and Van Loan (2003). The simplest case is when $\boldsymbol{\Omega}_1$ is diagonalizable with different eigenvalues, which holds with no significant loss of generality, see Geweke, Marshall, and Zarkin (1986), appealing to Dhrymes (1978). We then decompose $\boldsymbol{\Omega}_1 = \boldsymbol{S}\boldsymbol{D}_1\boldsymbol{S}^{-1}$, where $\boldsymbol{D}_1$ is a diagonal matrix with eigenvalues $\lambda_1, \ldots, \lambda_l$ of $\boldsymbol{\Omega}_1$ as its entries, and $\boldsymbol{S}$ is an invertible matrix consisting of the eigenvectors corresponding to the eigenvalues in $\boldsymbol{D}_1$. Then, we have

$$\begin{aligned}\exp(\boldsymbol{\Omega}_1 t) &= \sum_{k=0}^{\infty} \frac{\left(\boldsymbol{S}\boldsymbol{D}_1 t\boldsymbol{S}^{-1}\right)^k}{k!} = \boldsymbol{S}\left[\sum_{k=0}^{\infty} \frac{(\boldsymbol{D}_1 t)^k}{k!}\right]\boldsymbol{S}^{-1} \\ &= \boldsymbol{S}\begin{pmatrix} e^{\lambda_1 t|\boldsymbol{\nu},\boldsymbol{P}} & & \\ & \ddots & \\ & & e^{\lambda_l t|\boldsymbol{\nu},\boldsymbol{P}} \end{pmatrix}\boldsymbol{S}^{-1}.\end{aligned}$$

Should we not know in which ON state we are, based on the equilibrium distribution, we could use $\bar{\pi}_i|\boldsymbol{\nu},\boldsymbol{P} = \pi_i|\boldsymbol{\nu},\boldsymbol{P}/(\sum_{j=1}^{l}\pi_j\ |\boldsymbol{\nu},\boldsymbol{P})$, $i = 1,\ldots,l$ as the initial state probability vector $\boldsymbol{\pi}_{\text{ON}}^{(\mathbf{0})}|\boldsymbol{\nu},\boldsymbol{P}$, and, then,

$$R(t|\boldsymbol{\nu},\boldsymbol{P}) = (\boldsymbol{\pi}_{\text{ON}}^{(\mathbf{0})}|\boldsymbol{\nu},\boldsymbol{P})^T \exp(\boldsymbol{\Omega}_1\, t|\boldsymbol{\nu},\boldsymbol{P})\boldsymbol{e}.$$

We could, then, approximate the posterior reliability through

$$R(t|data) \simeq \frac{1}{N}\sum_{\eta=1}^{N}(\boldsymbol{\pi}_{\text{ON}}^{(\mathbf{0},\eta)})^T \exp(\boldsymbol{\Omega}_1^{(\eta)}\, t)\boldsymbol{e}, \tag{6}$$

where $\left\{\boldsymbol{\Omega}_1^{(\eta)},\, \boldsymbol{\pi}_{\text{ON}}^{(\mathbf{0},\eta)}\right\}_{\eta=1}^{N}$ are associated with samples from the posteriors of $\boldsymbol{\nu}$ and $\boldsymbol{P}$.

For computational issues concerning the exponentiation of a matrix, several software packages are available, see, e.g., Sidje (1998), who introduces the Expokit subroutine and pays special attention to the case of CTMCs. This subroutine is based on Arnoldi's matrix decomposition and allows us to compute the product of the exponentiated matrix times a column vector, without having to compute, explicitly, the matrix exponential in isolation. It is important to remark that the underlying principle in Expokit is to approximate $\exp(\boldsymbol{\Omega}_1 t)\boldsymbol{e}$ using the elements of the Krylov subspace

$$\mathcal{K}_d(\boldsymbol{\Omega}_1 t;\boldsymbol{e}) = \text{Span}\{\boldsymbol{e}, (\boldsymbol{\Omega}_1 t)\boldsymbol{e},\ldots,(\boldsymbol{\Omega}_1 t)^{d-1}\boldsymbol{e}\},$$

where d, the dimension of the Krylov subspace, is small compared with l, the order of the principal matrix (usually $d \leq 50$ while l can exceed many thousands). In practice, we use the version of the subroutine described in Sidje and Stewart (1999), which is shown to solve problems of order up to 66500 with fixed $d = 30$. This implies that, at every iteration, one has to compute the spectral decomposition of a matrix of such dimension d, something which is computationally affordable, see Sidje (1998) for details. The main computational effort within each call to the subroutine is then on Arnoldi's matrix decomposition, which involves an upper bound of $\mathcal{O}(d^2 l)$ flops, see Golub and Van Loan (1996) for details.

5 MAINTAINABILITY FORECASTING WITH CTMCs

Maintainability is defined as the probability of performing a successful repair action within a given time. If T' designates the random variable that represents the time passed until a system is repaired, then, conditional on $\boldsymbol{\nu}, \boldsymbol{P}$, and assuming that the system is initially in an OFF state j, maintainability is provided by

$$\mathcal{M}_j(t|\boldsymbol{\nu}, \boldsymbol{P}) = \Pr\{T' < t|\boldsymbol{\nu}, \boldsymbol{P}, X_0 = j\},\ j = l+1, \ldots, m.$$

As the repair process may be performed in several phases, maintainability should take into account all possible jumps between OFF states before a failure is fixed, that is, before an ON state is reached. Then, to study the maintainability of a system, we proceed in a similar way to that of Section 4.

For the general system introduced in Section 2, with transition probability matrix (3), we subsume all ON states into an absorbing state which we designate b. The intensity matrix of this system is:

$$\boldsymbol{\Lambda}_2 = \left(\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \boldsymbol{\omega}' & \boldsymbol{\Omega}_2 \end{array}\right),$$

where $\boldsymbol{\Omega}_2$ corresponds to OFF states, and $\omega_i' = -\sum_{j=l+1}^{m} r_{ij}$, $i, j = l+1, \ldots, m$, and $r_{ii} = -\nu_i$, $i = l+1, \ldots, m$.

The time until repair is the time until absorption into state b of our modified Markov chain, which is defined as $\tau' = \inf\{t \geq 0 | X_t = b\}$, and has a phase-type distribution $\tau'|\boldsymbol{\nu}, \boldsymbol{P} \sim PH(\boldsymbol{\pi}^{(\mathbf{0})}_{\text{OFF}}, \boldsymbol{\Omega}_2|\boldsymbol{\nu}, \boldsymbol{P})$, where $\boldsymbol{\pi}^{(\mathbf{0})}_{\text{OFF}}$ is an initial state probability vector over OFF states.

Assume now that we know from which OFF state we start, which, for simplicity will be $l+1$. Then, $\boldsymbol{\pi}^{(\mathbf{0})}_{\text{OFF}} = (1\ 0\ \cdots\ 0)^T$. Following a similar reasoning to that of Section 4, we have

$$\mathcal{M}_{l+1}(t|\boldsymbol{\nu}, \boldsymbol{P}) = 1 - \sum_{i=1}^{m-l} \epsilon_{1i}'|\boldsymbol{\nu}, \boldsymbol{P}, \tag{7}$$

where $\epsilon_{1i}'|\boldsymbol{\nu}, \boldsymbol{P}$ are the elements of the first row of $\exp(\boldsymbol{\Omega}_2\, t|\boldsymbol{\nu}, \boldsymbol{P})$, which is, again, the key computational task.

