Zorglogistiek

A Data-driven Decision Support System for Capacity Planning in Mental Healthcare

Providers of mental healthcare have been struggling with capacity problems for many years, in many cases leading to excessive waiting times for the patient-in-need.

D. Leca, B. van Ingen
Vrije Universiteit Amsterdam, Department of Mathematics, Amsterdam

R.J. Arntzen, R.D. van der Mei
Vrije Universiteit Amsterdam, Department of Mathematics, Amsterdam, Centre for Mathematics and Computer Science, Stochastics group

D. Moeke
HAN University of Applied Sciences, Research Group Logistics and Alliances, Arnhem

Abstract

Providers of mental healthcare have been struggling with capacity problems for many years, in many cases leading to excessive waiting times for the patient-in-need. This has raised the need for ways to make better use of the available capacity. To this end, it is crucial to develop models and tools to gain insight into the capacity bottlenecks, and to answer 'what-if' scenarios on the consequences and effectiveness of capacity decisions. Motivated by this, in this paper we address this problem by developing a data-driven Decision Support System (DSS) for a specific clinic of a Dutch mental health care provider. The DSS allows the management team to better understand (1) the characteristics and performance of their admission process, and (2) the consequences of capacity decisions in terms of service level and the occupancy rate of the clinics.

Introduction

About 40% of the Dutch population will face mental health problems at some point in their life [13]. In order to provide the right help and support, over 4 billion euros were spent on mental healthcare in the Netherlands in 2019 [14]. The causes and expressions of mental health problems vary widely, and as such, there are many different types of mental healthcare providers. Persons facing severe and complex mental health problems sometimes require admission to a mental health (psychiatric) inpatient unit for assessment and treatment. In 2020, over 5,000 Dutch adults were admitted to an inpatient unit.

Unfortunately, persons in need for specialized mental health care face long waiting times [16, 17]. As a result, there is a serious risk that they do not receive appropriate care in time. In order to prevent this, mental health care providers and health insurance companies have agreed on standards for acceptable maximum waiting times in the mental health sector, the so-called "Treeknorm". According to the Treeknorm, a person must be able to go for an intake interview with a mental health care provider within four weeks. The actual treatment should start ten weeks after the intake. However, at the time of writing of this paper (February 2022), for persons in need of specialized mental health care the Treeknorm is exceeded in 33% of the cases [18].

Mental healthcare providers have been battling capacity shortages for years. Moreover, since the COVID-crisis, the shortage has become even more problematic. Consequently, mental health care providers are diligently seeking ways to use their capacity more effectively. As we know from previous studies in the health care domain, capacity planning plays a key role in ensuring that a health care provider can respond in a timely and efficient manner to the level of demand experienced, see e.g. [1-5]. Furthermore, based on these studies, we conclude that health care providers could greatly benefit from a more data driven approach when it comes to capacity decisions. In our opinion, this is mainly due to the fact that still a large part of the capacity decisions is made without sound quantitative support.

In this paper, we address this problem by developing a data-driven capacity model implemented in a Decision Support System (DSS) for a specific clinic of a Dutch mental health care provider. The DSS allows the management team to better understand the (1) characteristics and performance of their admission process, and (2) the consequences of capacity decisions in terms of service-level and the occupancy rate of the clinics.

The presented DSS is based on a discrete-event simulation model. A simulation model is a representative model of a real-world process or system, which can be used to describe and analyze its behavior [6]. For the simulation of real-world service systems, discrete-event simulation (DES) is a commonly used approach. A DES-model can be defined as a mathematical model of a system that mimics that system's state changes at given points in simulated time. The state changes are "triggered" by events. There is a large body of scientific literature on the development DES-models for a variety of types of health care service systems (see e.g., [7]-[9]). However, to the best of our knowledge, this is the first paper in which a DES-driven DSS is presented specifically designed for the mental health care setting.

The remainder of this paper is structured as follows. In Section 2, the empirical context and the problem under consideration are discussed. Section 3 describes the results of a quantitative analysis of the arrival and service process. In section 4, the underlying DES-model is elaborated on. The functionalities of the DSS are discussed in section 5. Finally, this paper closes with conclusions and recommendations for further research.

Problem description

The organization under study is a Dutch mental health care provider in the Amsterdam region, that provides intramural treatment to clients with Serious Psychiatric Disorders (EPA). The treatment beds are spread across five clinics where clients are admitted on both compulsory and voluntary basis. Each clinic consists of multiple departments. Figure 1 presents a schematic overview of the main characteristics of the admission process.

