MSc thesis in Transport and Planning

Quantifying transmission risks of SARS-CoV-2 in pedestrian interactions at large events

Xinyi Wang

2021

TU Delft

QUANTIFYING TRANSMISSION RISKS OF SARS-COV-2 IN PEDESTRIAN INTERACTIONS AT LARGE EVENTS

A thesis submitted to the Delft University of Technology in partial fulfillment of the requirements for the degree of

Master of Science at the Delft University of Technology

to be defended publicly on Tuesday September 28, 2021 at 14:30

by

Xinyi Wang

Student number:	5130794
Project duration:	December 2nd, 2020 - September 28th, 2021
Thesis committee:	Dr. ir. D. C. Duives, TU Delft
	Dr. ir. E. Papadimitriou, TU Delft
	Dr. ir. QA. ten Bosch, WUR

Photograph cover: Property of Amsterdam Open Air Festival
An electronic version of this thesis is available at: https://repository.tudelft.nl.
This project was partly funded by SamenSlimOpen, a research project funded by ZonMw (Project No.10430022010018).

TU Delft ZonMw

SUMMARY

The COVID-19 global pandemic has influenced almost everyone's life on this planet, since the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) found its way to the human race. Due to its highly transmissible nature and high death rate, countries around the world have been taking preventive measures to protect their citizens from being infected, including social distancing, face mask obligations, and even lockdown.

Besides the health influence, the COVID-19 pandemic has also resulted in severe economic disruptions, due to supply shortages and lockdown of businesses. Most of the large events in the Netherlands have been cancelled since the first infection case was identified in February. The significant decrease of festivals has posed a serious negative impact on both the cultural and economical aspects of the event industry. However, it is not yet known what is the most appropriate restrictions to apply to event industry, as the exact risk of a visit to an event, such as a music festival, has not been thoroughly studied.

This research aims to close this research gap by developing a SARS-CoV-2 transmission risk analysis method for large events by modelling crowd interactions at different types of event spaces and quantifying the SARS-CoV-2 transmission risks in the process. The main research question to be answered by this study is formulated as follows:

How to model SARS-CoV-2 transmission risks based on pedestrian behavior and virus spread simulation at large events?

Through literature review, existing pedestrian modeling, virus spread modeling, and risk identification approaches are identified. The most suitable approaches for this research are selected as NOMAD, a force-based tactical to operational level pedestrian model, QVEmod, an agent-based virus spread model, and the dose-response model, which relates the amount of accumulated virus to the infection risk of individuals. A research gap is identified of simulating transmission risks at large events where people visit a set of different activity spaces during one day.

This research has proposed a method to connect activity scheduling, pedestrian route choice and movement modeling, virus spread modeling, and infection risk identification to determine the SARS-CoV-2 transmission risks at large events by probability method.

The first part of the proposed method is the NOMAD pedestrian model, which simulates pedestrian route choice and movement with the input of infrastructure layout, social force parameters, and demand pattern. With NOMAD, a number of activity spaces with potential SARS-CoV-2 transmission possibility can be simulated under different infrastructure layouts and physical interventions, such as 1.5m or 1m queue distance. NOMAD generates pedestrian trajectories with the accuracy of 0.1m at each 0.1-second time step.

Then the output of NOMAD is transformed into agent scripts which include the movement of pedestrians at the accuracy of 1m at each 20-second time step, the respiratory characteristics of the agents, whether wearing a face mask, the initial viral load, and other virus transmission related parameters. The second part of the proposed method is QVEmod, the model that simulates the virus spread in different activity spaces via 7 processes that the virus goes through, including emission, falling onto surfaces, decay, diffusion, inhalation, contaminating surfaces by touching, and being picked up from surfaces by touching. QVEmod generates the accumulated virus loads on agents via three routes, namely, droplets, aerosols, and fomites.

The third part of the method include e smaller steps of activity scheduling, possible exposure time and location identification, adding up virus obtained from different activity spaces, and infection risk estimation. The first step generates every visitor's activity schedule at the event, identifying where and when the infectious and susceptible individuals are located. The second step summaries the whereabouts of infectious individuals and derives the locations and duration of possible virus exposure. In the third step, the susceptible individuals' activity schedules are compared with the derived locations and duration. If a match exists, a virus exposure (obtained from the virus spread simulation) is assigned to this individual. For all the susceptible individuals, the virus exposure acquired throughout the activity schedule is added up. Finally, the accumulated virus exposure is fit into a dose-response model to calculate the general infection risk of visiting this event.

A few factors need to be identified before applying this method. First, after determining the type of event to evaluate and the general demand, the infrastructure types and activity spaces should be identified as the input information to the NOMAD model. Then the activity pattern data of such an event needs to be collected and analyzed, from which the activity schedules and demand pattern at activity spaces can be derived. The former will be used for NOMAD simulation and the latter for identification of virus locations and matching up the virus exposure on susceptible individuals.

In the case study application, the proposed method has shown its capability of revealing the general infection risk and the relation of influence factors to the transmission scale, and it also identifies risk-prone areas in an event.

Comparing the different scenarios at each activity space, the general trend is identified that the transmission of SARS-CoV-2 is limited when the facility is located outdoor, the queue distance is increased (from 1m to 1.5m), the density is lowered (from $5.76p/m^2$ to $0.4p/m^2$), and the respiratory activities are calmer (from 20%talking + 40%singing + 40%breathing to 20%talking + 80%breathing).

Among these four variables, the impact of increasing the queue distance from 1m to 1.5m is the smallest. It is observed that the virus transmission scale varies much more significantly when increasing the density from $1p/m^2$ (1m distance) to $2p/m^2$ (0.7m distance) and $5.76p/m^2$ (0.4m distance) compared to when increasing the density from $0.4p/m^2$ (1.5m distance) to $1p/m^2$ (1m distance), which indicates the capability of SARS-CoV-2 to transmit via airborne routes decays to a turning point at around 1m distance. Before reaching this point, interventions that increase the distance between people will have a strong impact on infection prevention, while after reaching this point, the impact is much smaller when the distance becomes longer.

Moving indoor events to an outdoor environment also has a positive influence on limiting the spread of SARS-CoV-2, halving the total amount of virus transmitted in the facility, as the viral-laden particles in aerosols and fomites decay much faster in an outdoor and UV-exposed environment.

Performing calm respiratory activities can also significantly reduce the virus transmission scale. When replacing 40% time spent on singing with breathing, the amount of virus picked up by individuals goes down by more than 85% in a 20-minute-duration 1000-participant music concert. The impact is even stronger in outdoor music stages, which indicates singing affects droplet transmission more than aerosol transmission.

It is also discovered that under the same distance between people, when staying for the same duration of time in the same facility, the queues, where people move to take up the place of the person in front of them, may be more risky than the music stages, where people stand close to each other but do not move.

The above findings correspond to the advice from outbreak management team (OMT), that keeping social distance, strictly limiting the scale of indoor activities/events, limiting the scale of outdoor activities/events (bigger scale than indoor), limiting the duration of festivals, applying fixed seating with low densities, requiring neg-

ative test results, and discouraging festival attendees from singing help limit the transmission scale of SARS-CoV-2.

A sensitivity analysis is conducted with different scenarios of activity space combinations, which gives infection numbers of a 10-infectee 10000-participant music festival ranging from 17.52 in the optimal scenario to 86.73 in the least expected scenario, which equals to 0.18% to 0.87% of the total susceptible population. The result of the least expected scenario falls into the infection probabilities of 2 smaller-scale real-life experimental events (less than 2000 participants), while another large-scale real-life event (10000 participants per day for two days) appeared to have a much larger infection scale. Underestimating as the research findings may be, such direct comparisons have very limited implications, due to the great differences between the simulated events and real-life events. Factors such as the event scale, infrastructure setting, crowd management measures, heterogeneity in the emission rate, activity schedules, group behavior, and number of infectious individuals, would directly influence the amount of virus transmitted in the event. Nevertheless, this research has shown when and where major risks can occur during an event. By comparing results of different scenarios, it also gives indications on crowd management measures and interventions that can help reduce the virus transmission scale.

