8-2019

Data-driven decision-support for process improvement through predictions of bed occupancy rates

Kar Way TAN
Singapore Management University, kwtan@smu.edu.sg

Qi You NG
Singapore Management University, qiyou.ng.2014@mais.smu.edu.sg

Francis Ngoc Hoang Long NGUYEN
Singapore Health Services

Sean Shao Wei LAM
Singapore Health Services

Citation

TAN, Kar Way; NG, Qi You; NGUYEN, Francis Ngoc Hoang Long; and LAM, Sean Shao Wei. Data-driven decision-support for process improvement through predictions of bed occupancy rates. (2019). *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE): Vancouver, BC, September 19: Proceedings*. 1-7.
Available at: https://ink.library.smu.edu.sg/sis_research/4684

Data-Driven Decision-Support for Process Improvement through Predictions of Bed Occupancy Rates

Kar Way Tan[1], Qi You Ng[1], Francis Ngoc Hoang Long Nguyen[2] and Sean Shao Wei Lam[3]

Abstract— **Managing bed utilization by ensuring the supply keeps up with demand is not an easy task in a large public hospital with many medical disciplines. The bed managers who make decisions on reserving and allocating beds centrally require high-dimensional data from several hospital information systems supporting processes in emergency room, specialized clinics and wards. In this work, we put together an automated process for cleaning, consolidating and integrating data from several information systems into reports required by the bed managers to analyse the bed occupancy situations across more than thirty medical disciplines. To prevent bed crunch situation where patients wait more than ten hours for beds, we developed two predictive models based on Principal Component Analysis and Multiple Linear Regression to provide hospital with the foreknowledge of the bed occupancy situations. Our aim is to move the hospital from reactive to proactive bed management. Our data-driven solution focuses on consistent and accurate reporting, content co-creation with the bed managers, and an approach which enables seamless transition towards proactive decision-making. The solution is implemented in a large public hospital in Asia. It provides high value to stakeholders in hospital by reducing the time-taken to get the information on hand for decision-making from at least four days to half a day. The selected prediction model for bed occupancy rate has achieved 80% accuracy (at an error tolerance of 5%). We hope our results will encourage and benefit hospitals with similar settings to adopt data-driven methods to tackle high bed occupancy situations in their premises.**

I. INTRODUCTION

Bed crunch is a situation to avoid in public hospitals around the world. In this study, we based our study on a large comprehensive public hospital in Asia, we referred to as Study Hospital (SH). Bed crunch is defined as the situation where at least one patient waits more than ten hours for a bed on that day. The local healthcare regulatory agency requires all public hospitals to report weekly, the number of patients who waited for more than 10 hours for a bed and the number of days that the hospital experienced bed crunch. Bed crunch is an important performance indicator tracked across the hospitals. SH faces the challenge of increasing occurrences of bed crunch situations. In 2013 and 2014, the occurrences of bed crunch at SH was below 8% but rose to about 15% in 2015. The situation was made worse by the increasing healthcare demands from aging population, a demographic challenge in many countries in Asia [1].

During the study, there was an incomplete overarching perspective of the bed situation due to lack of integration among disparate data sources, and the absence of a unified dashboard to display the key process indicators that can influence current and future bed occupancy rates. The Bed Management Department (BMD) of SH relied on a trending report which required data from three different sources and a period of four days to a week to consolidate the information. The resulting report was also susceptible to human errors. In general, SH receives inpatients from three major sources: Emergency Department (ED), Specialist Outpatient Clinics (SOCs), and Admissions Office (AO). Data consolidation across the three sources of admissions was not available as the hospital handles more than thirty different medical disciplines. Although data can be derived from various existing databases and the enterprise data warehouse, significant resources and time investment were required for the consolidation of data, tabulation of information and preparation of periodic historical reports. There were also lack of accurate forecast or predictive information for bed demands and bed occupancy situations.

Given the situation, we see a need to move towards proactive processes where hospital is to be equipped with accurate reporting of the bed situations and prediction of the bed occupancy rate to effectively tackle bed crunch. We worked with BMD to build a data-driven solution with two goals – firstly to be well-informed of the bed occupancy situation; and secondly to make better bed assignment decisions using a bed occupancy prediction model. We propose a two-phase approach where phase one includes process improvement tools in handling the data from three different hospital information systems to support regulatory reporting needs on bed situation; and phase two includes development of data-driven prediction models with forecast and key predictive variables for bed occupancy.

