



Theoretical bounds and approximation of the probability mass function of future hospital bed demand

Samuel Davis¹ · Nasser Fard¹

Received: 5 April 2018 / Accepted: 19 October 2018 / Published online: 6 November 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Failing to match the supply of resources to the demand for resources in a hospital can cause non-clinical transfers, diversions, safety risks, and expensive under-utilized resource capacity. Forecasting bed demand helps achieve appropriate safety standards and cost management by proactively adjusting staffing levels and patient flow protocols. This paper defines the theoretical bounds on optimal bed demand prediction accuracy and develops a flexible statistical model to approximate the probability mass function of future bed demand. A case study validates the model using blinded data from a mid-sized Massachusetts community hospital. This approach expands upon similar work by forecasting multiple days in advance instead of a single day, providing a probability mass function of demand instead of a point estimate, using the exact surgery schedule instead of assuming a cyclic schedule, and using patient-level duration-varying length-of-stay distributions instead of assuming patient homogeneity and exponential length of stay distributions. The primary results of this work are an accurate and lengthy forecast, which provides managers better information and more time to optimize short-term staffing adaptations to stochastic bed demand, and a derivation of the minimum mean absolute error of an ideal forecast.

Keywords Bed demand forecast · Patient flow · Length of stay distributions · Adaptive staffing

MSC Codes: 62H10 · 62H12 · 62H15

1 Introduction

1.1 Background

Significant cost and safety issues occur when the demand for hospital resources is not matched by the supply of these resources [1]. When demand exceeds supply, unsafe conditions arise due to less availability of resources, which can cause increased mortality and the rate of medical errors [2, 3]. Conversely, when supply exceeds demand, excess clinician staffing can produce significant cost with no benefit to patient care quality. Hospital management must actively manage the supply of fixed resources, such as beds and specialized

equipment, and variable resources, such as nurses, technicians, and providers, that affect care quality and ensure patient safety. Matching supply to demand is difficult because the future patient resource demands are random and unknown, while the resource supply is often set far in advance and is difficult to adjust in the short term significantly.

Uncertainty in patient resource demand is caused by several stochastic processes, including the number and timing of arrivals and discharges, length of stay (LOS), unit transfers, health improvement and deterioration, surgical complications, and same-day cancellations for outpatient and surgical appointments. Variability arises from disease seasonality, holidays, and the numerous types of specialized staff, physical resources, ailments, procedures, allergies, comorbidities, and surgical techniques. Additional variability is created when surgeries are scheduled without anticipating and managing the aggregate downstream resource requirements, such as the time needed in a post-anesthesia care unit (PACU) or intensive care unit (ICU). The aggregate patient resource demand several weeks in the future, which is the time period for which clinician shifts are being scheduled, is often significantly different from the current demand. Short-term staffing

✉ Samuel Davis
davis.sam@husky.neu.edu

Nasser Fard
n.fard@neu.edu

¹ Northeastern University, 334 Snell Engineering Center,
Boston, MA 02115, USA

adaptations are made to manage the mismatch of resource supply and demand, such as incorporating on-call and per diem staff and canceling unnecessary shifts.

Effective proactive adaptive staffing requires high-confidence forecasts of short-term resource demand to ensure the appropriate quantity and timing of staffing adjustments. Proactive is harder than reactive adaptation, in which the mismatch of supply and demand already exists, which is relatively simple to measure and align. The cost of a staffing adaptation is typically dependent upon its magnitude and timing, and thus requires a forecast that is both accurate and lengthy to be cost-effective. Developing and applying an accurate model to forecast patient resource demand for multiple time periods into the future improves both the cost and safety of providing care when coupled with an adaptive staffing strategy.

1.2 Literature review

Predicting and managing the demand for healthcare resources, such as beds, staff, and specialized equipment, is an extensive research topic. Settings have included emergency departments (EDs) [4–9], surgical suites and recovery beds [10–12], medical and surgical inpatient beds [13–16], ICUs [17–25], outpatient care [26–28] and long-term care [29, 30]. Methods have forecasted aggregate demand either by applying tools like regression [14, 18], time series analysis [8, 15, 31, 32], and neural networks [23, 31], or aggregating individual forecasts by applying tools like probability models [16, 29, 33–35], queueing theory [17, 36, 37], simulation [19, 20, 26, 27, 30, 38–46], and expert opinion [7, 47].

The wealth of literature on predicting the demand for hospital resources demonstrates the strong need to make accurate forecasts. Most models make simplifying assumptions to manage tractability and fit data availability, including assuming stationary or cyclic demand patterns [16, 17, 36], exponential-based inter-arrival times and LOS distributions [17, 36, 48, 49], patient homogeneity [10, 31, 36], single-day forecasts [47], and point estimates instead of probability mass functions (PMFs) [14, 31]. The model derived herein provides a generalized approach to forecasting bed demand that does not require these simplifying assumptions.

1.3 Objective

The goal of this study is to forecast an accurate approximation for the multi-period PMF of bed demand using the exact surgical schedule at the time of the forecast, non-stationary inter-arrival times, and patient-level duration-varying LOS distributions. Required to construct this model is an analysis of scheduled patient LOS distributions, including the probability of being an inpatient given information known at the time of the forecast.

The rest of this paper is organized as follows: first, the forecasting model is derived for each patient group. These

sub-forecasts are then aggregated, and their aggregate distribution is approximated, producing both a total patient PMF forecast and an inventory of the necessary parameters to construct the model. The theoretical minimum expected mean absolute error (MAE) for the forecast is then derived, which helps assess the quality of the model. A case study using blinded real hospital data demonstrates the construction and assessment of the model, followed by a discussion of results, limitations, and future work.

2 Forecasting model

2.1 Model structure

The forecasting model uses discrete time measured in periods and is run at period d . All patients currently in the hospital at that time have their discrete LOS incremented by one period. Let T represent the length of the complete forecast, and t be the specific period being forecasted, such that $d + 1 \leq t \leq d + T$. These time periods could correspond to days, non-overlapping shifts, or smaller increments of time. The model may be used to forecast a single unit or collection of hospital units. For instance, a forecast could be created for just an ICU, as well as the complete set of beds in the hospital. The first forecast would focus on ICU demand, while the second would provide information about the capacity of the hospital as a whole.

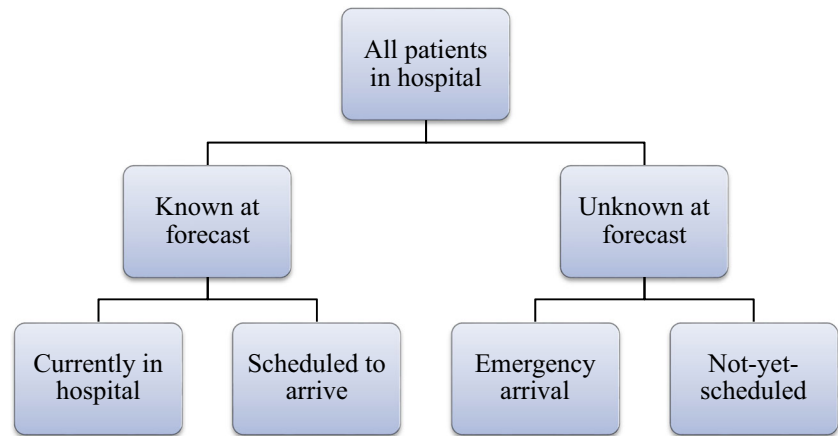
At any future period t , each patient in a hospital will either be known at current period d or unknown. At period d , all known patients existing at period t will either already be in the hospital or scheduled to arrive on or before period t . All remaining patients in the hospital at period t that were not known of at period d are either emergency arrivals, or patients who have not yet been added to the schedule, but will be scheduled to arrive on or before period t . Figure 1 shows these patient groups and their hierarchy.

Let $\gamma_d^{group}(t)$ represent the forecasted random variable of demand at period t made from period d for a given patient group, where $group \in \{curr, sch, emer, nysched\}$, representing current patients, scheduled patients, emergency patients, and not-yet-scheduled patients, respectively. Forecasts must be made for each of these patient groups individually and then aggregated to produce the overall patient demand forecast. All patient arrivals and LOS probabilities are assumed to be independent of one another.

