*Proceedings of the 2023 Winter Simulation Conference*
*C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.*

# FORECASTING PATIENT ARRIVALS AND OPTIMIZING PHYSICIAN SHIFT SCHEDULING IN EMERGENCY DEPARTMENTS

Vishnunarayan Girishan Prabhu

Industrial and Systems Engineering
University of North Carolina at Charlotte
314 Smith Hall
Charlotte, NC 28262, USA

Kevin Taaffe

Department of Industrial Engineering
Clemson University
100 Freeman Hall
Clemson, SC 29631, USA

Ronald Pirrallo
William Jackson
Michael Ramsay
Jessica Hobbs

Department of Emergency Medicine
Prisma Health - Upstate
701 Grove Road
Greenville, SC 29605, USA

## ABSTRACT

Emergency Departments (EDs) are the primary access points for millions of patients seeking medical care. The increasing patient demand and lack of long-term dynamic planning strain the EDs in providing timely patient care, leading to crowding. While a well-recognized problem, ED crowding is still prevalent, where suboptimal resource allocation is one significant contributing factor. In this research, we developed an end-to-end solution that first forecasted the patient arrivals to the partner ED and then used an optimization model to develop an optimal physician staffing schedule to minimize the combined cost of patient wait times, handoffs, and physician shifts. Finally, the new schedule was tested using the validated simulation model to evaluate the ED performance. By generating shift schedules based on forecasts and testing them in the validated simulation model, we observed that patient time in the ED and handoffs could be reduced by 5.6% and 9.2% compared to current practices.

## 1 INTRODUCTION

The Emergency Department (ED) is a critical division in a health system where patients receive care for various conditions, including life-threatening emergencies, chronic ailments, and non-emergent situations. This diverse nature of ED to care for patients and the Emergency Medical Treatment and Labor Act (EMTALA), which mandates ED to provide screening and stabilizing care to all patients regardless of their ability to pay, makes ED a primary access point for patients seeking care (Laxmisan et al. 2007; McDonnell et al. 2013). Over the last several years, patient arrivals to EDs in the US have increased from 96.5 million annual visits in 1995 to 115.3 million in 2005 and 151 million in 2019 (Cairns et al. 2019; Centers for Disease Control and Prevention 2010). At the same time, the number of EDs in the US has decreased by over 15% in the last decade (Hsia et al. 2011). The ever-increasing patient volume and the decreasing number of EDs lead to mismatch, predisposing EDs to crowding (Di Somma et al. 2015; George and Evridiki 2015; Kelen et al. 2021). The American College of Emergency Physicians (ACEP) defines ED crowding as the situation that "occurs when the identified need for emergency services exceeds available resources for patient care in the ED, hospital, or both" (American College of Emergency Physicians 2019). Crowding in ED is a global concern, and studies have often linked this as a factor leading to suboptimal care, delays in care, and higher chances of medical errors (Di Somma et al. 2015; Kulstad et al. 2010).

A few leading causes of ED crowding include high patient census (patient arrivals), inadequate resources (beds, medical devices, etc.), inadequate planning, and poor ED design (Morley et al. 2018; Moskop et al. 2009). Some of the most commonly adopted solutions to avoid ED crowding include expanding ED capacity, stopping boarding admitted patients in ED, adding hallway beds, increasing on-call providers, and adding temporary resources (Derlet and Richards 2008). While these solutions are effective temporary fixes, they can often be costly and negatively impact patient safety and physician wellbeing. A recent study investigating ED crowding identified that access to future patient demands (arrivals to ED) during shift scheduling and resource allocation can improve ED planning and potentially avoid crowding (Kelen et al. 2021). Most EDs, including our partner ED, build their clinician schedules about one quarter (3 months, 90 days) ahead. Hence, it is critical to have robust 90-day forecasts to assist ED administrators in planning clinician schedules to improve ED performance.

Researchers have used various forecasting methods to predict patient arrivals to the ED for different horizons (Aboagye-Sarfo et al. 2015; Batal et al. 2001; Carvalho-Silva et al. 2018; Choudhury and Urena 2020; Côté et al. 2013; Hertzum 2017; Jones et al. 2008; Kadri et al. 2014; Khaldi et al. 2019; Sun et al. 2009; Whitt and Zhang 2019; Xu et al. 2013; Zhang et al. 2022). Additionally, a few studies have focused on forecasting specific types of patient arrivals to the ED (primarily patients with respiratory diseases) (Becerra et al. 2020; Rosychuk et al. 2015). Regarding the methodology for forecasting patient arrivals to the ED, Autoregressive Moving Average (ARMA), Vector Autoregressive Moving Average (VARMA), Holt-Winters, linear regression, multiple linear regression (MLR), Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and machine learning models (ANNs and RNNs) have been used extensively. In terms of forecasting horizon, researchers have forecasted hourly, daily, and monthly patient arrivals. Finally, regarding the data used, some studies have reported better performance when using multiple input variables. However, these often require expert manipulation, making them difficult for ED administrators and stakeholders to use. A significant gap among prior studies that forecasted patient arrivals to the ED is that none included the Emergency Severity Index (ESI) levels of the patients while generating the forecasts. ESI is a standardized index assigned to each patient presenting to the ED across North America that varies from 1-5, representing the patient's severity and the expected resources the patient would require (Hossein et al. 2013). Hence, from a resource allocation standpoint, including ESI levels in the patient forecasts is critical for better planning the ED resource allocation.