The focus of this study lies in developing a risk evaluation tool for big events, comparing the potential infection risk under different scenarios (infrastructure design, physical intervention) and spotting where the most risk lies. It allows future exploration and comparison of the transmission scales of certain kinds of events, without posing ethical controversy of exposing people in infection risks. The method can further assist decision making on crowd management approaches and interventions to be used to reduce the infection risk at the simulated event. Yet, it has its limitations of not considering group behavior at the event, which is commonly observed at large events and may potentially increase the transmission scale. Another major limitation is its heavy dependency on detailed virus transmission parameters and activity patterns. The former needs to be validated under different scenarios, and the latter needs to be identified from the data collected at the same type of event.

ACKNOWLEDGEMENTS

This thesis marks the final step of my Master's program in Transport and Planning at TU Delft. It has been conducted in cooperation of the research project SamenSlimOpen, the objective of which is to assess and limit the risks of the spread of SARS-Cov-2 in indoor spaces. I genuinely hope that this report will provide some insights for practice and for future research.

The past 10 months has been an unforgettable and fruitful journey as I built my thesis mansion brick by brick. I have made great progress on the aspect of academical knowledge and skills, as well as mental strength. Being conducted during the COVID-19 pandemic, this thesis project has faced many challenges in the constantly changing environment. In this special time, I would like to express my gratitude and appreciation to everyone who has inspired and helped me throughout the journey.

First, I would like to thank my daily supervisor and chair of the committee, Dorine Duives, who has led me into this fascinating world of active modes and guided me on every aspect of this thesis project. Your professional insights and thorough feedback have always steered me to a clear path every time I encountered difficulties. Your encouragement and support have been fundamental to both my academic and mental strength, without which I would not be able to develop this project to this level. I will miss our fruitful biweekly thesis meetings and feedback documents full of insightful comments.

Next, I would like to thank Eleonora Papadimitriou and Quirine ten Bosch, my thesis committee members who gave me valuable advice on the aspect of risk evaluation and epidemiology. Your feedback brought a new light on the project and has inspired me to critically think about the structure and content of my work.

I would like to thank Martijn Sparnaaij and Yufei Yuan, who made the NOMAD model a reality, tutored me in making use of the simulation tool, and helped me with emerging problems in my scenarios. I also would like to thank You Chang, Busra Atamer Balkan, for the help in setting up the QVEmod simulations, exploring its capabilities and limitations, and looking into the problems I encountered along the way.

As this thesis involves professional knowledge and simulation tools from both pedestrian modelling and virus transmission modelling, great support has been needed and satisfied from the SamenSlimOpen project team. Therefore, I would like to express my gratitude to everyone from the team who helped me with my thesis and, more importantly, developed the simulation tool to benefit the society during this COVID-19 pandemic. Among them, I would like to give special thanks to Bas Dado, Colin Teberg, and Berend Wouda, who examined my python scripts and solved my problems with making adaptations. I also would like to thank the organization of the Amsterdam Open Air festival and Reinout de Roos for gathering and sharing the GPS tracing data that made my simulations possible with the baseline event information.

Finally, I would like to thank my parents, who have provided financial and emotional support along my journey, loved and encouraged me through the ups and downs. I would like to express my gratitude to my father for tutoring me in python and debugging my codes. I would like to thank my best friends from my Master's program, Mathijs, Lanxin, and Rui, who went through this challenging journey together with me. I also would like to thank all my friends who have always opened their arms for me and inspired me to become a better person.

Xinyi Wang
Delft, September 2021

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

OD origin and destination 8
DDCM dynamic discrete choice model 9
SEIR Susceptible - Exposed - Infectious - Recovered 19
MLR multiple linear regression 8
ARIMA Auto Regressive Integrated Moving Average 8
SVR support vector regression 8
R_0 basic reproduction number 23
R_e effective reproduction number 23
R_c controlled reproduction number 23
EGR exponential growth rate-based 23
SIR Susceptible - Infectious - Recovered 23
WKT Well-known text 48
NPI's non-pharmaceutical interventions 1
ACH air change rate 37
MOP measure of performance 41
OMT outbreak management team vi
RIVM Dutch National Institute for Public Health and the Environment 100

1 INTRODUCTION

A global outbreak of the COVID-19 pandemic has posed threats to the physical and mental health of people all over the world. Till August 24th, 2021, we have witnessed more than 213.7 million cases of COVID-19 and 4.4 million deaths [University, 2021]. Almost everyone across the globe has been influenced by the pandemic. For those who have been directly stricken by COVID-19, some lost their lives, others survived but may still suffer from its clinical sequelae Xiong et al. [2021]. For those who have been indirectly influenced by the pandemic, some grieve for the loss of family members, others find it difficult to maintain a healthy mental status during the lockdown and isolation [Rossi et al., 2020; Kaparounaki et al., 2020]. Besides its health influence, the COVID-19 pandemic also has a huge global economical impact on all kinds of industries. For instance, the production of goods has shrunken due to the government measures to protect public health and the consumption profile of people has changed due to the change of lifestyles shaped by the pandemic McKibbin and Fernando [2020]; Maital [2020]. The economical impact, in return, also poses an influence on people's lives and mental health status.

Caused by the highly transmissible SARS-CoV-2 virus, COVID-19 spreads between people when a healthy person gets in close contact with an infected person. The virus can be transmitted by respiratory droplets, indoor aerosols, contaminated surfaces, etc. [Harrison et al., 2020]. Due to the nature of its transmission mechanics, countries across the globe have introduced prevention measures aiming to limit the physical contact between people, such as lockdown on a national/provincial scale, curfew from dusk till dawn, restrictions for cross-border travel, closing down unessential public areas, etc. RIVM [2020]; AMT [2020]. To control the spread of COVID-19 in the Netherlands, the Dutch government has been constantly adapting the COVID-19 measures and advice to the public since April 2020. Before the vaccine coverage reaches the level that allows lifting all COVID-19 restrictions and recommendations, the control of the pandemic has mainly relied on the non-pharmaceutical interventions (NPI's), such as indoor mandatory face masks, 1.5-meter social-distancing, shutting down unessential public spaces, including restaurants, hairdressers, shopping malls, sports stadiums, museums, and schools RIVM [2020]. While these measures have contributed to controlling the spread of pandemic, they also pose a serious negative influence on the relevant industries. For instance, event organizers have been unable to operate their businesses as normal. In pre-pandemic time, 2019, 1117 festivals were held in the Netherlands, attracting more than 27 million visitors, while in 2020, due to the COVID-19 measures, the number of festivals dropped sharply to 190, with merely 1.4 million attendance [Statista, 2021]. Popular festivals, such as the Vierdaagsefeesten Nijmegen (International Four Day Marches) and the Amsterdam Dance Festival (ADE), has all been canceled due to the COVID-19 regulations. The significant decrease of festivals has posed a serious negative impact on both cultural and economical aspects of the event industry.

However, it is not yet known to what extent does an event contribute to the spread of COVID-19, as the exact risk of a visit to an event, such as a music festival, during this pandemic, has not been thoroughly studied. Numerous studies on the epidemiological mechanism of COVID-19 and SARS-CoV-2 virus have been conducted since the outbreak in December 2019 [Harrison et al., 2020; Gao et al., 2021; Arav et al., 2020; Bouchnita and Jebrane, 2020]. A number of experiments exploring the actual infection scale at events under different NPI's have been organized and analyzed in

the Netherlands in 2021 [FieldlabEvenementen, 2021; Ellyatt, 2021]. However, none of the existing studies has been able to reveal or estimate the transmission risks of SARS-CoV-2 at different types of large events. Simulation-based researches mainly focus on the transmission profile on the city or provincial scale [Tian et al., 2020] and the roughly estimated effectiveness of NPI's on transmission risks under simple circumstances, with assumptions of random movements of individuals [Bouchnita and Jebrane, 2020]. However, to reveal the transmission profiles of specific events more accurately, such as watching a football match or visiting a music festival, specific knowledge and assumptions about the event facilities, infection profile in the population, and people's activity schedule and actual movement in the space are required. Experimental events reveal the transmission scales of certain types of events, which are limited to the infection rate in the population at the time, the specific event facilities, the NPI's, etc. Moreover, this type of experiments obtain infection data by organizing events that potentially expose participants to infection risks, which may lead to ethical controversy. Media reported many people, including scientists, have criticized the scientific and ethical aspects of the experimental events [de Vrieze, 2021; RTLNieuws, 2021]. According to Ellyatt [2021], in July 2021, more than 1000 people caught COVID-19 after a two-day outdoor festival held in Utrecht, the Netherlands, which attracted more than 20000 people. Therefore, a research gap exists for an approach to simulate the transmission risks of SARS-CoV-2 via pedestrian modelling.