With these predictive indicators, the various hospital decision makers can make evidence-based decisions to have significant influence on the bed situation. This is depicted in Fig. 1. For example, the BMD can adjust the short-term ring-fenced beds for emergency and elective operation patients before bed crunch occurs. The key BOR predictors revealed in this study can be included in the dashboard and reporting system to be developed for the BMD for a better understanding and foreknowledge of any impending bed crunch scenario.

[1]Kar Way Tan and Qi You Ng are with School of Information Systems, Singapore Management University, kwtan@smu.edu.sg, qiyou.ng.2014@mais.smu.edu.sg

[2]Ngoc Hoang Long Nguyen is with the Health Services Research Centre at Singapore Health Services, nguyen.ngoc.hoang.long@singhealth.com.sg

[3]Sean Shao Wei Lam is with Lee Kong Chian School of Business, Singapore Management University, Health Services and Systems Research, Duke NUS Medical School and Health Services Research Centre at Singapore Health Services, seanlam@smu.edu.sg, gmslasws@nus.edu.sg

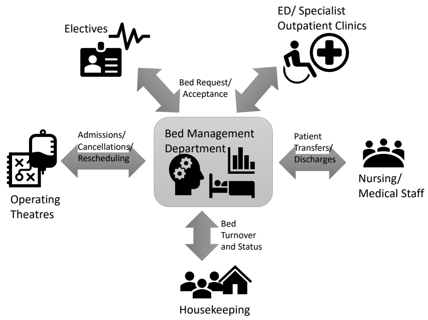

Fig. 1. Hospital stakeholders that can influence the bed occupancy situation in the study hospital

Our contribution is twofold. Firstly, we share an approach for a consistent data-driven method that combined data from heterogeneous sources for predicting bed utilization to support decision making. This approach enables practical and seamless transition to new solution for the bed management department. With an automated method, the process duration was cut from four to seven days to less than half a day. This means that the hospital can now make timely decisions on the daily basis based on more relevant and accurate information in the near-term to prevent bed crunch situations. Secondly, we provided two predictive models (and the comparisons among the models) for predicting bed occupancy rates (BOR) and showed how we can extend the prediction to a week. Our prediction models also allow us to uncover insights into the significant variables which affect bed occupancy rates, hence able to take proactive actions against bed crunch.

II. Literature Review

Forecasting hospital bed occupancy helps hospitals to efficiently manage their beds, improve resource utilization, efficiencies and rate of mortality. Drawing on the domain knowledge from bed managers whom we worked with, foreknowledge of bed occupancy can lead to high impact in preventing bed crunch by implementing short-term adjustments to the use of ring-fenced beds for emergency and elective operation patients [2].

From the perspective of providing sufficient supply of beds, Farmer et al. [3] used regression analysis with time series data to forecast the mean duration of stay or Length of Stay (LOS) to predict the number of surgical beds required in a hospital. They concluded that a joint estimation of mean duration of stay and admission rate is likely to provide in a more accurate forecast of beds. For a specific ward, Zernikow et al. [4] used multiple linear regression (MLR) and neural networks to predict LOS in preterm neonates.

From the perspective of matching demand with supply of bed, we reviewed several literature on bed occupancy and patient flow. Littig and Isken [5] developed an occupancy prediction model to predict short-term occupancy for a mid-size community hospital for each nursing unit and a specified horizon. The analysis includes stepwise regression and analysis of patient occupancy flow. This approach has its limitations in a large hospital where uplodging and overflowing across medical disciplines are possible. The other bed occupancy analysis include work by Jones et al. [6], who used ARIMA forecast and found that the daily number of occupied beds due to ED admissions is related to air temperature and influenza illness rate. This model focuses on the admissions contributed by a single source, i.e., ED admission. Scout and Tawney [7] used an Excel-based solution for evaluating bed occupancy level. Spreadsheet-based solution has limitations on the size and dimensions of the datasets. Zhu [8] and [9] created prediction procedures based on Discrete Event Simulation (DES) to predict Bed Occupancy Rate (BOR) and bed availability respectively. DES methods require collection of large number of parameters for modeling and often lead to higher computation cost.