Should we not know in which OFF state we are, based on the equilibrium distribution, we could use $\tilde{\pi}_i|\boldsymbol{\nu}, \boldsymbol{P} = \pi_i|\boldsymbol{\nu}, \boldsymbol{P}/(\sum_{j=l+1}^{m} \pi_j|\boldsymbol{\nu}, \boldsymbol{P})$, $i = l+1, \ldots, m$ as the initial state probability vector $\boldsymbol{\pi}^{(\mathbf{0})}_{\text{OFF}}|\boldsymbol{\nu}, \boldsymbol{P}$, and then, we could approach (7) through

$$\mathcal{M}(t|data) \simeq 1 - \frac{1}{N} \sum_{\eta=1}^{N} (\boldsymbol{\pi}^{(\mathbf{0},\eta)}_{\text{OFF}})^T \exp(\boldsymbol{\Omega}_2^{(\eta)}\, t)\boldsymbol{e}', \tag{8}$$

where $\left\{\boldsymbol{\Omega}_2^{(\eta)}, \boldsymbol{\pi}^{(\mathbf{0},\eta)}_{\text{OFF}}\right\}_{\eta=1}^{N}$ are associated with samples from the posteriors of $\boldsymbol{\nu}$ and $\boldsymbol{P}$, and $\boldsymbol{e}' = (1, \ldots, 1)^T$ is the $(m-l)$-vector of 1's.

6 AVAILABILITY FORECASTING WITH CTMCs

Availability is closely related to reliability and maintainability and is becoming a standard system performance measure, see Maciejewski and Caban (2008) and Lee (2000) for reviews.

From a steady-state point of view, availability is the sum of the limiting probabilities for ON states, conditional on $\boldsymbol{\nu}, \boldsymbol{P}$, that is

$$A|\boldsymbol{\nu}, \boldsymbol{P} = \sum_{i=1}^{l} \pi_i|\boldsymbol{\nu}, \boldsymbol{P}.$$

Then, as in Section 3.3, if the posterior distributions of ν_i, p_{ij}, are, respectively, very peaked about their posterior modes, $\hat{\nu}_i$ and $\hat{p}_{ij}$, we may use the approximate steady-state availability

$$\hat{A}|data = \sum_{i=1}^{l} \hat{\pi}_i|data.$$

Otherwise, we would obtain a posterior steady-state availability sample

$$\left\{ A^\eta|data = \sum_{i=1}^{l} \pi_i^\eta|data \right\}_{\eta=1}^{N},$$

and summarize it accordingly.

As discussed in Lee (2000), we may also be interested in a type of short-term availability, called interval availability. Define the indicator random variable θ_t

$$\theta_t|\boldsymbol{\nu}, \boldsymbol{P} = \begin{cases} 1, & \text{if } X_t|\boldsymbol{\nu}, \boldsymbol{P} \in \{1, 2, \ldots, l\} \\ 0, & \text{otherwise} \end{cases},$$

and

$$A_t|\boldsymbol{\nu}, \boldsymbol{P} = \frac{1}{t} \int_0^t \theta_s|\boldsymbol{\nu}, \boldsymbol{P} \; ds.$$

Then, assuming that the system is initially in an ON state $i \in \{1, \ldots, l\}$, the interval availability is

$$I_t|\boldsymbol{\nu}, \boldsymbol{P} = E[A_t|\boldsymbol{\nu}, \boldsymbol{P}] = \frac{1}{t} \sum_{j=1}^{l} \int_0^t \pi_j(s|\boldsymbol{\nu}, \boldsymbol{P}) \, ds. \tag{9}$$

We then define the posterior interval availability through

$$\begin{aligned} I_t|data &= \iint \left(I_t|\boldsymbol{\nu}, \boldsymbol{P}\right) \pi(\boldsymbol{\nu}, \boldsymbol{P}|data) \, d\boldsymbol{\nu} \, d\boldsymbol{P} \\ &\simeq \frac{1}{N} \sum_{\eta=1}^{N} I_t|\boldsymbol{\nu}^\eta, \boldsymbol{P}^\eta, \end{aligned} \tag{10}$$

for appropriate posterior samples $\{\boldsymbol{\nu}^\eta, \boldsymbol{P}^\eta\}_{\eta=1}^{N}$.

There are two issues concerning the computation of (9). Firstly, the computation of $E[A_t|\boldsymbol{\nu}, \boldsymbol{P}]$. To accomplish that, we solve the Chapman-Kolmogorov system of differential equations (Ross 2007)

$$\begin{aligned} &\boldsymbol{\pi}'(t|\boldsymbol{\nu}, \boldsymbol{P}) = \boldsymbol{\Lambda}|\boldsymbol{\nu}, \boldsymbol{P} \cdot \boldsymbol{\pi}(t|\boldsymbol{\nu}, \boldsymbol{P}); \quad t \in [0, T) \\ &\boldsymbol{\pi}(0|\boldsymbol{\nu}, \boldsymbol{P}) = \boldsymbol{\pi}^{(\mathbf{0})}, \end{aligned} \tag{11}$$

where $\boldsymbol{\pi}^{(\mathbf{0})} = \left(\pi_1^{(0)}, \pi_2^{(0)}, \ldots, \pi_m^{(0)}\right)^T$ is the initial state probability vector, and $\boldsymbol{\Lambda}|\boldsymbol{\nu}, \boldsymbol{P}$ is the intensity matrix defined in Section 3.4, conditional on $\boldsymbol{\nu}, \boldsymbol{P}$. Its analytic solution is $\boldsymbol{\pi}(t|\boldsymbol{\nu}, \boldsymbol{P}) = \exp(t \cdot \boldsymbol{\Lambda}|\boldsymbol{\nu}, \boldsymbol{P}) \cdot \boldsymbol{\pi}^{(\mathbf{0})}$, and represents the state probability vector, conditional on $\boldsymbol{\nu}, \boldsymbol{P}$, where $\pi_i(t|\boldsymbol{\nu}, \boldsymbol{P}) = \Pr\{X(t|\boldsymbol{\nu}, \boldsymbol{P}) = i\}$. Note again that the key operation is that of matrix exponentiation.

