Deep Reinforcement Learning Based Cost-Benefit Analysis for Hospital Capacity Planning

Anonymous
Anonymous
Anonymous
Anonymous, Anonymous, Anonymous
anonymous@anonymous.com

***Abstract*—This work proposes a deep reinforcement learning (RL) based model to devise a hospital augmentation plan for a particular region. It works on the cost-benefit over a range of geographic regions and proposes the best place to set up or augment hospital capacity. The stochastic nature of hospital bed demands makes it challenging to devise an appropriate augmentation scheme. We project hospital bed numbers as a capacity determiner and consider populations of different walks to analyze future demand economics. A particular geographic region is divided into several sub-regions based on the local administration body. The RL agent works on the age group, population growth, and current bed capacity to propose a sub-region where augmentation is necessary. We utilize the Advantage Actor-Critic (A2C) algorithm to minimize the cumulative cost. The simulation result with actual data testifies this approach's superiority over traditional per capita based and complain based policies.**

***Index Terms*—Hospital bed capacity, reinforcement learning, deep RL, markov decision process, agent based modeling.**

I. Introduction

The per capita hospital bed capacity of a geographic region can give us the primary impression of that region's healthcare capability. However, the number of people staying overnight in hospitals in the region is highly stochastic [1], [2]. Insufficient hospital capacity can cause overcrowding and deprived critical patients. Opening new hospitals or expanding the existing hospital capacity is so expensive that it is not feasible for the authority to maintain a large enough capacity to serve any number of patients. Hence, the health authority of a region need to perform a cost-benefit analysis to determine its expansion plans. The health authority often faces budgetary constraints for capacity expansion and should use an appropriate policy to prioritize the expansion needs [3]. Throughout the paper, we use the hospital augmentation plan to refer to opening new hospitals or expanding the existing hospital capacity.

Hospital bed capacity is subject to regular monitoring to support an area's growing healthcare needs. Many states in the US implemented a mechanism called Certificate of Needs (CON) to regulate the number of hospital beds in the respective state [3]. The CON mechanism depends on the projected population growth to determine the future bed demand, which lacks the consideration of different age groups, thus it is too general for such a complex and dynamic problem. A simple augmentation plan based on some flat number of target occupancy is strictly suboptimum due to the unaccounted variables. For instance, the large variability in patients' age in different hospital wards (i.e., subdivisions) makes it even harder for the authority to come up with a detailed augmentation plan [4], [5].

In this work, we consider seven counties of Florida, which is an attractive retirement home for an increasing number of elderly people. High population growth and the nationwide increase of median age [6] indicate more requirement for healthcare facilities. Besides, elderly people are supposed to stay longer in hospitals than younger people do. Other than this steady demand, there are no-notice incidents such as disease outbreaks, epidemics, increased drug abuse, which can cause a sudden surge in hospital bed demand. Such incidents can lead to unprecedented loss of life and wealth if the hospital preparedness measures are not adequate, efficient, and effective. Recent mishaps for COVID-19 have testified that existing prediction mechanisms are insufficient to tackle such demands. In addition to serving for the healthcare of a community, hospitals are a major revenue as well as employment provider. Improper decisions in the augmentation plan can bring havoc to a region's economy. Therefore, hospitals or healthcare authorities need appropriate methods to improve their forecasting to tackle such hazards in the future. A robust, dynamic, and detailed hospital augmentation plan can benefit both the government and private parties.

The Markov Decision Process (MDP) models are popular for sequential decision-making in complex systems [7]. Following a divide-and-conquer approach MDP implements an optimal policy by sequentially taking an optimum action at each state. This work formulates the hospital capacity planning problem as an MDP by considering the existing capacity, such as hospital beds in several regions, population in several age groups, and population growth. Due to the high-dimensional state space, we utilize a Deep Reinforcement Learning (RL) algorithm, Advantage Actor-Critic (A2C) [8], to learn the optimal policy for the MDP formulation [9]. Healthcare authority is the MDP agent that takes the hospital expansion actions through A2C. At each time step, its decision modifies the environment and incurs a cost. The main contributions of this work are:

- A novel MDP model to define the optimal policy for hospital capacity expansion.
- A deep RL algorithm that determines the optimal policy of the MDP to minimize the cumulative cost for a finite

time horizon.

- A comprehensive case study for the Tampa Bay area using real data to exhibit the advantage of the proposed policy over traditional myopic policies.

