

Forecasting the Demand for Emergency Medical Services

Krisjanis Steins
Department of Science and
Technology, Linköping
University, Sweden
krisjanis.steins@liu.se

Niki Matinrad
Department of Science and
Technology, Linköping
University, Sweden
niki.matinrad@liu.se

Tobias Andersson Granberg
Department of Science and
Technology, Linköping
University, Sweden
tobias.granberg@liu.se

Abstract

Accurate forecast of the demand for emergency medical services (EMS) can help in providing quick and efficient medical treatment and transportation of out-of-hospital patients. The aim of this research was to develop a forecasting model and investigate which factors are relevant to include in such model. The primary data used in this study was information about ambulance calls in three Swedish counties during the years 2013 and 2014. This information was processed, assigned to spatial grid zones and complemented with population and zone characteristics. A Zero-Inflated Poisson (ZIP) regression approach was then used to select significant factors and develop the forecasting model. The model was compared to the forecasting model that is currently incorporated in the EMS information system used by the ambulance dispatchers. The results show that the proposed model performs better than the existing one.

1 Introduction

Emergency Medical Services (EMS) are responsible for providing out-of-hospital medical care to people in need and also transporting them to a medical facility when needed. These services are time sensitive, i.e. the later the EMS arrive to the incident sites, the more severe the damage to the patients' health can become, especially in cases such as trauma and cardiac arrest [22]. Therefore, it is essential to have a system capable of providing timely services. Yet, when considering issues that emergency managers face, such as resource shortages and cutbacks, and long distances between the EMS and sparsely populated areas [26], without proper management of the existing limited resources, quick and efficient service would not always be possible. For achieving an efficient resource management, access to appropriate information about

the expected demand for service is necessary. With a reliable forecast of emergency events that can happen on a daily basis, it is possible to manage the resources efficiently, and proactively plan for an optimized management of the resources. This will help improving the emergency response system's performance.

The aim of this study was to develop a forecasting model for predicting the number of EMS calls per hour and geographical zone in an area.

The goal was to contribute to: (1) better management of existing emergency resources, and (2) the determination of factors affecting number of calls, which can be used to plan preventive action.

In collaboration with SOS Alarm, who operationally manages the major part of the ambulance resources in Sweden, different relevant factors that can be used in the forecasting model were identified. A forecasting model based on Zero-Inflated Poisson (ZIP) regression was constructed to determine where and when the emergency call(s) will occur. The results of the proposed model were compared to the current model used by SOS Alarm, which is an integrated part of their information system (IS), called CoordCom [5]. CoordCom is used for managing both calls for service and the resources required to provide the service. In particular, the current forecasting model is used as input to a set of ambulance dispatch support tools [3] in the geographical information system (GIS) part of CoordCom, called ResQMap [6].

The rest of the paper is organized as follows. A review of previous studies dealing with EMS call forecasting as well as positioning of this work with respect to the literature is presented in Section 2. Section 3 describes the data and methods, followed by the results of the statistical modelling and analysis in Section 4. The results and methodological issues are discussed in Section 5. Finally, conclusions and future research directions are outlined in Section 6.

2 Related work

The following is a brief review of previous works on emergency call forecasting, based on the methods used.

2.1 Regression models

In the literature considered in this study, eight papers use some form of regression modelling to make the required prediction. In the work of Aldrich et al. [2], based on ambulance trip data obtained from Los Angeles Central Receiving Hospital, a forecasting model was developed for which a least squares regression was adopted. In this model socioeconomic characteristics of the census tract, the type of public service, and the availability of alternative sources of care formed the independent variables. The result of their study indicated that aged people and single men generated more calls compared to the rest of the population. Using stepwise and multiple regression, and with the help of socioeconomic and geographic characteristics, Siler [20] constructed a model to predict the demand for ambulances in Los Angeles county. Unlike Aldrich et al. [2], who assumed a linear relationship between the dependent and the independent variables, Siler [20] adopted a nonlinear one. Similarly to [2] and [20], Kvålseth and Deems [14] considered socioeconomic and demographic characteristics to predict the ambulance demand, and used emergency medical calls and census tract data of the city of Atlanta to develop first-order and second-order regression models. In 1982, Kamenetzky et al. [13] also used first-order and second-order regression models to predict ambulance demand, and in their model they considered four independent variables to be of importance for the forecast: population, employment, logarithm of the percentage of white and married population, and the square of employment to population ratio, and considered the total number of emergency calls by geographic unit as the dependent variable. Their study used data from Southwestern Pennsylvania. Svensson [21], adopted Poisson regression to make an ambulance demand prediction for Kentucky. Independent variables included location (both rural and urban), level of availability of prehospital care, access to a 24-hour emergency department in the county, availability of 911 service, poverty (per capita income, percentage of households with no wage earner, and percentage of residents below the poverty level), education (percentage of those with less than ninth grade education), and availability of a telephone in the household. It was shown that age and poverty are significantly correlated with the demand for emergency services. Wong and Lai [25] utilized