The second issue is that of computing the integral that appears in (9). To do so, we may use the general Simpson's rule, dividing $[0, t)$ in $2M$ subintervals

$$\begin{aligned} I_t|\boldsymbol{\nu}, \boldsymbol{P} &\simeq \frac{1}{t}\frac{1}{3}h \sum_{j=1}^{l} \{\pi_j(t_0|\boldsymbol{\nu}, \boldsymbol{P}) \\ &+ 4[\pi_j(t_1|\boldsymbol{\nu}, \boldsymbol{P}) + \pi_j(t_3|\boldsymbol{\nu}, \boldsymbol{P}) + \cdots + \pi_j(t_{2M-1}|\boldsymbol{\nu}, \boldsymbol{P})] \\ &+ 2[\pi_j(t_2|\boldsymbol{\nu}, \boldsymbol{P}) + \pi_j(t_4|\boldsymbol{\nu}, \boldsymbol{P}) + \cdots + \pi_j(t_{2M-2}|\boldsymbol{\nu}, \boldsymbol{P})] \\ &+ \pi_j(t_{2M}|\boldsymbol{\nu}, \boldsymbol{P})\}, \end{aligned} \tag{12}$$

where $h = t/2M$, and $t \equiv t_{2M}$. In this way, we can obtain intermediate values at even times t_{2k}, $k = 1, \ldots, M$ within the process of integration, using the recursive relationship

$$\begin{aligned} I_{t_{2k}}|\boldsymbol{\nu}, \boldsymbol{P} &= \frac{1}{k}\Bigg\{(k-1)I_{t_{2k-2}}|\boldsymbol{\nu}, \boldsymbol{P} + \frac{1}{6}\sum_{j=1}^{l}[\pi_j(t_{2k-2}|\boldsymbol{\nu}, \boldsymbol{P}) \\ &+ 4\pi_j(t_{2k-1}|\boldsymbol{\nu}, \boldsymbol{P}) + \pi_j(t_{2k}|\boldsymbol{\nu}, \boldsymbol{P})]\Bigg\}. \end{aligned}$$

7 RAM ANALYSIS OF AN ERP CONFIGURATION

As a case study, we shall consider a multiserver system functioning to support our university Enterprise Resource Planner (ERP), which facilitates administrative processes to students (e.g., registering for a course), professors (e.g., checking resources concerning a grant) and staff (e.g., asking for a leave of absence), through the Internet. The general architecture of our ERP is shown in Figure 2. A user makes a petition to the system. The petition passes through an active/active web-cache (WC) cluster balancer (see Xie et al. (2004) for details on various cluster architectures), which distributes the load between four application (AP) servers. The balancer works if at least one of its two WC servers is up. The four AP servers work on a 2-out-of-4 basis, that is, it works if at least two out of four servers are up, see Shao and Lamberson (1991) for details. In this class of systems, each component is designed to carry only part of the total load. The AP servers access the database and complete the service to the user. Our ERP system works well in practice, attaining high availabilities and managing to offer high-quality services to over 32,000 users.

The transition diagram of our system is shown in Figure 3, along with the jumping intensities $r_{ij} = \nu_i p_{ij}$. The meaning of the nine states in our Markov chain is explained in Table 1. We have available failure and repair data, recorded by our ERP managers over an eight months period. A summary of the data is displayed in Table 2, which presents the transition counts n_{ij} and, in parentheses, the total sojourn time (in hours) that the system spent in state i before jumping into state j. For example, the second entry in the first row, 8(508.98), means that eight transitions have occurred from State 1 into State 2, and that the system spent a total time of 508.98 hours in State 1 before jumping into State 2, over the eight transitions.

The main features of our system are listed below:

1. The failure rate (number of failures per unit time) of each WC server is λ_{WC}. Our ERP managers believe that the WC server fails, roughly, once per fortnight.

2. The failure rate for the AP servers, λ_i^{AP}, $i = 2, 3, 4$, is constant and shared when there are i functioning servers, satisfying $\lambda_4^{\text{AP}} < \lambda_3^{\text{AP}} < \lambda_2^{\text{AP}}$: when the four AP servers function the load on each is smaller than when three are functioning, which, in turn, is smaller than when just two are working. Recall that when there is just one server on, the system is unable to cope with the load. Our ERP managers' beliefs about how often the AP servers fail are summarized as follows: if four AP servers are functioning, the failure rate for each server is reduced to 50% of the original level; if three AP servers are functioning, the failure rate is reduced to 65% of the original level; while if only two AP servers are functioning it drops to 75% of the original level, relative to the failure rate for a single AP server that would be ideally functioning alone, λ^{AP}, which is, roughly, one failure per week.

3. When a WC server fails, our ERP managers consider that, on average, it takes approximately three hours to fix it. A repaired WC server is assumed to be as good as new, and it is immediately reconnected to the system through hot plugging in. When both WC servers fail (Failure type I; State 7), the system fails; it is repaired and then restarts, as good as new, from State 1 (all currently non-functioning AP servers are also fixed). This task takes a longer time, which, on average, is approximately five hours, based on our ERP managers' experience.

4. The WC cluster manages load balancing. At least, two AP servers should be up in peak operation periods. Only one AP server can be repaired at a time, with an average repair time of two hours. A repaired AP server is assumed to be as good as new, and it is immediately reconnected to the system through hot plugging in. When more than two servers are down, the system fails, because of too heavy demand over the remaining server (Failure type II; State 8). When this occurs, the system is fixed (including any currently non-functioning WC server), and restarts, as good as new, from State 1. Based on our ERP managers' experience, this process takes an average time of, roughly, five hours.

5. The WC cluster balancer also detects and disconnects failed AP servers, but such detection process has a probability α of success. If the balancer cannot detect and disconnect a failed AP server, the whole system fails (Failure type III; State 9). Based on our system managers' experience, the average time needed to recover the system from such failure and restart, as good as new, is also, approximately, five hours. Our system managers are convinced that the WC cluster balancer performs successfully in around

99% of its detection operations. Detection is a programmed task which occurs every two hours, consuming a negligible time, and, practically, not affecting the overall system performance.

The transition probability matrix is

$$\boldsymbol{P} = \begin{array}{cc} & \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{array} & \left(\begin{array}{cccccc|ccc} 0 & p_{12} & p_{13} & 0 & 0 & 0 & 0 & 0 & p_{19} \\ p_{21} & 0 & 0 & p_{24} & 0 & 0 & p_{27} & 0 & p_{29} \\ p_{31} & 0 & 0 & p_{34} & 0 & p_{36} & 0 & 0 & p_{39} \\ 0 & p_{42} & p_{43} & 0 & p_{45} & 0 & p_{47} & 0 & p_{49} \\ 0 & 0 & 0 & p_{54} & 0 & p_{56} & p_{57} & p_{58} & 0 \\ 0 & 0 & p_{63} & 0 & p_{65} & 0 & 0 & p_{68} & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \end{array} .$$

The permanence rates are, respectively, $\nu_1 = 2\lambda_{\text{WC}} + 4\lambda_4^{\text{AP}}, \nu_2 = \lambda_{\text{WC}} + 4\lambda_4^{\text{AP}} + \mu_{\text{WC}}, \nu_3 = 2\lambda_{\text{WC}} + 3\lambda_3^{\text{AP}} + \mu_{\text{AP}}, \nu_4 = \lambda_{\text{WC}} + 3\lambda_3^{\text{AP}} + \mu_{\text{WC}} + \mu_{\text{AP}}, \nu_5 = \lambda_{\text{WC}} + 2\lambda_2^{\text{AP}} + \mu_{\text{WC}} + \mu_{\text{AP}}, \nu_6 = 2\lambda_{\text{WC}} + 2\lambda_2^{\text{AP}} + \mu_{\text{AP}}$, and $\nu_7 = \nu_8 = \nu_9 = \rho$, where μ_{WC} and μ_{AP} are the WC and AP repair rates, respectively, and ρ is the system repair rate when the system falls into any OFF state.