The remainder of the paper is organized as follows. Section II reviews related literature for hospital capacity planning. The MDP model is formulated in Section III, and the deep RL algorithm for the optimal policy is given in Section IV. The case study for the Tampa Bay area is presented in Section V. The proposed model and solution are discussed in Section VI. Finally, the paper is concluded in Section VII.

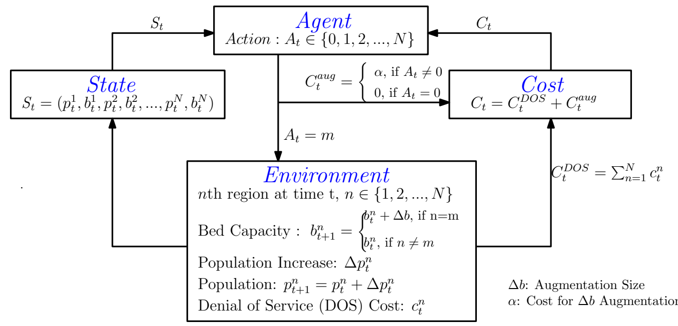

Fig. 1. Proposed MDP model.

II. Related Work

Several studies have been conducted considering a mathematical framework for determining the optimal number of beds within hospitals and surrounding regions [10], [11]. The hospital bed occupancy literature can be divided into two main groups. The first includes works related to forecasting hospital bed demands across different departments, Emergency Department visits, and other critical healthcare resources. These studies vary by time horizon (i.e., 1 hour to 8 hours after hospital admission), treatment type (i.e., critical care, scheduled treatment), population setting (i.e., age range, Medicare, and Medicaid), provider settings (i.e., primary care, secondary, and hospital-specific), and source of data (i.e., administrative claims data, real-time data, and clinical data) [12]–[18]. However, the main objective of these studies is to support the existing hospital critical resource allocation schemes, rather than long-term planning and bed expansion. In this paper, we are interested in the second group, which focuses on the long-term capacity planning and allocation of the hospital beds with a region. Most of the methods in this group are intended for medium- or long-term estimates at different regional settings. These models include the simple ratio method, Michigan's Bed Need model, Formula method, The Swiss Health Observatory (SHO) model, and Current Use Projection Model [11], [19], [20]. Some of the challenges investigated in the second group include low accuracy, overestimation of required bed numbers, and difficulty of demographic predictions [10]. The uncertainty and variability in estimating hospital occupancy levels due to the lack of a sophisticated prediction model might make capacity planning and hospital bed extension challenging in a region with a rapidly changing population [2].

More recently, a few studies presented promising data-driven approaches in forecasting bed occupancy in hospitals and surrounding regions [19], [21]. Most of these studies thus far are limited to forecasting critical care bed occupancy using machine learning (ML), in particular regression, methods. Among ML methods, several variants of neural networks have been utilized in predicting ICU bed occupancy, total hospital bed occupancy, and surgical operating room forecasting [20], [22]. However, ML-based hospital bed capacity planning and expansion forecast at the regional level has not been yet properly investigated. Only one study attempted to use an RL-based approach to provide hospital bed augmentation policy for Bangladesh, but without providing evidence to validate the proposed model [23]. RL methods have recently achieved significant theoretical and technical achievements, leading to a widespread application including gaming, finance, transportation, and healthcare [24]. In the healthcare domain, the application of RL has been focused on diagnosing conditions or forecasting outcomes, but not explicitly on healthcare policy-making [25]. RL empowers decision-making capabilities by generating agent-based scenarios through an MDP framework for the optimization of critical resources [26]. Hence, our study aims to develop a deep RL-based hospital capacity planning policy to minimize the total cost within a finite time horizon by dynamically providing the optimal number of beds in a geographical region.

III. MDP Model

MDP formulation is based on the Markov Property, which suggests that the future state is dependent only on the current state and action taken by the agent. In our MDP model shown in Fig. 1, healthcare authority is the MDP agent, and all the N regions (e.g., counties) served by it forms the environment. The state at time t, S_t is defined by the population p_t^n and the existing hospital bed capacity b_t^n of each region, where $n \in \{1, 2, ..., N\}$. At each time t, the agent takes action $A_t = m$, which refers to expanding the capacity of mth region out of the N regions, and receives cost C_t. Furthermore, it can also decide to do no expansion ($A_t = 0$), resulting in a total of $N + 1$ choices. The next state S_{t+1} and cost C_t are function of current state S_t and action A_t, satisfying the Markov property. We next explain the MDP model in detail.