multiple regression in two stages; in stage one, they investigated the relationship between ambulance data and weather variables, and in stage two, through forward variable selection, a series of regression models based on weather factors were created. The results showed that weather factors can be good predictors for ambulance demand for older people, patients with more severe conditions, hospital admitted cases, and comprehensive social security assistance recipients. Considering ambulance demand in Portland, Oregon Metro Area, Cramer et al. [9] used stepwise regression, which determined renters, businesses, jobs, people not in labor force, and college graduates as independent variables affecting ambulance call volume. A geographically weighted regression was used to investigate the effect of each identified variable on the number of ambulance calls. In 2011, Lowthian et al. [16] conducted a study to forecast the demand of ambulance calls in metropolitan Melbourne with linear regression. Demographic variables such as age and sex were used. The result of their study indicated that older people would make more ambulance calls compared to other age groups.

2.2 Time series models

Another type of forecasting method used to predict the ambulance calls is time series modelling. In 1986, Baker and Fitzpatrick [4] adopted Winters' exponential smoothing model to forecast ambulance calls using data from four South Carolina counties. In their work, they utilized a multistep approach for determining the optimal parameters of the exponential smoothing model and considering that the model would depict both emergency and non-emergency calls; goal programming was used to combine forecast results for both call types. Channouf et al. [7] presented two time series approaches for forecasting the volume of daily emergency calls: (1) autoregressive models of data obtained after eliminating trend and (2) doubly-seasonal autoregressive integrated moving average (ARIMA) models with special-day effects. For this purpose, they used the data from emergency calls in Calgary. They showed that the calls are affected by when people work, commute, sleep and celebrate. Matteson et al. [17] combined integer-valued time series models with a dynamic latent factor structure to predict ambulance calls, including all priority levels of emergency calls, at an hourly level in the city of Toronto. Based on data obtained from the Welsh Ambulance Service Trust, Vile et al. [23] explored a non-parametric technique for time series analysis known as Singular Spectrum Analysis and adopted it to predict demand at a daily level. As it has been shown that weather factors play a role as predictor in making

a forecast for ambulance calls in Hong Kong [25], Wong and Lai [24] employed ARIMA to predict daily ambulance demand using 7-day weather forecast data as predictor. Finally, to better manage the manpower in a call center responsible for answering and dispatching ambulances to emergency calls, Gijo and Balakrishna [10] developed a forecasting model based on seasonal ARIMA for call records obtained from one of India's states.

2.3 Spatial-temporal models

Besides models that are based on time series and regression, there are spatial-temporal models and other models capable of depicting both time and location simultaneously. Setzler et al. [19] utilized artificial neural networks to predict the emergency demand volume during different times of the day for specific geographical areas based on data of the Charlotte Mecklenburg (North Carolina) emergency service organization. Their model not only considered time of the day for making the prediction, but also season, day of the week, and month. Zhou and Matteson [27] used spatial-temporal kernel density estimation to make spatial and temporal ambulance call predictions for Toronto. Also for Toronto, on a continuous spatial domain and discretized temporal domain, Zhou et al. [28] adopted a bivariate Gaussian mixture model. In their model the component distributions are fixed through time while the mixture weights change over time. Nicoletta et al. [18] proposed a Bayesian model for predicting emergency call volumes in the city of Montréal; for describing the calls, a generalized linear mixed model was used in which the parameters' posterior densities were established by using Markov Chain Monte Carlo simulation. Their proposed model captured the time and location of future calls, and in their modeling, they had discretized time into 12 time slots, and location into several zones.

2.4 Positioning of the current study

In comparison to previous studies, the following points can be emphasized:

- We use a regression model that can deal with both count data and large number of zeros in the outcome variable. As far as we know, this is the first attempt to use ZIP regression to forecast EMS calls.
- In contrast to previous regression models, time (day of week, and hour) has been considered as independent factors in our model, in addition to other factors, such as socioeconomic and geographic.

- Our proposed model can handle geographical zones without population and with a very low call frequency, which is difficult for other models. This is important as some zones, being vacation spots or having roads crossing them, might have no historical calls, but still a positive risk for getting future calls.

3 Data and methods

The historical ambulance call data used in this study was provided by SOS Alarm, while data for explanatory factors (independent variables) was extracted from Statistics Sweden databases, the Swedish national road database and OpenStreetMap.