With the increased coverage of vaccination and constantly emerging variants of SARS-CoV-2, the society faces great challenges reopening businesses. Without detailed knowledge of the transmission risks at event spaces, policy makers face challenges when making crucial decisions on COVID-19 regulations, leading to (sometimes inconsistent) measures which may unnecessarily restrict business development, cause public doubts and non-compliance to measures, even social unrest. Therefore, a reliable transmission risk analysis tool is in urgent need. In this research, in order to better support the decision-making process of the government on regulations in the event industry, an infection risk estimation method which combines the pedestrian activity scheduling, pedestrian interactions modelling, and virus transmission modelling, will be proposed. This method will be applied to the risk estimation of a music festival under different infrastructure scenarios, the event scale, facility types, and activity schedules of which are based on the Amsterdam Open Air festival 2019. From the results, facilities and behavior that pose great risk will be identified, as well as the general risk of attending the festival. The proposed method can also be further adapted and applied to detect the infection risk at different types of large events, thus assist the policy makers in making crucial decisions on the event industry.

1.1 RESEARCH QUESTIONS

Following the above considerations, the research objective of this study is defined as follows:

To develop a SARS-CoV-2 transmission risk quantification method at large events by modelling activity scheduling, pedestrian behavior and SARS-CoV-2 transmissions at different types of event spaces, and quantifying the SARS-CoV-2 transmission risks throughout the entire visits to the event by individuals.

Based on the objective, the main research question is formulated as follows:

How to model SARS-CoV-2 transmission risks based on pedestrian behavior and virus spread simulation at large events?

To answer the main research question, the following sub-questions are put forward:

1. ***What are the behavior frameworks of visitors at large events?***

To answer this sub-question, this research will identify the type of event to be studied, the functional spaces in the event, and the behavior patterns of visitors at this type of event. The answer to this question will be revealed by the case study analyzing the real-life behavioral and infrastructural data collected from an event in Chapter 5.

2. ***What are the SARS-CoV-2 transmission routes among visitors at these events?***

   To answer this sub-question, this research will identify the transmission mechanics of SARS-CoV-2 and match the possible transmission routes that may occur during pedestrian interactions at the functional spaces of the events. The answer to this question will be revealed by the literature study on the state-of-art SARS-CoV-2 transmission studies in Chapter 2 and the theoretical analysis on transmission routes at the identified event spaces in Chapter 5.

3. ***How to make use of real-life pedestrian interaction data, pedestrian simulation models, and SARS-CoV-2 transmission models to identify the risk of SARS-CoV-2 transmission at large events?***

   To answer this sub-question, this study will research on the existing SARS-CoV-2 transmission modelling studies, analyze their data type and modelling approaches, evaluate the real-life data and simulation models to be used in this study, and identify their purposes based on the literature research and evaluation. Therefore, this question will be answered by the literature study reviewing pedestrian modelling methods, SARS-CoV-2 transmission risk modelling studies in Chapter 2, proposing an approach for SARS-CoV-2 transmission risk modelling in Chapter 3 and Chapter 4 based on literature studies, a case study analyzing the real-life data collected from an event and applying the proposed approach for an event in Chapter 5.

4. ***What is the SARS-CoV-2 transmission risk profile at large events?***

   This question will be answered in Chapter 6, by analyzing the results of the case study application of the proposed method, from which the transmission risk will be quantified in number of infections and the risk-prone event spaces will be identified as well.

5. ***How to apply the developed methodology in practice and how does it perform?***

   After researching on the state-of-art literature, developing the methodology, applying the proposed method in a case study, and analyzing the results, the process of the method application is presented. The capabilities and limitations of the developed methodology will be discussed in Chapter 8 and Chapter 9.

1.2 RESEARCH SCOPE

The scope of this research is described in this section. The following aspects are addressed: the type of events considered in this research, the level of pedestrian modelling taken into account in this research, the focus of transmission risk assessment, and the scope of development of the proposed transmission risk quantifying method.

Event type

Although all types of events, where close contact between people is unavoidable, are in urgent need for infection risk assessment to supervise the Covid-19 related regulations in the industry, the scale and activity type vary largely among different events. The risk assessment method involving specific pedestrian modelling in

event spaces is only capable of simulating the events which have similar functional areas and activity patterns. In the Netherlands, music festivals account for a large part of festival attendance, attracting both domestic and overseas visitors. Across different music festivals, the functional areas and activity patterns remain similar, which makes it possible to develop a general risk assessment tool for this type of event. Therefore, in this research, the development and application of transmission risk detection method will be mainly based on large music festivals, which may also shed light on the risk assessment of other types of large events, such as football matches and dance festivals.

Level of pedestrian modelling

Pedestrian modelling can be categorized in three levels, namely, strategical, tactical, and operational level, representing pedestrian behavior decisions at different stages [Hoogendoorn and Bovy, 2004]. From strategical to operational level, the choices become more specific as they develop from departure time choice and activity pattern choice to actual movement choice. More detailed explanations are presented in Section 2.1. In this research, all levels of pedestrian modelling are considered and reviewed, as one level has an influence on the next one and in the end affect the actual interactions between people. However, the pedestrian behavior considered in the method proposed by this thesis does not necessarily cover all the levels, as existing pedestrian behavior simulation tools mainly focus on one or two levels. The level(s) included in the simulation method will be determined according to the findings from literature review in Chapter 2. In Chapter 5, the choices of activity schedule and activity area at tactical level are analyzed based on real-life data collected from a music festival.

Focus of virus transmission risk assessment

In this research, the focus of virus transmission risk assessment lies in the general risk profile of visiting a certain event, which is translated from the accumulated virus dose of individuals who follow a general activity schedule and visit a number of activity spaces during this event. The probability of getting infected at different types of activity spaces will not be studied as it is not in a linear relation with the accumulated virus dose of individuals [Nicas and Sun, 2006]. Nevertheless, the infection risks of different types of activity spaces can be represented and analyzed qualitatively according to the accumulated virus doses within the facility.

Scope of proposed method

This research aims to develop and validate a SARS-CoV-2 transmission risk identification method which combines detailed pedestrian modelling with the transmission dynamics of SARS-CoV-2. However, defining and validating the model parameters does not fall into the scope of this research. The parameters representing pedestrian behavior, the essential parameters and assumptions for virus transmission analysis are based on existing studies. The focus of this research lies in quantifying the infection risk of a complete visit to an event by an average participant in specific event scenarios, by establishing a bridge connecting pedestrian modelling with virus transmission risks at large events.

1.3 THESIS OUTLINE

The approach for this research is illustrated in Figure 1.1.

As is shown below, the research begins with a background introduction in Chapter 1, which brings out the research questions, the scope and the structure of this thesis. Then a thorough literature review follows in Chapter 2, presenting the state-of-

art pedestrian behavior models, SARS-CoV-2 transmission mechanics, SARS-CoV-2 transmission models, and risk assessment methods. In Chapter 3, the reviewed models and methods are placed in a theoretical framework for the proposed SARS-CoV-2 transmission risk identification method, where the research gap is identified and made up for. Next, the proposed methodology is presented in Chapter 4, consisting of the detailed description of models and methods used in this research. In Chapter 5, a case study, which derives the activity pattern of a music festival from GPS data and applies the developed methodology for SARS-CoV-2 transmission risk identification at event spaces, is conducted, providing feedback to Chapter 4. After revising the methodology based on the findings from the case studies, the general results of the proposed method is presented in Chapter 6. A sensitivity analysis is then conducted in Chapter 7, exploring how variables (queue distance, indoor/outdoor space, etc.) influence the infection profile by comparing the results of different scenarios. In Chapter 8, other factors that might have an important influence on the results are discussed, such as assumptions and research scope. In the end, Chapter 9 draws the conclusions of this research, which conclude the result analysis, summarize the answer to the research question, discuss the inadequacy of this research, and provide recommendations for further research.

