Using agent based modelling to analyze student behaviour and space issues in Faculty of Science, University of Colombo

Lasal Hettiarachchi, Udesh Senevirathne†, Janith Abeyrathne†

[1]Department of Mathematics, University of Colombo, Kumaratunga Munidasa Mw , Colombo, 00300, Western Province, Sri Lanka.

Contributing authors: 2019s17363@stu.cmb.ac.lk; 2019s17478@stu.cmb.ac.lk; 2019s17298@stu.cmb.ac.lk;
†These authors contributed equally to this work.

Abstract

Agent-based modelling (ABM) is a computational modelling technique used in many domains to simulate complex systems made up of autonomous agents, commonly represented as individuals, and their interactions with their surroundings. This study uses ABM and the AnyLogic simulation platform to analyse student behaviour and space-related difficulties at the University of Colombo's Faculty of Science.

The primary goal of this study is to quantify the scarcity of locations available for students to spend their free time. The research intends to detect peak hours of space constraint and measure the success of an upcoming building construction project in addressing these difficulties by modelling the movements and preferences of individual students inside the AnyLogic framework.

Subsequently, the data that was gathered was thoroughly analysed to provide significant insights on optimising space utilisation within the Faculty of Science. The AnyLogic simulation platform enables a complete investigation of agent behaviours and preferences, allowing for a thorough knowledge of present spatial restrictions and their impact on student experiences.

The ultimate goal of this project is to improve the overall learning environment by resolving space constraints and improving the quality of student life, with AnyLogic serving as a strong simulation and analytic tool.

Keywords: Agent Based Modeling, Emergent Behaviour , Individual Based Models, Systems Dynamics Modelling , Mathematical Modelling , Anylogic

1 Introduction

The quality of the learning environment is critical in moulding students' educational experiences in academic institutions around the world. The availability and utilisation of physical spaces for students' leisure and study activities are significant factors determining this experience. The Faculty of Science at the University of Colombo, like many other educational institutions, faces spatial challenges that limit students' ability to locate appropriate locations for relaxation, study, and collaboration. This study investigates these issues and seeks solutions through the use of an innovative computer modelling technique known as Agent-Based Modelling (ABM). We investigate student behaviour, space limits, and upcoming infrastructure developments in collaboration with the advanced simulation platform AnyLogic to acquire insights into optimising space utilisation inside the Faculty.

The need for more adequate venues for students to spend their spare time is a widespread problem across colleges, resulting in overcrowding during peak hours and decreased student well-being. This study addresses this issue through a data-driven approach, utilising ABM and AnyLogic to mimic the complex interplay of student movements, preferences, and space availability. The main focus of this study is to quantify the spatial restrictions that are present for each of the subjected study areas, time-based analysis to identify the peak hours of resource utilization and finally understand how new infrastructure would shift the paradigm. This study seeks to give practical insights for improving the overall learning environment within the Faculty of Science by quantifying the spatial deficit, identifying peak hours of space shortages, and analysing the possible impact of a building development project.

1.1 Research objectives

This research is motivated by a set of clear and ambitious goals targeted at tackling the critical issue of space management within the Faculty of Science at the University of Colombo. Our goals are broad, encompassing both quantitative and qualitative aspects of space optimisation within the capabilities provided by anylogic simulation software under certain assumptions, with the overall goal of improving the learning environment and student well-being. The following are the primary goals of this research: 1. Quantify the Lack of Space Issue:- The primary objective of this research is to accurately measure the severity of the space problem in the Faculty of Science. By understanding the extent to which students are affected, the we can take informed measures.

2. Identify Peak Hours of Space Shortage:- Analyzing the data to pinpoint the specific hours of the day when the lack of space is most felt by the student body.

3. Test the New Building's Effectiveness:- The ongoing construction of a new building to address the space issue holds potential solutions. Through simulations, the model will assess whether the new facility can effectively serve to the student population.

1.2 Method

Our research aims to quantify the spatial constraints faced by the University of Colombo's Faculty of Science, identify critical time periods of space scarcity, assess the impact of an ongoing construction project, and ultimately optimise the utilisation of available spaces to improve the overall student experience. In the parts that follow, we delve into the organised approach that leads our analysis, using Agent-Based Modelling (ABM) and the adaptable AnyLogic simulation platform as our principal tools, and we explore the methodology behind this specific study.

1. Data Collection: - Academic timetables for students from each academic year will be collected to create representative agents.

2. Agent-Based Modeling in Anylogic: - The agent-based model will comprise of individuals with unique properties that will define them based on their academic years and their objectives. Each agent will have its unique academic timetable.

3. Rules and Behavior: - The behavior of individual agents will be influenced by their academic timetables. During their free time between lectures, students will attract towards open spaces and common areas. Their choices will be shaped by factors like closeness, capacity, and personal preferences.

4. Environment : - The faculty premises, including open spaces, common areas, and the library, will be represented in the model. AnyLogic, a simulation software, will be used to create the environment where agents interact and navigate.

5. Data Analysis and Visualization: - AnyLogic's built-in visualization tools will be used to create visual representations of the student behavior and space usage within the faculty premises.

1.3 Expected outcome and significance

The agent-based model developed in this research will provide a thorough understanding of student behavior around open spaces during peak hours. The simulation will offer valuable insights into how students interact with available spaces, helping the us to identify periods of high crowds and areas for improvement.

Additionally, by virtually testing the new construction, the model will assess its effectiveness in resolving the space issue.

The outcomes of this research project hold significant implications for the Faculty of Science . By optimizing space utilization, the university can enhance the overall student experience, create a more appealing learning environment, and promote academic success.

