Joint prediction of health status and demand for patient in home care services: a Bayesian approach.

Raffaele Argiento[1], Alessandra Guglielmi[2], Ettore Lanzarone[1], and Inad Nawajah[3]

[1]CNR-IMATI, [2]Politecnico di Milano and [3] Hebron University

Abstract

Estimation of uncertain future patients' demands is a key factor for appropriately planning human and material resources in health care facilities, where unplanned demand variations may deteriorate the quality of schedules and, consequently, of the provided service. This issue is even more important for health services provided outside hospitals, e.g., for home care services, where patients are assisted for a longer period and additional planning decisions related to service delivery in the territory must be taken. With the goal of helping home care management to take robust decisions, in this paper we propose a Bayesian model for estimating and predicting both the demand for care and the history of health conditions for patients in the charge of a home care service. In particular, we jointly model the temporal evolution of patients' care profile and the weekly number of visits required to nurses, and use a Markov chain Monte Carlo algorithm to compute posterior inference and prediction. The model is applied to data of one of the largest Italian home care providers, obtaining small prediction errors.

Keywords: Uncertain patients' demands; Home Care; Bayesian model; Multi-state process; Sojourn times, Random-effects model.

1 Introduction

A common feature in planning service delivery in health care facilities is the high uncertainty related to patients' demands. Typically, the number of assisted patients and their demand for care are unknown, and the service delivery must necessarily be dimensioned and planned taking this uncertainty into account. On the other hand, neglecting randomness may have a significant negative impact on the quality and feasibility of the plans, and consequently on the quality of the provided service.

This uncertainty is even more relevant for health services provided outside hospitals, where additional issues arise. For instance, in emergency vehicle location problems, uncertainty can

also be accounted for the availability of ambulances. In Home Care (HC) services, where patients are assisted for a usually longer period than in other facilities and additional planning choices related to service delivery in the territory are required, the impact of random demands is relevant. In particular, when continuity of care is pursued, the assignment of nurses to patients has an impact for a long period (see [1]), and accurate estimations of future patients' demands are fundamental for taking robust nurse-to-patient assignment decisions.

Specifically, HC refers to nursing, medical and social services provided to patients at their own domicile, without the necessity of bringing them to hospitals or nursing homes. Health services provided at home are usually less expensive (hospitalization costs are avoided) and improve patients' clinical, social and psychological conditions (they are cared at home in their familiar context). HC is a relevant sector of the health care domain in Western countries, and it is continuously growing because of the aging of population, the increase in chronic pathologies, the introduction of innovative technologies, and the continuous pressure of governments to contain health care costs. Many resources are involved in delivering HC services in the territory, including nurses, other types of operators, support staff, and also material resources. Appropriate resource planning is thus fundamental for avoiding process inefficiencies, delays, and overloaded operators. In addition, many random events may affect the service delivery and mine the feasibility of plans; see [1, 2, 3, 4]. As mentioned above, the most relevant randomness sources are patients' health conditions, which may determine a different demand for visits than the planned one, as reported in [5]. Hence, reliable estimation tools of future demands for visits required by the patients in the charge represent useful instruments to increase the quality of HC planning and service.

In particular, we are interested in predicting the future number of nurse visits (N) and the Care Profile (CP) of all patients. CP is a categorical variable usually adopted to summarize and represent the patient's health conditions and requirements, which is periodically assessed by a multidisciplinary health team composed by nurses, physicians, and other professional operators. Usually, a revision occurs every month, but the CP can be reassigned in advance in case of sudden variations in patient's conditions. Time is usually divided into discrete slots (e.g., the week or the day), and we are interested in estimating N and CP at each future slot. Here we focus on nurse visits because nurses manage the care pathway of patients and provide the largest number of visits to them.

In this paper, we apply the Bayesian approach to provide estimates, which allows a statistical analysis from a predictive point of view. It is particular useful for HC decision makers who can exploit the entire predictive distribution of each patient's demand and, thus, easily compute the predictive distribution of each nurse's workload and get the predictive probability that, in a future week, a nurse's workload exceeds the weekly nurse's capacity

(i.e., the working time without incurring overtime). We have already addressed this subject in [6], using a regression model with random effects for the number of nurse visits N at each time slot, i.e., for a univariate response, whereas the patient's CP was considered as a fixed covariate. Here, we significantly improve the model considering a bivariate response that includes both nurse visits N and care profile CP of the patients at each time slot. Indeed, we model N given CP as a generalized linear model with fixed and random effects on one hand, and the CP transition process by means of a multi-state process with transitions among visited states governed by a homogeneous Markov chain on the other. The model is applied to a dataset from one of the largest Italian public HC provider. Through a MCMC scheme, we compute posterior inference for all model parameters, and predictive distributions. The goodness-of-fit is also checked.

In the literature, multi-state models represent a useful approach for analyzing categorical longitudinal data, in particular for medical applications, where stages or levels of a disease can be easily represented by the states in the model. They have been used in a wide range of medical applications: see, for instance, [7] (breast cancer), [8] (bronchiolitis obliterans syndrome in lung transplant), [9] (post-heart-transplant cardiac allograft vasculopathy), [10] (Alzheimer's disease), [11] (papilloma virus infection) and [12] (review of frequentist modelling approaches for multi-state models). See also [13] as a reference textbook, and [14] for the description of an R package, called *msm*, to deal with multi-state processes under the frequentist approach.

More recently, multi-state models have been proved fruitful also in the context of resource management for health care facilities. For instance, Blanco [15] proposes a multi-state Markov model to estimate the cost of care provisioning to elderly people in order to help governments in efficiently and effectively allocating resources, whereas Gardiner *et* al. [16] adopt a Markov model to estimate the transition probabilities between health statuses to asses total treatment costs for cancer patients. On the other hand, for a HC dataset similar to the one we analyze here, Lanzarone *et* al. [5] proposes a frequentist approach based on Markov chains associated with a cost probability density function for the number of required visits in each state.

Multi-state processes have been discussed in the Bayesian literature too. Hui *et al.* [17] present Bayesian spatial continuous-time multi-state models for the analysis of geographically referenced event history data; Armero *et al.* [18] use survival analysis and multi-state models to assess survival times for lung cancer patients and the evolution of the disease over time. Other references are [19] and [20]. However, only few papers deal with the Bayesian approach for health care management purposes and, in the HC context, the only available example is our previous paper [6] to the best of our knowledge. Hence, besides the specific application to HC, our aim is to fill this gap and to show the benefits that can derive from applying

Bayesian approaches in this area.

The remainder of this paper is organized as follows. Section 2 describes the type of data at hand and shows the structure of the model. Then, Section 3 describes the dataset and some features of the HC provider supplying the data. In Section 4 we apply the model to the dataset, discussing in particular posterior inference of model parameters, Bayesian goodness-of-fit, prediction for a newly admitted patient, and comparison with the univariate model in [6]. Finally, a discussion and some conclusions are presented in Section 5.

2 Bayesian joint modeling of patient's demand and CP evolution

The first part of this section roughly describes the type of data at hand, in order to understand the model we are going to introduce in the second part.

We consider a sequence of time slots $(t = 1, \ldots, T)$ in which several patients $(i = 1, \ldots, n)$ are assisted, and we denote by $T_L(i)$ and $T_U(i)$ the time slots when patient i enters and exits the service, respectively. Here, each care pathway is entirely contained in the time window, i.e., $T_L(i) \geq 1$ and $T_U(i) \leq T$ for all i. Moreover, we assume that each patient enters and exit only once during his/her care pathway, i.e., we do not consider cases in which a patient is temporarily discharged and enters the service again. Data observed for each patient i at time slot $t \in \{T_L(i), \ldots, T_U(i)\}$ are:

- Number of nurse visits $N_{i,t}$ to patient i at time slot t (count data).

- Care Profile $\text{CP}_{i,t}$ of patient i at time slot t. It is a categorical variable with values in $\{1, 2, \ldots, R\}$.

Moreover, we take into account two covariates in the model:

- $age_{i,t}$: age of patient i at time t, expressed in terms of normalized age as follows:

$$age_{i,t} = \frac{age_{patient_{i,t}} - age_{mean}}{age_{max} - age_{min}}$$

where $age_{patient_{i,t}}$ is the age (in years) of patient i at time t, age_{mean}, age_{min} and age_{max} are the mean, the minimum and the maximum of ages of patients in the dataset at $t = 1$, respectively.

- sex_i: gender of patient i, expressed in terms of a binary variable equal to 0 if male, or 1 if female.