A. State, S_t

Each region's current population and hospital bed capacity collectively defines the state space. The population of the nth region at time t, $p_t^n = [p_t^{n1}, \dots, p_t^{nG}]$, is a vector of G age groups (e.g., 4 age groups: 0-17,18-44,44-65, and 65+ years in our simulations). The population at next time step for the region is $p_{t+1}^n = p_t^n + \Delta p_t^n$, where the population increase vector Δp_t^n is stochastic. So, the next state population depends on current population but is not controlled by the agent. To implement our model, we require the current age-grouped population and a reliable growth forecast for each region n.

The current hospital capacity for the nth region at time t is given by $b_t^n = b_{t-1}^n + \Delta b_t^n = b_0^n + \sum_{\tau=1}^{t} \Delta b_\tau^n$. Here $\Delta b_t^n = 0$ unless $A_t = n$, i.e., the region is selected for expansion. We set a suitable number Δb as a unit for bed capacity expansion size, and the selected region's capacity increases by Δb, i.e., $b_t^n = b_{t-1}^n + \Delta b$. b_0^n is the existing hospital capacity at the start ($t = 0$) of an episode for the region. Hence, b_t^n is controlled by the MDP agent deterministically.

B. Action, A_t

The action taken by the agent at time t, $A_t = m$, increases the capacity of the mth region by a fixed quantity Δb. The agent can also decide to do no expansion ($A_t = 0$), i.e., $A_t \in \{0, 1, 2, ..., N\}$. The expansion actions ($A_t \neq 0$) incurs augmentation cost, however reduces the future cost of denial of service (DOS) to the patients. The agent aims to learn the optimal policy to minimize its total cost over a period of time by striking a right balance between the augmentation cost and DOS cost.

C. Cost, C_n

The hospital bed capacity per 1000 people varies between 1.6 (Oregon) – 4.8 (South Dakota) in the US [27], and between 0.1 (Mali) – 13 (Japan) globally [28]. Thus, there is some probability for patients exceeding the hospital capacity in a given day, even for Japan. The places with lower per capita capacity are expected to face more frequent overflow of patients and thus denial of treatment. The paper [2] provides an insight about predicting the number of patient admission with different Poisson distribution mean for each day of the week, validated by real-world data. We use a similar approach with a dynamic Poisson mean for each day in each region based on the population vector that defines our state space. We model the number of beds required for a particular day in the nth region as follows:

$$r_t^n \sim \text{Poisson } (\lambda_t^n(p_t^n)). \tag{1}$$

Here λ_t^n is the dynamic Poisson mean for the nth region at time t, which is a function of the age-grouped population vector p_t^n of that region. We approximate this function from historical data of that region via regression analysis (e.g., decision tree regression).

When the existing capacity is less than the required beds for a day (i.e., $b_t^n < r_t^n$), $r_t^n - b_t^n$ number of patients will be deprived of treatment. To perform any cost-benefit analysis with limited resources, we need to represent the patients' inconvenience in monetary value. Regional economic data and existing literature are useful for determining this monetary value β for each untreated patient. For long term expansion planning, we can approximate the financial cost to be proportionate to the number of untreated patients. We calculate this daily cost as

$$c_t^n = \begin{cases} \beta(r_t^n - b_t^n), & \text{if } r_t^n - b_t^n > 0 \\ 0, & \text{otherwise.} \end{cases}$$

This cost gets summed up over all regions as the DOS cost for the time step t as

$$C_t^{DOS} = \sum_{n=1}^{N} c_t^n. \tag{2}$$

Without loss of generality, we assume the expansion amount Δb to be fixed as one unit expansion throughout the time horizon for all regions, which incurs the augmentation cost $C_t^{aug} = \alpha$. If the agent makes no expansion, then obviously the expansion cost is $C_t^{aug} = 0$. As a result, the total cost for time step t is given by

$$C_t = C_t^{aug} + C_t^{DOS}. \tag{3}$$

D. Next State, S_{t+1}

In our model, only the selected region's hospital capacity increases by Δb units in the next state; the others remain unchanged. So, the agent deterministically controls the bedding capacity of the next state.

$$b_t^{n+1} = \begin{cases} b_t^n + \Delta b, & \text{if} A_t = n \\ b_t^n, & \text{otherwise.} \end{cases}$$

Each region's population for the next state increases by its population growth projections. Specifically, we model the growth using a normal distribution with 20 % variation from the projected increase. Hence, the next state population is stochastic and is not controlled by the agent.

