Monte-Carlo Simulation Based on Patient-Individual Distributions for Supporting Intensive Care Occupancy Management

Arne Henning Witteborg
Bielefeld University
henning.witteborg@uni-bielefeld.de

Michael Römer
Bielefeld University
michael.roemer@uni-bielefeld.de

Markus Günther
Bielefeld University
markus.guenther@uni-bielefeld.de

Rainer Borgstedt
Hospital of Bethel Foundation
rainer.borgstedt@evkb.de

Gerrit Jansen
Hospital of Bethel Foundation
gerrit.jansen@evkb.de

Abstract

Managing Intensive Care Units (ICUs) in hospitals is a highly challenging endeavor. In particular, decisions such as admitting elective patients and discharging patients from the ICU have to be taken under a high level of uncertainty since the occupancy of ICUs does not only depend on these decisions but also on unknown parameters such emergency patient arrivals and lengths of stay of the patients in the ICU. In this paper, we develop a framework for supporting ICU occupation management by quantifying the impact of admission and discharge decisions on the probability of reaching critical ICU occupancy levels in a given planning horizon. A key component of this framework is the use of data-driven approaches for obtaining probability distributions for the parameters affected by uncertainty. In particular, we use standardized treatment and patient health state data to create patient-specific length-of-stay distributions with a Machine Learning approach. These patient-individual distributions are then validated and/or adjusted by medical experts. The validated distributions form the input to a Monte-Carlo Simulation that is used to approximate the probability distributions of the daily ICU occupancy levels resulting from ICU admission and discharge decisions. We experimentally evaluate our framework in a counterfactual simulation based on one year of historical data from 2019 from a medium-sized ICU in a German hospital. In that evaluation, we use a simple ICU management policy based on the probabilistic occupancy forecasts aiming at reducing the risk of running out of ICU capacity. The results show that following this policy would have avoided hitting critical occupancy levels by around 70% and would have had a smoothing effect on ICU occupancy levels.

1. Introduction

ICUs are one the most complex, sensitive and financially important resources of a hospital [1, 2]. Increasing capacity utilization without risking to run out of capacity, however, is a challenging endeavor: The ICU occupancy is affected by highly uncertain factors such as the number of new emergency patients or the length-of-stay (LoS) in the ICU which is not only uncertain for emergency patients, but also for elective patients and patients already staying in the ICU [3, 4]. For a given patient, the LoS is affected by various individual aspects such as age, health state, interventions performed and comorbidities [5, 6]. In this paper, we deal with the problem of short-term management of ICU occupancy by supporting the daily planning process for taking admission and discharge decisions for a horizon of around five days, which identified as the period of interest with our cooperation hospital. Our goal is to show that by combining patient-specific probabilistic LoS forecasting with Monte-Carlo Simulation, it is possible to provide probabilistic ICU occupancy projections of the consequences of ICU management decisions. These can be used to guide the decision-making process by providing a quantitative assessment of the risk of reaching critical ICU occupancy levels.

For a recent overview of the literature dealing with quantitative approaches for ICU management, see [4]. Acknowledging the essential influence of uncertainty on ICU occupancy, most of the research contributions dealing with ICU management rely on stochastic approaches such as queuing models [3], Markov Decision Processes [7, 1] or simulation-based methods [8]. In most cases, differences between patients affecting the LoS are accounted by forming groups for which separate distributions and/or transition probabilities are derived from historical data in form of

empirical distributions or fitted parametric distribution models.

When it comes to forecasting patient-specific LoS, e.g. based on patient health state data, most approaches discussed in the literature aim at providing point forecasts (for an example see [9] and an overview and a comparison of LoS prediction approaches is provided in [10]). They are only of very limited use for a probabilistic risk assessment of ICU management decisions. Point forecasts, however, cannot be used for a probabilistic risk assessment of ICU management decisions since they do not quantify the uncertainty associated with the the forecast while there is very little research on probabilistic LoS forecasting, that is on approaches providing LoS distributions that are conditioned on patient-specific information. Such an approach was recently presented in [11], where patient-related attributes such as gender, age and aggregated health state and interventions scores are used to obtain patient-individual distributions for the full LoS for COVID-19 patients. The method underlying that approach was introduced in [12] and provides additionally a good overview on distributional forecasting approaches.

In the paper at hand, we also devise a probabilistic LoS forecasting approach, which is embedded in a risk-oriented framework for supporting ICU management decisions where it is used to generate patient-specific LoS distributions. Other than [11] we aim at forecasting the remaining LoS of current patients within a relatively short time horizon of five days which allows for using more fine-granular characterizations of patient states and therapeutic interventions. For the forecasting we use the Machine Learning approach XGBoost [13]. To the best of our knowledge, our approach is the first to combine probabilistic Machine Learning for obtaining patient-specific LoS distributions with Monte-Carlo Simulation for supporting ICU management decisions.

The approach presented in this paper was developed in a close collaboration between the authors coming from a medium-sized German ICU with a capacity of 18 beds and from a group of OR / data science experts from an Business Department from a German university. In particular, all assumptions and design decisions were thoroughly discussed and checked from the clinical perspective by the partners from the ICU. The data used for training, fitting and calibrating the probabilistic models used in our approach are also used in the counterfactual simulation stem from the case study ICU.