3.1 Ambulance call data acquisition and processing

In this study, historical ambulance call data for 2013 and 2014 for three counties in Sweden was used. Each call registered at the emergency call center is assigned a case number and the time of the call. Additional information that is recorded includes the case priority (1-4, where Priority 1 are urgent life threatening calls and 4 are planned patient transportations), type of emergency, geographical coordinates, and time stamps including the time the call was registered, the time an emergency resource (e.g. ambulance) is dispatched, the time the resource starts its travel towards the call site, the time the resource reaches the call site, the time the patient is loaded, the time it reaches the hospital, and the time it returns to its station. It should be noted that since the proposed forecasting model considers only priority 1 and 2 cases, the primary data obtained from SOS Alarm did not contain priority 3 and 4 calls.

The following preprocessing steps were taken to create a clean dataset for further analysis. First, cancelled ambulances, i.e. entries that were missing data on "time of arrival to the call site", "time of loading a patient" and "time of reaching a hospital", were removed from the data. This step reduced the volume of the dataset by 6.32%.

Secondly, the difference between the time that the call was received and the time that the ambulance was dispatched to the site was calculated:

- Priority 1 calls with a time difference of more than 10 minutes and priority 2 calls with more than 30 minutes were removed. The reason for this was that consulted experts (ambulance dispatchers) stated that dispatch delay times longer than that were unreasonable, and probably due to faulty data.

- Entries with missing data on the time that the ambulance had been dispatched was also omitted from the dataset as it created a negative time difference.

As a result, an additional 7.9% of the data were removed.

Thirdly, each row in the data set corresponds to one ambulance dispatch. Thus, cases to which multiple ambulances were dispatched have multiple rows. As the case numbers indicate this, it was possible to calculate the number of ambulances dispatched to each case, and the number was added to the dataset as a new

calls, a literature study (described in Section 2) was performed to discover which factors had been used in previous models. In addition, a set of interviews with ambulance dispatchers employed by SOS Alarm was held. These resulted in a list of factors presented in Table 1. While the columns *Factor* and *Type* are self-explanatory, *Expected impact* describes expected influence of the factor on the number of ambulance calls. The value (low, medium, high) is set based on the results of interviews. Factors, the same or similar, which have been previously considered in the literature are referenced under *Reference*. Eventually, based on

Table 1. List of factors that may influence the number of ambulance calls

Factor	Type	Expected impact	Reference
Population (number of people)	Socioeconomic	High	[13]
Population characteristics: Age, sex, marital status, type of household, citizenship, land of birth, foreign background	Socioeconomic	Low - Medium	[13, 14, 16, 20, 25]
Income	Socioeconomic	Medium	[2, 13, 14, 21]
Education	Socioeconomic	Medium	[13, 21]
Employment (num. of pop employed)	Socioeconomic	Medium	[2, 13, 14, 20, 21]
Workplace (size, num. of employees)	Socioeconomic	High	[2, 13, 18]
Workplace (type of business)	Socioeconomic	Low	[9]
Housing type (house, apartment, etc)	Socioeconomic	Low	[2, 14]
Health (long time sick leave)	Socioeconomic	Medium	Interview
Temperature	Weather	Medium	[24, 25]
Extreme weather	Weather	High	[24, 25]
Time of day	Time	High	[19, 27]
Day of week	Time	High (for weekends)	[7, 17, 19, 23]
Month/Season	Time	Medium	[7, 19]
Holidays	Time	High	[7, 23]
Traffic flow	Traffic	High	[14]
Road type	Traffic	Medium	[2]
Special events	Events	Medium	[7, 23]
Distance from city center	Place	Low	[9]
Distance from ambulance station	Place	Low	[21]

column. Then for each case, the number of patients that was loaded and transferred to a hospital was calculated.

Thus, the data was transformed from a dataset that contained information about the dispatched ambulances to a more compact dataset with only information about single events that resulted in a dispatch of one or more ambulances. The information for each remaining row of data after these steps, included the number of loaded patients and the number of dispatched ambulances, besides the primary information such as case priority and time stamps.

3.2 Selection of candidate explanatory factors

To obtain information about factors that might explain the temporal and spatial variation of ambulance

the expected impact and the availability of data for the factors, explanatory factors that were tested in the model were selected, and are presented in Table 2.

3.3 Spatial grid zone characteristics

The forecast currently used by SOS Alarm (further described in Section 3.4) is spatially structured in grid zones (squares) with varying side lengths in different counties in Sweden. For each zone, there exists one forecast for each hour of the week, e.g. expected number of calls per hour on a Tuesday between 8 and 9 am. To be able to compare the new model to the currently used one, the data had to be structured in the same format.

The factors in Table 2 were processed using the GIS software ArcMap, to fit the existing grid zones, used for the current forecast. Thus, for every grid zone,

the number of people living there, the median income in the zone, the sum of all road lengths traversing the zone and the number of shops and restaurants were calculated.