A variety of operations research approaches, including mathematical optimization models, queuing theory models, simulation modeling, and probabilistic models, have been used to address various challenges observed in the ED, including resource allocation (Ahsan et al. 2019; Connelly and Bair 2004; Elalouf and Wachtel 2021; Rais and Vianaa 2011). While simulation is popular for addressing operational issues, its outcome is a realization and not an optimal solution. However, by identifying a specific objective, researchers have developed mathematical models to identify optimal staffing levels, generate schedules, determine optimal bed or other resource requirements, etc. One study used a mixed-integer linear programming (MILP) model to minimize understaffing with respect to patient volumes which resulted in significant improvements to median length of stay, door-to-provider time, and door-to-bed time (Sir et al. 2017). Researchers have also used a combination of simulation-optimization models to identify optimal solutions and test them in the simulation model to validate them (Ghanes et al. 2015). Most of the ED studies have focused on identifying solutions that reduce patient waiting time, reduce length of stay or improve ED throughput. While a few studies have used patient wait times as surrogates for patient safety, we introduce a new metric – handoffs, directly quantifying patient safety (Maughan et al. 2011).

Patient safety is a crucial part of the ED as continuous patient flow and interactions with multiple departments and providers make it prone to errors. Additionally, researchers have observed ED as one of the hospital departments with high error rates. Among different issues that lead to medical errors, studies have identified handoffs, transfer of a patient's care, and responsibility from one physician to another as major patient safety issues (Maughan et al. 2011; Venkatesh et al. 2015). Specifically, studies investigating ED handoffs observed that the vital signs were not communicated for approximately 75% of the patients, and errors were observed in about 60% of cases (Venkatesh et al. 2015). Hence, while developing the ED physician shift schedules, it is crucial to consider patient safety metrics such as handoffs.

Our literature review shows that researchers have used time series forecasting, mathematical models, and simulations to improve ED operations. However, to our knowledge, these studies have not combined these approaches and included patient severity in their forecast. Moreover, as noted earlier, few have considered a direct patient safety metric in the mathematical model. To address these research gaps, we first develop forecasting models for predicting the 90-day patient arrivals to the ED, including their ESI level, which is then inputted into the mathematical model for developing the schedule. The objective of the mathematical model was to identify optimal shift schedules that minimize the combined cost of patient wait times, handoffs, and physician shifts, thus considering the patient flow, patient safety, and staffing budget to generate schedules. Further, to test the impact of the generated schedules on the ED performance, we used our validated simulation model (Girishan Prabhu et al. 2022).

## 2 DATA

Input data for developing the forecasting model included two specific data points: the time of the day and the ESI levels assigned to the patient presenting to the ED. However, for developing the optimization and models, other data points, including the number of beds, allowable physician shifts, patient arrivals, ESI level of the patients, patient time in the ED, and the number of interactions between physicians and patients, were gathered from the PRISMA Health Greenville Memorial Hospital (GMH), Greenville, SC. Additionally, the research team included ED physicians working in GMH, SC for guidance and addressing any other physician-dependent activities in the ED to be included in the model. PRISMA Health is the largest healthcare provider in South Carolina and serves as a tertiary referral center. The flagship GMH academic ED is an Adult Level 1 Trauma Center seeing over 106,000 patients annually. Figure 1, below, represents the patient arrivals to the GMH ED averaged for 2017-2019 used for our forecasting model.

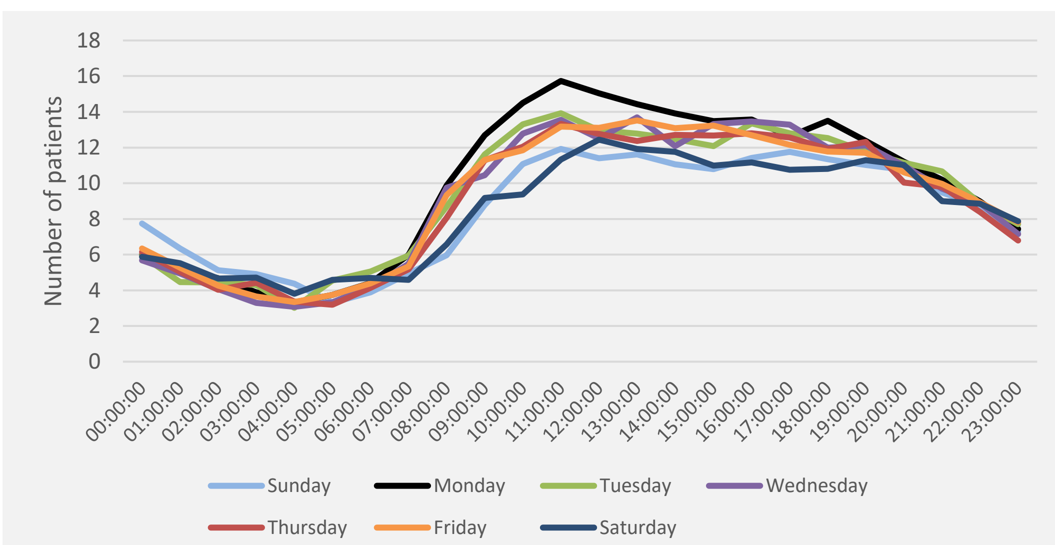


Figure 1: Patient arrivals to the GMH ED.

One of the first patterns that can be noticed is the impact of hours of the day where the patient arrivals are low during the early hours and slowly start picking up from 7:00 am until 12:00 pm when they reach the maximum and stay the same until 7:00 pm. This patient arrival trend aligns with some prior studies (Alvarez et al. 2009; Whitt and Zhang 2019). Another pattern that can be noticed is the difference between weekdays and weekends, where weekdays have higher patient arrivals compared to weekends, and Mondays have the highest patient arrivals. These observations clearly suggest the need for long-term planning as each day and hour of the day require different staffing levels. In terms of ESI levels, we observed that about 50% of patient arrivals to the ED are ESI-3 patients, followed by ESI 2 and ESI 4, which contributed 25% and 20% of the patient arrivals. Finally, ESI 1 and 5 each contributed only 2-3% of the total arrivals. ESI 1 refers to severely unstable patients who need immediate intervention, and ESI 5 patients are the most stable patients and may be treated non-urgently and mostly require the least resources. ESI levels are critical during forecasting as each ESI demands different resources.