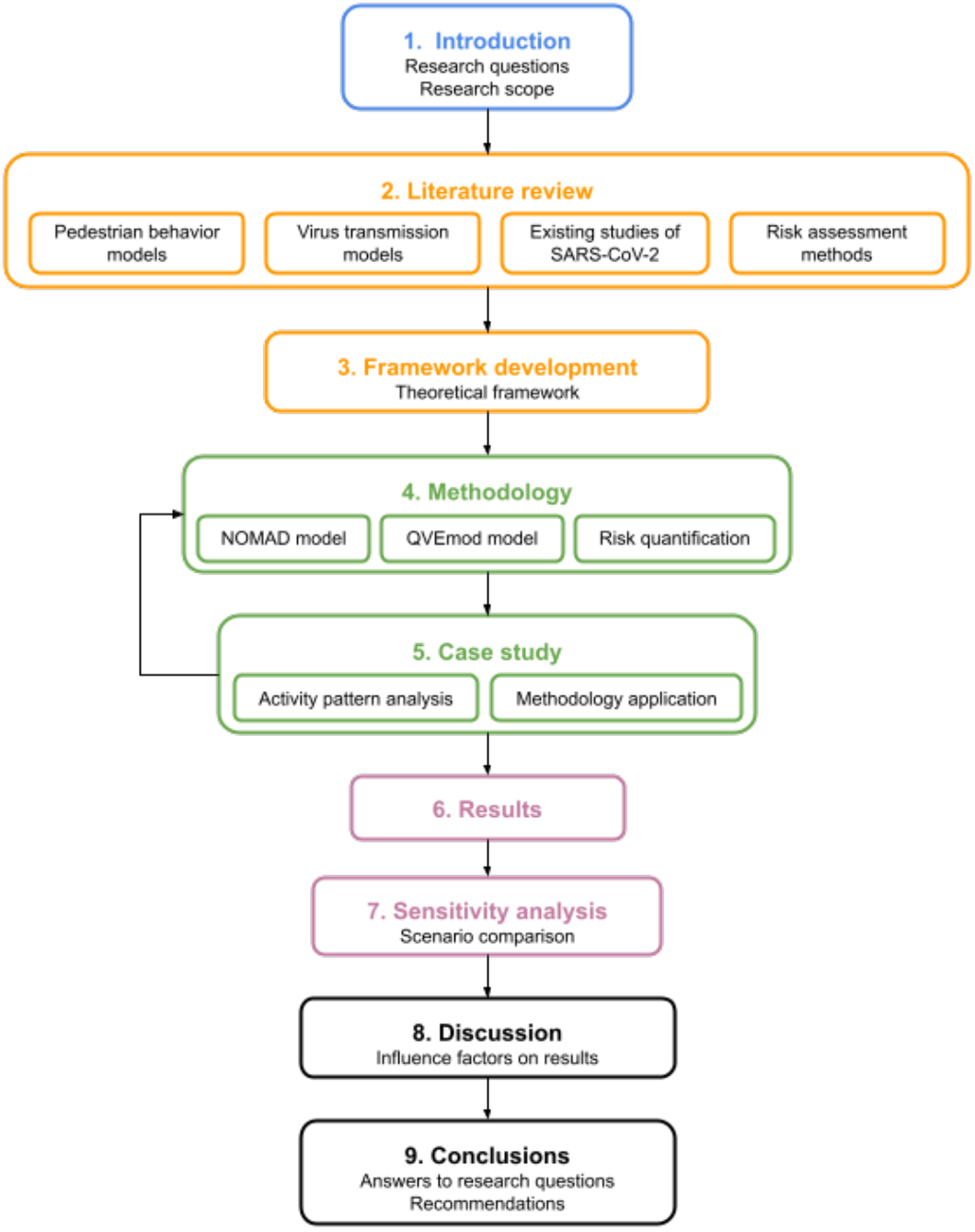


Figure 1.1: Thesis outline

2 LITERATURE REVIEW

A review of literature regarding studies of pedestrian behavior models, virus transmission models, and risk assessment methods is presented in this chapter.

The goal of the literature review presented here is threefold. First, to identify which methods of pedestrian modelling and virus transmission modelling have been proposed so far and what level of detail they are capable to achieve. Second, to identify suitable methods for development of an infection risk model of SARS-CoV-2 involving pedestrian interaction. Third, to identify a risk assessment method that quantifies and visualize the model results. An overview of the state-of-the-art pedestrian modelling methods is discussed in Section 2.1. The respiratory virus transmission models, together with existing studies on SARS-CoV-2 transmission are reviewed in Section 2.2. The risk assessment methods are presented in Section 2.3.

2.1 PEDESTRIAN BEHAVIOR MODELS

Pedestrian behavior has been studied by researchers since 1950s [Hankin and Wright, 1958]. Numerous models have been put forward to account for pedestrian's travel patterns from different angles. Some focus on the decisions of trips and activity schedules. Some study the influence factors of pedestrian's actual movements. To present different types of models in an organized manner, this paper categorizes them into three choice levels, namely strategical, tactical, and operational level, each representing a different stage of pedestrian travel behavior. According to [Hoogendoorn and Bovy, 2004], the strategical level corresponds to the activity, destination, and departure time choices; the tactical level models the activity schedule, activity area, mode and route choice; the operational level explains the movement choice at the presence of objects and other pedestrians. Although obtaining the choice at a higher level is essential for generating the next-level choice, all of the existing models are not capable of simultaneously simulating choices at more than one level. In practice, the input of the next-level model is usually assumed by the researchers or randomly generated. In this section, pedestrian behavior models are reviewed separately at each level. It is important to be aware that overlaps exist among models at different levels, as some models can be implemented at more than one level, with different inputs and adjusted settings.

The theories and features of the reviewed pedestrian models will be explained in the following subsections, Section 2.1.1, Section 2.1.2, Section 2.1.3. A conclusion will be reached in subsection Section 2.1.4.

2.1.1 Strategical level

At strategical level, pedestrian behavioral choices include departure time choice and activity pattern choice [Hoogendoorn and Bovy, 2004]. Although departure time and activity choice play an important role in pedestrian behavior, they receive less attention compared to route choice and operational movement, as the latter forms the ultimate presenting behavior. However, pedestrian choices at strategical level have a crucial influence on the choices at tactical level, as they determine when, where, who will be present. In other words, strategical choices influence

the pedestrian demand, population characteristics, and walking environment at an activity space. Therefore, pedestrian modeling at strategical level are of vital importance to derive location-based pedestrian behavior. For instance, they have often been developed for public transport stations and busy urban areas to assist crowd management [Lai and Kontokosta, 2018; Desyllas et al., 2003].

Three types of models can be applied at strategical level, namely data-driven models, equilibrium assignment models, and discrete choice models. It is important to be aware that these models do not work exclusively for pedestrians. Instead, at strategical level, mode choice is usually modelled together with departure time and activity pattern choice.

Data-driven models

Data-driven models are defined as a type of macroscopic model, as it focuses on deriving the trend from the historical number of people making the choice to visit a certain location under certain circumstances. They predict pedestrians' departure time and activity pattern choice by analyzing the relation between the historical demand and exogenous factors, such as attributes of activities and environmental factors. Depending on the mathematical relation applied in the model, they can be further categorized into linear models, such as multiple linear regression (MLR) models and Auto Regressive Integrated Moving Average (ARIMA) model, non-linear models, such as support vector regression (SVR) models. Data-driven models for pedestrian demand estimation have mainly been adapted from vehicular traffic estimation, where they have been extensively developed [Barros et al., 2015; Li et al., 2017]. As for application in pedestrian researches, Lai and Kontokosta [2018] developed multivariate MLR models to estimate pedestrian activities in urban regions, based on the historical data from New York City. The models consider the influence of contextual features and time-varying situational indicators on pedestrian activity across time of day, day of the week, season, and year. The relation between pedestrian activity and land use, building density, transportation infrastructure, and other factors commonly associated with urban walkability are quantified in the study. An ARIMA model is developed by Wang et al. [2017] to predict the city foot traffic at multiple locations in the City of Melbourne. By describing the autocorrelation in the data, the ARIMA model is able to capture the trend in the time series and make accurate long-term and short-term predictions without looking into external factors.