In order to assign prior distributions to the failure and repair rates, we have considered two different scenarios:

A. We are relatively sure about the failure and repair rates average values, as recorded by our ERP managers. In this case, we assign peaked gamma priors around such values, with small variances.

B. We might be less sure about the rates involved. We, then, build in more uncertainty, and, therefore, assign more diffuse gamma priors (with the same average values).

With this in mind, we have summarized our prior choices in Table 3. For convenience, we have expressed all rates in terms of failures/repairs per fortnight. As an example, let us consider λ_{WC}. In the first scenario, we are relatively sure that there will be around one failure per fortnight. We assume a prior $\lambda_{\text{WC}} \sim \mathcal{G}(10, 10)$, whose mean and variance are equal to 1. In the second scenario, we take into account more uncertainty, and, therefore, our prior choice is $\lambda_{\text{WC}} \sim \mathcal{G}(0.1, 0.1)$, with mean 1 and variance 10.

Based on the data in Table 2, we get the posterior parameters described in Table 4. To compute the posterior of, e.g., λ_{WC}, the only transitions we have to take into account are (1,2), (2,7), (3,4), (4,7), (5,7), and (6,5), see Figure 3. Once we have the posteriors for the failure and repair rates, we can obtain samples from the posteriors for the ν's and the p_{ij}'s, using their relationships to the failure and repair rates. Then, to obtain a sample from the posterior equilibrium distribution, we iteratively solve system (2), for various posterior samples of $\boldsymbol{\nu}$ and $\boldsymbol{P}$, which, in this example, is

$$\begin{cases} \nu_1\pi_1 = r_{21}\pi_2 + r_{31}\pi_3 + r_{71}\pi_7 + r_{81}\pi_8 + r_{91}\pi_9, \\ \nu_2\pi_2 = r_{12}\pi_1 + r_{42}\pi_4, \\ \nu_3\pi_3 = r_{13}\pi_1 + r_{43}\pi_4 + r_{63}\pi_6, \\ \nu_4\pi_4 = r_{24}\pi_2 + r_{34}\pi_3 + r_{54}\pi_5, \\ \nu_5\pi_5 = r_{45}\pi_4 + r_{65}\pi_6, \\ \nu_6\pi_6 = r_{36}\pi_3 + r_{56}\pi_5, \\ \nu_7\pi_7 = r_{27}\pi_2 + r_{47}\pi_4 + r_{57}\pi_5, \\ \nu_8\pi_8 = r_{58}\pi_5 + r_{68}\pi_6, \\ \nu_9\pi_9 = r_{19}\pi_1 + r_{29}\pi_2 + r_{39}\pi_3 + r_{49}\pi_4, \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 + \pi_7 + \pi_8 + \pi_9 = 1, \\ \pi_i \geq 0. \end{cases}$$

Figure 4a shows density plots for the posterior equilibrium distribution obtained when using peaked priors (Scenario A). We may summarize it through its mean which is

$$\begin{array}{lll} \hat{\pi}_1 = 0.9455; & \hat{\pi}_2 = 0.0330; & \hat{\pi}_3 = 0.0134; \\ \hat{\pi}_4 = 0.0005; & \hat{\pi}_5 = 9.422 \cdot 10^{-6}; & \hat{\pi}_6 = 0.0003; \\ \hat{\pi}_7 = 0.0010; & \hat{\pi}_8 = 1.309 \cdot 10^{-5}; & \hat{\pi}_9 = 0.0063. \end{array}$$

When more diffuse priors are used, there is also more uncertainty in the posterior equilibrium distribution, see Figure 4b. In this scenario, the system is less prone to be in State 1, and there is a significant increase in the value of the limiting probability associated with State 2 (one WC server down). The limiting probabilities corresponding to OFF states, π_7, π_8, π_9, although still relatively small, have increased significantly their respective values, and, e.g., the system would spend now, in the long run, around 1% of the time in State 9, due to failures in the detection process.

Assuming that the initial state is 1, and considering peaked priors for the failure and repair rates, the reliability conditional on $(\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}})$ (the posterior modes) is obtained by plugging in such estimates in (5)

$$\begin{aligned} R_1(t|\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}}) &= \epsilon_{11}|\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}} + \epsilon_{12}|\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}} + \epsilon_{13}|\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}} \\ &\quad + \epsilon_{14}|\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}} + \epsilon_{15}|\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}} + \epsilon_{16}|\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}} \\ &= -3.44 \cdot 10^{-9} e^{-312.60t} + 1.76 \cdot 10^{-8} e^{-264.86t} \\ &\quad + 6.00 \cdot 10^{-6} e^{-196.18t} - 3.06 \cdot 10^{-5} e^{-148.45t} \\ &\quad - 5.75 \cdot 10^{-4} e^{-116.90t} + 1.00 e^{-0.49t}. \end{aligned}$$

Time units have been expressed, for simplicity, in fortnights, although we will plot the time axis in hours in all the graphics below. In Figure 5a, we have plotted $R_1(t|\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{P}})$, plugging in the MLEs (dashed line) and the posterior modes (solid line) as estimates of $(\boldsymbol{\nu}, \boldsymbol{P})$, along with .95 predictive bands (dotted lines) around the mean reliabilities (dotted-dashed line), as computed from (6). Note that there is a lot of uncertainty which is ignored through the standard approach based on plugging in parameter estimates. As we can observe, the predictive bands are quite separated from the posterior mean reliability, with relative errors close to 100% for time values around 1000 hours (approx. 40 days). The values of the reliability conditional on the posterior modes of $(\boldsymbol{\nu}, \boldsymbol{P})$ are, in practice, the same as those of the posterior mean reliability. On the other hand, the reliability conditional on the MLEs of $(\boldsymbol{\nu}, \boldsymbol{P})$ drops drastically to zero, something which indicates the poor quality of the MLEs when little data are available, as in this case. It should be noted that when peaked priors are used, they tend to dominate the analysis. Nevertheless, we should also remark that our aim is to provide additional information to that of reliability point estimation, based on MLEs, posterior modes or similar, by plotting predictive bands.

The same set of graphics is plotted in Figure 5b, when more diffuse priors are used for the failure and repair rates. As we can observe, the expected reliability has decreased with respect to its value in Figure 5a, and the uncertainty is much larger now. In this case, there is little difference between the results obtained when using the posterior modes, the MLEs or the posterior mean reliability, due to the relatively low importance of the prior distribution of the failure and repair rates. Finally, we should note that similar results are obtained for both scenarios when starting from a different ON state, as the repair rates are much higher than the failure rates, and, therefore, the systems quickly tends to be in State 1. A similar behavior is also observed when the initial ON state is unknown, as the system reliability in this case is a weighted sum of the reliabilities when starting at the different ON states, with weights depending on the limiting probabilities π_i, $i = 1, \ldots, 6$, as was mentioned in Section 4.

