

# Forecasting Daily Arrivals and Peak Occupancy in a Combined Emergency Department

**Jalmari Tuominen** (✉ [jalmari.tuominen@tuni.fi](mailto:jalmari.tuominen@tuni.fi))

Faculty of Medicine and Health Technology, Tampere University

**Antti Roine**

Department of Surgery, Faculty of Medicine and Health Technology, Tampere University

**Taavi Saviauk**

Unitary Healthcare Ltd

**Anssi Seppo**

Unitary Healthcare Ltd

**Miika Pihlaja**

Unitary Healthcare Ltd

**Jani Ovaska**

Emergency Division of Pirkanmaa Hospital District, Tampere University Hospital and Tampere University

**Satu-Liisa Pauniahho**

Emergency Division of Pirkanmaa Hospital District, Tampere University Hospital and Tampere University

**Niku Oksala**

Faculty of Medicine and Health Technology, Tampere University

---

## Research Article

**Keywords:** Emergency department, Overcrowding, Forecasting, Occupancy

**Posted Date:** January 8th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-138768/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

**Objective** Emergency department (ED) crowding is a global problem associated with negative patient outcomes such as mortality and prolonged length of stay. Forecasting overcrowding would enable pre-emptive strategical maneuvers and is a subject of constant academic interest. However, most studies focus on forecasting arrivals in United States ED setting. We propose a novel and intuitive crowding metric called daily peak occupancy and assess forecasting ED crowding in a Nordic Combined ED using both established and novel predictive algorithms.

**Methods** All episodes of care in Tampere University Hospital ED were acquired from December 1, 2014 to June 19, 2019, amounting to 488 167 individual events. Predictability of two target variables was investigated: total daily arrivals (TDA) and daily peak occupancy (DPO) with forecast horizon of one day. Three models were investigated: Seasonal Autoregressive Moving Average (SARIMA), Facebook's Prophet algorithm, and General Linear Model (GLM). Calendar variables were used as independent variables.

**Results** SARIMA outperformed other models in predicting both total daily arrivals and daily peak occupancy with mean absolute percentage errors of 6.6% ( $\pm 5.3$ ) and 12.4% ( $\pm 10.7$ ) respectively. Next day overcrowding can be predicted using SARIMA with an AUC of 0.74 and accuracy of 79 %.

**Conclusion** Predictive models can be utilized in a Nordic emergency medicine setting with similar or better accuracy as previously reported. Predicting future occupancy is possible but more challenging than predicting arrivals. Predicting peaks in demand remains a significant challenge. Future work should focus on investigating the value of exogenous variables in increasing model sensitivity.

## 1 Introduction

Emergency department (ED) crowding is a global problem associated with negative patient outcomes such as mortality, increased number of medical errors and prolonged length of stay (1–3). ED crowding is a complex phenomenon that can be divided into three components as described by Asplin et al. 2003: 1) patient input, 2) throughput, 3) output (4). Major factors affecting input are the number of seriously ill and injured patients, the availability of ambulatory care, and socioeconomic characteristics of the population. In some systems, input can also be adjusted by diverting patients to other centers. Throughput is controlled by the speed of triage and diagnostic evaluation, whereas output is controlled by the availability of patient beds both inside and outside the hospital. Improving patient input or output requires broad, multi-stakeholder effort throughout the healthcare system. Throughput can be directly modulated by the ER management by changes in resourcing and is thus a prime target for optimization (5).

A challenge to resource optimization is the difficulty of predicting the load the ED is exposed to. Much research has been done on predicting the number of arrivals or the number of admissions. In order to better forecast future patient volumes, predictive algorithms with different forecasting horizons and exogenous variables have been investigated (6,7). Typical inputs for predictive algorithms utilize calendar and meteorological variables as input. Time series tools such as Seasonal Autoregressive Integrated Moving Average (SARIMA) have achieved mean average percentage errors (MAPE) of 4–15% on a one day ahead forecast horizon (6). The ability to better predict ED loading would enable hospital administrators to make informed decisions on resourcing and improve ED quality metrics. However, most of the research has been conducted in the United States using arrivals as the target variable with less emphasis on occupancy (7). In this study we aimed to provide a novel perspective on ED forecasting by introducing a new target variable called daily peak occupancy (DPO). DPO is affected by all three components of the Asplin's model in contrast to daily arrivals, which may prove useful when building multivariate models with several exogenous independent variables. Moreover, we hypothesize that DPO may be a more clinically relevant target variable than daily arrivals since high input does not necessarily develop into crowding if compensated by efficient throughput and output.

This study has three goals. First we investigate the applicability of well-documented forecasting models in a Nordic setting. Second, we document the performance of novel Prophet algorithm in the context of ED forecasting. Third, we introduce a novel target variable called DPO and investigate its predictability in terms of both continuous and discrete error measures.