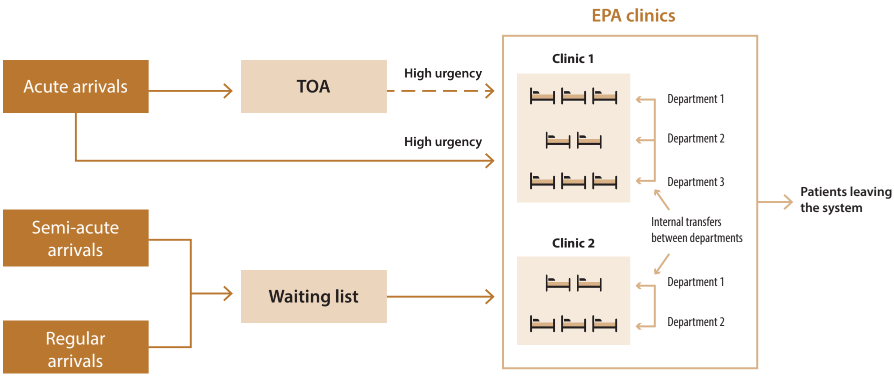

Figure 1 Overview of the admission process

At the start of the admission process, clients are being prioritized based on urgency (i.e., acute, semi-acute and regular clients). Consequently, acute clients are always admitted first. If no bed is available in the clinics for an acute arrival, the client will be temporarily admitted to the Temporary Transitional Admissions Department (TOA). This TOA serves the purpose of ensuring that acute patients can always be admitted as soon as they arrive (i.e., admission without waiting).

From an emergency perspective, a practical guideline is to always have at least three TOA-beds available. However, due to the pressing demand for mental healthcare, in practice both the clinics and the TOA are often fully occupied. In addition, the number of client discharges varies greatly over time, which makes it difficult to plan them in advance.

Given these challenges, the organization under study recognized the need for more data-driven insight and decision making. This recognition has prompted the organization to take the initiative to develop a DSS.

Data analysis

For the purpose of this study, a dataset containing the log data of 8,537 events (which took place between 2018 and 2020) for a total of 3,173 clients have been analyzed. The data mainly includes information on the arrival of clients and transfers between departments. For an impression of the content of the dataset, see Table 1. It shows the most relevant log data for two events, one new intake and one interdepartmental transfer.

Table 1 Example of available log data for two events

Variables	Example intake	Example transfer
Patient	2	2
Type of intake	New intake	Transfer
Movement from/to	intake at B2	from TOA to 3A
Intake on department	2B	3A
Date on waiting list	27-06-18	-
Start date intake	04-07-18	24-01-19
Transfer to (after)	-	4A
Duration of intake (days)	22	19
Urgence of intake	Semi-acute	-

For the DSS to be useful in practice, it must be ensured that the underlying simulation model is based on realistic assumptions. With this in mind, the following subsections discuss the results of an analysis of the arrival and service process.

Patient arrival process

Arrival rates

First, the demand patterns have been analyzed. Figures 2, 3 and 4 provide an overview of the total number of patient arrivals on a monthly, weekly and daily basis, respectively.

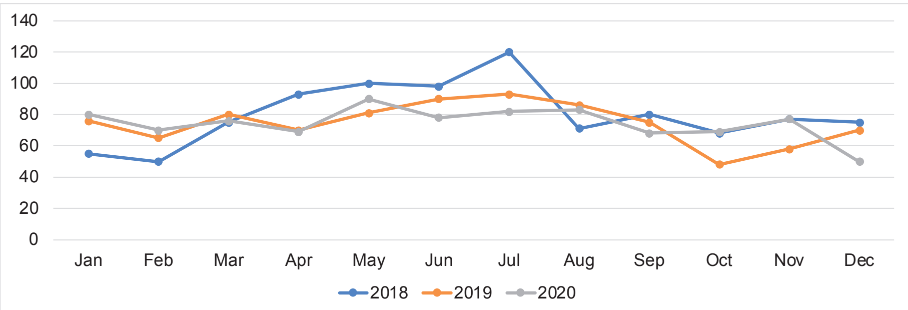

Figure 2 Total monthly arrivals per year.

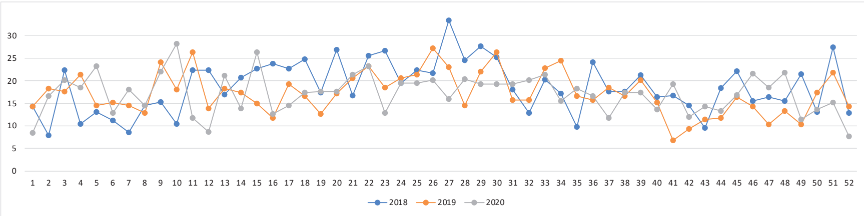

Figure 3 Total weekly arrivals per year.

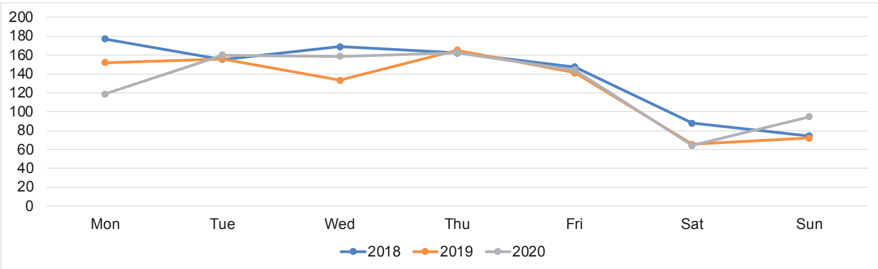

Figure 4 Total arrivals per weekday per year.