IV. Solution Approach

Our MDP agent aims to minimize the discounted total cost in T time steps,

$$C_T = \sum_{t=0}^{T} \gamma^t C_t, \tag{4}$$

where $\gamma \in (0, 1)$ is the discount factor for future cost. To obtain the optimal policy $\{A_t\}$ we need to solve the following Bellman equation. After i iterations, the agent's value function at time step t becomes

$$V^i(b_t, p_t) = \min_{A_t} \Big\{ \mathsf{E}\left[C_t + \gamma V^{i-1}(b_{t+1}, p_{t+1})\right] \Big\},$$

where $b_t = [b_t^1, \ldots, b_t^N]$ represents the hospital capacity of all regions collectively. Similarly, $p_t = [p_t^1, \ldots, p_t^N)$ represents the population vector of all regions.

Since model-based dynamic programming solutions are not feasible for this high-dimensional MDP, we follow a model-free deep RL approach.

The Advantage Actor-Critic (A2C) algorithm, which is an adaptation of policy gradient-based algorithm REINFORCE [8], fits a continuous state environment. The A2C uses the advantage functions for policy update, which reduces the REINFORCE algorithm's variance.

The actor-network, also known as the policy network, outputs probability for each action value through a softmax function. It aims to find the gradient of expected return $J(\pi_\phi)$

Algorithm 1 A2C algorithm for hospital capacity planning

Input: discount factor γ, learning rate, number of regions R and number of episodes E
Initialize Actor network with random weights ϕ and critic network with random weights θ
for episode $= 1, 2, ..., E$ **do**
  Initialize Health care state $S_0 = (b_0, p_0)$
  for t $= 1, 2, ..., T$ **do**
    for Regions $= 1, 2, ..., R$ **do**
      for d $= 1, 2, ..., days$ **do**
        Generate the number of patient requiring overnight accommodation from Eq. (1).
        Calculate cost due to unattended patients from Eq. (2).
      end for
    end for
    Select action A_t using Eq. (5).
    Execute action A_t and observe cost C_t from Eq. (3).
    Store transitions (S_t, A_t, C_t, S_{t+1}).
    Update actor network ϕ via Eq. (6).
    Update critic network θ through back propagation.
  end for
end for

of the policy π_ϕ with respect to the weights ϕ of the neural network through the following equation:

$$\nabla_\phi J(\pi_\phi) = \mathsf{E}_{\pi_\phi}[\nabla_\phi \log(\pi_\phi(A_t|S_t))\mathcal{A}(S_t; A_t)], \quad (5)$$

where the advantage function is given by

$$\mathcal{A}(S_t; A_t) = \gamma V^{\pi_\phi(S_{t+1};\theta)} - V^{\pi_\phi(S_t;\theta)}. \quad (6)$$

The critic-network learns the value function for each state-action pair. In Eq. (6), $V^{\pi_\phi(S_t;\theta)}$ is the output of critic network for weight matrix θ. A pseudocode for the A2C algorithm is given in Algorithm 1.

V. Case Study

We evaluate our deep RL-based policy using real data from the Tampa Bay region, one of Florida's largest metropolitan areas situated along the Gulf of Mexico. Around 3 million people live in its seven counties with a high elderly population. This metropolitan area is growing fast with a yearly population growth of 1.5% [29], providing a suitable testbed for our case study.

A. Hospital Occupancy Forecasting

Hospital bed requirement for a region depends on multiple factors that might be infeasible to capture. However, for an area of considerable size (e.g., with inhabitants more than 100,000), we can model it through age-segregated population data. We use the age-segregated population and hospital admission data for all the 7 counties of Tampa Bay between 2010 and 2017 [30]. Table I shows that on average more people are admitted to hospital on weekdays than weekends, thus we use separate models to fit weekdays and weekend hospital admission. Higher than 90% accuracy is achieved in predicting hospital admission for 2017, as shown in Fig. 2, by training different regression algorithms on data from 2010-2016. Decision tree regression with Mean Absolute Error (MAE) achieves around 94% accuracy for both weekdays and weekend data. Further investigation of regression analysis and more factors to predict hospital admission may provide better results. However, to keep the focus on our deep RL-based policy, we proceed with the decision tree regression with MAE for predicting the dynamic Poisson mean in Eq. 1 based on the concurrent age-segregated population projections from [31]. We further add 20% variance to the Poisson mean to account for the day-to-day variation. The fact sheet in [31] states the average length of stay per admission as 4.7 days. Hence, we model the number of people staying at hospital as 4.7 times the random number generated by the Poisson distribution for a given day.

