Delft University of Technology

Using pedestrian modelling to inform virus transmission mitigation policies

A novel activity scheduling model to enable virus transmission risk assessment in a restaurant environment

Sparnaaij, Martijn; Yuan, Yufei; Daamen, Winnie; Duives, Dorine C.

DOI
10.1016/j.physa.2023.129395

Publication date
2023

Document Version
Final published version

Published in
Physica A: Statistical Mechanics and its Applications

Citation (APA)
Sparnaaij, M., Yuan, Y., Daamen, W., & Duives, D. C. (2023). Using pedestrian modelling to inform virus transmission mitigation policies: A novel activity scheduling model to enable virus transmission risk assessment in a restaurant environment. *Physica A: Statistical Mechanics and its Applications*, *633*, Article 129395. https://doi.org/10.1016/j.physa.2023.129395

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Physica A 633 (2024) 129395

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Using pedestrian modelling to inform virus transmission mitigation policies: A novel activity scheduling model to enable virus transmission risk assessment in a restaurant environment

Martijn Sparnaaij *, Yufei Yuan, Winnie Daamen, Dorine C. Duives

Department of Transport & Planning, Delft University of Technology, Stevinweg 1, Delft, 2628 CN, The Netherlands

ARTICLE INFO

Dataset link: Restaurant activity scheduling data (Original data)

Keywords:
Pedestrians
Microscopic pedestrian model
Virus transmission risks
COVID-19
Restaurant operations
Activity scheduling

ABSTRACT

The Covid-19 pandemic has had a large impact on the world. The virus spreads especially easily among people in indoor spaces such as restaurants. Hence, tools that can assess how different restaurant settings can impact the potential spread of an airborne virus and that can assess the effectiveness of mitigation policies are of high value. Microscopic pedestrian models provide the tools necessary to assess the detailed movements of people in a restaurant and with that the risk of virus transmission. This paper presents the application of a microscopic pedestrian model, including a novel activity choice and scheduling model, to assess virus transmission risks in restaurants. Simulation experiments identify that different factors impact virus transmission risks in a restaurant. Contacts between restaurant staff and customers are the driving factor for virus transmission in a restaurant whereby especially staff presents a big risk. Hence, mitigation policies focussing on these interactions and on preventing staff from transmitting the virus can be highly effective. The results also show that different restaurant layouts and setups lead to distinctly different transmission risks. Therefore, insights obtained from simulating one restaurant cannot be just transferred to any other restaurant. Together, these results show the added value of including pedestrian models in disease transmission risk modelling exercises to mitigate the impact of a pandemic caused by an airborne virus. However, the research also shows that, to better utilize the potential of pedestrian models for disease transmission risk modelling, future research of pedestrian activity scheduling behaviour in indoor spaces is necessary.

1. Introduction

Since its introduction at the end of 2019, the SARS-CoV-2 virus has had a major impact on the world. Next to the direct impact of, among others, a worldwide death toll in the millions and a many-fold of hospitalizations [1], measures to curb the spread of the virus such as lock downs heavily affected many societies [2]. Next to vaccination, a group of measures, collectively known as non-pharmaceutical interventions (NPI's), are the primary tools of countries worldwide to reduce virus spread. Well-known examples of these NPI's are physical distancing regulations and face masks. A lot of these measures focus on limiting the virus transmission in indoor environment, such as shops, restaurants, public transport and schools, as this is the primary location where virus transmission takes place [3]. However, the question is, how effective are these different measures in preventing or reducing the transmission of the virus in these indoor environments?

* Corresponding author.
E-mail address: m.sparnaaij@tudelft.nl (M. Sparnaaij).

https://doi.org/10.1016/j.physa.2023.129395
Received 28 July 2022; Received in revised form 18 October 2023
Available online 22 November 2023
0378-4371/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Modelling of virus transmission is a common method within the field of epidemiology to obtain insights into the spread of a virus and how different measures, both pharmaceutical and non-pharmaceutical, might affect this. In the case of the SARS-CoV-2 virus, the value of mathematical modelling is also widely recognized and models are widely used especially to inform policy makers [4–6]. These models come in various forms, such as compartmental, network or agent-based models [7], all of which have their strengths and limitations. What these risk assessment models have in common is that they lack the ability to provide detailed insight into how the activity and movement dynamics of people in indoor spaces affect the virus transmission, and therefore also the potential impact different measures can have on the virus spread. Yet, more specific insights on the impact of movement dynamics can be of added value to policy makers by providing more detailed information about how the specific function, layout and setup of an indoor space relates to the probability of virus transmission and the effectiveness of different mitigation strategies. Furthermore, these insights could assist owner/operators of indoor spaces, such as restaurants, schools and museums in setting up their indoor space as safely as possible.

Models from the field of transportation, and in specific the field of pedestrian dynamics, can reproduce these detailed behaviours of people in indoor spaces. For different types of indoor environments these attempts have been made. Xu and Chraibi [8], Harweg et al. [9] and Tsukanov et al. [10] all use pedestrian models to estimate virus transmission risks in a supermarket whilst Alvarez Castro and Ford [11] and Li and Yin [12] focus on university campuses and Xiao et al. [13] and Sajjadi et al. [14] model generic public spaces. These studies use different ways to estimate the transmission risk including the SEIR model [11,14], exposure time [12,13] or distance related metrics [8,9].

Restaurants form another category of indoor spaces that can be a major source of COVID-19 transmission [15] and that have been hit hard by the NPI measures to dampen the pandemic [16]. Therefore, detailed insight into the risk of the virus spreading in a restaurant environment can be of value when determining when it is safe to open restaurants and how mitigation policies can assist in achieving this. Restaurants also form a very different type of indoor environment compared to the indoor environments that have been studied (supermarkets, campuses and generic indoor spaces) regarding the movement dynamics. That is, most people are seated at a specific location for most of the time (i.e. the customers), whilst a minority is moving around a lot (i.e. the waiting staff). Therefore, the aforementioned models are not suitable for modelling restaurants, and to the authors' knowledge, no models have been presented as of yet that can model this specific type of indoor space. Our objective is threefold. Firstly, to present a pedestrian model for restaurant environments that can provide detailed insight into the risk of the virus spreading in a restaurant. Secondly, to show how such a model can be used by policymakers and restaurant owners to inform policies to mitigate virus transmission. And thirdly, to investigate what information on virus transmission risks a pedestrian model in and of itself can and cannot provide understand what the added value of a virus transmission model would be.

We solely focus on the pedestrian model in this paper although the modelling effort is part of a bigger effort to fully integrate pedestrian and virus transmission modelling [17]. We ignore the virus transmission model for two reasons. Firstly, it reduces the complexity of the overall model enabling a better understanding of the behaviour of the new activity model and a better understanding of how different activity choice and scheduling patterns can impact the contact patterns between people in a restaurant setting and with that the virus transmission risk. Secondly, it provides insight into what a pedestrian model in and of itself can already provide as insights for policy makers. This also provides a baseline for future research into what the added value is of adding more complexity in the form of a virus transmission model. Not only with respect to the effort in [17] but also in general.

To model the movements of people in a restaurant we need to model different behaviours. In pedestrian modelling these are typically classified into: (1) The *strategic* behaviour, which involves activity choice. (2) The *tactical* behaviour, which involves activity scheduling and route choice. And, (3) the *operational* behaviour, which involves the step-by-step walking behaviour [18]. A multitude of pedestrian model exist that model the route choice and operational behaviour of pedestrians [19,20]. However, in general, pedestrian models do not contain integrated activity choice and scheduling models but consider this an input. As the next section shows, to the authors' knowledge, no activity choice and scheduling model has yet been presented for a restaurant environment. This presents a challenge because creating these activity schedules for all customers and staff requires time and modelling expertise. Certainly, because for the model to make an impact, ideally it should be used by policy makers and restaurant owners neither of which necessarily have this modelling expertise and/or time. To solve this challenge, we present a novel activity choice and scheduling model for restaurants in this paper before investigating the use of pedestrian models to investigate the virus transmission risk in restaurants.

The remainder of the paper is setup as follows. Section 2 presents the state-of-the-art in pedestrian activity choice and scheduling models. Section 3 then introduces the novel activity choice and scheduling model. Following this, Section 4 presents the data that is used for estimating the parameter values of the activity choice and scheduling model and the procedure used to collect it. Then, Section 5 presents the setup of the simulation experiments whose results are discussed in Section 6. Lastly, in Section 7, the main conclusions are presented, the strengths and limitations of the model are discussed and its implications for practice are presented.

2. State-of-the-art in pedestrian activity choice and scheduling models

Activity choice and scheduling are essential parts of a pedestrian simulation model. However, thus far it has not been studied comprehensively [18] and is mainly seen as exogenous to the model [Chap. 3 21]. That is, it is usually specified as an (essential) input to the model. In the literature, few examples can be found of activity models that focus on pedestrians and on the microscopic scale (i.e. a single pedestrian facility, such as, a restaurant or a small section of a city like a shopping street).

Both Timmermans et al. [22] and Zhu and Timmermans [23] use a discrete choice model approach to model the activity schedules of pedestrians in a shopping street. The schedules are created dynamically whereby after every activity a new activity is chosen based on its utility. The aforementioned papers on modelling pedestrian movements in a supermarket in the COVID-19 context [8–10] all use a similar approach though without the utility concept of the discrete choice model. These models result in activity schedules for each pedestrian that result in (semi-) random walks through the environment and are independent of the schedules of other pedestrians. Though a valid simplification for a shopping street or supermarket environment, we expect that the activity schedules of customers and staff in a restaurant are not well represented by this kind of simplification. For example, we expect that customers going to the toilet return to the same seat, we expect customers belonging to the same group to arrive and leave at approximately the same time and staff to serve occupied tables instead of empty ones.

Airports are another environment that have received some attention from the modelling community. Liu et al. [24] use a nested logit model to choose the activities a departing passenger performs between entering the airport and boarding the plane. The model is limited as it does not actually schedule the activities but only chooses which activities to perform in one of the three predefined time periods (before check-in, before security and before boarding). To our knowledge, only the work by Usher et al. [25] presents a model capable of creating a complete activity schedule for departing passengers. They use a rule-based approach centred around a set of activities which are all classified as being mandatory, auxiliary or discretionary. Based on the remaining available time till boarding passengers update their activity schedule dynamically whereby they can skip non-mandatory activities if time is pressing.

Shelat et al. [26] present a model in the context of an office environment. To create the activity schedule of office workers they rely on a combination of an event scheduler which schedules all time-dependent activities and a Markov-chain scheduler to schedule all other activities between the time-dependent activities. This model still requires extensive inputs in the form of the time-dependent activities such as all scheduled meetings.

Lastly, two papers [27,28] from the field of operations research present activity models focussed on restaurants. However, these papers (1) focus on optimization of customer satisfaction or limiting waiting times for a table instead of the number and duration of contact between people in a restaurant environment and (2) the model description and evaluation are very limited. Hence, it is not possible to reproduce these models nor judge their validity.

Overall, a review of recent literature shows that activity choice and scheduling models for pedestrians have received little attention. However, the limited research that is available shows multiple methods to tackle the problem. Yet, no activity choice and scheduling model was found that allows us to model the activity choice and scheduling of customers and waiting staff in restaurants, in such a way that realistic estimates of contacts between different people can be obtained.

3. An activity choice and scheduling model for a restaurant

This section describes the new restaurant activity choice and scheduling model. First, the intended goal and use of the model will be explained, directly followed by the set of practical limitations faced by the researchers (Section 3.1). Based on the goals and requirements and the insights from the previously discussed literature, the basic design of the model is presented. Afterwards, a detailed explanation of the model is provided. Here, we distinguish a customer scheduling model (Section 3.2) and a staff scheduling model (Section 3.3).

3.1. Requirements of an activity choice and scheduling model

3.1.1. The goal of the activity choice and scheduling model

The intended goal of the model is to provide activity schedules for all customers and waiting staff in a restaurant. Together with an existing route choice and movement dynamics model, these schedules can provide a realistic estimate of the contacts between different people within a typical dine-in restaurant setting. The focus of this modelling endeavour is on the customer–customer contacts and the staff–customer contacts, rather than the overall capacity or operation of the spaces. Please note, our focus is not on quantifying staff–staff contacts in non-public areas such as the kitchen as this is a completely different type of environment compared to the public side of a restaurant.

3.1.2. The application context

The context within which this model is intended to be applied is dine-in restaurants. The main features that distinguish typical dine-in restaurants from other types of restaurants, such as take-out, fast-food or buffet-styled restaurants, is that (1) people eat the ordered food in the restaurant at one predefined location (table & seat) and (2) waiting staff serves the food. Furthermore, the intended users of the risk assessment (and as such the scheduler) model are restaurant owners, who neither have much experience or expertise in pedestrian and virus transmission modelling nor much time to spend on setting up these simulations. This mainly implies that the inputs should be simple and few. That is, they should be easily understood by policy makers and restaurant owners, take little time to provide and require only those details that a restaurant owner could reasonably know based on their experience managing their restaurant.

