

A Bayesian Model for Describing and Predicting the Stochastic Demand of Emergency Calls

Vittorio Nicoletta, Ettore Lanzarone, Alessandra Guglielmi,
Valérie Bélanger and Angel Ruiz

Abstract Emergency Medical Service (EMS) systems aim at providing immediate medical care in case of emergency. A careful planning is a major prerequisite for the success of an EMS system, in particular to reduce the response time to emergency calls. Unfortunately, the demand for emergency services is highly variable and uncertainty should not be neglected while planning the activities. Thus, it is of fundamental importance to predict the number of future emergency calls and their interarrival times to support the decision-making process. In this paper, we propose a Bayesian model to predict the number of emergency calls in future time periods. Calls are described by means of a generalized linear mixed model, whose posterior densities of parameters are obtained through Markov Chain Monte Carlo simulation. Moreover, predictions are given in terms of their posterior predictive probabilities. Results from the application to a relevant real case show the applicability of the model in the practice and validate the approach.

Keywords Emergency medical services · Demand prediction · Generalized linear mixed model · Posterior predictive probabilities · Markov chain Monte Carlo

V. Nicoletta (✉) · A. Ruiz

Department of Operations and Decision Systems, Université Laval,
Quebec, Canada
e-mail: vittorio.nicoletta.1@ulaval.ca

A. Ruiz

e-mail: angel.ruiz@osd.ulaval.ca

E. Lanzarone

Consiglio Nazionale delle Ricerche (CNR), Istituto di Matematica Applicata
e Tecnologie Informatiche (IMATI), Milan, Italy
e-mail: ettore.lanzarone@cnr.it

A. Guglielmi

Dipartimento di Matematica, Politecnico di Milano, Milan, Italy
e-mail: alessandra.guglielmi@polimi.it

V. Bélanger

Department of Logistics and Operations Management, HEC Montréal,
Montreal, Canada
e-mail: valerie.belanger@cirrelt.ca

1 Introduction

Emergency Medical Service (EMS) consists of pre-hospital medical care and transport to a medical facility. Almost all EMS requests arrive by phone, through calls to an emergency number. The urgency of each request is evaluated and the location is obtained. Then, an ambulance is dispatched to the call site and, if needed, the patient is transported to a medical facility. Demand for such services is constantly increasing throughout the world, according to population growth and aging, while we observe a continuous pressure of governments to reduce health care costs; thus, an efficient use of resources is fundamental to guarantee a good quality of the service while maintaining the economic sustainability.

Several optimization planning models have been developed in the literature for EMS systems (see Bélanger et al. [3] for an extensive review). Unfortunately, the EMS demand is highly variable, and the uncertainty should not be neglected while planning the activities. Hence, it is fundamental to fairly predict the future number of emergency calls and their interarrival times.

The goal of this paper is thus to propose and validate a Bayesian model to predict the number of emergency calls in future time slots. The number of calls is described by means of a generalized linear mixed model, and the inference is based on the posterior density of model parameters, which is obtained through a Markov Chain Monte Carlo simulation scheme. Then, predictions are given in terms of their posterior predictive probabilities.

We demonstrate the applicability of the approach using the information available from the city of Montréal, Québec, Canada. Results show the convergence of the approach, good fitting, and low prediction errors.

The paper is organized as follows. A review of previous works dealing with stochastic modeling of EMS calls is presented in Sect. 2; the general features of an EMS system and the typical structure of the demand dataset are described in Sect. 3. Then, the Bayesian model is proposed in Sect. 4, and its application to the Montréal case is presented in Sect. 5. Conclusions of the work are finally given in Sect. 6.

2 Literature Review

Several studies deal with EMS calls prediction under a frequentist approach. An interesting survey of the works dated before 1982 can be found in Kamenetsky et al. [10]. In addition, the authors also presented regression models to predict EMS demand as a function of population, employment, and other demographic variables. Socio-economic parameters such as median income and percentage of people living below poverty line have been considered by Cadigan and Bugarin [5]. More recently, McConnell and Wilson [12] focused on the increasing impact of age distribution on EMS demand, while Channouf et al. [6] developed ARIMA models.

To the best of our knowledge, Bayesian approaches have not been considered for the EMS demand yet, even though they have been successfully applied in the health

care literature. In fact, Bayesian approaches allow combining the available data with prior information within a solid theoretical framework, and results can be used as prior information once new data are available, which are important features in health applications. A good example of application to another health care service (i.e., the home care service) can be found in [1, 2].