2.2 Current patients

Each current patient $i \in M$, where M represents the number of current patient types indexed by i , currently in the hospital at period d , either is or is not still in the hospital at future period t , and can thus be represented by a Bernoulli random variable

Fig. 1 Patient groups in the bed demand forecast



where success is equal to the probability of having a remaining LOS from strictly greater than $t - d$. Let $f_i(h)$ represent the probability of patient i staying exactly h periods from d , meaning the PMF of their remaining LOS is $f_i(h)$. Let $F_i(h)$ be the cumulative distribution function (CDF) of $f_i(h)$, and $R_i(h) = 1 - F_i(h)$. The probability of patient i still being present at period t when forecasted from period d is represented by $C_{d,i}(t)$, as shown in Eq. 1.

$$C_{d,i}(t) = 1 - \sum_{h=1}^{t-d} f_i(h) = 1 - F_i(t-d) = R_i(t-d) \quad (1)$$

The convolution of all individual patient probabilities shown in Eq. 1, represented by $\gamma_d^{curr}(t)$, is the random variable representing the number of current patients at period d still present at period t . If these patients all have the same probability of being present at period t , then $\gamma_d^{curr}(t)$ is a Binomial random variable. However, if these probabilities are non-homogeneous, then their convolution follows a Poisson Binomial (PB) distribution [50]. Thus $\gamma_d^{curr}(t)$, which represents the random variable representing all current patients at period d being present at period t , follows a PB distribution, with mean and variance as shown in Eqs. 2 and 3, respectively.

$$E[\gamma_d^{curr}(t)] = \sum_{i=1}^M C_{d,i}(t) = \sum_{i=1}^M R_i(t-d) \quad (2)$$

$$Var[\gamma_d^{curr}(t)] = \sum_{i=1}^M [C_{d,i}(t) \times (1 - C_{d,i}(t))] = \sum_{i=1}^M [F_i(t-d) \times R_i(t-d)] \quad (3)$$

2.3 Scheduled patients

A patient $j \in N$, where N represents the number of scheduled patient types indexed by j , still being present at the hospital at period t can be represented by a Bernoulli random variable, similar to current patients. Let s_j represent the future period on which the patient is scheduled to arrive, such that $d + 1 \leq s_j \leq t$. Let $f_j(h)$ represent the PMF of staying exactly h periods, $F_j(h)$ be the CDF of $f_j(h)$, and $R_j(h) = 1 - F_j(h)$. Then, as represented

in Eq. 4, the probability of patient j , scheduled for period s_j , being present at period t , when forecasted from period d , is represented by $S_{d,j}(t)$. Note that h may equal 0.

$$S_{d,j}(t) = 1 - \sum_{h=0}^{t-s_j} f_j(h) = 1 - F_j(t-s_j) = R_j(t-s_j) \quad (4)$$

Similar to current patients, $\gamma_d^{sch}(t)$ is the convolution of the individual patient distributions shown in Eq. 4, which is the sum of non-homogeneous Bernoulli random variables, and thus PB. The mean and variance of this random variable are represented by Eqs. 5 and 6, respectively.

$$E[\gamma_d^{sch}(t)] = \sum_{j=1}^N S_{d,j}(t) = \sum_{j=1}^N R_j(t-s_j) \quad (5)$$

$$Var[\gamma_d^{sch}(t)] = \sum_{j=1}^N [S_{d,j}(t) \times (1 - S_{d,j}(t))] = \sum_{j=1}^N [F_j(t-s_j) \times R_j(t-s_j)] \quad (6)$$

2.4 Emergency patients

Emergency patients arrive at the hospital without ever being scheduled in advance, and are traditionally modeled as Poisson random variables in the literature [51], with a time-varying arrival rate. The arrival rate should be determined for each period, and should thus vary across time. Let the expected number of arrivals on period s , where $d + 1 \leq s \leq t$, of patient type $u \in U$, where U represents the number of emergency patient types indexed by u , be represented by $\lambda_u(s)$, and the probability of staying exactly h nights be represented by $f_u(h)$, where $f_u(0) = 0$ for all u . Let $F_u(h)$ be the CDF of $f_u(h)$, and $R_u(h) = 1 - F_u(h)$. Because both thinning and summing Poisson random variables produces a Poisson random variable, the sum of the expected remaining patients from each previous period's arrivals is the expected value of a Poisson-distributed variable of the total remaining emergency patients. The number of future emergency patients present at period t , when forecasted from period d , represented by Poisson random

variable $\gamma_d^{emer}(t)$, has a mean equal to its variance, which is shown in Eq. 7.

$$\begin{aligned}
 E[\gamma_d^{emer}(t)] &= Var[\gamma_d^{emer}(t)] \\
 &= \sum_{s=1}^t \sum_u^U \left[\lambda_u(s) \times \left(1 - \sum_{h=0}^{t-s} f_u(h) \right) \right] \\
 &= \sum_{s=1}^t \sum_u^U [\lambda_u(s) \times (1 - F_u(t-s))] \\
 &= \sum_{s=1}^t \sum_u^U [\lambda_u(s) \times R_u(t-s)] \tag{7}
 \end{aligned}$$

2.5 Not-yet-scheduled patients

Patients booked at period b to arrive on period s , such that $d + 1 \leq b \leq s \leq t$, are not-yet-scheduled, but may be present at period t . These patients are unknown, and can thus not be modeled with Bernoulli random variables. There is little research on this topic, though linear regression has been shown to reasonably predict the number of not-yet-scheduled patients [52]. Suppose the number of patients of type $v \in V$, where V represents the number of emergency patient types indexed by v , scheduled on a given booking period b for scheduled period s is Poisson distributed and represented by $\lambda_v(b, s)$. The probability of patients of type v staying exactly h nights is represented by $f_v(h)$, where $f_v(0) = 0$ for all v . Let $F_v(h)$ be the CDF of $f_v(h)$, and $R_v(h) = 1 - F_v(h)$. The random variable representing not-yet-scheduled patients present at period t , with mean and variance as shown in Eq. 8, is thus represented by $\gamma_d^{mysch}(t)$.

$$\begin{aligned}
 E[\gamma_d^{mysch}(t)] &= Var[\gamma_d^{mysch}(t)] = \sum_{s=1}^t \sum_{b=1}^s \sum_v^V \left[\lambda_v(b, s) \times \left(1 - \sum_{h=0}^{t-s} f_v(h) \right) \right] \\
 &= \sum_{s=1}^t \sum_{b=1}^s \sum_v^V [\lambda_v(b, s) \times (1 - F_v(t-s))] \\
 &= \sum_{s=1}^t \sum_{b=1}^s \sum_v^V [\lambda_v(b, s) \times R_v(t-s)] \tag{8}
 \end{aligned}$$

2.6 Total demand

Determining the distribution of total demand at period t from forecast period d , represented by $\Gamma_d(t)$, requires a convolution of known and unknown patients present at period t . The PMF of current and scheduled patients are PB distributions, meaning the distribution of known patients is also PB, as the sum of two sums of heterogeneous Bernoulli random variables is also a sum of heterogeneous Bernoulli random variables. The PMF of emergency and not-yet-scheduled patients are Poisson, meaning the distribution of unknown patients is also Poisson.

Calculating the full convolution of total patient demand $\Gamma_d(t)$, which is the sum of a PB random variable representing

known patients, and a Poisson random variable representing unknown patients, is computationally expensive to calculate [53]. This is partly because it has an infinite domain and the number of unique combinations of the PB distribution is equal to $2^{(M+N)}$, where $M+N$ is the number of known patients, which exceeds one trillion with just 40 known patients. Instead of calculating the complete convolution of $\Gamma_d(t)$, an approximation using the normal distribution is obtained using the mean and variance of forecasted patient demand. The mean, as shown in Eq. 9, is the sum of Eqs. 2, 5, 7, and 8, and the variance, as shown in Eq. 10, is the sum of Eqs. 3, 6, 7, and 8. This Normal distribution, as well as the underlying convolution that includes a Poisson distribution, both have infinite domains. These distributions represent the demand for beds, and can thus exceed the hospital unit capacity. When calculating the bed demand, this Normal distribution should be truncated at zero, and when predicting the bed occupancy, this Normal distribution should be truncated at both zero and the hospital unit capacity.