In our research, we draw inspirations from the hospitality industry where Principal Component Analysis (PCA) was used to investigating patterns in hotel bed occupancy [10]. We saw the opportunity for applying PCA as a dimension-reduction method to correlate high dimensional data from several information systems into manageable set of parameters suitable for predicting bed occupancy at the hospital. We then combine the forecasted information with the PCA prediction model for predicting 7-day bed crunch indicators for BMD. Our solution is compared with a model based on MLR, which many existing studies have shown promising results.

III. Case Scenario

A. Scope

During the time of the study, the SH had about 2,000 licensed beds serving more than 30 medical disciplines. In 2014, SH had approximately 85 thousand admissions to its licensed beds. Distribution by admission sources were 40% elective patients from AO and SOCs and, 60% of admissions from ED. The admission-attendance ratio from ED was approximately 37.5%. There were more than 10 types of beds in SH. The general wards Class A, B1, B2, and C made up to about 80% of the beds across disciplines. The remaining 20% of the beds were specialized beds such as those in Intensive Care Unit (ICU), High Dependency (HD) and Infectious Disease (ID) wards. We focus our work on the *general wards* because the specialized wards followed different bed allocation policy for minority of the patients requiring critical and/or specialized medical care.

SH was required to report its daily Bed Occupancy Rate (BOR) and other key indicators to the local healthcare regulatory agency on a weekly basis. Bed occupancy rate is a common indicator of resource utilization in hospitals. BOR is calculated by dividing the total number of inpatient beds occupied by the total number of beds in service. SH experienced high occupancy rates ranging between 75% to 99% on their licensed beds in the yearly trend except for a few exceptions such as festive holidays that lowered the rate of AO admissions. On the weekly basis, bed occupancy rates varied 10-20% between the lowest and highest rates.

The fluctuations posed resource optimization challenges that SH grappled with. Fluctuating occupancy would invariably impact the previously planned manpower decisions (doctors and nurses rostering), reservations on the ring-fenced beds, and the demand for ancillary services such as housekeeping, laboratory, pharmacy, radiology, surgical and therapy services. It also impacted wait time in other medical units because the ancillary services were shared across disciplines and functions within the hospital. Before this study, bed managers relied on the trending report generated monthly to augment their bed allocation decisions. The bed allocation and planning were done in advanced and were unable to react to the fluctuating demand situations. With this study, by having the ability to predict the BOR and other key indicators, a variety of short-term operational measures can then be considered. Some of the measures include the management of inflowing patients (from ED, SOC and surgical elective patients), collaborating with step-down care providers for low-risk patients requiring extended medical care, and managing outflowing patients with active discharge policies.

B. Data

The sources of data we used for BMD's reporting to support bed occupancy analysis were based on data from Bed Management System (BMS), SH Enterprise Data Warehouse (SEDW) and the Accident & Emergency Department Database system (EDDB) over a 12-month study period in 2014, as shown in Table I.

TABLE I
DATASETS OBTAINED FROM THE 3 HOSPITAL INFORMATION SYSTEMS

No.	Name of Dataset	System
1.	Admissions	SEDW
2.	Bed Acceptance	BMS
3.	Bed Occupancy	BMS
4.	Bed Turnover	BMS
5.	Patient Length-of-Stay	BMS
6.	(Patient) Overflow	BMS
7.	Response Time to Bed Request	BMS
8.	Stepdown Transfers	BMS
9.	Waiting Time of Patients	BMS & EDDB

The data collected from the BMS, SEDW and EDDB were of different time intervals – daily, weekly and monthly. For the daily dataset, there were 30 Excel sheets with each day's records which were to be consolidated into 1 output files with 30 worksheets. The task was extremely labor-intensive and error-prone. These files were typically consolidated by a non-IT-trained executive using Excel spreadsheets.

IV. Approach

We present our two-phase approach as shown in Fig. 2. The approach includes a business process improvement automation for SH in Phase 1 and development of BOR predictive model in Phase 2. The objective is to allow SH to move from a reactive mode in response to bed crunch to a proactive process where action can be taken prior to bed crunch.