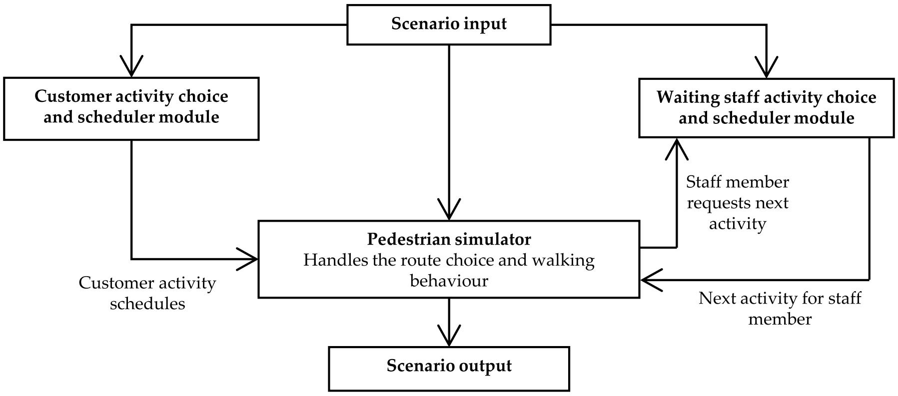

Fig. 1. Overview of the integration of the customer and waiting staff activity choice and scheduling model into a typical pedestrian model.

3.1.3. *The structure of the activity choice and scheduling model*

Fig. 1 presents an overview of how the newly designed model integrates with a typical pedestrian model which simulates the route choice and walking behaviour of all customers and staff in the restaurant. The model features two sub-models, one for the customers and one for the waiting staff. Each sub-model has a different approach and interacts differently with the pedestrian simulator. The core difference between both sub-models is that the customer sub-model provides static activity schedules for the customers to the pedestrian model at the start of the simulation. The staff sub-model, on the other hand, provides on-the-fly assignments to staff members assigning their next activity dynamically during the simulation. The next two subsections provide more detail about the design and design choices of both these sub-models.

3.2. *Customer activity choice and scheduling model*

The design and implementation of the customer sub-model will be explained in three steps. First, the input is discussed whereby the focus is upon the information that restaurant owners can realistically and easily provide regarding the customer and staff behaviour in their particular restaurant. Second, the scheduling of the visit of every group of customers is explained (i.e. when do they start their visit, at what table and when do they leave). And third, we show how the activity pattern of each individual member of a group is determined based on the input and the schedule of each group.

3.2.1. *Inputs*

Restaurant owners are the intended group of users of the model, therefore the inputs to the model should align with the inputs they can easily provide. Based on discussions with people in the restaurant industry and our own insights into what inputs are necessary, we selected the following inputs:

1. The restaurant layout. Specifically the number of tables and the number of chairs per table.
2. The time period that should be simulated.
3. The demand pattern: This input divides the overall time period into smaller time slots and for each of those defines how many groups of customers will visit the restaurant during that time.
4. The expected duration of the customer visits $t_{visit;expected}$.

3.2.2. *Group scheduling*

Based on the input described above, the groups are scheduled. Yet, before the implementation is presented the expected behaviour is discussed. The scheduler is expected to schedule the groups such that:

- The groups start and end their visit within the time slot defined by the demand pattern.
- Time slots can overlap. That is, the next time slot can start before the end time of the previous time slot.
- Time slots cannot have more groups than the number of tables in the restaurant.
- If, during the time slot, no unoccupied table is available for at least the duration defined by the expected visit duration, the group is not scheduled to visit the restaurant.
- A group should be able to occupy their assigned table for at least a duration equal to the expected visit duration although they can stay shorter or longer if they choose to and the final schedule allows for this.
- Each table is only occupied by one group and a group only occupies one table.

To achieve the above-mentioned scheduling behaviour, the scheduling process is divided into two steps: The first step creates a provisional schedule, which details which groups sit at which table in which order. The second step further details the provisional schedule and determines the actual start and end time of each group's visit.

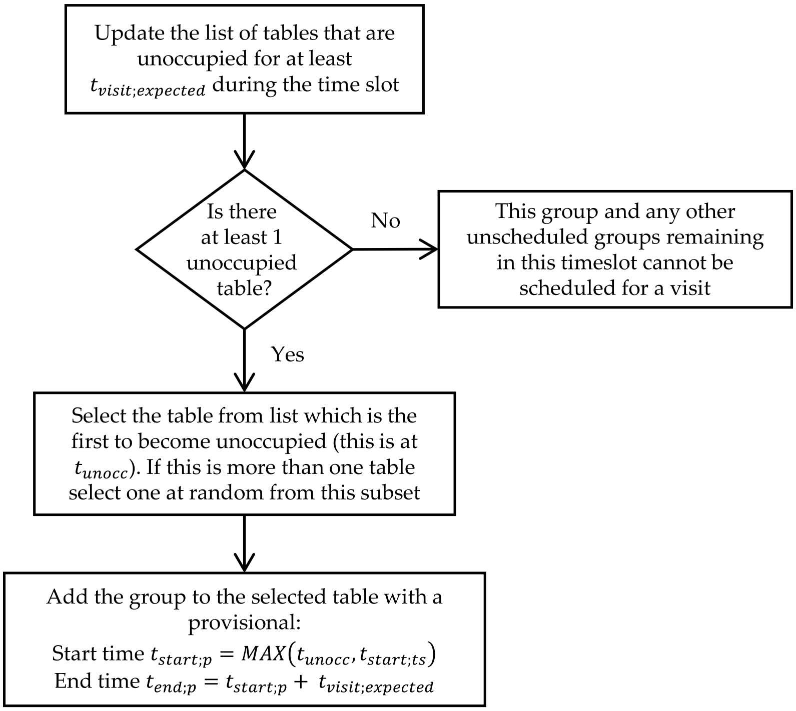

Fig. 2. The algorithm that assigns a group from a given time slot ($t_{start;ts} - t_{end;ts}$) a table and a provisional start and end time for the visit.

To create the provisional schedule the algorithm loops over the demand pattern time slots in order of increasing start times. The start time of a time slot is given by $t_{start;ts}$. For every time slot, the algorithm loops over the number of groups expected to visit the restaurant within this time slot and assigns them to a table according to the logic displayed in Fig. 2.

The second step sets the actual start and end time of each group's visit by adding some variation in the duration of the visit and the precise start time of the visit. This ensures that the groups within a time slot have slightly different arrival times and visit durations. To achieve this, the algorithm loops over all groups per table in reverse order of arrival. Eqs. (1)–(4) are used to calculate the start and end time ($t_{start;i}$ and $t_{end;i}$) for each group.

$$d_{visit;max;i} = \mathrm{MIN}(t_{end;ts;i}, t_{start;i+1}) - t_{start;p;i} \tag{1}$$

$$d_{visit;i} = \mathrm{MIN}(d_{visit;max;i}, d_{visit}) \tag{2}$$

$$t_{start;i} = t_{start;p;i} + \mathcal{U}(0, d_{visit;max;i} - d_{visit;i}) \tag{3}$$

$$t_{end;i} = t_{start;i} + d_{visit;i} \tag{4}$$

Eq. (1) computes the maximum possible visit time for the group ($d_{visit;max;i}$) based on the preliminary start time ($t_{start;p;i}$) and the end of the time slot ($t_{end;ts;i}$) or the start time of the next group that sits at the same table ($t_{start;i+1}$), whichever comes first. In case the group is the last to sit at the table the value $t_{start;i+1}$ is infinite. Eq. (2) then computes the actual visit duration ($d_{visit;i}$) by drawing a visit duration value (d_{visit}) from a predefined distribution and ensuring the duration does not exceed the maximum possible visit time. The start time is then computed by Eq. (3) whereby a uniform distribution ($\mathcal{U}$) is used to create some variation in the starting times of the different groups. Finally, the end time is computed using Eq. (4). The result of this first step is a group assignment (group id, time start, time end, table no.).

3.2.3. Individual activity schedule creation

After a group has been assigned a table and the start and end times of the visit have been computed, individual customers can be assigned to the group and their individual activity schedules can be created. The number of individuals that is assigned to a group is equal to the number of seats at the table the group was assigned to. That is, per table we expect 100% capacity. To create activity schedules for all individuals, we identified the activities that people perform at a restaurant. The following list of activities are/can be part of the schedule of an individual visiting a restaurant:

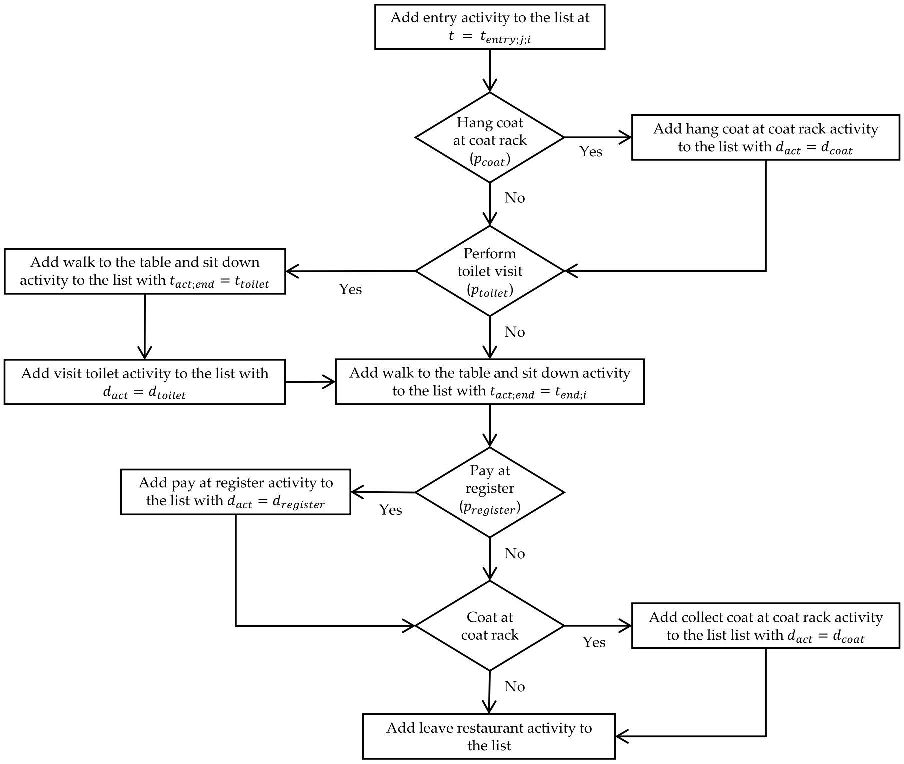

Fig. 3. The algorithm that creates the activity schedule for each individual j of group i.

- *Enter the restaurant:* This is always the first activity and a mandatory activity.
- *Hang coat at the coat rack:* This is an optional activity with a certain duration d_{coat} performed after entering the restaurant. Based on a certain probability p_{coat} the customer chooses to use it or not. In case no coat rack is available $p_{coat} = 0$.
- *Sit at the table:* A mandatory activity performed after entering the restaurant or using the coat rack.
- *Go to the toilet:* An optional activity with a certain duration d_{toilet} and which a customer chooses to perform given a certain probability p_{toilet}.
- *Pay at the register:* A conditional activity with a certain duration $d_{register}$ assigned to only one member of a group provided the payment is not performed at the table. Whether or not a payment is performed at the register is determined by a certain probability $p_{register}$.
- *Pick up coat from the coat rack:* A conditional activity provided the customer chose to hang their coat at the coat rack when entering the restaurant.
- *Leave the restaurant:* The last activity and a mandatory one.

The list only includes activities related to the customers' movements. Activities at the table, such as ordering food, have no influence on the movement pattern of a customer and are hence not included. They do impact the movements of staff which is discussed in Section 3.3.

Fig. 3 provides an overview of the algorithm that creates the activity schedule for each individual j of group i. Each activity starts after the previous activity has finished. This can be either at a set end time $t_{act;end}$ or after a certain duration d_{act}. The exception is the first activity which has a set starting time. By default this starting time is the group starting time $t_{start;i}$ but it can be individualized to simulate customers of the same group arriving at different times. If a customer decides to visit the toilet a time at which this toilet visit should take place t_{toilet} is scheduled. The toilet visits of all individuals are scheduled such that there are no more toilet visits scheduled at the same time than there are toilets. The underlying assumption is that customers know how busy the toilets are (global knowledge) and plan accordingly. Short queues and waiting times can still occur as the dynamics in the simulation (e.g. walking time) are not taken into account.

3.3. Waiting staff activity choice and scheduling model

The model that schedules the activities for the customers is static. That is, all activities are computed and assigned at the start of the simulation. In contrast, the model that chooses and schedules the activities of the waiting staff in the restaurant is an event-based model that dynamically adds activities to be performed and also dynamically distributes the activities over the staff members.

3.3.1. Inputs

The inputs are again chosen such that restaurant owners are only required to provide few and simple inputs. In this case those inputs are:

- The number of waiting staff
- Serving neighbourhoods: A serving neighbourhood is a set of tables that is primarily served by one or more appointed members of staff.
- The average number of times the staff performs an activity at the table per group that sits at the table. These activities feature, among other things, serving food or drinks, taking orders, asking for feedback or taking empty plates away or any logical combination.
- The staff base area: This is a location in the restaurant where the staff waits if they have no activity to perform and also the location where they collect the food and drinks that are to be served and/or drop empty plates.