2 Conceptual Model

The following section describes the necessity of the model and the implementation of the proposed model.

2.1 Overview

2.1.1 Purpose

The primary objective of this research is to assess the existing space allocation within the University of Colombo, focusing on the Faculty of Science. A central concern addressed by this study is the critical issue of insufficient space within the university campus. This research aims to delve into the intricate details of how resources are currently distributed and utilized within the Faculty of Science, and subsequently, to project the potential effects and benefits brought about by the construction of the new building that is presently underway. Through a comprehensive analysis, this study seeks to provide valuable insights into optimizing resource allocation, enhancing spatial efficiency, and forecasting the trans-formative impact of the upcoming infrastructure development on the university's overall functioning.

2.1.2 State variable and scales

Although we didn't use anykind of a state diagram , there were clearly few states an student was in while he or she is inside the model. An student is spending their free time , and they attend lectures and they take lunch in dining areas, interacting with each other.

Central to the model's architecture is the representation of the environment, intricately mirroring the physical layout of the Faculty of Science. The scale of this environment is intricately determined by the dimensions of the faculty itself. Within this encompassing representation, diverse elements are encapsulated, ranging from lecture halls and open spaces to essential infrastructure elements such as pathways and roads.

The following section discusses the parameters pertaining to the student agent type that uniquely identifies the agents.

1. **ishungry**: There are a fraction of students who'd go to eat breakfast before their lectures starts. To model this scenario we introduced the parameter called "ishungry". It is a boolean variable. We assigned each student's "ishungry" to be true with a chance of 0.1 as they enter the university premises.
2. **independentStudy**: There is a clear distinction between choosing a library over other open spaces to indicate that a student prefers independent study."indendentStudy" is yet another Boolean variable indicating a student's study preference. Based on a probability we assigned it to be True.
3. **distanceFromHome**: Another significant factor we identified that affects student's behaviour is the distance from their residencies. This will affect their decisions on whether to stay in the university till evening or leave in the afternoon to head home. We assigned a parameter called "distanceFromHome" to each student and let it be a normal distribution with mean and variance. This parameter was used in pedestrian flowcharts to influence student's decisions.
4. **nextLecture**: If a student from a particular department/program has multiple lectures (maximum 2 in most cases) to attend it is less likely for that student to spend a long time in the open space and for the library it is even less likely .

"nextLecture" is another Boolean which can be called in the model to make such students to spend less time or avoid such behaviour.

2.1.3 Process overview and scheduling

Before diving into the model-specific details we need to make one important remark. This project was initially designed as an agent-based model. Where each individual has it's own set of characteristics that will decide it's behaviour in the system. An important (if not the most important) characteristic is the timetable-based behaviour of each agent (student). We wanted to model it in such a way that we would be able to monitor each student's routine without disturbance until that agent is out of the system. However, due to the limitations of the analogic software (35 different agents or processing blocks), we had to give up on that approach and use the tables of students to make estimations.

The key to understanding our model is to understand the pedestrian library of the Anylogic software. As the name suggests, it is a library that is used to model pedestrians. Although it bears the name pedestrians, this library can be used to model which involve crowds other than ordinary pedestrians. Some examples of this can be found in the public models of Analogic software.

Although there are numerous components of the library that are worth looking into we shall focus on the ones we have utilized in our model. For each of them first, we will introduce the component and then what we did with each in our model.

1. **Ped Source**

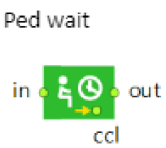
This Generates pedestrians. It is usually used as a starting point of the pedestrian flow. Can produce pedestrians of a custom pedestrian type with arbitrary flow intensity. This block also can be used to generate groups of pedestrians. Allows defining multiple criteria for group creation — group size, form, group arrival rate, pedestrian interarrival delay, etc. There are multiple of ways to define when and how many pedestrians (or groups of pedestrians) should be generated Apart from that it allows us to control where the pedestrian should appear. In our project, we have used x of such sources. 1 of them was used to model the students appearing at the entrance of the university. We modelled students to arrive in groups of size 2 at maximum. All the others were used to model the students departing from lecture halls situated in various departments. Each of them is used to form groups of students of sizes between x and x. Here we let the students form groups of size 1 to 5. We used a rate schedule to tell the sources how many students it should generate in given time intervals. These time intervals and rates were determined by making estimates based on the Time tables of 1st year, 2nd year and 3rd-year general degree students time tables. The timetables and the Excel sheet of statistical estimates xare found in supplementary documents.

2. **Ped to go**

Causes pedestrians to go to the specified location. Location can be defined by a number of space marking elements in pedestrian library. The pedestrian has to reach a given point, or any point of the specified line, or node. Pedestrians find path to the specified destination. We used these blocks to direct the pedistrians towards the open spaces. Namely

(a) Library
(b) Information & Learning Center (ILC)
(c) Open space in front of the Dean's office (referred to as Deanfront)
(d) New building
(e) Canteen of the University of Colombo School of Computing (UCSCC)
(f) Science Canteen

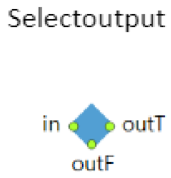
3. **Ped wait**



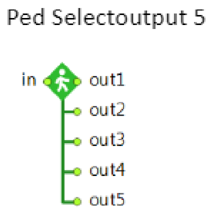
Causes pedestrians to go to the specified location and wait there for a specified period of time. Pedestrians can wait for some specified amount of time, calculated from different events (e.g. reaching the waiting point, entering the area, or even manually), or until the user manual calls the block function free().