The remainder of this paper is structured as follows: In Section 2 we provide an overview of the proposed decision-support framework. The approach used for the data-driven generation of the input distributions including the approach for deriving patient-specific LoS distributions is described in greater detail in Section 3 and Section 4 discusses how these distributions are used in a Monte-Carlo Simulation to obtain a risk assessment for ICU management decisions. In Section 5, we present results from a first simple impact analysis for our approach in form of a counterfactual historical simulation. These results and the general potential of the approach are discussed in Section 6.

2. A Framework for Risk-Oriented ICU Occupancy Management

To support the daily decision-making process in an ICU affecting the occupancy levels the following days, we propose a framework that allows for supporting decision-makers (ICU managers) to cope with the inherent uncertainty involved in the planning process stemming from (i) the (unknown) number of new patients as well as (ii) the (unknown) patient-individual length-of-stay distributions (for an overview of the framework refer to Figure 1). These distributions need to be estimated for (i) current patients in the ICU, (ii) emergency patients that enter or re-enter the ICU and (iii) elective patients scheduled for surgeries from other departments.

The proposed planning and decision-making process consists of four phases, which will be procedurally described in the remainder of this section in order to give a holistic overview. The technical details of the two data/algorithm-driven phases 1 and 4 are provided in the subsequent Sections 3 and 4.

Phase 1: Data-driven Generation of Input Distributions In the first phase, distributions for the uncertain inputs are automatically generated, that is for the arrival of emergency patients and for the lengths of stay for each patient from the three groups.

For the *current patients* already staying in the ICU at the time of planning, the patient-individual distributions of the LoS are estimated using a probabilistic Machine Learning approach and based on two daily updated individual patient health data sources: the individual items of the "Simplified Acute Physiology Score II" (SAPS II) for the patients severity of illness and the items of the "Therapeutic Intervention Scoring System" (TISS) for the patients interventions (more information on the machine learning approach will be given in Section 3.2). These standardized medical values and treatment activities score items have proven their usefulness predictive purposes in several studies
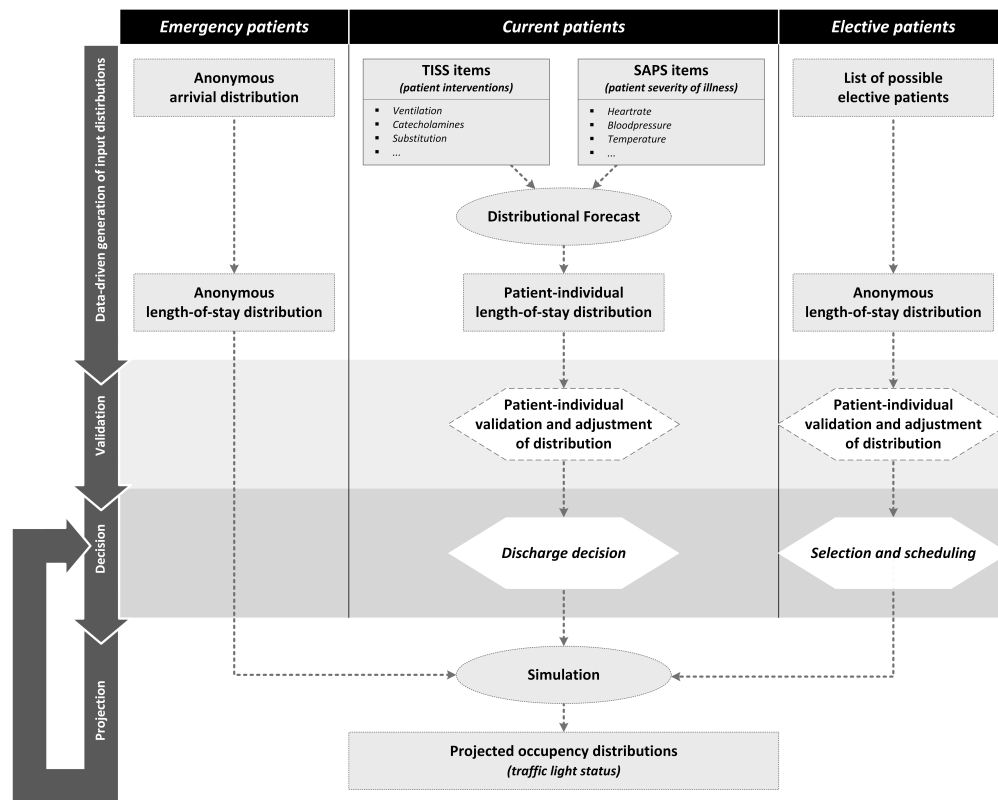
Figure 1. Framework

before [14, 15]. Therefore, these diagnostic data of a patient collected on a daily basis throughout the current hospital stay can provide valid indication for the expected occupancy in the ICU [16].

As emergencies cannot be postponed or influenced by the ICU or any other specialist department of the hospital, only anonymous arrival and anonymous length-of-stay distributions (compared to the non-anonymous and patient-individual distributions of current patients) can be used in order to consider the arrival of new (external as well as internal) *emergency patients* for the evaluation of the ICUs capacity. In order to get those anonymous distributions, an analysis of historical data of emergencies was performed (for methodological information see Section 3.1).

