International Journal of Health Sciences
Available online at www.sciencescholar.us
Vol. 9 No. 1, April 2025, pages: 230-243
E-ISSN: 2550-696X
https://doi.org/10.53730/ijhs.v9n1.15567

Modeling and Forecasting Emergency Department Crowding using SARIMA, Holt Winter method, and Prophet models

Mansoor AL Yarubi [a], Nur Arina Bazilah Kamisan [b], Siti Mariam Norrulashikin [c]

Manuscript submitted: 09 November 2024, Manuscript revised: 27 January 2025, Accepted for publication: 18 February 2025

Corresponding Author [c]

Keywords

emergency department;
Holt winter method;
Prophet model;
SARIMA model;
univariate time series;

Abstract

Emergency department (ED) crowding in health care is linked with longer wait times, high mortality rates, and low healthcare quality. Univariate time series models such as Seasonal Autoregressive Integrated Moving Average (SARIMA), Holt-Winters method (HW), and Prophet model (PM) have been widely employed to predict ED crowding. However, there is no consensus on the best fit time series model for ED crowding forecasting. This study compared the predictive precision of three univariate time series models, SARIMA, HW, and PM, in predicting ED crowding at Nizwa Hospital in Oman. The study used hourly patient visits at ED from January to December 2023. The model selection was based on minimizing Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). ED visits showed irregular trends and seasonal effects due to time and day of the week effects. The 24-hour ED visits depicted two peak phases: noon (local maximum) and around 10 PM to midnight (global maximum). The prophet model had better accuracy than the SARIMA and HW models. Adopting the Prophet model predictions can help avoid unexpected ED crowding, reduce waiting times, and improve quality health care management.

Contents

[a] Ministry of Health, Nizwa, Oman
[b] Universiti Teknologi Malaysia, Johor Bahru, Malaysia
[c] Universiti Teknologi Malaysia, Johor Bahru, Malaysia

1 Introduction

Emergency department (ED) crowding is a significant challenge in healthcare worldwide. ED crowding is linked with treatment delays, increased patient mortality rates, operational inefficiencies (Hoot & Aronsky, 2008; Hoot et al., 2009; Pines et al., 2008), admission rates (than discharge), and disease severity (Chen et al., 2020). For instance, delayed attendance to patients with critical conditions such as pneumonia due to ED crowding can increase patient mortality rates (Hoot et al.., 2009). Besides, delays in pain treatment for patients with bone fractures can worsen pain and discomfort (Pines et al., 2006). Overcrowding in EDs can impede efforts to prepare for and respond to acute and emergency health issues like heart attack, pneumonia, and stroke (Chiu et al., 2018). Thus, addressing ED crowding is crucial for improving healthcare delivery and outcomes. Hospitals need emergency preparedness programs to minimize overcrowding in EDs (Jeyaraman et al., 2021). Predictions of ED crowding play a crucial in informing emergency preparedness decisions. Accurate ED demand forecasts can help hospital management optimize staff allocation to handle emergency service demands during peak periods. Factors such as seasonal illnesses and public health emergencies such as road accidents make patient crowding in ED unpredictable (Hoot & Aronsky, 2008).

Several univariate seasonal times series models have been adopted in several studies to forecast ED crowding. The SARIMA and HW are renowned traditional models for predicting seasonal time series data. The two models capture trend and seasonal patterns in time series data but have different efficiencies. SARIMA extends the traditional ARIMA framework by incorporating a seasonal component (Box *et al.*, 2015). Literature shows that SARIMA can effectively predict ED patient visits (Calegari et al., 2016), monthly rates of Chronic obstructive pulmonary disease ED visits in Canada (Rosychuk et al.,2016), and monthly prehospital diabetic emergencies in Austria (Villani et al., 2017), and ED admissions due to respiratory diseases in Chile (Becerra et al., 2020). Among these studies, there is consensus that the single-order SARIMA model is generalizable to all ED visits. SARIMA model orders often differ depending on the periodicity of the data (hourly, monthly, or annual).

The Holt-Winters model is often suitable for short-run forecasting. HW employs smoothing algorithms to account for trends and seasonality. HW is often suitable for short-term forecasting in dynamic environments like EDs since the weights exponentially decrease as the lagged order increases (Hyndman & Athanasopoulos, 2018). Some studies show that the Holt-Winters model outperforms the SARIMA model in predicting ED volumes (Vieira et al., 2023). For instance, Vieira et al. (2023), established Triple Exponential Smoothing; (TES) with 8.16% outperforms SARIMA (MAPE = 8.46%) in forecast weekly admissions to the ED of the Hospital of Braga, Portugal. However, the Holt-Winters method is less effective for predicting ED visits over longer horizons since it places less weight on distance-lagged series observation (Etu et al., 2022).

The Prophet model is an additive model suitable for forecasting seasonal time series data, which was recently developed by Facebook Inc. in 2017. It has been adopted in time series forecasting because it can effectively account for strong seasonal and holiday effects and handle missing data and outliers (Taylor & Letham, 2018). Prophet accommodates seasonal and holiday impact, making it suitable for forecasting ED visits. ED visits are influenced by specific seasonal events such as road accidents or periods such as winter. Recent studies have shown that Prophet forecasts reliable ED visits (Duarte & Faerman, 2019; Feng et al., 2022).

There is a lack of consensus on whether the novel PM outperforms the traditional SARIMA and HW models. PM has outperformed SARIMA in forecasting monthly road traffic injury inpatients of general hospitals in Northeast China (Feng et al., 2022). However, there is a lack of consensus, as other studies have shown that the SARIMA model outperforms the Prophet model. For instance, Tuominen et al. (2021), established that PM (MAPE = 6.7) outperformed SARIMA (MAPE = 7.0) in forecasting total daily arrivals. However, SARIMA was the best (MAPE = 12.4) in forecasting ED crowding using daily peak occupancy, outperforming PM (MAPE = 13.1).