This block is important for the model for two reasons. We could model the capacity of each open space by limiting the number of students it could serve at a given time. Based on our knowledge and external sources we estimated the number of students that can be held in each open space mentioned. The values are estimated under the assumption of the existing available seating capacity data which may be subject to change.

4. **Selectoutput**



Routes the incoming agents to one of the two output ports depending on (probabilistic or deterministic) condition. The condition may depend on the agent as well as on any external factors.



Functions pretty much the same as Pedselectoutput , only difference being it sends incoming pedestrians to one of several (up to five) processes depending on specified ratios or conditions. Each process has to be defined with a flowchart connected to one of the block's output ports.The block works in three modes: Probabilities, Conditions, or Exit number.Pedestrians spend zero time in the PedSelectOutput block.

5. **Ped Sink**

Disposes of incoming pedestrians. Is usually used as an endpoint of the pedestrian flow.

Along with above motivations several events are defined in the model for transitions. In Anylogic , Event is the simplest way to schedule some action in the model. We used this feature to model several important characteristics of the student population. The events in the conceptual model are defined as follows:

(a) **noon**: This event causes the students to leave their open spaces and move to a where they can take lunch. We could see how this sudden increase in student flow will affect the capacity of dining areas.

(b) **hourlydrop**: This even was designed to model a certain behaviour that can be seen in open spaces. It's usually the case that a portion of students will leave an open space as clock time reaches a full hour mark such as 8.00 AM, 9.00 AM, 10.00 AM,etc.Via this event, we free students from waiting blocks by calling the builtin command “free()”. And we free 0.2 portions of students from each open space at each hour.

2.2 Design concepts

2.2.1 Observation

At a global level, we oversee various metrics to track the distribution of specific activities among the agents. Additionally, we closely monitor the population of agents present in particular locations, observing the count of agents at these places during different time intervals. The driving force behind the model’s dynamics lies in the interactions between agents and their environment. Agents engage in a variety of interactions, such as sharing resources like computers, as well as engaging with other agents who may be their friends or colleagues. These interactions are quantified through statistical measures, allowing us to assess how effectively resources are being allocated within the environment at any given moment. In essence, our focus is on maintaining a comprehensive understanding of how agents interact with each other and their surroundings, leveraging statistical analysis to evaluate the efficiency of resource utilization within this dynamic system.

Within the model at the global level, we monitor the following statistics: the proportion of the total number of agents carrying out specific activities. We also monitor the crowd at specific places (the number of agents at a given place at any given time). With respect to Interactions, the dynamics of the model are driven by the interactions of the agents with their environment. Agents share resources (e.g. computers), they also interact with other agents who are their friends or colleagues. These dynamics are modeled in terms of statistics to identify the optimal resource allocation with the environment at any given time.

2.3 Details

2.3.1 Initialization

The initialization of the model relies on spatial data of the study area. Some of the parameter values require reviewing relevant literature. Where ranges are used, it is appealing to apply a distribution to model certain scenarios.

Parameters and variables summarized

Parameters	Details	Value
Arrival rates	Student's arrival rates from each source	Estimated from the student timetables
Waiting times at the library, ILC ,deanfront & dining areas	The time a student spend at a given resource	Assumed to have a normal distribution with varying mean and variation
Capacity of the Library	Available seating capacity inside the library	500
Capacity of the ILC	Available seating capacity in the Ground floor of Information and Learning center	48
Capacity of the Deanfront	Available seating capacity in the open space infront of dean office	80
Capacity of the UCSC Canteen	Available dining capacity in the Canteen of UCSC	150
Capacity of the Science Canteen	Available dining capacity in the Canteen of Science faculty	60
Capacity of the Dining area of New building	Available dining capacity estimated for the new building's dining area which is currenty under construction	400
Capacity of the Study zone of New building	Available study area capacity estimated for the new building's dining area which is currenty under construction	400
ratemodifier	One of the major considerations was to decide what portion of students will choose to remain in the university after the lecture. We introduced a parameter called "ratemodifier" to address this question. It's a variable that can take values between 0 and 1, and we implemented it so that we could control the input rates of each source before running the simulation. There was no way for us to surely take a value (without doing a detailed survey) but 0.5 seemed to be a good choice from our experience.	0.5

2.3.2 Inputs

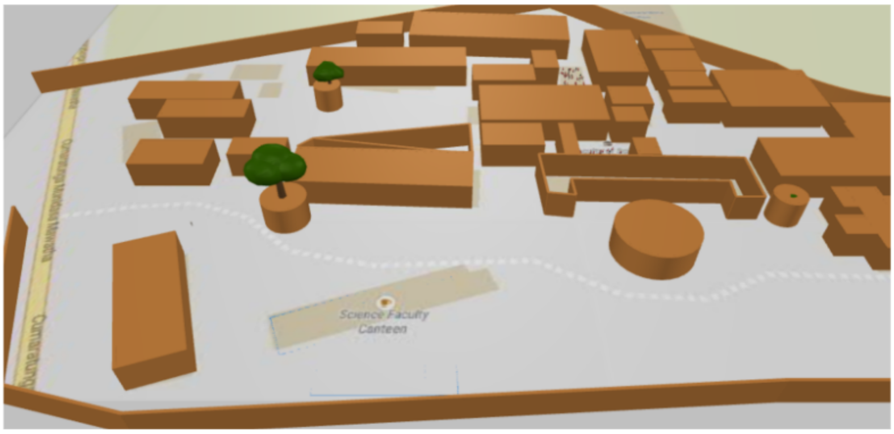
Time tables of academic levels are accesible to anyone via science faculty website. If possible with permission we might be able to get data about students featuring below factors that will define the behaviour of a student. Parameters defining each individual:

1. Year of study
2. Degree program
3. Friend association
4. Study motivation

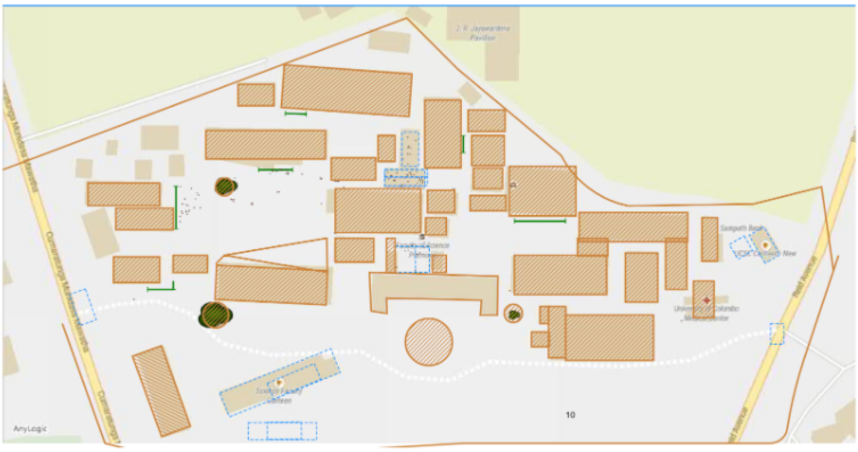
And finally the GIS (Geographic Information system) data available on Science faculty of University of Colombo, which can be directly accessed through Anylogic itself.

2.3.3 Sub-models

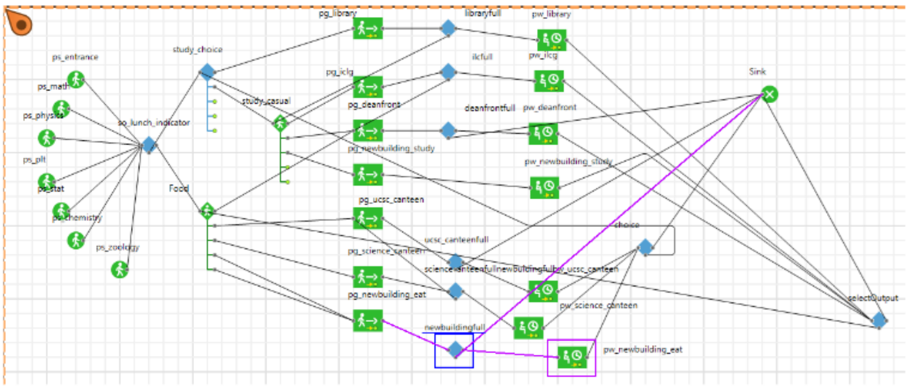
Using the above components we designed the below model to suit our purpose. The provided visualisation depicts an elaborate portrayal of the spatial layout within the Faculty of Science, with a particular emphasis on the distribution of lecture halls. This vital knowledge is painstakingly generated using Geographic Information System (GIS) technology, which allows us to thoroughly analyse and comprehend the changing patterns of space utilisation within the academic institution. We acquire a thorough grasp of which locations are likely to encounter the highest levels of activity during peak hours by studying the intricate spatial arrangement. This not only aids in efficient resource allocation but also improves overall facility management. The GIS library is an essential instrument in this endeavour, enabling for the smooth integration of spatial data to create a comprehensive and visually appealing representation of space utilisation patterns, facilitating informed decision-making processes.



3D visualization of the spatial resources inside UOC

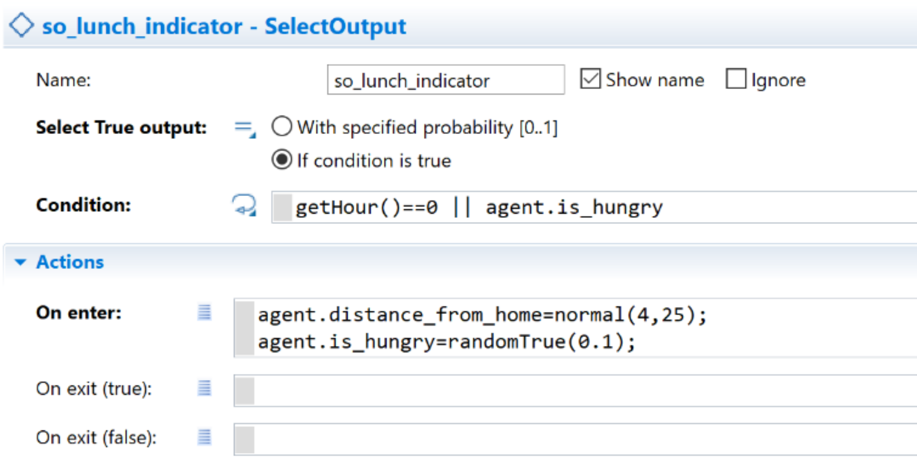


2D visualization of the spatial resources inside UOC



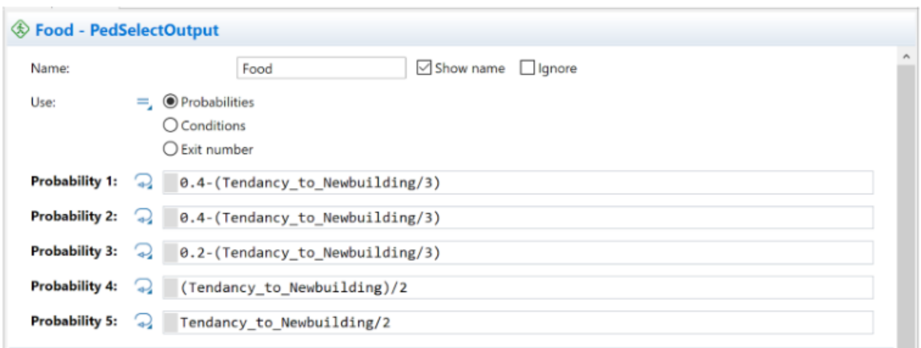
Logical model of the spatial resources inside UOC

Here all the pedsorces have the naming convention of starting with "ps", for readability. Same goes for pedgoto blocks and pedwait blocks with conventions "pg" and "pw" respectively. So if we pick a student arriving from any source, that student will go through the "lucnch_indicator" selectoutputblock. The purpose of this block is to route the student to studying purposes if the time is not between 12.00 pm and 1.00 pm or if they are not hungry.