From Figure 2, it can be observed that arrival rates per month are fairly stable over the year. When looking at 2018, one could argue that there are slightly more arrivals during the summer, but this is less visible in the next two years. Also, the total number of arrivals per year only differ slightly from each other. When it comes to the weekly arrival rates (see Figure 3), there does not seem to be a clear pattern either. Figure 4 shows the total number of arrivals per day of the week. The weekend days seem to have a consistently lower number of arrivals than the weekdays.

Interarrival times

The interarrival times are the time spans between the arrival moments of two consecutive clients. In order to identify the underlying distribution of the interarrival times, a (visual) comparison of the real data to common probability distributions was performed. It was concluded that the interarrival times can be reasonably approximated by an exponential distribution. By way of illustration, in Figure 5 an exponential distribution is fitted on the empirical distributions of the interarrival times of one of the departments, and for the three urgency levels.

Based on this, we assume that the arrival processes are Poisson processes, which is a common modeling assumption for health care processes (see e.g., [10]), based on the memoryless property of the exponential distribution.

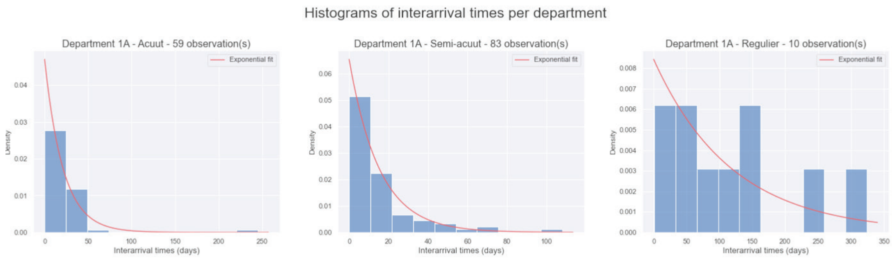

Figure 5 Frequency distribution of the interarrival times for the three urgency levels.

Service process

Another key model parameter is the probability distribution of the length-of-stay (LoS). To find the best fitting probability distribution, a comparison to the exponential, log-normal and gamma distributions was performed as these distributions are often used for modeling the LoS of hospital patients (see e.g., [11, 12]). To illustrate, Figure 6 below gives the empirical distribution of the LoS in a given department and the best-fitting probability distributions. For all departments (including the TOA), and each of the three urgency levels, the gamma distribution appears to give the best fit.

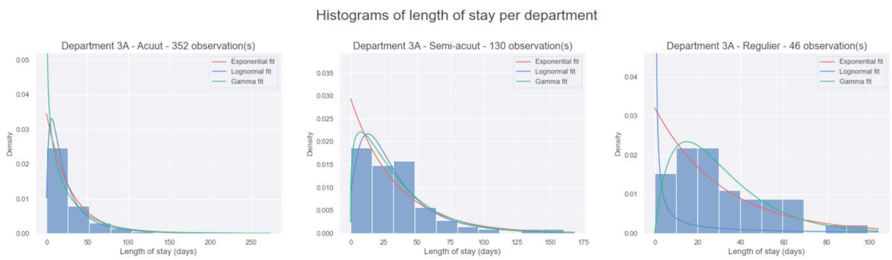

Figure 6 Frequency distribution of the length-of-stay for the three urgency levels.

Discrete-event simulation model

As mentioned in the introduction section, the underlying model of the DSS is a DES-model. In this section, the main characteristics of the model are discussed.

Entities

The DES-model consists of the following entities:

- **Patient –** representing the clients in the system. Clients are generated by the arrival processes. They enter the system on arrival and exit the system upon (the final) departure. Each client has a particular urgency-level: *acute, semi-acute or regular*.
- **Hospital bed** – representing the available beds in the system. The number of beds in every department is set at the beginning of the simulation. Each bed in the system can have one of the following statuses: *OpenFree* (free to accept a patient) or *OpenBusy* (currently occupied by a patient).
- **Department** – representing a homogeneous collection of beds. It is assumed that a department has a single type of bed which is suitable for any type of client in terms of urgency-level. For each department, the LoS is modeled as an urgency-dependent gamma distribution. Each department manages two types of queues:
    1. **QueueInternalTransfer** – representing the queue of clients which are waiting for an interdepartmental transfer (including TOA-clients). The transfer of TOA-clients is prioritized over all other clients in the internal queue. Non-TOA-clients are assigned by order of arrival (i.e., first come first served).
    2. **WaitingList** – representing the queue of clients outside the system waiting for admission. Clients are assigned based on urgency-level; within each urgency class, priority is assigned by the order of arrival.
- **Clinic** – representing a collection of departments.
- **TOA** – representing the department for Temporary Transitional Admissions. In essence, TOA is a department that handles the acute clients who could not be admitted directly to one of the departments due to lack of free beds.

Simulation process

The DSS helps to gain insight into effects of altering the capacity in a selected clinic. However, the resulting patient flow is also determined by the availability of beds in the TOA, which is the result of a stochastic process including patients for other clinics, the so-called 'external flow'.