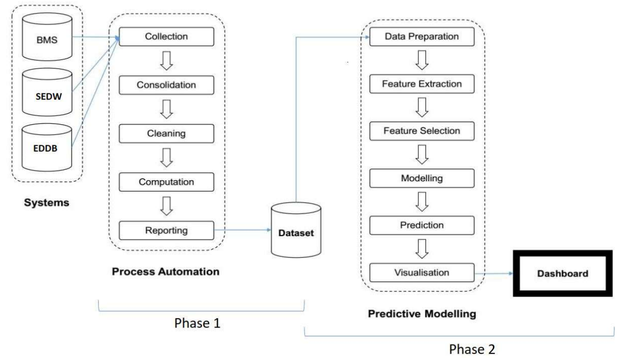

Fig. 2. Two-phase approach to proactive management of bed crunch

A. Phase 1: Process Improvement For Evidence-based Bed Occupancy Analysis and Reporting

Phase 1 of this work focuses on improving the process for evidence-based bed occupancy analysis and reporting. We propose an analytic solution based a 5-step automation process for coordinating heterogenous datasets from the various systems (see Fig. 2 Phase 1). Firstly, the data from hospital information systems was exported into files by scheduled tasks. The data in the files were then combined for data cleaning. Next, summary statistics such as such central tendency metrics were computed and store in a database table. Charts and reporting dashboard were built based on the cleaned data and the computed statistical metrics. The automated improved process eliminated the risk of human error, reduced manpower and cut down the cycle time from four days to half a day.

In order to ensure successful delivery and practical adoption of the solution by the hospital users, we put in place change management strategies to ensure the solution was aligned with the needs. Our process improvement strategies followed these criteria:

1) Co-create the solution with BMD to ensure continuity and adoption
2) Consistent reporting which satisfy the needs of BMD and regulatory agency
3) Maintain business-as-usual, without changes to IT systems for seamless transition towards proactive process

1) Co-Creation: While deciding the solution approach for consolidation, data cleaning, computation of statistical measures and reporting, we focused on adoption and sustainability, i.e., the users must be comfortable in using and the continual development of the solution. The collection part of the process was kept the same, extracting the data from 3 different systems and nearly transparent to the users. It was also important that the solution was co-created with the users, giving them sense of ownership, be sufficiently comfortable to take control, to make changes and to use it on the day-to-day basis without help of the IT department. As the users preferred a bespoke solution for transparent analysis over a black box solution without the ability to customize, R was selected due to its statistical computation and graphics capabilities. The BMD users were relatively comfortable with the language due to its gentle learning curve and we worked out a change management process to train the users to manage the automated solution based on R.

2) Consistency: Outputs exported from the 3 information systems (BMS, SEDW, EDDB) were of different formats including some legacy ones. The reporting periods were not consistent either, some were daily, some were weekly, and some were monthly. One of our consolidation efforts required us to combine datasets and to format the data to a consistent one using our new solution. At the calculation step, we carefully designed the right statistical measures for each of the reported measures. For example, average was used for discharges and median for understanding the wait-time. This provided consistency in reporting and a common understanding across the hospital decision makers and healthcare authority reporting requirements.

3) Business-as-Usual: Next, we designed the improvement process such that both the originating dataset and the content of the final output to the healthcare authority and hospital's management were the same. There was no change to the hospital information systems required, while the visuals were improved. Readers of the report who were comfortable with the as-is report format need not adjust to a new reporting format. With these changes, BMD transitioned to the new solution with ease and adopted the improvements readily.

B. Phase 1 Process Improvement Results

The first improvement between the to-be and the as-is processes was the significant number of automation tasks. The tasks from data consolidation to reporting (as per Phase 1 in Fig. 2) were automated with R programs.

1) Result 1: Improved Cycle Time for Report Generation: Time taken to generate a report had dropped from days to hours. Before automation, it took between 4 to 7 days. With the new solution, it took only between 2 to 4 hours to generate the necessary reports.

2) Result 2: Improved Accuracy of Key Performance Indicators: With automation, reports were generated with less (human) errors. Based on the 2014 dataset that BMD provided, we generated results for 12 months and compared them with the original manually consolidated numbers. There were about 20% discrepancies found. Fixes were made to the reports to correct the discrepancies. This was a great positive impact on the operation of BMD.