Whereas the number as well as the LoS of emergency patients remains highly uncertain, more information is available for the third group of patients, the *(potential) elective patients*. As they are scheduled for a surgery by other medical departments of the hospital due to a planned intervention, the number of patients as well as the types of interventions are known. Yet, it is hard to obtain a valid patient-specific LoS for the elective patients from data and thus, the proposed anonymous distributions need to be adjusted by the

decision-makers in the next phase.

Phase 2: Validation In the second phase, the patient-individual LoS-distributions of current and (potential) elective patients are (manually) validated and (potentially) adjusted by the decision makers. This allows for considering any additional (medical) information not incorporated in the current data set so far (e.g., patient-individual impact of planned interventions), before the patient-individual LoS-distributions of both patient groups are used in the simulation. Note that the effort for the validation phase turns out to be manageable: On the one hand, mainly the LoS distributions for the elective patients will call for an adjustment and the number of these patients is rather small. On the other hand, the LoS distributions are discrete distributions with a small sample space since all lengths of stays exceeding the planning horizon are summarized in a single outcome. Finally, our prototype offers functionality that simplifies the allocation of probabilities to the different outcomes from a user perspective.

Phase 3: Decision In the third phase, ICU-managers are assumed to take two decisions. The first decision is whether current patients should be discharged from the ICU due to improvements of their health status. Consequently, discharged patients as well as the deceased patients will be excluded from the set of patients considered in the simulation. The second type of decision affects the admission of patients in the ICU: A subset of the potential elective patients is selected and scheduled in a way that critical ICU occupancy levels are avoided. Other than emergency patients, planned stays of elective patients in the ICU can be scheduled with some flexibility and planned admissions can often be postponed without serious medical consequences, e.g. in case of a high occupancy. Scheduling elective admissions is the central decision in the overall process as it as it has a direct managerial impact on the overall capacity of the ICU. Therefore, this selection step forms the starting point for the iterative decision loop, which allows for evaluating the impact of different sets of elective patients on the capacity.

Phase 4: Projection Based on the random distributions obtained from data and validated by medical experts, the impact of the ICU management decisions (discharge and selection an scheduling of elective patients) on the (probabilistic) occupancy of the ICU can now be assessed using a simulation-based approach (for more information on the simulation see Section 4). The outcomes of the Monte-Carlo Simulation, that is, the estimated distributions of the occupancy levels on each day in the planning horizon, support the decision-makers with useful information allowing to assess the risk of a given set of ICU management decisions. For instance, a simple traffic light systems indicates the risk of reaching critical occupancy levels throughout the planning horizon that can be used to guide the complex decision process in ICU management.

Iterative Decision-Making Process The output of phase 4 aims at helping to identify requirements or potentials for adjusting admission decisions taken in phase 3 with the focus on ICU capacity. Consequently, the preliminary decisions (e.g. number of elective patients) can be altered and the impact of these changes becomes instantly visible in the iterative decision-making process by looping between the decision and the projection phase. Like this, various possibilities can be evaluated interactively. For example, as surgical interventions that are not critical from a medical point of view can be postponed for elective patients, decision-makers can adjust their decision if

the traffic light indicates high risk of entering a red occupancy level the following days. This would allow adjusting the risk of reaching the maximum capacity by reserving capacities for possible emergencies or longer periods of stay for the patients currently in the ICU. In a nutshell, the information output provided by the simulation can then be used in order to alter preliminary decisions by looping between the decision and the projection phase in this iterative decision process performed each day with the focus of the overall goal to avoid the risk of reaching the maximum capacity.

3. Data-Driven Generation of Input Distributions

In this section, we focus on the methodological description of the automatic and data-driven generation of emergency arrival and length-of-stay distributions using historical data generated in phase 1 of our framework. We start with the distributions for patients arriving during the planning period, that is, emergency and elective patients, followed by the description of the generation of the patient-individual LoS distributions for patients who already stay in the ICU at the time of planning.

3.1. Patients Admitted During The Planning Period

In our approach, we use historical data to obtain the distributions regarding the patients admitted during the planning period, that is, the arrival of emergency cases and the length of stay of both emergency and elective patients. For the case study described in Section 5, we used an anonymized data set from the case study hospital containing ICU admission and discharge dates, admission reasons (e.g. emergency, post-surgery).

Emergency Arrivals The number of emergency cases arriving per day (we do not distinguish between internal and external emergency cases) form an integer-valued discrete distribution denoted as E^{arr}. In the historical data from our case study hospital spanning around two years, the maximum number of emergency cases per day was 2. Given the relatively small size of the data set, we decided to refrain from conditioning the emergency arrival distribution on calendar-based covariates (e.g. day of week or month) and simply use the relative frequencies from the historical data to form a single empirical emergency arrival distribution that is used for each day.

Length of Stay for Emergency Cases In addition, we do not distinguish between the severities of the emergency cases. As a result, we use a single length-of-stay distribution L^{em} for emergency patients which we also estimate as an empirical distribution from the historical data. As for all length-of-stay distributions considered in this paper, this distribution is discrete and its sample space reflects the planning horizon T and is $\{0, 1 \ldots \geq |T| - 1\}$ where $|T|$ is the number of days in the planning horizon.