The current study seeks to establish the best forecasting approach for ED crowding among the three competing: SARIMA, HW, and PM. The findings contribute to the growing body of literature on forecasting ED crowding. Results help achieve a consensus on what is feasible for emergency medicine, aiming to improve operational efficiency and patient care in ED settings. If adopted, PM forecasts can also help healthcare administrators at Nizwa Hospital plan for peak ED demands.

2 Materials and Methods

The study used a cross-section design to use hourly data on ED patient visits. The forecasting models were derived from hourly patient crowding in the ED. The data collection was performed in the period January to December 2023. The scientific and ethical committee in the Ministry of Health in Oman approved the study. The forecasting method of historical data of EDs was designed and shown in Figure 1.

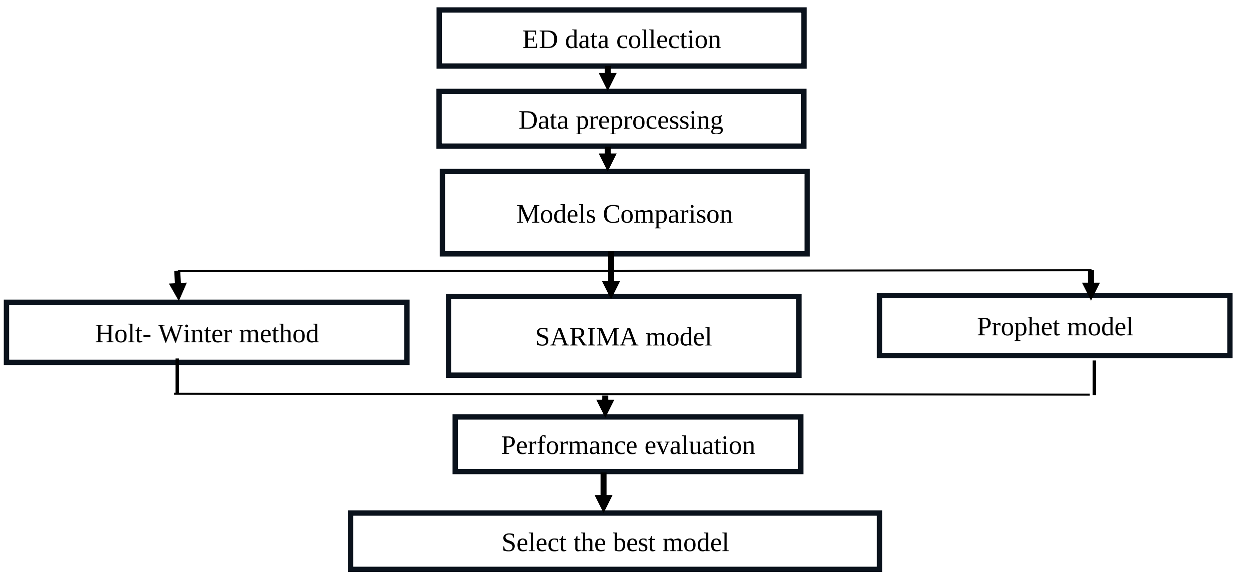

Figure 1. Research modeling framework for ED crowding

Figure 1 illustrates the overall research framework of this study. The first step involves gathering historical data on emergency department crowding from a healthcare institution. This dataset included ED crowding in Nizwa Hospital in Oman for each hour from 1st January 2023 to 31st December 2023.

The collected data is preprocessed before analysis. It entails checking and handling missing observation outliers and normalizing the data. Time series decomposition was done to identify the data's underlying trends and seasonal patterns. The dataset was split into training and testing sets, typically using a 99/1 ratio, to facilitate the evaluation of training and test data. The train and test set have 8,672 and 88 observations, respectively. Three forecasting models were developed and trained on the training dataset: HW, SARIMA, and PM. The final phase involves a comparative analysis of the forecasting accuracy of each model. The best model is one that minimizes the prediction errors of ED crowding.

Holt-Winter methods

Holt-Winters exponential smoothing is a time series forecasting technique suitable for data that exhibits trend and seasonality. The Holt-Winters' Additive Models have a set of four equations for the observation, Y_t, level, L_t, trend, T_t, and seasonal,S_t, components as in Equations 1 to 3.

$$L_t = \alpha(Y_t - Y_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1}) \quad [1]$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad [2]$$

$$S_t = \gamma(S_t - L_t) + (1-\gamma)\, S_{t-s} \quad [3]$$

The model predictions over h-step periods ahead, $\hat{Y}_{t+h}$, are made using Equations 4.

$$\hat{Y}_{t+h} = L_t + hT_t + S_{t+h-s} \quad [4]$$

where α, β, and γ are the smoothing parameters for the level, trend, and seasonal components. All three parameters range between 0 and 1.
Unlike moving average models, which assign equal weights to past observations within the moving average window, HW uses exponentially decreasing weights for distant past observations (Hyndman & Athanasopoulos, 2018). Thus, the HW model is more suitable for short-term forecasts since recent observations carry higher weights than distance observations.

SARIMA Model:

The SARIMA model is a traditional univariate time series model suitable for data that exhibit non-seasonal and seasonal patterns. The SARIMA model denoted as SARIMA $(p,d,q)(P,D,Q)_s$, is an extension of ARIMA that includes seasonal terms Autoregressive (AR) and Moving Average (MA) models. Where p,d,q are the non-seasonal AR, differencing, and MA orders; s is the seasonal period, P,D,Q are the seasonal AR, differencing, and MA orders, respectively. Algebraically, the SARIMA model is expressed in Equation 5.

$$Y_t = \mu + \emptyset_1 Y_{t-1} + \ldots + \emptyset_p Y_{t-p} + \Phi_1 Y_{t-s} + \ldots + \Phi_p Y_{t-Ps} + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q} + \Theta_1 \in_{t-1} + \ldots + \Theta_Q \in_{t-Qs} + \in_t \quad [5]$$

where μ is the mean of the series, $\emptyset_i$ and $\theta_i$re the parameters of the AR and MA parts, $\in_t$ is the white noise or the error term; Φ_i and $\Theta_i$are the parameters of the seasonal AR and MA parts.