For training the forecasting model, we used 2017, 2018, and the first six months of 2019 data, and the predictions were made on the last six months divided into clusters of 3 months. Rather than using an entire year of patient arrival and using it for physician scheduling, we created clusters of 3 months and used the cluster with the highest patient arrivals for this research (July 2019 – September 2019). Using the whole year's data was impossible because of the variability. However, relying on daily or weekly data will fail to account for operational biases such as leave of absence, vacations, etc., which could impact patient times in the ED. Clustering the data by quarters allowed us to address these issues. Additionally, based on expert opinions from the ED physicians, we wanted to use the pre-COVID-19 data as the patient arrivals varied significantly during 2020. Another reason for using this specific period was to test the optimal schedule in our validated simulation model that used the same patient arrivals. However, both models were developed such that any patient arrivals can be used to generate a weekly schedule. Next, we introduce Table 1, which represents the time a patient spends in the ED based on their ESI levels.

Table 1: Patient time in the ED.

| **Severity** | **Bed to Disposition (mins)** | **Disposition to ED Departure (mins)** | **Total Time (mins)** |
|---|---|---|---|
| ESI 1 | 115 | 121 | 236 |
| ESI 2 | 186 | 86 | 272 |
| ESI 3 | 175 | 54 | 229 |
| ESI 4 | 90 | 24 | 114 |
| ESI 5 | 107 | 15 | 122 |

As seen above, we split the data into two parts: "Bed to Disposition" and "Disposition to ED Departure." Bed to disposition represents the time a patient occupies an ED bed and is provided care by physicians and other medical providers, including performing tests, providing medicines, blood draws, etc. Although patients will be waiting in their beds during this period without receiving direct care, all these delays are due to waiting for their test results, medicines, etc. In general, this represents the period a patient first occupies a bed in the ED until the physicians make a disposition decision (admit, discharge, or transfer). The second part, "Disposition to Departure," is the period for which a patient occupies the ED bed from the time the physician makes a disposition decision until they are physically moved from the ED (discharged, admitted, or transferred). Hence, these are logistical delays where a patient can be either waiting until a bed is available in the hospital (admission) or waiting for transportation (discharged or transfer). While we primarily focus on the bed-to-departure time for this study, our model still accounts for delays before assigning a bed in the ED, where the patients wait in a waiting room until the beds are available, similar to an actual setting. As mentioned earlier, the entire bed-to-disposition time of a patient is not spent with a physician as it includes other activities. Based on literature and discussions with ED physicians, we used between 15-30% of total time as the care time where a patient would be cared for by a physician (Füchtbauer et al. 2013). The percentages were assigned based on severity such that the total time spent with an ESI-1 patient was the highest and that with an ESI-5 patient was the lowest. This approach was used mainly used because of the lack of detailed visit-by-visit data available to support detailed modeling.

Further, to build a model representative of ED operations where a physician visits patients multiple times based on their severity (ESI- level), we split the care time into multiple smaller windows. Based on our past observational studies and discussion with ED faculties and physicians, on average, an ES1-1 patient was visited four times by a physician, ESI-2 and 3 were visited three times, and ES1-4 and 5 were visited two times (Girishan Prabhu et al. 2020). The physician's time with a patient for each visit was a constant time block of 15 minutes, as the MILP modeling approach considers time as a discrete block of events.

## 3 MODEL DEVELOPMENT

### 3.1 Forecasting Models

In this study, we used the moving average naive model as our benchmark model to compare the forecasts from other models. Based on literature and data visualizations, we decided to develop both ARIMA and SARIMA models as these models are effective in forecasting time series data, especially SARIMA, when the data is considered to have seasonality (Carvalho-Silva et al. 2018; Choudhury and Urena 2020; Hertzum 2017; Kadri et al. 2014; Sun et al. 2009; Whitt and Zhang 2019). Additionally, we developed a Holt-Winters forecasting model as this approach can account for the level, trend, and seasonality component in the time-series data. Finally, we also developed two machine learning models: Extreme Gradient Boosting (XGBoost) and Random Forest Regression model. Both are decision tree machine learning algorithms and require a supervised learning approach where each input requires an output pair within the training model for the model to learn and predict. However, the foundation of each algorithm is different, where Random Forest Regression uses a bagging technique, whereas the XGBoost uses a boosting technique for learning.

To avoid the potential issue of overfitting with the machine learning models and tuning the hyperparameters, we used a blocked-crossed validation approach. This approach was preferred as it can avoid memorizing the patterns and leakage from future data, which is unavoidable when using k-fold cross-validation with time-series data. Specifically, the blocked-cross-validation approach accounts for this issue by adding margins at two positions: (i) between the training and validation folds and (ii) between the folds used at each iteration. Finally, to evaluate the performance of each model, we used Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percent Error (MAPE).

### 3.2 Optimization Model

We formulate the ED physician shift scheduling problem as a MILP problem. The primary goal is to identify the optimal staffing levels of ED physicians such that the patient onboarding time, waiting time after ED admission, and patient handoffs are minimized while considering the physician staffing cost. To compare the impact of each factor, we use dollar amount as the common scale.