Table I: Average hospital admission per day for the 7 counties of the Tampa Bay region.

Year	2015		2016		2017	
County	Wkday	Wkend	Wkday	Wkend	Wkday	Wkend
Hernando	88	60	91	64	94	66
Hillsborough	516	356	527	357	361	522
Manatee	137	89	137	89	143	94
Pasco	225	158	229	161	230	161
Pinellas	417	299	432	298	425	300
Polk	291	205	295	203	304	207
Sarasota	151	92	152	94	153	94

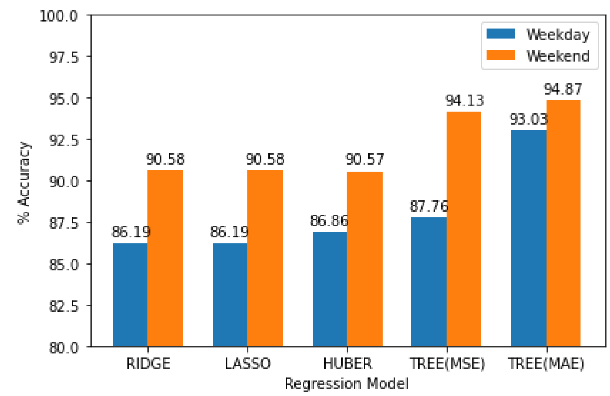

Fig. 2. Average accuracy in terms of predicting hospital admission for different regression models for weekdays and weekends. Results are obtained using the data from [30].

B. Policies

Apart from the deep RL-based policy, we discuss two other myopic policies for a 30-year scheme. All the policies take yearly decisions to augment a maximum of one county with $\Delta b = 120$ hospital bed capacity. We estimate the cost of adding 120 hospital beds to be $\alpha = 100$M USD [32], [33]. The cost of not attending a patient is a complex estimation. We estimate it as $\beta = 0.04$M USD based on [34].

1) Per Capita-Based Policy: The number of hospital beds per 1,000 people is a useful metric to represent a region's healthcare condition. Among the US states, Florida ranks moderately with 2.6 beds per thousand people [27]. However,

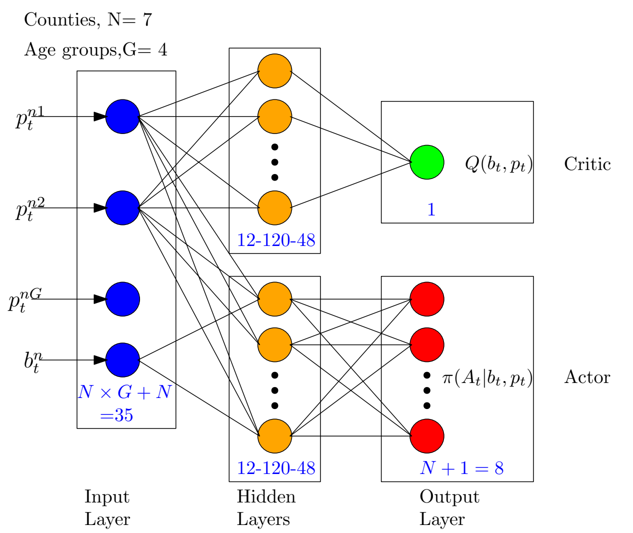

Fig. 3. Neural network architecture for the proposed deep RL-based policy.

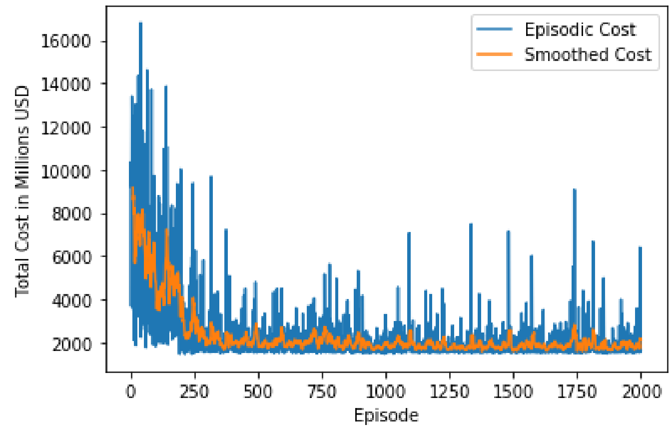

Fig. 4. Convergence of the deep RL-based policy.