## 2 Materials And Methods

### 2.1 Data

Tampere University Hospital is an academic hospital located in Tampere, Finland serving a population of 900,251 patients and providing level 1 trauma center equivalent capabilities. The ED has 70 beds and approximately 100,000 patients are treated annually. For this study, data on all registered ED visits were obtained from hospital database created during the sample period from December 1, 2014 to June 19, 2019. The requested dataset contained timestamps for admission and discharge, diagnosis, patient's age, gender, identifier of the physician and patient's municipality of residence. Overcrowding is relatively common and an internal protocol has been implemented to help the staff to react to these situations. Two levels of overcrowding are recognized: level 1 when occupancy exceeds 80 patients (70% occupancy) and level 2 when occupancy exceeds 100 patients (88% occupancy). Overcrowding annotations were acquired from a patient logistics system (Uoma, Unitary Healthcare Ltd., Tampere, Finland) and these were used to validate the occupancy reconstruction as described below. This was a retrospective study using anonymized data acquired from the hospital electronic health record. According to Medical Research Act of Finland, retrospective studies base on health records are exempted from review by Ethics committee and it was approved by Tampere University Hospital Research, Development and Innovation Centre.

### 2.2 Study design

This study consist of two phases, both of which aim to predict events one day ahead. In the first phase we perform predictions for total daily arrivals. In the second phase DPO is used as the target variable. *Daily Peak Occupancy (DPO)* is defined as the highest occupancy observed on any given day. Occupancy of the emergency department is not continuously recorded in the hospital database and thus it was not contained in the requested dataset. We reconstructed the hourly occupancy by utilizing an algorithm that used the timestamps of admission and discharge as inputs to output hourly occupancy. Once the hourly occupancy had been reconstructed we resampled the data into daily resolution by extracting the maximum occupancy of each day. The resulting vector was validated by comparing it to overcrowding alerts given by ED staff and recorded using a dedicated patient logistics system as described above. The busiest 25% of the days were labeled as crowded since this threshold has been previously documented to be associated with increased 10-day mortality (8).

Data was missing for a time period from December 1, 2014 to May 1, 2015. This time period was excluded and the time period from June 1, 2015 to June 19, 2019 was used in further analysis. We also observed significant discontinuity in the reconstructed hourly occupancy for a time period from December 1, 2014 to June 1, 2016 potentially due to changes in the internal coding of the different sections of the ED. Thus, the time period from June 1, 2016 to June 19, 2019 was used in phase 2.

### 2.3 Analysis

Statistical analyses were performed using Python 3.7.3. Jupyter Lab 1.0.4 powered by iPython 7.8.0 was used in the exploratory data analysis (9). Analytical tools and data structures were provided by Pandas 0.25.0 (10). Visualizations were created using Matplotlib 3.1.1. Predictive models were created using StatsModels 0.10.1 (11), Scikit-learn (12) and Prophet (13). SARIMA hyperparameters were optimized using pmdarima (14). The dependencies were managed using Anaconda 4.7.12. An exhaustive list of all dependencies is provided in the Appendix.

### 2.4 Predictive models

Both *simple naïve* and *seasonal naïve* models were included as benchmarks. In case of simple naïve, the value of the previous day is used as the prediction for tomorrow. Seasonal naïve on the other hand works by observing the corresponding value a season ago and using this value for predicting the future. For example, when forecasting the occupancy peak of the coming

Tuesday, the peak of last Tuesday was used. Some earlier publications utilize the seasonal naïve as the baseline to compare different methods to (15). Others suggest linear regression as the benchmark method (16).

*Seasonal Autoregressive Moving Average (SARIMA)* is widely used in both econometrics and ED forecasting and has been established as a benchmark model in international forecasting competitions (17)(7). SARIMA consists of 7 independent components denoted by  $(p, d, q) \times (P, D, Q)_s$ . The parameters  $p, d,$  and  $q$  describe the autocorrelation, differencing, and size of the moving average window of the non-seasonal component of the models, whereas  $P, D,$  and  $Q$  describe the same values for the seasonal component.  $S$  parameter describes the length of the season and was assigned a value of 7 due to the well-known weekly seasonality of the target variables. We report results for both SARIMA and SARIMAX, the latter of which was trained using weekdays and months as additional independent variables.

*General Linear Model (GLM)* has been widely used in predicting ED visits and is regarded as the benchmark against which other models are compared (16). Weekdays and months were used as independent variables and a Poisson distribution was assumed for both target variables.

*Prophet* is a modular regression model created by Facebook. It has three main components: trend, seasonality, and holidays and also supports adding exogenous predictors into the linear component of the model. The potential of Prophet in ED forecasting has been mentioned but as far as we know no results have so far been published in the peer-reviewed literature although results in flu epidemic peak intensity prediction have been promising (18,19). Prophet is tested both with and without calendar variables.

## 2.5 Cross-validation

Both SARIMA and Prophet evaluate the complete historical vector to perform predictions into the future. We utilized a time series cross-validation to simulate moving forward in time while iteratively updating the models with new data. In both phases the data was divided into train and validation sets using December 31, 2017 as the split date. The resulting train sets were used to evaluate the initial hyperparameters of SARIMA and Prophet. We then iterated through the resulting validation set from January 1, 2018 to June 19, 2019 with steps of one day, performing predictions one day forward on each iteration and updating model parameters. GLM was trained using a simple holdout using the train and validation sets described above.

## 2.6 Accuracy measures

*Mean Absolute Percentage Error (MAPE)* is the most used accuracy metric in the field of ED forecasting (7) and is widely used in time series forecasting in general. MAPE has the advantage of being scale-independent, thus allowing model comparisons between different datasets. For these reasons, MAPE was selected as the preferred continuous error measure in phases 1 and 2. The formula for MAPE is as follows:

$$MAPE = \frac{100}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}$$

where  $n$  = number of samples,  $y_i$  = ground truth,  $\hat{y}_i$  = prediction.