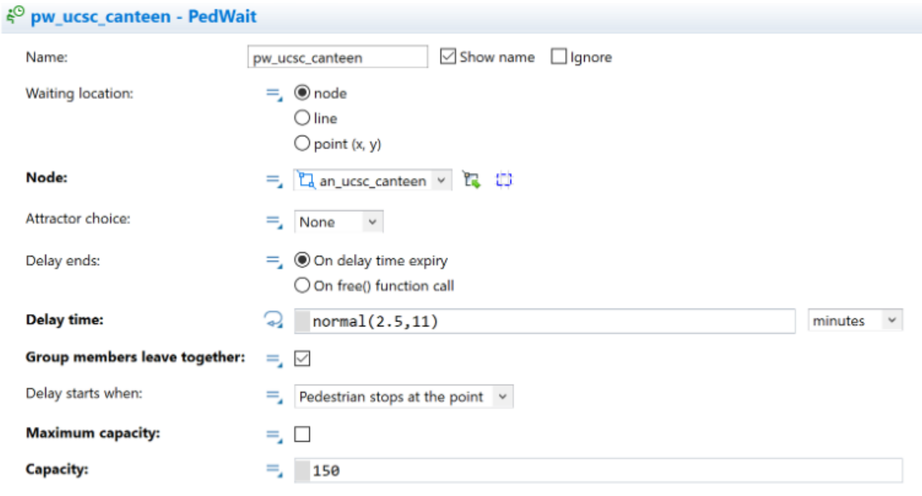
lunch indicator parameters and values

(also the parameter called distance from home is also assigned to the agent at the enter to this block) Otherwise the student will go to the “food“ , selectoutput5 option and make a probabilistic choice for a place to dine. These probabilities are parameterized with the parameter “Tendencay_To_new_building” for the ease of doing model experiments with and without newbulding.



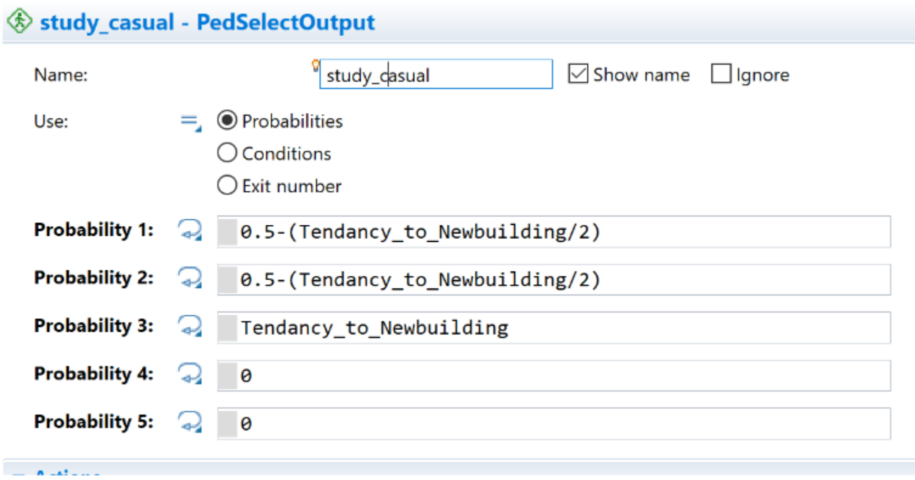
probability values of the 5 output

(Although there are 5 options please note the last two options both lead to “pg_newbuilding_eat”. This seemingly inappropriate choice of connection was made to avoid certain errors put forth by the Anylogic software.) Note that if “Tendencay_To_new_building” is 0 , student will choose UCSC and Deanfront spaces to dine at the probabilities 0.4 each. A small probability was assigned to Science canteen because it is at a corner of the university premises , away from the entrances and generally lack resources with respect to other two available choices. each of “pedgoto” blocks are connected to a selectoutput block to check weather is is fully occupied . If it is the case the student is sinked via “pedsink”. This is the implication that the student will have to turn towards facilities outside of science faculty. Otherwise the student will enter to the location (pedwait blocks) and spend time in minutes according to a normal distribution. The below figure shows the case for UCSC canteen.