Prophet model

The PM is an additive model for forecasting seasonal time series data recently developed by Facebook Inc. in 2017 (Taylor & Letham, 2018). It decomposes time series data into trends, seasonality, and holidays/events. The model is expressed in Equation 6.

$$Y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad [6]$$

where *Y(t)* is the forecasted value at time *t*; *g(t)* and *s(t)* are the trend and seasonal components, respectively; *h(t)* is holiday effects; and ϵ_t is the error term (presumed to be normally distributed).

Model Performance Evaluation Criteria

Univariate forecasting models were compared using two performance metrics: the mean absolute error (MAE), indicating the average of the absolute differences between the actual and the predicted values, and root mean squared error (RMSE), indicating the measure of the square root of the average of the squared differences between the actual and the predicted values. The selection criterion is that the best model provides lower MAE RMSE values, as it indicates that the provided predicted model has the potential to fit the time-series data.

The mathematical equations for the two measurements are:

$$\text{MAE} = \frac{\sum_{i=1}^{n} |\widehat{v_i} - v_i|}{n} \quad [7]$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\widehat{v_i} - v_i)^2} \quad [8]$$

where n is the sample size, $\widehat{v_i}$ is the predicted value to respect with i and v_i is the actual value to respect with i. The smallest measurement value indicates the best forecasting model.

3 Results and Discussions

A total of 65,967 patient visits in Nizwa Hospital in Oman from 1st January to 31 December 2023 were recorded and analyzed using univariate forecasting models. Figure 2 shows an hourly analysis of CROWD levels, revealing clear patient flow patterns throughout the day. The early morning hours (00:00 - 06:00) are generally quieter, with CROWD levels gradually decreasing until around 06:00. However, from 07:00 onwards, the crowding levels rise, peaking in the late afternoon and evening hours. The busiest period is between 19:00 and 23:00, with higher CROWD levels. This pattern indicates that the emergency department faces its greatest challenges in managing patient flow during the evening, requiring adequate staffing and resource allocation to handle the surge in patient numbers effectively.

Figure 2 illustrates Saturday and Sunday with the highest CROWD levels, especially on Saturday. The weekdays (from Monday to Friday) show moderate crowding levels. There is a gradual increase in ED visits as the week progresses, peaking on Saturdays. On average, weekends have a higher likelihood of higher ED crowding. Nonetheless, the daily averages might not be significantly different across the week., Thus, hourly effects are responsible for significant ED peaks daily of over 40 in a given hour. This implies that hospital administration should be vigilant daily to ensure adequate staff between mid-day and midnight. Generally, hourly seasonal effects are stronger than daily effects.

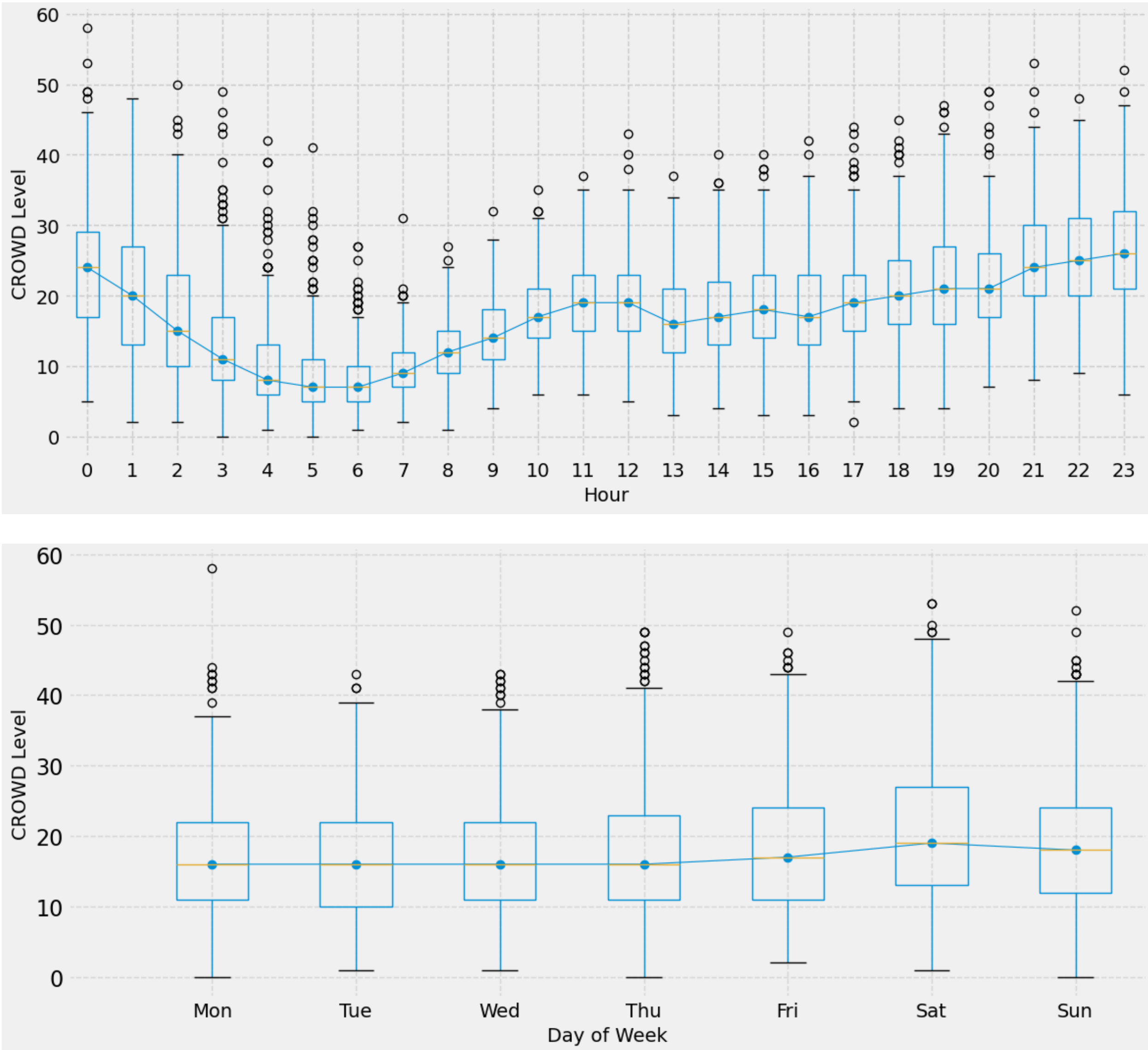

Figure 2. Trends of Crowding levels by hour and day of the week at Nizwa Hospital in Oman from January to December 2023

From Figure 3 late afternoon and evening hours experience high ED visits. The ED visits are often high in the summer months. ED visits are high in June, July, October, November, and December. Winter effects seem not strong as December only registers a high influx of ED visits, whereas January is the lowest. A high ED is crowding from October to December and drastically declines in January.