Before formulating the problem, we first list the key ED operational activities to replicate the partner ED. The first was accounting for the varying patient arrivals to the ED, including patient ESI levels. The second was modeling multiple patient-physician interactions based on the patient's ESI, accounting for minimum delays between patient-physician interactions for secondary care (imaging, blood draws, etc.), ensuring that the same physician provides care for the patient unless the physician ends their shift (handoff), physician shift length is limited to 8 hours. Next, we define the notation used in the MILP model. The model included four sets and corresponding indices as follows: (i) $I$ represents the set of patient arrivals to the GMH ED indexed by $I$, (ii) $K$ represents the set of possible physicians that can be staffed for a day indexed by $k$, (iii)$T$ represents the set of time slots considered for staff scheduling indexed by $t$ and (iv) $M$ represents the set of physician visits required by a patient indexed by $m$.

Here, set $I$ include all the unique patient arrivals to the GMH ED for a week, which totals more than 1500. Set $K$ consists of the unique physician identification number that can start an ED shift for a day with an upper threshold of 25 physicians per day. Further, $T$ represents timeslots for an entire week (which varies based on slot length). Finally, set $M$ includes values from 1 through 4, representing the patient interaction with a physician. Next, we introduce the parameters considered in the model. Most of the parameters represent various patient characteristics, including severity, arrival time, physician visits, and fixed time slots that should be avoided for calculating patient wait time as these delays are inherent and one parameter defining the ED bed capacity.

- $\alpha_i$ represents the time slot of arrival for patient $i$.
- $\beta_i$ represents the severity level of patient $i$.
- $\gamma_i$ represents the total number of visits required by patient $i$.
- $w_i$ represents the total time slots for patient $i$ that should not be considered for waiting cost.
- $C$ represents the total bed capacity of the GMH ED.

Finally, we introduce the decision variables in the model:

- $\mathrm{U}_{ik}$ $\begin{Bmatrix} 1, \text{ If patient i served by physician k} \\ 0, \text{ otherwise} \end{Bmatrix}$
- $\mathrm{Ystart}_{kt}$ $\begin{Bmatrix} 1, \text{ If physician k starts their shift at time slot t} \\ 0, \text{ otherwise} \end{Bmatrix}$
- $\mathrm{Y}_{kt}$ $\begin{Bmatrix} 1, \text{ If physician k is available for service at time slot t} \\ 0, \text{ otherwise} \end{Bmatrix}$
- $\mathrm{X}_{iktm}$ $\begin{Bmatrix} 1, \text{ If patient i is served by physician k at time slot t for their visit m} \\ 0, \text{ otherwise} \end{Bmatrix}$

Minimize:

$$\mathrm{SC}^* \sum_{kt} \mathrm{Ystart}_{kt} + \mathrm{OC}^* \sum_{ikt} t^*\mathrm{X}_{ikt1} - \alpha_i + \mathrm{OC}^*\mathrm{F}^* \sum_{ikt} (t^*\mathrm{X}_{ikt\gamma_i} - t^*\mathrm{X}_{ikt1} - w_i) + \mathrm{HC}^* \sum_{ik} \mathrm{U}_{ik}$$

Subject to:

$$\sum_{kt} t * \mathrm{X}_{ikt1} \geq \alpha_i \qquad \forall i \in I$$

$$\sum_{ktm} \mathrm{X}_{iktm} = \gamma_i \qquad \forall i \in I$$

$$\sum_{km} \mathrm{X}_{iktm} \leq 2 \quad \forall i \in I, \forall t \in T$$

$$\sum_{kt} \mathrm{X}_{iktm} = 1 \quad \forall i \in I, \forall m \in M$$

$$\sum_{ikm} \mathrm{X}_{iktm} \le C \qquad \forall t \in T$$

$$\sum_{kt} t^{*}\mathrm{X}_{iktm} \le \sum_{kt} t^{*}\mathrm{X}_{iktm+1} \qquad \forall i \in I$$

$$\sum_{mt} \mathrm{X}_{iktm} \le 4^{*}\mathrm{U}_{ik} \qquad \forall i \in I, \forall k \in K$$

$$\sum_{im} \mathrm{X}_{iktm} \le 4^{*}\mathrm{Y}_{kt} \qquad \forall k \in K, \forall t \in T$$

$$\sum_{t} \mathrm{Ystrt}_{kt} \le 1 \qquad \forall k \in K$$

$$\sum_{kt} \mathrm{Ystrt}_{kt} \le K$$

$$8^{*}\mathrm{Ystrt}_{kt} \le \sum_{q=t}^{\mathrm{Min}(168,\, t+7)} \mathrm{Y}_{kq} \qquad \forall k \in K, \forall t \in T$$

$$\mathrm{U}_{ik}, \mathrm{Ystart}_{kt}, \mathrm{Y}_{kt}, \mathrm{X}_{iktm} \in \{0, 1\}$$

In the formulation, the objective function minimizes the cost of staffing the ED physicians, handoffs, patient onboarding, and patient waiting time in the ED. The cost of staffing an ED physician (*SC*) was based using the national average rate for ED physicians, and the onboarding cost (*OC*) for patients based on their ESI level was derived from the literature (Salary.com 2021; Woodworth and Holmes 2020). However, because of the lack of data on the cost of patient waiting once admitted, we used a factor value (*F*) between 0 and 1 and multiplied it by the *OC* to calculate the waiting cost. Finally, for the handoff cost (*HC*), we used high values ($1,000) to avoid any possible handoffs.