$$\begin{aligned}
 E[\Gamma_d(t)] &= E[\gamma_d^{curr}(t)] + E[\gamma_d^{sch}(t)] + E[\gamma_d^{emer}(t)] \\
 &\quad + E[\gamma_d^{mysch}(t)] \\
 &= \sum_{i=1}^M R_i(t-d) + \sum_{j=1}^N R_j(t-s_j) \\
 &\quad + \sum_{s=1}^t \sum_u^U [\lambda_u(s) \times R_u(t-s)] \\
 &\quad + \sum_{s=1}^t \sum_{b=1}^s \sum_v^V [\lambda_v(b, s) \times R_v(t-s)] \tag{9}
 \end{aligned}$$

$$\begin{aligned}
 Var[\Gamma_d(t)] &= Var[\gamma_d^{curr}(t)] + Var[\gamma_d^{sch}(t)] \\
 &\quad + Var[\gamma_d^{emer}(t)] + Var[\gamma_d^{mysch}(t)] \\
 &= \sum_{i=1}^M [F_i(t-d) \times R_i(t-d)] \\
 &\quad + \sum_{j=1}^N [F_j(t-s_j) \times R_j(t-s_j)] \\
 &\quad + \sum_{s=1}^t \sum_u^U [\lambda_u(s) \times R_u(t-s)] \\
 &\quad + \sum_{s=1}^t \sum_{b=1}^s \sum_v^V [\lambda_v(b, s) \times R_v(t-s)] \tag{10}
 \end{aligned}$$

2.7 Error calculations

Let the error of the forecast for period t made from period d be represented by $e_{d, t}$. If A_t is the actual census for period t , then Eq. 11 shows the calculation for $e_{d, t}$. The Z-score of the error, represented in Eq. 12 by $Z_{d, t}$, equals the number of standard

deviations away from the forecasted mean and the standard deviation is equal to the square root of the forecast variance.

$$e_{d,t} = A_t - E[\Gamma_d(t)] \tag{11}$$

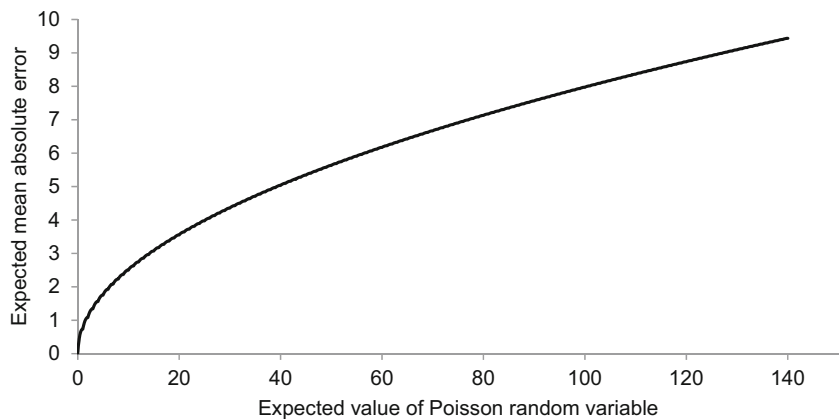
$$Z_{d,t} = \frac{e_{d,t}}{\sqrt{\text{Var}[\Gamma_d(t)]}} \tag{12}$$

The case where all LOS distributions are deterministic, meaning no uncertainty in LOS, such that all variance comes from unknown future patients, is the theoretical lowest-variance forecast. In this case, the mean and variance of the PB random variable would become an integer constant and zero, respectively. Thus, the distribution of the forecasted PMF with the lowest variance is a translated Poisson random variable that is translated by the amount equal to the expected value of the PB random variable. If a forecast is made equal to the mean of this translated Poisson random variable, then the expected MAE of the forecasted PMF with the lowest variance can be obtained as shown in Eq. 13, where $\lambda + c$ and λ represent the mean and variance of the translated Poisson (TP) random variable, respectively. The proof is shown in the appendix. Note that $\lfloor \lambda \rfloor$ represents the largest integer less than or equal to λ . Translating the Poisson random variable does not affect the expected MAE, and the result in Eq. 13 can be further reduced, as shown in Eq. 14, where f_P represents the PMF of a standard Poisson random variable with a mean of λ . To calculate the theoretical best expected value of a multi-period forecast with non-homogenous Poisson random variables, the value in Eq. 14 must be averaged for all periods, using the actual number of arrivals to estimate λ . Figure 2 shows the expected MAE for a single Poisson random variable as a function of λ .

$$E[MAE_{TP}] = \sum_{k=c}^{\lfloor \lambda \rfloor + c} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) (\lambda + c - k) + \sum_{k=\lfloor \lambda \rfloor + c + 1}^{\infty} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) (k - \lambda - c) = \frac{2e^{-\lambda} \lambda^{\lfloor \lambda \rfloor + 1}}{\lfloor \lambda \rfloor!} \tag{13}$$

$$E[MAE_P] = \frac{2e^{-\lambda} \lambda^{\lfloor \lambda \rfloor + 1}}{\lfloor \lambda \rfloor!} = 2\lambda \frac{e^{-\lambda} \lambda^{\lfloor \lambda \rfloor}}{\lfloor \lambda \rfloor!} = 2\lambda f_P(\lfloor \lambda \rfloor) \tag{14}$$

Fig. 2 Theoretical minimum expected mean absolute deviation of a Poisson random variable



2.8 Parameterization

Populating all the necessary parameters for a forecast of T total periods requires significant parameterization and calculation. Table 1 shows the six groups of parameters that must be developed, along with their corresponding indices. These parameters can either be determined from the empirical distribution of the dataset or produced using more powerful methods, such as machine learning. For instance, historical data could populate a logistic regression or decision tree model to estimate the probability of staying any number of nights. Several other works have thoroughly analyzed LOS predictions, and are good resources for this important aspect of forecasting bed demand [13, 49, 37]. Unlike for the other patient groups, $f_j(0)$ for scheduled patients must be determined.

3 Case study

Blinded de-identified data were obtained for all patients scheduled to arrive or having stayed between 10/31/2015 and 9/16/2016 at a mid-sized Massachusetts community hospital. The next three subsections describe the raw data, develop the model parameters, and produce and analyze the forecast results. This particular application of the model used days as periods to determine the midnight census, and combined all hospital units into a single entity. To assess the performance of the proposed model, a 7-period moving average (MA) model was also developed.

3.1 Descriptive analytics

Included in the hospital dataset were about 10,000 patients, 36% of whom were scheduled to arrive, versus 64% of whom were never scheduled. Of the scheduled patients, 26% had a LOS of at least one night, meaning the dataset included about 1000 scheduled inpatient surgeries. The dataset included fields for arrival date, departure date, if the patient was scheduled or not, and for all scheduled patients, fields for booking

Table 1 Parameters requiring calculation for a forecast of length T from period d

Parameters	Indices
$f_i^a(h)$	$\forall i \text{ and } h \geq 1$
$f_j^e(h)$	$\forall j \text{ and } h \geq 0$
$f_u^s(h)$	$\forall u \text{ and } h \geq 1$
$f_v^s(h)$	$\forall v \text{ and } h \geq 1$
$\lambda_u(s)$	$\forall u \text{ and } d+1 \leq s \leq d+T$
$\lambda_v(b, s)$	$\forall v \text{ and } d+1 \leq b \leq s \leq d+T$

date, surgeon, and procedure. The mean LOS for scheduled inpatients and emergency patients were 2.5 and 4.1 days, respectfully.

The nightly midnight census is set equal to the sum of new arrivals that day and the current patients that morning, minus the departed patients. Figure 3 shows the mean arrivals, current patients, and departures by DOW in waterfall form, where the bottom of the third bar represents the mean midnight census for that particular DOW. About one-third of all inpatient

surgeries occur on Mondays, and taper across the week, with the highest mean of surgical census occurring on Wednesday. Mean emergency patient arrivals also peak on Monday, with the census peaking on Wednesday. Few scheduled inpatient surgeries occur on weekends, whereas the mean number of emergency arrivals is roughly constant by DOW.

3.2 Model parameterization

Construction of the forecasting model required calculation of the flow probabilities and arrival rates in Table 1. In this case study, seven primary patient types were differentiated: emergency patients and six surgical patients grouped by orthopedic, general surgery, urology, gynecology, neurosurgery, and all other specialties combined.