In phase 2 we evaluate the accuracy of the models as binary predictors with the goal of forecasting whether the following day will be crowded or not which calls for discrete error measures. We report *Receiver operating characteristics (ROC)* and calculate the *Area under the curve (AUC)* measure along with reporting overall *accuracy, sensitivity* and *specificity* of the models.

## 3 Results

### 3.1 Descriptive statistics

In the sample period we observed 488,167 episodes of care and 131,644 individual patients. Of these patients 48% were male, 52% female and the average age was 52.7 years (95% CI 52.6–52.7). On average there were 261 ( $\pm 25$ ) daily arrivals and the average daily peak occupancy was 75 ( $\pm 13$ ). The 75th percentile was at the occupancy of 85 patients. A weekly seasonality was observed in the target variables and an hourly seasonality in the underlying hourly vectors. Number of arrivals was lowest at 5 a.m. with 3.5 ( $\pm 2.1$ ) patients on average, and highest at 2 p.m. with 17.9 ( $\pm 4.5$ ) patients. Lowest occupancy was observed at 6 a.m. with 17.6 ( $\pm 5.6$ ) patients and highest at 4 p.m. with 71.6 ( $\pm 14.3$ ) patients. In terms of weekly seasonality, Thursdays had the lowest number of daily arrivals with 246 ( $\pm 19.4$ ) patients on average and Saturdays were the busiest with 276 ( $\pm 23.9$ ) patients. On average, Sundays had the lowest DPO's with 64.4 ( $\pm 10.6$ ) patients whereas Mondays had the highest with 85 ( $\pm 11.6$ ) patients. Please see Fig. 1 for details.

## 3.2 Model performance

In phase 1  $SARIMAX(1, 0, 1)$  with calendar variables was identified as the most accurate model in predicting TDA with MAPE of 6.6% ( $\pm 5.3$ ). Prophet performed with MAPE 6.7% ( $\pm 6.4$ ), which was superior to GLM, which performed with MAPE of 8.1% ( $\pm 6.6$ ). Omitting calendar variables resulted in better performance when using Prophet (6.7 vs. 6.9 %). All models bested the benchmark models (Table 1, Fig. 2).

Table 1

Model accuracies in terms of absolute percentage errors when predicting total daily arrivals. Calvars = calendar variables.  $SARIMA(X)$  = Seasonal Autoregressive Integrated Moving Average (Exogenous), GLM = General Linear Model.

	Naive	Seasonal Naive	GLM:calvars	SARIMA:none	Prophet:calvars	Prophet:none	SARIMAX:calvars
Count	535	535	535	535	535	535	535
Mean	8.4	8.2	8.1	7.0	6.9	6.7	6.6
Std	6.4	6.6	6.6	5.4	5.6	5.4	5.3
25%	3.5	3.2	3.1	2.8	2.7	2.7	2.6
50%	6.9	6.6	6.8	5.7	5.6	5.6	5.4
75%	12.3	11.6	11.5	10.2	9.9	9.9	9.5
Max	36.4	41.7	40.3	33.6	41.5	41.8	31.8

In phase 2  $SARIMA(2, 1, 1) \times (1, 0, [1, 2], 7)$  without calendar variables was identified as the most accurate model with out-of-sample MAPE of 12.4% ( $\pm 10.7$ ). Prophet performed with 13.1% ( $\pm 11.6$ ) and GLM with 14.5% ( $\pm 13$ ) respectively. All models outperformed the benchmarking models. SARIMA, with definitions described above, was the best predictor in terms of binary performance as well, providing an accuracy 79 %. Both Prophet and SARIMA without exogenous variables provided an AUC of 0.74 as did GLM although with temporally inconsistent manner. In terms of binary prediction, all models were high on specificity but low on sensitivity with unadjusted sensitivities ranging from 50 % on GLM to 37 % on SARIMA and specificities ranging from 89 % on SARIMA to 77 % on GLM. (Table 2, Figs. 3 and 4). Detailed model specifications are available in the Appendix.

Table 2

Model accuracies when predicting daily peak occupancy. Continuous error measures are reported in terms of absolute percentage errors. Calvars = calendar variables. SARIMA(X) = Seasonal Autoregressive Integrated Moving Average (Exogenous), GLM = General Linear Model.

	Naive	Seasonal Naive	GLM:calvars	Prophet:calvars	Prophet:none	SARIMAX:calvars	SARIMA:none
Count	535	535	535	535	535	535	535
Mean	18.8	16.0	14.5	13.5	13.1	12.5	12.4
Std	15.0	12.9	13.0	11.9	11.5	11.1	10.7
25%	6.7	6.5	5.4	5.4	5.5	4.7	5.0
50%	15.8	12.7	11.5	10.8	10.1	9.5	9.7
75%	26.7	22.6	19.7	18.7	18.2	16.8	16.6
Max	100.0	83.0	94.3	81.9	79.4	72.5	72.9
Accuracy	0.69	0.72	0.77	0.76	0.76	0.78	0.79
Sensitivity	0.25	0.34	0.36	0.33	0.36	0.31	0.29
Specificity	0.80	0.82	0.87	0.88	0.87	0.90	0.92

## 4 Limitations

This study is limited by its retrospective nature and the generalizability of the results should be confirmed in a prospective setting. Despite validating the reconstructed occupancy as described above, some uncertainty remains on the correctness of the target variable. Only few independent variables were used and it is likely that the performance would improve e.g. by incorporating holiday variables. However, this was out of scope of our current investigation.