In a similar way, we perform prediction for the system maintainability when peaked priors are used for the failure and repair rates, based on (7). In Figure 6a, we have plotted system maintainability when the initial OFF State is 7, plugging in the posterior modes of $(\boldsymbol{\nu}, \boldsymbol{P})$ (solid line) and the MLEs (dashed line) in (7). We have also plotted .95 predictive bands (dotted lines) around the posterior mean maintainability (dotted-dashed line), as computed from (8). Note that, in this case, less uncertainty is present, as the sojourn time at any OFF state depends only on the value of ρ, which has a peaked distribution in the first scenario. Again, the predicted maintainability when plugging in the MLEs of $(\boldsymbol{\nu}, \boldsymbol{P})$ is much worse than when plugging in the posterior modes, or when computing the posterior mean maintainability. When more diffuse priors for the failure and repair rates are used, the results are quite similar, see Figure 6b. However, the expected maintainability is worse than in the previous case, and the uncertainty has now significantly larger values. No difference is practically found when we start from the OFF states 8 or 9, as we have assumed that the repair rate is the same for states 7, 8 and 9.

Finally, to forecast system availability, we make use of (12), following (11), to compute the value of the state probability vector $\boldsymbol{\pi}(t|\boldsymbol{\nu}, \boldsymbol{P})$ at each point of the interval $[0, t)$, which has been divided in 200 subintervals. We have plotted the system availability in Figure 7a, plugging in the posterior modes (solid line) and the MLEs (dashed line) of $(\boldsymbol{\nu}, \boldsymbol{P})$, when starting in State 1. We have also plotted .95 predictive bands (dotted lines) around the posterior mean availabilities, as computed from (10), when the system was initially in State 1 (dotted-dashed line) or in State 5 (long-dashed line). As we can observe, the posterior mean availability has, in the transient regime, a lower value when the system starts in State 5 than when it starts in State 1. This is reasonable, as in State 1 all the components are functioning, whereas State 5 is a critical one, as one WC server and two AP servers are down. However, in the long run, the mean availability tends to the steady-state availability, regardless of the initial ON state. Again, the availability computed when plugging in the MLEs underrepresents the values obtained when plugging in the posterior modes, or when the posterior mean availabilities, starting from States 1 or 5, are computed. We can observe that, in this case, the uncertainty is, in practice, negligible, with relative errors less than 1%. This is reasonable, as our ERP system is designed as a high-availability device. Its availability will tend, in the long run, to a steady-state availability, whose value is the sum of the limiting probabilities at the ON states, π_i, $i = 1, \ldots, 6$. Such probabilities have peaked posteriors, see Figure 4a, and, therefore, there will be little uncertainty when computing the posterior availability.

When we use more diffuse priors for the failure and repair rates, we observe a similar behavior, see Figure 7b. More uncertainty is present, especially when the initial State is 5. In this case, the posterior mean availability goes below .90 in the transient regime (at a time around 10 hours), although it gradually reaches a steady value around .95, as time runs, merging asymptotically with the availabilities obtained for the other cases.

8 CONCLUSIONS

Forecasting reliability, availability and maintainability of hardware components in safety critical systems is becoming increasingly important. Although there are several software packages available to support reasonably complex RAM analyses, such as Relex, CASRE or SHARPE, among others, they have several limitations. One of the most important ones is the fact that they provide little support to Bayesian analysis. With this drawback in mind, we have focused, in this paper, on Bayesian analysis of such systems.

We have provided a description of hardware systems through Continuous Time Markov Chains. We have followed a Bayesian approach to estimate the reliability, availability and maintainability of such systems, devoting our attention to three key items: description of the CTMC and the underlying discrete Markov process; inference and forecasting for the CTMC parameters and the posterior equilibrium distribution; and system RAM forecasting.

We have considered the role of prior uncertainty, and whether this is more or less tight, and how this influences the results, especially in cases in which there is little data, as in our case.

We have shown how the key computational issue refers to Arnoldi's matrix decomposition, within the subroutine of matrix exponentiation, and discussed its potential inefficiency when the number of states becomes too large. In such case, we could alternatively use reduced order models, see Grigoriu (2009) and references therein. These models would approximate the posterior on $\boldsymbol{\nu}$, $\boldsymbol{P}$ through k points $\boldsymbol{\nu}^k$, $\boldsymbol{P}^k$ with weights q^k, appropriately chosen, and then approximate $R(t|data)$ by $\sum q^k R(t|\boldsymbol{\nu}^k, \boldsymbol{P}^k)$, thus performing only k matrix exponentiations, with k chosen based on our computational budget.

Another issue of interest refers to the actual convenience of using CTMCs to model the system. Should the holding times not follow exponential distributions, we could rely on semimarkovian processes to analyze the system. We should note, however, that although the equilibrium analysis is as easy as we have described here, the transient analysis, of special interest to us, becomes more difficult. See Marín, Plà, and Ríos Insua (2005) for an example in a different application area. This will be the subject of future work.

ACKNOWLEDGMENTS

Work supported by projects from the Spanish Ministry of Science and Innovation program MTM2009-14087-C04-01, and the Government of Madrid RIESGOS-CM program S2009/ESP-1685. We are grateful to Luis Redondo, from MTP, for suggesting us this research area. We also acknowledge José Manuel Dévora, manager of URJC Corporative Applications Division, for his invaluable advice and support throughout the discussion and preparation of the case study. The first version of the paper was prepared while the third author was visiting SAMSI. We are also grateful to the referees and editors for their very helpful comments.

REFERENCES

Baier, C., Haverkort, B., Hermanns, H., and Katoen, J. (2003), "Model-Checking Algorithms for Continuous-Time Markov Chains," *IEEE Transactions on Software Engineering 29*(6), 524–541.

Berger, J., and Ríos Insua, D. (1998), "Recent Developments in Bayesian Inference with Applications in Hydrology," in *Statistical and Bayesian Methods in Hydrological Sciences*, eds. E. Parent, P. Hubert, B. Bobée, and J. Miquel, Paris: UNESCO Press, pp. 56–80.

Bladt, M. (2005), "A Review on Phase-Type Distributions and Their Use in Risk Theory," *Astin Bulletin 35*(1), 145–161.

Dhrymes, P. (1978), *Mathematics for Econometrics*, Springer.

Diaconis, P., and Rolles, S. W. W. (2006), "Bayesian Analysis for Reversible Markov Chains," *Annals of Statistics 34*(3), 1270–1292.

Eusgeld, I., Fechner, B., Salfner, F., Walter, M., Limbourg, M., and Zhang, L. (2008), "Hardware Reliability," in *Dependability Metrics*, eds. I. Eusgeld, F. C. Freiling, and R. Reussner, Berlin: Springer, pp. 59–103.

French, S., and Ríos Insua, D. (2000), *Statistical Decision Theory*, Arnold.

Geweke, J., Marshall, R., and Zarkin, G. (1986), "Mobility Indices in Continuous Time Markov Chains," *Econometrica 54*(6), 1407–1423.

Glynn, P. (1986), "Problems in Bayesian Analysis of Stochastic Simulation," in *Proceedings of the 18th Conference on Winter Simulation*, pp. 376–379. ACM New York, NY, USA.

Golub, G. H., and Van Loan, C. F. (1996), *Matrix Computations*, Johns Hopkins University Press.

Grigoriu, M. (2009), "Reduced Order Models for Random Functions. Application to Stochastic Problems," *Applied Mathematical Modelling 33*(1), 161–175.