Length of Stay for Elective Patients Regarding the elective patients, that is, those with a post-operative stay in the ICU, we once again use a single empirical distribution L^{el} based on historical data. While we considered conditioning the LoS distribution on information regarding the intervention such as operation and procedure codes used in Germany, it turned out that the case numbers for identical or similar codes in our data set were too low. In addition, several patient-specific aspects such as health state, age and comorbidities and their impact in conjunction with the uncertain outcome of an intervention such as a surgery make it hard to obtain a valid data-driven patient-specific LoS distribution for elective patients. As a result, we agreed with the ICU decision-makers to use a single distribution for all elective cases as a first proposal, which is then adjusted in the validation and adjustment step described in Section 2. As a consequence, our simulation assumes that the length-of-stay distributions of elective patients are patient-specific; the distribution for a given patient i is denoted with L_i^{el}.

3.2. Patient-Individual Distributions for Current Patients

The majority of the patients to be considered during the planning horizon already stay in the ICU at the time of planning (from the first day after the admission onwards, each patient is a current patient). For these patients, there is typically a lot of data describing their current therapeutic interventions and health status which can used as covariates for creating an individualized distribution for the remaining LoS in the ICU from the current day onwards.

Data Set Our framework makes use of the already existing and daily collected data used for the two standardized scores SAPS II and TISS. While these (aggregated) scores are originally used to measure the treatment effort for billing purposes, they—as well as similar scores—are also useful for predictive purposes, e.g. for forecasting mortality rates [17], ICU admission

probabilities [16] or lengths of stay [18, 10, 11]. Note that our approach makes use of the data items underlying the calculation of these scores.

The items contributing to the SAPS II encompass a wide range of measurements used to characterize the health state of a patient [19]. The values measured in each item (e.g. the heart rate; for a full list see Table 1) are translated into integer-valued points which are then aggregated to the daily SAPS II score.

Item	Explanation
Heart rate	Beats per minute
Blood pressure	Systolic blood pressure
Temperature	Highest temperature in 24 hours
Oxygen ratio	PaO_2/FiO_2 ratio
Urine	Urine output, mL/day
Urea	Dialysis/Filtration
Leucocytes	White blood cells/mm^3
Potassium	Worst value in 24 hours
Sodium	Worst value in 24 hours
Bicarbonate	Lowest value in 24 hours
Bilirubin	Highest value in 24 hours
Admission state	Scheduled surgical, medical, unscheduled surgical
Chronical disease	Chronical disease (Cancer e.g.)
Age	Age in years

Table 1. Simplified Acute Physiology Score (SAPS)

The purpose of the TISS scoring system is—according to its conception [20, 21]—to quantify the daily care and therapy effort required for a patient (the items contributing to the TISS core are listed in Table 2). In a TISS-related data set for a patient on a given day, each of the contributing items is assigned a standardized integer value between 0 and 5 which are then summed up to the aggregated TISS score.

SAPS II and TISS data is collected in a standardized way each day for each patient on a German ICU and, thus, form a good basis for applying our framework in German hospitals. For our case study ICU, we had a corresponding data set containing daily SAPS II and TISS entries and LoS information for each ICU patient from November 2017 to December 2019. In the following, we describe how we used this data to train a model for generating patient-specific LoS distributions to be used in the Monte-Carlo Simulation.

Generating Patient-Individual Length of Stay Distributions While most of the existing LoS-forecasting approaches aim at providing point forecasts, we need to predict length-of-stay distributions.

Item	Explanation
Ventilation	Mechanical ventilation
Catecholamines	Infusion of multiple catecholamines
Substitution	Fluid intake
Arterial	Peripheral Catheter (arterial System)
Cardiac	Catheter (pulmonary artery)
Renal	Dialysis/Filtration
Cranial	Intracranial pressure measurement
Metabolic	Acidosis/Alkalosis
Interventions	Interventions on ICU
Actions	Actions outside ICU

Table 2. Therapeutic Intervention Scoring System (TISS)

Our system uses the values (e.g. the number of points corresponding to a measured heart rate) assigned to each item depicted in Tables 1 and 2 as covariates to generate patient-individual distributions $L_{t,i}^{\text{remain}}$ for the remaining LoS of an existing patient i already being in the ICU at t. In other words, the LoS distribution for a current patient is a conditional distribution based on covariates describing the health state and the interventions on the day before, that is on day $t^0 - 1$.

We estimate the length-of-stay distribution using an XGBoost [13] classifier. XGBoost is a Machine Learning approach based on an ensemble of decision trees. In particular, it makes use of gradient boosting to iteratively add trees to the ensemble until no improvements can be observed. Additional techniques such as tree pruning and parallelization are used to keep the size of the model and the training time moderate. In order to obtain a probabilistic classification with XGBoost, we use its *softprob* loss function; otherwise we all of XGBoosts parameters are kept at their standard values. We train the classifier using the historical data described above as follows: For each patient i and for each day t the patient stayed in the ICU, we create one data point consisting of the SAPS II and TISS item values of i on day t as features / covariates and the remaining LoS (the time from t to the discharge day of patient i) as label (as explained above, all LoS labels exceeding the planning horizon receive the same class label).