## 5 Discussion

In this study we showed that readily available algorithms can be used to predict not only future arrivals but also future peak occupancies on a daily resolution in a Nordic ED. SARIMA defended its place as the best performing model at least when the number of exogenous features is limited. In terms of discrete error measures, all models demonstrated poor sensitivity but acceptable specificity.

Comparing our results to those of earlier studies is challenging due to significant differences in both study settings and reporting practices. However, we document superior performance of 6.6 % in predicting total daily arrivals compared to Whitt et al. 2019 (20) in which MAPE 8.4 % was achieved using a SARIMAX model. Ekström et al. 2015 documented MAPE 6.5 % in a comparable hospital setting using GLM with internet activity of the preceding day and weekdays as the independent variables (21). It is noteworthy that similar forecasting accuracy was achieved in this study using only static calendar variables as an input.

Our findings demonstrate that forecasting occupancy is more challenging than forecasting arrivals with MAPE 12.4 % compared to 6.6 % using a SARIMA model. This is theoretically sound; DPO is dependent not only on ED input, but also on throughput and output capacities (4). Whether the performance documented in this study is sufficient to provide actionable information for hospital management is debatable. We hypothesize that inclusion of other exogenous variables in a multivariate model with DPO as the target variable could substantially improve the performance of the forecasts. We consider the availability of inpatient beds, number of website visits and ED throughput efficiency characteristics to be especially interesting potential predictors. We suggest that future work should focus on investigating the combined value of these predictors in improving model sensitivity.

DPO as an ED crowding metric is a relevant parameter as negative outcomes are related to occupancy of the ED, not input or throughput *per se*. This is also demonstrated by the fact that many validated scoring systems for ED crowding use parameters such as proportion of beds occupied and longest waiting time as input. The challenge with ED scoring systems is that in many cases their inputs are not readily available or require substantial integration effort to bring them into practice (22). Additionally, many scoring systems are generalizable only to a subset of ED's. For example, NEDOCS uses number of patients in the ventilator as one input of the score (23). Even though undoubtedly correlating with increased mortality, having a significant number of mechanically ventilated patients simultaneously in the emergency department is rare in the Finnish setting as these patients are typically promptly transferred to dedicated intensive care units and may not reflect the perceived crowding in a more general setting. In the future, the predictive power of DPO for negative outcomes and perceived crowding should be investigated.

The majority of attempts to forecast ED visits look at patient population as one surrogate (7). However, patients presenting at the ED vary widely and have different underlying conditions. It could therefore be beneficial to divide the ED into relevant subsections e.g. by department, acuity, diagnosis or method of arrival. An example of meaningful stratification of ED visits was done by Aroua et al. (24). They stratified patients according to 25 classes of major diagnostic categories and achieved low MAPEs of 7–9% for cardiovascular, respiratory, and gastrointestinal diagnoses, whereas MAPE was as high as 68 % for mental and substance abuse disorders. If implemented into a practical model, these findings would have obvious utility in administrative decision-making. In addition, training individual models for a number of relevant subgroups may prove beneficial in increasing model accuracy since both seasonality and the effect of predictors likely differ significantly across different subgroups.

## 5.1 Conclusions

In this study we showed that readily available time series forecasting tools can be used in a Nordic combined ED with similar or better accuracy as previously reported. We also documented the performance of these models as binary predictors of high daily peak occupancy and demonstrated their poor sensitivity in this task. Further work should focus on increasing model sensitivity by identifying a set of useful independent variables and validating their use in a prospective setup.

### List Of Abbreviations

SARIMA(X) Seasonal Autoregressive Integrated Moving Average (Exogenous)

GLM General Linear Model

ED Emergency Department

MAPE Mean Absolute Percentage Error

AUC Area Under the Curve

ROC Receiver Operating Characteristics

TDA Total Daily Arrivals

DPO Daily Peak Occupancy

### Declarations

### Ethics approval

Since our study was retrospective in nature, an approval from the ethics committee was not required. An institutional approval was acquired prior to data collection with following specifications.

Name: Potilaslogistiikan häiriötekijöiden tunnistaminen ja mallintaminen

Number: PSHP/R19565

Date: June 16, 2019

## Consent for publication

Not applicable

## Data availability statement

Not available

## Funding

This study was financially supported by The Finnish Medical Foundation and the Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital and Pirkanmaa Hospital District Grants 9X040.

## Author contributions statements

JT performed statistical analysis with assistance from MP and AR. AS, TS, MP, AR and NO designed the study. SP and NO acquired the data. JO helped interpreting both the data and the results. JT and AR wrote the manuscript. JT, AS, TS, MP, AR, NO, JO and SP commented the manuscript and accepted the final version.