The LOS distribution for emergency patients was set equal to the empirical LOS distribution for all emergency patients. Patients not spending a night in the hospital were excluded from the analysis due to not being part of the midnight

Fig. 3 Average inpatient arrivals and current inpatients and inpatient departures by DOW and split by scheduled and emergency and all patients combined

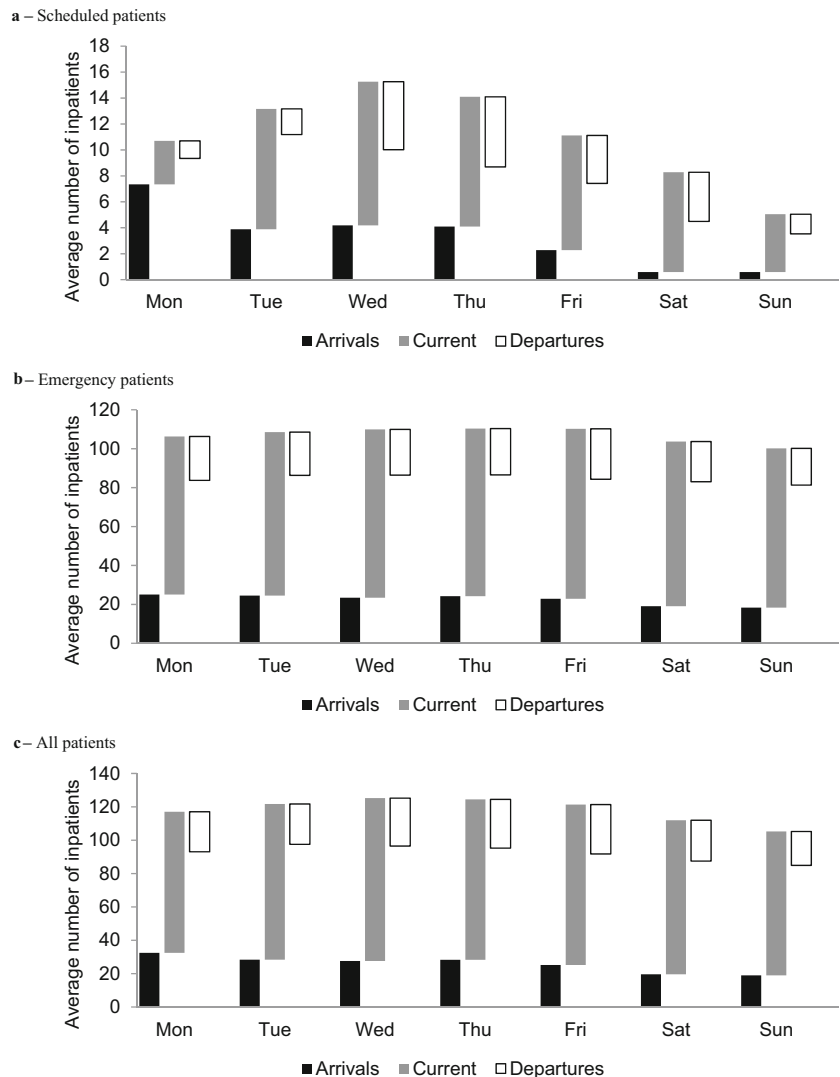
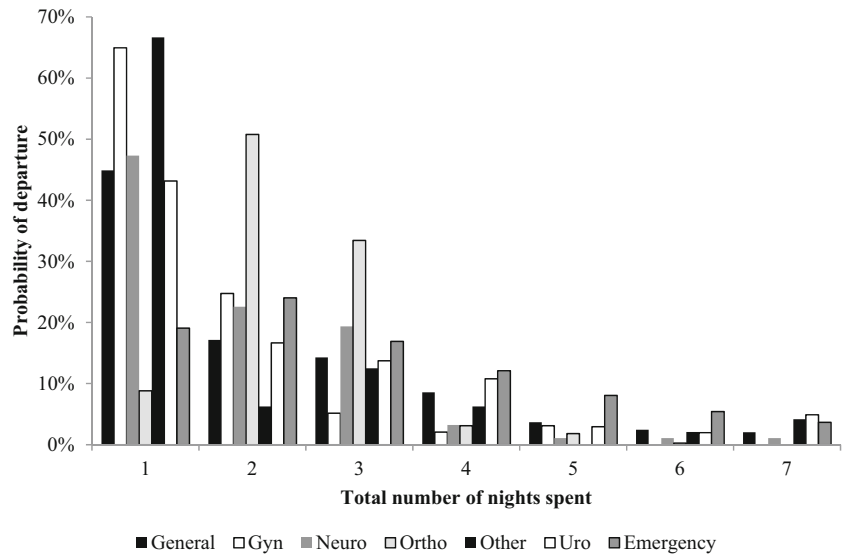


Fig. 4 LOS distributions conditioned on staying at least one night and truncated after seven nights



inpatient bed census. The LOS distributions for each surgical group were calculated assuming an inpatient stay of at least one night. The probability of a scheduled surgery becoming an inpatient stay was based on the proportion of patients having that particular procedure being an inpatient to the number scheduled. Hence, the LOS distribution was conditioned on staying at least one night, based purely on the procedure, then determined based on the surgical group. Figure 4 shows a bar chart containing the LOS distributions for each patient type, conditioned on staying at least one night, and truncated after seven nights. Due to data availability, the LOS distributions were not updated at each subsequent time period for each patient.

The arrival rates for emergency patients were set equal to the mean number of daily arrivals by DOW. Figure 5 contains box plots showing the arrival distributions by DOW. The bottom and top whiskers represent the minimum and maximum values, respectively, the centerlines represent the medians, and the bottom and top of the boxes represent the 25th and 75th

percentiles. Arrival rates for not-yet-scheduled patients were set to the empirical distribution of scheduled patients by surgical group, appointment DOW, and the number of days between the booking date and the appointment date.

3.3 Forecast results

From each date between 10/31/2015 and 8/26/2016, the patient volume was forecasted 14 days into the future, representing 301 forecasts for each of the 14 lengths. Figure 6a shows the forecasted 1-day PMF of demand from 11/4/2015 to 11/5/2015, based on the discrete Normal approximation using the mean and variance of demand. Figure 6b shows an example of a single 14-day forecast starting from 11/4/2015 versus the actual values, and Fig. 6c shows fourteen consecutive 1-day forecasts starting from 11/4/2015 versus the actual values. The lower control limits (LCL) and upper control limits (UCL) are drawn two standard deviations below and above the means in both figures, respectively, where the