Policy	Cumulative cost	Reduction	Complaints	Reduction
Per capita-based	4323$	0 %	32775	0 %
Complaint-based	2708 $	37.36 %	28272	13.74 %
Deep RL-based	2238 $	48.23 %	21566	34.2 %

Table II: Cumulative cost and complaints over a 30-year timeline.

the hospital capacity is not uniformly distributed. Some of the counties have low per capita hospital capacity, making them good candidates for capacity expansion. Selecting a county based on its per capita capacity is a meaningful approach for the healthcare authority. Moreover, if each county's per capita capacity is above a threshold, the authority might decide not to make any augmentation. We do a grid search between 1 to 5 per capita capacity to find the optimal threshold to be 2.9 beds per 1,000 people, which minimizes the 30-year cost for this policy.

2) Complaint-Based Policy: It is often complicated to determine the critical healthcare conditions for making capacity expansion decisions. The number of unattended patients for the previous year for different counties can provide a reasonable guideline for the healthcare authority. They may use this number of deprived patients or complaints as a decision-making criterion. The authority will augment the capacity of the county with the most number of yearly complaints, given that the number is above a threshold. In our simulation setup, we find that the optimal threshold for this policy is 70 complaints.

3) Proposed Deep RL-Based Policy: The actor and critic networks for the deep RL-based policy discussed in Section IV are shown in Fig. 3. We use a learning rate of 0.0003 and discount factor $\gamma = 0.99$ for the simulation. For the $N = 7$ counties, there are $G = 4$ age group populations and also the hospital capacity data, which makes the state a 35-dimensional input for the neural network. The actor and critic networks do not share any hidden layers, but share the common input. We use three hidden layers consisting of 12, 120 and 48 neurons, respectively. The critic network outputs one value for the state space. However, the actor-network outputs $N+1 = 8$ softmax values representing the probability of each action for the given state. Fig. 4 shows that the proposed neural networks converge within 1,000 episodes and learn the optimal policy.

C. Comparative Analysis

We perform comparative analysis among the policies explained above in terms of the total cost and total complaints received due to unattended patients for the 30-year timeline (Fig. 5). The per capita policy incurs maximum cost and complaints compared to the other two policies. This result testifies that the number of hospital admissions is better captured by age-grouped population data, consistent with the general understanding that some age groups require more medical attention (especially children and older people). Per capita-based policy, without age-grouped population distribution incurs 4,323 million USD cost and 32,775 complaints on average for 30 years. We use these values to evaluate the benefit of the other policies.

The complaint-based policy works well to bring down the cost to 2,708 million USD and complaints to 28,272. The complaint-based policy is successful to some extent as it focuses on the outcome by making decisions based on complaints. However, this reactive policy is always one step behind since its decision is based on previous years experience. This is where the proposed deep RL-based policy outperforms all the other approach. It intuitively puts optimal weights to the corresponding feature that determines the future cost. This policy gives the best result by minimizing the cost to 2,238 million USD and complaints to 21,566. The comparative analysis is summarized in Tab. II.

VI. Discussions

We proposed an MDP framework to devise a hospital augmentation plan and illustrated it for the Tampa Bay Area. The proposed model works on age-segregated population data, which is more useful than straightforward population growth rate-based methods. This work is also more realistic

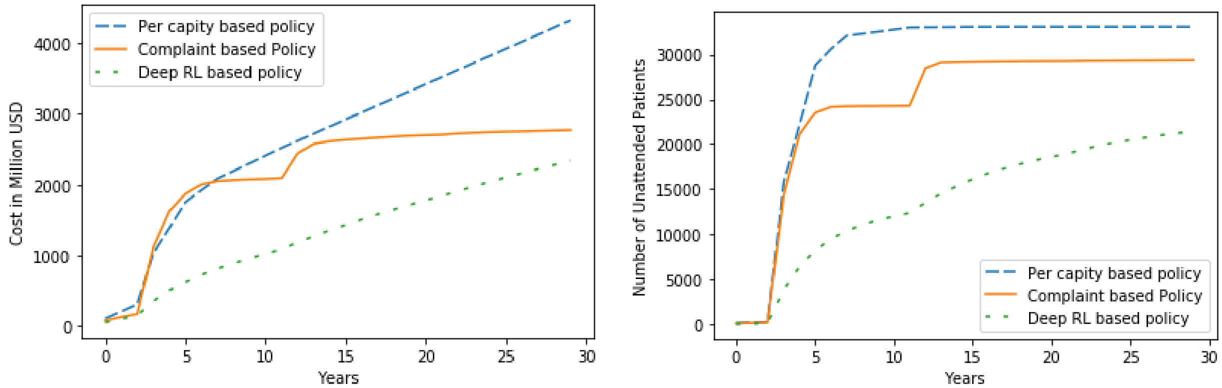

Fig. 5. Comparison between the per capita-based, the complaint-based, and the proposed deep RL-based policy in terms of cumulative cost (left), and number of cumulative complaints (right).