## Competing interests statement

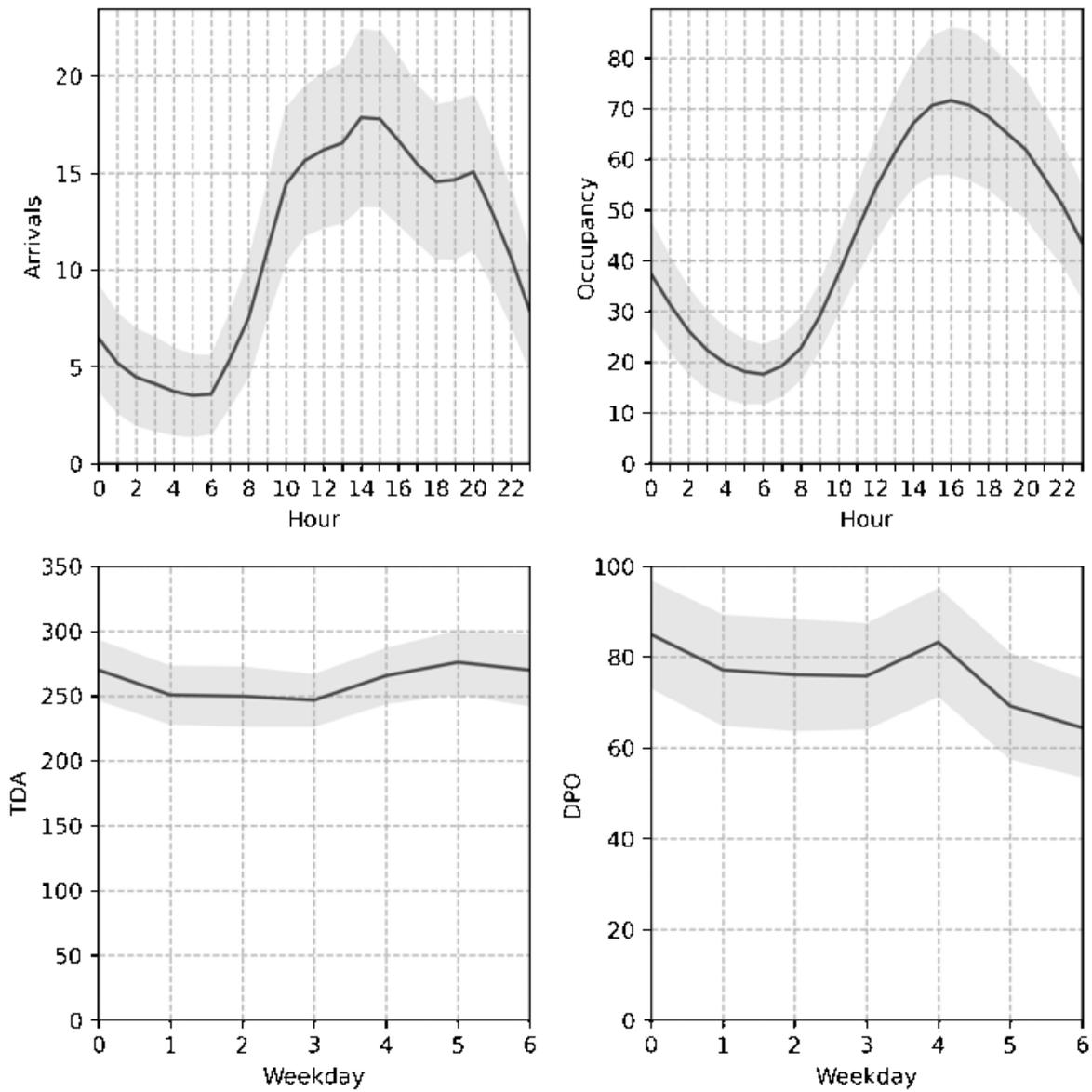
JT, SP, AR and JO declare no conflicts of interest. MP, TS, AS and NO are shareholders of Unitary Healthcare Ltd. which has developed patient logistics system currently used in the study emergency department.

## References

1. Bernstein, S. L. *et al.* The effect of emergency department crowding on clinically oriented outcomes. *Acad. Emerg. Med.* **vol. 16** (no. 1), 1–10 (2009).
2. Eriksson, C. O., Stoner, R. C., Eden, K. B., Newgard, C. D. & Guise, J. M. The Association Between Hospital Capacity Strain and Inpatient Outcomes in Highly Developed Countries: A Systematic Review. *J. Gen. Intern. Med.* **vol. 32** (no. 6), 686–696 (2017).
3. Morley, C., Unwin, M., Peterson, G. M., Stankovich, J. & Kinsman, L. *Emergency department crowding: A systematic review of causes, consequences and solutions*, vol. 13, no.8. 2018.
4. Asplin, B. R. *et al.* A conceptual model of emergency department crowding. *Ann. Emerg. Med.* **vol. 42** (no. 2), 173–180 (2003).
5. Estey, A., Ness, K., Saunders, L. D., Alibhai, A. & Bear, R. A. Understanding the causes of overcrowding in emergency departments in the Capital Health Region in Alberta: A focus group study. *Can. J. Emerg. Med.* **vol. 5** (no. 2), 87–94 (2003).
6. Wargon, M., Guidet, B., Hoang, T. D. & Hejblum, G. A systematic review of models for forecasting the number of emergency department visits. *Emerg. Med. J.* **vol. 26** (no. 6), 395–399 (2009).
7. Gul, M. & Celik, E. An exhaustive review and analysis on applications of statistical forecasting in hospital emergency departments. *Heal. Syst.* **vol. 00** (no. 00), 1–22 (2018).