3 Problem Description

An EMS system consists of an operations center and a certain number of ambulances, including the related staff. Ambulances are located in predetermined sites, ready to serve EMS requests. Requests arrive at the operations center via telephone, where they are evaluated. If a call requires an intervention, it is assigned to one of the available vehicles. The aim of an EMS is to serve all calls as fast as possible, maximizing the number of calls served within a given threshold that depends on the type of area (urban or rural).

For this purpose, due to the high uncertainty related to EMS calls, the decision maker needs accurate estimates of the demand as input for any optimization model underlying ambulance dispatching.

The typical EMS dataset includes several information about the calls and the provided service. For the aim of developing a prediction model, we focus on the calls. Three types of information are available:

- *Type*: required service and patient characteristics; this information is usually summarized into a priority associated to the call.
- *Arrival time*: day and time of the call.
- *Coordinates*: latitude and longitude of the call, or alternatively the address.

Usually, for managing purposes, the territory is divided into zones; thus, coordinates are translated into the zone z ($z = 1, \dots, Z$) of the call. Moreover, concerning the arrival times, in this work we group the time into slots. Thus, day i ($i = 1, \dots, I$) and slot t ($t = 1, \dots, T$) are associated to the call, and for each day i we register the number of calls $N_{z,t}^i$ arisen in slot t and zone z . In particular, we consider slots of two hours, i.e., $T = 12$.

4 The Bayesian Model

We propose the following generalized linear mixed model for the number of calls $N_{z,t}^i$:

$$N_{z,t}^i | \lambda_{z,t}^i \stackrel{ind}{\sim} Poisson(\lambda_{z,t}^i) \tag{1}$$

$$\log(\lambda_{z,t}^i) = \beta_1 p_z + \beta_2 a_z + \sum_{k=1}^K \beta_{3,k} \phi_{k,z} + \beta_4 h_i + \gamma_t \tag{2}$$

where: p_z and a_z are the population and the area of zone z , respectively; h_i is a binary covariate equal to 1 if day i is holiday and 0 otherwise; $\Phi_z = [\phi_{k,z}]$ is a dummy vector of dimension K describing the type of zone z .

Zones z are classified into $K + 1$ types (e.g., residential, commercial, industrial); $\phi_{k,z} = 1$ if zone z is of type k (with $k = 1, \dots, K$) and 0 otherwise, while $\phi_{K+1,z}$ is always equal to 0 if zone z is of type $K + 1$, to avoid identifiability problems.

Model (1) and (2) is a generalized linear mixed model with four fixed-effects parameters $\beta_1, \beta_2, \beta_3$ and β_4 (where β_3 is K -dimensional), and a random-effects parameter γ_t . The latter takes into account the similarity of the number of calls in different zones during the same time slot t . In this formulation $\lambda_{z,t}^i$ is the parameter responsible for EMS calls: the higher the parameter $\lambda_{z,t}^i$ is, the higher the expected number of calls is.

Finally, independent non-informative priors, i.e., Gaussian distributions with 0 mean and large variance equal to 100, are chosen for $\beta_1, \beta_2, \beta_4, \gamma_t$, and for the components of vector β_3 :

$$\begin{aligned} \beta_j &\stackrel{iid}{\sim} \mathcal{N}(0, 100) & j = 1, 2, 4 \\ \beta_{3,k} &\stackrel{iid}{\sim} \mathcal{N}(0, 100) & k = 1, \dots, K \\ \gamma_t &\stackrel{iid}{\sim} \mathcal{N}(0, 100) & \forall t \end{aligned}$$

5 Application to the Dataset

Data adopted in this work are those adopted in [4, 7, 11]. They refer to EMS calls arisen in the city of Montréal and the near suburb of Laval, Québec, Canada, i.e., a region with about 2.3 million of inhabitants and a territory of 744 km². According to these data, the region is divided into $Z = 595$ demand zones. In addition to the EMS data, information from Municipality of Montréal have been used to define the vector Φ_z for each zone. Eleven different types of zone are present, as described in Table 1; moreover, to avoid collinearity due to the low number of zones belonging to some types, types are regrouped as follows:

- *Residential* ($k = 1$);
- *Workplace*, regrouping commercial, office, industrial and institutional ($k = 2$);
- *Street* ($k = 3$);
- *Other*, regrouping park, agricultural, empty space, water, and golf field.

Finally, data about population has been divided by 1,000 to be of the same order of magnitude of the other covariates.