To get an estimate of the number of people in the zone during day time (*Day population*), the total population in the zone was subtracted by the number of people who can be expected to leave the zone during the day (estimated as the population above 16 years of age with a job). Then, the number of people who have a registered work place in the zone was added.

The number of people in the zone during night time (*Night population*) was estimated to be the same as the total population, which is based on the national registration of people.

3.4 Current forecasting system used by SOS Alarm

The current forecasting model is based on moving average with seasonality weights (one season per week hour). As updates of the forecast are done only once per year, this practically translates to calculating averages for one year, and using these as forecasts for the next.

First, a temporal forecast is made for the whole county, giving one value for each season, i.e. $7 \cdot 24 = 168$ time slots. For each time slot, the average number of calls, based on the previous years' data, is calculated, giving a temporal forecast for the expected number of calls per each hour of the week.

In the next stage, to obtain a spatial distribution of the forecast, the weight for each geographical grid zone is calculated. The population in each zone is broken down into "day population" and "night population" to depict where people work and where they live. Day population is assumed to be in effect from 07:00 to 17:59, and night population between 18:00 and 06:59. By dividing the zone population with the total population, fractions for each zone (both day and night population) are determined. Eventually, the population fraction for each zone is multiplied with the forecasted number of calls per week hour.

Thus, the current model uses the following predictors for the forecast in each zone: day of the week, hour of the day, and day- or night population.

3.5 Zero-inflated Poisson (ZIP) regression

When developing models for forecasts of count data, commonly used regression models, such as multiple linear regression, are not able to perform well due to a skewed frequency distribution. Besides this, these methods can produce biased standard errors if the mean of outcome variable is low [8]. Hence other approaches, such as Poisson regression and negative binomial regression have to be utilized. However, if there are an excessive number of zeros among the count data, standard Poisson and negative binomial regression models will not perform well either. For this reason, zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) regression models have been

Table 2. Explanatory factors tested in the model

Factor	Original data format	Source	Data format, additional information
Population in different age groups	Number of people in squares with 250-1000 m sides	Statistics Sweden	Number of people in each of the following age groups: 0-6, 7-15, 16-19, 20-24, 25-44, 45-64, 65- years
Day population	Number of people in squares with 250-1000 m sides	Statistics Sweden	Estimated number of people in the square during day time
Night population	Number of people in squares with 250-1000 m sides	Statistics Sweden	Estimated number of people in the square during night time
Employment	Number of people in squares with 250-1000 m sides	Statistics Sweden	Number of people in the square between 20-64 that have a job
Median income	Swedish <i>kronor</i> for squares with 250-1000 m sides	Statistics Sweden	The median income per household in the square
Place of birth	Number of people in squares with 250-1000 m sides	Statistics Sweden	Number of people living in the square for each of the four place-of-birth categories: Sweden, the Nordic countries, EU, elsewhere
Road length	Polyline	The Swedish national road database	The total road length in km
Shops	Point	OpenStreetMap	Number of points of interest classified as typical day time attractions, e.g. shops, libraries, malls, banks
Restaurants	Point	OpenStreetMap	Number of points of interest classified as typical night time attractions, e.g. restaurants, bars, pubs

suggested [12]. One of the first comprehensive approaches to ZIP regression modelling was presented by Lambert [15].

Unlike ZINB, standard ZIP cannot model the data properly if the data is overdispersed, in the sense that estimates of standard errors and confidence limits will be too small, and the significance will be overestimated [8].

Since the data used in this study is count data (i.e. number of emergency calls), and not overdispersed, and due to its structure contains excessive zeros, ZIP regression is suitable. Even though the selected method suits the data and problem at hand very well, it also shares many of the same weaknesses as other regression modelling methods, such as potential multicollinearity or difficulty in selecting relevant predictors.

A separate ZIP model was developed for each of the three counties, and parameters estimated using data for year 2013. Different model versions (in terms of included variables, their transformations and interactions) were compared using mean error (ME), mean absolute error (MAE), and root mean squared error (RMSE). The model performance was evaluated comparing the forecast to a validation dataset based on data for year 2014. Similarly, a forecast using the current model (described in Section 3.4) was developed using data from 2013 and then evaluated using data from 2014, thus enabling a comparison between the current and the new model.

3.6 Accuracy measures: ME, MAE and RMSE

There are several criteria in the literature used for measuring the accuracy of predictions, and a model's goodness of fit. As Hyndman and Koehler [11] suggest, these can be categorized into (1) scale-dependent measures, (2) measures based on percentage errors, (3) measures based on relative errors, and (4) relative measures. From these four categories, we employ three of the most commonly used scale-dependent measures: ME, MAE and RMSE. These measures are useful when comparing the results of one model with another on the same dataset [11], which we do here. Furthermore, they are in the same scale as the dataset, making their interpretation and understanding easier.