The first constraint ensures that a patient is served their first visit (m=1) only after their arrival at the ED. The second constraint ensures that the patient is provided with all their required visits before discharge. As mentioned earlier, each hour represents a time slot, but from observations and discussions with physicians, we assume that a physician can visit four patients in an hour. However, the same patient cannot be visited four times in an hour as that is not realistic as patients wait to get their tests, imaging, radiology, etc., completed. The third constraint ensures that a patient can be visited at most twice by a physician in an hour. The fourth constraint assures that each visit *m* for a patient cannot exceed 1, making sure that each visit is completed fully during a physician visit. The next constraint ensures that at any given time *t* the patients served cannot exceed the ED bed capacity. As patients have multiple interactions with physicians during an ED stay, these visits must be ordered such that a later visit (m+1) follows the prior visit (m) in terms of time slot, and our sixth constraint ensures the visits are ordered. The next two constraints ensure that a patient can be visited a maximum of four times by a physician, and a physician can visit up to four patients during any given time slot (1-hour block). The next two constraints ensure that a physician starts their shift only once a day and that the total number of physicians staffed per day does not exceed the maximum number of possible physicians that can work for a day based on health system budget constraints. To ensure that a physician shift, once started lasts for eight hours, we use the second to the last constraint. Finally, the last constraint defines the variable types, which are all binary in this case.

### 3.3 Simulation Model

We utilized a novel hybrid modeling approach to develop the discrete event simulation where both patients and physicians are represented as agents with unique attributes. This approach allowed us to simulate the actual patient arrivals to the PRISMA health ED with specific features, including severity level, arrival time, etc. Moreover, the main reason to adopt this modeling methodology was to replicate the physician activities in the ED in a realistic manner, including starting a shift at a particular time, spending time in their workstation ordering tests, updating a patient record, visiting patients' multiple times, and finally handing off a patient to the next physician when the shift ends. These activities would have been challenging to replicate using a traditional modeling approach where physicians are denoted as simple resources. A detailed description of the simulation model is discussed in our prior work (Girishan Prabhu et al. 2022).

After developing the simulation model of PRISMA Health ED, the next step was validating the model against actual data. We used the patient time in the ED for each ESI level and daily number of handoffs as the validation metrics to ensure that the model is representative of the partner ED. The model was simulated for a three-week schedule with an additional two-day warm-up period for the model to attain equilibrium. A total of 60 replications were performed, such that the margin of error on time in the ED metric was ± 10 minutes (at $\alpha$=0.05). We first compared the simulated weekly throughput and the daily number of handoffs to the retrospective data from partner ED. On performing a t-test, we observed that the actual weekly throughput (1508±8.2) and simulated weekly throughput (1505±5.3) values did not vary significantly (p-value = 0.90). Similarly, we observed that the actual daily handoffs (92±4.2) and simulated daily handoffs (93±5.8) values did not vary significantly (p-value = 0.11). Finally, comparing the simulated average time in the ED to the actual data for each ESI, we did not observe any significant differences (p-value > 0.05). Table 2 below represents the simulated and actual data for each ESI level.

Table 2: Simulation model validation.

| Severity | Actual Time in ED (mins) | Simulated Time in ED (mins) | Percent Difference |
|---|---|---|---|
| ESI 1 | 236 | 218 | -7.6% |
| ESI 2 | 272 | 281 | 3.3% |
| ESI 3 | 229 | 216 | -5.7% |
| ESI 4 | 114 | 121 | 6.1% |
| ESI 5 | 122 | 122 | 0% |

## 4 RESULTS

Upon developing the model and tuning the parameters on the training data, the next step was to use these models to forecast patient arrivals to the ED. We first discuss the findings from the long-term forecasting model – ED patient arrivals for the next 90 days. Table 3 represents the performance metrics score for each model output for the long-term forecasts. It is evident that both machine learning models outperformed the naïve model and other traditional time series models. However, it is interesting to notice that the Holt-Winters approach outperformed the ARIMA model, and this can be primarily attributed to the fact that the Holt-Winters model can account for seasonality. However, comparing the SARIMA model to the Holt-Winters model, SARIMA was slightly better. The most significant improvements were observed with the machine learning models, where the MAPE value was reduced by half compared to the traditional time series forecasting model. Among the two machine learning models, XGBoost outperformed the Random Forest model for all the performance metrics. One of the key observations here is the high RMSE values irrespective of the forecasting approach, which could be caused by extreme values (outliers). Even with a significant change in patient demands and arrivals, the machine learning models forecast was robust (based on RMSE, MAE, and MAPE), as models with a MAPE value of 5.0% are considered excellent. However, to avoid bias and over-relying on one value, we look at RMSE (16.6), which is comparatively low given the daily arrivals vary from 150 patients a day to as high as 270 patients.

After identifying the best-performing model, the next step was to look at the ESI predictions for the 90-day forecast. Table 4 above represents the performance metrics score for each ESI level from the XGBoost forecast. The first thing to notice is the varying RMSE, MAE, and MAPE values across the ESI levels. Specifically, it can be noticed that MAPE values are high for ESI 1 and 5 and minimum for ESI 3, whereas the RMSE and MAE behave vice versa. This represents the bias associated with each metric where MAPE penalizes heavily when the forecasted values are smaller as it is a percentage value. However, by using a combination of performance metrics, we can identify that the ESI-level forecasts from the model are robust.