3.3.2. Choice and scheduling model

The staff sub-model makes use of activity stacks which contain the information about *which* activities still need to be performed, *when* these activities should be performed and in *which order*. Every neighbourhood has its own stack. When the entire restaurant features only one neighbourhood, a single stack is used. There are four events that either add or remove activities from the stack:

1. *Group entry:* A group sits down at their table for the first time. This is the only event that adds activities to the stack.
2. *New activity:* A member of staff has finished its current activity and request the next activity from the stack.
3. *New activity after waiting:* A staff member is waiting for their next activity and the start time of one or more activities on the stack is equal to the current time. The waiting staff member get assigned the top-most activity on the stack that has a start time equal to the current time.
4. *Aiding colleagues:* A staff member is waiting for their next activity and another neighbourhood gets overloaded with activities that should already have been performed but were not. The waiting staff member get assigned the top-most activity on the stack of the other neighbourhood.

Before going into detail how the activities are added and removed from the stack, we discuss the activities themselves first.

Activities One can imagine a wide array of activities that staff members typically perform in a dine-in restaurant in relation to the visit of a group of customers. Examples are, serving the food, taking orders, cleaning the tables and from time-to-time checking-in with the customers to see if everything is in order and to their liking. However, difference between restaurants or even countries can exist regarding what this exact set of typical activities is. To deal with these differences, we have simplified the activities that staff members perform into two abstract activities types, namely:

- Walk to a specific location (i.e. a table or the base area) and wait there for a given amount of time to perform the activity at hand (e.g. serve the food, collect empty plates, checking-in on the customers).
- Walk to a specific location (i.e. the base area) and wait there for the next activity to be assigned.

The first type of abstract activity is the type of activity that goes onto the activity stack whilst the second type is assigned to a staff member when none of the activities on the stack can be performed at the moment. Both activity types have a location property which defines where the activity should be performed. The first activity type also has an activity duration property. Furthermore, to determine when these activities should be performed and in what order they should be placed on the stack every activity has a start time. Lastly, an activity can also be a set of chained activities. Within the model there are three chaining options, next to just performing a single activity. First a chain of activities starting with going to the base area, then to a table and then returning to the base area. This could entail collecting an order to serve out, serving it at the table and then returning some empty plates back to the base area. The other two options are a subset of this, namely: (1) First going to the base area and then to the table and after this continuing with the next activity (e.g. serving out food or drinks). And, (2) first going to the table and then going to the base area after which the next activity can be started (e.g. picking up empty plates and returning them to the kitchen). Only the first activity of the chain is pushed onto the stack whilst the other one or two activities are always performed directly after the previous activity in the chain has finished. After finishing the last activity in the chain, a staff member can request a new activity from the stack.

Adding activities to the stack When a group sits down at their table for the first time, the algorithm determines how many activities should be performed by the waiting staff in relation to the visit of this group, how long each activity is, when it should be performed and if it is chained with collecting things at or returning things to the base area. This ensure that staff members only start visiting a table when guests are present at the table. The number of activities that are performed per group is defined by the input and most but not all of these activities are performed while the customers are at the table. A small number of activities, such as cleaning, are performed after the customers have left the table. The exact number of activities that is performed after customers have left the table is defined by the parameter N_{after}. For each of the activities a duration value is drawn from a distribution of activity duration. This distribution is a parameter. The start time of each activity is determined as follows:

$$t_{start;i;k} = \begin{cases} t + \Delta t & \text{if the activity is the first activity} \\ t_{leave;i} & \text{if the activity is performed after customers have left the table} \\ t + \Delta t + \sum_{m=2}^{M} g_m & \text{otherwise} \end{cases} \tag{5}$$

where t is the current time, Δt the time step, g_m the gap between activity m and $m-1$ and M the number of activities that are performed while the customers are at the table. The vector of the gaps between the start times of the activities are computed using:

$$\mathbf{g} = T_{\text{gap}} * \mathbf{f}_{\text{gaps}} \tag{6}$$

$$T_{\text{gap}} = t_{\text{end};i} - d_{\text{act};M} - b_{\text{last:act}} - t_{\text{start};i;1} \tag{7}$$

where T_{gap} the total gap time, $\mathbf{f}_{\text{gaps}}$ is a set of $M-1$ fractions that determine how long each of the $M-1$ gaps between the M activities is, $d_{\text{act};M}$ is the duration of the last activity that is performed while the customers are at the table, $b_{\text{last:act}}$ is a buffer time parameter ensuring that the last activity is actually performed while the customers are still at the table and $t_{\text{start};i;1}$ is the start time of the first activity To ensure that all M activities are performed whilst the customers are still at the table it follows from Eq. (7) that $\sum_{m=1}^{M-1} g(m) \leq T_{\text{gap}}$ and thus $\sum_{m=1}^{M-1} f_{\text{gaps}}(m) \leq 1$. This problem can be defined as a stick-breaking process [29] which redefines the problem as $\sum_{m=1}^{M-1} f_{\text{gaps}}(m) = 1$. That is, it assumes that the last activity will always be started at $t_{\text{end};i} - d_{\text{act};M} - b_{\text{last:act}}$. By making $b_{\text{last:act}}$ a random variable instead of a constant value we regain the original problem. Defining the problem as a stick-breaking problem has a big advantage as the stick-breaking problem has a well-defined solution. Namely, using N value drawn from a Dirichlet distribution to divide the stick in $N+1$ parts. The Dirichlet distribution is a beta distribution that is generalized to multiple variables which has the useful property that a set of N fractions drawn from it sums to 1. Therefore, we use this Dirichlet distribution (see Eq. (9)) to draw the $M-1$ fractions which determine the duration of each gap whereby each gap can have its own average duration.

$$f_X(x) = \frac{1}{B(\boldsymbol{\theta})} \prod_{i=1}^{N} x_i^{\theta_i - 1} \tag{8}$$

$$B(\boldsymbol{\theta}) = \frac{\prod_{i=1}^{N} \Gamma(\theta_i)}{\Gamma\left(\sum_{i=1}^{N} \theta_i\right)} \tag{9}$$

where $B(\boldsymbol{\theta})$ is a multivariate beta function expressed in the terms of a gamma function.

For each activity, the main location is the table. In some cases the activity will be chained with the activity of going to the base area to collect and/or return things. For each activity the probabilities p_{before} and p_{after} determine if an activity is chained with both first going to the base area and afterwards returning to the base area, with either of the two or with neither of the two. The activity duration at the base area is provided by the parameter $d_{act:base}$.

After the set of activities is created it is pushed onto the stack and the stack is sorted in ascending order of the activity start times.

Requesting an activity from the stack There are three manners in which an activity on an activity stack can be assigned to a staff member. The first is after a staff member has finished its current activity it requests an activity from the activity stack of its own neighbourhood (or the global activity stack if there are no neighbourhoods). If the activity on the top of the stack has a start time smaller or equal to the current time, this activity is assigned to the staff member. If there are no activities that need to be performed at this moment the staff member walks to the base area to wait for its next activity.

The second way in which a staff member can receive an activity is when the staff member is waiting at the base area for its next activity (or on its way to there to wait) and the current time is equal to the start time of the activity on top of the stack. The top activity is then assigned to this staff member. Lastly, a waiting staff member can also receive an activity from another neighbourhood. This only happens when one of two conditions is met for the other neighbourhood:

1. The number of activities that should already have been started (i.e. the start time is smaller than the current time) exceeds the threshold value $N_{late:activites}$.
2. The difference between the start time of the activity on top of the stack and the current time is greater than the threshold value $\Delta T_{top:activity}$.

4. Empirical data collection and parameter estimation

The previous section shows that the novel activity choice and scheduling model contains a number of parameters which influence the activity schedules of both customers and staff. To get realistic values for the different parameters of the activity choice and scheduling model, data was collected. This section presents firstly the details of the data collection followed by how parameter values are derived from the collected data.

4.1. Data collection

The data collection took place in the summer of 2021 in a dine-in restaurant in the city of Delft, The Netherlands. The focus of the data collection was solely on the indoor area of the restaurant. Fig. 4 presents a sketch of the indoor area and Fig. 5 a snapshot. The restaurant has 20 tables and additionally three locations at the bar where groups can sit and dine. The figure also marks the location between the bar and kitchen where staff picks up the food and drinks to serve out and where they return the empty plates and glasses. The restaurant did not have a coat rack at the time of observation and customers paid at the table.

We collected data during a Saturday evening between 17:00 and 22:00. The restaurant was fully booked for the evening and was visited by groups of various sizes which mainly consisted of adults. Three observers recorded the different visit and activity patterns

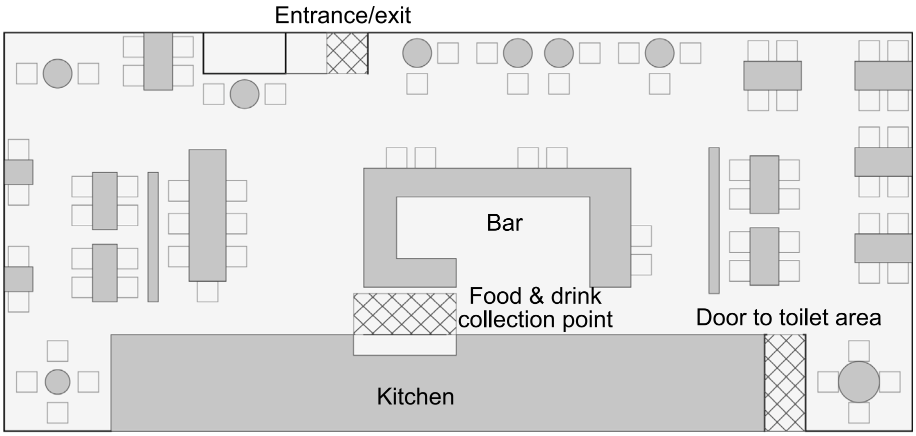

Fig. 4. An sketch of the restaurant layout. The drawing is not on scale but in correspondence with the then applicable COVID-19 regulations, all tables are spaced 1.5 m apart.

Fig. 5. A snapshot of the restaurant.

by hand whereby no personal information was recorded to ensure privacy of both customers and staff. The following elements were recorded:

- The arrival and departure times of all customers per table, whereby the start time is the first time a customer sits down at the table and the departure time is the time the customer leaves the table and does not return any more.
- The number of toilet visits in relation to the number of customers present in the restaurant (including the customers present in the outdoor area)
- The duration of the toilet visits, defined as the time between entering and exiting the toilet area.
- The activities, performed by staff members, per table including:
  - The start and end time of each activity, defined as the staff member arriving at the table and then leaving again.
  - The type of each activity whereby three main types are identified: (1) Serving food and drinks, (2) Collecting empty plates, and (3) Taking orders, asking feedback and others.

4.2. Parameter derivation from the empirical data

The empirical data is used to estimate the parameter values of the scheduling model. For the customer sub-model, data is available to estimate the parameter pertaining to the toilet visits. The probability of a customer visiting the toilet during their visit is computed by $p_{toilet} = \frac{\#\text{ toilet visits}}{\text{average }\#\text{ customers inside and outside}}$ which results in a value of 0.4. To obtain the distribution which best fits the observed distribution of toilet visit durations, all the different continuous distributions of the OpenTURNS Python package [30] are fitted to the data. To ensure that all durations that can be drawn from the fitted distribution are realistic (i.e. positive, finite and falling within a reasonable range), truncated versions of the continuous distributions are used. The lower and upper truncation boundaries are set to respectively the minimum (0.3 min) and maximum (7.3 min) measured durations. Out of all tested continuous truncated

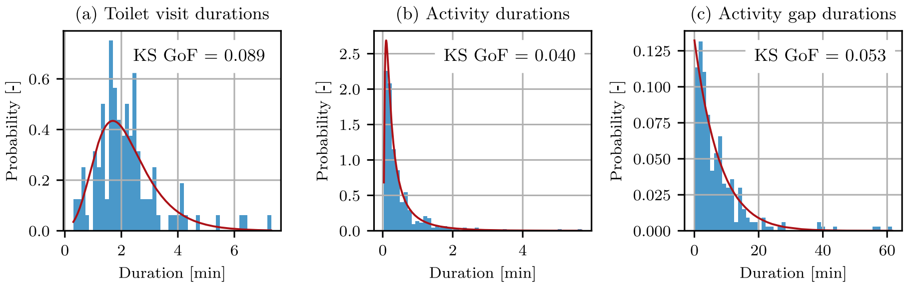

Fig. 6. The empirical data (histogram), the best fitting distribution (line) and the GoF value for respectively: (a) The distribution of toilet visit durations, (b) The distribution of activity durations, and (c) The distribution of the gaps between the activities.

distributions, the truncated version of the Weibull Maximum Extreme Value (see Eq. (10)) fitted best to the empirical data with a Kolmogorov–Smirnov (KS) test value of 0.089. This fit is achieved with the parameters $\alpha = 5879.16$ $\beta = 5016.68$ and $\gamma = 5018.38$. Graph (a) of Fig. 6 presents both the histogram of the empirical distribution and the line representing the fitted distribution.

$$f_X(x) = \frac{\alpha}{\beta}\left(-\frac{x-\gamma}{\beta}\right)^{\alpha-1} \exp\left(-\left(-\frac{x-\gamma}{\beta}\right)^{\alpha}\right), \quad x \in (-\infty;\gamma] \tag{10}$$

For the staff sub-model, the values for five different parameters can be derived from the empirical data. The distribution of activity durations is derived using the same method as is used to derive the toilet visit distribution. That is, fit all truncated versions of the continuous distributions of the OpenTURNS package and selecting the best fitting distribution. The truncation values are again the minimum and maximum values measured in the data which are respectively 0.03 and 5.7 min. Graph (b) of Fig. 6 presents the empirical data and the fitted distribution. In this case a truncated lognormal distribution (see Eq. (11)) produces the best fit with the parameters $\mu = 2.74$, $\sigma = 0.31$ and $\gamma = 1.12$.