Early morning hours (3:00 AM to 9:00 AM) persistently register the least ED visits ill all days across the year. The early morning hours registered the lowest average CROWD levels between 6 to 16 patients than other hours. Fewer incidences occur at this time as most of the public is asleep, minimizing common ED visits due to accidents and injuries at work or on roads. The few cases could be common health conditions such as phenominia, sleep apnea, heart attacks, strokes, seizures, restless leg syndrome, and arrhythmia, which can be triggered or worsened by disrupted sleep patterns during the night (Rémi et al., 2019). They often lead to complications like reduced oxygen levels or irregular blood pressure, which require immediate medical emergency attention. They are usually less than emergencies from injuries and accidents that occur during the day.

The CROWD levels rise steadily from noon, peaking at midnight to 1. a.m. Afternoon to night hours shows elevated CROWD levels, with the highest levels observed on Friday and Saturday evenings. On average, ED visits peak between 9 PM and midnight on Mondays, Tuesdays, and Wednesdays. Thursday, Friday, and Saturday show an early onset of ED crowding above 20, occurring from 7 PM to 4 PM and 3 PM, respectively. The average crowding levels are 20.6 at 6 PM on Thursday, 20.3 on Friday at 3 PM, and 20.5 on Saturday at 2 PM.

By day of the week, Saturdays had the highest average CROWD levels, ranging from 10 to 30. Like on other days, ED visits on Saturdays are low in the early morning, rise in the evening, and peak at midnight. Fridays and Sundays are the second and third leading days with high ED visits, whereas Mondays to Wednesdays register the least ED visits. There is a general increasing trend of ED visits from Mondays peaking on Saturdays. Several factors can contribute to increased ED visits over the weekend. The peak ED trends can be associated with accidents and injuries since weekends have heavy traffic and sporting activities compared to working days. Besides, the peaks could be attributed to regular medical appointments, which increases overall congestion because most medical staff are likely to be preoccupied with other patients.

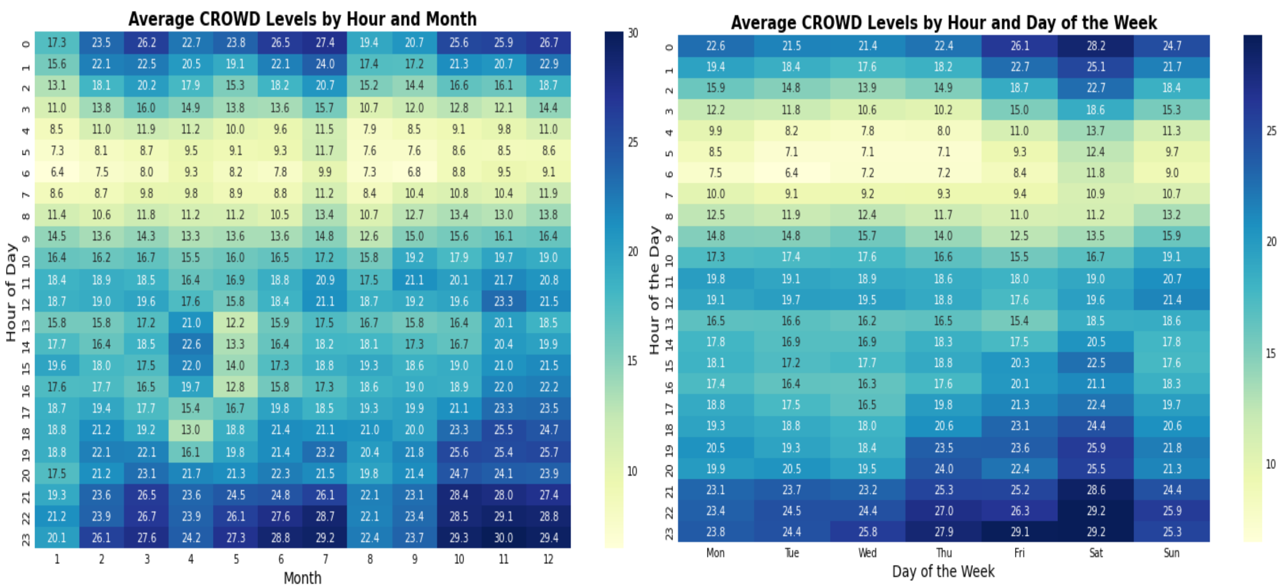

Figure 3. Time series plots of historical ED presentations by hour and Month (12 months), hour and day (7 days)

Descriptive statistics

Table 1 shows that the visits were predominantly on working days, accounting for 65.6% with a total of 43,302 visits, while non-working days comprised 34.4% with a total of 22,665 visits. The patient demographic shows a higher proportion of 56.8% of male patients than the proportion of 43.2% of female patients. Most patients were Omani nationals with a percentage of 94.1, with a small percentage being non-Omani at 5.9%. In terms of age distribution, the largest group of patients was aged between 15 and 59 years with a percentage of 57.64, followed by children aged 1-4 years (11.45%) and those aged 5-14 years (14.55%). The ED zones were primarily busy in the Green zone with a higher percentage of 49.14 and a Yellow zone percentage of 44 indicating a high volume of patients requiring urgent care. On average, the ED sees about 180 patients daily, with a mean patient age of 31.75 years and a standard deviation of 23.289, reflecting a diverse age range among the patients treated.