The forecast error e_{ijk} is calculated as: $e_{ijk} = Y_{ijk} - X_{ijk}$, where (for a given county) Y_{ijk} is the observed number of calls in zone i , day of week j and hour k , and X_{ijk} is the forecasted number of calls in zone i , day of week j and hour k . From this we can calculate ME, MAE and RMSE in a standard way. It should be noted that while RMSE has theoretical relevance in statistical

modelling, it is more sensitive to outliers compared to MAE [11]. The main reasons for selecting RMSE rather than mean squared error (MSE) is that RMSE has the same scale as the main data and hence is easier to interpret. The other two accuracy measures complement the RMSE in that MAE will not be as sensitive to possible outliers and ME will allow detection of systematic overestimation or underestimation by the model.

4 Results

4.1 Call data

The call data was first analyzed to discover any seasonal variations or trends.

While we could not establish significant seasonality trends in call volumes over larger periods of time (month or year), there was a significant variation pattern present in call volumes per day of week and hour of day, similar for all three counties (Figure 1). The total number of calls per county in year 2013 are displayed in Table 3.

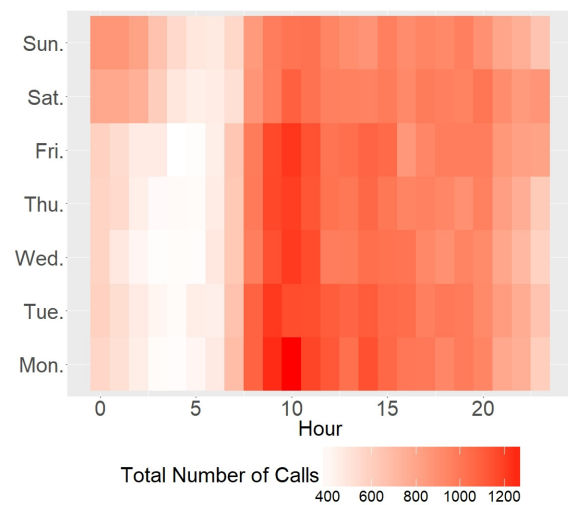


Figure 1. Ambulance call volume per day of week and hour in county O, year 2013.

4.2 Zone data

The spatial grid zones are of different size and different population density in the three counties studied in this paper. All three counties contain both cities and rural areas, but overall population density varies substantially between the counties and many spatial grid zones contain little or no population at all (Table 3).

Table 3. Call and population data (year 2013)

	County J	County E	County O
Total number of calls	21123	37396	141536
Number of zones	1768	1438	1950
Zone size	6x6 km	3x3 km	4x4 km
Nr. of zones (%), pop. = 0	744 (42%)	222 (15%)	274 (14%)
Nr. of zones (%), pop. < 100	1557 (88%)	1124 (78%)	890 (46%)
Median pop. per zone	5	27	118
Mean pop. per zone	146.74	303.33	822.51
Total population	259429	436195	1603889

4.3 Forecasting model results

The resulting dataset that was used for model development contained 297 024, 241 584 and 327 600 rows for counties J, E, and O respectively. The percentage of rows with zero calls was 97%, 94% and 87%, which confirms the need for a model capable of handling excess zeros.

After testing different sets of variables, their transformations and interactions, the following variables were included in the final model based on their statistical significance and overall model performance. For the Poisson regression part of the ZIP model, a log of the population (number of people) in different age groups was included as the main group of parameters, complemented by median income and total length of roads in the spatial grid zone. The number of nightlife spots, modeled as the number of restaurants, was also added to the model. This variable was in effect only during the weekends. Finally, day of the week and hour of the day were included as categorical variables. For the logistic regression (zero inflation) part of the ZIP model, total population in the spatial grid zone was included along with the road length. The resulting coefficients after fitting the ZIP model for the three counties separately are shown in Table 4 and 5.

The model parameters (regression coefficients) can be interpreted as follows for Poisson regression (Table 4): for one unit increase in non-categorical independent variable (e.g. weekend nightlife spots) the predicted number of calls is multiplied by e^{coeff} ($e^{0.0057} = 1.0057$ for county J) holding all other variables constant. The coefficients for categorical variables (day of week and hour) are interpreted relative to the reference (0=Sunday; hour 0 = 00:00 until 01:00). Thus during hour 1 (from 01:00 until 02:00) there will be $e^{-0.052} = 0.9496$ (for county J) times calls compared to the number of calls during the reference hour.