Differently from [6], in this model we consider $\mathrm{CP}_{i,t}$ as a response to be modeled and estimated together with $N_{i,t}$. Therefore, we jointly model the distribution of $\{N_{i,t}, \mathrm{CP}_{i,t}\}_{i,t}$. In particular, we assume that, for each patient i, the transitions between visited CPs are regulated by a multi-state Markov Chain, whereas the holding time (alternatively, sojourn time) in a visited CP state depends on all the CP history up to that time. To be more precise, let $\boldsymbol{\eta_i} = (\eta_{i,1}, \eta_{i,2}, \ldots, \eta_{i,J(i)})$ be the sequence of all the $J(i)$ different categories assumed by the CP history of patient i during the time window. Let also $H_{i,j}$ be the number of times the CP of patient i remains in his/her j-th visited state $\eta_{i,j}$. In this way each patient is characterized by the sequence of visited care profiles $\boldsymbol{\eta_i} = (\eta_{i,1}, \ldots, \eta_{i,J(i)})$, the sequence of holding times $\boldsymbol{H_i} = (H_{i,1}, \ldots, H_{i,J(i)})$, and the sequence of nurse visits $\boldsymbol{N_i} = (N_{i,T_L(i)}, N_{i,T_L(i)+1}, \ldots, N_{i,T_U(i)})$. Observe that the number of components in $\boldsymbol{\eta_i}$ and $\boldsymbol{H_i}$) is different from that in $\boldsymbol{N_i}$. The CP trajectory $\{\mathrm{CP}_{i,t}, t = T_L(i), \ldots, T_U(i)\}$ can be represented by two vectors $\boldsymbol{\eta_i}$ and $\boldsymbol{H_i}$, such that we model $\mathcal{L}(\{\mathrm{CP}_{i,t}\}_t)$ by assigning $\mathcal{L}(\{\eta_{i,j}, H_{i,j}\}_j)$. For each patient $i = 1, \ldots, n$, we assume

(1)
$$\mathcal{L}(\boldsymbol{\eta_i}, \boldsymbol{H_i}) = \mathcal{L}(\eta_{i,1})\mathcal{L}(H_{i,1}|\eta_{i,1})\mathcal{L}(\eta_{i,2}|\eta_{i,1})\mathcal{L}(H_{i,2}|H_{i,1}, \eta_{i,2}) \ldots \mathcal{L}(\eta_{i,J}|\eta_{i,J-1})$$
$$\times \mathcal{L}(H_{i,J}|H_{i,1}, \ldots, H_{i,J-1}, \eta_{i,J}).$$

We also assume conditional independence among patients. The contribution to the likelihood of each patient i is (conditionally to covariates and parameters, not explicitly reported here) as follows:

(2)
$$\begin{aligned}
\mathcal{L}(\{\mathrm{CP}_{i,t}\}_t, \{N_{i,t}\}_t) &= \mathcal{L}(\{\mathrm{CP}_{i,t}\}_t) \times \mathcal{L}(\{N_{i,t}\}_t|\{\mathrm{CP}_{i,t}\}_t) \\
&= \mathcal{L}(\{\mathrm{CP}_{i,t}\}_t) \times \prod_t \mathcal{L}(N_{i,t}|\mathrm{CP}_{i,t})
\end{aligned}$$

where $\mathcal{L}(\{\mathrm{CP}_{i,t}\}_t)$ is the law of the process described in (1) and $\mathcal{L}(N_{i,t}|\mathrm{CP}_{i,t})$ is a generalized linear mixed effects model (GLMM). In particular, we model the number of visits $N_{i,t}$ for patient i at time t as a Poisson distribution with an average rate which depends on the current $\mathrm{CP}_{i,t}$ and covariates, i.e.,

(3)
$$N_{i,t}|\mathrm{CP}_{i,t} = r \sim \mathrm{Pois}(\lambda_r e^{\boldsymbol{x'}_{i,t}\boldsymbol{\gamma}}), \quad r = 1, 2, \ldots, R, \qquad T_L(i) \leq t \leq T_U(i)$$

where $\boldsymbol{x}_{i,t}$ is the covariate vector of patient i at time t ($age_{i,t}$ and sex_i) and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$ is the corresponding regression parameter vector.

As far as the law of the CP trajectories is concerned, we assume that the holding times $H_{i,j}$s are distributed according to a Negative Binomial distribution on $\{1, 2, \ldots\}$, i.e., $\mathrm{NB}(z, q)$:

(4)
$$H_{i,j}|\eta_{i,j}, H_{i,j-1}, \ldots, H_{i,1} \sim \mathrm{NB}(z[\eta_{i,j}], q_{i,j}) \qquad 1 \leq j \leq J(i).$$

Of course, for $j = 1$ the formula above is meant without the conditioning event. Here $\text{NB}(z, q)$ denotes the Negative Binomial distribution with probability of "success" q and "number of successes" z, and its expectation is $1 + r(1 - q)/q$. The $q_{i,j}$s are modelled through a logit regression of the form

$$
\begin{aligned}
\text{logit}(q_{i,1}) &= \log \frac{q_{i,1}}{1 - q_{i,1}} = \beta_1[\eta_{i,1}] \\
\text{logit}(q_{i,j}) &= \log \frac{q_{i,j}}{1 - q_{i,j}} = \beta_1[\eta_{i,j}] + \beta_2(H_{i,1} + \ldots + H_{i,j-1}) \qquad 1 < j \leq J(i).
\end{aligned}
$$

(5)

Note that, because of (5), we are assuming that $q_{i,j}$ depends on the current value of patient's CP, as well as on the time the patient has spent in the service from admission to the last CP change (i.e., the change from $\eta_{i,j-1}$ to $\eta_{i,j}$). Parameters $\beta_{1,r} = \beta_1[\eta_{i,j} = r]$ and $z_r = z[\eta_{i,j} = r]$ describe the random effects of the patient as a function of the current $\eta_{i,j}$, λ_r represents the random effect (health status) of a patient with CP$= r$, whereas β_2, γ_1, γ_2 are fixed-effects parameters. Moreover, we have assumed that holding times of each patient are not independent, but depend on all the previous ones.

As far as modelling of $\boldsymbol{\eta_i}$ is concerned, the visited CPs are described by a (conditionally) homogeneous Markov chain, with states $\{1, \ldots, R + 1\}$. States from 1 to R correspond to CP categories, whereas the last state $R + 1$ represents an exit state from the service (if the patient dies, or leaves the service for any different reason). Obviously, $R + 1$ is an absorbing state. By $\mathbf{P} = [P_{r,s}]$ we denote a $(R + 1) \times (R + 1)$ matrix, and $P_{r,s}$ is the probability that the visited care profile (at any time) moves from state r to s. We assume the initial state $\eta_{i,1}$ to have a discrete (categorical) distribution with weights $(\pi_1, \pi_2, \ldots, \pi_R, \pi_{R+1})$ where $\pi_r \geq 0$ $\forall r$ and $\sum_{r=1}^{R+1} \pi_r = 1$.

As far as the prior is concerned, all parameters $\boldsymbol{P}, \gamma_1, \gamma_2, \beta_2, (\lambda_1, \ldots, \lambda_R), (\beta_{1,1}, \ldots, \beta_{1,R}), (z_1, \ldots, z_R)$ are assumed a priori (conditionally) independent. In particular, we assume the rows of the transition matrix $\boldsymbol{P}$ to be independent, being each row

(6) $\qquad (P_{r,1}, \ldots, P_{r,R+1}) \sim \text{Dirichlet}(a_1, \ldots, a_{R+1}) \qquad r = 1, \ldots, R,$

where $\text{Dirichlet}(\cdot)$ denotes the (absolute continuous) Dirichlet distribution on $S_R = \{(x_1, \ldots, x_R) : x_j \in [0, 1] \ \forall j, 0 < x_1 + \ldots + x_R < 1\}$. Self transitions are not allowed $(P_{r,s} = 0, r = s)$ as this agrees with the definition of holding times between state changes, whereas $P_{R+1,R+1} = 1$ almost surely. For the rest of the parameters, we make standard assumptions: the marginal priors for the random effects are assumed exchangeable, according to the Bayesian hierarchal approach, and the fixed effects have weakly informative marginal priors, i.e., in this case, Gaussian distributions, centered at 0, with large variance. Specifically,

we assume:

(7)
$$\log(\lambda_r)|\mu_\lambda, \sigma_\lambda^2 \overset{\text{iid}}{\sim} N(\mu_\lambda, \sigma_\lambda^2), \ r = 1, 2, \ldots, R,$$

$$\mu_\lambda \sim N(0, 100), \ \sigma_\lambda \sim U(0, 10), \ \mu_\lambda, \sigma_\lambda \text{ independent},$$

(8)
$$\beta_{1,r}|\sigma_{\beta_1}^2 \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\beta_1}^2), \ r = 1, \ldots, R, \quad \sigma_{\beta_1} \sim U(0, 10)$$

(9)
$$z_r \overset{\text{iid}}{\sim} Gamma(2, 2), \ r = 1, \ldots, R$$

(10)
$$\gamma_1, \gamma_2, \beta_2 \overset{\text{iid}}{\sim} \mathcal{N}(0, 1000).$$

Marginal uniform priors for the standard deviation parameters as in (7) or (8) represent a reasonable choice in Bayesian hierarchical models, when a weakly informative prior is used, as in our case (see [21]). Further details on the hyperparameters above will be given in Section 4.

All inference is based on the posterior distribution of $\boldsymbol{\theta} = (\gamma_1, \gamma_2, \beta_2, \mu_\lambda, \sigma_\lambda^2, (\lambda_1, \ldots, \lambda_R),$ $(P_{r,1}, \ldots, P_{r,R}), (\beta_{1,1}, \ldots, \beta_{1,R}), (z_1, \ldots, z_R))$, given $\boldsymbol{N}, \mathbf{CP}$ and the covariates:

$$\pi(\boldsymbol{\theta}|\boldsymbol{N}, \mathbf{CP}, covariate) \propto \pi(\boldsymbol{\theta})Lik(\boldsymbol{\theta}),$$

where the joint likelihood $Lik(\boldsymbol{\theta})$ can be recovered from (1)-(2) as

$$Lik(\boldsymbol{\theta}) = \prod_{i=1}^{n} \Big\{ \mathcal{L}(\eta_{i,1})\mathcal{L}(H_{i,1}|\eta_{i,1})$$
$$\times \prod_{t=T_{L(i)}}^{T_{U(i)}} \mathcal{L}(N_{i,t}|\mathrm{CP}_{i,t}) \prod_{j=2}^{J(i)} \mathcal{L}(H_{i,j}|\eta_{i,j}, H_{i,j-1}, \ldots, H_{i,1})\mathcal{L}(\eta_{i,j}|\eta_{i,j-1}) \Big\}.$$

Of course, we resort to Markov Chain Monte Carlo (MCMC) algorithms to compute the posterior distribution.