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma(x-\gamma)} \exp\left(-\frac{1}{2}\left(\frac{\log(x-\gamma)-\mu}{\sigma}\right)^2\right) \quad x \in [\gamma;+\infty[\tag{11}$$

From the sequence of activities per table and the accompanying start and end times, the gaps between the activities can be computed. As the previous section explains, a Dirichlet distribution is used to compute the gaps between a sequence of activities. For the Dirichlet distribution we would ideally like to estimate a separate parameter (θ) for each gap in a sequence of activities (i.e. a parameter for the first gap, for the second gap and so on). However, the sample (23 gap sequences of different lengths) is too small to do so, hence we assume that the gap duration is independent of its location in the sequence. Fitting the Dirichlet distribution to the sample of all gaps results in an estimated parameter value $\theta = 1$ (see Eq. (9)). The empirical distribution and the fitted distribution are presented in Fig. 5, in this case in graph (c).

The recorded type of each activity provides a basis for estimating the number of activities that require a staff member to first pick-up something at a base area (bar, kitchen etc.) before heading to a table and/or require a staff member to drop-off something afterwards. This was not recorded explicitly, however, observations showed that in most cases that something was served or empty plates and glasses were taken, this entailed a staff member respectively first going to the base area or returning to it before engaging in a subsequent activity. So, we consider the number of serving activities divided by the total number of activities as a reasonable estimate of the probability that a staff member first goes to the base area. Based on the observations this probability is 0.4. For the probability that a staff member goes to the base afterwards the same logic is used but now the probability is estimated by dividing the number of collecting empty plates activities by the total number of activities. This results in a probability of 0.3.

Lastly, the empirical data also provides information on the number of activities performed after customers have left. The data shows that either two or three activities are performed by the staff after the customers have left the table. Furthermore, it is equally likely that this number of activities is two or three. So, in the model the value is drawn from the set of these two numbers whereby each has an equal probability of getting selected.

5. Simulation experiments: Methodology

This section introduces the setup of the simulation experiments, whose results are presented in the next section. First the NOMAD pedestrian models is briefly introduced. Accordingly, the scenarios are presented followed by the description of the quantification of the virus transmission risk. The goal of these simulation experiments is to investigate the model's ability to reproduce the movement dynamics in a restaurant and especially to investigate how these insights can be used to inform policy.

5.1. The NOMAD pedestrian model

The NOMAD pedestrian model [31] is used to model the route choice and operational walking behaviour. We chose NOMAD because of our extensive experience with this microscopic model and the fact that it has been extensively calibrated and

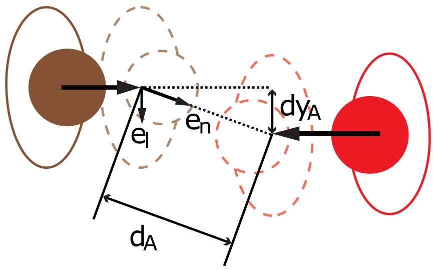

Fig. 7. NOMAD uses the anticipated positions of the pedestrian to determine the anticipated distance between the pedestrians (d_A) and in case of pedestrians moving along the anticipated lateral distance (d_{y_A}).

validated [32] but any other microscopic model with an integrated route choice model would also suffice. The route choice modelling of NOMAD is based on static precomputed directional fields. Each destination has such a field and for each location in the walkable space the field provides a directional vector. The directional vectors point towards the shortest route that also ensures enough separation from any obstacle.

The operational walking model is a force-based model which at every time step computes the acceleration of a pedestrian using three social forces. The model also includes a physical force which pedestrians experience when coming into physical contact with an obstacle or another pedestrian. This generally only occurs in high density situations which a restaurant is not. Eq. (12) shows how the three social forces and a small random fluctuation term ($\boldsymbol{\eta}$) result in the acceleration.

$$\mathbf{a} = \mathbf{a}_f + \mathbf{a}_o + \mathbf{a}_r + \boldsymbol{\eta} \tag{12}$$

The first of the three social forces is the path following force ($\mathbf{a}_f$). This force ensures that a pedestrian follows their desired path to their destination (Eq. (13)).

$$\mathbf{a}_f = \frac{v_{des} \cdot \mathbf{e}_{des} - \mathbf{v}}{\tau} \tag{13}$$

where $\mathbf{v}$ is the pedestrian's current velocity, the v_{des} is the pedestrian's desired speed, $\mathbf{e}_{des}$ the desired direction as obtained from the direction field and τ is the relaxation parameter which affects how stringently the pedestrian tries to stay on their desired path.

The obstacle force ($\mathbf{a}_o$) prevents pedestrians from walking into obstacles (Eq. (14)).

$$\mathbf{a}_o = a_W \sum_{o \in O} -\mathbf{e}_o \begin{cases} 1 & \text{if } 0 < d \leq d_{shy}/2 \\ 2(1 - d/d_{shy}) & \text{if } d_{shy}/2 < d \leq d_{shy} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where a_W is the parameter defining the size of the force in relation to the other forces, O is the set of all obstacles, $\mathbf{e}_o$ is the unit vector pointing to the closest point on the obstacle in relation to the pedestrian's current position, d in the distance to the closest point on the obstacle and d_{shy} is the parameter which defines the minimum distance a pedestrian ideally wants to keep to any obstacle.

The $\mathbf{a}_r$ force is the social force a pedestrian experiences from other pedestrians and this force prevents pedestrians from bumping into each other (Eq. (15)). The computation of the pedestrian interaction forces are based on the anticipated positions of the pedestrians (see Fig. 7).

$$\mathbf{a}_r = -\sum_{p \in P} a_0 \mathbf{e}_n \cdot e^{-d_A/r_0} + \begin{cases} a_1 \mathbf{e}_l \cdot e^{-d_A d_{y_A}/r_1} & \text{if moving in opposing directions} \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

where a_0 is the parameter defining the size of the force in relation to the other forces, P is the set of all other pedestrians in the neighbourhood, $\mathbf{e}_n$ is the unit vector pointing from the pedestrian's anticipated position to the anticipated position of the other pedestrian, d_A is the anticipated distance between the pedestrians and r_0 is the parameter which relates the distance between the pedestrians to the strength of the force. In the case of pedestrians moving in opposing directions an additional lateral force is added where a_1 is the parameter defining the size of the force, $\mathbf{e}_l$ is a unit vector orthogonal to the pedestrian's velocity, d_{y_A} is the anticipated lateral distance between the pedestrians and r_1 is the parameter which relates the lateral distance between the pedestrians to the strength of the force.

Finally, by integrating the computed acceleration twice, using Euler's method, the new position of the pedestrian is obtained. For a more detailed description of the NOMAD model the reader is referred to Campanella [31]. In this research, the python implementation PyNOMAD is used [33]. For more details and access to the software, the reader is referred to Sparnaaij et al. [33].

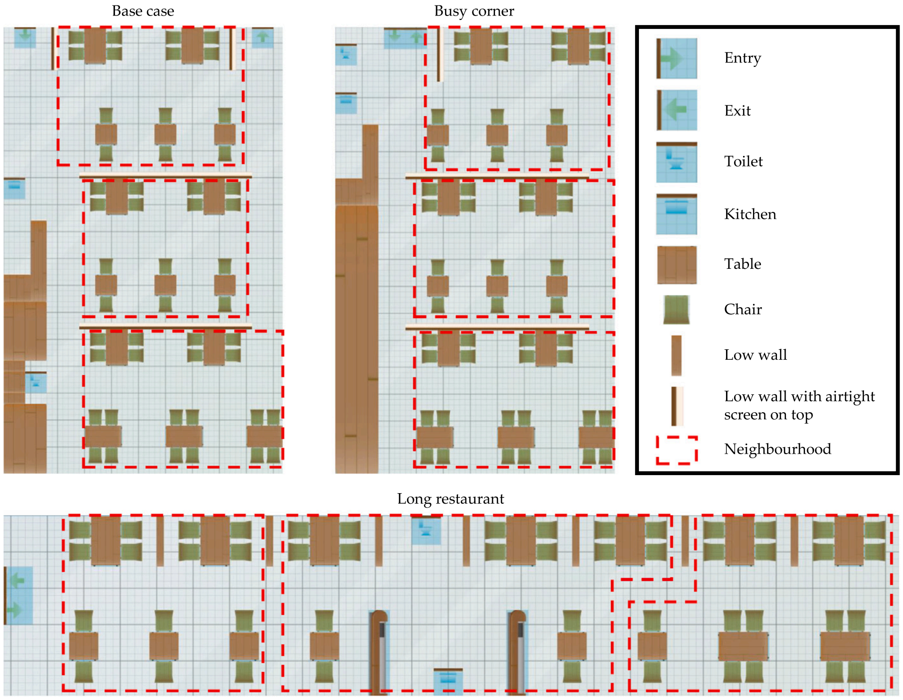

Fig. 8. The three different restaurant layouts used for the simulation experiments.

5.2. Scenarios

A number of scenarios with different layouts and setups are used to gain insights into the following question: How can the model be used to identify factors that generally affect the virus transmission risk in a restaurant, and thus form good targets for general policy, and to identify what factors are more restaurant specific? Additionally, the results from the scenarios are used to get insight into the behaviour of the new activity model. A total of 15 scenarios are used, which are the combination of three different layouts and five different scheduling setups.

The three layouts are presented in Fig. 8, and set up to feature different movement scenarios that might occur in a restaurant.

Base case layout: This layout is a spacious layout with separate entry and exit locations. There are also some low walls present as separation whereby an airtight screen has been added on top to prevent contacts. This base case is designed to limit the number of interactions between the visitors.

Busy corner layout: This layout is very similar to the base case layout except that all activities (entry, exit, kitchen and toilet) are all located in the same corner of the restaurant. This is expected to lead to more contacts but the question is if this also leads to a higher probability of virus transmission.

Long restaurant layout: This layout has all the tables placed along a single corridor which means more people are likely to pass them which in turn likely increases the probability of contacts. This is especially the case for the tables situated near the entrance and exit and near the middle where the toilet and the kitchen are.

To ensure that the simulation outcomes can be compared, the three layouts all have the same number of tables and chairs in the same configuration, namely nine tables with four chairs and six tables with two chairs. Furthermore, all tables are placed such that, in line with the Dutch regulation surrounding Covid-19, sitting customers are always more than 1.5 m separated from customers sitting at other tables.

The activity choice and scheduling model provides a number of inputs (see 3.2.1 and 3.3.1) to represent different restaurant setups. Based on these inputs, five different setups have been designed. One base setup and four variations. To ensure that the setups can be easily compared they all have three staff members, the visits are all planned within a 5-h time period, they all have the same total visit time (the sum of the visit times of all groups) and the average number of times staff members visit a table per hour in relation to a group's visit is also equal. This last number was set at around 8.8 visits an hour based on the empirical data. Table 1 presents the five setups and the inputs which define them.

Base setup: In this setup, 25 customer groups are randomly scheduled to visit within the 5-h time period whereby they have an average visit time of about 2.4 h during which the staff will visit the table 21 times and there are no serving neighbourhoods.

Neighbourhoods setup: This setup is the same as the base setup in every respect except that each of the three staff members has their own neighbourhood. Fig. 8 presents these neighbourhoods. We expect that using neighbourhoods will lead to customers having

Table 1
The five different restaurant setups. "–" indicates that the value does not differ from the base setup.

Name	Average visit time [h]	Group count	# Time slots	Neighbourhoods	# Staff visits
Base	2.4	25	1	No	21
Neighbourhoods	–	–	–	Yes	–
Time slots	–	–	2	–	–
Shorter visits	1.7	35	–	–	15
Longer visits	4	15	–	–	35

contact with fewer staff members but that the customers will interact with the staff member serving their particular neighbourhood more often compared to the base setup.

Time slots setup: This setup is the same as the base setup in every respect except that two time slots are used. This setup schedules the customer groups in one of two time slots whereby there is a gap between the time slots during which there are no customers present in the restaurant. We expect that customers will have fewer interactions with other customers as they can only interact with customers who are visiting the restaurant in the same time slot.

Shorter visits setup: The shorter visit setup has more customer groups visiting the restaurant but has each group stay for a shorter duration and has staff members performing fewer activities at the table. We expect that both staff members and customers will have come into contact with more people but that the durations of these contacts will be shorter compared to the base setup.