Table 1
Characteristics of study patients

Characteristic	Indicator	Number (n)	Percent (%)
ED Visits	Total patient visits	65967	100
Working days / non-working days visits	Working day visits	43302	65.6
	Nonworking days visits	22665	34.4
Sex	Male	37448	56.8
	Female	28519	43.2
Nationality	Omani	62046	94.1
	Non-Omani	3921	5.9
Age group	Under 7 days	120	0.18
	7-27 days	207	0.31
	28-1 years	2176	3.3
	1- 4 years	7550	11.45
	5-14 years	9596	14.55
	15-49 years	38025	57.64
	>60 years	8293	12.57
ED Zone	Black	89	0.13
	Blue	96	0.15
	Green	32419	49.14
	Red	3311	5.02
	White	1027	1.56
	Yellow	29025	44.00

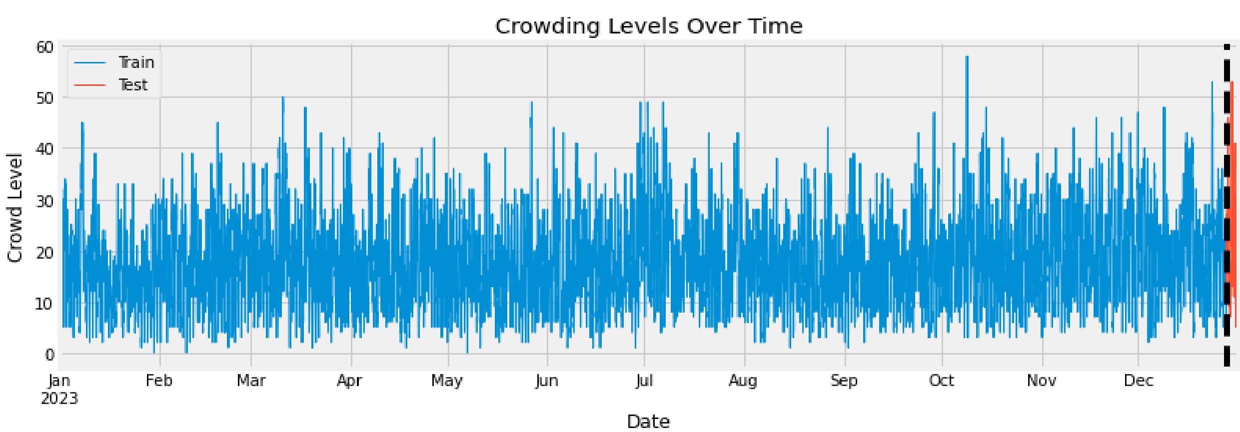

Figure 4. Emergency Department crowding levels (Train and Test data) over time

The total time series data in Figure 4 was 8760 and was split into train and test sets in observation of 8,672 / 88 for training and testing sets. The result of the Mann- Kendall trend test reveals that the crowding levels in the ED are increasing over time, with a statistically significant, albeit weak, upward trend ($p - value < 0.05$). This trend highlights the growing demand for emergency services, which may require the hospital to adjust its resource allocation and operational strategies to manage the increasing number of patients effectively.

Table 2
Mann-Kendall trend test output

Mann-Kendall Trend Test	Value
Trend	INCREASING
P − value	0
Tau	0.065568
Z	9.162896

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test in Table 3 indicated that the crowding levels are non-stationary with respect to unit roots ($p - value < 0.05$). After differencing the time series, The KPSS statistic for the first-differenced series for 24 periods as the data is hourly base is 0.003555, which is well below the critical values at the 1%, 5%, and 10% levels. The KPSS test does not detect any trend or varying variance in the differenced series, further confirming its stationarity.

Table 3
KPSS test output

KPSS TEST	BEFORE DIFFERENCING	AFTER DIFFERENCING
STATISTIC	1.641237	0.003555
$P - value$	0.01	0.1
STATIONARITY	Non-stationary	Stationary

Table 4
Kruskal – Wallis H test output

KRUSKAL-WALLIS H TEST	HOURLY SEASONALITY	WEEKLY SEASONALITY
STATISTIC	3404.54	149.86
$P - value$	0	8.27E-30
SEASONALITY	Seasonality	Seasonality

The Kruskal-Wallis H-test in Table 4 showed a test statistic of 3404.54 with $p - value < 0.05$, which indicates that the variation in crowding levels by hour is statistically significant. The differences observed in the CROWD levels during various hours of the day are not due to random fluctuations. However, they are substantial enough to reject the null hypothesis (H_0) that assumes no difference between hours. The Kruskal-Wallis H-test was also applied to evaluate the differences in CROWD levels across the days of the week. The test yielded a statistic of 149.86 and $p - value < 0.05$. There are statistically significant differences in crowding levels between the different days of the week, leading us to reject the null hypothesis (H_0), which assumes no variation in crowding between the days.