In this study, our final goal is to predict the demand for visits at future time slots, given covariates and data via predictive distributions. This is very important for HC decision makers, who are interested in planning the service and assigning nurses to patient over a future planning horizon to improve service efficiency. If we observe the care pathway of a patient until time t, $\{\mathrm{CP}^*(\tau), N^*(\tau), T_L \leq \tau \leq t\}$, then his/her predictive distribution is:

(11) $\quad \mathcal{L}(N^*(t+1), \mathrm{CP}^*(t+1)|\boldsymbol{x}^*, \boldsymbol{N}, \mathbf{CP}) =$
$$= \int \mathcal{L}(N^*(t+1)|\mathrm{CP}^*(t+1)) \mathcal{L}(\mathrm{CP}^*(t+1)|\{\mathrm{CP}^*\}, \boldsymbol{\theta}) \pi(\mathrm{d}\boldsymbol{\theta}|\boldsymbol{N}, \mathbf{CP}, \boldsymbol{x}^*)$$

where $\boldsymbol{x}^*$ is the covariate vector, and $N^*(t+1)$ and $\mathrm{CP}^*(t+1)$ are the number of nurse visits and the CP of the patient at time $t+1$, respectively. Evaluation of (11) is usually achieved through the MCMC strategy. In particular, in this work, differently from Argiento *et al.* [6], there is no need to condition on an hypothetical "future" trajectory of the care profile path, since here this is part of our response variable and can be predicted according to the model.

Furthermore, we are able to compute also the posterior predictive distribution for a newly admitted patient; in this case, differently for (11), no information is available from previous time slots. Let us denote by i^* the new patient with covariate vector x^* and care profile trajectory $\mathbf{CP}_{i^*}$; then the predictive distribution of i^* is computed integrating the conditional joint distribution of $\boldsymbol{N}_{i^*} = (N_{i^*,T_{L(i^*)}}, \ldots, N_{i^*,T_{U(i^*)}})$ and $\mathbf{CP}_{i^*} = (CP_{i^*,T_{L(i^*)}}, \ldots, CP_{i^*,T_{U(i^*)}})$ as in (3)-(5) with covariate x^*, with respect to the posterior distribution of $\boldsymbol{\theta}$.

3 Exploratory analysis of the HC dataset

We consider data from one of the largest Italian public Home Care providers; data from this provider have already been analyzed according to frequentist [5] and Bayesian approaches [6]. This provider operates in the north of Italy, covering a region of about 800 km^2, with about 1000 patients assisted at the same time [2, 3]. Moreover, the human resource organization and the patient classification adopted by this provider can be considered general and common to several other HC providers as underlined by Matta *et al.* [1], so that it is representative of a general class of providers. The provider is divided into three divisions, and the analysis refers to the largest one.

Patients assisted by this HC provider are mainly clustered into two groups: *palliative* and non-palliative patients. Non-palliative patients are further divided into two groups, denoted by *extemporary care* (with very low frequency of visits) and *integrated care* (with higher frequency). Each class is then divided into CPs based on the care intensity required by the patients. CPs related to palliative care refer to a homogeneous class of terminal patients whose pathology is in a terminal state. On the other hand, for non–palliative care, each CP includes a wide range of patients in terms of age, pathology and social context; however, patients belonging to the same CP are characterized by similar therapeutic projects with similar levels of demands. With respect to the classification adopted by the HC provider, we have regrouped very similar CPs, thus reducing the number to 9, as in Table 1. As mentioned before, we have added one additional state (CP=10) for patients exited from the service, which is absorbing.

The time slot here is the week. In fact, for the provider analyzed here, as for many others, the assignment phase of the planning process is carried out over a weekly basis.

3.1 Dataset description

We started the analysis considering the same dataset as in [5]. The time horizon is pretty wide, i.e., 252 weeks long, from January 2004 to March 2008. We filtered the dataset, considering only patients in the provider's largest division who entered and exited the service only

Table 1: Classification of CPs and associated patients and pathologies. CPs within each category are listed in increasing order of complexity and expected demand for visits.

Types of care	Associated pathologies	CP
Extemporary Care	Heterogeneous class of patients assisted by HC even without a specific pathology, who require generic nursing and medical assistance: patients with the same CP are characterised by similar demands with low frequency of visits	1 9
Integrated Care	Heterogeneous class of non-palliative patients with different pathologies: patients with the same CP are characterised by similar therapeutic projects with high frequency of visits	8 7 2 3 4 5
Palliative Care	Homogeneous class of terminal patients generally affected by oncological diseases	6

once within the observed period without any interruption of the service (e.g., hospitalization periods with an interruption of the HC assistance), and whose care pathway is entirely contained in the whole time window.

In this way, we got a large dataset consisting in $n = 2358$ patients with an overall number of (bivariate) observations between $T_L(i)$ and $T_U(i)$ equal to 34390. The dataset includes 1006 men (43%) and 1352 women (57%). The age (at the entrance in the service) ranges from a minimum of 1 year to a maximum of 101 years, while the empirical mean and standard deviation are 73.74 and 14.43 years; empirical means and standard deviations for male and female patients are 70.60, 14.53, and 76.08, 13.90, respectively. The boxplots of age grouped by gender, not included here, give evidence that age and gender are not independent. The overall average number of nurse visits (i.e., the ratio between 42259, the sum of all numbers of weekly visits, and 34390 observations) is equal to 1.23. To give insight to these data, the total number of observations between $T_L(i)$ and $T_U(i)$ and the average number of visits, grouped by CP, age and gender, are reported in Table 2 and Table 3, respectively. Moreover, sample histograms of the holding times $H_{i,j}$ grouped by $\eta_{i,j} = r$ (with $r = 1, \ldots, 9$) are depicted in Figure 1. From the figure we see that most of the histograms seem heavily skewed right.

Finally, we summarize the multi-state data $(\eta_{ij})_{i,j}$ as a frequency table of pairs of consecutive states (Table 4): this counts over all individuals, for each state r and s, the number

Table 2: Number of observations and average number of visits grouped by CP.

group of HC	Extemporary Care		Integrated Care						Palliative Care
	CP = 1	CP = 9	CP = 8	CP = 7	CP = 2	CP = 3	CP = 4	CP = 5	CP = 6
total no. of obs	9606	963	776	7061	1032	2563	2631	3706	6052
aver. no. of visits	0.45	0.89	0.37	0.83	1.38	1.37	1.19	2.36	2.34

Table 3: Number of observations and average number of visits grouped by age and gender.

age	men		women	
	no. of obs	average no. of visits	no. of obs	average no. of visits
≤ 50	949	1.01	600	1.55
$(50, 60]$	1246	1.12	1203	1.07
$(60, 70]$	2803	1.22	2683	1.65
$(70, 80]$	4047	1.27	5925	1.29
$(80, 90]$	4361	0.97	7571	1.12
> 90	781	0.81	2217	1.65

of times the care profile moved from state r (at any time t) to state s (at time $t + 1$).

4 Bayesian inference for the HC dataset

As we mentioned in Section 2, a weakly informative prior for hierarchical standard deviation parameters (here σ_λ and σ_{β_1}) is the uniform on a bounded large interval $(0, \sigma_{max})$. Here we report the inference when $\sigma_{max} = 10$, but we checked that we got the same posterior distribution as when $\sigma_{max} = 100$. In both cases, the marginal posterior distributions of σ_λ and σ_{β_1} are concentrated on values much smaller that 10. On the other hand, we set the marginal prior expectation $\mathbb{E}(z_r)$ for the random effect z_r equal to 1, since this corresponds to the Geometric distribution for the holding times, which has the memoryless property, as it is well known. The prior variance $\text{Var}(z_r) = 0.5$ corresponds to a prior standard deviation equal to $1/\sqrt{2}$ (neither too small nor too large). In addition, under no further prior information, we set the hyperparameters of the distribution of the initial CP state $\pi_r = 1/(R + 1)$ and all a_r in (6) equal to 1.