Table 4. Results, model parameters for Poisson regression

Variables	Model parameters		
	County J	County E	County O
Log of population			
age group 0-6	-0.138	-0.009	-0.096
age group 7-15	-0.085	-0.061	-0.043
age group 16-19	0.062	0.061	0.055
age group 20-24	0.036	0.088	0.218
age group 25-44	0.076	0.012	0.039
age group 45-64	0.182	0.444	0.118
age group 65-100	0.836	0.511	0.731
Median income	-9.87E-07	-3.07E-06	-4.27E-06
Road length	-3.40E-07	-1.24E-06	-4.65E-07
Weekend nightlife spots	0.0057	0.0084	0.0005
Day of week (0 = Sunday as reference)			
1	-0.017	0.068	0.035
2	-0.022	0.049	0.037
3	-0.003	0.012	-0.016
4	-0.046	0.042	-0.003
5	0.010	0.075	0.020
6	0.033	0.013	0.017
Hour (0 as reference)			
1	-0.052	-0.086	-0.072
2	-0.115	-0.189	-0.196
3	-0.233	-0.266	-0.329
4	-0.380	-0.446	-0.452
5	-0.308	-0.351	-0.417
6	-0.097	-0.281	-0.318
7	0.191	0.013	-0.022
8	0.569	0.531	0.441
9	0.634	0.618	0.587
10	0.624	0.703	0.626
11	0.701	0.743	0.577
12	0.592	0.603	0.487
13	0.541	0.601	0.465
14	0.479	0.626	0.503
15	0.599	0.568	0.485
16	0.517	0.580	0.424
17	0.474	0.427	0.423
18	0.529	0.447	0.413
19	0.432	0.443	0.420
20	0.476	0.343	0.409
21	0.370	0.303	0.279
22	0.236	0.260	0.206
23	0.131	0.122	0.047
constant	-5.958	-5.698	-5.159

Table 5. Results, model parameters for logistic regression

Variables	Model parameters		
	County J	County E	County O
Total population	-0.0018	-0.0004	-0.0023
Road length	-1.3E-05	-3.5E-05	-1.4E-05
constant	1.1661	0.4151	0.5004

The total number of calls predicted by the ZIP model is compared to the actual number of calls in 2014 and to the forecast of the current model in Table 6.

Table 6. Total number of calls (2014), actual and forecasted

	Number of calls		
	County J	County E	County O
Actual	21123	37396	141536
ZIP model	20337	38565	133773
Current model	20223	39055	136061

The model performance was also evaluated by calculating the forecast errors for each spatial grid zone and time slot (day of week * hour) aggregated for the entire 2014 (Table 7).

Table 7. Forecast errors

	Number of calls		
	County J	County E	County O
ME			
ZIP model	0.0026	-0.0048	0.0237
Current model	0.0030	-0.0069	0.0167
MAE			
ZIP model	0.0542	0.1042	0.2600
Current model	0.0565	0.1149	0.2789
RMSE			
ZIP model	0.2875	0.4455	0.7869
Current model	0.3438	0.5200	0.8950

5 Discussion

The resulting model parameters in Table 4 show some interesting features, e.g. looking at the age groups, the older population (65-100) seem to generate more ambulance calls, something that also was stated in the interviews with the ambulance dispatchers. That the population between 0 and 15 would be the group with the least need of ambulances is less intuitive. Possibly this might be explained by the fact that parents of small children often have access to a car and can drive to the hospital. While the median income having a small negative impact on the number of calls can be confirmed by previous studies (e.g. [1]), it seems counterintuitive that the road length also would have a negatively correlated relationship to the dependent variable.

Some of the variables that are not used in the ZIP model, even though data exists, are day and night population, birth place and employment rate. When trying to use day population to estimate the number of people in a zone during day time, the forecast was worse than when just using the total population, which most likely is due to the unreliable data used for day population estimates. Birth place and employment rate

did not improve the results either, and thus were removed.

In Table 5, the parameters should be positive if a high variable value increases the likelihood of having zero calls. Thus, negative values for both population and road length are to be expected. Possibly, the use of road length in both the Poisson regression and the logistic regression, can explain the unexpected parameter values in the Poisson regression. As the variable already influences the forecast in the logistic regression, decreasing the probability of having zero calls where there is a lot of roads, this is compensated by reducing expected number of calls that can be explained by road length in Poisson regression part of the model.

Looking at the total number of actual and forecasted calls for 2014 (Table 6), the ZIP model performs slightly better than the current model for two of the counties (J and E). Both forecasting models either underestimates (J and O) or overestimates (E) the number of calls. The underestimation is probably due to the models' inability to capture a positive trend, as the number of calls increases from 2013 to 2014, while the overestimation is due to a negative trend in County E.

From a practical perspective, it is interesting to discuss whether under- or overestimations are more serious than the other. Should the forecast be used for planning the number of necessary resources, it could be argued that it is important not to underestimate the number of calls and risk resource shortages. In our case, however, the forecast is used to make operational decisions, and thus trade-offs between different time periods and geographical areas. Therefore, it does not matter very much if the forecast is under- or overestimating the aggregated number of calls. If however, systematic under- or overestimations for specific zones or time periods are discovered, these should be corrected as soon as possible.