Table 3: Model performance for the 90-day forecast.

| Model | RMSE | MAE | MAPE |
|---|---|---|---|
| MA | 30.1 | 23.6 | 14.2% |
| ARIMA | 27.2 | 21.6 | 10.6% |
| SARIMA | 25.6 | 19.2 | 9.9% |
| Holt-Winters | 26.8 | 19.8 | 10.0% |
| XGBoost | 16.6 | 14.1 | 5.9% |
| Random Forest | 17.4 | 14.6 | 6.4% |

Table 4: XGBoost ESI level 90-day forecast.

| ESI | RMSE | MAE | MAPE |
|---|---|---|---|
| ESI 1 | 3.1 | 2.8 | 38.0% |
| ESI 2 | 8.9 | 7.0 | 12.5% |
| ESI 3 | 13.4 | 10.1 | 8.4% |
| ESI 4 | 6.0 | 5.1 | 15.5% |
| ESI 5 | 1.9 | 1.4 | 46.1% |

The forecasted patient arrivals, along with other data, were inputted into the mathematical model to identify staffing schedules that can minimize patient handoffs, physician shifts, and patient wait times while considering the staffing budget. Table 5 below represents the physician shift start times for the week. While the table below represents the shift start times for the whole week, the mathematical model output provides specific start times and the number of shifts for each day, as some days require more resources.

Table 5: Weekly physician shift start times.

| Time | 00 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Current | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 22 | 4 | 20 | 5 | 0 | 2 | 0 | 0 | 21 | 4 | 20 | 0 | 0 | 0 | 0 | 15 | 21 |
| New | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 14 | 14 | 7 | 7 | 0 | 7 | 7 | 14 | 21 | 0 | 0 | 0 | 0 | 0 | 0 | 21 |

The first thing to notice here is how the start times are restricted to certain time frames that are very similar to the baseline policy, as these are operational policies where physicians cannot start shifts at certain hours (12:00 am, 2:00 am, etc.). Although the shift start windows are the same, one of the interesting factors to notice is how the schedule generated by the mathematical model recommends starting a shift in a staggering approach as opposed to starting shifts only at particular time frames (e.g., 7:00, 9:00, etc.) as observed in the current policy. Finally, in terms of total hours staffed for the week, there were no significant differences. After generating the new schedule, the next step was to compare the new policy to the current (baseline) policy in the validated simulation model. We used two ED performance metrics: number of handoffs and patient time in the ED. The first two metrics evaluate patient flow, and the third evaluates patient safety. Each policy was simulated for a three-week schedule and replicated until the margin of error on time in the ED metric was ± 10 minutes (at $\alpha$=0.05).

Table 6: Simulation model results.

| Policy | Weekly Throughput | # handoffs per day | Time in the ED (mins) | Change in hours/week (FTEs) |
|---|---|---|---|---|
| Baseline | 1505 | 93 | 213±4.6 | 0 |
| New | 1503 | 84 | 201±5.9 | +6 (+0.2 FTEs) |

From Table 6 above, on running an independent t-test, we observed that both handoffs per day and patient time in the ED varied significantly between the two policies (p-value <0.05), where the new policy outperformed the baseline policy. Compared to the baseline policy, the new policy reduced the patient time in the ED by 5.6% and handoffs by 9.6% with a slight non-significant (p-value = 0.21) increase in FTE.

## 5 CONCLUSIONS

Protecting the ED from crowding is one of the highest public health priorities to ensure timely patient care and patient safety. Although most ED across the US plan in advance to avoid ED crowding, studies have observed that most EDs still rely on quick temporary fixes. Although sometimes ad-hoc actions are required because of unexpected issues such as evacuations and natural disasters, most of the time, these are required because of inadequate short and long-term planning. One of the most important inputs required for robust planning is the future patient census (arrivals) to the ED. Over the last decades, several studies have applied numerous approaches for forecasting patient arrivals to the ED and have generated acceptable results. However, most of these studies have focused on predicting daily patient arrivals to the ED, except for two recent studies that have explored hourly patient arrival forecasting (Choudhury and Urena 2020; Zhang et al. 2022). Surprisingly, none of these studies has included ESI levels of forecasted patient arrivals which significantly influences resource requirements.

This research developed traditional time-series models and machine learning models to forecast long-term (90 days ahead) patient arrivals to the partner ED with the patient's ESI levels. XGBoost algorithm generated the best long-term forecasts with MAPE values of 5.9%, outperforming prior studies. Moreover, we forecast ESI levels of these arrivals with a maximum RMSE value of 13.4. These findings are promising, given the simple input variables and the long-term forecasts. Further, we utilized these forecasts in the mathematical model to allocate an optimal number of physicians while minimizing onboarding time, waiting time after ED admission, and patient handoffs. Finally, the generated schedule was tested in our validated simulation model representative of the partner ED, and we observed that the new policy reduced the patient time in the ED by 5.6% and handoffs by 9.6% with a non-significant increase in FTE.

Future research will focus on fine-tuning the forecasting model by incorporating other simple parameters that can be exported from EHR to investigate if the model predictions can be improved. Additionally, a hierarchical forecasting approach with an optimization function could potentially improve ESI-level forecasting. Additionally, a major limitation in the mathematical and simulation model is that we are not representing ancillary resources to the ED, including labs, consults, and nurses, as separate resources. However, the model still accounts for these delays. In future work, we plan to include the impact of these ancillary resources and processes to better represent the partner ED.

## ACKNOWLEDGMENTS

This study was supported by a Prisma Health grant and Harriet and Jerry Dempsey Professorship in Industrial Engineering at Clemson University.

## REFERENCES

Aboagye-Sarfo, P., Q. Mai, F. M. Sanfilippo, D. B. Preen, L. M. Stewart, and D. M. Fatovich. 2015. “A Comparison of Multivariate and Univariate Time Series Approaches to Modelling and Forecasting Emergency Department Demand in Western Australia”. *Journal of Biomedical Informatics* 57(1):62–73.

Ahsan, K. B., M. R. Alam, D. G. Morel, and M. A. Karim. 2019. “Emergency Department Resource Optimisation for Improved Performance: A Review”. *Journal of Industrial Engineering International* 15(1):253–266.