We also evaluated other approaches such as a distributional regression and other classifiers such as a logistic regression. However, even after calibrating these other classifiers, XGBoost reached the best performance in terms of measures for distribution fitting such as log likelihood and Brier Score. We assume that two aspects make it a good choice: First, we only have a couple of discrete outcome values to predict, and as a consequence the fact that a classifier does not impose the order that is present in the outcomes does not seem to matter too much. Second, the fact that XGBoost relies on decision trees may help to cope with the fact that there are two reasons for a relatively short remaining length of stay, which are associated with very different covariate values: A patient gets better and can be transferred to a non-ICU bed, or a patient is in a bad state and deceases.

4. Projecting the Effects of ICU Management Decisions on Occupancy Levels

In our framework, the distributions automatically generated from the data and possibly adjusted by the ICU managers form the basis for a Monte-Carlo Simulation used in the projection phase (phase 4) in order to to evaluate the ICU management decisions affecting a planning period $T = \{t^0, t^1, \ldots t^{|T|-1}\}$. The result of the simulation is an approximate probability distribution B_t for the ICU occupancy level on each day $t \in T$. In this section, we first describe this simulation, followed by a description of a "traffic light system" that uses the distributions B_t to visually indicate the risk of exceeding critical occupancy thresholds.

Monte-Carlo Simulation We assume that the Monte-Carlo Simulation is used to evaluate the following decisions taken in phase 3 of our framework in the morning of day t^0, the first day of the planning period: (i) The schedule of surgeries determining the sets $\mathcal{I}_t^{\text{el-sched}}$ of elective patients to be admitted in the ICU on each day $t \in T$ and (ii) the patients $\mathcal{I}_{t^0}^{\text{dis}}$ to be discharged from the ICU at day t^0. Note, we assume in our simulation that discharge decisions are taken only for discharges at day t^0 and that the discharge of patients during the rest of the simulation horizon is assumed to depend only on the patient's LoS distribution and is not affected on (path-dependent) ICU occupation levels.

The Monte-Carlo Simulation approximates the ICU occupancy distributions B_t for each $t \in T$ by computing a set of N sample paths where each sample path n forms one possible realisation of the planning week, yielding one observation b_t^n of the occupancy per day $t \in T$. The vector $[b_t^n]_{n=1..N}$ of occupation samples then serves as a sample approximation for the ICU occupancy distribution B_t for day t. The standard setting in our software implementation for N is 1 000 000, yielding a running time of about 15-20 seconds on

a standard notebook.

A single sample path n is computed as follows: First, we draw the remaining lengths of stays for each existing patient who is not discharged, that is, for each $i \in \mathcal{I}^{\text{cur}} \setminus \mathcal{I}^{\text{dis}}_{t^0}$, we draw a sample l^n_i from the patient-individual LoS distribution $L^{\text{remain}}_{t^0,i}$. For the patients in $\mathcal{I}^{\text{dis}}_{t^0}$, we set $l^n_i = 0$. Then, we draw a length of stay l^n_i for each elective patient $i \in \mathcal{I}^{\text{el-sched}}_{t^0}$ admitted at day t^0 from the distribution $L^{\text{el}}_{t^0,i}$. Finally, we draw a number e^n_t of emergency cases arriving at t_0 from the distribution E^{arr}. Based on the number of arriving cases, we create the set of admitted emergency patients $\mathcal{I}^{\text{em-ad},n}_{t^0}$ at t^0 by admitting as many of the arriving emergency cases as as there is open ICU capacity on day t^0 in sample path n. For each of the admitted emergency patients $i \in \mathcal{I}^{\text{em-ad},n}_{t^0}$, we then draw a length of stay l^n_i from the respective distribution L^{em}. We are then ready to compute the occupancy $b^n_{t^0}$ on day t^0 consisting of the non-discharged existing patients, the elective patients and the emergency patients admitted at day t^0.

For each of the subsequent days $t \in \{t^1, \dots t^4\}$ in the planning horizon, we proceed by first computing the occupancy b'^n_t before any new admission. It is composed of the patients $I^{\text{cur}}_{t^0}$ whose stay includes t (that is, those patients $i \in \mathcal{I}^{\text{cur}}_{t^0}$ for whom $t^0 + l^n_i - 1 \geq t$), and of the emergency and elective cases admitted on a day $t' < t$ and whose stay includes t (that is, $t' + l^n_i - 1 \geq t$. Given this intermediate occupancy b'^n_t, as many of the scheduled elective patients $\mathcal{I}^{\text{el-sched}}_t$ are admitted at day t as there is remaining capacity; yielding the set $\mathcal{I}^{\text{el-ad},n}_t$. Any patient who is not admitted is postponed and added to the set $\mathcal{I}^{\text{el-sched}}_{t+1}$. For each of the patients $i \in \mathcal{I}^{\text{el-ad},n}_t$, we draw a length of stay l^n_i. Finally, we draw a number e^n_t of arriving emergency patients, admit as many as there is remaining capacity to obtain the set $\mathcal{I}^{\text{em-ad},n}_t$ and also draw a length of stay l^n_i for each of them. The ICU occupation b^n_t for day t in sample path n can then be computed as $b'^n_t + |\mathcal{I}^{\text{el-ad},n}_t| + |\mathcal{I}^{\text{em-ad},n}_t|$.