We analyze the posterior distribution of the parameter vectors via point estimates and credible intervals. Moreover, Bayesian prediction for patients already in charge and for a
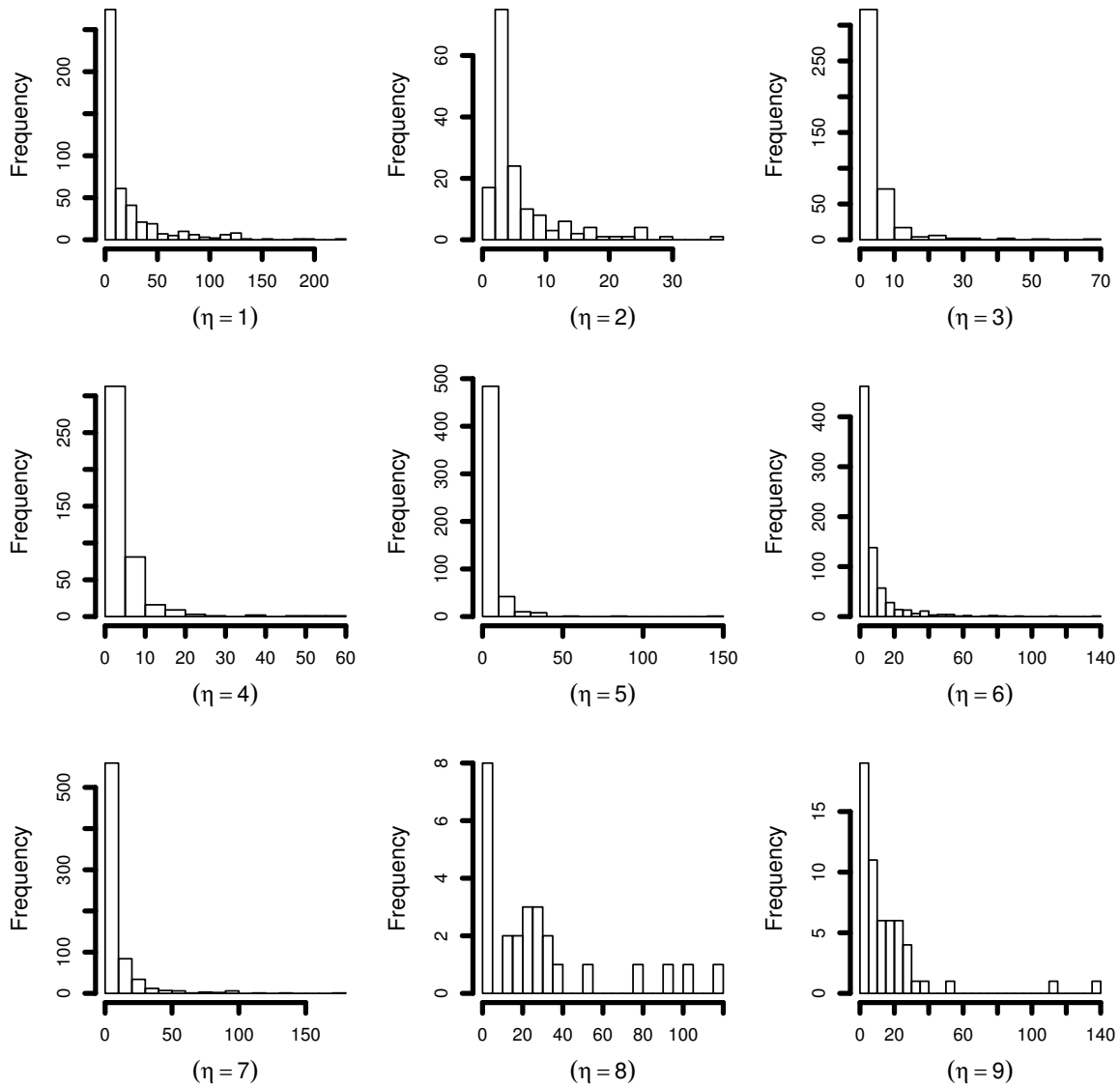
Figure 1: Histograms of holding times grouped by care profile η.

newly admitted patient are presented here. The posterior inference of the model parameters is reported in Section 4.1, predictive goodness-of-fit in Section 4.2, predictions for newly admitted patients and the comparison with model in [6] are shown in Section 4.3 and Section 4.4, respectively.

4.1 Bayesian inference of the parameters

In order to compute the Bayesian estimates, the model was implemented in Jags [22], with the support of R [23], with chains consisting of 255000 iterations with a burn-in of 5000 and a

Table 4: Number of observed transitions between CP states.

r \ s	1	2	3	4	5	6	7	8	9	10
1	0	4	18	10	12	16	13	0	2	393
2	5	0	6	3	4	5	53	3	0	79
3	12	31	0	30	20	3	160	1	1	170
4	8	32	93	0	22	6	133	2	0	132
5	3	15	97	125	0	11	101	0	1	194
6	2	0	0	0	1	0	0	0	0	746
7	42	12	35	15	13	10	0	15	0	575
8	4	0	2	0	2	1	2	0	0	15
9	1	0	0	1	0	1	0	0	0	54

thinning of 50 iterations, yielding a final sample size of 5000 iterations. Standard convergence diagnostics in the CODA package (see [24]) were checked. Moreover, we monitored traceplots, autocorrelations, and bivariate scatterplots for all the parameters, indicating that the MCMC algorithm converged. Code is available from the authors upon request.

To give an indication of the mixing of the chain, Figure 2 displays traceplots of the $\beta_{1,r}$ parameters. Figure 3 displays the 95% posterior credibility intervals (with posterior medians) of random-effects parameters $z_r, \beta_{1,r}$ and $\log(\lambda_r)$, $r = 1, \ldots, 9$. Its clear from Figure 3 that the random-effects parameters are significantly different with respect to the care profile. In particular, we see that the holding times $H_{i,j}$ strongly depend on patient's CP through z_r and $\beta_{1,r}$. Moreover, from the model we have assumed we expect that the larger $\beta_{1,r}$ is, the smaller $\mathbb{E}(H_{ij})$ is; this is exactly the feature that we can see from the figure, where the highest CIs for $\beta_{1,r}$ are obtained for $r = 2,3,4,5$, which are exactly the groups corresponding to smaller sample means. Furthermore, we remark that larger credibility intervals in Figure 3 are generally obtained for profiles with a smaller number of observations, as, for instance, when CP= 8 and CP= 9. Despite we have already displayed posterior CIs for the z_r parameters, in Figure 4 we display the marginal posterior distributions, together with the marginal prior (gamma(2,2) distribution). It is apparent that the posterior marginal distributions of the z_rs are pretty different, and also different from the prior. Therefore, it has been reasonable to model z as a random effect parameters (i.e. considering 9 different parameters $z_1, \ldots, z_9$). Note that the posterior density of z_1 and z_6 has significantly flatter tails. We also report posterior means and standard deviations of the patients health statuses $\lambda_1, \ldots, \lambda_9$ in the
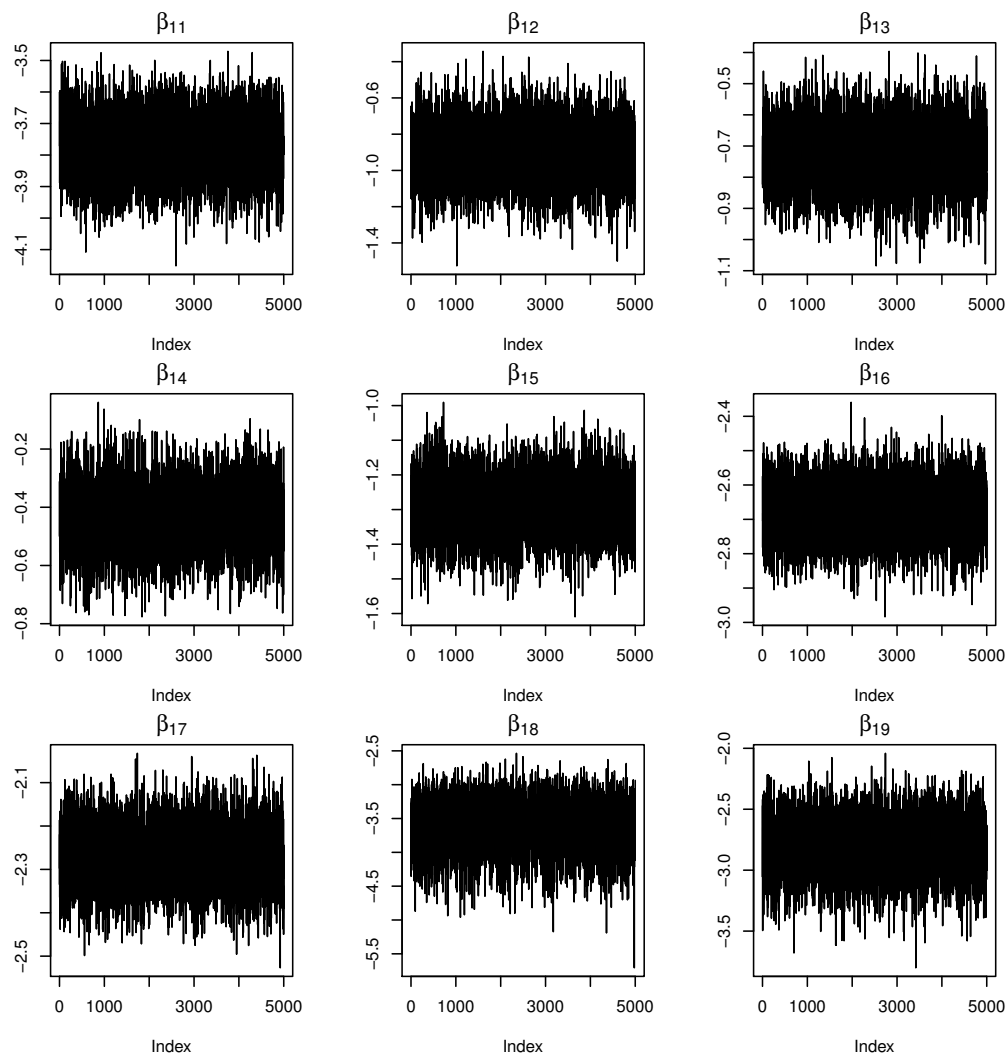
Figure 2: Traceplots of MCMC sampled values of $\beta_{1,r}$, for $r = 1, \ldots, 9$.
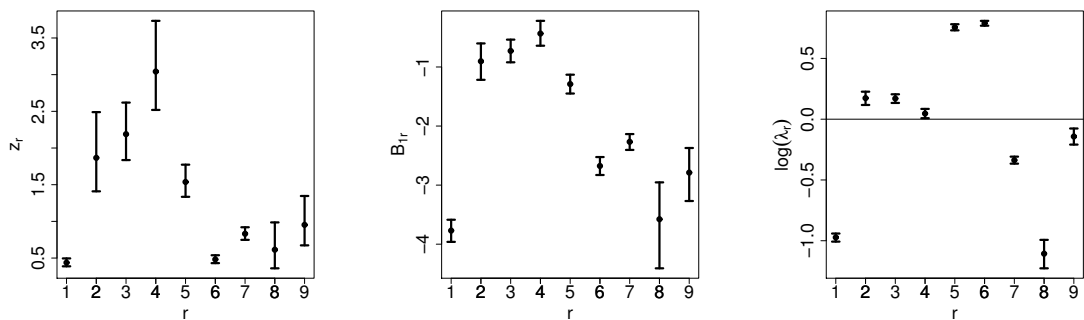