Table 1 Total number of calls and empirical mean divided by the type of zone

Type of zone	Number of zones	Total number of calls	Mean number of calls
Residential	266	95,394	358.62
Commercial	14	4,325	308.93
Office	7	3,361	480.14
Industrial	19	4,359	229.42
Institutional	46	18,004	391.39
Park	30	8,003	266.77
Street	184	67,738	368.14
Agricultural	4	506	126.50
Empty space	19	4,190	220.53
Water	2	405	202.50
Golf field	4	615	153.75

Table 2 Empirical mean and standard deviation of the number of calls divided by time slot

Time slot	Mean number of observations	Standard deviation of the number of observations
1	0.0606	0.2513
2	0.0535	0.2416
3	0.0409	0.2040
4	0.0529	0.2334
5	0.0908	0.3098
6	0.1039	0.3328
7	0.1010	0.3272
8	0.0991	0.3244
9	0.0937	0.3143
10	0.0915	0.3089
11	0.0893	0.3067
12	0.0776	0.2851

5.1 Descriptive Statistics

The dataset consists of 2,606,100 observations for $N_{z,t}^i$ ($I = 365$ days, $Z = 595$ zones and $T = 12$ slots) together with the related covariates.

Tables 1 and 2 report the main information about the data. Moreover, Fig. 1 shows a map of the territory together with the number of calls.

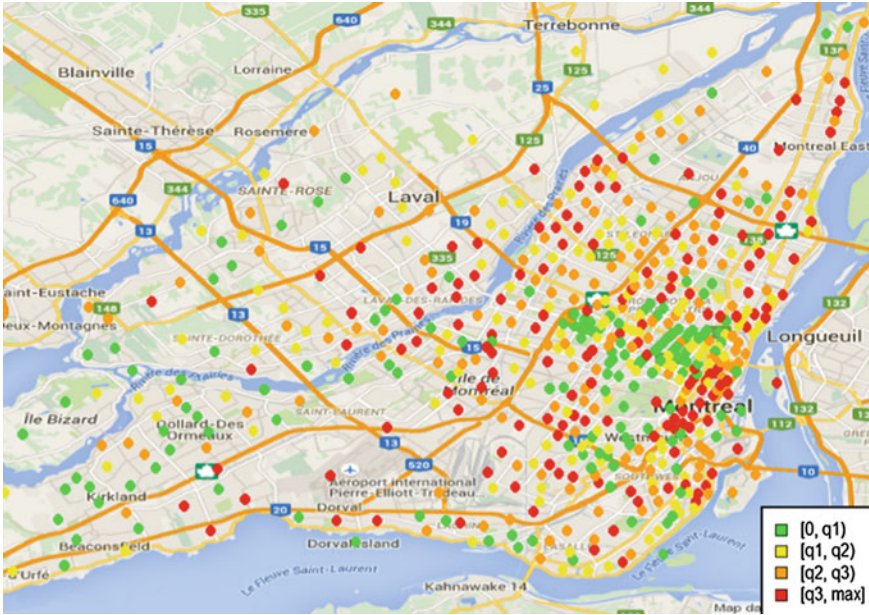


Fig. 1 Map of the city of Montréal together with the total number of calls. The number of calls for each zone is represented by a point in the center of the zone. *Green points* correspond to lower numbers of EMS calls, while *red points* represent higher numbers of EMS calls, according to the division in quartiles reported in the legend

5.2 Posterior Analysis

5.2.1 Convergence Analysis

The model is implemented in STAN (<http://mc-stan.org/>), which uses the Hamiltonian Monte Carlo algorithm to reduce the correlation and obtain faster convergence of the chains. Hence, 5,000 MCMC iterations have been run, with a burn-in of 1,000 iterations and a final sample size of 4,000.

Traceplots, autocorrelations and the Gelman–Rubin convergence statistics (\hat{R}) have been considered to verify that convergence is achieved. Moreover, we have estimated the Monte Carlo Standard Error (MCSE) with the MC error, the Naive SE and the Batch SE. See [8, 9] for further information.

Results show that \hat{R} is equal to 1 and that the MCSE is always less than the 5% of the standard deviation for all parameters. Moreover, nice traceplots and autocorrelations are obtained, showing that the convergence of the chain is satisfactory.

5.2.2 Credible Intervals of Model Parameters

Inference for each model parameter is reported in terms of the posterior 95% credible interval (CI).