For that reason, we set up the simulation process as a combination of two patient flows:

- **A flow of "regular" clients through the system**. For each client, the LoS is sampled from its empirical distribution upon arrival. The path that a client follows through the system depends on the availability of beds in the departments. For example, if an acute client enters the system when all beds in de clinic are occupied, then the client will be directed to the TOA. A client who is due to leave a department and enter another department is delayed in the department of origin until a bed becomes available at the destination. This flow is referred to as the *foreground* flow.

- **An external flow of patients that enter and leave the TOA.** Recall that the purpose of this flow – also referred to as the *background* flow - is to mimic the reduction of available beds in the TOA due to the shared nature of the location. These patients enter the TOA, spend a certain amount of time (drawn from a distribution that is computed from historic data by filtering out the patients within the network), and exit the system.

Assumptions

For the DES-model, the following assumptions are made:

- Within each department the available beds are considered interchangeable.
- For each client entering the system, only one eventual transfer is considered.
- It is assumed that the arrivals and the service times are time-homogeneous, i.e., seasonality and time-dependency are not considered.
- The arrival processes are independent and Poisson and the LoS is gamma-distributed.
- All departures from a department scheduled on day d are handled at 8:00 in the morning of the next possible opportunity (day d if they are due on day d before 8:00, d + 1 otherwise).
- Urgency levels (i.e., acute, semi-acute and regular) are assigned when a client enters the system and do not change over time.
- The number of beds in a department (i.e., the maximum capacity) is fixed over time.

We emphasize that these assumptions are mainly made for ease of the presentation, and can be easily relaxed if required.

Model validation

The model has been validated by comparing the performance results to the simulations. Our simulation experiments showed that the simulation-based results are very close to performance metrics observed in our out-of-sample data set. The details of the validation experiments are omitted here for compactness of the paper.

Decision Support System

In this section we present the DSS which can be used to interactively explore a wide variety of 'what-if' scenarios by changing the input parameters. The DSS requires the user to choose the clinic of interest. After the desired clinic is selected, the user sees the overview of tabs and input choices as presented in the screenshot in Figure 7.

For each clinic, the user can change the settings for four different elements. Namely, as can be seen in the tabs in Figure 7, the user can select input parameters for the *Departments* of the clinic, for the *TOA* and for the clients' *Transfers*. Lastly, settings are required for the *Simulation*, for which the input can be found by scrolling down in the DSS. We discuss the four setting elements here below.

Department settings – As can be seen in Figure 7, the input tab for the departments consists of multiple sections or columns. The sections are the following:

1. **General setup** – For the general setup, the user is required to input the number of beds available in the selected department. Also, the expected proportion of rejected patients has to be given here. This is a proportion of patients that are denied resources for any reason (other than the clinic being full). An example of why this may occur is that a patient has been admitted into another department/clinic. This option can be used to limit arrival rates without manually adjusting each slider for the arrival process.
2. **Arrival process** – Here, the user can adjust the arrival rate for patients of each severity (acute, semi-acute and regular). The arrival rates used in the model are counted in patients per day. Changing the arrival rates also immediately shows the distributions of the number of monthly arrivals of patients of each severity in a bar chart directly below the sliders. This gives the user an indication of how their choices influence the model.
3. **Length of stay** – This section allows the user to make changes to the probability distribution of the length of stay of patient of all severities. Both the means and standard deviations can be adjusted. This allows the user to see how making the treatment better/faster or more predictable, thus lowering the variability i.e. the standard deviation, would influence the performance indicators used in the simulation.