Alvarez, R., G. A. Sandoval, S. Quijada, and A. D. Brown. 2009. “A Simulation Study to Analyze the Impact of Different Emergency Physician Shift Structures in an Emergency Department”. *Proceedings of the 35th International Conference on Operational Research Applied to Health Services*, July 12$^{th}$ -17$^{th}$, Leuven, Belgium, 900-902.

American College of Emergency Physicians. 2019. “Crowding”. Policy Statement, American College of Emergency Physicians, Irving, Texas.

Batal, H., J. Tench, S. McMillan, J. Adams, and P. S. Mehler. 2001. “Predicting Patient Visits to an Urgent Care Clinic using Calendar Variables”. *Academic Emergency Medicine* 8(1):48–53.

Becerra, M., A. Jerez, B. Aballay, H. O. Garcés, and A. Fuentes. 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* 7(6):968-978.

Cairns, C., J. J. Ashman, and K. Kang. 2019. "Emergency Department Visit Rates by Selected Characteristics: United States". *NCHS data brief* 1(434): 1–8.

Carvalho-Silva, M., M. T. T. Monteiro, F. de Sá-Soares, and S. Dória-Nóbrega. 2018. "Assessment of Forecasting Models for Patients Arrival at Emergency Department". *Operations Research for Health Care* 18(4):112–118.

Centers for Disease Control and Prevention. 2010. NCHS Pressroom - Fact Sheet - Emergency Department Visits. https://www.cdc.gov/nchs/pressroom/04facts/emergencydept.htm, accessed 8th April 2023.

Choudhury, A., and E. Urena. 2020. "Forecasting Hourly Emergency Department Arrival using Time Series Analysis". *British Journal of Health Care Management* 26(1):34–43.

Connelly, L. G., and A. E. Bair. 2004. "Discrete Event Simulation of Emergency Department Activity: A Platform for System-level Operations Research". *Academic Emergency Medicine* 11(11):1177–1185.

Côté, M. J., M. A. Smith, D. R. Eitel, and E. Akçali. 2013. "Forecasting Emergency Department Arrivals: A Tutorial for Emergency Department Directors". *Hospital Topics* 91(1):9–19.

Derlet, R. W., and J. R. Richards. 2008. "Ten Solutions for Emergency Department Crowding". *Western Journal of Emergency Medicine* 9(1):24-36.

Di Somma, S., L. Paladino, L. Vaughan, I. Lalle, L. Magrini, and M. Magnanti. 2015. "Overcrowding in Emergency Eepartment: An International Issue". *Internal and Emergency Medicine* 10(2):171–175.

Elalouf, A., and G. Wachtel. 2021. "Queueing Problems in Emergency Departments: A Review of Practical Approaches and Research Methodologies". *Operations Research Forum* 3(1):1–46.

Füchtbauer, L. M., B. Nørgaard, and C. B. Mogensen. 2013. "Emergency Department Physicians Spend only 25% of their Working Time on Direct Patient Care". *Danish Medical Journal* 60(1):120-128.

George, F., and K. Evridiki. 2015. "The Effect of Emergency Department Crowding on Patient Outcomes Results". *Health Science Journal* 9(1):1–6.

Ghanes, K., O. Jouini, A. Diakogiannis, M. Wargon, Z. Jemai, R. Hellmann, V. Thomas, and G. Koole. 2015. "Simulation-based Optimization of Staffing Levels in an Emergency Department". *SIMULATION* 91(10):942–953.

Girishan Prabhu, V., K. Taaffe, R. Pirrallo, and D. Shvorin. 2020. "Stress and Burnout among Attending and Resident Physicians in the ED: A Comparative Study". *IISE Transactions on Healthcare Systems Engineering* 11(1):1–19.

Girishan Prabhu, V., K. Taaffe, R. G. Pirrallo, W. Jackson, and M. Ramsay. 2022. "Overlapping Shifts to Improve Patient Safety and Patient Flow in Emergency Departments". *SIMULATION* 98(11):961–978.

Hertzum, M. 2017. "Forecasting Hourly Patient Visits in the Emergency Department to Counteract Crowding". *The Ergonomics Open Journal* 10(1):1–13.

Hossein, A., J. Rouhi, S. Sardashti, A. Taghizadieh, H. Soleimanpour, and M. Barzegar. 2013. "Emergency Severity Index (ESI): A Triage Tool for Emergency Department". *International Journal of Emergency Medicine* 12(2):92-106.

Hsia, R. Y., A. L. Kellermann, and Y. C. Shen. 2011. "Factors Associated with Closures of Emergency Eepartments in the United States". *Journal of the American Medical Association* 305(19):1978–1985.

Jones, S. S., A. Thomas, R. S. Evans, S. J. Welch, P. J. Haug, and G. L. Snow. 2008. "Forecasting Daily Patient Volumes in the Emergency Department". *Academic Emergency Medicine* 15(2):159–170.

Kadri, F., F. Harrou, S. Chaabane, and C. Tahon. 2014. Time Series Modelling and Forecasting of Emergency Department Overcrowding. *Journal of Medical Systems* 38(9):1–20.

Kelen, G. D., R. Wolfe, G. D'onofrio, A. M. Mills, D. Diercks, S. A. Stern, M. C. Wadman, and P. E. Sokolove. 2021. "Emergency Department Crowding: The Canary in the Health Care System". *New England Journal of Medicine Catalyst*. 2(5):1–26.

Khaldi, R., A. El Afia, and R. Chiheb. 2019. "Forecasting of Weekly Patient Visits to Emergency Department: Real Case Study". *Procedia Computer Science* 14(8):532–541.