Kim, H., Lee, H., and Lee, K. (2005), "The Design and Analysis of AVTMR (All Voting Triple Modular Redundancy) and Dual-Duplex System," *Reliability Engineering & System Safety, 88*(3), 291–300.

Lee, K. (2000), "Stochastic Models for Random-Request Availability," *IEEE Transactions on Reliability 49*(1), 80–84.

Maciejewski, H., and Caban, D. (2008), "Estimation of repairable System Availability Within Fixed Time Horizon," *Reliability Engineering and System Safety 93*(1), 100–106.

Marín, J., Plà, L., and Ríos Insua, D. (2005), "Forecasting for Some Stochastic Process Models Related to Sow Farm Management," *Journal of Applied Statistics 32*(8), 797–812.

Moler, C., and Van Loan, C. (2003), "Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later," *SIAM Review 45*(1), 3–49.

Prowell, S., and Poore, J. (2004), "Computing system Reliability Using Markov Chain Usage Models," *The Journal of Systems & Software 73*(2), 219–225.

Pukite, J., and Pukite, P. (1998), *Modeling for Reliability Analysis: Markov Modeling for Reliability, Maintainability, Safety, and Supportability Analyses of Complex Computer Systems*, IEEE Press.

Ross, S. M. (2007), *Introduction to Probability Models*, Academic Press.

Shao, J., and Lamberson, L. (1991), "Modeling a Shared-Load k-out-of-n:G System," *IEEE Transactions on Reliability 40*(2), 205–209.

Sidje, R. B. (1998), "Expokit: a Software Package for Computing Matrix Exponentials," *ACM Transactions on Mathematical Software 24*(1), 130–156.

Sidje, R. B., and Stewart, W. J. (1999), "A Numerical Study of Large Sparse Matrix Exponentials Arising in Markov Chains," *Computational Statistics & Data Analysis 29*(3), 345–368.

Xie, M., Dai, Y., and Poh, K. (2004), *Computing Systems Reliability: Models and Analysis*, Kluwer Academic/Plenum Publishers.

TABLES

Table 1: Meaning of the states of the Markov chain

State	(#WC servers down,#AP servers down)
1	(0,0)
2	(1,0)
3	(0,1)
4	(1,1)
5	(1,2)
6	(0,2)
7	Failure type I: all WC servers down
8	Failure type II: 3 or 4 AP servers down
9	Failure type III: erroneous balancer detection

Table 2: Transition counts and sojourn times

		Final state								
		1	2	3	4	5	6	7	8	9
Initial state	1	–	8 (508.98)	6 (5007.89)	–	–	–	–	–	2 (31.68)
	2	6 (68.18)	–	–	–	–	–	2 (39.11)	–	–
	3	4 (27.72)	–	–	1 (13.35)	–	1 (4.09)	–	–	–
	4	–	–	1 (1.57)	–	–	–	–	–	–
	5	–	–	–	–	–	–	–	1 (0.02)	–
	6	–	–	–	–	1 (8.86)	–	–	–	–
	7	2 (29.36)	–	–	–	–	–	–	–	–
	8	1 (2.21)	–	–	–	–	–	–	–	–
	9	2 (19.13)	–	–	–	–	–	–	–	–

Table 3: Prior parameters of the failure and repair rates

	Small variance		Large variance	
	α_{sv}	β_{sv}	α_{lv}	β_{lv}
λ_{WC}	10	10	0.1	0.1
λ_4^{AP}	9.2	9.6	0.092	0.096
λ_3^{AP}	16	13	0.16	0.13
λ_2^{AP}	22.5	15	0.225	0.15
μ_{WC}	1300	11	13	0.11
μ_{AP}	2800	17	28	0.17
ρ	450	6.7	4.5	0.067

Table 4: Posterior parameters of the failure and repair rates

	Small variance		Large variance	
	$\alpha_{\mathrm{sv}}^{\mathrm{post}}$	$\beta_{\mathrm{sv}}^{\mathrm{post}}$	$\alpha_{\mathrm{lv}}^{\mathrm{post}}$	$\beta_{\mathrm{lv}}^{\mathrm{post}}$
λ_{WC}	23	11.70	13.1	1.80
λ_4^{AP}	17.2	24.6	8.092	15.095
λ_3^{AP}	17	13.01	1.16	0.14
λ_2^{AP}	23.5	15.00	1.225	0.15
μ_{WC}	1307	11.21	20	0.32
μ_{AP}	2804	17.08	32	0.25
ρ	455	6.85	9.5	0.218