Figure 3: Posterior 95% credible intervals and medians of $z_r, \beta_{1,r}$ and $\log(\lambda_r)$, for $r = 1, \ldots, 9$.
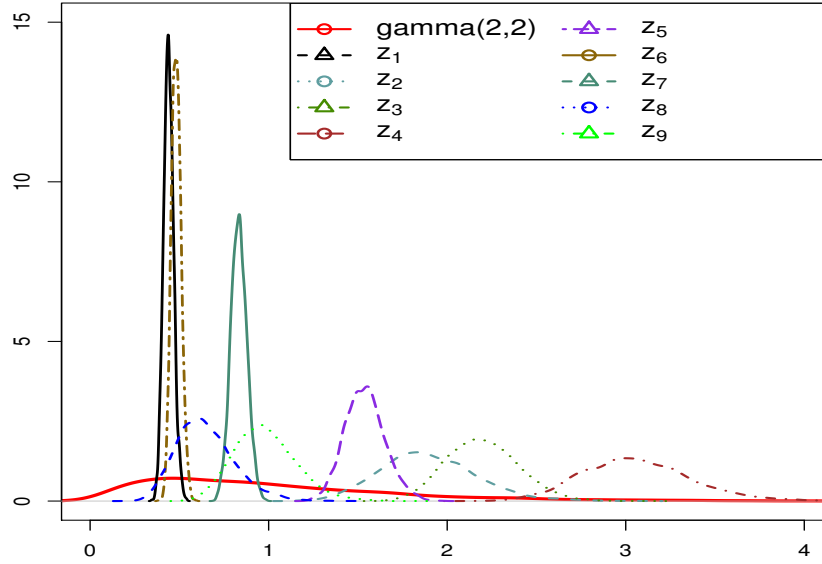
Figure 4: Marginal posterior distributions of z_r, $r = 1, \ldots, 9$; gamma(2,2) is the marginal prior of any z_r.

original scale (see Table 5). Once again, by looking at the values in the table, it is clear that our estimates reflect the knowledge inherent to the data itself. In fact, since the largest value of the posterior means of the λ_r parameters is obtained by the palliative group (λ_6), these patients manifest the worst health status and require the highest number of visits. In contrast, extemporary care patients correspond to the smallest number of visits and the best health status, while integrated care patients have intermediate needs.

As far as fixed effects are considered, Table 6 reports posterior quantiles of γ_1, γ_2 and β_2. The marginal posterior distributions of γ_1 and γ_2 are mostly constrained on positive values, yielding that the demand of visits increases with age and it is larger for female patients (the

Table 5: Posterior means and standard deviations of $\lambda_r, r = 1, \ldots, 9$; the estimates are listed according the increasing order within each category as reported in Table 1).

group of HC	Extemporary Care		Integrated Care						Palliative Care
	λ_1	λ_9	λ_8	λ_7	λ_2	λ_3	λ_4	λ_5	λ_6
Mean	0.378	0.868	0.331	0.714	1.189	1.185	1.047	2.129	2.203
sd	0.006	0.030	0.020	0.011	0.033	0.021	0.020	.027	0.022

Table 6: Posterior quantiles of the fixed-effects parameters.

	2.5%	50%	97.5%
(age) γ_1	0.906	0.980	1.057
(gender) γ_2	0.098	0.119	0.138
(past time) β_2	-0.012	-0.010	-0.008

Table 7: Posterior means of each element in the transition probability matrix $\mathbf{P}$.

r \ s	1	2	3	4	5	6	7	8	9	10
1	0.0000	0.0105	0.0398	0.0230	0.0274	0.0356	0.0293	0.0021	0.0063	0.8261
2	0.0359	0.0000	0.0417	0.0237	0.0298	0.0357	0.3231	0.0241	0.0059	0.4802
3	0.0300	0.0734	0.0000	0.0707	0.0478	0.0091	0.3682	0.0046	0.0046	0.3916
4	0.0207	0.0758	0.2149	0.0000	0.0525	0.0159	0.3065	0.0068	0.0024	0.3046
5	0.0073	0.0288	0.1765	0.2268	0.0000	0.0215	0.1835	0.0018	0.0036	0.3503
6	0.0039	0.0013	0.0013	0.0013	0.0026	0.0000	0.0013	0.0013	0.0013	0.9855
7	0.0592	0.0180	0.0495	0.0221	0.0192	0.0151	0.0000	0.0220	0.0014	0.7936
8	0.1429	0.0283	0.0863	0.0287	0.0860	0.0583	0.0859	0.0000	0.0282	0.4554
9	0.0302	0.0152	0.0151	0.0304	0.0154	0.0303	0.0154	0.0152	0.0000	0.8328

women are 57% of our dataset). On the other hand β_2 is a posteriori constrained on negative values i.e., the holding time at j increases when the summation of all holding times until $(j-1)$ decreases.

Now, let us make some comments on the Bayesian estimates of the transition probability matrix $\mathbf{P}$. Table 7 reports posterior means of all $P_{r,s}$. It shows that palliative patients (CP= 6) leave the service with higher probability than non palliative patients, in agreement with empirical evidence (see Table 4); of course, estimates of $P_{r,r}, r = 1, \ldots, 9$ are zero, since they are zero a priori. Except for the last column, representing probabilities of exiting the study, the transition probability estimates do not seem to strongly depend on CPs. Moreover, extemporary patients (CP= 1, 9) exit the service with high estimated probabilities. To gain additional insight into the transition matrix, 95% credibility intervals for each row of the transition matrix $\mathbf{P}$ are reported in nine panels in Figure 5. It is clear that all palliative patients will leave the service sooner or later, since the posterior variance is quite small and the mean is fairly high (0.99); indeed our dataset shows very strong empirical evidence that palliative patients died (if they changed CP). Once again, larger variability in the estimates
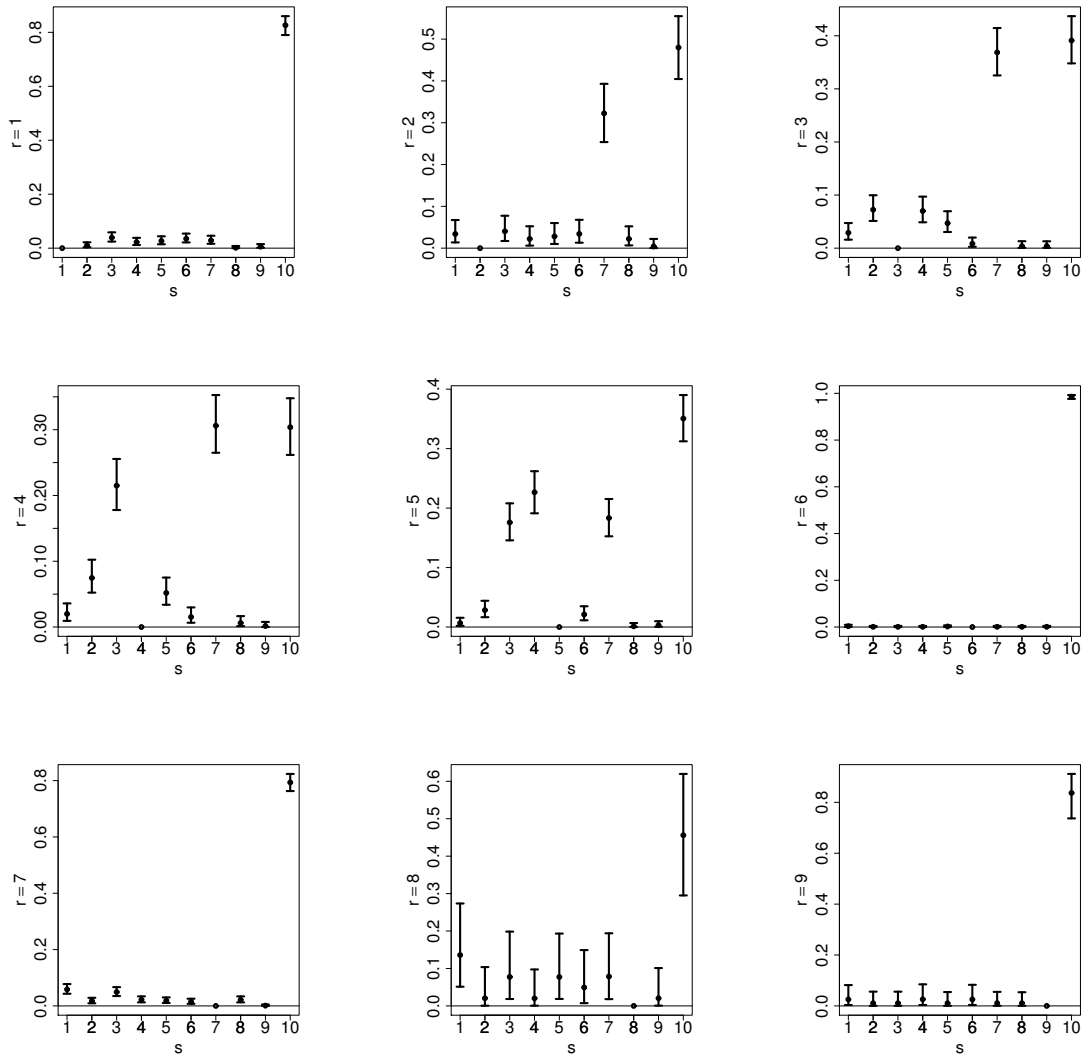
Figure 5: 95% posterior CIs for $(P_{r,s}, s = 1, \ldots, 10)$, $r = 1, \ldots, 9$. The label on the y-axis on each panel identifies r.