As part of the process automation, we provided better visuals which allowed the same set of data to obtain more insights. Fig. 3 shows an example of this improvement in which we have included multivariate data in a chart to keep track of the two important key performance indicators (KPIs), i.e., the number of patients who waited for 4 hours and 10 hours.

3) Result 3: Discovery of Hidden Insights: With the analysis capability of improved visuals, we gathered some counter-intuitive insights for increasing BMD's awareness

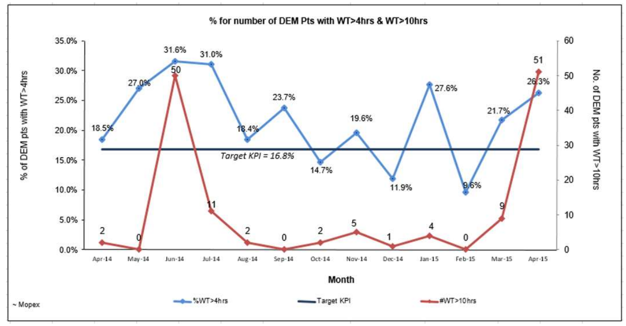

Fig. 3. Keeping track of the 4-hour and 10-hour wait time KPI

of the situation on the ground. The common intuition was that higher admission numbers will result in longer waiting time. However, when we plotted admission numbers from Department of Emergency Medicine (DEM) and waiting time, we found that days with admission numbers lower than the expected value may end up with longer waiting time for a bed ("long wait" was defined as having waited for more than 4 hours) and vice versa. We deployed the new visuals for subsequent year in 2015 and plotted the chart presented in Fig. 4. The height of the bubble indicates the number of admissions, the size of the bubble indicates the number of patients who had waited for more than 4 hours, and the colour intensity of the bubble indicates the number of patients who had waited more than 10 hours. For example, on 20th July, the admissions were higher than average, the bubble was small and shade was light, indicating low number of patients who had waited for more than 4 hours and 10 hours respectively. This day, the admission number was high but few patients waited more than 4 hours and 10 hours. These results were highlighted to the SH for better understanding of other factors that could potentially affect the waiting times for beds. Through further analysis, other factors that were previously not considered, e.g., number of longstayers and total discharges were being added into considerations for BOR prediction.

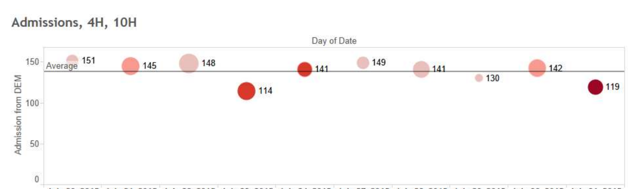

Fig. 4. Wait time and number of admissions after deployment in year 2015

4) Result 4: Generated Inputs for Phase 2 Predictive Model: The output from the Phase 1 automation captured a set of measures and KPIs which were then served as the inputs for the predictive modeling in Phase 2, including additional data such as long-stayers and discharges. The list of datasets used for Phase 2 are shown in Table II.

TABLE II

DESCRIPTIVE ANALYTICS OUTPUT

No.	Name of Dataset	Reporting Frequency
1	Patients waiting in ED for 4h/10h	Weekly
2	Admissions and Discharges	Monthly
3	Bed Occupancy	Monthly
4	Bed Turnover (BTO)	Monthly
5	Longstayers	Monthly
6	Overflow at 7am census	Monthly
7	Overflow at Point of Admission	Daily
8	Pending discharges	Monthly
9	Response Time and Bed Acceptance	Monthly
10	Transfer Stepdown	Monthly
11	Waiting Time of patients	Monthly

C. Phase 2: Predictive Model for BOR

In the second phase, two predictive models were proposed to provide prediction of BOR, one based on Principal Component Analysis; and another based on Multiple Linear Regression model. For building the predictive models, a 6-step approach was used (see Fig. 2 Phase 2).