FIGURES

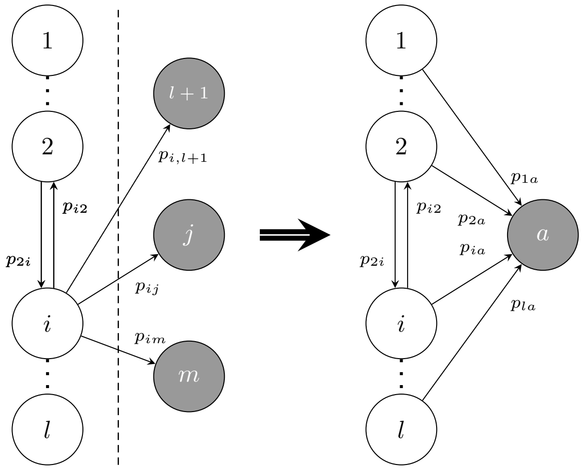


Figure 1: Subsuming the OFF states in state a

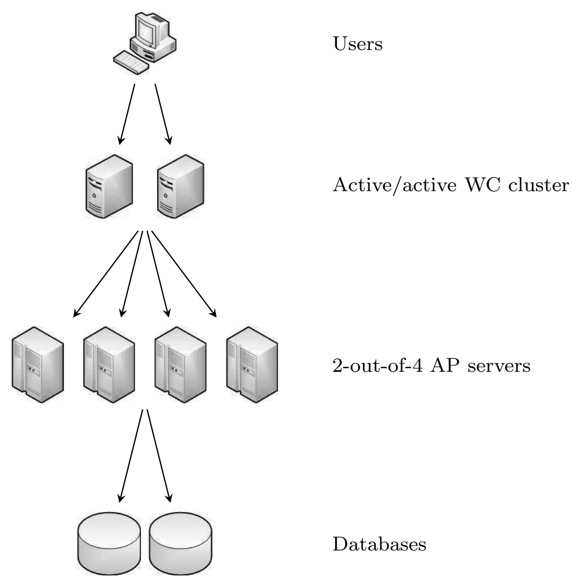


Figure 2: Architecture of our university ERP

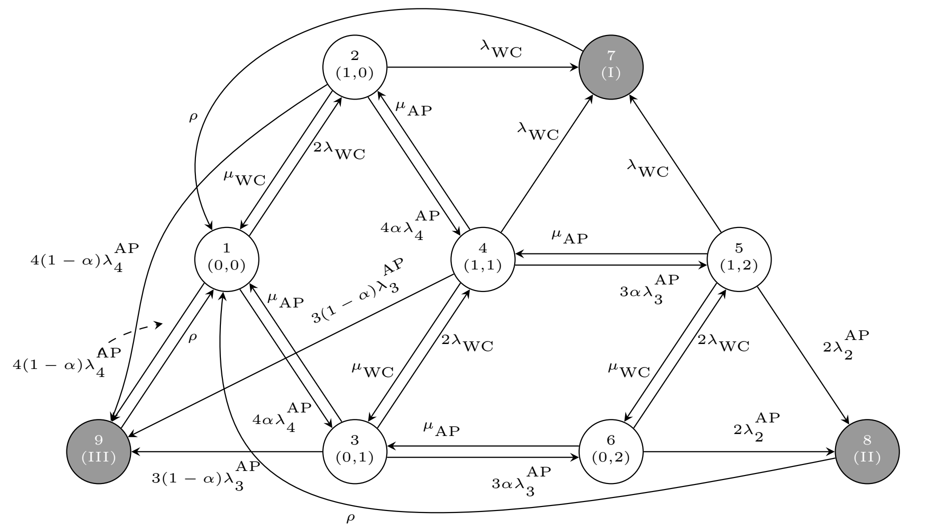
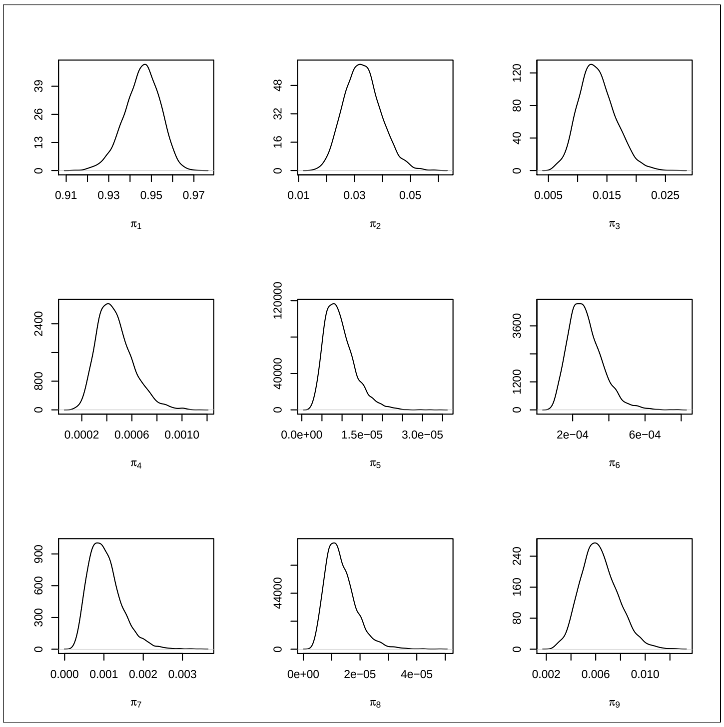
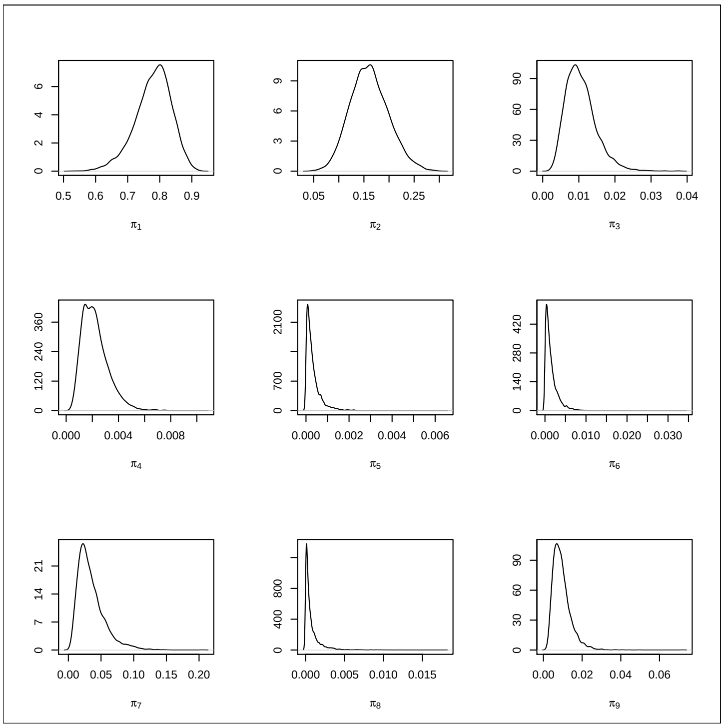


Figure 3: Transition diagram for university ERP infrastructure with jumping intensities