holds for CP=8, and for CP=9 to a smaller extent, since these groups are smaller. As regards the rest of integrated care CPs (R =7,2,3,4,5), the figure shows that there is a strong tendency to move from a visited CP to the next one.

4.2 Bayesian goodness of fit

Some graphical and numerical tools for assessing the goodness of fit of the Bayesian model considered are presented here. In this section we adopt a predictive point of view, focusing our attention on patients who are in charge at a certain time t and predicting their number

Table 8: MAE of the number of visits at successive weeks.

	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$t+6$	$t+7$	$t+8$
MAE_{99}	1.20	1.13	1.12	1.11	1.08	1.06	1.06	1.05
MAE_{149}	1.01	1.05	1.06	0.95	0.95	0.94	0.94	0.93
MAE_{175}	1.12	1.12	1.13	1.09	1.10	1.10	1.10	1.08
MAE_{234}	1.00	0.98	1.07	0.97	0.98	0.97	0.95	0.95

of nurse visits and care profiles at time $t + 1$.

For goodness-of-fit purposes only, we divided the dataset into a training set and a testing set according to a predictive cross-validation approach. Patients who are in charge of the provider at week t are in the testing set, whereas all the others are in the training set. We computed the posterior densities of model parameters again, considering only the training set as "data". Then, we computed the predictive distributions as in (11) for patients in the testing set by means of the posterior of parameter $\boldsymbol{\theta}$ obtained under the training set. Finally, we checked predictions and observed data of the testing sample. This validation procedure was applied at some of the 252 weeks in the time window. For each of them, we computed the joint predictive distribution of the number of nurse visits and the care profile for all the patients in the charge at that week.

As in [6], the accuracy of the prediction for nurse visits was evaluated in terms of the mean absolute error (MAE), that is

$$(12) \qquad MAE_{t+1} = \frac{\sum_{i=1}^{m_t} |n_{i,t+1} - \hat{N}_{i,t+1}|}{m_t},$$

where m_t is the number of active patients at week t, $n_{i,t+1}$ is the observed number of nurse visits at time $t + 1$ and $\hat{N}_{i,t+1}$ is the Bayesian prediction of the nurse visits for each patients at week $t + 1$. Here, $\hat{N}_{i,t+1}$ is taken as the mean of the predictive distribution of $N_{i,t+1}$. The lowest is the MAE_{t+1}, the highest is the accuracy of our prediction at time $t+1$. We computed MAE as in (12) at four different weeks $t = 99$, $t = 149$, $t = 175$ and $t = 234$, where the number of patients in the charge at those weeks are $m_{99} = 158$, $m_{149} = 143$, $m_{175} = 165$ and $m_{234} = 108$, respectively; Table 8 displays our estimates. In order to calibrate the values we obtain, we also computed the empirical MAE as in (12) substituting the Bayesian prediction $\hat{N}_{i,t+1}$ with $n^*_{i,t+1}$, the sample average of the number of nurse visits for patients in the training set with care profile equal to $CP_{i,t}$; these values are 1.29, 1.01, 1.15, 0.88 for week 100, 150, 176, 235, respectively. Hence, comparing values of MAE in the first column in Table 8 to these above, we conclude that all the values are quite close, so that the model show a rather good fit to the data.

To assess the accuracy of the estimates and to check the presence of asymmetric errors in the predictions at week $t + 1$, we also plot the differences $n_{i,t+1} - \hat{N}_{i,t+1}$ for each patient in the testing sets. Figure 6 shows small errors, but our predictive estimates are mostly
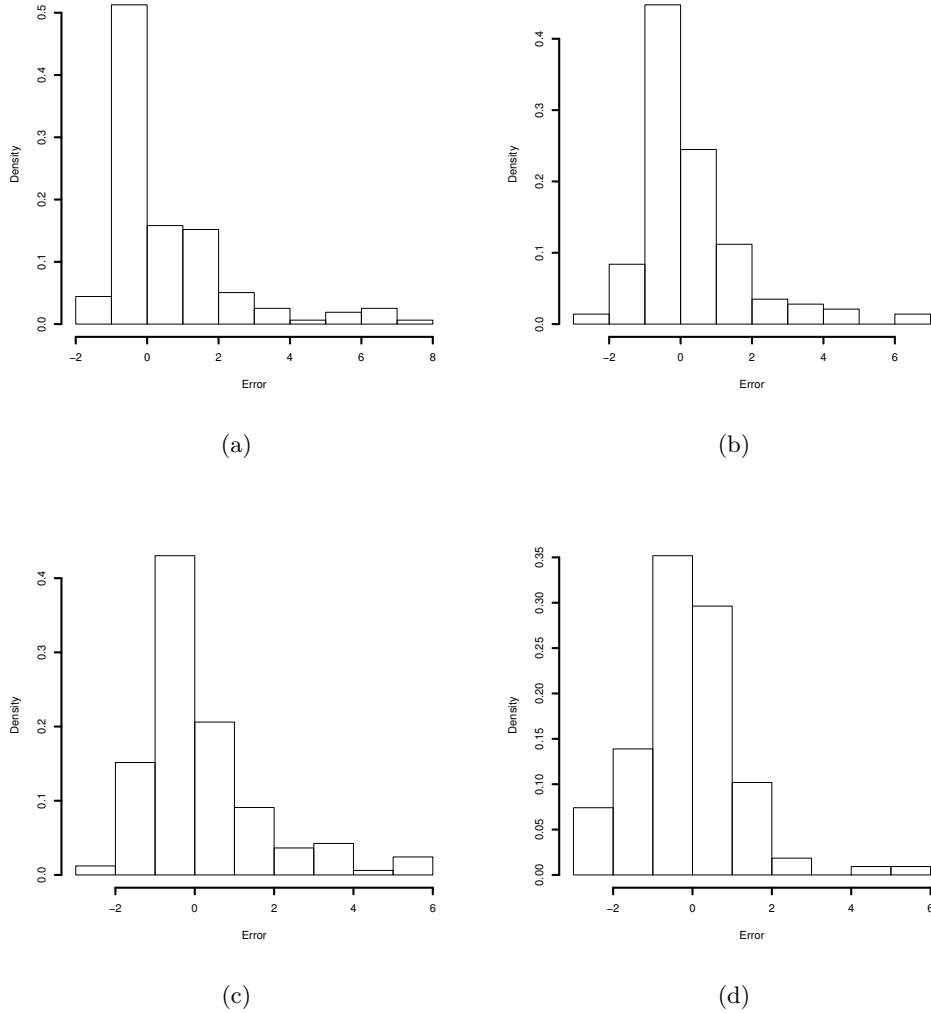

Figure 6: Sample histograms of the errors for predictions of nurse visits at $t + 1 = 100$ (a), $t + 1 = 150$ (b), $t + 1 = 176$ (c) and $t + 1 = 235$ (d).

overestimating the number of visits effectively administered to the patients.

Now, let us turn our attention to the prediction of the care profile at $t + 1 = 100$, 150, 176 and 235. In this case, since CP is a categorical variable, the Bayesian prediction we considered is the predictive mode. Table 9 displays Bayesian prediction of the care profile at week $t + 1$ for those patients who were in the study at week t. The prediction is displayed for patients within the three groups of care profile. For example, at week $t = 99$ we consider

158 patients, 41 of them with CPs in the extemporary care group, 86 in the integrated care group and 31 palliative (CP=6). Our prediction at $t = 100$ yields that 38 patients remain in the extemporary care group, while the observed frequency is 40, and 3 patients leave the service (while the observed frequency is 1). Analogously, 46 of the integrated patients are predicted by our model to remain in the same group although the observed frequency is 77, while, out of 31 of palliative patients, we expect that 11 of them to remain palliative (the observed frequency is 28) and 20 to exit the service (the observed frequency is 3). In all, these predictions seem accurate.