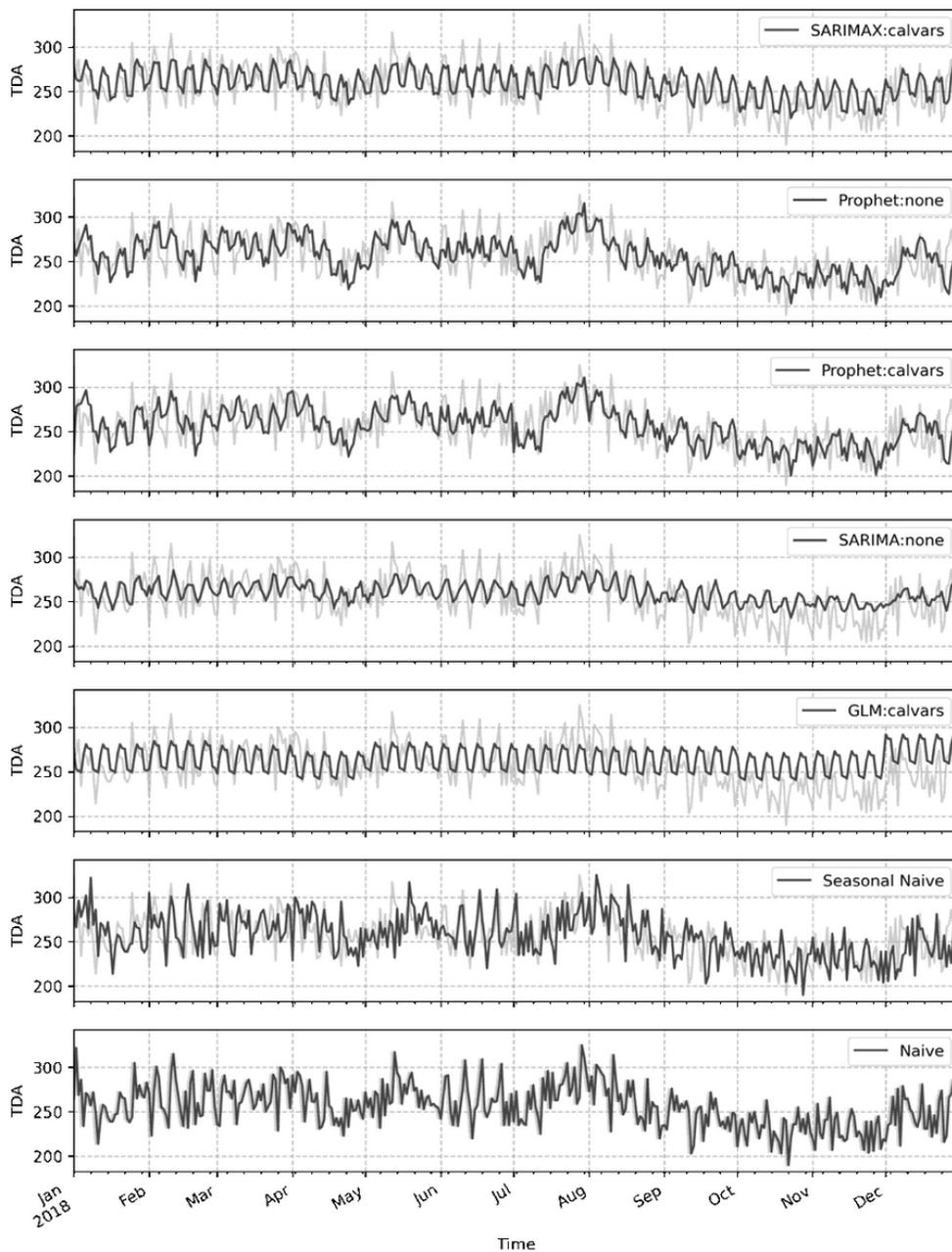
8. Pérez, F. & Granger, B. E. “{IP}ython: a System for Interactive Scientific Computing. [\url{http://ipython.org}](http://ipython.org) Accessed: May 10, 2014.” *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 21–29 2007.
9. McKinney, W. “Data Structures for Statistical Computing in Python,” in *Proceedings of the 9th Python in Science Conference 2010*, pp. 51–56.
10. Hunter, J. D. & Matplotlib A 2D graphics environment. *Comput. Sci. Eng.* **vol. 9** (no. 3), 99–104 (2007).
11. Seabold, S. & Perktold, J. “Statsmodels: Econometric and statistical modeling with python,” in *9th Python in Science Conference 2010*.
12. Pedregosa, F. *et al.* “Scikit-learn: Machine Learning in Python,” *J. Mach. Learn. Res.*, vol. 12, no. Oct, pp. 2825–2830, 2011.
13. Taylor, S. J. & Letham, B. Forecasting at Scale. *Am. Stat.* **vol. 72** (no. 1), 37–45 (2018).
14. T. G. Smith and others, “{pmdarima}: ARIMA estimators for {Python}.”
15. Hertzum, M. Forecasting Hourly Patient Visits in the Emergency Department to Counteract Crowding. *Ergon. Open J.* **vol. 10** (no. 1), 1–13 (2017).
16. Jones, S. S. *et al.* Forecasting daily patient volumes in the emergency department. *Acad. Emerg. Med.* **vol. 15** (no. 2), 159–170 (2008).
17. Park, J. C., Chang, B. P. & Mok, N. 144 Time Series Analysis and Forecasting Daily Emergency Department Visits Utilizing Facebook’s Prophet Method. *Ann. Emerg. Med.* **vol. 74** (no. 4), 57 (2019).
18. Lu, J. & Meyer, S. “Forecasting flu activity in the united states: Benchmarking an endemic-epidemic beta model,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 4, 2020.
19. Richardson, D. B. Increase in patient mortality at 10 days associated with emergency department overcrowding. *Med. J. Aust.* **vol. 184** (no. 5), 213–216 (2006).
20. Whitt, W. & Zhang, X. Forecasting arrivals and occupancy levels in an emergency department. *Oper. Res. Heal. Care.* **vol. 21**, 1–18 (2019).
21. Ekström, A., Kurland, L., Farrokhnia, N., Castrén, M. & Nordberg, M. Forecasting emergency department visits using internet data. *Ann. Emerg. Med.* **vol. 65** (no. 4), 436–442 (2015).
22. Hwang, U. *et al.* Measures of crowding in the emergency department: A systematic review. *Acad. Emerg. Med.* **vol. 18** (no. 5), 527–538 (2011).
23. Weiss, S. J. *et al.* “Clinical practice: NEDOCS,” *Acad. Emerg. Med.*, no. April, pp. 38–50, 2002.
24. Aroua, A. & Abdul-Nour, G. “Forecast emergency room visits-a major diagnostic categories based approach,” *Int. J. Metrol. Qual. Eng.*, vol. 6, no. 2, 2015.