Longer visits setup: This setup is the opposite of the shorter visit setup with longer visit durations, more staff activities at the table but fewer customer groups visiting the restaurant. The expected difference compared to the base setup are also the opposite of the shorter visit setup. Namely, fewer but longer contacts between people in the restaurant.

Together, these five setups provide insight into how the input of the activity model and thus the restaurant setup affect contacts between people. And, in combination with the three layouts, how this might interact with the restaurant layout.

5.3. Pedestrian contact quantification

As we state in the introduction, we solely focus on the pedestrian model's contributions and capabilities in assessing virus transmission risks in restaurants. So, to assess the virus transmission probabilities and the activity model's behaviour we quantify the contacts between people in the restaurant and especially the close contacts between people. The concept of close contacts is used in this case to obtain insight into the probability of virus spread because the model does not model virus spread explicitly. It does provide proximity data however and health agencies world-wide [34–36] use this kind of proximity information in their contact tracing to determine who is at high risk of having been infected after a contact with an infectious person (i.e. who is a close contact). Therefore, we deem the concept of close contacts to be useful in determining the relation between proximity and the probability of virus spread.

The choice of using the concept of close contacts to assess the probability of virus spread has a number of consequences for the setup of the simulations and analysis of the results. Therefore, it is important to define what we consider to be a close contact in the context of this paper. The exact definition that health agencies use to determine if someone is a close contact differs slightly per country. But the consistent elements are:

1. Being a member of the same household as the infected person.
2. Having been in contact at less than 1.5 to 2 m for more than 15 min. This is cumulative for a 24-h period.
3. Having been in contact at less than 1 m for one or more consecutive minutes.
4. Having been in high risk contact which includes having been coughed upon, having had direct physical contact such as kissing or having had a face-to-face conversation at less than one metre distance.

The model can provide the information required by the second and third elements. Regarding the first, it is assumed that customers that sit at the same table will very likely either be people from the same household, have face-to-face conversations and/or be within 1.5 to 2.0 m of each other for more than 15 min. Hence, contacts between these people are left out of the analysis as we presume these to be close contacts of each other anyway. The same holds for the contacts between staff members as we presume them to work in close proximity for prolonged periods.

The model only provides which contacts would be close contacts in case one of the two people would be infectious. That is, which contacts are risky but for which it is unknown if either of the people is infectious and thus if virus transmission occurred. This is a limitation as the virus spread risk is also dependent on the number of infectious people in the restaurant. However, knowing how many risky contacts occur is still valuable as this, combined with an estimate of the likelihood that one or more people in the restaurant are infectious, provides insight into the likelihood that virus transmission occurred. That is, the more risky contacts and the more likely that one or more people in the restaurant are infectious, the higher the probability of close contacts occurring and the higher the probability that the virus is transmitted to one or more people that are not sitting on the same table.

To capture the contacts between people and to determine if these contacts are risky contacts a weighted connectivity graph is used. The graph is computed as follows:

$$G = \left\{ W\left(\{a_i, a_j\}\right) | a_i, a_j \in A \text{ and } i \neq j \text{ and } a_i \notin g(a_j) \text{ and } \exists\, t \mid d_{i;j}(t) \leq D_{cutoff} \right\} \tag{16}$$

where A is the set off all agents in the simulation, $g(a_j)$ the set of agents forming the group to which agent j belongs, t the time, $d_{i;j}(t)$ the distance between pedestrian i and j at time t and D_{cutoff} the distance cut-off value which is defined by the close contact definition.

The weight function $W\left(\{a_i, a_j\}\right)$ captures the duration of the contact (Eq. (17)). This duration can be cumulative over a 24-h period, in line with the second element of the close contact definition. Or it is the maximum consecutive time two people were in contact with each other, in line with the third element of the close contact definition.

$$W\left(\{a_i, a_j\}\right) = \begin{cases} \sum_{t \in T} \Delta t_{\{d_{i;j}(t) \le D_{cutoff}\}}(t) & \text{if cumulative} \\ \max \sum_{t \in T} \Delta t_{\{d_{i;j}(t) \le D_{cutoff} \text{ and } (d_{i;j}(t+\Delta t) \le D_{cutoff} \text{ or } d_{i;j}(t-\Delta t) \le D_{cutoff})\}}(t) & \text{if consecutive} \end{cases} \tag{17}$$

where Δt is the time step size and T is the set of time steps of the simulation. Within this paper, any contact is considered a risky contact if the contact takes place at less than 1.5 m for more than 15 min cumulatively and/or takes place within 1 m for 1 or more minute consecutively.

Both the NOMAD model and the novel activity model are stochastic in nature. By replicating each scenario 140 times we ensure, with a high degree of confidence (99%), that all connectivity graphs of all scenarios have converged. This ensures that we can have high confidence that any differences we find during the analysis are not the result of the stochastic nature but are actual differences.

6. Result from the simulation experiments

The results from the 15 scenarios provide insight into how likely is it that people experience risky contacts in a restaurant setting. We investigate how this likelihood depends on various aspects. Firstly, we look into the question if the likelihood of experiencing risky contacts is different for customers and staff. Secondly, we investigate how the layout and setup of the restaurant impact the likelihood of risky contacts. Thirdly, we look at the sensitivity of the results to the close contact definition. And lastly, we investigate the sensitivity to the number of infectious people in the restaurant.

During the analysis different results are compared to each other. To determine differences are statistically significant and relevant three supporting values are computed:

1. The p-value (p): The Cramer–von-Mises test is used to determine if the two results being compared are samples drawn from the same distribution. The null hypothesis is that they are and this hypothesis is strongly rejected when $p < 0.01$ and not rejected at all when $p \geq 0.05$.
2. Cliff's delta (δ): This is an effect size value between −1 and 1, whereby the closer to 0, the smaller the difference in central tendency and the less relevant the difference is (even if statistically significant).
3. The Kolmogorov–Smirnov (KS) statistic (D): A value between 0 and 1, whereby the larger the value the bigger the largest difference in the cumulative distributions of the two results being compared. This value is used to supplement Cliff's delta. For example, in case the effect size computed by Cliff's delta is small, and the difference in central tendency is small and not relevant, but the KS statistic is larger, the difference is still relevant as the results are clearly distributed differently.

6.1. The likelihood of risky contacts and the difference between customers and staff

Customers and staff have very different activity schedules and therefore very different movement patterns in restaurants. The question is how much this affects their likelihood to experience risky contacts and additionally, for customers, if this is primarily caused by their interactions with the staff or with customers from other groups. To investigate this we identify three types of interactions:

1. *Customer–customer:* These are the interactions between customers from different groups and these provide insight into the risk customers pose to each other.
2. *Customer–staff:* These are the interactions between customers and waiting staff from the perspective of the customer. These interactions provide insight into the risk that staff members pose to customers.
3. *Staff–customer:* These are the interactions between customers and waiting staff from the perspective of the staff. These interactions provide insight into the risk that customers pose to staff members.

The simulation data shows that in none of the 15 scenarios risky contacts occur between customers. This is not due to the fact that customers never come within 1.5 m of each other as Fig. 9 shows. The graphs show that the majority of customers spend some time within 1.5 m of many other customers. Yet, for all these contacts it holds that the cumulative time is well below 1 min and in most cases even just a few seconds. This suggest that the contacts are the result of customers briefly passing each other when one or both are moving through the restaurant.

The interactions between customers and staff are far more likely to lead to risky contacts. Fig. 10 presents the average number of risky contacts from the viewpoint of the customers and staff in respectively graph (a) and (b). From these graphs it is clear that the average waiter has around the 10 to 15 risky contacts during one shift whilst the average customer has only a 30 to 80 percent chance to have just one risky contact. This shows that customers, as a group, pose a far larger risk to the staff than vice-versa. This is also not unexpected as there are far more customers than staff.

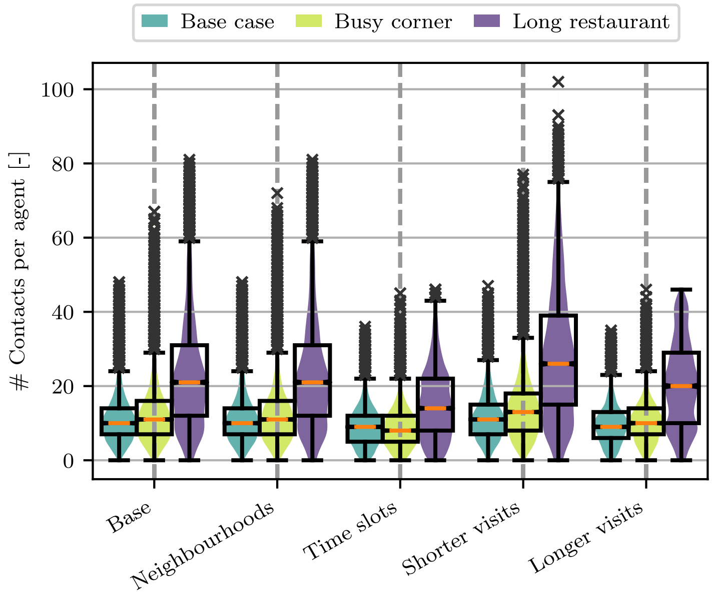

Fig. 9. The distribution of the number of contacts each customer has with customers from other groups. A contact in this context means that the customers have been within 1.5 m of each other at least once during their stay in the restaurant. The box plot present the median value (the orange line), the locations of Q1 and Q3 (the bottom and top of the box), and the outliers (x) which are more than 1.5 times the IQR removed from the Q1 or Q3 border. The violin plot below the box plot presents roughly the shape of the distribution.

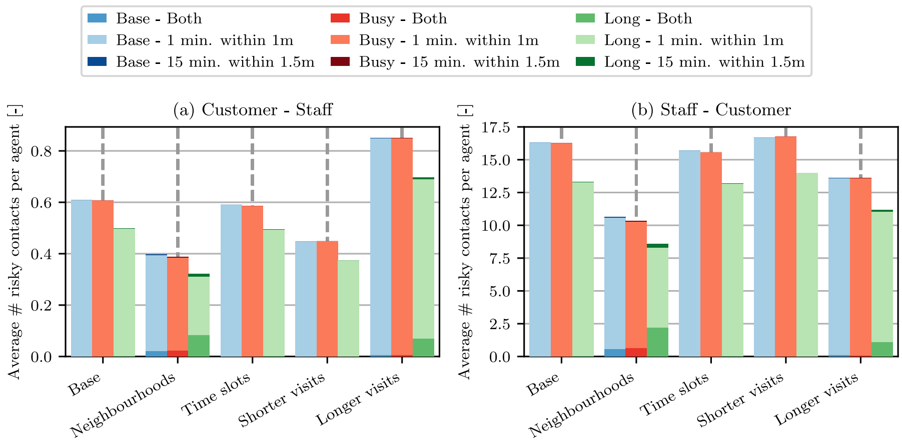

Fig. 10. (a) The average number of risky contacts a customer has with members of the staff. (b) The average number of risky contacts a staff member has with customers. Each bar has three parts displaying the fraction of the risky contacts that is caused by (1) both the 1-min and 15-min criteria being met, (2) only the 1 min criterion being met and (3) only the 15 min criterion being met.

A second observation is that by far most of the risky contacts between customers and staff are due to them being within 1 m of each other for a minute or more. Only in a few cases is it the case that the risky contact is caused by them being with 1.5 m of each other for cumulatively more than 15 min or that both these criteria are met.

The graphs in Fig. 10 provide the average number of risky contacts per agent but do not provide information on how equally these contacts are distributed over the customers and staff. Graph (a) in Fig. 11 presents how these risky contacts are distributed over the customers. In all 15 scenarios there are 3 staff members so the number of risky contacts a customer can have with a member of staff is a number between 0 and 3. The bars show the fraction of customers that have risky contacts with respectively 1, 2 or all 3 staff members. The remainder of the customers has no risky contacts with any staff member. For example, in the base case layout and base setup (the left most bar in graph (a)) about 18% of the customers has a risky contact with only 1 staff member, about 15% has a risky contact with 2 staff members, about 4% with three staff members and about 63% has no risky contact with any of the staff members. These bars show that in all cases a minority of the customers has one or more risky contacts with staff members

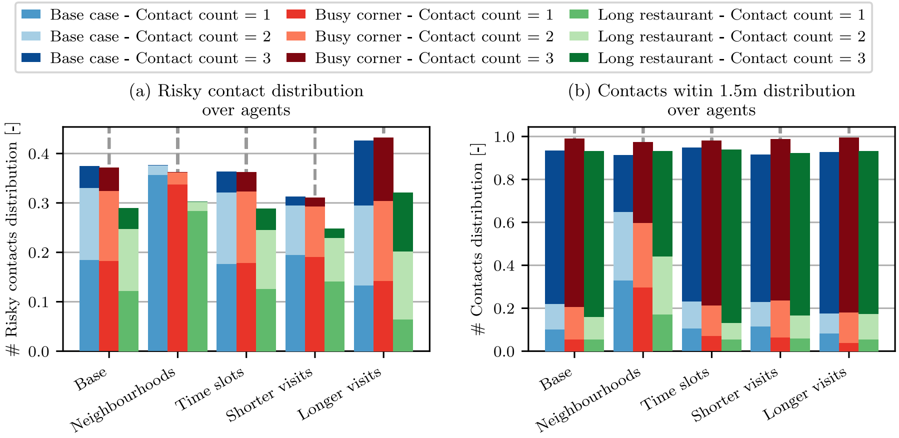

Fig. 11. The distribution of the number of (a) risky contacts and (b) contacts within 1.5 m for customer–staff interactions. Each bar presents the fraction of cases where a customer had (a risky) contact with respectively 1, 2 or all three of the staff members. In the remainder of the cases the customer had no contact at all with any of the staff members.