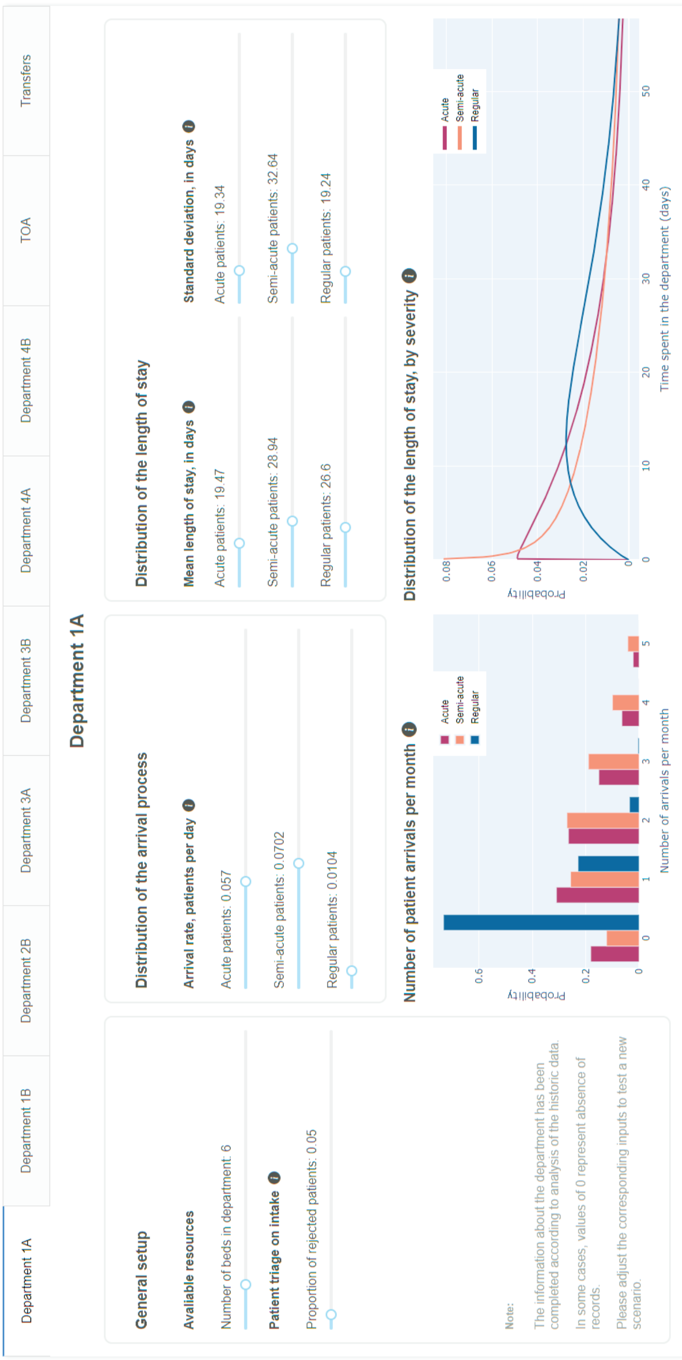

Figure 7 The input-screen for the department settings

TOA settings – The input tab for the TOA is similar to the department setting with two main differences. The page is shown in Figure 8 below.

1. **General setup** – The user is asked to give a maximum length of stay in the TOA, besides the inputs which are also required for the departments. At the moment this value is only used as a benchmark value to report the number of patients exceeding the desired maximum number of days in the TOA. In the code however, there is the possibility to extend the design of the TOA with, for example, speeding up the admission to department for the patients that exceed the maximum length of stay, by advancing the departures scheduled in the following days.
2. **Arrival rate** – The user is required to input the arrivals in the TOA that are not from the selected clinic. Since the TOA is a department that is shared between clinics, we cannot simply allocate all beds to the clinic we are simulating. Instead we let the user tell the model how often patients arrive via one of the other clinics.

> See figure 8: The input-screen for the TOA-settings , page 22

Transfer settings – The transfers tab redirects the user to a table in which the movement of patients can be described. In this table, the cells indicating transfers from one department to another are all variable. Each cell represents the probability that a patient is transferred from one department to another. The final column represents the portion of patients that are dismissed from the department. These values are not variable and are programmed to make sure the total probabilities of the transfers add up to 1. A screenshot can be found in Figure 9.

Table 2 The input-screen for the transfer-settings

	Department							
From/To	**1A**	**1B**	**2B**	**3A**	**3B**	**4A**	**4B**	**Discharged**
Department 1A	0	0.02	0.09	0.49	0.06	0.09	0.01	0.24
Department 1B	0.04	0	0.2	0.04	0.18	0.03	0.2	0.31
Department 2B	0.03	0	0	0	0.16	0	0.19	0.62
Department 3A	0.14	0	0	0	0	0.35	0.03	0.48
Department 3B	0.01	0.04	0	0.01	0	0	0.27	0.67
Department 4A	0.01	0.02	0.01	0.05	0.01	0	0.01	0.89
Department 4B	0	0.05	0	0.04	0.02	0.02	0	0.87

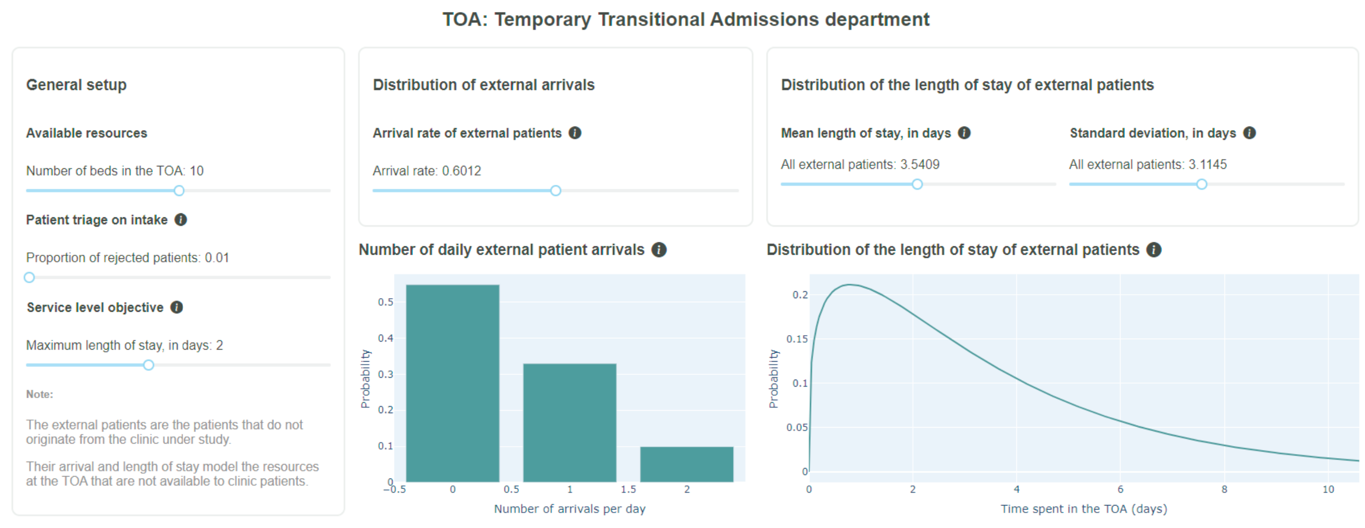

Figure 8 The input-screen for the TOA-settings.