Data driven models have been used to explore the important drivers of local pedestrian activities and provide scientific foundations for the improvement of pedestrian experience in urban areas. By deriving trends from a large amount of data, they are capable of producing realistic predictions with very limited assumptions of people's decision making process. On one hand, they provide a method to analyse pedestrian demand patterns without related knowledge. On the other hand, they ignore the influence factors of choices at the individual level, which limits their results to be further applied in models at the next level.

Equilibrium assignment models

The equilibrium assignment model is developed based on the 4-step transport modelling framework, traditionally used for assignment of trips among origin and destination (OD) pairs [de Dios Ortúzar and Willumsen, 2011]. For the application in pedestrian modelling, it has been developed for departure time choice at the strategic level. In this type of model, pedestrians select departure time of a certain OD pair, according to a logit formula involving the predetermined departure time costs and the equilibrium OD walking costs. The number of pedestrians choosing a certain departure time is determined by the equilibrium used in the model, such as deterministic user equilibrium and stochastic user equilibrium [de Dios Ortúzar and Willumsen, 2011]. Huang and Lam [2002] developed a dynamic pedestrian equilibrium assignment model for departure time and path choice. The model assumes a

pedestrian network with a set of links where pedestrians are assigned to a departure time of a link based on their subjective disutility to the departure time. The disutility is derived from the historical traffic conditions, the activity schedule and utility in origins and destinations, and the subjective preference for the departure time. By considering the time-dependent OD demands as endogenous variables, this model also takes into account of inter-elasticity of OD demands between time intervals, which enables flexible departure times.

Although this type of model involves the subjective disutility from pedestrians' perspective, it does not consider the individual difference. Therefore, it falls into the category of macroscopic models. With its macroscopic assumption, the application of equilibrium assignment models in pedestrian-related research is very limited compared to the transport modelling field.

Discrete choice models

Discrete choice models fall into the category of microscopic models, as they simulate choices of individuals, taking into account of personal characteristics. This type of model determines the choice of an individual person or a group of people, according to their own attributes and a set of rules, among an finite number of choices. Discrete choice models can be further categorized by the decision rules used in the model, such as multinomial logit and nested logit [Hall, 2012]. Discrete choice models have been widely applied in all aspects of transport modelling [Aloulou, 2018]. In pedestrian related studies, numerous models have been developed and applied at strategic level. Dekker et al. [2014] incorporated needs of satisfaction in a discrete choice model for leisure activities. Besides conventional attributes, such as activity costs and accessibility, the individual's anticipation of activity satisfaction also counts for the expected utility. Each individual aims for the activity with the maximum utility. Västberg et al. [2020] developed a dynamic discrete choice model (DDCM) for daily activity-travel planning. In this model, a sequence of decisions of when, where, why, and how to travel compose a daily activity–travel pattern. The sum of the utility of all trips and activities determines the individuals' preferences for a certain activity–travel pattern. Individuals make choices at each decision stage to maximize the expected utility of the remainder of the day. This model allows for a detailed treatment of timing decisions consistent with other choice dimensions, respects time–space constraints, and enables the inclusion of explicitly modeled uncertainties in travel time. It is capable of accurately reproducing the activity patterns from choices of 1,240 locations, four modes, and six activities. A group of such DDCMs have been developed at strategical level, with subtle adjustments. For instance, Karlström et al. [2009] developed a DDCM for mode choice and departure time modelling, where the order of activities is considered dynamically. To explore the effective dimensions in the real-life problem, the Restricted Boltzmann Machine is used to realize a dimensionality reduction without losing accuracy.

In general, discrete choice models provide an individual-based modelling approach, with preliminary assumptions of internal and external choice drivers. They have been applied for both multimodal modelling and pedestrian modelling. They can also be adapted to explore the important factors in individuals' decision making process.

Conclusion

In conclusion, models at strategical level provide both general and individual-based approaches to simulate pedestrian departure time choice and activity pattern choice, together or separately. They each have their own advantages and disadvantages. Data-driven models make realistic reproductions of pedestrian activity without detailed assumptions. Equilibrium assignment models are capable of simulating departure time choice and path choice at the same time. However, they ignore individ-

ual differences and it is difficult to take their results for models at the next level. In discrete choice models for activity pattern modelling, personal attributes are taken into account, which increases the computational efforts to achieve a realistic estimation.

In this study, there are two types of activity area and activity schedule to be considered. One takes into account the entire visit to the music festival, the activity areas of which consist of all types of activity spaces. Its activity schedule refers to the schedule to visit a set of activity areas. The corresponding strategical choices include the activity choice of visiting this music festival and the departure time to visit this music festival. On this aspect, the activity choice is predefined and the departure time choice will be derived by a data-driven method in Chapter 5, which analyzes the scattered real-life location data collected from a musical festival. The other type considers one visit to one activity space in this event, for instance, one visit to a bar. In this case, the activity area is the bar and the schedule is the schedule to enter the bar, visit the counter, sit at the table and leave. The corresponding strategical choices include the activity choice of visiting this bar and the departure time to visit this bar. On this aspect, the both activity choice and the departure time choice are determined by the the activity schedule of the entire visit to this music festival. Therefore, except the data-driven method to derive the activity choice and departure time choice to visit activity spaces, no further derivation is required in this study.

2.1.2 Tactical level

The tactical level of pedestrian behavior describes the choices of activity schedule, activity area, and route to reach activity areas [Hoogendoorn and Bovy, 2004]. The decision making at this level is determined by how pedestrians perceive the environment. The influence factors include external factors, internal factors, and expected traffic conditions. The external factors include the presence of obstacles and stimulation of the environment, while the internal factors include time pressure and attitudes of the pedestrian. The expected traffic conditions both influence and are influenced by pedestrian travel demands and walking behavior [Hoogendoorn and Bovy, 2004].

A variety of existing models have been developed to simulate pedestrian behavior according to these factors. They can be categorized into 4 types, namely, network models, queuing models, force-based models, and discrete choice models.

Network models

Network models are a type of macroscopic model, describing pedestrian flows by aggregate density and speed. They simulate the pedestrian route choice within a predefined network, where space is represented by links and pedestrians are simplified as continuous flows. Daamen [2002] developed a simulation tool for pedestrian flow modelling at large transfer stations, partly based on a network model. This model assigns pedestrians with certain OD pairs to the optimal routes in a network, based on the shortest time principle. It derives the density in an area from the number of pedestrians on a certain link and develops macroscopic relations between density and pedestrian speeds. Another application of network models is conducted by Hänseler et al. [2017]. Following the principles of the well-known cell transmission model [Daganzo, 1994] which discretizes time and space to simulate vehicular traffic, Hänseler et al. [2017] developed a dynamic network loading model that bases on a multi-directional-discretized formulation of a pedestrian fundamental diagram, adopted from the cell transmission model.

This type of model is capable of reproducing empirical walking time distributions in counter-flow and cross-flow experiments. Network models have also shown advantages simulating large crowds at low densities.

Queuing models

Queuing models [Løvås, 1994] describe individual pedestrian behavior in a queuing network, where pedestrians move between nodes (rooms) through links (doorways) in the network. Each link is randomly assigned with a waiting time, as a queue builds up when the demand exceeds the node capacity. The space taken by queuing pedestrians is not taken into account in the model, which let vertical queues build up. Each pedestrian follows a prespecified evacuation plan, which is determined by the rules of perceiving short routes and personal attributes. Different from network models, queuing models simulate individual behavior rather than pedestrian flows. Queuing models have been developed by other researchers as well Watts Jr [1987]; Yuhaski and Smith [1989]; Rahman et al. [2013], with different network and queue settings.

Queuing models have limited applications for its single-purpose assumptions, which ignore the dynamics inside each node (room) Johansson and Kretz [2012]. They have been used mostly for simulating pedestrian evacuation behavior from buildings to assist and evaluate emergency infrastructure design.

NOMAD model

Hoogendoorn and Bovy [2004] developed a normative theory of pedestrian behavior choice at the tactical level based on utility maximization under uncertainty. It assumes that each pedestrian continuously optimizes his/her own utility function for route choice, activity area choice, and activity scheduling. The function consists of the utility gained from performing activities at a specific location, the predicted cost of walking subject to the physical limitations of the pedestrians, and the kinematics of the pedestrian. The theory is different from discrete choice models as an infinite number of alternatives are available and the uncertainty of alternatives is considered. The theory has been applied in a microscopic model NOMAD to simulate pedestrian behavior in Schiphol Plaza, where the combined choice of route and activity area to exit Schiphol Plaza is simulated according to the shortest path based on minimum perceived disutility.