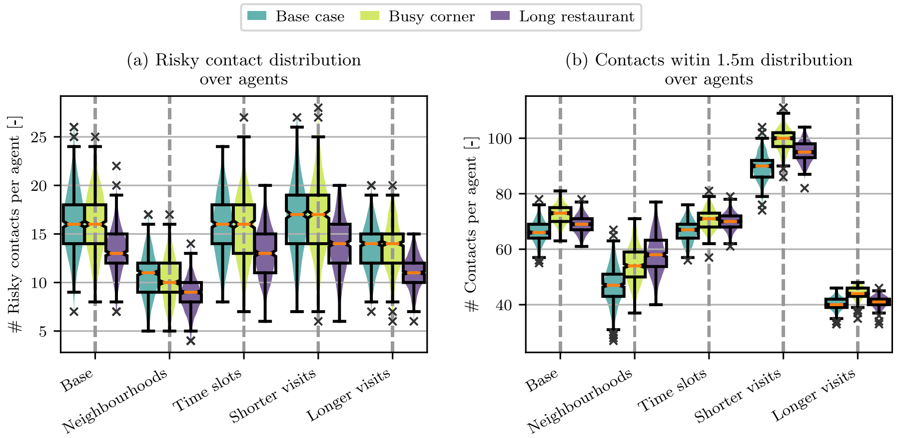

Fig. 12. The distribution of the number of (a) risky contacts and (b) contacts within 1.5 m for staff–customer interactions.

as all bars have a height below 0.5. This despite the fact that, as graph (b) Fig. 11 shows, most customers come within 1.5 m of all three staff member during their stay. Furthermore, they also show that it is not uncommon for some customers to have 2 or 3 risky contacts with the waiters.

Graph (a) of Fig. 12 shows that the risky contacts between the customers and the staff are also not necessarily equally distributed over the staff members. The box/violin plots show that the majority of cases are centred around the median value which indicates that in most cases the differences are in the order of a few contacts. However, in all scenarios the difference between the maximum and minimum number of risky contacts and the median number of risky contacts is large (i.e. in the order of 5 to 10 contacts). This means that some members of staff run a far higher risk of getting infected due to the larger number of risky contacts than other staff members. Comparing these distributions to the distribution of the number of contacts per staff member (Fig. 12.b), it is clear that the variance in the number of contacts is, relatively, much smaller than the variance in the number of risky contacts.

Now, observing that:

1. the distributions of risky contacts between customers and staff do not correlate well with distributions of the number of contacts
2. most of these risky contacts are caused by the 1-min criterion

it suggests that these contacts occur when a member of staff is standing close to the table for more than a minute (which is a rather long activity such as explaining the card and taking an order for example). In this case most people at the table are probably at a distance of more than 1 m from the staff member but clearly not all which would explain why many have no risky contact with a member of staff and some with all three. It would also explain why in some cases some staff members have a far larger number of risky contacts compared to others. If these staff members happen to mostly perform the longer activities (i.e. those longer than a minute) than they would logically also have the most risky contacts.

Table 2
The statistical test values for significance and effect size for the comparison between the different restaurant layouts.

		Busy corner			Long restaurant		
		p	δ	D	p	δ	D
Customer–staff	Base	Stat. insig.			0.000	0.077	0.086
	Neighbourhoods	0.000	0.013	0.015	0.000	0.073	0.074
	Time slots	Stat. insig.			0.000	0.068	0.074
	Shorter visits	Stat. insig.			0.000	0.059	0.064
	Longer visits	Stat. insig.			0.000	0.089	0.105
Staff–customer	Base	Stat. insig.			0.000	0.552	0.445
	Neighbourhoods	Stat. insig.			0.000	0.539	0.431
	Time slots	Stat. insig.			0.000	0.467	0.358
	Shorter visits	Stat. insig.			0.000	0.450	0.362
	Longer visits	Stat. insig.			0.000	0.613	0.486

So, in short, the data suggest that the driver of the risky contacts in a restaurant setting are the contacts between customers and staff in the case that a staff member is standing close to the table for a minute or more. Furthermore, the data show that staff members run the largest risk of getting infected, provided someone in the restaurant is infectious, as well as pose the largest risk to infect visitors in a restaurant at a large scale.

How can these results be used to inform policy? First and foremost, the results suggest that any virus transmission mitigation policy should focus on preventing virus transmission from the staff to the customers. This because any infectious staff member can potentially infect many customers (i.e. due to having many risky contacts) whilst an infectious customer is not very likely to infect staff members (more than half of the customers does not even have a single risky contact). These mitigation strategies could, for example, entail staff getting tested regularly to prevent them from working while potentially infectious or staff wearing face masks to reduce the probability of them transmitting the virus if they are unknowingly infectious. Policies aimed at preventing customers from infecting staff member can also be relevant as the likelihood of this happening is not completely unlikely. This could, for example, involve customers having to show a recent negative test result before entering the restaurant. Lastly, any measure that could prevent staff members standing within 1 m of a customer for a minute or more would greatly reduce the risks of virus transmission in both ways.

6.2. The impact of the restaurant layout and setup on the likelihood of experiencing risky contacts

Fig. 10 shows that both the restaurant layout and the setup impact the number of risky contacts between customers and staff. Additionally, the results shows that customers have no risky contacts with each other. Therefore, the focus of this part of the analysis is solely on the interactions between customers and staff.

The first thing that the graphs (a) and (b) in Fig. 10 and graph (a) in Figs. 11 and 12 tell us is that the busy corner layout does not result in a significant different number of risky contact or a different distribution of those risky contacts over the agents compared to the base case layout. This is also supported by the fact that the differences are statistically insignificant and/or the effect sizes are very small (see Table 2).

The long restaurant layout has more impact. This layout results in fewer risky contacts than the other two layouts. From the customer's point of view this effect is still very small as the values in the table show. From the staff's point of view the difference is much larger. This is likely due to the difference in group size (3 staff members versus 48 to 112 customers depending on the setup). As to the reason why the long restaurant layout results in fewer risky contacts, this is likely due to the relation between the layout and the location where staff stands when performing an activity at a table. In the long restaurant layout, this location is generally such that the staff member is standing close to fewer customers than is the case for the other two layouts.

For all three layouts, the restaurant setup can have an effect on the number of risky contacts between the customers and the staff. The use of time slots or shorter but more visits has no significant or a very small effects as is shown in Table 3. Although in the shorter visit case it must be noted that the effect, though small, is larger from the customer's point of view than from the staff's point of view. This is caused by the fact that there are fewer interactions per visit between the staff and customers but more visits. So, the customers experience slightly fewer risky contacts with the staff whilst the staff experiences and equal number of risky contacts.

The use of neighbourhoods has an impact on the number of risky contacts as well and especially on the distribution over the agents. From the customer's point of view, the number of risky contacts drops slightly (Fig. 10.a) and it is more likely that those risky contacts are partly caused by being with 1.5 m for more than 15 min. However, the impact of the neighbourhoods is best seen in how the risky contacts are distributed over the customers. Compared to the other setup, the overwhelming majority of the customers in the neighbourhood setup only have a risky contact with one member of staff (provided they have at least one risky contact) and only very few have two or three risky contacts. This is also supported by the fact that the Kolmogorov–Smirnov (KS) test statistic is larger than the cliff's delta value indicating that the difference in distribution shape is larger than the difference in central tendency (Table 3).

From the staff's point of view, neighbourhoods have a larger impact compared to the impact for the customers. Graph (b) in Fig. 10 and graph (a) in Fig. 12 show that the number of risky contacts decreases significantly when the neighbourhood setup is

Table 3
The statistical test values for significance and effect size for the comparison between the different restaurant setups.

		Customer–staff			Staff–customer		
		p	δ	D	p	δ	D
Base case	Neighbourhoods	0.000	0.063	0.170	0.000	0.857	0.695
	Time slots	0.030	0.012	0.012	0.019	0.100	0.084
	Shorter visits	0.000	0.080	0.073	0.076	−0.067	0.071
	Longer visits	0.000	−0.092	0.103	0.000	0.496	0.367
Busy corner	Neighbourhoods	0.000	0.070	0.164	0.000	0.875	0.736
	Time slots	0.000	0.011	0.009	0.007	0.117	0.091
	Shorter visits	0.000	0.077	0.069	0.052	−0.078	0.071
	Longer visits	0.000	−0.097	0.101	0.000	0.485	0.383
Long restaurant	Neighbourhoods	0.000	0.032	0.148	0.000	0.905	0.786
	Time slots	0.181	0.002	0.005	0.116	0.029	0.064
	Shorter visits	0.000	0.053	0.060	0.000	−0.154	0.164
	Longer visits	0.000	−0.061	0.089	0.000	0.571	0.474

compared to the base setup. Table 3 additionally shows that it is mainly a change in central tendency of the distribution and not in the shape (Cliff's delta is larger than the KS statistic). The fact that the impact of using neighbourhoods is larger for the staff than for the customers can be explained by the big difference in group size.

Longer visits also impact the number of risky contacts between customers and staff. The impact is the opposite of the impact shorter visits have. For customer the number of risky contacts increases slightly and also the distribution of those risky contacts changes with more customers having two or three risky contacts. For staff the number of risky contacts decreases significantly though less strongly than in the neighbourhood case. The main explanation for this trend is the fact that the probability that a customer experiences a risky contact is closely tied to the visit duration of that customer. For the staff it is also tied to the number of different people visiting the restaurant during the shift of that staff member.

Overall, the data shows that both the restaurant layout and the setup can influence the number of risky contacts and that this influence is larger on the number of risky contacts staff experience than the number of risky contacts customers experience. Furthermore, the main driver for the difference between customers and staff are the number of possible risky contacts (maximally 3 versus maximally 48 to 112 depending on the setup).

How can these results be used to inform policy? Layout can impact the number of risky contacts but not necessarily. This means that a policy that is effective in one restaurant is not necessarily as effective in a restaurant with another layout. It also means that results obtained using one specific layout cannot necessarily be transferred to another restaurant with another layout. Hence, it is relevant to model individual restaurants to get more precise insights or to model a wide range of commonly occurring layouts. The setup has a bigger and more consistent impact on the number of risky contacts. Especially the use of neighbourhoods can reduce the risks of virus transmission in a restaurant. This suggest that advising restaurants to use serving neighbourhoods might be one of the possible generally applicable mitigation policies. Furthermore, the number of activities staff performs at a table in relation to the visit of a group of customers correlates positively with the number of risky contacts. So, any policy that can reduce the number of times that staff has to visit the table will reduce the risks.

6.3. The sensitivity to the close contact definition

The definition of a close contact depends on two threshold values; one time and one distance threshold. These values differ slightly per country which indicates some uncertainty about the exact values of these thresholds. Therefore, this part of the analysis focusses on the question how sensitive the results are to the exact thresholds. Here, we use the following set of alternative threshold settings:

1. Removing the within 1 m for a minute or more criterion leaving only the 15 min or more cumulatively at 1.5 m or less criterion.
2. Increasing the distance threshold for the within 1.5 m for 15 min or more (cumulatively) criterion to 1.8 m or 2.0 m
3. Decreasing the cumulative time threshold for the within 1.5 m for 15 min or more (cumulatively) to 10 min or 5 min
4. Increasing the distance threshold and decreasing the cumulative time threshold. These are the four combinations of the two variations described above (i.e. d = 1.8 & t = 10, d = 1.8 & t = 5, d = 2.0 & t = 10, d = 2.0 & t = 5).

The three graphs in Fig. 13 present the relative change in the average number of risky contacts between customers and staff for different variations to the close contact definition. The graphs show that some variations to the close contact definition would result in significant changes to the number of risky contacts. Decreasing the cumulative time threshold to 5 min and increasing the distance threshold results in two to almost four-fold increases in the average number of risky contacts between customers and staff when compared to the base definition of a close contact. Contrary, removing the 1 min or more within 1 m criterion leads to a 100% decrease of the number of risky contacts (i.e. no more risky contacts) in most cases. This is expected because the data in Fig. 10 shows that the overwhelming majority of risky contacts is caused by this criterion.