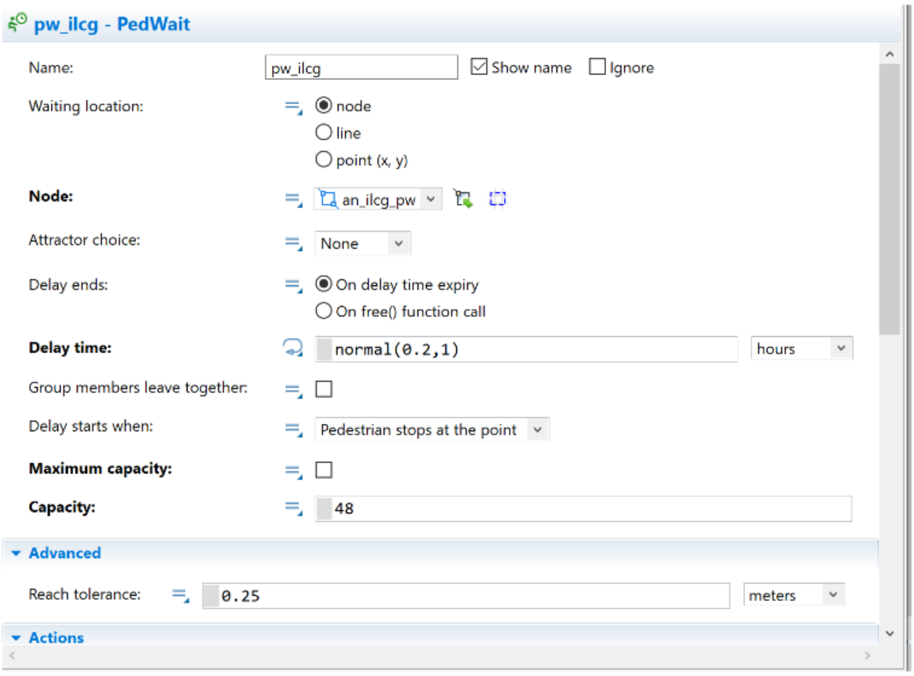
parameter value used for UCSC canteen

After dining the student will leave pedwait blocks to another selectoutputblock. From here student is allowed to ‘sink’ if a threeway AND condition is satisfied. first condition is checking if it’s after 12.00 pm. Second if the distance from home is greater than 25 km to the university. And the third is if the agent does not have a lecture to attend. Otherwise the student will be routed to “study_choice” selectoutput5 block to make a choice for a location to study, based on their preferences. Let’s shift our attention to the upper branch of the model now.Here student will choose the library if they doesn’t have a lecture to attend and if they prefer independent quiet study sessions. Otherwise they will go to the block “study_casual” to choose an open area to study. The probabilities of “study_casual” block is again parameterized with the parameter “Tendencay_To_new_building” for the reasons stated before.



probabilities of “study casual” block is again parameterize

Here we have taken that students are equally likely to pick deanfront and ILC bottom floor to spend their freetime. And you can see from the graph that the students who initially went to the library , will go through a selectoutputblock called “libraryfull”. Its task is to send the students who didn’t find a place within library to an open space. So they are rerouted to “study_casual” block to make a choice again. Each of the pedgoto blocks to the open areas are also connected to a such a select outputblock to see weather if they are full. If a student goes to ILC and it is full, they will turn back to dean office. If the both options failed they’ll leave the system via sink. If they aren’t full they will spend there a time distributed as a normal distribution. Here is an example from the block pedwait block to ILC (pw_ILCg)



pedwait block to ILC (pw ILCg)

Once they leave the pedwaits , they go through one last selectoutputblock to check weather the time is between 12.00 pm and 1 pm. If so they'll head to "Food " selectoutput5 block, whose functionality was explained earlier. That was brief overview of the logic behind our model and most of its anylogic components and submodels.

3 Results

During the runtime of the model following data were gathered in each 5 minute lapse. The number of students currently occupying,

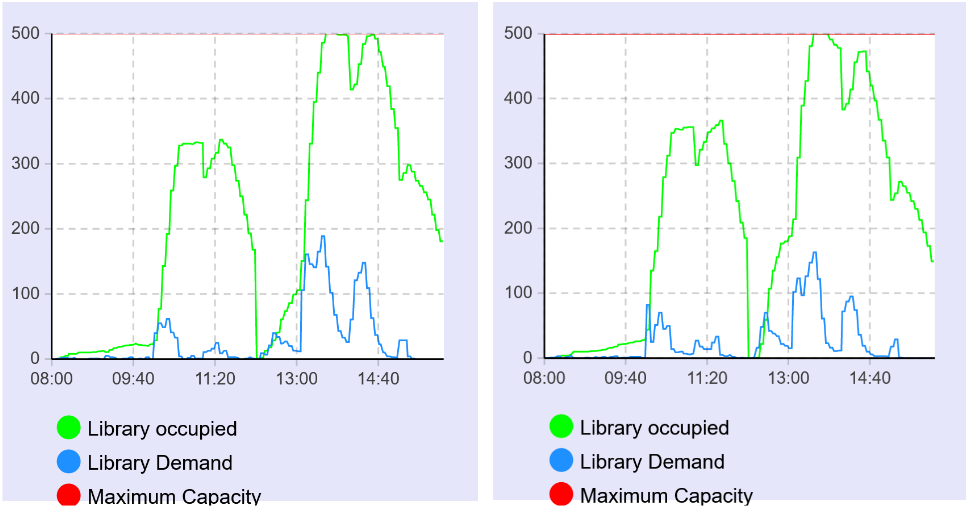
1. Library
2. Bottom floor of ILC
3. Deanfront
4. Science Cafeteria
5. UCSC Cafeteria
6. New building's Cafeteria
7. New building's Study Zone

And they were plotted against time during the model runtime to recognize the time slots where the capacity of each space reach its critical level. First we ran the model without the newbuilding's presence. This was achived by setting "Tendancy_to_new_building" parameter to 0. Used rate schedules for each source.

Next we ran the model with the presence of the new building. From our point of view (as current students) it is logical to assume that once it is opened in upcoming months , it will be the go-to place for the students for both casual studying and to take lunch. To model this high preference , we set the parameter of "Tendancy_to_new_building" to 0.7.