Simulation settings – Finally, the user is being asked if they want to run the simulation using distributions that were fit on the data, or if they want to use values that were bootstrapped from the data directly. Figure 10 shows the selection options in the DSS.

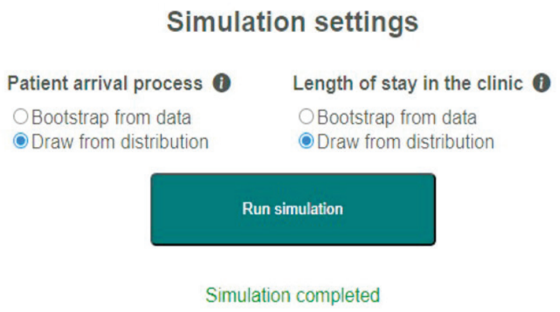

Figure 9 Screenshot of the input page for the simulation settings.

Simulation output

After the simulation is done, the output is presented. Figure 10 gives a screenshot of an example of the outputs. As can be seen in Figure 11, the DSS reports performance indicators from different perspectives:

System view – The first tab in the simulation outputs shows the key performance indicators from the point of view of the clinics a whole. This means that it takes all departments into account at the same time. The page can be divided into three sections or columns, that is, the tables and graphs below them correspond to the same output. These are the following:

1. **Waiting list** – The first KPI's one sees are the waiting list KPI's. The minimum, maximum, mean, median and standard deviation of the total number of people on the waiting list is given in a simple table together with the 75%-, 90%-, and 95%-percentiles.

- Below that, three different graphs can be chosen to inspect. First, there is a tab 'Total' showing the total number of patients on the waiting list over time. The second tab gives a graph depicting the number of patients awaiting inter-departmental transfers over time. Lastly, the user can choose to see the graph depicting the number of patients staying in the TOA waiting to be transferred to the clinic. These graphs are not to be taken as deterministic events in time, but more as an indication of the general trend of the number of patients on waiting list.

2. **Clinic bed occupancy** – The middle table shows the clinic bed occupancy containing a number of key indicators. The minimum, maximum, mean and standard deviation of the bed occupancy over all the beds in the clinic are reported numerically. These values are supported by the histogram below the table, depicting the distribution of the bed occupancy of the beds during the duration of the simulation.
3. **TOA bed occupancy** – Lastly, similar to the clinic bed occupancy, we report the TOA bed occupancy. These metrics can be used to get an idea of if there were enough beds in the TOA as a buffer for the clinic being full.

Department view – From the department point of view, the output is somewhat similar to the system point of view. The main difference is that the statistics are split *per department*. Also, since the TOA does not belong to a specific department, there is no distinct section for TOA statistics.

What we do report for each department separately is the waiting list KPI's for patients waiting to enter the selected department. The user can also find two graphs showing the course of the number of patients waiting to be admitted into the department over time. The first graph shows the number of new patients waiting on the waiting list. The second graph depicts how many patients are waiting on a transfer into the department either from the TOA or from another department.

The user can also find a small table for the department bed occupancy KPI's together with a histogram showing the distribution of the occupancy of all beds in the department during the simulation. Figure 12 shows how the KPI's are presented in the DSS.

Figure 10 Screenshot of the System performance dashboard.