Kulstad, E. B., R. Sikka, R. T. Sweis, K. M. Kelley, and K. H. Rzechula. 2010. "ED Overcrowding is Associated with an Increased Frequency of Medication Errors". *American Journal of Emergency Medicine* 28(3):304–309.

Laxmisan, A., F. Hakimzada, O. R. Sayan, R. A. Green, J. Zhang, and V. L. Patel. 2007. "The Multitasking Clinician: Decision-making and Cognitive Demand during Team Handoffs in Emergency Care". *International Journal of Medical Informatics* 76(1):801–11.

Maughan, B. C., L. Lei, and R. K. Cydulka. 2011. "ED handoffs: Observed Practices and Communication Errors". *American Journal of Emergency Medicine* 29(5):502–511.

McDonnell, W. M., C. A. Gee, N. Mecham, J. Dahl-Olsen, and E. Guenther. 2013. "Does the Emergency Medical Treatment and Labor Act Affect Emergency Department Use?". *The Journal of Emergency Medicine* 44(1):209–216.

Morley, C., M. Unwin, G. M. Peterson, J. Stankovich, and L. Kinsman. 2018. "Emergency Department Crowding: A Systematic Review of Causes, Consequences and Solutions". *PLoS ONE* 13(8):1–42.

Moskop, J. C., D. P. Sklar, J. M. Geiderman, R. M. Schears, and K. J. Bookman. 2009. "Emergency Department Crowding, Concept, Causes, and Moral Consequences". *Annals of Emergency Medicine* 53(5):605–611.

Rais, A., and A. Vianaa. 2011. "Operations Research in Healthcare: A survey". *International Transactions in Operational Research* 18(1):1–31.

Rosychuk, R. J., E. Youngson, and B. H. Rowe. 2015. "Presentations to Alberta Emergency Departments for Asthma: A Time Series Analysis". *Academic Emergency Medicine* 22(8):942–949.

Salary.com. 2021. Physician - Emergency Room Salary in the United States. www.salary.com/research/salary/benchmark/er-doctor-salary, accessed 10th April 2023.

Sir, M. Y., D. Nestler, T. Hellmich, D. Das, M. J. Laughlin, M. C. Dohlman, and K. Pasupathy. 2017. "Optimization of Multidisciplinary Staffing Improves Patient Experiences at the Mayo Clinic". *INFORMS Journal on Applied Analytics* 47(5):425–441.

Sun, Y., B. H. Heng, Y. T. Seow, and E. Seow. 2009. "Forecasting Daily Attendances at an Emergency Department to Aid Resource Planning". *BMC Emergency Medicine* 9(1):1–9.

Venkatesh, A. K., D. Curley, Y. Chang, and S. W. Liu. 2015. "Communication of Vital Signs at Emergency Department Handoff: Opportunities for Improvement". *Annals of Emergency Medicine* 66(2):125–130.

Whitt, W., and X. Zhang. 2019. "Forecasting Arrivals and Occupancy Levels in an Emergency Department". *Operations Research for Health Care* 21(2):1–18.

Woodworth, L., and J. F. Holmes. 2020. "Just a Minute: The Effect of Emergency Department Wait Time on the Cost of Care. Economic". *Inquiry* 58(2):698–716.

Xu, M., T. C. Wong, and K. S. Chin. 2013. "Modeling Daily Patient Arrivals at Emergency Department and Quantifying the Relative Importance of Contributing Variables using Artificial Neural Network". *Decision Support Systems* 54(3):1488–1498.

Zhang, Y., J. Zhang, M. Tao, J. Shu, and D. Zhu. 2022. "Forecasting Patient Arrivals at Emergency Department using Calendar and Meteorological Information". *Applied Intelligence* 52(10):11232–11243.

## AUTHOR BIOGRAPHIES

**VISHNUNARAYAN GIRISHAN PRABHU** is an Assistant Professor in the Industrial and Systems Engineering at the University of North Carolina at Charlotte. His research interests include developing mathematical models, simulation models and machine learning models to improve healthcare. His email address is vgirisha@charlotte.edu.

**KEVIN TAAFFE** is the Harriet and Jerry Professor and Chair of the Dept. of Industrial Engineering at Clemson University. His teaching and research interests include the application of simulation and optimization in healthcare, production, and transportation logistics. Dr. Taaffe focuses on healthcare logistics problems that range from patient flow to operating room scheduling. He enjoys working on projects that bridge the gap between theoretical research and application. His email address is taaffe@clemson.edu.

**RONALD G PIRRALLO** is an emergency physician and the Vice-Chair for Academic Affairs for the Upstate Prisma Health Department of Emergency Medicine. He also holds an appointment as a Professor of Emergency Medicine at the University of South Caroline School of Medicine – Greenville. His email address is Ronald.Pirrallo@Prismahealth.org.

**WILLIAM JACKSON** is an emergency physician and the Vice-Chair of Clinical Operations and Clinical Affairs for the Upstate Prisma Health Department of Emergency Medicine where he holds an appointment as an Assistant Professor of Emergency Medicine at the University of South Carolina, School of Medicine Greenville. His experience includes managing clinical operations, disaster preparedness, healthcare systems quality programming, managing large group practice risk, creating, and implementing ED design, and developing innovations in patient care delivery. His email address is William.Jackson@Prismahealth.org.

**MICHAEL RAMSAY** is an emergency medicine physician and the Director for Informatics in the Department of Emergency Medicine at Prisma Health Greenville Memorial Hospital. His email address is Michael.Ramsay@Prismahealth.org.

**JESSICA HOBBS** is an emergency medicine physician and the Medical Director for the Department of Emergency Medicine at Prisma Health Greenville Memorial Hospital. Her email address is Jessica.Hobbs@Prismahealth.org.