D. Data Preparation

We first applied near-zero variance analysis and pairwise correlation analysis to discover variables with low volatility and high possibility of collinearity among the variables. We started off with a time series dataset with 205 variables. The variables contained attendances and admissions of Emergency Department (ED), admissions and discharges of the various wards and long-stayers in the wards across the thirty medical disciplines. Coupled with domain knowledge and the near-zero variance analysis, we removed 40 variables. The removed variables either had too many zero-value records or had little variances in the values. Exercising caution with domain knowledge, we noticed examples of these variables were attendances and admissions of patients with very low acuity at the ED or cases from specialist clinics with very low admission rates. Next, we used the Alias and Variance Inflation Factor functions in R to detect and remove variables with high collinearity. We further found another 11 variables related to admissions, discharges and netflow in specialized clinics to be highly correlated and removed 11 more variables. Finally, split our dataset into train (70%) and test (30%) sets for model building and validation by random sampling. In total, the training set had 254 observation points and test set had 109 observation points across 154 dimensions.

Next, we performed feature extraction and modeling. Let T be the prediction day. In actual deployment of the analytic model, T refers to the next operation day. We used the lagged variables of T–1 for predicting the BOR of T.

E. Feature Extraction, Selection and Modeling using Principal Component Analysis

In Feature Extraction step, we took the pre-processed dataset and applied Principal Component Analysis (PCA), a popular dimensionality reduction method. Using BOR as the dependent variable, we used the predict() function on the pre-processed data and the trained dataset to arrive at the trained principal components (PCs). Our PCA resulted in 10 PCs.

In Feature Selection step, we took the principal components (PCs) and applied scree plot and summary statistics and then checked the standard deviation associated with each PC, the proportion of the variance explained and the cumulative proportion of explained variance by each component. Next, we computed the correlation strengths against the Bed Occupancy Rate to identify the significant PCs. When we ranked the significance of the PCs in descending order (i.e., the first PC being the most significant), we found that only 2 PCs accounted for about 88% of the variance and the top 7 PCs accounted for more than 95% of the variance of the data. Hence, to form our prediction model, we selected the 2 most significant PCs (PC_1 and PC_2) to be part of the model.

Due to sensitivity of the healthcare data, we shall broadly describe the variables in the PCA loading of PC_1 and PC_2. The variables which were affecting BOR were indicators of ED admissions of patients from various acuity levels 1 to 3 (there were 4 acuity levels with 1 being the most severe), ED attendances of patients with various acuity levels 1 to 3, overall elective admissions, admissions in B2 and C wards, and the net discharges from various disciplinary wards. We proceeded with a PCA model by using only 2 significant PCs. Our PCA model for predicting BOR is $BOR = PC_1 + PC_2 + \epsilon$ where ϵ is the error term. The R^2 value of the PCA model was found to be 0.53.

F. Feature Extraction, Selection and Modeling using Multiple Linear Regression

We built a second model based on Multiple Linear Regression (MLR). The MLR model started with Feature Selection step after data preparation. We performed stepwise Multiple Linear Regression (at 0.05 significance level) on the dataset to identify the significant variables and used them as predictors in the MLR model. We discovered that significant variables contributing towards the variances in BOR were indicators related to attendances and admissions of patients with acuity level 2 from the ED, discharges from a few specific disciplinary wards such as Endocrinology and Neurology, number of long-stayers from several types of wards such as Dermatology and Hepato-pancreato-biliary. We fitted the significant variables into the MLR model and the resulting model provided a R^2 value of 0.4152.

G. Results Comparison Between PCA and MLR

Both the PCA and MLR models provided insights to the variables contributing towards BOR. In both models, ED attendances and admissions (in particular patients with acuity 2) were part of the significant variables or key loading variables in the significant principal components. As part of our evaluation of which prediction model to use, we compared the predictive capabilities of both the models. Both PCA and MLR predictive models were validated with the 30% test data. We considered two types of results analysis, *collective* and *individual* prediction performance. The objective was not for one bad prediction to affect the overall performance of the model by considering collective performance only.

