Forecasting emergency admissions due to respiratory diseases in high variability scenarios using time series: A case study in Chile

Miguel Becerra a, Alejandro Jerez a,∗, Bastián Aballay a, Hugo O. Garcés b, Andrés Fuentes a

a Departamento de Industrias, Universidad Técnica Federico Santa María, Av. España 1680, Casilla 110-V, Valparaíso, Chile
b Computer Science Department, Universidad Católica de la Santísima Concepción, Alonso de Ribera 2850, Concepción 4090541, Chile

HIGHLIGHTS

• Fulfillment emergency admissions due respiratory diseases is challenging in winter season.
• Exploratory data analysis and hyperparameters grid search was implemented.
• SARIMA modelling is an adequate modelling approach for Chilean time series.
• MAPE is under 20% considering 30-day ahead forecasting process.
• Improvement hospital resource management eases dealing with admission variability.

ABSTRACT

Respiratory diseases are ranked in the top ten group of the most frequent illness in the globe. Emergency admissions are proof of this issue, especially in the winter season. For this study, the city of Santiago de Chile was chosen because of the high variability of the time series for admissions, the quality of data collected in the governmental repository DEIS (selected period: 2014–2018), and the poor ventilation conditions of the city, which in winter contributes to increase the pollution level, and therefore, respiratory emergency admissions. Different forecasting models were reviewed using the Akaike Information Criteria (AIC) with other error estimators, such as the Root Mean Square Error (RMSE), for selecting the best approach. At the end, Seasonal Autoregressive Integrated Moving Average (SARIMA) model, with parameters \(p(d,q)(P.D.Q)_s = (2,1,3)(3,0,2)_7\), was selected. The Mean Average Percentage Error (MAPE) for this model was 7.81%. After selection, an investigation of its performance was made using a cross-validation through a rolling window analysis, forecasting up to 30 days ahead (testing period of one year). The results showed that error do not exceed a MAPE of 20%. This allows taking better resource managing decisions in real scenarios: reactive staff hiring is avoided given the reduction of uncertainty for the medium term forecast, which translates into lower costs. Finally, a methodology for the selection of forecasting models is proposed, which includes other constraints from resource management, as well as the different impacts for social well-being.

Keywords:
Time series
Emergency admissions
Respiratory diseases
Forecasting
Hospital management
SARIMA

Article info

Article history:
Received 21 July 2019
Received in revised form 12 October 2019
Accepted 13 October 2019
Available online 23 November 2019

Editor: Scott Sheridan

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

The diseases related to respiratory problems belong to the top ten group of the most frequent in the world (Forum of
The capital of Chile, Santiago, is an excellent case for studying the impact of unhealthy air quality. The population is approximately 5.2 million (Instituto Nacional de Estadística (INE), 2017) and the city has geographic conditions that prevent adequate ventilation (Ragsdale et al., 2013). This last factor is relevant, as poor air quality conditions are sustained for up to 5 days (Mazzeo et al., 2018). This triggers the activation of many protocols already put in place (Plan of Atmosphere Decontamination). For instance, sports activities are suspended in schools, vehicular constraints are in place, and the city has geographic conditions that prevent adequate ventilation (Ragsdale et al., 2013). This last factor is relevant, as poor air quality conditions are sustained for up to 5 days (Mazzeo et al., 2018).

The public health service shows capacity problems when high user demand appears, especially in the Emergency Service. For example, patients have to wait several hours to get medical assistance. Also, medical staff work long shifts, having insufficient infrastructure, and out-of-date equipment (Goic, 2015).

The most critical period is winter, due to the increase of respiratory diseases given the environment pollution of Santiago (Franck et al., 2015). Therefore, the Hospital Administration needs to manage properly the staff requirement (e.g. winter vaccination campaigns), medical supplies and stretchers (Carvalho-Silva et al., 2018). Furthermore, given the fact that 30–40% of the total admissions of the Emergency Service presents respiratory causes (Departamento de Estadísticas e Información de Salud, 2019), the possibility of forecasting would become a useful tool. This allows the peaks of the season to be mitigated, and to absorb the inherent variability of users.

This topic is not novel in the literature. Several researchers have worked in performing hospital admission forecasting for different diseases. Generalized Autoregressive Moving Average (GARMA) models have been used to forecast weekly counts of patients with influenza at Hopkins Hospital Emergency Department in Baltimore (Jalalpour et al., 2015). Carvalho-Silva et al. (2018) performed a comparison between Autoregressive Integrated Moving Average (ARIMA) and Exponential Smoothing forecasting models for patients’s arrivals at the Emergency Department, up to a week ahead in Braga Hospital, Portugal. ARIMA models have been compared to Artificial Intelligence approaches, such as: Artificial Neural Network (ANN), Random Forest (RF) and Gradient Boosting Machines (GBM), in order to forecast daily hospital admissions due to respiratory causes, across the region of Madrid (Navares et al., 2018). Also, Whitt and Zhang tested a Seasonal ARIMA with exogenous regressors (SARIMAX), such as temperature and calendar variables (holidays), to forecast daily arrivals and hourly occupancy at Rambam Hospital in Haifa, Israel (Whitt and Zhang, 2019).

Currently, the approaches evaluate the predictive capacity of the models for short term forecast: one day or week ahead at the most. This restricts resource management, making it difficult to take an optimal decision for medium term (weeks and months). Also, the use of exogenous variables requires the prior development of an independent procedure to forecast them, as well as a reliable source of information that provides them.

Therefore, this research aims to identify a reliable modelling approach for forecasting the respiratory emergency admissions in the medium term. This would allow cost to be reduced and support to be given to other initiatives, as well to respond better to society’s demand in the season peak.

This work is organized as follows: Section 2 presents a description of the geographic area and the data collection used in this study. Section 3 establishes the methodology and data processing. In Section 4, the results and performance analysis are exhibited. The discussions are addressed in Section 5, where recommendations to improve the current approaches are made. Fig. 1 summarizes this standard linear approach for forecasting research. Finally, Section 6 presents the main conclusions of the work and future research guidelines are suggested.