The forecast errors in Table 7 confirm that the ZIP model performs better than the current model. The ME for County O is the only measure where the current model has a better value. Still, the current model is not much worse compared to the more advanced ZIP model. The reason may be that the population, which is used as an independent variable in both models, is so dominant compared to the other (spatial) factors included in the ZIP model. Concerning temporal factors, both models handle these in a similar way, and can consequently be expected to give similar results.

Even if the new model gives better results, it is a relevant question whether it is worth the effort to switch to a new, more complex model. The current model has the advantage that it is easy to calculate and update and takes into account population changes

between the years. Advantages for the ZIP model, apart from the improved forecasts, is that it can give some information about which factors actually influence the number of ambulance calls, which can be used in preventive work, trying to lower the number of calls. This information can also be used in comparative studies between the counties. If the same parameter values are used, the ZIP model can be easily updated when new input data is available. If new parameter values are desired, the model has to be calibrated using adequate software, something that probably is more complicated than updating the current model.

It may be noted that the forecast accuracy depends on the spatial and temporal resolution; it is easier to accurately forecast the number of calls for a large area than for a small one. However, the difference in zone size between the three countries does not affect the results in a conclusive way. While the errors in general are smallest for the county with the largest zones (County J with 6x6 km zones), County O, which has larger zones than County E, has much larger errors than County E. Evidently, other aspects in the county, such as the population distribution and the number of zones with zero historical calls, affect the errors.

6 Conclusions and future work

The forecasting model developed using a zero-inflated Poisson regression approach and described in this paper has shown to be more accurate than the current model implemented in the EMS information system used by ambulance dispatchers in Sweden. The model development has also helped to identify the factors contributing to temporal and spatial variations in the demand for EMS. During the modelling work, several directions for future work and model improvement were outlined.

Some temporal component estimating the trend could be added to the model. Even if the ZIP model can take into account changes in the explanatory variables, such as the population, this data is often already old when made available. When trying to estimate the number of calls for next year, there will not exist data on the number of people in each zone next year, possibly only for the last year, which then will be two years inaccurate. Also, the information about the population numbers in each zone is very static and does not fully capture where people actually are at any given time. That is why both the current and the new model will miss the increased number of calls generated in a zone due to temporary gathering of people (e.g. large sports and entertainment events).

A remedy for this would be to acquire real time data on the estimated number of people in a zone (e.g. through mobile phone tracking), and make a real time

forecast of the expected number of ambulance calls based on the current situation.

Another issue with the described forecasting models is that for several factors, pointed out as important in the interviews with the ambulance dispatchers, it was not possible to find data retrospectively in the required spatial-temporal resolution. This includes traffic flow and weather data where the former should have a natural correlation with number of ambulance calls, especially for calls to traffic accidents (accounting for approximately 2% of calls in the dataset) but also indicating where people are, rather than where they supposedly live. Weather has been previously proven to affect the number of calls [25], and while it may be possible to use real time weather data, such as the temperature, it is probably easier to use weather forecasts as input variables, as in [24].

On the other hand, a likely future development will be to complement a forecast based on historical data with real time data, which is currently becoming more and more accessible, e.g. for traffic flow, weather, location of the population or special events. This could then be used as input to the EMS information system, and automatically updated based on the forecast accuracy.

7 References

- [1] N. E. Adler, W. T. Boyce, M. A. Chesney, S. Folkman, and S. L. Syme, "Socioeconomic inequalities in health. No easy solution", *JAMA*, vol. 269, pp. 3140-5, Jun 23-30 1993.
- [2] C. A. Aldrich, J. C. Hisserich, and L. B. Lave, "An analysis of the demand for emergency ambulance service in an urban area", *American Journal of Public Health*, vol. 61, pp. 1156-1169, 1971.
- [3] T. Andersson and P. Vårbrand, "Decision Support Tools for Ambulance Dispatch and Relocation", in *Operational Research for Emergency Planning in Healthcare: Volume 1*, N. Mustafee, Ed., 1st ed London: Palgrave Macmillan UK, 2016.
- [4] J. Baker and K. Fitzpatrick, "Determination of an optimal forecast model for ambulance demand using goal programming", *Journal of the Operational Research Society*, vol. 37, pp. 1047-1059, 1986.
- [5] Carmenta. (2018, June 14). *Carmenta CoordCom*. Available: <https://www.carmenta.com/en/products/carmenta-coordcom>
- [6] Carmenta. (2018, June 14). *Carmenta ResQMap*. Available: <https://www.carmenta.com/en/products/carmenta-resqmap>
- [7] N. Channouf, P. L'Ecuyer, A. Ingolfsson, and A. N. Avramidis, "The application of forecasting techniques to modeling emergency medical system