In collective prediction performance analysis, we considered all predictions made by a model. We summed the deviations and averaged the values by the number of predictions to evaluate the performance of the predictions against the actuals. In this case, we used Mean Absolute Deviation (MAD) as the measure. In individual prediction performance analysis, we considered each prediction deviation by itself and grouped them into bin based on the deviation from actual. Each bin represented a percentage of deviation from the actual. By aggregating the number of predictions based on their deviations from the actuals, we computed the proportion of good predictions over all predictions. In essence, we computed the ratio of correct predictions to the total number of predictions over different ranges of deviations. In this analysis, we set the percentage of deviation to 5% and used the term *percentage of accuracy* to mean the percentage of predictions which deviated from the actual for less than or equal to 5%. The results of PCA and MLR models are shown in Table III. The evaluations based on MAD and individual deviations showed that PCA model provides more promising BOR predictions compared to MLR model. For illustration, we have included in Fig. 5 and Fig. 6, showing the scatter plots of the deviations between the individual predicted and actual BOR for the PCA and MLR models respectively.

TABLE III
COMPARISON BETWEEN PCA AND MLR MODELS

Description	**PCA**	**MLR**
No of predictors in model	2	14
R^2	0.531	0.4152
Mean Absolute Deviation (MAD)	3.2%	4.8%
Percentage of Accuracy	80.7%	60.5%

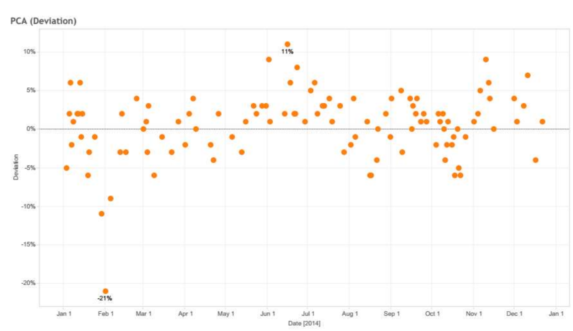

Fig. 5. Scatter plot of the Predicted vs Actual BOR based on PCA model

H. Extended Prediction Model and Visualization for Deployment

Using the PCA prediction model, we developed a 7-day BOR prediction model. As our prediction for day T is based on variables of T–1, the 7-day prediction model was built based on time-series forecast of the significant variables in the chosen principal components for the subsequent days. Based on BMD's requirements, we derived a simple colored

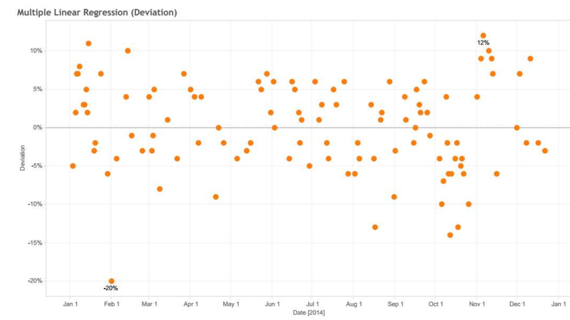

Fig. 6. Scatter plot of the Predicted vs Actual BOR based on MLR model

bed occupancy indicators for indicating potential bed crunch situation. The Bed Occupancy Indicator displays red if the predicted BOR is 95% or more, yellow if between 82% (inclusive) and 95% (not inclusive) and green if below 82%.

With the forecast of key contributing variables and the predictive bed occupancy indicators, we put together an interactive dashboards for our bed managers at BMD as shown in Fig. 7. The dashboards can be easily customizable for other departments and needs (e.g., comparison of the previous week's data with the lookahead 7-day BOR predictions). Another sample dashboard with previous week's and lookahead indicators is shown in Fig. 8. BMD and other stakeholders can now proactively make better decisions on managing the beds by using the dashboards.

Fig. 7. 7-day BOR indicators and time series forecast on few key parameters in a single dashboard

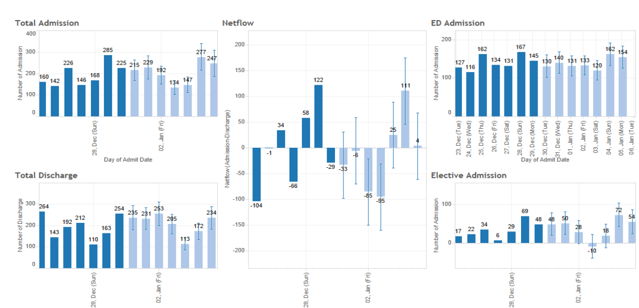

Fig. 8. 14-day BOR indicators and time series forecast on few key parameters in a single dashboard

We understand the limitations of the bed indicator predictions for days $T+1$ to $T+6$ are based on forecasted values and may affect the accuracy of the predictions. The dashboard is a rolling horizon, meaning the bed occupancy for $T+1$ will be predicted again using the actual values on day T, adjusting the predicted values on continuous basis. We will also in our future work, look into analysis and predictions of the key contributing variables which contributed towards the prediction of BOR.