Fig. 5 Arrival distributions for emergency patients

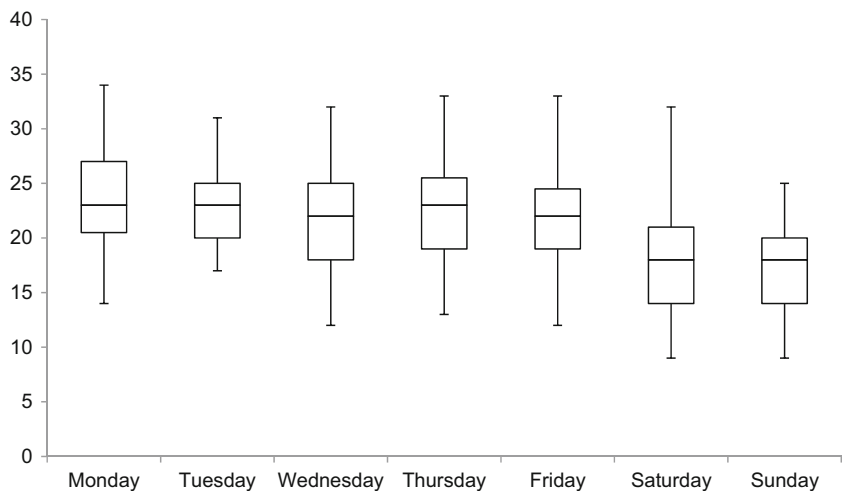
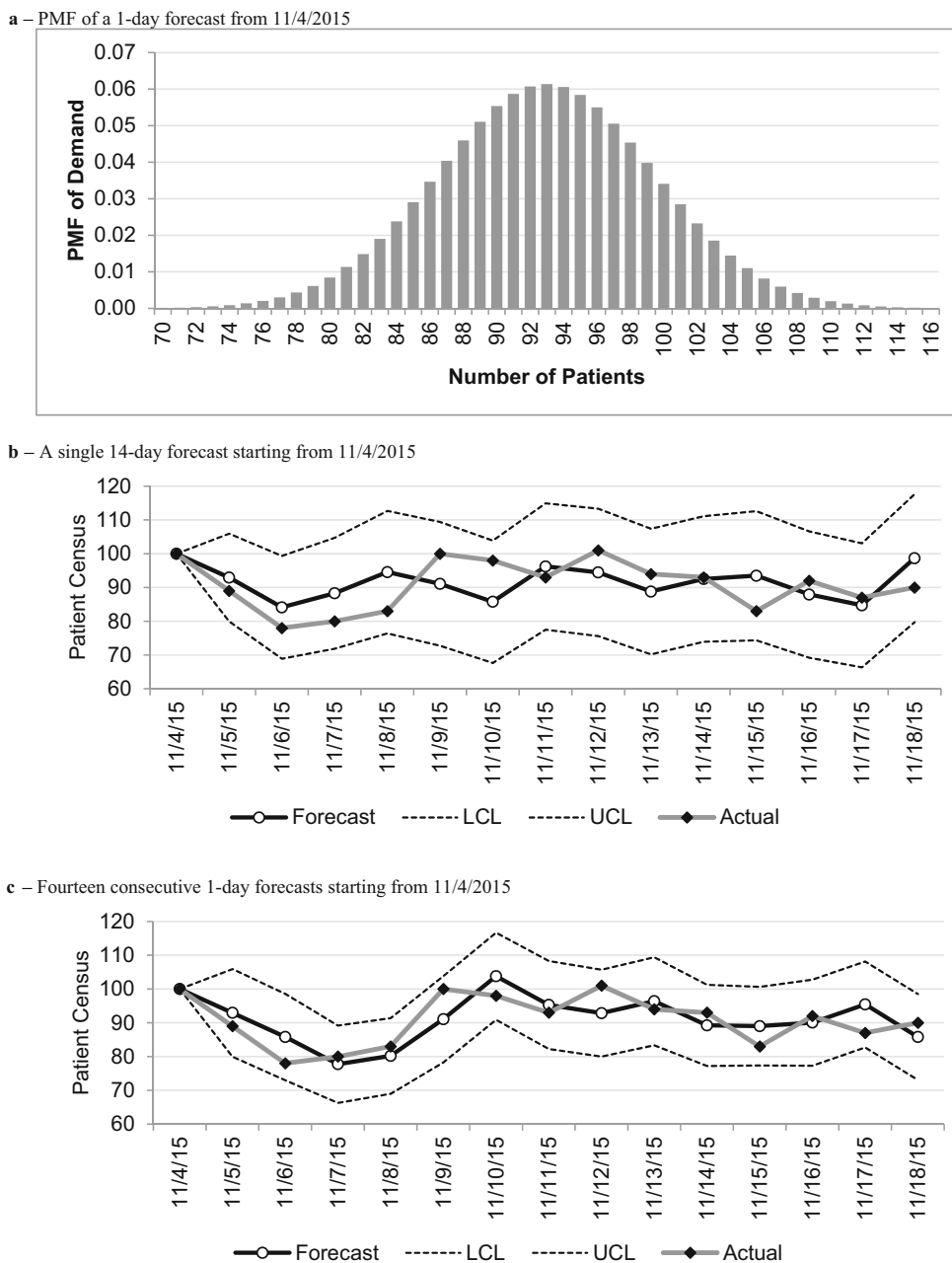


Fig. 6 A 1-day forecasted PMF followed by a comparison of a single 14-day forecast to fourteen consecutive 1-day forecasts for the same period



standard deviation is set to the square root of the forecasted variance.

From each day d , forecasts were made with a total length of 14 days, meaning $T = 14$. For each of the forecasted lengths, the MAE was calculated, as well as the mean and standard deviation of the Z-scores of the errors. These results are shown in Table 2, in addition to the theoretical minimum expected MAE as determined by Eq. 14. The theoretical MAE is calculated by setting λ equal to the actual number of unknown daily arrivals and taking the mean of the daily expected MAE. The simple 7-period MA model had a MAE of 6.0 for a 1-period forecast, representing a 17% increase in error compared to the proposed model.

Two-sided Z-tests were applied to each forecast length to compare the means of the forecast error Z-scores to zero. They all passed with $\alpha = .05$, meaning none of the means of the Z-score errors were statistically different than zero. Also using two-sided Z-tests, the means of the squared values of the Z-scores were compared to one for each forecast length. The difference between these values and one were not found to be statistically significant for any of the forecasted lengths. Given that the standard deviation of a random variable is equal to $\sqrt{E[X^2] - E[X]^2}$, with $E[X^2]$ not statistically different from one, and $E[X]$ not statistically different from zero, it is reasonable to conclude that the standard deviations of the forecasts

Table 2 Experimental results of the 14 forecast lengths with 301 forecasts each

Forecast Length	MAE of daily expected value			Mean of Z-scores	SD of Z-scores
	Observed	Theoretical	Difference		
1	5.1	3.7	1.4	-0.01	1.02
2	6.5	5.0	1.5	-0.02	1.02
3	6.9	5.7	1.2	-0.02	0.98
4	7.1	6.2	0.9	-0.03	0.96
5	7.3	6.5	0.8	-0.04	0.99
6	7.5	6.7	0.8	-0.05	1.00
7	7.6	6.8	0.7	-0.05	1.02
8	7.7	7.0	0.7	-0.06	1.03
9	7.7	7.0	0.7	-0.06	1.04
10	7.8	7.1	0.7	-0.08	1.05
11	7.8	7.1	0.7	-0.08	1.05
12	7.8	7.2	0.6	-0.08	1.06
13	7.8	7.2	0.6	-0.09	1.06
14	7.8	7.3	0.6	-0.09	1.06

were not statistically different from one. This finding supports the assumption of patient bed demand following a normal distribution.

In addition to testing the means and standard deviations of the Z-scores of the error, each of the 14 forecast lengths was tested for normality using the Kolmogorov-Smirnov test, in which the null hypothesis was that the forecast distribution was from a normal distribution with mean and variance equal to the forecasted mean and variance. If the test statistic exceeds the critical value, then the null hypothesis is rejected. Otherwise, if the null hypothesis is embraced, then statistical support exists for a normal distribution being a good fit for the true distribution. For all fourteen forecasts lengths, the null hypothesis was embraced at the 5% significance level, which supports the use of normal approximations for future bed demand.

To visually assess the overall fit of the normal approximation, Q-Q plots were created to identify outliers. Q-Q plots are used to compare the empirical distribution percentiles to the theoretical distribution percentiles, as shown in Fig. 7. The x-axis in each plot corresponds to the theoretical normal distribution percentiles, and the y-axis in each plot corresponds to the empirical distribution of the Z-scores of the forecast errors. A 45-degree line is drawn in each chart to show the ideal fit of the two sets of quantiles.

This forecasting model assumes that all arrival rates and LOS distributions are independent, and thus does not take into account system effects like congestion and negative feedback loops. A chi-squared test for independence was prepared to test if the 1-day forecast errors were independent of the census at the time of the forecast. Five bins were created for both the error and the census at the time of the forecast, centered on their respective means, creating a total of 25 bins in the joint