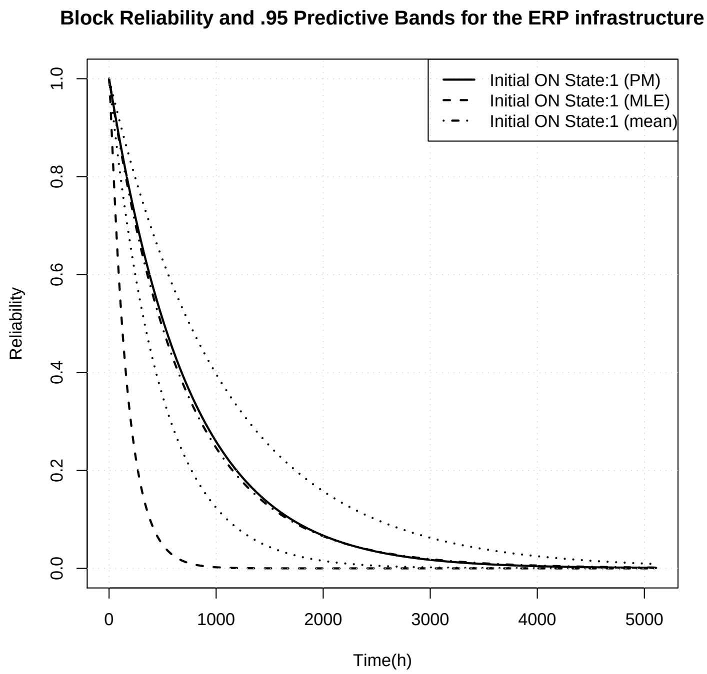


(a) Peaked priors

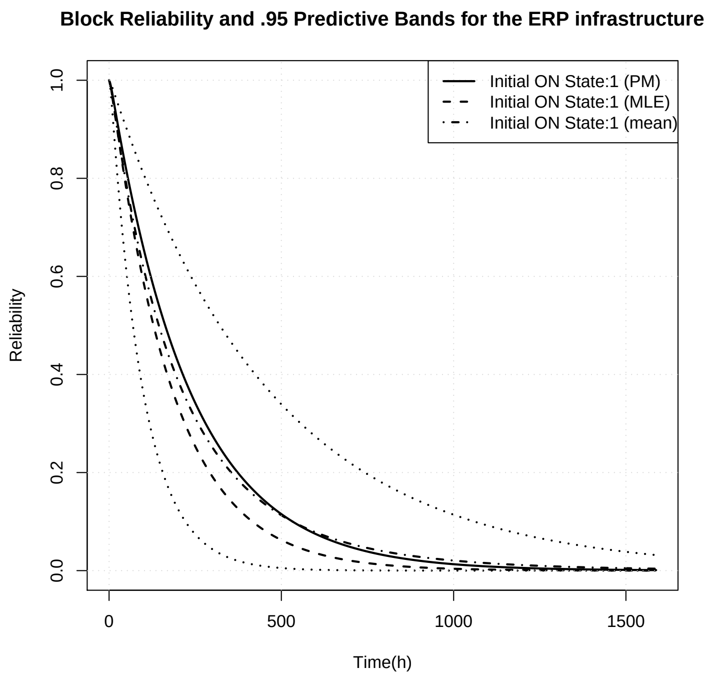


(b) Diffuse priors

Figure 4: Posterior equilibrium distribution

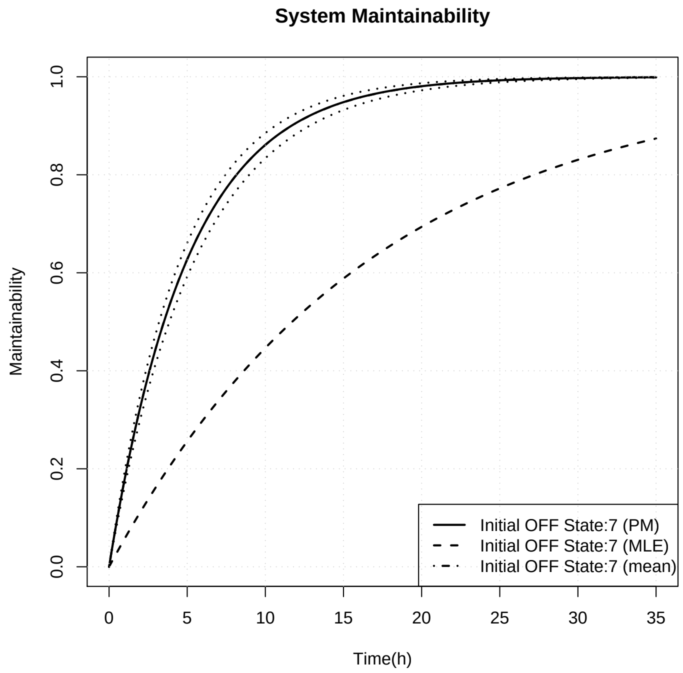


(a) Peaked priors

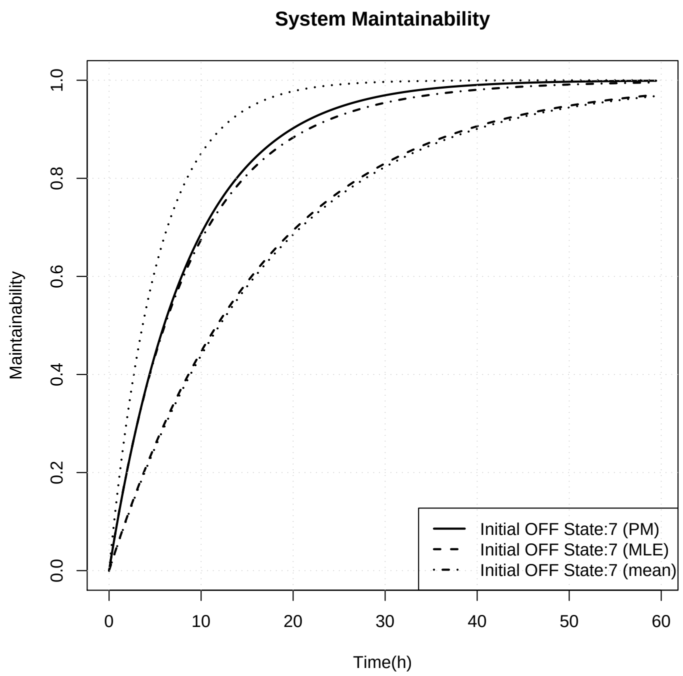


(b) Diffuse priors

Figure 5: System reliability

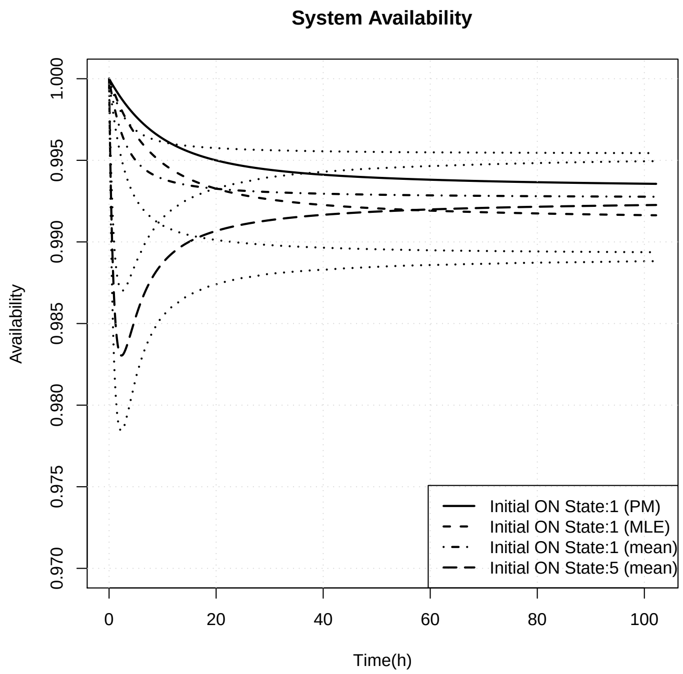


(a) Peaked priors

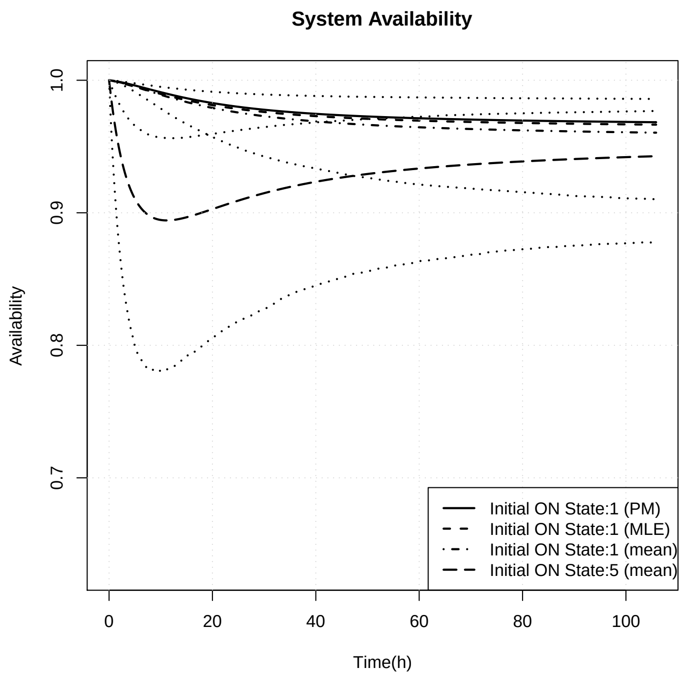


(b) Diffuse priors

Figure 6: System maintainability



(a) Peaked priors



(b) Diffuse priors

Figure 7: System availability