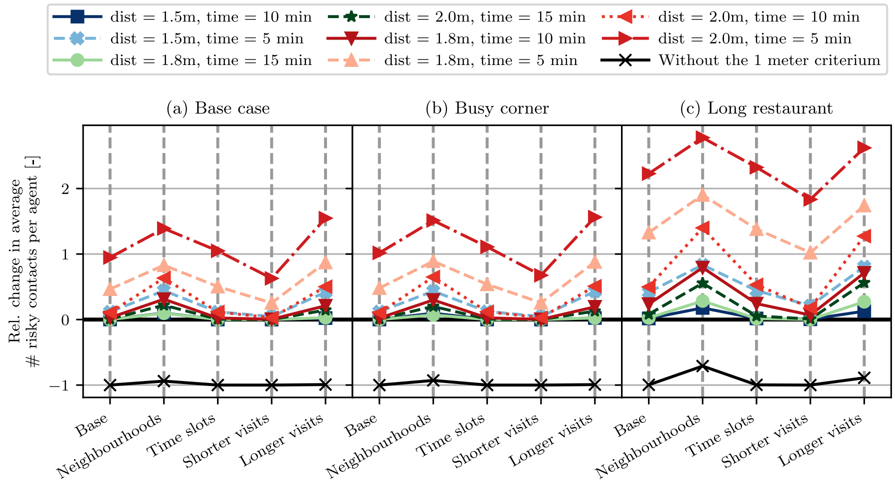

Fig. 13. The sensitivity of the number of risky contacts between customers and staff to the close contact definition. A change of 1 indicates a 100% increase in the number of risky contacts compared to the base definition of close contact.

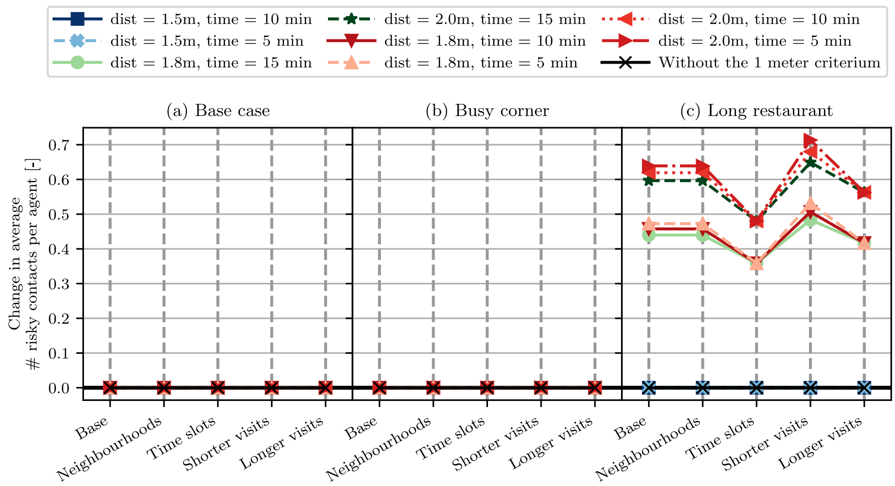

Fig. 14. The sensitivity of the number of risky contacts between customers to the close contact definition.

The graphs also show two other notable patterns. Firstly, the sensitivity depends on the setup of the restaurant independent of the layout. For all three layouts we see that the neighbourhood and longer visit setups are more sensitive to the changes in the definition and the shorter visit setup less sensitive. Secondly, the sensitivity also depends on the layout but not necessarily. The base case layout and the busy corner layout show a similar sensitivity to the change in the close contact definition. The long restaurant layout is generally more sensitive to changes in the close contact definition. The main reason for the higher sensitivity of the long restaurant layout and the neighbourhood and longer visit setups is that in those case the chance that a customer has many interactions with a particular staff member is higher compared to the other scenarios. The more interactions between a customer and a particular staff member, the more likely that the cumulative contact time exceeds the threshold value.

All in all, the graphs identify that the number of risky contacts between customers and staff is sensitive to the exact definition of a close contact. To what degree the results are sensitive to the close contact definition is context dependent with certain layouts and setups being more sensitive than others.

Fig. 14 presents how much the average number of risky contacts per agent would increase or decrease if the close contact definition changes. The graphs show that a change in the close contact definition would only impact the long restaurant layout. In this case the results are primarily sensitive to an increase in the distance threshold. This is caused by the fact that the chairs in the long restaurant layout are closer to each other than in the other two layout. So, despite the fact that the chairs are placed such that customers from different groups are seated at more than 1.5 m of each other a certain fraction is not placed such that they are 1.8 or 2.0 m from each other. This shows that the chance that one customer infects another customer is highly dependent on how accurate the value for the distance threshold is. If this value is underestimated and a restaurant layout is based on this value, the risk of virus transmission might be underestimated. Certainly, because our analyses in Section 6.1 show that customer customer interactions play no role in the number of risky contacts presuming that this distance threshold is accurate. It also has to be noted that the time slots setup has a noticeable impact in this case due to the fact that tables can only be occupied by a single group during a certain time slot. This leads to fewer possible contacts between customers when they are both seated and this to fewer risky contacts.

Overall, the analysis shows that the number of risky contacts is sensitive to changes in the close contact definition. To which degree the results are sensitive to the close contact definition depends on the layout whereby an important factor is the location where staff stands when performing an activity at a table and how often the same staff member serves the same groups at the same table. Another important factor driving the sensitivity of the results to changes in the close contact definition is the separation between the tables. If the spacing between the tables is based on an underestimated value of the distance threshold, contacts between customers can become risky contacts thus increasing the virus transmission risk in a restaurant.

How can these results be used to inform policy? The results show that the properties of a virus that determine its transferability, for which the close contact definition is a proxy, impact the virus transmission risk in a restaurant. Therefore, the results suggest that a change in the transferability properties of a virus, for example due to a new variant, requires an update of the mitigation policies. The results also show that the model can provide these insights by reinterpreting the simulation results using an updated close contact definition (provided that the layout and setup of the restaurant remain the same). Furthermore, the results show that the mitigation policy of physically distancing guest of different groups by ensuring they are seated with the required physical distance between them is very sensitive to the physical distancing value. If this value is not accurate and too small, the measure becomes ineffective.

6.4. The sensitivity to the number of infectious people in the restaurant

Logically, the number of risky contacts in and of themselves do not provide insight into the probability of a virus being transmitted in the restaurant. It also depends on if and how many of the people present in the restaurant are infectious. In this paper we do not make any assumptions regarding this aspect. However, the number of infectious people can also impact the number of risky contacts more directly. Namely, via the fact that part of the close contacts definition is based on a cumulative time spend within 1.5 m of someone else. The cumulative part of the definition does require it to be within 24 h but does not require it to be cumulative over the contact with a single person. So, imagine there are two infectious people in the restaurant with whom you have been in contact with within 1.5 m but that the cumulative duration of those contacts is individually below the 15 min threshold but combined above it. According to the close contact definition you would still be at risk.

By looking at all possible combinations of each customer's and staff member's contacts we can identify how much their likelihood of experiencing risky contacts increases for different combination sizes. To ensure a fair comparison the number of risky contacts is divided by the number of combinations. The analysis is performed using combination sizes of 1, 2, 3 and 4 (i.e. assuming 1, 2, 3 or 4 people are infectious). This shows that the increase in the number of risky contacts that both staff and customers experience when we take into account the possibility of more than one person being infectious is very small. Only the longer visit setup shows any real effect but even in that case the average number of risky contacts increases only by about 0.2 to 0.3 contacts. So, from this we can conclude that the number of risky contacts is not really sensitive to the number of infectious persons.

6.5. The validity of the activity scheduling model

The data presented so far can also provide insight into the validity of the novel activity scheduling model. The five different restaurant setups each produce different activity schedules for the customers and/or staff. These should also affect the contacts among customers and the contacts between customers and staff. Seven hypotheses about the expected impact of the activity scheduler can be tested when comparing the setups to the base setup. For the interactions between customers one would expect that:

1. The neighbourhood setup does not influence the contact patterns between customers because neighbourhoods only influence the activity patterns of the staff.
2. The shorter visit setup should increase the number of contacts between customer due to more customers visiting the restaurant, and thus, more people moving around in the restaurant. For the longer visit setup the opposite is expected.
3. The time slot setup should decrease the number of contacts between customers. This is expected because there are fewer opportunities for groups to come in contact with each other as each group of customers can only come in contact with groups that share the same time slot.

Based on Fig. 9 and Table 4 we can conclude that all these hypotheses hold. For all layouts, the difference between the neighbourhood setup and the base setup is not significant. Similarly, the time slots setup and the longer visit setup show an increase in the number of contacts compared to the base setup and vice versa for the shorter visits.

For the interactions between customers and staff one would expect that:

4. The time slot setup does not influence the contact patterns between customers and staff. This is expected because this only impacts when the groups can visit the restaurant but not the visit time nor the number of activities of the staff per table.

Table 4
The statistical test values for significance and effect size for the comparison of the number of contacts between the different restaurant setups.

		Base case		Busy corner		Long restaurant	
		p	δ	p	δ	p	δ
Customer–customer	Neighbourhoods	1.000	−0.001	0.927	−0.003	1.000	0.002
	Time slots	0.000	0.169	0.000	0.253	0.000	0.298
	Shorter visits	0.000	−0.067	0.000	−0.148	0.000	−0.168
	Longer visits	0.000	0.112	0.000	0.096	0.000	0.061
Customer–staff	Neighbourhoods	0.000	0.434	0.000	0.435	0.000	0.266
	Time slots	0.000	−0.006	0.000	0.021	0.000	−0.031
	Shorter visits	0.000	0.034	0.000	0.033	0.000	0.018
	Longer visits	0.000	−0.032	0.000	−0.032	0.000	0.013
Staff–customer	Neighbourhoods	0.000	0.988	0.000	0.988	0.000	0.809
	Time slots	0.347	−0.033	0.000	0.314	0.000	−0.218
	Shorter visits	0.000	−1.000	0.000	−1.000	0.000	−1.000
	Longer visits	0.000	1.000	0.000	1.000	0.000	1.000

5. The neighbourhoods setup should decrease the number of contacts between staff and customers because most customers will only be served by one staff member instead of being served by all three.
6. The shorter visits setup should increase the number of contacts between staff and customers from the staff's point of view and for the longer visits setup vice versa. This is excepted due to the number of customers being larger in the shorter visit setup and smaller in the longer visit setup.
7. The shorter and longer visit setups should have a negligible effect on the number of contacts between staff and customers from the customer's point of view. This is due to the fact that given the number of times the staff comes to the table, even in the short visit setup, it is likely that all three staff members will serve the table at least once during the visit.

Graph (b) of Figs. 11 and 12 and the data in Table 4 illustrate that these hypotheses hold. The effect of the time slot setup on the number of contacts is either not significant or has such a small effect size that the difference is not relevant. The graphs and table also show that both the neighbourhood and longer visit setup will significantly decrease the number of contacts from the staff's point of view. And also that this is vice versa for the shorter visit setup. Lastly, the values in Table 4 show that from the customer's point of view the number of contacts with the staff differs very little in the case of the shorter and longer visit setups compared to the base case.

The results consistently show the patterns we would expect. Therefore, it is likely that the results provide a realistic view of the impact different elements have on the likelihood of risky contacts occurring in a restaurant setting. Yet, more extensive modelling study is advised to quantitatively validate the novel activity choice and scheduling model.

7. Conclusions, implications and discussion

This paper presents the application of a pedestrian model to understand the risk of virus transmission in a restaurant. Furthermore, it provides insight into the added value of using pedestrian models to both assess the risk of virus transmission in a restaurant and the effectiveness of mitigation policies. Moreover, it provides a baseline for investigating what the added value is of adding an explicit virus transmission model. Lastly, to accomplish this, the paper presents a novel pedestrian activity choice and scheduling model designed to easily provide realistic movement patterns of waiting staff and customers in a restaurant.

Based on the results of previous section we conclude that a pedestrian model can provide insight into different elements that influence the risk of virus transmission in a restaurant. The results show that the investigated elements influence the virus transmission risk in different ways and to differing degrees. From the results we can conclude that the main factor driving the transmission risk in a restaurant are the interactions between staff and customers that are within 1 m and take more than 1 min. These interactions are the source of the overwhelming majority of all risky contacts occurring in a restaurant. This is also independent of the layout and setup of the restaurant although both the layout and the setup can impact the number of risky contacts. From this we also conclude that mitigation policies focussing on preventing virus transmission through these interactions are most effective. And especially preventing staff from transmitting the virus as any infectious staff member is a potential super spreader.

From the results we also conclude that modelling the virus transmission risks in a restaurant using a pedestrian model has added value. The model enables a quick but detailed insight into where the transmission risks occur in a specific restaurant and thus which mitigation strategies might be effective for this particular restaurant. Furthermore, it enables testing different layouts and setups to see if any changes to the layout or setup can reduce the transmission risk. On a larger scale, by modelling different restaurant layouts and setups (as was done in this study), we can obtain insight into what drives the virus transmission risk in restaurants in general. This can inform both policy makers and restaurant owners about what measures are effective in reducing the transmission risks in general.