in assessing the hospitalization demands as it works on weekdays and weekend demand separately. It is able to provide a dynamic data-driven policy over a long period. However, there are several limitations for the proposed model. First of all, age-segregated population data might not be enough to predict the future hospital bed requirements. So, looking for other critical factors will make the model more robust. Secondly, this work assumes a fixed cost of α and β for all regions, which is limited in the sense that building the same structure in a crowded city is far costlier than building it in a rural area. Another shortcoming of the model is that it assumes that one region's patient does not seek medical help from a neighboring region. Addressing these shortcomings may provide many future research directions.

VII. Conclusions

This paper showed that hospital bed augmentation plans need dynamic and robust mechanisms to deal with future demand. We proposed an MDP formulation to cost-benefit analysis for augmentation plans for a particular geographic region. We utilized the actor-critic (A2C) deep RL framework to find the optimal policy for the Tampa Bay region. The proposed policy significantly outperforms the myopic policies which depend on a fixed per capita hospitalization estimate and complaints from the previous year. Although illustrated only for Tampa, Florida, the proposed model is applicable to other regions where sufficient data are available.

References

[1] L. Seematter-Bagnoud, S. Fustinoni, D. Dung, B. Santos-Eggimann, V. Koehn, R. Bize, A. Oettli, and J.-B. Wasserfallen, "Comparison of different methods to forecast hospital bed needs," *European Geriatric Medicine*, vol. 6, no. 3, pp. 262–266, 2015.

[2] G. W. Harrison, A. Shafer, and M. Mackay, "Modelling variability in hospital bed occupancy," *Health Care Management Science*, vol. 8, no. 4, pp. 325–334, 2005.

[3] L. V. Green, "How many hospital beds?" *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, vol. 39, no. 4, pp. 400–412, 2002.

[4] W. Shonick, "A stochastic model for occupancy-related random variables in general-acute hospitals," *Journal of the American Statistical Association*, vol. 65, no. 332, pp. 1474–1500, 1970.

[5] W. M. Hancock, J. B. Martin, and R. H. Storer, "Simulation-based occupancy recommendations for adult medical/surgical units using admissions scheduling systems," *Inquiry*, vol. 15, no. 1, pp. 25–32, 1978.

[6] Worldometer, "United states demographics," https://www.worldometers.info/demographics/us-demographics/#median-age, Accessed: 2020-02-09.

[7] R. Bellman, "A markovian decision process," *Journal of mathematics and mechanics*, pp. 679–684, 1957.

[8] P.-H. Su, P. Budzianowski, S. Ultes, M. Gasic, and S. Young, "Sample-efficient actor-critic reinforcement learning with supervised data for dialogue management," *arXiv preprint arXiv:1707.00130*, 2017.

[9] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[10] H. Ravaghi, S. Alidoost, R. Mannion, and V. D. Bélorgeot, "Models and methods for determining the optimal number of beds in hospitals and regions: A systematic scoping review," *BMC Health Services Research*, vol. 20, no. 1, p. 186, 2020.

[11] L. Seematter-Bagnoud, S. Fustinoni, D. H. Dung, B. Santos-Eggimann, V. Koehn, R. Bize, A. Oettli, and J. B. Wasserfallen, "Comparison of different methods to forecast hospital bed needs," *European Geriatric Medicine*, vol. 6, no. 3, pp. 262–266, jun 2015.

[12] S. Sitepu, H. Mawengkang, and I. Husein, "Optimization model for capacity management and bed scheduling for hospital," in *IOP Conference Series: Materials Science and Engineering*, vol. 300, no. 1, 2018, p. 012016.

[13] B. G. Barnes and N. L. Harp, "The U.S. Medicare Disproportionate Share Hospital program and capacity planning," *Journal of Accounting and Public Policy*, vol. 37, no. 4, pp. 335–351, 2018.