Traffic Light-Based Characterization of Risk Levels
In order to obtain a relatively simple characterization of the level of risk resulting from ICU management decisions, we agreed to use two occupancy threshold to partition the occupancy levels into the three "traffic light" categories green, yellow and red. As an example, for our case study hospital, we used the thresholds 70% and 85% of the available ICU capacity: An occupancy of up to 70% of the total capacity corresponds to green, an occupancy between 70% and 85% corresponds to yellow and an occupancy above 85% corresponds to red.

The ICU occupancy distributions B_t can then be used to characterize the probabilities of ending up in the three categories. As an example, a certain ICU admission schedule may lead to a probability of exceeding the threshold for the category yellow of 10% and a probability of 2% of ending up in category red on a given day t in the planning horizon. One way of using this information for decision-making is to define decision policies based on probability thresholds for one of for both categories. As an example, one may consider a probability of 1% of reaching the category red as acceptable. As a result, if a given ICU admission schedule exhibits a probability of 5% of reaching the level red on a given day, the ICU decision-maker would re-enter the decision phase described in Section 2, e.g. to cancel or postpone certain elective surgeries. If, on the other hand, the risk of exceeding the yellow threshold is relatively low, one may consider scheduling additional elective surgeries. In any case, after adjusting the ICU management decisions, the simulation can be run again to evaluate their impact on the risk of reaching a yellow and/or red occupancy level.

5. Impact Analysis based on Historical Data

In order to get an impression of the potential of our approach for supporting risk-oriented decision-making in ICU management, we performed a simplified counterfactual impact analysis. In this analysis, we use historical data from the ICU of our case study hospital to see if a simple postponement policy based on the traffic light-based risk assessment described above could have avoided critical occupancy levels. In other words, the following analysis emulates what would have happened if the historical decision-making process would have followed the traffic light-based risk assessment our system would have provided at the time of planning.

Data, Decision Policy and Logic of the Counterfactual Simulation For our analysis, we use historical data from our case study ICU covering the year 2019 (which was not yet affected by the COVID-19 pandemic). The data set contains the admission and discharge dates as well as the admission reasons (emergency vs. elective) of all ICU patients. In addition, the data set contains the item values for the SAPS II and TISS scores for each patient and for each day the patient stayed in the ICU. Recall that these item values are needed for generating the patient-specific LoS distributions.

In our case the ICU has a total capacity of 18 beds. Using the percentages underlying the traffic light categories described in the last section (70% and 85%),
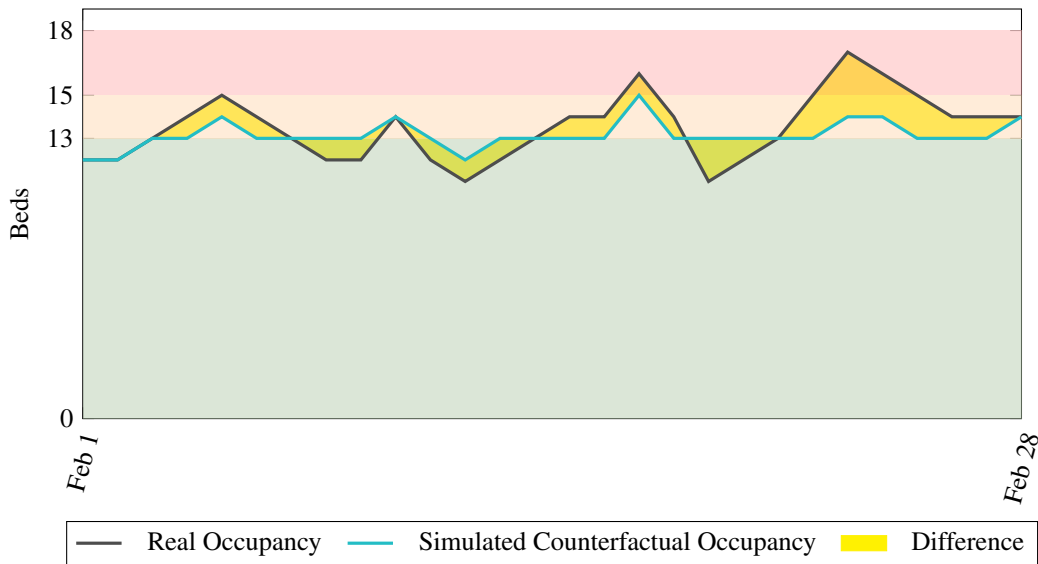
Figure 2. ICU occupancy in February 2019: History vs. Counterfactual Simulation

the yellow range starts at an occupancy of 13 beds, and an occupancy of ≥ 15 beds means that we are in the red, that is, in the critical range. For our counterfactual simulation, we basically simply "replay" everything as it historically happened (arrival of emergency patients, LoS of each patient) with the exception of the admission date of the elective patients which may be postponed according to the policy that elective patients will only be admitted as long as the probability of entering the red occupancy range is $\leq 5\%$ for the next day. This policy is applied to the historical data as follows: For a given day t^0 on which the planning occurs, we consider elective patients historically admitted at day t^0 as scheduled elective patients for day t^0. We then generate all relevant distributions. In particular, we generate the patient-specific LoS distributions for the current patients using the probabilistic prediction model described in Section 3.2 based on SAPS II and TISS item values for those patients that were available at day t^0. Based on this information the Monte-Carlo Simulation described in Section 4 is executed and used to provide the probability of reaching a critical occupancy. If this probability is $\geq 5\%$, then one of the elective patients historically admitted on day t^0 is postponed and added to the set of elective patients scheduled at day $t^0 + 1$. Then, the emulated decision-making process for day t^0 starts again with the reduced set of elective patients until either the admission decision is accepted or the set of elective patients for t^0 is empty. After that, the same overall procedure is applied to the next day $t^0 + 1$, assuming that the set of elective patients consists of those historically admitted at day $t^0 + 1$ *and* the

patients for which the admission was postponed the day before (these may include patients whose admission was postponed multiple times). Note that in this experimental setup, stochasticity only plays a role in the Monte-Carlo Simulation used for deriving the admission / postponement decision; the counterfactual occupancy levels form a single simulation path resulting from postponing admissions given the policy described above; all other parameters such as emergency cases and the patient's LoS are the same as in the historical data.