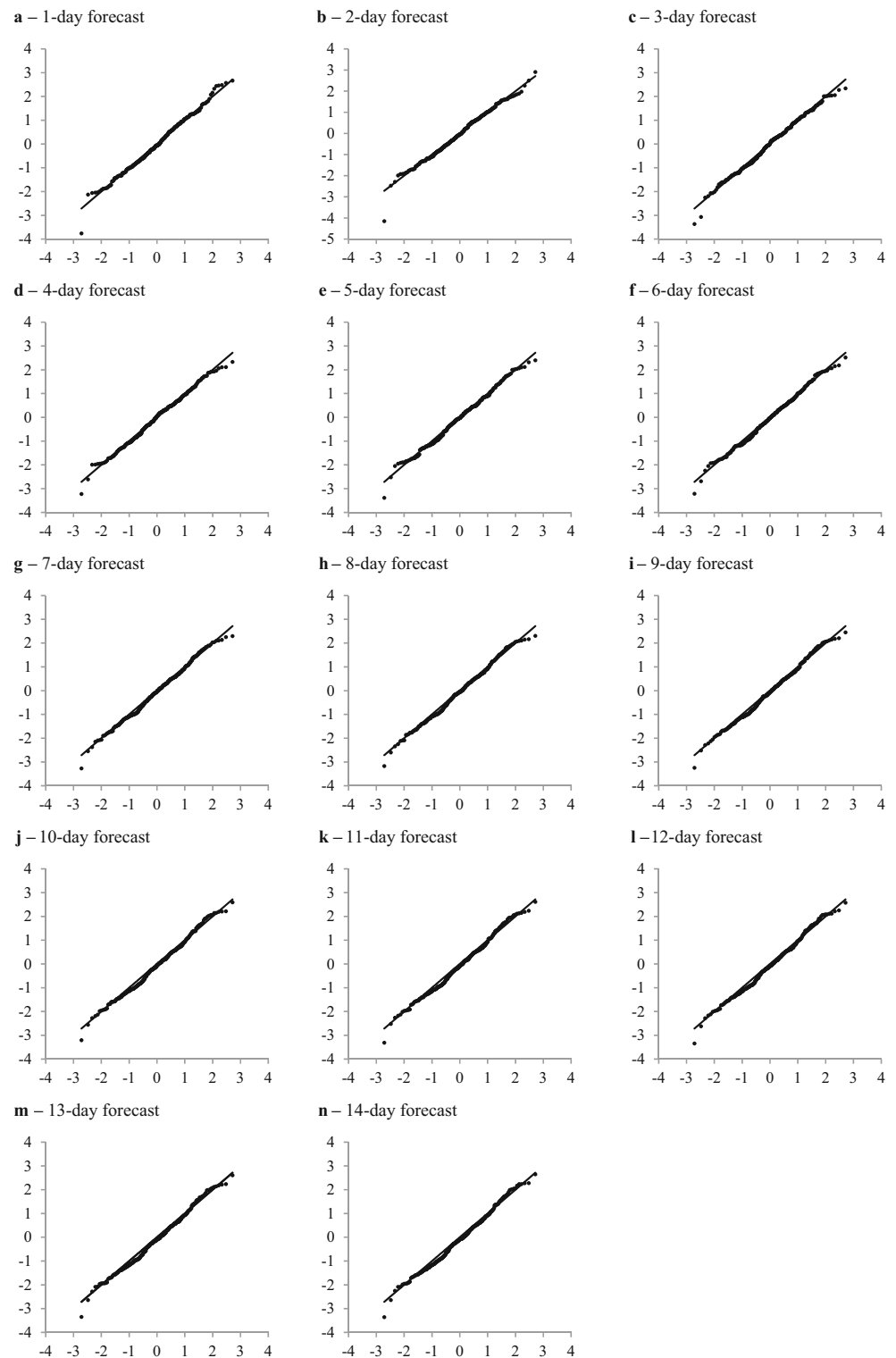
distribution. The middle bins represented the mean plus or minus half the standard deviation. The next bin on each side also had a width of a single standard deviation. Values more than 1.5 standard deviations above or below the mean were placed in the highest and lowest bins, respectively. The chi-squared test statistic was equal to 16.4, which with 16 degrees of freedom and $\alpha = .05$, was not found to be statistically significant, and hence, there was insufficient evidence to demonstrate non-independence. The forecast residuals versus the census at the forecast times are shown in Fig. 8, the distribution of observed values is shown in Table 3, and the distribution of expected values is shown in Table 4.

The total population of patients is a combination of known and unknown patients, represented by a PB and Poisson random variable, respectively. Figure 9 shows the average ratio of known patients to all patients from each forecast length, representing the proportion of the forecast that could theoretically be known with close to certainty if LOS distribution predictions improve and approach a smaller variance. As $t \rightarrow \infty$, when all known patients will have departed, and only unknown patients remain, the forecast converges to a Poisson random variable. On the first day of the forecast, nearly 80% of the mean of bed demand was from known patients, compared to just 20% on the seventh day.

4 Discussion

This paper demonstrates a novel and accurate method to produce a multi-day forecast of the PMF of hospital bed demand. The case study results are promising, including an MAE close to the theoretical minimum expected MAE, means of error Z-scores close to zero, standard deviations of error Z-scores

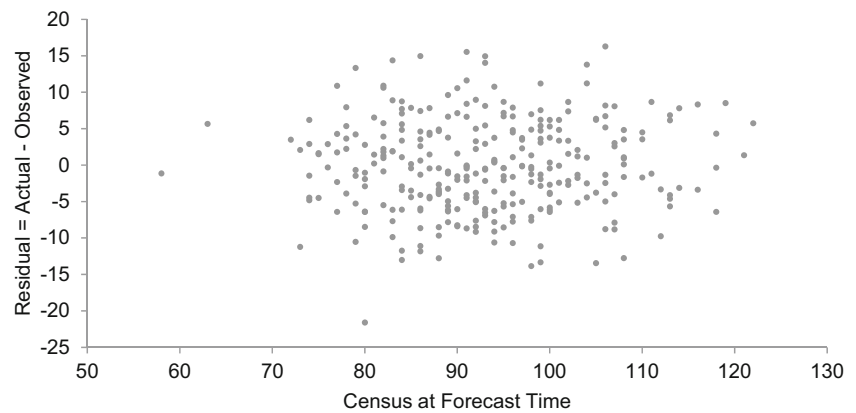
Fig. 7 Q-Q plots of case study forecasts for 1–14 day lengths where the x-axes represent the theoretical quantiles and the y-axes represent the empirical quantiles



close to one, and successful Kolmogorov-Smirnov tests for all forecasted lengths. In addition, the chi-squared test for independence demonstrated that forecast residuals were independent of the bed demand at the time of the forecasting, confirming that the assumption of independent patient flow variables was reasonable.

Predictably, the forecasts increase in MAE as their lengths increase, though the difference between the prediction and the theoretical minimum expected MAE decreases. This suggests that although there is room for improvement in predicting the arrival rates for emergency patients and not-yet-scheduled patients, the more significant forecast improvement opportunity

Fig. 8 1-day forecast residuals versus census at forecast time



lies in predicting (i.e., reducing the variance of the PMF of) the LOS distributions for current and scheduled patients. Improving LOS distributions could be accomplished via real-time updating from electronic medical records, including data on age, gender, medical history, and list of orders, to drive personalized classification and regression models. Furthermore, as natural language processing improves, valuable information from clinician notes may become available for predictive models to improve personalization.

Only 13% of the inpatient stays at this hospital were scheduled, representing only 8% of total nights spent. This relatively low proportion of scheduled patients significantly increases the variance of the theoretical best quality forecast due to the dominance of the Poisson random variable for unknown patients relative to the PB random variable for known patients. Applying the forecasting model to a hospital that predominantly sees scheduled patients relative to emergency patients could significantly reduce the MAE of the predictions. In addition, because the curve in Fig. 2 is concave down, the rate of expected MAE increase is slower than the rate of the increase of the mean of unknown patients, meaning this forecasting model should produce a smaller ratio of MAE to mean census at larger hospitals with similar LOS distributions.

The Q-Q plots in Fig. 7 demonstrate that a normal distribution with the forecasted mean and variance is a suitable estimation of the full probability convolution. The estimations hold well within the range of about two standard deviations,

indicated by little deviation from the 45-degree line. However, the 1-day high outliers are above the 45-degree line, indicating that the upper tail of the normal distribution estimate may be too thin to fit the real data. This makes sense given that Poisson distributions have positive skew and most of the forecast is driven by the underlying Poisson random variable for unknown patients. The Q-Q plots for other forecast lengths tend to have high outliers below the 45-degree line, suggesting that consistent multi-day volume increases may be mitigated over several days via reduced LOS or fewer scheduled surgeries. A prominent outlier occurs on the night of June 4th, 2016, when the 1-day forecast from the previous day had an error Z-score of -3.8 , and a 2-day forecast Z-score of -4.2 . The following day, June 5th, 2016, also clearly shows up on the 3-day Q-Q plot. Although an explanation for this outlier likely exists, tracing its root-cause was beyond the scope of the study.

A potential limitation of this work is the assumption that all random variables are independent of each other, as it is possible that some processes adjust when the census is extreme. For instance, a full hospital may prohibit non-critical short-term scheduling, or demonstrate a greater urgency to discharge current patients to make room for new ones. Although the chi-squared test demonstrated independence between the census at the time of the forecast and the next day's forecast residual, incorporating system-level adjustments in future forecast model expansion work may enhance the quality of the predictions in extreme cases.