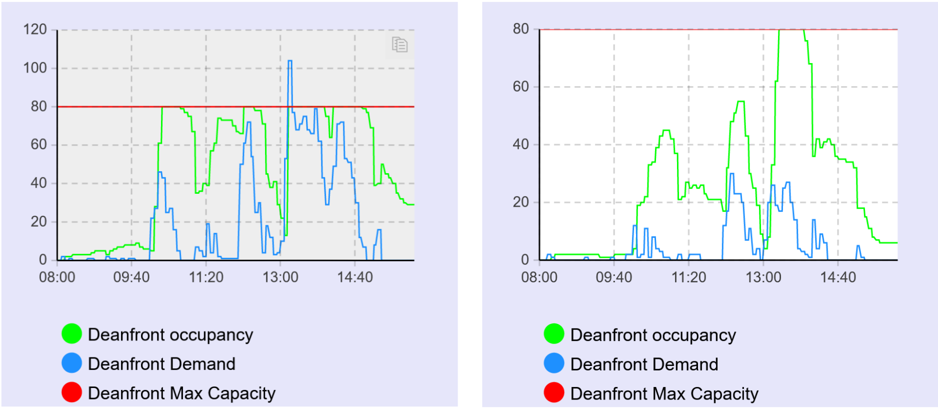
Here we present the output for only a one set of parameters Capacity of the New building's Cafeteria = 350 Tendancy to newbuilding = 0.7 Capacity of the New building's Study zone = 400 ratemodifer = 0.6 Chosen parameters , Capacity of both the study zone and the Cafeteria

3.1 Library



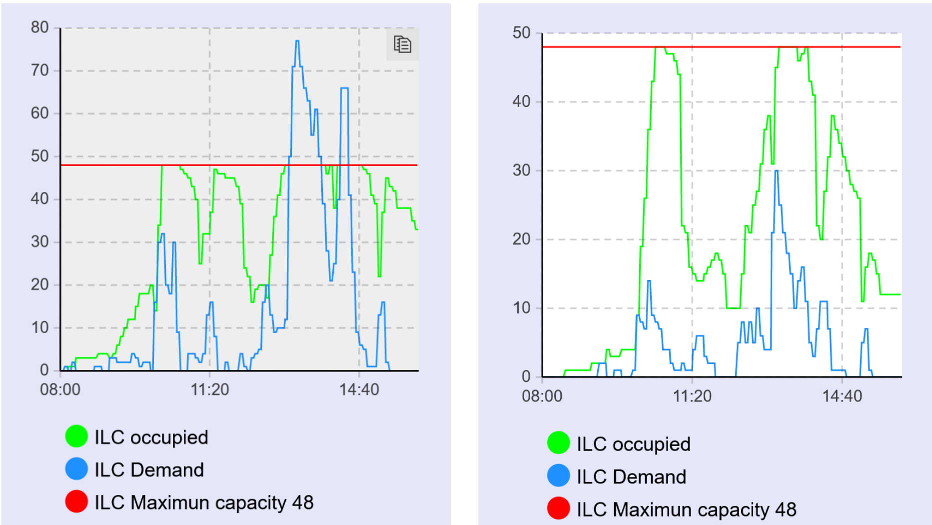
The state of the Library without (Left) and with (Right) the presence of New Building

3.2 In front of Dean's Office



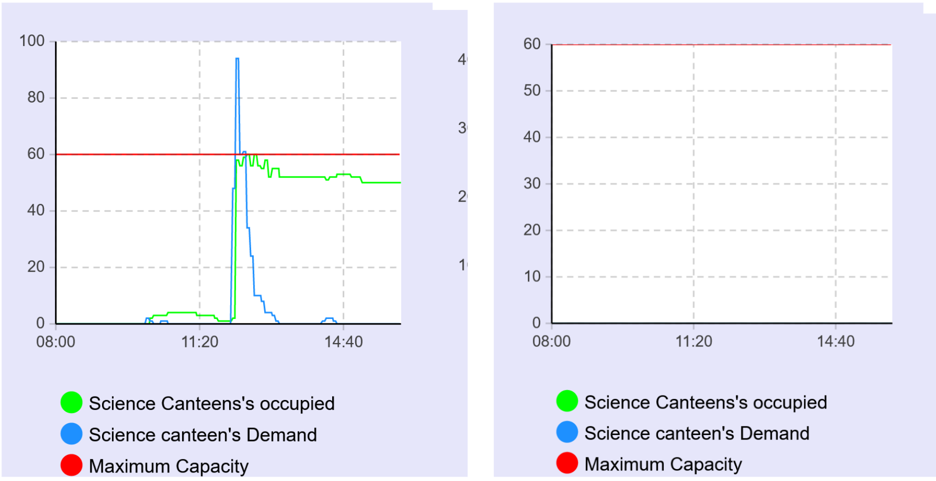
The state of the in front of the dean's office without (Left) and with (Right) the presence of New Building

3.3 Information and Learning Center



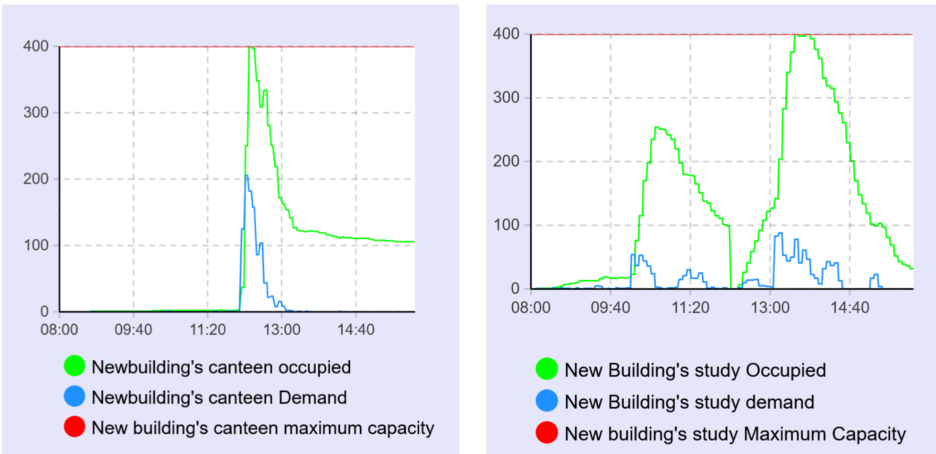
The state of the ILC without (Left) and with (Right) the presence of New Building