[14] A. Larsson and A. Fredriksson, "Tactical capacity planning in hospital departments," pp. 233–245, feb 2019.

[15] N. Bouckaert, K. Van den Heede, and C. Van de Voorde, "Improving the forecasting of hospital services: A comparison between projections and actual utilization of hospital services," *Health policy*, vol. 122, no. 7, pp. 728–736, 2018.

[16] J. Matos and P. P. Rodrigues, "Modeling decisions for hospital bed management - A Review," in *Proceedings of the International Conference on Health Informatics*, vol. 2, 2011, pp. 504–507.

[17] Z. Zhu, B. H. Hen, and K. L. Teow, "Estimating ICU bed capacity using discrete event simulation," *International Journal of Health Care Quality Assurance*, vol. 25, no. 2, pp. 134–144, 2012.

[18] A. X. Costa, S. A. Ridley, A. K. Shahani, P. R. Harper, V. De Senna, and M. S. Nielsen, "Mathematical modelling and simulation for planning critical care capacity," *Anaesthesia*, vol. 58, no. 4, pp. 320–327, 2003.

[19] R. B. Tate, L. MacWilliam, and G. S. Finlayson, "A Methodology for Estimating Hospital Bed Need in Manitoba in 2020," *Canadian Journal on Aging / La Revue canadienne du vieillissement*, vol. 24, no. S1, pp. 141–151, 2005.

[20] J. Schiele, T. Koperna, and J. O. Brunner, "Predicting intensive care unit bed occupancy for integrated operating room scheduling via neural networks," *Naval Research Logistics (NRL)*, vol. 68, no. 1, pp. 65–88, feb 2021.

[21] A. Kumar, R. J. Jiao, and S. J. Shim, "Predicting bed requirement for a hospital using regression models," in *IEEE International Conference on Industrial Engineering and Engineering Management*, 2008, pp. 665–669.

[22] E. Kutafina, I. Bechtold, K. Kabino, and S. M. Jonas, "Recursive neural networks in hospital bed occupancy forecasting," *BMC medical informatics and decision making*, vol. 19, no. 1, p. 39, 2019.

[23] S. S. Shuvo, M. R. Ahmed, S. B. Kabir, and S. A. Shetu, "Application of Machine Learning based hospital up-gradation policy for Bangladesh," in *ACM International Conference Proceeding Series*. New York, NY, USA: Association for Computing Machinery, dec 2020, pp. 18–24.

[24] "Reinforcement learning model, algorithms and its application," in *Proceedings 2011 International Conference on Mechatronic Science, Electric Engineering and Computer, MEC 2011*, 2011, pp. 1143–1146.

[25] C. Yu, J. Liu, and S. Nemati, "Reinforcement Learning in Healthcare: A Survey," *arXiv*, aug 2019. [Online]. Available: http://arxiv.org/abs/1908.08796

[26] O. Gottesman, F. Johansson, M. Komorowski, A. Faisal, D. Sontag, F. Doshi-Velez, and L. A. Celi, "Guidelines for reinforcement learning in healthcare," *Nature Medicine*, vol. 25, no. 1, pp. 16–18, jan 2019.

[27] A. H. Association *et al.*, "Trendwatch chartbook 2018: trends affecting hospitals and health systems, 2018," 2019.

[28] W. Bank, "Hospital beds per 1000 people," ://-data.worldbank.org/indicator/SH.MED.BEDS.ZS.

[29] Wikipedia, "Tampa bay area," 2021. [Online]. Available: https://en.wikipedia.org/wiki/Tampa_Bay_area

[30] "HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2010-2017." [Online]. Available: www.hcup-us.ahrq.gov/sidoverview.jsp

[31] S. Rayer and Y. Wang, "Projections of florida population by county, 2020–2045, with estimates for 2019," *Bureau of Economics and Business Research, Florida Population Stuides, Bulletin 186*, 2020.

[32] N. S. Kalman, B. G. Hammill, K. A. Schulman, and B. R. Shah, "Hospital overhead costs: The neglected driver of health care spending?" *Journal of Health Care Finance*, vol. 41, no. 4, 2015.

[33] "Cost to Build a Hospital — Hospital Construction Cost." [Online]. Available: https://www.fixr.com/costs/build-hospital

[34] D. Farrell and F. E. Greig, "Paying out-of-pocket: The healthcare spending of 2 million us families," *SSRN Electronic Journal*, 2017.