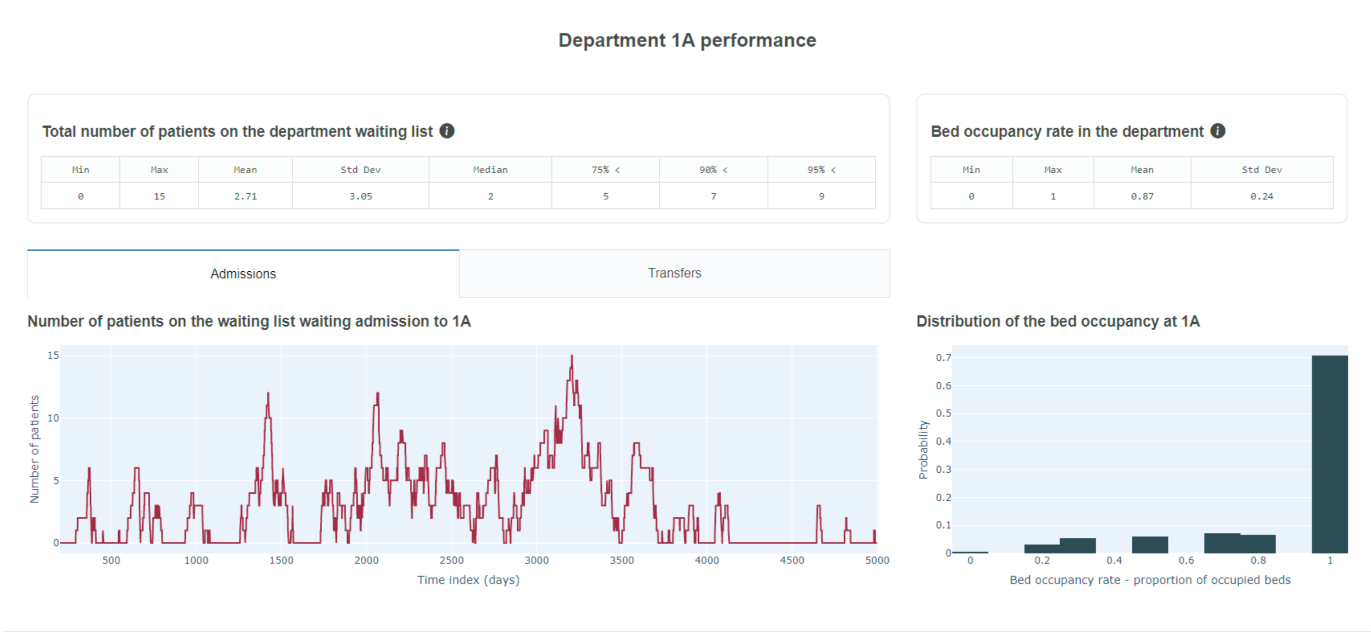

Figure 11 Screenshot of the Department performance dashboard.

Patient view – The performance indicators from the patients' point of view are presented in a different way. Since we mainly report percentages here, it was decided to use pie charts, since they offer an intuitive way of showing proportions. All pie charts are supported by a short text for clarity. The patient view is also split into three sections:

1. **General summary** – Here we show two pie charts. The first simply shows the number of generated patients for each severity that arrived into the system during the simulation. The text next to the second chart reports on the number of patients from whom the total time spent in to TOA is known. The chart itself gives an idea of how many of these patients stayed longer in the TOA than the desired maximum stay target time.
2. **Waiting list** – For the waiting list statistics, we report what percentage of patients who were not rejected by triage, had to wait on the waiting list before being admitted into the clinic. Secondly, a pie chart shows how many of the patients that were redirected to the waiting list have entered a department before dropping out of the waiting list for whatever reason (e.g., being admitted into a different clinic). Lastly, a small table is presented with the 50- and 90% percentiles for the number of days patients spent on the waiting list. A third column shows what percent of patients on the waiting list exceeded 100 days waiting.
3. **Rejections** – The final three pie charts show the rejection rates. The first chart informs on the number of acute patients that had to be directed to admission via the TOA in the simulation. The second chart makes a distinction between patients that were let into the TOA, those who were rejected by triage at the TOA, and those who were rejected because the TOA was full. The final chart reports the number of (non-acute) patients that were rejected from entering the clinic on arrival by the triage process.

> See figure 12: Screenshot of the Patient performance dashboard, page 28

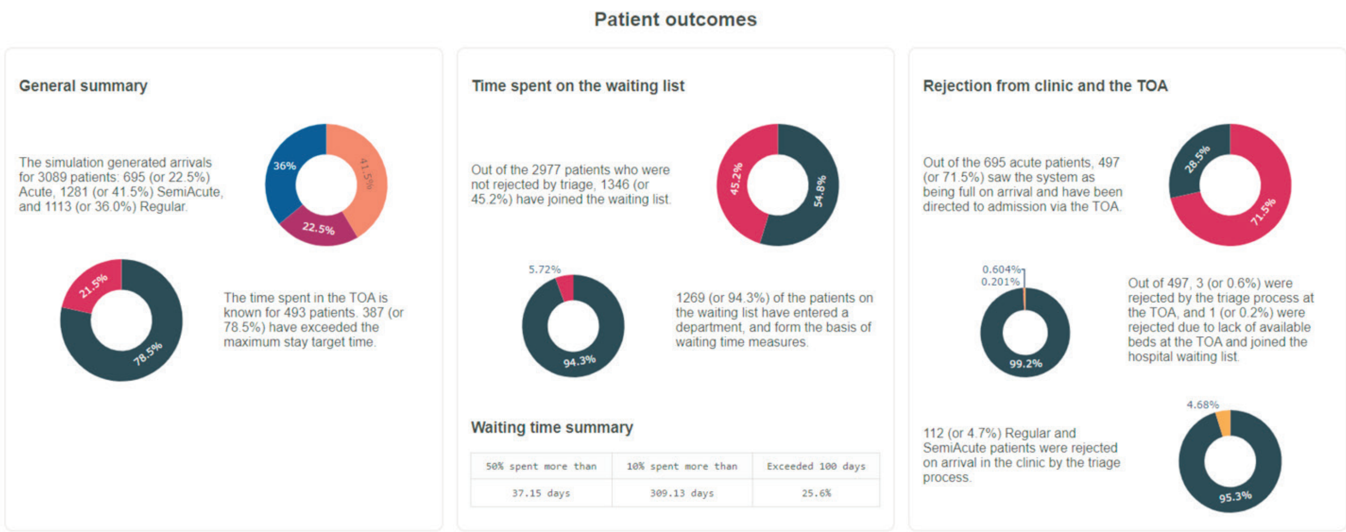

Figure 12 Screenshot of the Patient performance dashboard.