V. Conclusion

Bed crunch is an important problem plaguing many major public healthcare systems which hospitals need to take proactive action against it. However, the hospital information systems with diversity of data required laborious efforts to put together holistic view of the bed situations for accurate and timely decision making. Making decisions on bed allocation is a fundamentally challenging problem dependent on the decisions of multiple stakeholders operating in different departments. In this work, we shared our experiences and strategies for improving processes in the bed management and to build an evidence-based data analytics solution for supporting bed occupancy analysis. Our solution used not just a single dataset, but high-dimensional heterogenous datasets from three hospital information systems. The study enabled SH to gain insights into the factors with high impact on the bed occupancy in the hospital. Our automated analytic solution reduced the cycle time from at least four days to less than half a day, and empowered bed management decision-makers with forecast of key indicators and prediction of bed occupancy rate for better decision-making. The solution allows preventive actions to be taken before an occurrence of bed crunch. Our approach to process improvement and prediction modeling was designed with consistency in mind, co-created with the bed management users and had led to positive adoption.

Moving forward, to improve the predictions of bed occupancy rates, we will focus on building extended prediction models to include the predictions of the attendances and admissions coming from the related medical and surgical departments. We hope our experiences will encourage and also benefit the other hospitals to adopt data-driven methods for bed management. The decision-support system can make use of updated knowledge of the current and imminent bed occupancy situation to make better decisions.

Acknowledgment

The authors would like to thank the support of staff members in the Bed Management Department of the study hospital.

References

[1] Asian Development Bank, "Population and aging, asia and the pacific datasets," 2017, https://data.adb.org/dataset/population-and-aging-asia-and-pacific, Last accessed on 2018-08-04.

[2] P. Dimitriadis, S. Iyer, and E. Evgeniou, "The challenge of cancellations on the day of surgery," *International Journal of Surgery*, vol. 11, no. 10, pp. 1126 – 1130, 2013.

[3] R. D. T. Farmer and J. Emami, "Models for forecasting hospital bed requirements in the acute sector," *Journal of Epidemiology and Community Health (1979-)*, vol. 44, no. 4, pp. 307–312, 1990.

[4] B. Zernikow, K. Holtmannsptter, E. Michel, F. Hornschuh, K. Groote, and K.-H. Hennecke, "Predicting length-of-stay in preterm neonates," *European Journal of Pediatrics*, vol. 158, no. 1, pp. 59–62, 1999.

[5] S. J. Littig and M. W. Isken, "Short term hospital occupancy prediction," *Health Care Management Science*, vol. 10, no. 1, pp. 47–66, 2007.

[6] S. A. Jones, M. P. Joy, and J. Pearson, "Forecasting demand of emergency care," *Healthcare Management Science*, vol. 5, no. 4, pp. 297–305, 2002.

[7] W. A. Stout and B. Tawney, "An excel forecasting model to aid in decision making that affects hospital resource/bed utilization - hospital capability to admit emergency room patients," in *2005 IEEE Design Symposium, Systems and Information Engineering*, April 2005, pp. 222–228.

[8] Z. Zhecheng, "An online short-term bed occupancy rate prediction procedure based on discrete event simulation," *Journal of Hospital Administration*, vol. 3, no. 4, pp. 37–42, 2014.

[9] W. Cahill and M. Render, "Dynamic simulation modeling of icu bed availability," in *Simulation Conference Proceedings, 1999 Winter*, vol. 2, 1999, pp. 1573–1576.

[10] S. D. Cantis, M. Ferrante, and F. Vaccina, "Seasonal pattern and amplitude a logical framework to analyse seasonality in tourism: An application to bed occupancy in sicilian hotels," *Tourism Economics*, vol. 17, no. 3, pp. 655–675, 2011.