In addition, we have taken into account computation of posterior predictive p-values for our model, which are among the most popular tools to assess goodness-of-fit in the Bayesian context; specifically,

$$p - \text{value}_{i,t+1} = \min(P(N_{i,t+1}^{new} > n_{i,t+1}|data), P(N_{i,t+1}^{new} \leq n_{i,t+1}|data)),$$

defined in terms of the predictive distribution of $N_{i,t+1}^{new}$, where $N_{i,t+1}^{new}$ is the i-th "replicated data" (see [25], Section 6.3). An extremely small value of the Bayesian $p - value$ indicates that the data are unlikely under the model. However we do not report here these values, since they would not give further insight. We would like only mentioning that predictive $p - values$ of the model seem to be uniformly distributed on to $(0, 0.5)$, which indicates a good fit of our Bayesian model (for more details, see [26]).

4.3 Bayesian prediction for a newly admitted patient

As a second goal of Bayesian prediction, the interest in this subsection is forecast of the number of nurse visits required to a new patient just admitted into the service. In particular we follow the approach described at the end of Section 2, for a new female patient, who is 77 (i.e. the overall sample mean) years old at the first week of the study. We simulated the whole trajectory of this patient, i.e. $N_{i^*,t}$ and $CP_{i^*,t}$ $t = 1, \ldots, T_U(i^*)$. In Figure 7 we display the posterior predictive probabilities of $N_{i^*,t}$ and $CP_{i^*,t}$ when $t =$4, 8 and 12. From the right column in the figure, we see that low intensity profiles 1,7,8,9 have predictive masses which do not change over time, while the rest of the predictive masses decreases in time in favor of the exit state (CP=10). This means that low intensity profile patients have sojourn times larger that at least 12 weeks, whereas higher intensity profile patients show lower sojourn times. On the other hand, the predictive distributions of the number of visits do not significantly change over time.

Table 9: Comparison between observed and predicted values of CP within the three groups of home care for all patients in the charge at week t.

			$t+1 = 100$, $m_{99} = 158$							
			Extemporary		Integrated		Palliative		Exit	
			obs	pred	obs	pred	obs	pred	obs	pred
	Extemporary	41	40	38	0	0	0	0	1	3
obs at $t=99$	Integrated	86	1	0	77	46	1	0	7	40
	Palliative	31	0	0	0	0	28	11	3	20

			$t+1 = 150$, $m_{149} = 143$							
			Extemporary		Integrated		Palliative		Exit	
			obs	pred	obs	pred	obs	pred	obs	pred
	Extemporary	42	39	41	1	0	0	0	2	1
obs at $t=149$	Integrated	74	0	0	70	48	0	0	4	26
	Palliative	27	1	0	0	0	21	16	5	11

			$t+1 = 176$, $m_{175} = 165$							
			Extemporary		Integrated		Palliative		Exit	
			obs	pred	obs	pred	obs	pred	obs	pred
	Extemporary	41	39	39	0	0	0	0	2	2
obs at $t=175$	Integrated	101	0	0	92	60	0	0	9	41
	Palliative	23	0	0	0	0	21	12	2	11

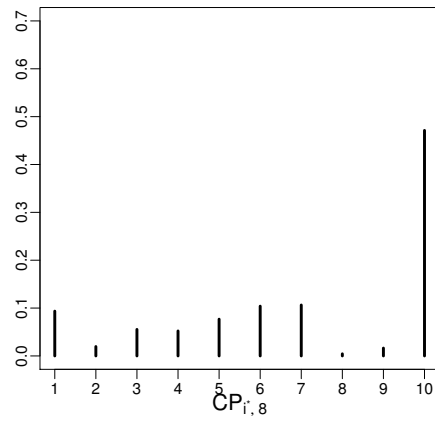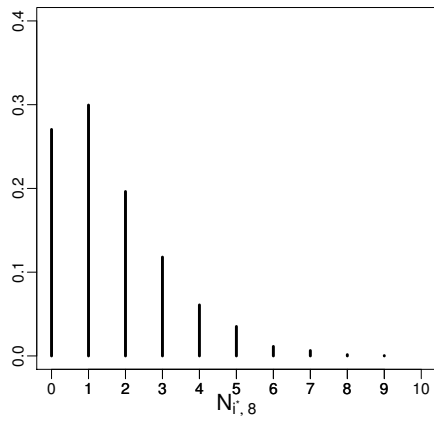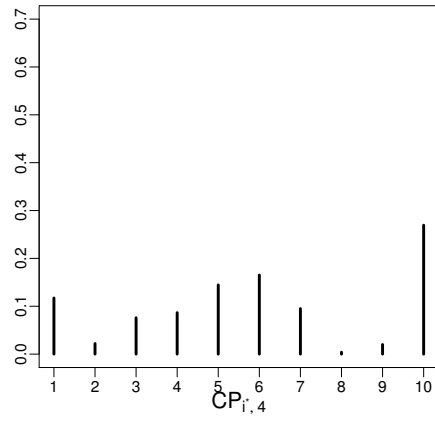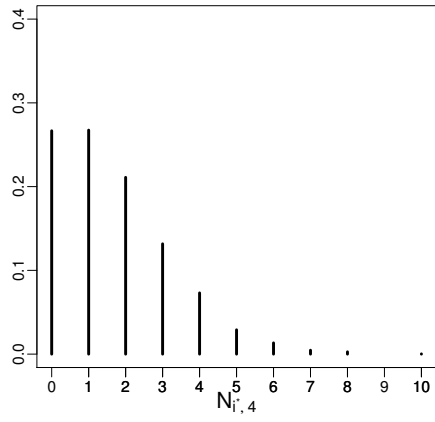|			$t+1 = 235$, $m_{234} = 108$							
			Extemporary		Integrated		Palliative		Exit	
			obs	pred	obs	pred	obs	pred	obs	pred
	Extemporary	38	35	38	0	0	0	0	3	0
obs at $t=234$	Integrated	52	1	0	43	46	0	0	8	6
	Palliative	18	0	0	0	0	15	17	3	1
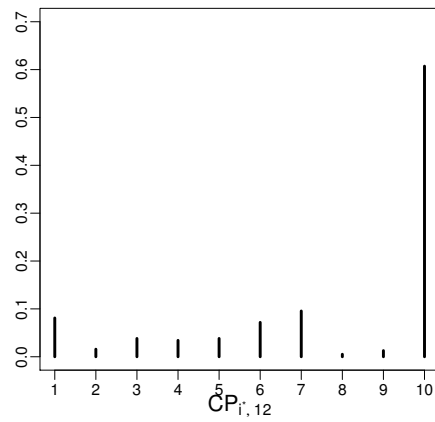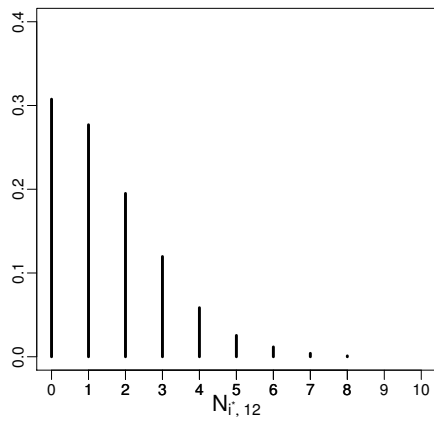
Figure 7: Posterior predictive probabilities of the number of nurse visits and care profile at some weeks for a new patient.

4.4 Comparison with the model in Argiento *et al.* [6]

In this section we aim at making a comparison between the model presented here and that in Argiento *et* al. [6]; they both are Bayesian models. Of course, the first difference between the two is the dimension of the response: in this work, we model jointly visits demand and evolution of care profile of the patient, while in [6] it is only patient's demand which is accounted for by the model, with the assumption that patient's care profile is known all along the time window. Therefore the model here allows a higher degree of flexibility. Moreover, note that in [6] the mean $\lambda_{i,t}$ of the number of visits $N_{i,t}$ has an autoregression formulation, while here it depends only on the patient's current care profile which, however, is a response variable itself. Further, Argiento *et al.* [6] considered the age of each patient at the beginning of the study as a fixed covariate, while here we consider it as a time-varying covariate.

In order to compare goodness-of-fit of the two models, we run the model in [6] with the same data we analyzed here, and compared values of MAE of nurse visits as in Table 8, for eight next weeks from t, when, as before, $t =$99, 149, 175 and 234. Figure 8 displays comparison among the values of MAEs under the two models. The two lines in all panels are quite close each other till $t + 4$, but then the MAEs under the model here are smaller and seem slightly decreasing in time, whereas the others rapidly increase. This points out that the present model is fairly efficient on the long-term prediction.

5 Discussion

In this paper, we have analyzed health profile and demand for home care patients in Italy. We have proposed a bivariate Bayesian model to represent the evolution in (discrete) time of the number of nurse visits and care profile of patients in the service. We have computed estimates of the parameters, as well as prediction of either a new patient, or patients already in the charge. Our final aim is helping the HC decision makers to organize nurses workload through a whole probability distribution, so that they could be able to compute the (predictive) probability that, in a future week, nurses do not have to work overtime (consequently yielding higher costs) to fullfill patients' requirements.

The model formulation we have presented here is very general and extremely flexible. Unlike the previous work [6], the patients care profile is not a fixed covariate, so that this new model, being bivariate, is fairly flexible, though more complex.