Experimental Results The key result from performing the counterfactual simulation described above with the data from the year 2019 is that compared to the true historical occupancy, the number of days on which the critical occupancy was reached could be reduced by more than 70%.

Figure 2 compares the course of the occupancy levels in the counterfactual simulation based on the traffic light policy with the historical one for February 2019. During that month, the ICU occupancy exhibited three peaks, which reached the critical red occupancy range. In total, the red range was reached on six days. In the counterfactual simulation, the critical range (its lower bound of 15) was only observed on a single day. On that day, two emergency cases were admitted to the ICU, a relatively rare event that happens just in a few days in this month. Additionally, a smoothing effect resulting from postponing elective patients can be observed. For instance, those patients whose admission was postponed to avoid reaching the first peak and got admitted a few

days later yield a higher occupancy for these days in the counterfactual simulation than in the historical data. At the end of the month it can be observed that there are still a number of postponed patients waiting to be admitted when the traffic light-based admission policy permits.

The fact that the postponed patients were admitted on days with a relatively low risk of yielding a critical occupancy level according to the "original" historical admission schedule points to a second important use case for our simulation-based risk assessment: The identification of days with a potential for admitting additional patients without running into a risky situation.

6.   Discussion

Although the design of the counterfactual simulation is somewhat simplistic, it gives a clear indication that the simulation-based approach is able to recognize admission decisions that induce a high risk of leading to a critical occupancy level. However, the simple postponement strategy assumed in the counterfactual simulation—that a postponed admission is scheduled day-to-day until the patient can be admitted according to the simple traffic light policy—is not very realistic. Moreover, this postponement strategy is an immediate consequence of the myopic admission policy assumed in the counterfactual simulation: On each day, only the decisions to be taken on that day and its consequences on the next day are considered. In a more realistic setting, the planning horizon covers multiple days which allows for considering the admission decisions for each day of this time span. In such a setting, critical situations are anticipated (at least to a certain extent) by the multi-day risk assessment based on the Monte-Carlo Simulation, allowing to avoid postponements in most cases. Due to the fact that the risk assessment is intended to be performed on each work day means that the whole admission schedule can be adapted in case that a low-probability event such as the arrival of multiple emergency cases has occurred the day before or in case that the state of a patient has worsened and a longer LoS becomes more likely.

The last example highlights the relevance of having daily updated high-quality patient-specific LoS distributions (generated automatically and/or resulting from the validation and adaption step performed by the medical experts): If the worsened state of a given patient is reflected in her LoS distribution, this has an impact on the occupancy distribution on later days of the planning period. This, in turn may trigger the re-scheduling of an elective surgery that was scheduled later in the planning period. As a result, short-term postponements or the occurrence of critical occupancy levels may be avoided.

At the time of this writing, the approach presented in this paper is starting to be evaluated in the practical real-world planning process. Note that our case study ICU is currently not affected by COVID-19 cases which allows to use the XGBoost-algorithm with training data from pre-pandemic times. Anyhow, special conditions like a pandemic situation would demand for further adaptions. The first impression based on our prototypical implementation is that the process described in Figure 1 may indeed provide valuable support in ICU decision-making. An interesting side effect of the presence of a quantitative risk assessment is that it provides a somewhat "objective" basis for supporting the decision-making process involving multiple stakeholders from both the ICU and from the specialties scheduling surgeries with post-operative ICU stays who often have diverging objectives.

7.   Conclusion

In this paper we presented a framework for supporting risk-oriented decision-making in ICU management. In particular, this framework combines probabilistic Machine Learning for obtaining daily updated patient-individual LoS distributions reflecting the patients' health state and interventions using a Monte-Carlo Simulation for assessing the risk of reaching critical ICU occupancy levels. The LoS prediction is based on standardized and daily updated data involving 24 features describing the health state and the types and severity of interventions. Since these items form the basis for computing the so-called SAPS II and TISS scores that have to be reported by every German ICU, our approach should be applicable in principle in any German ICU. Results from the counterfactual impact analysis with historical data from 2019 of a medium-sized German ICU confirms the impression that our risk assessment tool—although relying on certain simplifying assumptions—can be effectively used to avoid reaching critical occupancy levels.

The Monte-Carlo-based risk assessment is embedded in a decision-making process emphasizing the role and expertise of the medical experts. First, they are given the possibility to adjust the LoS distributions according to their own state of knowledge, e.g. with regard to the LoS after a surgery given the comorbidities of a patient. Second, the decision-making process is not automated, but our approach is merely used to provide a quick risk assessment of decisions taken by human experts.