Both free flow and congested traffic can be simulated with the NOMAD model and stand-alone applications also allow simulating route choice in infrastructure facilities, such as transfer stations and shopping malls. Based on its formulation, NOMAD model is defined as a type of force-based model, which will be further explained in Section 2.1.3.

Discrete choice models

At tactical level, discrete choice models simulate the probability of an individual's route choice by considering the utilities of all alternative routes. Lue and Miller [2019] developed a path size logit model with stochastic route choice generation choice set based on the revealed preference from GPS data collected in Toronto. The model takes into account multiple route attributes, including length, number of turns, intersections, etc. and network characteristics, including percentage of links with sidewalks, road type, etc. A generalized path size factor is introduced in the model to correct for the correlation from overlapping alternatives. Attributes such as route distance, the number of turns, the number of signalized intersections, and distance along links with sidewalks on both sides of the street are proven to be significant in this model. Discrete choice models have also been applied in smaller-scale route planning, such as in a metro station where individuals' decisions of exit and routes to exits are simulated [Stubenschrott et al., 2014].

Based on the utilities taken into account in the model, different types of discrete choice models can be categorized. For instance, the shortest-distance model can be seen as the simplest model, as the utility of a route only consists of its distance.

Compared to more sophisticated discrete choice models, which require more computational efforts and a large amount of data to generalize, this shortest-distance model have been commonly adopted in simulation methods with the advantage of fast computation [Stubenschrott et al., 2014].

Discrete choice models provide a method to incorporate various internal and external influence factors in the decision making process of individuals. However, at tactical level, they usually merely focus on the route choice and are not able to simulate activity space and scheduling simultaneously.

Route choice decision principle

Route choice decision principle is a fundamental property of pedestrian route choice models. It defines the driven factors of why pedestrians choose a certain route. The decision principle can be different in different route choice models. For instance, in forced-based models, the decision principle is usually based on the shortest path, which is measured by time, distance, or utility, influences by the velocity and position of the pedestrian, people and obstacles in the vicinity, etc. [Hoogendoorn and Bovy, 2004]. In discrete choice models, the decision principle can more flexible, as more environmental and personal attributes, such as number of intersections, can be included in the model [Lue and Miller, 2019]. In some other empirical studies, the probability of route choice is derived from historical data and used as the decision principal [Ton et al., 2015]. Although this method is more commonly applied in vehicular and bicycle traffic studies, its potential in pedestrian modeling is yet to be discovered.

Conclusion

To conclude, models at tactical level are capable of different choice modeling, most of which mainly focus on route choice. Whereas the normative force-based model Hoogendoorn and Bovy [2004] provides an approach to simulate route choice, activity area choice, and activity scheduling simultaneously. This model also entitles other advantages such as taking into account individual differences and an infinite number of alternatives.

2.1.3 Operational level

The operational level choice consists of the actual walking behavior of pedestrians [Hoogendoorn and Bovy, 2004]. In previous studies, sometimes the route choice and movement choice are not clearly extinguished, causing confusion to readers. To specify the different, this paper defines the movement models as models representing how pedestrians react to the changing environment while moving along the chosen routes.

Existing models at this level can be classified into 5 types, which consist of continuum models, cellular automata models, force-based models, velocity-based models, and discrete choice models.

Continuum models

Continuum models [Treuille et al., 2006] are a type of macroscopic model, which simulates continuous pedestrian flows in discrete space cells in two dimensions. Global navigation and local collision avoidance are integrated while individual variability are ignored to achieve real-time crowd simulation with minimal computational effort.

This type of model is capable of simulating real-time crowd motions of thousands of individuals with intersecting paths. However, as individual characteristics are not taken into consideration, continuous models can only simulate homogeneous crowds and cannot cope with multidirectional pedestrian traffic.

Cellular automata models

Cellular automata models were first introduced to model pedestrian behavior by Blue and Adler [2001], studied and adapted further by Meyer-König et al. [2002]; Iltanen [2012]. In cellular automata models, three modes of bidirectional pedestrian flow are modeled based on predetermined local behavioral rules. This type of model often discretizes the space and time in which pedestrians move and interactions are determined based on a set of rules at each iteration. It is often used to evaluate the effects of behavioral rules or crowd management measures.

This rule-based strategy saves computational efforts compared to equation-based models and makes it easier to develop its algorithm. However, it is difficult to incorporate the randomness of pedestrian behavior and the fixed behavioral rules can lead to unrealistic movement in simulations.

Force-based models

The force-based models [Helbing and Molnar, 1995] assume that pedestrian behavior is influenced by a multitude of social forces that either attract or repulse the person, such as obstacles and the movement of other pedestrians. The acceleration of each person is determined by the sum of forces at each discrete time step. The movement space is continuous. This type of model has been widely applied and modified with extended force interpretations, such as collision prediction rules [Zanlungo et al., 2011] and self-stopping mechanism [Parisi et al., 2009].

The force-based models usually simulate fairly realistic movements and interactions between pedestrians. It is possible to incorporate the environmental features, such as light conditions and exit signs, in the analysis. However, higher computational efforts are required for force-based models.

Velocity-based models

The idea of velocity-based model is chronologically first put forward by Paris et al. [2007]. Based on the same principle, other velocity-based models have been developed over the years [Van Den Berg et al., 2008; Karamouzas and Overmars, 2010]. This type of model assumes that a pedestrian optimizes the usage of the available space and attempts to avoid collisions. Same as force-based models, velocity-based models assume pedestrians move in a continuous space and make decisions at each discrete time step. The velocity of a pedestrian is determined by the person's current path, available space, and the trajectories of other surrounding pedestrians. This type of model is often used to analyze evacuation behavior and crowd behavior at events.

Usually natural movements and pedestrian interactions are presented in velocity-based models, supported by high computational efforts. The downsides of velocity-based models also lie in the difficulty to model friction and incorporate environmental features.

Discrete choice models

Discrete choice models have also been developed at operational level to model the short-term behavior of individuals as a response to the immediate environment and the presence of other pedestrians. Antonini et al. [2006] modeled pedestrian walking process as a sequence of short-time choices. The destinations and routes are known, generated by models at tactical level. Each pedestrian makes choices out of a set of walking alternatives, based on utility maximization. The utility is calculated as the weighted sum of various elements of the potential next positions, such as change in kinetic energy, collision risk, availability, distance to other people, and distance to destination. The model was calibrated with real-life data and has been successfully applied in video surveillance applications for automatic tracking of

pedestrians in video sequences. Further adaption of this model was made by Robin et al. [2009], who introduced more data for revealed walking behavior, captured leader–follower and collision-avoidance patterns in simulation, and validated the model on another experimental data set.

Discrete choice models are not widely applied at operational level of pedestrian modelling, as the interaction complexity of pedestrians cannot be easily captured in the discrete choice model framework.

Conclusion

To sum up, at operational level, various types of models provide both macroscopic and microscopic approaches to simulate pedestrian movement, each having their own limitations. Continuous models require the minimum computational efforts but cannot cope with multidirectional pedestrian traffic. The rule-based cellular automata models either produce unrealistic movement or become very complicated with a large number of rules. The discrete choice models with decades of development in diverse fields encounter difficulties on revealing the complex movement dynamics of pedestrians. The force-based and velocity-based models generate natural movements but are computationally expensive. Operational behavior plays an important role for the detailed pedestrian interation modelling.

As mentioned before, most existing models do not simulate choices at more than one level, which makes it difficult to incorporate operational movements with tactical route choices in a simulation. Whereas force-based models have this capability to include choices at the tactical level explicitly. The NOMAD model, for instance, can consider "forces" from moving pedestrians in the process of routing simulation. Therefore, in this study, the force-based model, NOMAD will be used for pedestrian simulation.

2.1.4 Conclusion

To summarize, the reviewed models are categorized in Table 2.1.

	Macroscopic models	Microscopic models
Strategical level	Data-driven models Equilibrium assignment models	Discrete choice models
Tactical level	Network models	Queuing models Force-based models Discrete choice models
Operational level	Continuum models	Cellular automata models Force-based models Velocity-based models Discrete choice models

Table 2.1: Pedestrian behavior model overview

This study aims to reveal the dynamics of pedestrian interactions to assess the virus transmission risks of individuals. Therefore, the simulation scope falls into the category of microscopic models, which are also described as agent-based models, as the behavior is computed separately for each individual, i.e. the agent.