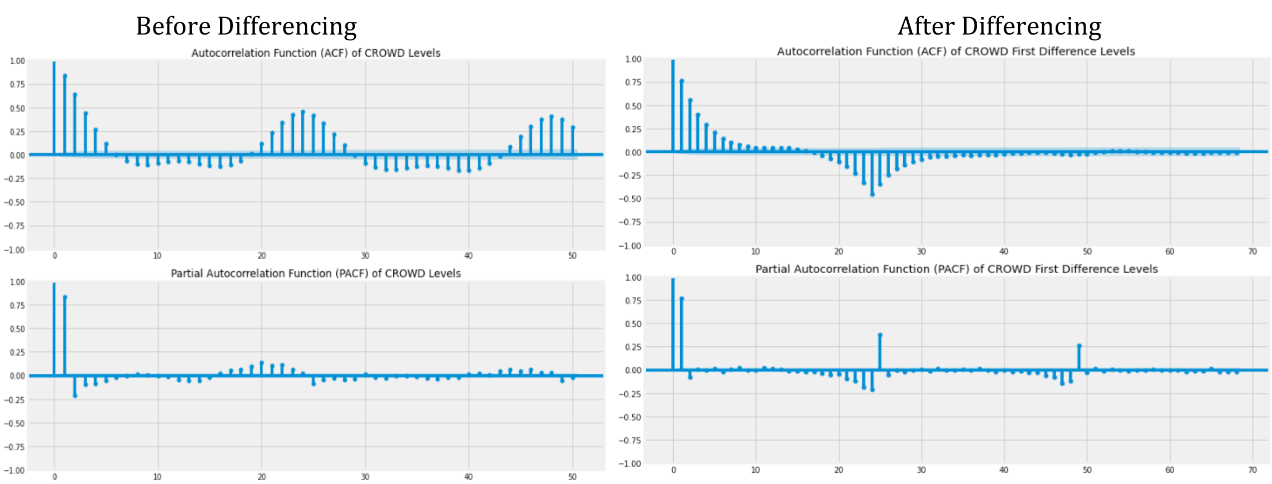

Figure 5. CROWD Time Series Autocorrelation and Partial Autocorrelation

Figure 5 shows that the ACF plot of the level data tails off, as well as seasonal waves. The significant decaying spikes with repeated cycles indicate the presence of autocorrelation and seasonality in the ED crowding series. The significant autocorrelations indicated that past crowd levels can be used to forecast future levels.

The crowd series is non-stationary, given its higher-order, non-seasonal autoregressive nature in the ACF plot. Thus, the differencing was required to provide a stationary process. Therefore, the order of integration is set at 1. In both differenced series, significant positive spikes up to lag 2, suggesting a positive autocorrelation, which can be modelled using an AR (2) process. For the MA (5) model, the ACF plot "cuts off" at a lag (5), but the PACF does not "tail off" gradually, implying that q (non-seasonal part) can be 0. After 1st differencing, the autocorrelation coefficients of the ACF rapidly decline, indicating that the differenced series has become stationary. Significant spikes may still appear at seasonal lags, suggesting the model needs seasonal components. The regular spikes suggested the presence of a recurring seasonal pattern. From the ACF and PACF of the 1st differenced series, there is a significant spike at lag 24 and 48.

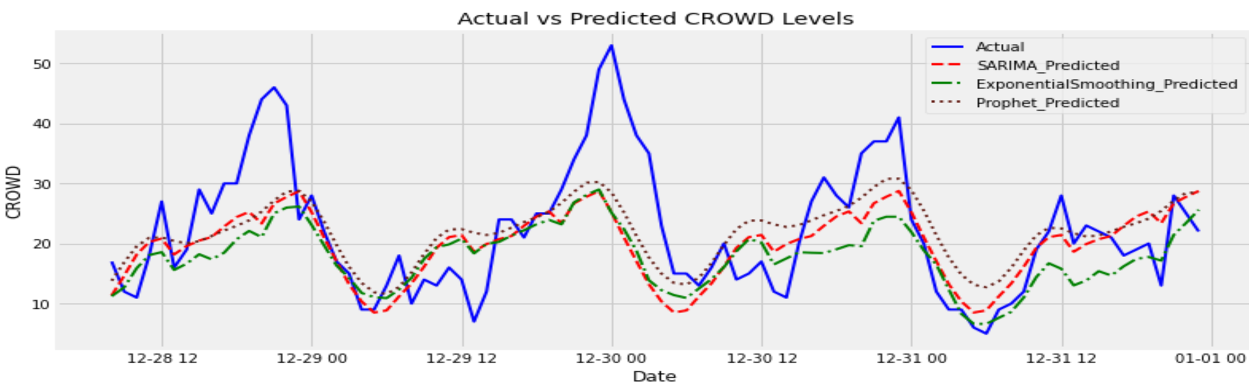

Figure 6. Actual and predicted ED crowding in three forecasting models

Figure 6 shows the time series fit of the forecasting models, demonstrating its ability to track observed crowd levels over time using only historical patient counts. The blue line represents the observed data points, showing real-time crowd-level fluctuations. The shaded region's red, green, and black dots reflect the forecasted values generated by SARIMA, the HW method, and the Prophet model. Table 5 shows that the optimal HW parameters indicate the trend and seasonality in the CROWD levels.

Table 5
The optimized Holt Winter's parameters

Initial level	Smoothing level (alpha)	Initial trend	Smoothing trend (beta)	Smoothing seasonal (gamma)
17.235	0.9599	0.394	0.0027	0.0395

Notes. Initial seasons: - 0= 6.4429, 1 = 5.0654, 2 = 2.9547, 3 = 0.9658, 4 = -3.2799, 5 = -4.2415, 6 = -5.2393; The Akaike Information Criterion (AIC) = 26,456.215; Bayesian Information Criterion (BIC) = 27,671.886.

The selection of the additive HW is coherent with the evolution of the CROWD levels in (Figure 6). The series exhibits a linear trend growth and no exponential growth. The additive model shows a constant increase in crowding from morning hours to midnight and a decrease thereafter to about 9 AM, where it troughs. The seasonal and regular comments also have the same magnitude over time, with repeating patterns packing at about midnight and low in the morning. Generally, the seasonal variation of the CROWD levels at the ED remains roughly constant regardless of the trend level. The high smoothing level of 0.9599 indicates that the model gives more weight to recent observations. This shows a persistent short-run (hourly) trend effect with an initial slope of 0.094, which changes slowly over time, as indicated by the small value of the smoothing trend parameter of 0.0027.

The smoothing seasonal parameter of 0.0395 reflects the regular adjustments (initialized at 6 points) to account for the seasonality of the ED crowding component within a seasonal cycle of 96 periods (over a 4-day cycle) (Figure 6). The HW underestimates the values compared to the SARIMA and Prophet models. SARIMA seems to be an average of the HW and the Prophet. One of the assumptions of the HW is the additive components, which assume constant fluctuation over time, failing to account for weekend effects where ED was established to be high over the weekends (Figure 3). The SARIMA (1, 0, 1) x (0, 1, [1], 24) was the best among competing SARIMA models. Table 6 shows the model parameters.