At the moment, we are starting to evaluate the

framework which involves a prototypical interactive software implementation in the real-world planning process of our case study ICU. After possible adjustments and changes in the user interface, we plan to conduct a study for evaluating our framework over a longer period of time and at different ICUs.

References

[1] J. Bai, A. Fügener, J. Gönsch, J. O. Brunner, and M. Blobner, "Managing admission and discharge processes in intensive care units," *Health care management science*, pp. 1–20, 2021.

[2] O. Moerer, E. Plock, U. Mgbor, A. Schmid, H. Schneider, M. B. Wischnewsky, and H. Burchardi, "A german national prevalence study on the cost of intensive care: an evaluation from 51 intensive care units," *Critical Care*, vol. 11, no. 3, 2007.

[3] O. Bittencourt, V. Verter, and M. Yalovsky, "Hospital capacity management based on the queueing theory," *International Journal of Productivity and Performance Management*, pp. 224–23, 2018.

[4] J. Bai, A. Fügener, J. Schoenfelder, and J. Brunner, "Operations research in intensive care unit management: a literature review," *Health care management science*, vol. 21, no. 1, pp. 1–24, 2018.

[5] G. Robinson, L. Davis, and R. Leifer, "Prediction of hospital length of stay," *Health services research*, vol. 1, no. 3, p. 287—300, 1966.

[6] T. Johnson, R. Mcnutt, R. Odwazny, D. Patel, and S. Baker, "Discrepancy between admission and discharge diagnoses as a predictor of hospital length of stay," *Journal of hospital medicine : an official publication of the Society of Hospital Medicine*, vol. 4, pp. 234–9, 04 2009.

[7] C. Chan, V. Farias, and N. Bambos, "Optimizing intensive care unit discharge decisions with patient readmissions," *Operations Research*, vol. 60, pp. 1323–1341, 12 2012.

[8] J. Barado, J. M. Guergué, L. Esparza, C. Azcárate, F. Mallor, and S. Ochoa, "A mathematical model for simulating daily bed occupancy in an intensive care unit," *Critical Care Medicine*, vol. 40, no. 4, pp. 1098–1104, 2012.

[9] L. Turgeman, J. May, and R. Sciulli, "Insights from a machine learning model for predicting the hospital length of stay (los) at the time of admission," *Expert Systems with Applications*, p. 376–385, 2017.

[10] I. Verburg, N. de Keizer, E. de Jonge, and N. Peek, "Comparison of regression methods for modeling intensive care length of stay," *PLOS ONE*, vol. 9, no. 10, 2014.

[11] A. Henzi, G.-R. Kleger, M. P. Hilty, P. D. W. Garcia, J. F. Ziegel, and o. b. o. R.-.-I. I. for Switzerland, "Probabilistic analysis of COVID-19 patients' individual length of stay in Swiss intensive care units," *PLOS ONE*, vol. 16, p. e0247265, Feb. 2021.

[12] A. Henzi, G.-R. Kleger, and J. F. Ziegel, "Distributional (single) index models," *Journal of the American Statistical Association*, vol. 0, no. 0, pp. 1–15, 2021.

[13] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, (New York, NY, USA), pp. 785–794, ACM, 2016.

[14] B. Misset and I. Ouanes, "Chapter 89 - scoring systems for comparison of disease severity in intensive care unit patients," in *Critical Care Secrets (Fifth Edition)*, pp. 603–608, Philadelphia: Mosby, fifth edition ed., 2013.

[15] J. Vincent, "Chapter 9 - general illness severity scores," in *Critical Care Nephrology (Second Edition)*, pp. 55–60, Philadelphia: W.B. Saunders, second edition ed., 2009.

[16] S. Jauk, D. Kramer, G. Stark, K. Hasiba, W. Leodolter, S. Schulz, and J. Kainz, "Development of a machine learning model predicting an icu admission forpatients with elective surgery and its prospective validation in clinical practice," *MedInfo*, pp. 173–177, 2019.

[17] L. A. G. Celi, R. J. Tang, M. C. Villarroel, G. A. Davidzon, W. T. Lester, and H. C. Chueh, "A clinical database-driven approach to decision support: Predicting mortality among patients with acute kidney injury," *Journal of healthcare engineering*, vol. 2, no. 1, pp. 97–110, 2011.

[18] J. Moran, P. Bristow, P. Solomon, C. George, and G. Hart, "Mortality and length-of-stay outcomes, 1993–2003, in the binational australian and new zealand intensive care adult patient database," *Critical Care Medicine*, pp. 46–61, 2008.

[19] J. Le Gall, S. Lemeshow, and F. Saulnier, "A New Simplified Acute Physiology Score (SAPS II) Based on a European/North American Multicenter Study," *JAMA*, vol. 270, no. 24, pp. 2957–2963, 1993.

[20] D. J. Cullen, J. Civetta, B. Briggs, and L. C. Ferrara, "Therapeutic intervention scoring system: a method for quantitative comparison of patient care," *Critical Care Medicine*, 1974.

[21] D. R. Miranda, A. de Rijk, and W. Schaufeli, "Simplified therapeutic intervention scoring system: the tiss-28 items–results from a multicenter study," *Critical Care Medicine*, vol. 24, no. 1, pp. 64–73, 1996.