It is clear from Table 2.1 that agent-based models dominant the pedestrian modeling field. They are popular for several reasons [Treuille et al., 2006]. Firstly, the agent-based assumption is realistic, as real crowds clearly operate with each individual making independent decisions. Such models can capture each person's unique situation: visibility, proximity of other pedestrians, and other local factors. Therefore, an intuitive interpretation of equations is allowed. Secondly, different

simulation parameters may be defined for each crowd member, yielding complex heterogeneous decisions. However, these advantages also lead to drawbacks. For instance, these models are often stochastic, which requires multiple runs of simulation to get the "average". Due to the large number of parameters and the stochastic nature, it is difficult to develop and calibrate behavioral rules that consistently produce realistic decisions. Moreover, global path planning for each agent quickly becomes computationally expensive, particularly in real-time contexts. As a result, most microscopic models separate local collision avoidance from global path planning [Parisi et al., 2009; Zanlungo et al., 2011], and conflicts inevitably arise between these two competing goals.

As mentioned in Section 2.1.3, only the microscopic force-based model, NOMAD, fulfills the simulation requirements of this study. It is capable of incorporating tactical choices with operational choices with different 'forces'. Therefore, the NOMAD model will be used for pedestrian simulation. As for the strategical choices discussed in Section 2.1.1, the activity schedule of the entire visit to the simulated event in this study will be derived by a data-driven method from the analysis of real-life data. By determining the activity schedule in another model, the computational efforts can be saved as the pedestrian behavior at different activity spaces of this event are independently simulated.

2.2 VIRUS TRANSMISSION

To understand how SARS-CoV-2 virus transmits during interactions of people, its transmission routes are reviewed in Section 2.2.1. The models explaining the transmission of respiratory viruses in the crowd are reviewed in Section 2.2.2. Existing studies on SARS-CoV-2 virus transmission modelling are introduced in Section 2.2.3. As this study aims to incorporate pedestrian modelling with the spread of SARS-CoV-2 in the crowds at large events, the scope of the literature review on SARS-CoV-2 transmission is limited to the agent-based level, excluding the studies that model SARS-CoV-2 transmission based on the reproduction rate in the entire population.

2.2.1 Transmission routes

The spread of SARS-CoV-2 occur primarily through 3 routes, namely respiratory droplets, aerosols, and fomites [Harrison et al., 2020]. Although cases of fecal–oral transmission have also been reported, it is not likely to happen in event spaces and will not be discussed in this study.

Among the three primary routes, droplet spread is considered as the main route of transmission. Droplets are large liquid particles (usually defined as diameter larger than 5 μm) loaded with viruses that spread into the air by infected people when sneezing, coughing, talking, or breathing. These particles directly project onto the mucous membranes or upper respiratory tract of a susceptible individual through the person's mouth, nose, or eyes [Morawska and Cao, 2020]. As droplets can travel over limited distances, close, concurrent contacts are required for droplet transmission to occur.

Aerosols and fomites transmission are defined as indirect routes, as they build up the potential of virus in the environment. Aerosols are formed by small virus-laden liquid particles (usually defined as diameter smaller than or equal to 5 μm) evaporated from the droplets of infected people. These particles are so small that transport by air current affects them more than gravitation. Therefore, they are free to travel in the air and carry their viral content to meters and tens of meters away from where they originated [Morawska and Cao, 2020]. Moreover, compared to droplets, it usually takes much longer for aerosols in space (especially indoor space with poor ventilation) to expire, which means that the possibility exists that individ-

uals can be infected with SARS-CoV-2 by being in the same space where infected people have stayed, spread the virus, and left. Studies have claimed that SARS-CoV-2 virus RNA could be maintained in respirable-sized aerosols for up to 16 hours [Fears et al., 2020]. Fomites transmission occurs when an individual touches a contaminated surface where the virus has landed on via droplets or aerosols from infected people. Among the three primary routes, SARS-CoV-2 stays infectious the longest on contaminated surfaces. Viable virus was detected up to 72 hours after laboratory application to plastic and stainless steel surfaces [Van Doremalen et al., 2020].

A summary of possible transmission routes are shown in Figure 2.1. It is important to consider different routes at different event spaces to assess the overall risks of individuals participating in the event. For instance, transmission by fomites and aerosols is not likely to occur at outdoor music stages where people do not touch any surface and natural ventilation prevents aerosols from coming into being. However, indoor spaces, such as toilets and indoor music stages, pose great risks of aerosols transmission as indoor ventilation does not eliminate the possibility of long-existing aerosols. At bars and food stands, fomites transmission plays and important role as people cannot avoid touching different kinds of surfaces. Therefore, droplets, aerosols, and fomites are considered as the main transmission routes of SARS-CoV-2 in event spaces, but each presents a different degree of risk at different spaces. The virus transmission models related to these routes will be presented in the next subsection.

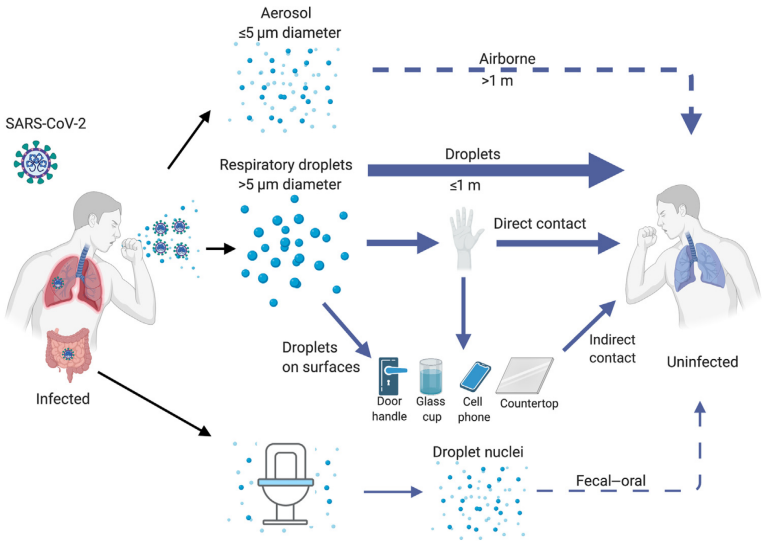


Figure 2.1: Transmission routes of SARS-CoV-2 (Source: Harrison et al. [2020])

2.2.2 Transmission models

In the scientific field of virus transmission, a large range of models have been developed to describe the transmission mechanisms of different types of viruses. This study specifically aims to understand the infection risk of SARS-CoV-2 virus among visitors at event spaces. Therefore, only models explaining the transmission mechanisms of respiratory viruses will be reviewed in this subsection. They are the Wells-Riley model, dose-response models, and multi-route transmission models. The first two mainly focus on airborne transmission.

Wells-Riley model

The foundation of airborne transmission models is the Wells-Riley model, proposed by Wells et al. [1955]. By assuming that infectious particles are well mixed and steady in space, the Wells-Riley equation makes prediction of the number of new cases in a certain time period based on the number of infected and susceptible individuals in an indoor space, the ventilation rate, and the quantity of infectious material in the air. The concept of "quantum of infection" is used in the Wells-Riley model, implicitly taking into account of the infectivity, infectious source strength, biological decay of pathogens, etc. [Zhang and Lin, 2020]. A quantum unit is defined as the number of infectious particles needed to cause 63% of individuals getting infected. Its number is usually obtained through comparing infection data in a well-defined indoor outbreak. Providing a simple and quick evaluation method of airborne infection risk, the Wells-Riley model has been widely applied and adapted in infectious respiratory disease studies. However, the model's assumption of spatially and temporally uniform distribution of virus-laden aerosols does not allow heterogeneous infectivity of individuals as everyone has the same risk of infection at any time. Therefore, the Wells-Riley model does not satisfy the analysis of individual infection risk in this study.