Conclusions and discussion

This paper is motivated by the need to solve capacity problems in the mental healthcare sector. We propose a performance and capacity model that enables providers to answer important 'what-if' questions, which makes it possible to assess the effectiveness of managerial decision *a priori*, before to their actual implementation. For this, the model has been implemented in a DSS, with a simulation engine and a user-friendly dashboard interface. The tool provides providers and decision a powerful means to understand the consequences of managerial decisions with respect to waiting times, capacity bottlenecks and occupancy rates of the clinics.

This paper can be seen as a significant first step in facing capacity problems in mental healthcare, and leaves many challenges for follow-up research. Let us address some next steps. First, the first assumption about the interchangeability of the available beds may be unrealistic in many situations, call for a more accurate model description, which comes at the price of additional model complexity. Second, the assumption time-homogeneity may be relaxed to better describe situations where arrivals processes and service times do exhibit seasonal behavior. Third, in the model the routing probabilities of the patients' flows through the system are assumed to be known (see Figure 9 for an illustration). However, in practice this assumption may be unrealistic. An interesting venue for further research is to use the concept of process mining [19] to get a better understanding of the patient flows through the system.

Acknowledgement: the authors would like to thank Bianca van Daalen for providing us the data and for our valuable discussions about the topic.

References

[1] Bekker, R., Moeke, D., & Schmidt, B. (2019). Keeping pace with the ebbs and flows in daily nursing home operations. *Health Care Management Science*, *22*(2), 350-363.

[2] Clapper, Y., Bekker, R., Moeke, D., & Eijkelkamp, W. (2021). De meerwaarde van datagedreven capaciteitsplanning in de thuiszorg. *Logistiek+ Tijdschrift voor Toegepaste Logistiek* (special edition), 94-103.

[3] Laheij, G., Moeke, D., Westerman, R., Hertman, F., & Migchielsen, A. (2019). Onderzoek naar de logistieke inrichting van een acute opname afdeling bij het CWZ: hoeveel bedden zijn er nodig? *Logistiek+ Tijdschrift voor Toegepaste Logistiek* 8, 47-61.

[4] Markhorst, B., Zver, T., Malbasic, N., Dijkstra, R., Otto, D., van der Mei, R., & Moeke, D. (2021). A data-driven digital application to enhance the capacity planning of the covid-19 vaccination process. *Vaccines* 9, 1181-1193.

[5] Moeke, D., & Bekker, R. (2020). Capacity planning in healthcare: finding solutions for healthy planning in nursing home care. In *Integrating the Organization of Health Services, Worker Wellbeing and Quality of Care* (pp. 171-195). Springer, Cham.

[6] Banks, J. (1998). Principles of simulation. *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*, 3-30.

[7] Jacobson, S. H., Hall, S. N., & Swisher, J. R. (2006). Discrete-event simulation of health care systems. In *Patient flow: Reducing delay in healthcare delivery* (pp. 211-252). Springer, Boston, MA.

[8] Günal, M. M., & Pidd, M. (2010). Discrete event simulation for performance modelling in health care: a review of the literature. *Journal of Simulation*, *4*(1), 42-51.

[9] Jun, J. B., Jacobson, S. H., & Swisher, J. R. (1999). Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society*, *50*(2), 109-123.

[10] De Bruin, A.M., Bekker, R., Van Zanten, L., & Koole, G.M. (2010). Dimensioning hospital wards using the Erlang loss model. *Annals of Operations Research*, *178*(1), 23-43.

[11] Ruffieux, C., Marazzi, A. & Paccaud, F. "Exploring models for the length of stay distribution". In: Sozial-und PrÅNaventivmedizin 38.2 (1993), pp. 77–82.

[12] Marazzi, A., et al. "Fitting the distributions of length of stay by parametric models". In: Medical care (1998), pp. 915–927.

[13] Website of Rijksoverheid, www.rijksoverheid.nl.

[14] Website of Nederlandse Zorgautoriteit, ww.nza.nl.

[15] https://www.staatvenz.nl/kerncijfers/zorg-met-verblijf-ggz-indicatie.

[16] https://www.pwc.nl/en/insights-and-publications/services-and-industries/public-sector/social-determinants-of-health/waiting-times-in-mental-healthcare.html,

[17] https://www.igj.nl/zorgsectoren/geestelijke-gezondheidszorg/wachtijden-in-de-ggz

[18] https://www.staatvenz.nl/kerncijfers/wachttijd-gespecialiseerde-ggz-overschrijding-treeknorm-en-gemiddelde-wachttijd

[19] Van der Aalst, W., Process Mining – Data Science in Action. Springer. 2018.