The application of solely a pedestrian model to investigate virus transmission risks in restaurants also has limitations. First and foremost, it does not explicitly models the virus transmission but relies on the close contact definition. This means that it inexplicitly relies on the close contact definition being a good proxy for the virus transmission risk. That is, only direct contacts between people are considered as a potential sources of virus transmission and transmission via contact surfaces or aerosols in combination with bad ventilation are not taken into account. In future research we aim to look into the added value of integrating this pedestrian model with a virus transmission model. This limitation has its advantages as well though. It does not require and extensive calibration of a virus transmission model and thus requires less data. Furthermore, it can produce results more quickly as it does not require to run (many replications) of a virus transmission model for each pedestrian model run. Another limitation of the study is lack of validation. We currently lack data that relates the movement dynamics in a restaurant to the virus transmission risks. Again, this means that the application of a pedestrian model must rely on a proxy, such as the close contact definition, correlating well with the risk of transmission.

We also conclude that, in order for the model to make an impact, it should be used by the right people. In this case that would be policy makers and restaurant owners. To account for their (probable) lack in modelling skills and lack of time, the model should require few and simple inputs. To achieve this, we present a novel activity choice and scheduling model that requires few and simple inputs and that, coupled with an existing pedestrian model, can reproduce the movement dynamics of both customers and waiting for a wide range of dine-in restaurants. One other strength of this new model is that the data required to estimate the few parameters of the model can be collected using low-tech and privacy friendly methods.

This study also has some limitations with regard to the new model. First and foremost, we currently lack detailed movement data in relation to the activities of both customer and staff in a restaurant. Therefore, we cannot yet quantitatively validate the model. However, the simulation experiments show that different restaurant setups and the resulting different contact patterns can be reproduced in line with expectations. Secondly, the model currently assumes that every table is fully occupied. This means that the results present the worst-case scenario especially in the case of the shorter visits setup where all table are occupied most of the time. However, an advantage of this assumption is that the results will present the maximum risk for each restaurant which makes it easily interpretable.

All in all, this paper shows that pedestrian models can be a valuable tool for policy makers and restaurant owners to obtain detailed insight into virus transmission risks in a restaurant setting and to inform mitigation strategies. The paper also adds to the growing body of literature showing how pedestrian models can aid in developing insights into the virus spread dynamics. And with this, how pedestrian models can be a valuable tool in reducing the impact of a pandemic caused by an airborne virus. Lastly, the paper also identifies a lack of models and data pertaining to activity schedules of pedestrian in common indoor spaces, such as restaurants. Therefore, in order to extent the number of indoor spaces where microscopic pedestrian models can be applied for virus transmission risks assessment, more research into the topic of pedestrian activity scheduling behaviour is necessary.

CRediT authorship contribution statement

Martijn Sparnaaij: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft. **Yufei Yuan:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Winnie Daamen:** Conceptualization, Writing – review & editing. **Dorine C. Duives:** Conceptualization, Methodology, Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Restaurant activity scheduling data (Original data) (4TU.RESEARCHDATA)

.

Acknowledgements

This publication is part of the project SamenSlimOpen (with project number 10430022010018) of the research programme COVID-19 Programma, which is financed by the Dutch Research Council (NWO) and ZonMw. We thank Quirine Ten Bosch, Büsra Atamer Balkan, Colin Teberg, Berend Wouda, Doris Boschma, Linda van Veen, Marjon Koopmans and Wim van der Poel for their active involvement & fruitful discussions in the SamenSlimOpen project.

References

[1] WHO: World Health Organization, WHO coronavirus (COVID-19) dashboard, 2022, https://covid19.who.int/. Accessed: 2022-03-08.
[2] A. Brodeur, D. Gray, A. Islam, S. Bhuiyan, A literature review of the economics of COVID-19, J. Econ. Surv. 35 (4) (2021) 1007–1044, http://dx.doi.org/10.1111/joes.12423, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12423.
[3] B. Fouda, H.P. Tram, O.M. Makram, A.S. Abdalla, T. Singh, I.-C. Hung, A. Raut, L. Hemmeda, M. Alahmar, A.S. ElHawary, D.M. Awad, N.T. Huy, Identifying SARS-CoV2 transmission cluster category: An analysis of country government database, J. Infect. Public Health 14 (4) (2021) 461–467, http://dx.doi.org/10.1016/j.jiph.2021.01.006.
[4] C.S. Currie, J.W. Fowler, K. Kotiadis, T. Monks, B.S. Onggo, D.A. Robertson, A.A. Tako, How simulation modelling can help reduce the impact of COVID-19, J. Simul. 14 (2) (2020) 83–97, http://dx.doi.org/10.1080/17477778.2020.1751570.
[5] R. Padmanabhan, H.S. Abed, N. Meskin, T. Khattab, M. Shraim, M.A. Al-Hitmi, A review of mathematical model-based scenario analysis and interventions for COVID-19, Comput. Methods Programs Biomed. 209 (2021) 106301, http://dx.doi.org/10.1016/j.cmpb.2021.106301.
[6] E.S. McBryde, M.T. Meehan, O.A. Adegboye, A.I. Adekunle, J.M. Caldwell, A. Pak, D.P. Rojas, B.M. Williams, J.M. Trauer, Role of modelling in COVID-19 policy development, Paediatric Respir. Rev. 35 (2020) 57–60, http://dx.doi.org/10.1016/j.prrv.2020.06.013.
[7] R. Capobianco, V. Kompella, J. Ault, G. Sharon, S. Jong, S. Fox, L. Meyers, P.R. Wurman, P. Stone, Agent-based Markov modeling for improved COVID-19 mitigation policies, J. Artif. Intell. Res. 71 (2021) 953–992, http://dx.doi.org/10.1613/jair.1.12632.
[8] Q. Xu, M. Chraibi, On the effectiveness of the measures in supermarkets for reducing contact among customers during COVID-19 period, Sustainability 12 (22) (2020) http://dx.doi.org/10.3390/su12229385.
[9] T. Harweg, D. Bachmann, F. Weichert, Agent-based simulation of pedestrian dynamics for exposure time estimation in epidemic risk assessment, J. Public Health (2021) http://dx.doi.org/10.1007/s10389-021-01489-y.
[10] A.A. Tsukanov, A.M. Senjkevich, M.V. Fedorov, N.V. Brilliantov, How risky is it to visit a supermarket during the pandemic? PLoS One 16 (7) (2021) 1–18, http://dx.doi.org/10.1371/journal.pone.0253835.
[11] D. Alvarez Castro, A. Ford, 3D agent-based model of pedestrian movements for simulating COVID-19 transmission in university students, ISPRS Int. J. Geo-Inf. 10 (8) (2021) http://dx.doi.org/10.3390/ijgi10080509.
[12] C.-Y. Li, J. Yin, A pedestrian-based model for simulating COVID-19 transmission on college campus, Transp. A: Transp. Sci. (2021) 1–25, http://dx.doi.org/10.1080/23249935.2021.2005182.
[13] Y. Xiao, M. Yang, Z. Zhu, H. Yang, L. Zhang, S. Ghader, Modeling indoor-level non-pharmaceutical interventions during the COVID-19 pandemic: A pedestrian dynamics-based microscopic simulation approach, Transp. Policy 109 (2021) 12–23, http://dx.doi.org/10.1016/j.tranpol.2021.05.004.
[14] S. Sajjadi, A. Hashemi, F. Ghanbarnejad, Social distancing in pedestrian dynamics and its effect on disease spreading, Phys. Rev. E 104 (2021) 014313, http://dx.doi.org/10.1103/PhysRevE.104.014313.
[15] J. Lu, J. Gu, K. Li, C. Xu, W. Su, Z. Lai, D. Zhou, C. Yu, B. Xu, Z. Yang, COVID-19 outbreak associated with air conditioning in restaurant, Guangzhou, China, 2020, Emerg. Infect. Dis. 26 (7) (2020) 1628.
[16] K. Dube, G. Nhamo, D. Chikodzi, COVID-19 cripples global restaurant and hospitality industry, Curr. Issues Tour. 24 (11) (2021) 1487–1490.
[17] B.A. Balkan, Y. Chang, M. Sparnaaij, B. Wouda, D. Boschma, Y. Liu, Y. Yuan, W. Daamen, M.C. de Jong, C. Teberg, K. Schachtschneider, R.S. Sikkema, L. van Veen, D. Duives, Q.A. ten Bosch, The multi-dimensional challenges of controlling respiratory virus transmission in indoor spaces: Insights from the linkage of a microscopic pedestrian simulation and SARS-CoV-2 transmission model, 2023, http://dx.doi.org/10.1101/2021.04.12.21255349, medRxiv. arXiv:https://www.medrxiv.org/content/early/2023/07/21/2021.04.12.21255349.full.pdf.
[18] S.P. Hoogendoorn, P.H.L. Bovy, Pedestrian route-choice and activity scheduling theory and models, Transp. Res. B 38 (2) (2004) 169–190, http://dx.doi.org/10.1016/S0191-2615(03)00007-9.
[19] D.C. Duives, W. Daamen, S.P. Hoogendoorn, State-of-the-art crowd motion simulation models, Transp. Res. C 37 (2013) 193–209, http://dx.doi.org/10.1016/j.trc.2013.02.005.
[20] F. Martinez-Gil, M. Lozano, I. Garcîa-Fernández, F. Fernández, Modeling, evaluation, and scale on artificial pedestrians: A literature review, ACM Comput. Surv. 50 (5) (2017) 1–35, http://dx.doi.org/10.1145/3117808.
[21] W. Daamen, Modelling Passenger Flows in Public Transport Facilities (Ph.D. thesis), Delft University of Technology, 2004, URL http://resolver.tudelft.nl/uuid:e65fb66c-1e55-4e63-8c49-5199d40f60e1.
[22] H. Timmermans, X. van der Hagen, A. Borgers, Transportation systems, retail environments and pedestrian trip chaining behaviour: Modelling issues and applications, Transp. Res. B 26 (1) (1992) 45–59, http://dx.doi.org/10.1016/0191-2615(92)90019-S.
[23] W. Zhu, H. Timmermans, Modeling pedestrian shopping behavior using principles of bounded rationality: model comparison and validation, J. Geogr. Syst. 13 (2) (2011) 101–126.
[24] X. Liu, J.M. Usher, L. Strawderman, An analysis of activity scheduling behavior of airport travelers, Comput. Ind. Eng. 74 (2014) 208–218, http://dx.doi.org/10.1016/j.cie.2014.05.016.
[25] J.M. Usher, E. Kolstad, X. Liu, Simulation of pedestrian behavior in intermodal facilities, Int. J. Agent Technol. Syst. (IJATS) 2 (3) (2010) 66–82.
[26] S. Shelat, W. Daamen, B. Kaag, D. Duives, S. Hoogendoorn, A Markov-chain activity-based model for pedestrians in office buildings, Collect. Dyn. 5 (2020) 423–430.
[27] W. Huan, S. Yu-Qiang, Simulation and optimization of service system for restaurant, in: 2011 IEEE 18th International Conference on Industrial Engineering and Engineering Management, Vol. Part 2, 2011, pp. 1390–1392, http://dx.doi.org/10.1109/ICIEEM.2011.6035415.
[28] T. Tanizaki, T. Shimmura, Modeling and analysis method of restaurant service process, Procedia CIRP 62 (2017) 84–89, http://dx.doi.org/10.1016/j.procir.2016.06.076, 10th CIRP Conference on Intelligent Computation in Manufacturing Engineering - CIRP ICME '16. [Edited by: Roberto Teti, Manager Editor: Doriana M. D'Addona].
[29] P.E. Rossi, Bayesian non- and semi-parametric methods and applications, J. Econ. 115 (2014) 195–197.
[30] OpenTURNS, Openturns, 2022, https://openturns.github.io/openturns/master/contents.html. Accessed: 2022-03-01.
[31] M.C. Campanella, Microscopic Modelling of Walking Behaviour (Ph.D. thesis), Delft University of Technology, 2016, http://dx.doi.org/10.4233/uuid:b65e6e12-a85e-4846-9122-0bb9be47a762.
[32] M.C. Campanella, S.P. Hoogendoorn, W. Daamen, Improving the Nomad microscopic walker model, IFAC Proc. Vol. 42 (15) (2009) 12–18, http://dx.doi.org/10.3182/20090902-3-US-2007.0076.
[33] M. Sparnaaij, M.C. Campanella, W. Daamen, S.P. Hoogendoorn, Pynomad: A microscopic pedestrian simulation model, 2023, http://dx.doi.org/10.4121/543A7EEB-652A-40FE-8A08-2DB5FF9E222B.V1, URL https://data.4tu.nl/datasets/543a7eeb-652a-40fe-8a08-2db5ff9e222b/1. (Version 1) [Computer software].
[34] UK-HSA: UK Health Security Agency, What is meant by a contact, 2022, https://www.gov.uk/government/publications/guidance-for-contacts-of-people-with-possible-or-confirmed-coronavirus-covid-19-infection-who-do-not-live-with-the-person/guidance-for-contacts-of-people-with-possible-or-confirmed-coronavirus-covid-19-infection-who-do-not-live-with-the-person#what-is-meant-by-a-contact. Accessed: 2022-03-09.
[35] CDC: Centers for Disease Control and Prevention, How to determine a close contact for COVID-19, 2022, (https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/determine-close-contacts.html). Accessed: 2022-03-09.
[36] RIVM: Rijksinstituut voor Volksgezondheid en Milieu, Protocol bron- en contactonderzoek COVID-19, 2022, https://lci.rivm.nl/COVID-19-bco. Accessed: 2022-03-09.