Table 3 Observed distribution 2-way table showing the joint distribution of 1-day residuals and census at forecast times

{Min, Max}		Residual = Actual - Observed					Total
		< -9.5	-9.5, -3.2	-3.2, 3.1	3.1, 9.4	> 9.4	
Census at Forecast Time	< 76.7	1	3	8	3	0	15
	76.7, 87.6	7	19	27	22	6	81
	87.6, 98.4	5	40	36	24	7	112
	98.4, 109.3	4	15	25	23	4	71
	> 109.3	1	6	5	10	0	22
	Total	18	83	101	82	17	301

Table 4 Expected distribution 2-way table showing the joint distribution of 1-day residuals and census at forecast times

	{Min, Max}	Residual = Actual - Observed					Total
		< -9.5	-9.5, -3.2	-3.2, 3.1	3.1, 9.4	> 9.4	
Census at Forecast Time	< 76.7	0.9	4.1	5.0	4.1	0.8	15
	76.7, 87.6	4.8	22.3	27.2	22.1	4.6	81
	87.6, 98.4	6.7	30.9	37.6	30.5	6.3	112
	98.4, 109.3	4.2	19.6	23.8	19.3	4.0	71
	> 109.3	1.3	6.1	7.4	6.0	1.2	22
	Total	18	83	101	82	17	301

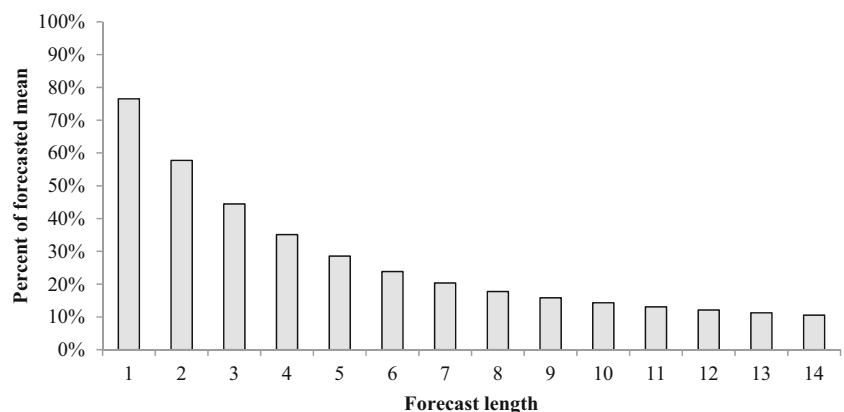
Another limitation is that the parameterization could benefit from the application of machine learning algorithms and additional data, especially data that is updated during the patient visit. For instance, learning that a patient has moved to the ICU or is on a ventilator would significantly increase the expected remaining LOS. Incorporating natural language processing to read the clinician notes would further improve the accuracy of the patient-level parameters. The difference in accuracy between the case study accuracy and theoretical best accuracy demonstrates the value in improving these parameters. As these parameter errors compound, the gap between the theoretical best accuracy and the model accuracy widens. Furthermore, error can exist within the input parameters themselves, in addition to the inherent error of any probability not equaling zero or one. This model assumes that probabilistic input parameters have been accurately determined, and an opportunity for future work would be to analyze the effect of poor parameter estimation on the aggregate demand prediction.

Future work could improve this model by incorporating unit-level flows and a joint distribution of demand across these units by addressing unit-level transfers, as well as including resource demand ratios to produce a PMF of resource demand. For instance, beyond predicting the demand for beds, which is a one-to-one relationship to patients, the model could be expanded to predict the need for nurses and technicians in

specific units, who can cover multiple patients based on severity and needs. The ultimate intent is to apply a stochastic optimization model that sets base level staffing and adjusts it based on a pool of shared resources based on a joint probability distribution of demand across all units. In addition to expanding the detail of the model structure and parameters, the full convolution of demand could be integrated to improve the predicted PMF accuracy.

5 Conclusion

An accurate forecasting model has been derived and implemented that generalizes previous models in the literature by allowing multi-period forecasts, PMFs of demand instead of point estimates, non-cyclic surgery schedules, time-varying LOS distributions, and patient heterogeneity. In a case study, the forecasts held up to statistical tests of independence and normality. This forecasting method allows bed managers significantly more information, both in accuracy and timeliness, to adapt staffing levels to stochastic demand, potentially leading to cost reduction and safety improvements. Most significantly, this work demonstrates a theoretical minimum error in a single-unit bed demand forecast, which can be used to assess the accuracy of future work in this domain.

Fig. 9 Percent of the mean from known patients for each forecast length

Appendix

Calculation of expected MAE of a translated Poisson random variable given its mean $(\lambda + c)$ and variance (λ) .

$$E[MAE_{TP}] = \sum_{k=c}^{[\lambda]+c} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) (\lambda + c - k) + \sum_{k=[\lambda]+c+1}^{\infty} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) (k - \lambda - c)$$

Split the sums along $\lambda + c$ and k .

$$= (\lambda + c) \sum_{k=c}^{[\lambda]+c} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) - (\lambda + c) \sum_{k=[\lambda]+c+1}^{\infty} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) - \sum_{k=c}^{[\lambda]+c} k \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) + \sum_{k=[\lambda]+c+1}^{\infty} k \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right)$$

Combine the sums without the extra k in the sum.

$$= -(\lambda + c) + 2(\lambda + c) \sum_{k=c}^{[\lambda]+c} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) - \sum_{k=c}^{[\lambda]+c} k \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) + \sum_{k=[\lambda]+c+1}^{\infty} k \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right)$$

Add and subtract identical sums to get $(k - c)$ in the sums that used to have just k .

$$= -(\lambda + c) + 2(\lambda + c) \sum_{k=c}^{[\lambda]+c} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) - \sum_{k=c}^{[\lambda]+c} (k-c) \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) + \sum_{k=[\lambda]+c+1}^{\infty} (k-c) \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) - c \sum_{k=c}^{[\lambda]+c} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right) + c \sum_{k=[\lambda]+c+1}^{\infty} \left(\frac{e^{-\lambda} \lambda^{(k-c)}}{(k-c)!} \right)$$

Adjust sums by c .

$$= -(\lambda + c) + 2(\lambda + c) \sum_{k=0}^{[\lambda]} \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) - \sum_{k=0}^{[\lambda]} k \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) + \sum_{k=[\lambda]+1}^{\infty} k \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) - c \sum_{k=0}^{[\lambda]} \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) + c \sum_{k=[\lambda]+1}^{\infty} \left(\frac{e^{-\lambda} \lambda^k}{k!} \right)$$

Combine terms.

$$= -(\lambda + c) + 2(\lambda + c) \sum_{k=0}^{[\lambda]} \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) + \lambda \left[1 - 2 \sum_{k=0}^{[\lambda]} \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) \right] + c \left[1 - 2 \sum_{k=0}^{[\lambda]} \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) \right]$$

Simplify.

$$= 2\lambda \left[\sum_{k=0}^{[\lambda]} \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) - \sum_{k=0}^{[\lambda]-1} \left(\frac{e^{-\lambda} \lambda^k}{k!} \right) \right]$$

Simplify.

$$= 2\lambda \left(\frac{e^{-\lambda} \lambda^{[\lambda]}}{[\lambda]!} \right)$$

Simplify.

$$= \frac{2e^{-\lambda} \lambda^{[\lambda]+1}}{[\lambda]!}$$

To calculate the expected value of the daily expected MAE across multiple days, sum this value across all days and divide by n .

$$E[E[MAE]] = \frac{1}{n} \sum_{d=1}^n \left(\frac{2e^{-\lambda_d} \lambda_d^{[\lambda_d]+1}}{[\lambda_d]!} \right)$$

References

- GRUEN R (2001) To Err Is Human. Eds. L. T. Kohn, J. M. Corrigan and M. S. Donaldson. National Academy Press Washington, D.C. 2000. Pp. 287. \$34.95. *Epidemiol Infect* 126(1):312
- Needleman J, Buerhaus P, Pankratz VS, Leibson CL, Stevens SR, Harris M (2011) Nurse Staffing and Inpatient Hospital Mortality. *N Engl J Med* 364(11):1037–1045
- McGillis Hall L, Doran D, Pink GH (2004) Nurse Staffing Models, Nursing Hours, and Patient Safety Outcomes. *J Nurs Adm* 34(1): 41–45
- Cochran JK, Broyles JR (2010) Developing nonlinear queuing regressions to increase emergency department patient safety: Approximating reneging with balking. *Comput Ind Eng* 59(3): 378–386
- Hoot NR, Zhou C, Jones I, Aronsky D (2007) Measuring and Forecasting Emergency Department Crowding in Real Time. *Ann Emerg Med* 49(6):747–755
- Zeng Z, Ma X, Hu Y, Li J, Bryant D (2012) A Simulation Study to Improve Quality of Care in the Emergency Department of a Community Hospital. *J Emerg Nurs* 38(4):322–328
- Peck JS, Benneyan JC, Nightingale DJ, Gaehde SA (2012) Predicting emergency department inpatient admissions to improve same-day patient flow. *Acad Emerg Med* 19(9):1045–1054
- Kadri F, Harrou F, Chaabane S, Tahon C (2014) Time series modeling and forecasting of emergency department overcrowding. *J Med Syst* 38(9):107
- Connelly LG, Bair AE (2004) Discrete event simulation of emergency department activity: A platform for system-level operations research. *Acad Emerg Med* 11(11):1177–1185
- Azari-Rad S, Yontef A, Aleman DM, Urbach DR (2014) A simulation model for perioperative process improvement. *Oper Res Heal Care* 3(1):22–30
- Wang B, Han X, Zhang X, Zhang S (2015) Predictive-reactive scheduling for single surgical suite subject to random emergency surgery. *J Comb Optim* 30(4):949–966