Table 6
The optimized SARIMA (1,0,1) x (0,1,1,24) parameters

	Coefficient	*SE*	*Z*	$P - value$
AR (1)	0.7397	0.008	95.889	0.000
MA (1)	0.01194	0.013	9.469	0.000
Seasonal MA (L24)	-1.0328	0.003	-360.889	0.000
Sigma^2	16.4827	0.239	69.096	0.000

Notes. AIC = -24618.6, BIC = 49245.2

HW failed to incorporate the error component accounted for by the SARIMA model. The AR (1) is statistically significant, b = 0.7397,$p - value < 0.001$, indicating that the CROWD level from the previous hour strongly influences the current level.

The MA (1) term is also significant, b = 0.1194, $p - value < 0.001$, suggesting that the error from the previous period positively contributes to the current CROWD level. Thus, unexpected changes or shocks in crowding levels tend to propagate to the crowding levels the following day. The significant seasonal MA (1), b =-1.0328, $p - value < 0.001$, indicates strong adverse seasonality effects of crowding level occurring on a 24-hour cycle. Crowding at a given time is negatively correlated with the error from the same time on the previous day. The seasonality reflects the regular daily cycles in hospital activity. As shown in Figure 3, ED crowding peaks at midnight and begins a new phase in the morning around 9 AM. The Prophet is more conservative and gives minimal errors in the out-sample predictions.

Forecasting performance metrics

Table 7 indicates the performance metrics of the three predicted forecasting models. Two error metrics, MAE and RMSE, were used. Model selection is based on the minimization of these error metrics. The Prophet model provides the most reliable forecasts for ED crowd levels at Nizwa Hospital. The HW had the least predictive accuracy among the competing models (MAE = 6.64; RMSE = 9.01). The SARIMA model had better precision (MAE = 6.13; RMSE = 8.42) than HW. The Prophet model was the best (MAE = 5.494; RMSE = 7.274).

Table 7
Model performance metrics

Forecasting Model	MAE	RMSE
HW method	6.644	9.014
SARIMA model	6.127	8.419
Prophet model	5.494	7.274

Discussion

The paper compared three univariate time series models, SARIMA, PM, and HW models, to forecast ED crowding in a public hospital in Oman. The study used hourly patient visits at ED from January to December 2023. Based on the minimization of RMSE and MAE, PM was established as the best-fit model for the ED visits forecast at the Nizwa Hospital. The HW model had the least predictive accuracy. The HW is often suitable for short-run forecasting since recent observations carry higher weights than distance observations. Thus, the observed cycles occurring at lag 24 and 48 might not be captured by HW. Therefore, persistent seasonal trends in ED across the year make HW inadequate in forecasting daily ED crowding over a one-month horizon. Given that SARIMA can account for such cyclical aspects using the seasonal AR and MA orders, the SARIMA model had better precision than HW.

The Prophet model has better performance in forecasting ED crowding than SARIMAX and HW methods. Its performance can be attributed to its capability to model non-linear trends, multiple seasonality, and holiday effects (Taylor & Letham, 2018). Thus, FBPM is more robust in modeling hour and day-of-the-week effects present in ED crowd levels than the traditional SARIMA and HW models. As observed in Figure 2. and Figure 3. hourly and day effects were significant. Hourly variations were non-linear, with crowding levels low in the early morning (3:00 to 9:00 AM), rising in the afternoon, and peaking at midnight. Early morning hours persistently register the least ED visits on all days across the year. By day, ED was high on Saturdays, followed by Fridays and Sundays. ED visits are high in June, July, October, November, and December. In contrast, January records the least average ED visits. The model PM accounts for such strong day-of-week effects and non-linear trends, making it more accurate than HW and SARIMA models. The finding aligns with existing literature that suggests that the Prophet model outperforms HW and SARIMA models in ED crowding forecasting.

While there is a general tendency to adoption of the Prophet model in ED forecasting, univariate models might not be reliable. The current study fitted univariate models, which do not account for external variables that may influence crowd levels. There is a possibility for other factors, such as total patient arrivals, weather conditions, and the use of dummy calendar regressors such as month or day, public health crises, seasonal illnesses, and local events. The current study is univariate; hence, future studies can explore the possibility of incorporating external regressors to the SARIMA and Prophet models to improve accuracy. Thus, the Prophet model is recommended for ED crowding prediction at Nizwa Hospital with caution. The hospital can consider improving with internalized metrics such as patient visits and the number of staff to provide better forecasting and help in planning and resource allocation to handle future ED crowding.

4 Conclusion

The univariate Prophet model provides more reliable ED crowing forecasts than the SARIMA and HW models. The Prophet model can model non-linear trends and hour and daily effects present in ED crowd levels compared to the traditional SARIMA and HW models. Early morning hours (3:00 AM to 9:00 AM) persistently register the least ED visits on all days across the year and peak between 10 PM and midnight. By day, ED was high on Saturdays, followed by Fridays and Sundays. By month, ED visits are high in June, July, October, November, and December. January records the least average ED visits.

The current study suffers from omitted variable bias. In particular, Nizwa Hospital can adapt and improve the robustness of Prophet with internalized metrics such as patient visits and number of staff as external regressors to provide better forecasting. Despite the Prophet's best performance, there is a substantial deviation from the Prophet model predictions from the actual values. None of the models have a nearly perfect symmetry of the exact pattern of crowding (Figure 6). The significant spikes in the actual ED crowding levels relative to the model predictions may be due to excluded covariates such as weekends and other exogenous factors such as holidays, which can be accounted for by multivariate models. Better forecasts can help plan and allocate resources to handle future ED crowding. Future studies can explore incorporating external regressors to the SARIMA and Prophet models to improve accuracy. Additionally, further investigation into the long-term trends and their implications for ED crowding on patient outcomes such as mortality and admission is crucial.

Acknowledgments
We want to thank and appreciate the computer technician at Nizwa Hospital in Oman for his help in collecting data.

References

Becerra, M., Jerez, A., Aballay, B., Garcés, H. O., & Fuentes, A. (2020). Forecasting emergency admissions due to respiratory diseases in high variability scenarios using time series: A case study in Chile. *Science of the Total Environment*, *706*, 134978. https://doi.org/10.1016/j.scitotenv.2019.134978

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time series analysis: Forecasting and control (5th ed.). Wiley.