## Figures



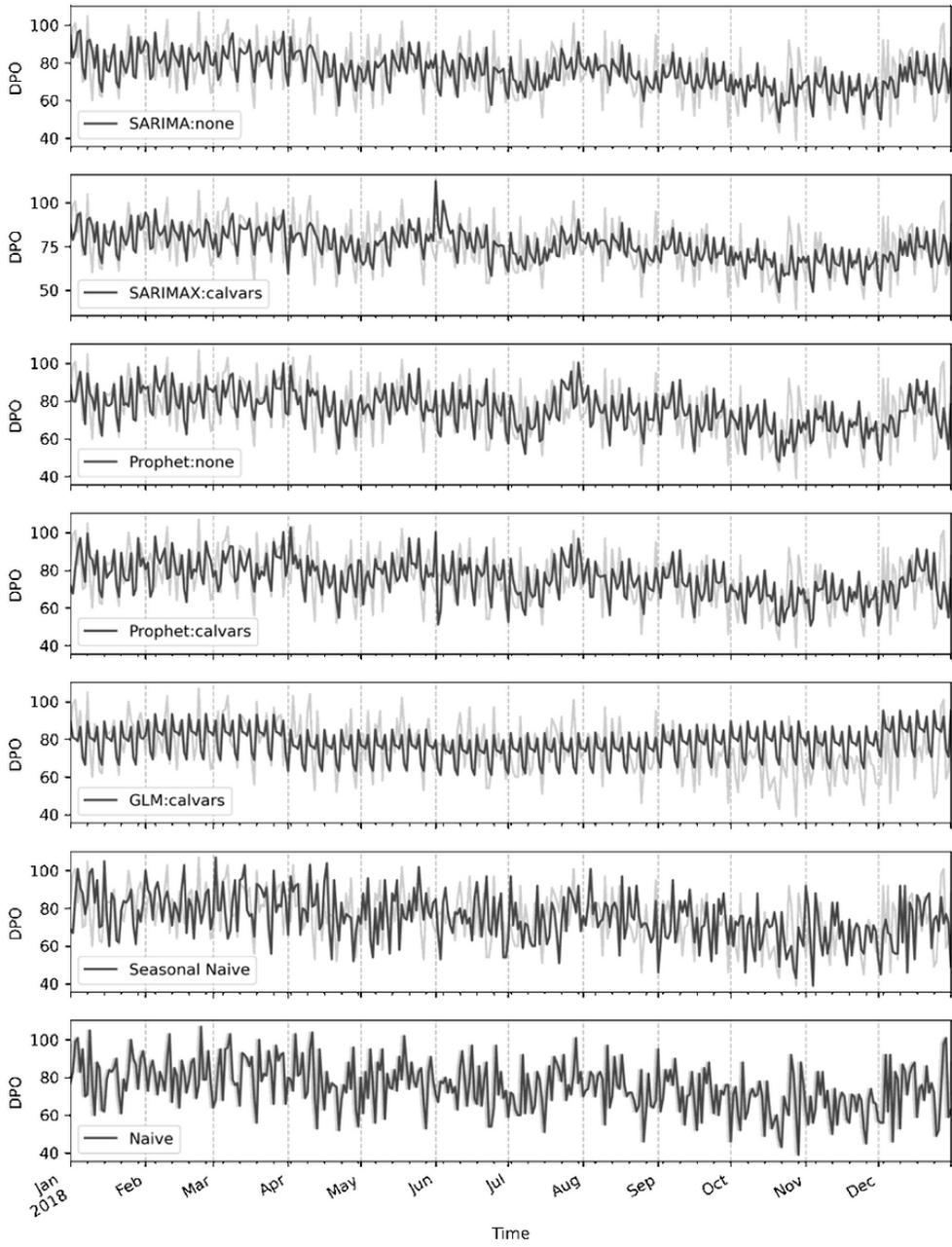
**Figure 1**

Seasonalities of the target variables and the underlying hourly statistics. Dark grey line = mean, light grey area = standard deviation. TDA = Total Daily Arrivals, DPO = Daily Peak Occupancy.