CIs of the fixed-effects parameters are reported in Table 3. The population parameter β_1 yields a positive effect, thus increasing number of calls, while the area parameter β_2 gives a negative effect. This is in agreement with the considered data, in which zones with large areas have small population densities; thus, the higher the population density of a zone is, the higher the number of calls is. Vector β_3 gives the effect of the zone; results show that workplace zones and streets have more EMS calls, followed by Residential Zones. Finally, CI of parameter β_4 suggests that a lower number of calls is to be expected during holidays.

Posterior CIs for the random-effects vector γ_t are reported in Fig. 2. They suggest a clear distinction of the time slots: a higher number of calls arrive during the day (slots $t = 5, \dots, 11$), while a lower number of calls arrive during night hours.

Table 3 95% CIs for the fixed-effects parameters

Parameter	Covariate	2.5%	50%	97.5%
β_1	Population	0.087	0.090	0.093
β_2	Area	-0.049	-0.047	-0.044
β_{31}	Residential	0.277	0.297	0.316
β_{32}	Workplace	0.347	0.369	0.389
β_{33}	Street	0.332	0.352	0.371
β_4	Holiday	-0.067	-0.055	-0.044

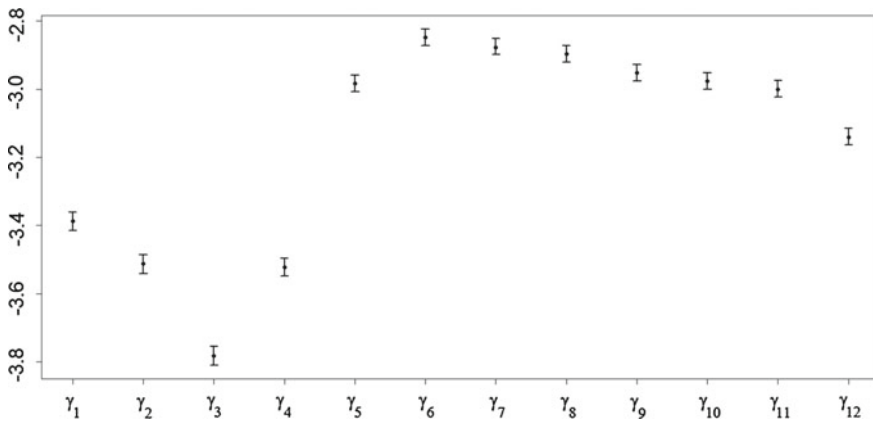


Fig. 2 95% CIs for the random-effects vector γ_t

5.2.3 Cross-Validation Prediction

A cross-validation approach is adopted to validate the model, by partitioning the complete dataset. The first 90% of the days (with $i = 1, \dots, I - \tilde{I}$) is used as *training set* to determine the posterior density, while the remaining 10% (with $i = I - \tilde{I} + 1, \dots, I$) is used as *testing set*. The predictive distributions of each $N_{z,t}^i$ (with $i = I - \tilde{I} + 1, \dots, I$) are computed, and the predictions are checked with the corresponding observed data.

The accuracy of the predictions is evaluated in terms of the global Mean Absolute Error (MAE), defined as:

$$MAE = \frac{1}{\tilde{I} Z T} \sum_{i=I-\tilde{I}+1}^I \sum_{z=1}^Z \sum_{t=1}^T \left| N_{z,t}^{i\ obs} - \hat{N}_{z,t}^i \right|$$

where the product $\tilde{I} Z T$ is the numerousness of the sample in the testing set, and $N_{z,t}^{i\ obs}$ and $\hat{N}_{z,t}^i$ represent the observed number of calls and the number predicted by the model (median of the predictive distribution) at day i , zone z and slot t , respectively. The obtained value is 0.078, which is two orders of magnitude lower than the unit, showing a good fit of the model.

We have also detailed the MAE for each combination of type of zone k and time slot t . Results in Table 4 show quite similar values, whose maximum is 0.111, confirming a good fit of the model that does not significantly deteriorate for any pair k and t .

Table 4 MAE for each combination of type of zone k and time slot t

Time slot t	Type of zone k				
	0	1	2	3	All
1	0.035	0.061	0.060	0.062	0.059
2	0.036	0.048	0.046	0.051	0.048
3	0.025	0.043	0.036	0.043	0.040
4	0.033	0.060	0.057	0.056	0.056
5	0.066	0.103	0.100	0.096	0.097
6	0.063	0.111	0.111	0.104	0.104
7	0.054	0.105	0.104	0.102	0.099
8	0.052	0.102	0.103	0.104	0.098
9	0.061	0.093	0.093	0.097	0.091
10	0.057	0.093	0.085	0.092	0.088
11	0.063	0.087	0.083	0.089	0.084
12	0.047	0.077	0.074	0.078	0.074
All	0.049	0.082	0.079	0.081	0.078