12. Fügener A, Hans EW, Kolisch R, Kortbeek N, Vanberkel PT (2014) Master surgery scheduling with consideration of multiple downstream units. *Eur J Oper Res* 239(1):227–236
13. Littig SJ, Isken MW (2007) Short term hospital occupancy prediction. *Health Care Manag Sci* 10(1):47–66
14. Kumar A, Jiao RJ, Shim SJ (2008) Predicting bed requirement for a hospital using regression models. In: 2008 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2008, pp. 665–669
15. Abraham G, Byrnes GB, Bain CA (2009) Short-term forecasting of emergency inpatient flow. *IEEE Trans Inf Technol Biomed* 13(3):380–388
16. Kortbeek N, Braaksma A, Smeenk FHF, Bakker PJM, Boucherie RJ (2015) Integral resource capacity planning for inpatient care services based on bed census predictions by hour. *J Oper Res Soc* 66(7):1061–1076
17. McManus ML, Long MC, Cooper A, Litvak E (2004) Queuing Theory Accurately Models the Need for Critical Care Resources. *Anesthesiology* 100(5):1271–1276
18. Verburg IWM, De Keizer NF, De Jonge E, Peek N (2014) Comparison of regression methods for modeling intensive care length of stay. *PLoS One* 9(10):e109684
19. Cahill W, Render M (1999) Dynamic simulation modeling of ICU bed availability. *WSC'99 1999 Winter Simul Conf Proceedings 'Simulation - A Bridge to Futur (Cat No99CH37038)* 2:1573–1576
20. Troy PM, Rosenberg L (2009) Using simulation to determine the need for ICU beds for surgery patients. *Surgery* 146(4):608–620
21. McConnell KJ, Richards CF, Daya M, Bernell SL, Weathers CC, Lowe RA (2005) Effect of increased ICU capacity on emergency department length of stay and ambulance diversion. *Ann Emerg Med* 45(5):471–478
22. Zhu Z, Hoon Hen B, Liang Teow K (2012) Estimating ICU bed capacity using discrete event simulation. *Int J Health Care Qual Assur* 25(2):134–144
23. Gholipour C, Rahim F, Fakhree A, Ziapour B (2015) Using an artificial neural networks (ANNS) model for prediction of intensive care unit (ICU) outcome and length of stay at hospital in traumatic patients. *J Clin Diagn Res* 9(4):19–23
24. Tierney LT, Conroy KM (2014) Optimal occupancy in the ICU: A literature review. *Aust Crit Care* 27(2):77–84
25. Marmor YN, Rohleder TR, Cook DJ, Huschka TR, Thompson JE (2013) Recovery bed planning in cardiovascular surgery: a simulation case study. *Health Care Manag Sci* 16(4):314–327
26. Rohleder TR, Lewkonja P, Bischak DP, Duffy P, Hendijani R (2011) Using simulation modeling to improve patient flow at an outpatient orthopedic clinic. *Health Care Manag Sci* 14(2):135–145
27. Pan C, Zhang D, Kon AWM, Wai CSL, Ang WB (2015) Patient flow improvement for an ophthalmic specialist outpatient clinic with aid of discrete event simulation and design of experiment. *Health Care Manag Sci* 18(2):137–155
28. Liu N, Ziya S, Kulkarni VG (2010) Dynamic Scheduling of Outpatient Appointments Under Patient No-Shows and Cancellations. *Manuf Serv Oper Manag* 12(2):347–364
29. Xie H, Chaussalet TJ, Millard PH (2005) A continuous time Markov model for the length of stay of elderly people in institutional long-term care. *J R Stat Soc Ser A Stat Soc* 168(1):51–61
30. Zhang Y, Puterman ML, Nelson M, Atkins D (2012) A Simulation Optimization Approach to Long-Term Care Capacity Planning A Simulation Optimization Approach to Long-Term Care Capacity Planning
31. Joy MP, Jones S (2005) Predicting bed demand in a hospital using neural networks and ARIMA models: a hybrid approach. In: *13th European Symposium on Artificial Neural Networks*, pp. 27–29
32. Xu Q, Tsui KL, Jiang W, Guo H (2016) A Hybrid Approach for Forecasting Patient Visits in Emergency Department. *Qual Reliab Eng Int* 32(8):2751–2759
33. Weiss EN, Cohen MA, Hershey JC (1982) An Iterative Estimation and Validation Procedure for Specification of Semi-Markov Models with Application to Hospital Patient Flow. *Oper Res* 30(6):1082–1104
34. Andersen AR, Nielsen BF, Reinhardt LB (2017) Optimization of hospital ward resources with patient relocation using Markov chain modeling. *Eur J Oper Res* 260(3):1152–1163
35. Pagel C et al (2017) Development, implementation and evaluation of a tool for forecasting short term demand for beds in an intensive care unit. *Oper Res Heal Care* 15:19–31
36. Green LV, Nguyen V (2001) Strategies for cutting hospital beds: The impact on patient service. *Health Serv Res* 36(2):421–442
37. Armony M, Israelit S, Mandelbaum A, Marmor YN, Tseytlin Y, Yom-Tov GB (2015) On Patient Flow in Hospitals: A Data-Based Queueing-Science Perspective. *Stoch Syst* 5(1):146–194
38. Harrison GW, Shafer A, Mackay M (2005) Modelling Variability in Hospital Bed Occupancy. *Health Care Manag Sci* 8(4):325–334
39. Steins K, Walther SM (2013) A generic simulation model for planning critical care resource requirements. *Anaesthesia* 68(11):1148–1155
40. Konrad R et al (2013) Modeling the impact of changing patient flow processes in an emergency department: Insights from a computer simulation study. *Oper Res Heal Care* 2(4):66–74
41. Choon OH, Dali Z, Beng PT, Magdalene CPY (2014) Uncovering effective process improvement strategies in an emergency department using discrete event simulation. *Heal Syst* 3(2):93–104
42. Dexter F, Macario A, Traub RD, Hopwood M, Lubarsky DA (1999) An operating room scheduling strategy to maximize the use of operating room block time: Computer simulation of patient scheduling and survey of patients' preferences for surgical waiting time. *Anesth Analg* 89(1):7–20
43. Blasak RE, Armel WS, Starks DW, Hayduk MC (2003) The use of simulation to evaluate hospital operations between the emergency department and a medical telemetry unit. *Proc 2003 Winter Simul Conf 2003* 2:1887–1893
44. Günel MM, Pidd M (2010) Discrete event simulation for performance modelling in health care: A review of the literature. *J Simul* 4(1):42–51
45. Day TE et al (2014) Addressing the variation of post-surgical inpatient census with computer simulation. *Pediatr Surg Int* 30(4):449–456
46. Hoot NR et al (2008) Forecasting Emergency Department Crowding : A Discrete Event Simulation. *Ann Emerg Med* 52(2):116–125
47. Resar R, Nolan K, Kaczynski D, Jensen K (2011) Using real-time demand capacity management to improve hospitalwide patient flow. *Jt Comm J Qual Patient Saf* 37(5):217–227
48. Marshall A, Vasilakis C, El-Darzi E (2005) Length of stay-based patient flow models: Recent developments and future directions. *Health Care Manag Sci* 8(3):213–220
49. Harrison GW, Escobar GJ (2010) Length of stay and imminent discharge probability distributions from multistage models: variation by diagnosis, severity of illness, and hospital. *Health Care Manag Sci* 13(3):268–279
50. Hoëffding W (1956) On the distribution of the number of success in independent Trials. *The Annals of Mathematical Statistics*, vol. 3. Institute of Statistical Science, Academia Sinica, pp. 295–312
51. Young JP (1965) Stabilisation of inpatient bed occupancy through control of admissions. *Hospitals (Lond)* 39(19):41–48
52. Graue RM (2013) Prediction and optimization techniques to streamline surgical scheduling
53. Hong Y (2013) On computing the distribution function for the Poisson binomial distribution. *Comput Stat Data Anal* 59:41–51