- calls in Calgary, Alberta", *Health Care Management Science*, vol. 10, pp. 25-45, 2007.
- [8] S. Coxe, S. G. West, and L. S. Aiken, "The analysis of count data: a gentle introduction to poisson regression and its alternatives", *J Pers Assess*, vol. 91, pp. 121-36, Mar 2009.
- [9] D. Cramer, A. A. Brown, and G. Hu, "Predicting 911 calls using spatial analysis", in *Software engineering research, management and applications 2011*, ed: Springer, 2012, pp. 15-26.
- [10] E. Gijo and N. Balakrishna, "SARIMA models for forecasting call volume in emergency services", *International Journal of Business Excellence*, vol. 10, pp. 545-561, 2016.
- [11] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy", *International Journal of Forecasting*, vol. 22, pp. 679-688, 2006.
- [12] H. Jang, S. Lee, and S. W. Kim, "Bayesian analysis for zero-inflated regression models with the power prior: applications to road safety countermeasures", *Accid Anal Prev*, vol. 42, pp. 540-7, Mar 2010.
- [13] R. D. Kamenetzky, L. J. Shuman, and H. Wolfe, "Estimating need and demand for prehospital care", *Operations research*, vol. 30, pp. 1148-1167, 1982.
- [14] T. O. Kvålseth and J. M. Deems, "Statistical models of the demand for emergency medical services in an urban area", *American Journal of Public Health*, vol. 69, pp. 250-255, 1979.
- [15] D. Lambert, "Zero-inflated Poisson regression, with an application to defects in manufacturing", *Technometrics*, vol. 34, pp. 1-14, 1992.
- [16] J. A. Lowthian, D. J. Jolley, A. J. Curtis, A. Currell, P. A. Cameron, J. U. Stoelwinder, *et al.*, "The challenges of population ageing: accelerating demand for emergency ambulance services by older patients, 1995-2015", *Medical Journal of Australia*, vol. 194, p. 574, 2011.
- [17] D. S. Matteson, M. W. McLean, D. B. Woodard, and S. G. Henderson, "Forecasting emergency medical service call arrival rates", *The Annals of Applied Statistics*, pp. 1379-1406, 2011.
- [18] V. Nicoletta, E. Lanzarone, A. Guglielmi, V. Bélanger, and A. Ruiz, "A Bayesian Model for Describing and Predicting the Stochastic Demand of Emergency Calls", in *International Conference on Bayesian Statistics in Action*, 2016, pp. 203-212.
- [19] H. Setzler, C. Saydam, and S. Park, "EMS call volume predictions: A comparative study", *Computers & Operations Research*, vol. 36, pp. 1843-1851, 2009.
- [20] K. F. Siler, "Predicting demand for publicly dispatched ambulances in a metropolitan area", *Health Services Research*, vol. 10, p. 254, 1975.
- [21] J. E. Svenson, "Patterns of use of emergency medical transport: a population-based study", *The American journal of emergency medicine*, vol. 18, pp. 130-134, 2000.
- [22] R. A. Waalewijn, J. G. P. Tijssen, R. W. Koster, and R. De Vos, "Survival models for out-of-hospital cardiopulmonary resuscitation from the perspectives of the bystander, the first responder, and the paramedic", *Resuscitation*, vol. 51, pp. 113-122, 2001.
- [23] J. L. Vile, J. Gillard, P. R. Harper, and V. A. Knight, "Predicting ambulance demand using singular spectrum analysis", *Journal of the Operational Research Society*, vol. 63, pp. 1556-1565, 2012.
- [24] H.-T. Wong and P.-C. Lai, "Weather factors in the short-term forecasting of daily ambulance calls", *International journal of biometeorology*, vol. 58, pp. 669-678, 2014.
- [25] H. Wong and P. Lai, "Weather inference and daily demand for emergency ambulance services", *Emergency Medicine Journal*, vol. 29, pp. 60-64, 2010.
- [26] K. Yousefi Mojir and S. Pilemalm, "Actor-centred emergency response systems: a framework for needs analysis and information systems development", *International Journal of Emergency Management*, p. 403, 2016.
- [27] Z. Zhou and D. S. Matteson, "Predicting Ambulance Demand", in *21th ACM SIGKDD international conference on knowledge discovery and data mining*, 2015, pp. 2297-2303.
- [28] Z. Zhou, D. S. Matteson, D. B. Woodard, S. G. Henderson, and A. C. Micheas, "A spatio-temporal point process model for ambulance demand", *Journal of the American Statistical Association*, vol. 110, pp. 6-15, 2015.