Table 5 Comparison of the MAE between the proposed Bayesian model and the mean frequentist approach, grouped by type of zone z , time slot t , and holiday h

		Bayesian model	Frequentist mean
Type of zone	0	0.049	0.098
	1	0.082	0.152
	2	0.079	0.148
	3	0.081	0.153
Time slot	1	0.059	0.112
	2	0.048	0.096
	3	0.040	0.078
	4	0.056	0.103
	5	0.097	0.171
	6	0.104	0.188
	7	0.099	0.181
	8	0.098	0.178
	9	0.091	0.169
	10	0.088	0.164
	11	0.084	0.159
	12	0.074	0.141
Holiday	0	0.078	0.147
	1	0.077	0.142

5.2.4 Comparison with the Mean Estimate

In this Section we compare the outcomes of the proposed model with those of a very simple frequentist approach, in which the predictions are simply given by the historical means. This approach gives as a predictor the mean number of calls for the specific combination of type of zone z , time slot t and holiday parameter h . MAE values are computed considering the same training and testing sets as in Sect. 5.2.3.

The global MAE of the frequentist approach is equal to 0.145, while the values grouped by z , t and h are reported in Table 5. Results show that the MAE under the frequentist approach is higher, being about the double than the MAE under the proposed Bayesian approach. This further confirms the good fit of the proposed model to the data.

6 Conclusions

This paper presents a first attempt to deal with stochasticity in the EMS calls by using the Bayesian approach. A generalized linear mixed model has been proposed, with the aim of identifying relevant effects that influence the calls and giving predictions of future EMS calls.

Results from the Montréal case suggest that population, area and type of zone have a strong impact. Moreover, as expected, the time slot has a relevant effect, showing lower predicted values of number of calls during the night. Finally, the model shows good performance when used to make predictions, and documented by the low MAE values under cross-validation.

Moreover, the model is general, and can be easily applied to describe EMS demand in other cities. On the contrary, as for rural situations, we expect that some modifications are necessary to include the presence of rare events in an environment characterized by a usually low demand. Another extension will be to consider the area of each zone as an offset/exposure term of the Poisson regression.

References

1. Argiento, R., Guglielmi, A., Lanzarone, E., Nawajah, I.: A Bayesian framework for describing and predicting the stochastic demand of home care patients. *Flex. Serv. Manuf. J.* **28**, 254–279 (2016)
2. Argiento, R., Guglielmi, A., Lanzarone, E., Nawajah, I.: Bayesian joint modeling of the health profile and demand of home care patients. *IMA J. Manag. Math.* (2016). doi:[10.1093/imaman/dpw001](https://doi.org/10.1093/imaman/dpw001)
3. Bélanger, V., Ruiz, A., Soriano, P.: Recent advances in emergency medical services management. Working document, Faculty of Business Administration of Laval University (2015)
4. Bélanger, V., Kergosien, Y., Ruiz, A., Soriano, P.: An empirical comparison of relocation strategies in real-time ambulance fleet management. *Comput. Ind. Eng.* **94**, 216–229 (2016)
5. Cadigan, R.T., Bugarin, C.E.: Predicting demand for emergency ambulance service. *Ann. Emerg. Med.* **18**, 618–621 (1989)
6. Channouf, N., L'Ecuyer, P., Ingolfsson, A., Avramidis, A.N.: The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta. *Health Care Manag. Sci.* **10**, 25–45 (2007)
7. Crowe, J.: Une approche hybride pour la confection d'horaires des paramédics. Master thesis, HEC Montréal (2007)
8. Gelman, A., Rubin, D.B.: Inference from iterative simulation using multiple sequences (with discussion). *Stat. Sci.* **7**, 457–511 (1992)
9. Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B.: *Bayesian Data Analysis*, 2nd edn. Chapman and Hall/CRC, Boca Raton (2003)
10. Kamenetsky, R., Shuman, L., Wolfe, H.: Estimating need and demand for prehospital care. *Oper. Res.* **30**, 1148–1167 (1982)
11. Kergosien, Y., Bélanger, V., Soriano, P., Gendreau, M., Ruiz, A.: A generic and flexible simulation-based analysis tool for EMS management. *Int. J. Prod. Res.* **53**, 7299–7316 (2015)
12. McConnell, C.E., Wilson, R.W.: The demand for prehospital emergency services in an aging society. *Soc. Sci. Med.* **46**, 1027–1031 (1998)