Of course, different modelizations could have been assumed. For instance, one could wonder why we have assumed the negative binomial to model the holding times (see (4)). Indeed, we tried different distributions, like the geometric and the Poisson distributions. However, as it is well known, conditionally on its parameter, the geometric distribution has
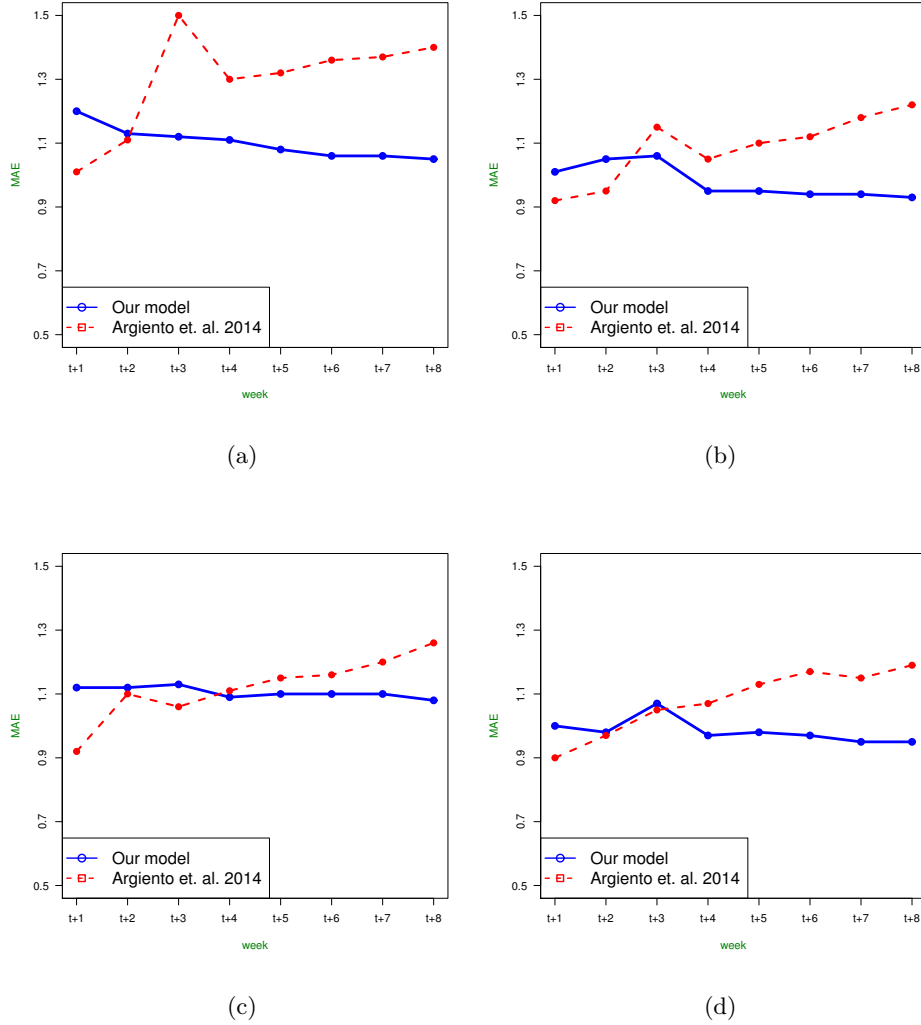
Figure 8: Comparison between MAE of the number of nurse visits under our model (solid line) and that in [6] (dashed) at $t = 99$ (a), $t = 149$ (b), $t = 175$ (c) and $t = 234$ (d).

the *memoryless property*, while the Poisson does not account for overdispersion, or more generally for different mean and variance. Both assumptions were not reasonable for our dataset. As far as modelization of $logit(q_{i,j})$ is concerned, instead of (5), we tried different alternatives: (i) one more random effect parameter $\beta_2[\eta_{i,j}]$ in place of the fixed effect β_2, (ii) modelling the dependence through time only via $H_{i,j-1}$, or (iii) only via the sum of the previous holding times spent in the care profile equals to $\eta_{i,j}$. By the way, assumption (i) is equivalent to assume that the bivariate discrete time process of the holding times and the care profile states is a semi-Markov process (see [27]). We also considered modelling the parameters $z_{i,j}$ in the negative binomial distribution of $\{H_{i,j}\}_j$ as a regression model on the

log-scale (iv) with a random intercept and a fixed effect term taking into account the sum of the previous holding times spent in the care profile equals to $\eta_{i,j}$, or (v) with a random intercept and two fixed-effects parameters for age at time t and gender (of course removing these covariates from the distribution of $N_{i,t}$). Summing up, none of the alternatives produced better goodness-of-fit nor gave new insight to the corresponding parameterization.

In conclusion, the main advantage of this model, compared to the other Bayesian paper of ours [6], is the opportunity to handle a longer time horizon, since the estimates remain good for a longer number of time slots. This is particularly useful in presence of middle- and long-term decisions, e.g., the nurse-to-patient assignment under continuity of care. Such a result justifies the higher complexity of this model, introduced to take into account the care profile CP as a response variable.

Acknowledgments

References

[1] Matta A, Chahed S, Sahin E, Dallery Y. Modeling home care organizations from an operations management prospective. *Flexible Service and Manufacturing Journal* 2013; doi:10.1007/S10696-012-9157-0.

[2] Lanzarone E, Matta A, Sahin E. Operations management applied to home care services: the problem of assigning human resources to patients. *IEEE Transations on System Man Cybernetics A* 2012; **42**(6):1346-1363.

[3] E. Lanzarone, A. Matta. Robust nurse-to-patient assignment in home care services to minimize overtimes under continuity of care. *Operations Research for Health Care* 2014; **3**: 48-58.

[4] Carello G, Lanzarone E. A cardinality-constrained robust model for the assignment problem in Home Care services. *European Journal of Operational Research* 2014; **236**(2):748-762.

[5] Lanzarone E, Matta A, Saccabarozzi G. A patient stochastic model to support human resource planning in home care. *Production Planning and Control* 2010; 21:3-25.

[6] Argiento R, Guglielmi A, Lanzarone E, Nawajah I. A Bayesian framework for describing and predicting the stochastic demand for care in home care patients. *Flexible Service and Manufacturing Journal*, 2014; under second minor review.

[7] Perez-Ocon R, Ruiz-Castro JE, Gamiz-Perez ML. Non-homogeneous Markov models in the analysis of survival after breast cancer. *Journal of the Royal Statistical Society Series C-Applied Statistics*, 2001; 50:111124

[8] Jackson C, Sharples L. Hidden Markov models for the onset and progression of bronchiolitis obliterans syndrome in lung transplant recipients. *Statistics in Medicine*, 2002; **21**(1):113-128.

[9] Sharples L, Jackson C, Parameshwar J, Wallwork J, Large S. Diagnostic accuracy of coronary angiography and risk factors for post-heart-transplant cardiac allograft vasculopathy. *Transplantation*, 2003; **76**(4):679-682.

[10] Commenges D, Joly P, Letenneur L, Dartigues JF. Incidence and mortality of alzheimer's disease or dementia using an illnessdeath model. *Statistics in Medicine*, 2004; 23:199210.

[11] Kang M, Lagakos S. Statistical methods for panel data from a semi-Markov process, with application to HPV. *Biostatistics* 2007; **8**(2):252-264.

[12] Meira-Machado LP, de Uña-Álvarez J, Cadarso-Suárez C, Andersen P. Multi-state models for the analysis of time-to-event data. *Statistical methods in medical research*, 2008.

[13] Beyersmann J, Schumacher M, Allignol A. Multistate modelling of competing risks. Competing Risks and Multistate Models with R. Springer New York, 2012. 41-53.

[14] Jackson C. Multi-state models for panel data: the msm package for r. *Journal of statistical Software* 2011; **38**(8):1-29.

[15] Blanco FJ. A multi state Markov model for projecting health care spending. *European Scientific Journal* 2013; **9**(21):745-752.

[16] Gardiner JC, Luo Z, Bradley CJ, Sirbu CM, Given C W. A dynamic model for estimating changes in health status and costs. *Statistics in medicine*, 2006; 25: 3648-3667.

[17] Hui-Min Wu G, Chang S-H, Hsiu-Hsi Chen T. A Bayesian Random-Effects Markov Model for Tumor Progression in Women with a Family History of Breast Cancer. *Biometrics* 2008; **64**(4):1231-1237.

[18] Armero C, Cabras S, Castellanos ME, Perra S, Quirós A, Oruezábal MJ, Sánchez-Rubio J. (2012). Bayesian analysis of a disability model for lung cancer survival. *Statistical methods in medical research* 2012; 0962280212452803, first published on-line on July 5, 2012.

[19] Nathoo F, Dean C. Spatial multi-state transitional models for longitudinal event data. *Biometrics* 2008; **64**(1):271-279.

[20] Kneib T, Hennerfeind A. Bayesian semi parametric multi-state models. *Statistical Modelling* 2012; **8**(2):169-198.

[21] Gelman A. Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). *Bayesian Analysis* 2006; **1**:515-534.

[22] Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing* 2003, 20–22.

[23] R Development Core Team. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing* 2003, Vienna.

[24] Plummer M, Best N, Cowles K, Vines K. CODA: Convergence diagnosis and output analysis for MCMC. *R News* 2006; 6:7-11.

[25] Gelman A, Carlin JB, Stern H, Dunson D, Vehtari A, Rubin D. Bayesian Data Analysis. Chapman & Hall: Boca Ration, Florida 2004.

[26] Gelfand A, Dey Dk, Chang H. Model Determination Using Predictive Distributions with Implementation via Sampling-Based Methods. In *Bayesian Statistics*, Volume 4, J. Bernardo et al. (eds), 147-167. Oxford: Oxford University Press.

[27] Barbu VS, Limnios N. Semi-Markov chains and hidden semi-Markov models toward applications. Springer: New York, 2008.