Calegari, R., Fogliatto, F. S., Lucini, F. R., Neyeloff, J., Kuchenbecker, R. S., & Schaan, B. D. (2016). Forecasting daily volume and acuity of patients in the emergency department. *Computational and Mathematical Methods in Medicine, 2016*, 3863268.

Chen, W., Linthicum, B., Argon, N. T., Bohrmann, T., Lopiano, K., Mehrotra, A., Travers, D., & Ziya, S. (2020). The effects of emergency department crowding on triage and hospital admission decisions. *The American Journal of Emergency Medicine, 38*(4), 774–779. https://doi.org/10.1016/j.ajem.2019.06.039

Chiu, I.-M., Lin, Y.-R., Syue, Y.-J., Kung, C.-T., Wu, K.-H., & Li, C.-J. (2018). The influence of crowding on clinical practice in the emergency department. *The American Journal of Emergency Medicine, 36*(1), 56–60. https://doi.org/10.1016/j.ajem.2017.07.011

Duarte, D., & Faerman, J. (2019). Comparison of time series prediction of healthcare emergency department indicators with ARIMA and Prophet. *Proceedings of the Computer Science and Information Technology Conference*, 123-133.

Etu, E. E., Monplaisir, L., Masoud, S., Arslanturk, S., Emakhu, J., Tenebe, I., ... & Krupp, S. (2022, June). A comparison of univariate and multivariate forecasting models predicting emergency department patient arrivals during the COVID-19 pandemic. In *Healthcare* (Vol. 10, No. 6, p. 1120). MDPI.

Feng, T., Zheng, Z., Xu, J., Liu, M., Li, M., Jia, H., & Yu, X. (2022). The comparative analysis of SARIMA, Facebook Prophet, and LSTM for road traffic injury prediction in Northeast China. *Frontiers in public health*, *10*, 946563.

Hoot, N. R., & Aronsky, D. (2008). Systematic review of emergency department crowding: causes, effects, and solutions. *Annals of emergency medicine*, *52*(2), 126-136. https://doi.org/10.1016/j.annemergmed.2008.03.014

Hoot, N. R., LeBlanc, L. J., Jones, I., Levin, S. R., Zhou, C., Gadd, C. S., & Aronsky, D. (2009). Forecasting emergency department crowding: A prospective, real-time evaluation. *Journal of the American Medical Informatics Association, 16*(3), 338–345.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Jeyaraman, M. M., Copstein, L., Al-Yousif, N., Alder, R. N., Kirkland, S. W., Al-Yousif, Y., Suss, R., Zarychanski, R., Doupe, M. B., Berthelot, S., Mireault, J., Tardif, P., Askin, N., Buchel, T., Rabbani, R., Beaudry, T., Hartwell, M., Shimmin, C., Edwards, J., … Abou-Setta, A. M. (2021). Interventions and strategies involving primary healthcare professionals to manage emergency department overcrowding: A scoping review. *BMJ Open, 11*(5), e048613.

Pines, J. M., Hollander, J. E., Localio, A. R., & Metlay, J. P. (2006). The association between emergency department crowding and hospital performance on antibiotic timing for pneumonia and percutaneous intervention for myocardial infarction. *Academic Emergency Medicine*, *13*(8), 873-878.

Pines, J. M., Iyer, S., Disbot, M., Hollander, J. E., Shofer, F. S., & Datner, E. M. (2008). The effect of emergency department crowding on patient satisfaction for admitted patients. *Academic Emergency Medicine*, *15*(9), 825-831.

Rémi, J., Pollmächer, T., Spiegelhalder, K., Trenkwalder, C., & Young, P. (2019). Sleep-related disorders in neurology and psychiatry. *Deutsches Ärzteblatt International*, *116*(41), 681.

Rosychuk, R. J., Youngson, E., & Rowe, B. H. (2016). Presentations to emergency departments for COPD: A time series analysis. *Canadian respiratory journal*, *2016*(1), 1382434.

Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician, 72*(1), 37-45.

Tuominen, J., Roine, A., Saviauk, T., Seppo, A., Pihlaja, M., Ovaska, J., ... & Oksala, N. (2021). Forecasting Daily Arrivals and Peak Occupancy in a Combined Emergency Department.

Vieira, A., Sousa, I., & Dória-Nóbrega, S. (2023). Forecasting daily admissions to an emergency department considering single and multiple seasonal patterns. *Healthcare Analytics*, *3*, 100146. https://doi.org/10.1016/j.health.2023.100146

Villani, M., Earnest, A., Nanayakkara, N., Smith, K., De Courten, B., & Zoungas, S. (2017). Time series modelling to forecast prehospital EMS demand for diabetic emergencies. *BMC health services research*, *17*, 1-9.

Biography of Authors

	**Mansoor AL Yarubi** has been working in the Ministry of Health in Oman as Head of the Statistics section since 2004. Before that, I was working as a statistician at Nizwa Health Center in the Ministry of Health in Oman. In addition to my position, I am a lecturer in statistics part-time at the University of Nizwa in Oman. My research interest is the analysis of time series models and machine learning methods in health care, especially emergency department crowding. *Email: dakhresearch@gmail.com*
	**Dr. Nur Arina Bazilah Kamisan** is a Senior Lecturer at Universiti Teknologi Malaysia (UTM) teaching Mathematics and Statistics subjects such as Engineering Mathematics, Engineering Statistics, and Time Series. Her research specializes in time series, missing values, and directional statistics and modeling in various data applications such as climate, economics, and education. *Email: nurarinabazilah@utm.my*
	**Dr. Siti Mariam Norrulashikin** is a Senior Lecturer at Universiti Teknologi Malaysia (UTM), specializing in time series analysis and modeling. She has co-authored numerous papers from many areas related to time series. Dr. Siti Mariam holds a PhD in Mathematics from UTM and is currently exploring predictive models in the machine-learning field. *Email: sitimariam@utm.my*