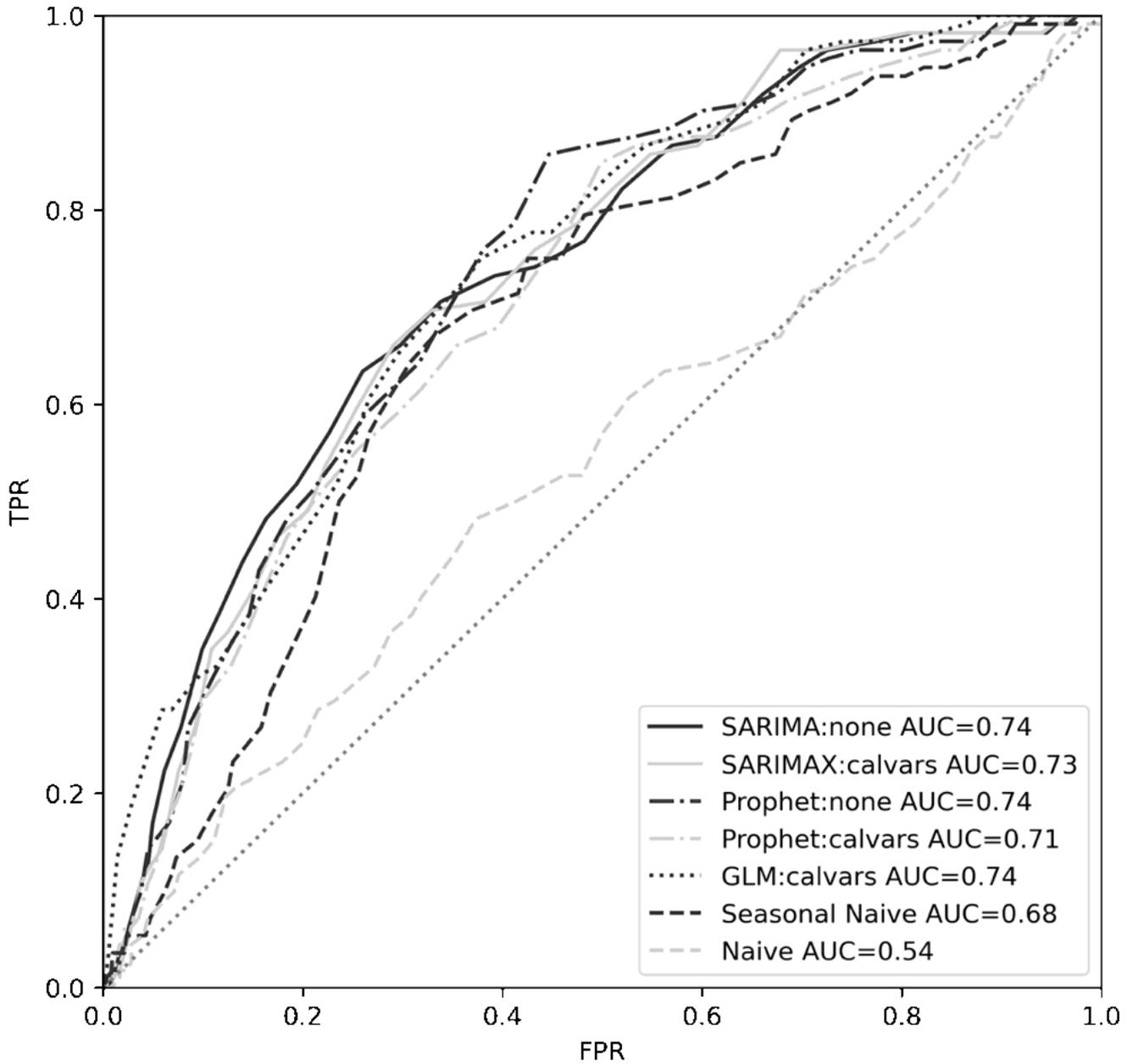
**Figure 2**

Visual representation of model performance when predicting total daily arrivals (TDA). Dark grey line = prediction, light grey line = ground truth. SARIMA(X) = Seasonal Autoregressive Integrated Moving Average (Exogenous), GLM = General Linear Model.



**Figure 3**

Visual representation of model performance when predicting daily peak occupancy (DPO). Dark grey line = prediction, light grey line = ground truth. SARIMA(X) = Seasonal Autoregressive Integrated Moving Average (Exogenous), GLM = General Linear Model.



**Figure 4**

Receiving operating characteristics when predicting next day peak occupancy. AUC = Area Under the Curve, SARIMA(X) = Seasonal Autoregressive Integrated Moving Average (Exogenous), GLM = General Linear Model.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementaryinformation.html](#)