3.4 Science Canteen



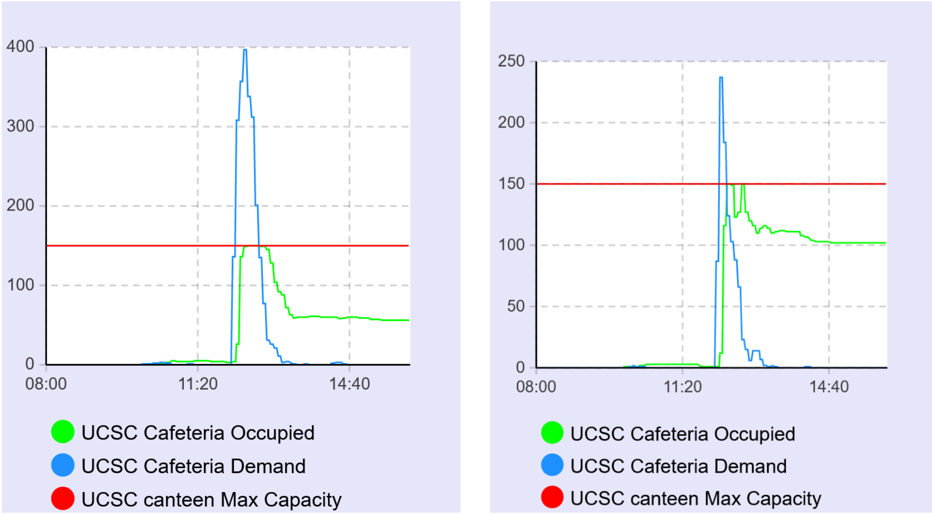
The state of the Science Canteen without (Left) and with (Right) the presence of New Building

3.5 New building



The state of student behaviour in the New building's Cafeteria (Left) and Study Zone (Right)

3.6 UCSC canteen



The state of the UCSC canteen withoyt (Left) and with (Right) the presence of New Building

4 Discussion

4.1 Without the presence of new building

After lunch, about 1.30 pm the library seemed to have a surge of incoming students. And it remains the same up until about 2.40 pm. The open space infront of dean's office seemed to be a busy place after 9.40 am. The reason for this could be because unlike other open spaces students are allowed to take lunch in this area. Apart from lunch hours (10 am- 12 pm 1 pm to 3 pm) , the crowd at the Information Learning Center also behaves in a very similar manner to the Deanfront. This is due to the fact that these two have become the primary candidates for students to spend their free time or to do casual study. When we consider Science Canteen and the UCSC canteen we could see that there is a high demand from students –nearly 400 students at times, for students, to find a space to take lunch there. Science canteen also shows the same behaviour but only for a shorter period of time (about half and hour between 12.00 pm and 12.30 pm).

4.2 With the presence of new building

The Library area seemed to be unaffected by the presence of the newbuilding. Which makes sense because we assumed the newbuilding does not offer the independent and quite study zones like the ones the Library has. Compared to the previous experiment , the open space infront of the deans office shows a large improvement in managing its crowd. But There is still a short period of time between 13.30 pm and 2.00 pm the space fails to accommodate the incoming flow of students. The observation for the ILC is the mostly same but there are two are two time slots where te crowd became too much to handle. The demand for the UCSC canteen has been reduced by the new building and somehow the high preference to the New building's cafeteria has made the science canteen go empty throughout the day. Also it should be noted that both these places the number of occupying students remains the same after the peak. We consider this to be an error occurred during our modelling phase, but there is enough evidence to conclude that this behaviour does not interfere with the model's behaviour during the rush hours. The newbuilding's study zone is handling the needs of majority of the student body as we observe from its two zones. Although the cafeteria reaches its maximum capacity at 12 pm and the study zone reaches its critical capacity at about 2.00 pm a both time slots are smaller time periods. (30 minutes at maximum).

5 Conclusions

From the parameters set we used, it appears to be the case that while the new building is highly effective on reducing the space crisis occurring in the busy hours on the day it does not solves the issue at hand completely. Also the Science canteen didn't have any students due to the student's high preference towards Newly constructed building.Hence it might be a financially good decision to close the cafeteria , and convert it to an open space as it's still a prevailing problem from what we observe from these outputs.

Simulating a university proved to be a challenging but interesting task for us. There was no previous attempt to do such a modelling as far as we know. That being said, it should be accepted that using parameters without sufficiently being backed by data is hardly a good practice. Even if that was done , it is difficult to capture human behaviour in modelling software with a number of limitations on a free user . So further and more complex and accurate development of this implementation is possible via a suitable software or adopting another technique to achieve what we achieved in anylogic. The code for this project and all other relevant information and supplementary documents can be found in the following public Github repository.

6 Code availability

The anylogic files and other relevant documents to this project can be found in the following public Github repository.

https://github.com/CandyPanda-LS/Agent-based-modeling-

7 Supplementary Documents

Student rate calculations based on time table data