Dose-response models

Another quantitative evaluation method is the dose-response model, which allows individual heterogeneity in the analysis of infectivity. This type of model describes the magnitude of the response, namely risk of infection, as a function of exposure to different doses of a pathogen after a certain exposure duration [Crump et al., 1976; Nicas, 1996]. The exposure dose is the amount of pathogens that reach the susceptible individuals under the circumstance of certain infectious people's emission rates, room volume, recipients' pulmonary ventilation rate (i.e. the amount of air inhaled per unit time), and exposure duration. Based on the type of the mathematical function, dose-response models can be categorized into different types, for instance, exponential and beta-Poisson models [Watanabe et al., 2010]. The dose-response models are more flexible than the Wells-Riley model, as individual characteristics can be taken into account to determine different infection risks for different people. However, the limitation lies in the assumption of a uniform spatial distribution of infectious particles, by which the risks of pedestrian interaction are directly simplified as being or having been in the same indoor space. This study aims to evaluate the dynamic infection risks of individuals when moving and interacting with each other. It is expected that the distances between people and their trajectories feature different levels of risk. Therefore, the basic-assumption dose-response models do not meet the evaluation requirements of this study.

Multi-route transmission models

Multi-route transmission models incorporate airborne transmission routes with transmission by formites. This type of model has been developed for different types of respiratory viruses. One of the widely applied frameworks is developed by Nicas and Sun [2006], which integrates different source-environment-receptor pathways and their physical elements in a discrete-time Markov chain model. Initial pathogen loads are assumed on textile and nontextile surfaces and in the room air, depending on the amount of pathogens emitted from the patient in events such as coughing and body fluid discharges. A uniform distribution of aerosols in the room is assumed. The pathogens can exchange between the surface and the air, due to particle settling and resuspension. These pathogens follow different pathways to end up in people's respiratory tract, mucous membranes, losing viability, or being exhausted from the room. The model estimates the rate of transfer at each step in the pathway and the probability of a pathogen moving from one "state" to another "state" by

the end of a specified time interval. As a result, the expected pathogen dose to an individual's mucous membranes and respiratory tract is estimated. A nonthreshold infectious dose model is used to relate the expected dose to infection risk. This framework is composed of discrete events and therefore can be expanded for more transmission pathways in detail.

Conlusion

In conclusion, either the Wells-Reily model or a dose-response model alone is not capable of detailed virus transmission modelling. They need to be incorporated with other models or be adjusted with specific settings for more complicated airborne transmission situations. To include multiple transmission routes, the study by Nicas and Sun [2006] provides a modelling framework of discrete time and events, by which the detailed transmission processes can be modeled. More multi-route transmission models related to SARS-CoV-2 transmission will be reviewed in the next section.

2.2.3 Existing studies on SARS-CoV-2 transmission in the crowd

Numerous studies revealing the transmission dynamics of SARS-CoV-2 have been conducted since the start of the COVID-19 pandemic in December 2019, most of which focus on evaluating the spread of the disease on a relatively large scale. In this subsection, the models developed for SARS-CoV-2 transmission which incorporate multiple transmission routes and describing the transmission dynamics at the crowd interaction level will be reviewed.

Arav et al. [2020] developed a multi-route quantitative mechanistic mathematical model for pre-symptomatic transmission of SARS-CoV-2. The model tracks the transmission dynamics based on individual activities, focusing on three transmission routes between people, namely, direct physical contact, fomites, and aerosols (droplet nuclei). Instead of the common threshold of 5 μm, 100 μm is used as the cutoff size between droplets and aerosols, as droplets smaller than 100 μm evaporate to their droplet nucleus size before hitting the ground. With this conservative assumption of the cutoff size, an overestimation of aerosols transmission contribution is expected. Moreover, droplet transmission is not considered in the study. This stochastic model simulates a number of scenarios by generating an ensemble of realizations via Monte-Carlo simulation. In each realization, the primary (infetor) and secondary (infectee) individuals perform a series of randomized actions such as touching each other, touching fomites, or touching their own faces. The hygienic and behavioral parameters are obtained by other empirical studies. The duration of each realization is the incubation time of the primary individual, which is on average 5 days, during which the viral load of the person increases exponentially until reaching the level of showing symptoms. The probability of a secondary individual being infected is inferred from the dose-response curve reported for SARS-CoV-1 [Watanabe et al., 2010].

This method provides a relatively detailed modelling framework of the agent-based transmission process. However, the main limitation lies in the randomly generated pedestrian interaction, which does not consider the actual movement of pedestrians. As a result, heterogeneous aerosols distribution cannot be incorporated in the model.

Gao et al. [2021] developed a multi-route transmission model considering (1)long-range airborne transmission, (2)short-range airborne transmission, (3)direction inhalation of medium droplets or droplet nuclei, (4)direct deposition of droplets of all sizes, (5)indirect contact. The model combines the Wells-Riley equation and the dose-response model to calculate the infection risk of an individual from all trans-

mission routes in a certain location. Different dose-response coefficients are defined for different routes to account for their potential different dose-response rates. The exposure doses of the 5 considered routes are based on the following factors: (1) For long-range airborne transmission, the exposure dose is calculated according to the cumulative deposition infectious dose in the susceptible's respiratory tract, taking into account of pulmonary ventilation rate, exposure time of the infectee, room volume, the air change rate in the room, particle loss rate, etc. (2) As for short-range airborne transmission, a respiratory jet cone with a certain spreading angle is assumed to transmit the droplet nuclei to the infectee. The exposure dose is calculated according to the distance to the infector, initial concentration and dilution rate along the cone, face-to-face contact time, etc. (3) The exposure dose due to direct inhalation of medium droplets or droplet nuclei is based on similar factors of (2), considering additional parameters for larger droplets, modelled by combing the buoyant round jet model and droplet evaporation and motion models. (4) The exposure dose due to direct deposition in the facial membranes follows the assumption of the respiratory jet cone and is estimated based on a similar parameter set. (5) As for hand-surface contact, factors such as the frequency of the hand touching facial membranes, transmission rate of droplets from hand to facial membranes, contact area of the hand to mucous membranes, etc. are used to calculate the exposure dose.

Although the parameter estimation of this model is based on influenza, the detailed and realistic consideration of the multi-route transmission is applicable to different types of respiratory infections, including SARS-CoV-2.

Bouchnita and Jebrane [2020] developed a hybrid model combining an Susceptible - Exposed - Infectious - Recovered (SEIR) model and a social force model to describe the transmission of SARS-CoV-2 in 250 individuals. The SEIR model is a type of compartmental models, which models the disease transmission in a closed population, where four classes of people are identified, namely susceptible, exposed, infectious, and recovered. This type of model is widely applied to describe the spread of diseases with a long incubation period, such as COVID-19. The agent-based social force model allows transmission tracking at the individual level. In this study, it is assumed that people, all potentially symptomatic, move in random directions in a square space of $250 \times 250 m^2$. A pre-symptomatic individual can transmit the virus to other people via two routes, direct contact (droplets) and indirect contact (formites). The threshold of direct contact is set as 1 meter. It is assumed that the virus is transmitted when certain interactions such as sneezing, coughing, or handshaking happen. A Bernoulli distribution is assumed for person-person transmission. The formites transmission rate is based on the estimated lifetime of SARS-CoV-2 surviving on a hard surface and the probability of an individual touching the hard surface. The demographic characteristics determine the mortality risk of infected agents. The models simulate the spread of COVID-19 with and without nonpharmaceutical interventions and are calibrated with real-life data in two situations.

Though the model development is based on a number of unrealistic assumptions, such as enclosing individuals for 90 days and hard surfaces existing everywhere, this hybrid framework provides an inspiring solution to incorporate macroscopic epidemiological models with agent-based pedestrian models for virus transmission modelling.

An ongoing research conducted by TU Delft and Wageningen university [Duives et al., 2021], combines 4 models to simulate the risk of SARS-CoV-2 infection of individuals during a visit to a restaurant, the model chain of which is illustrated in Figure 2.2. The models include an activity scheduler, a microscopic pedestrian model, NOMAD, a virus spread model, QVEmod, and a risk identification model. The activity scheduler determines the strategical choices of individuals, including activity choices, destination choices, and departure time choices, based on the con-

text, spatial layout, population, demand, and COVID-19 control measures. The NOMAD model, as discussed in Section 2.1, simulates pedestrian routing and movement dynamics based on the minimum walking cost principle. Its output consists of a set of trajectories which pertain the coordinates and velocity of each individual at each time step of the simulation. The QVEmod model then takes the output of NOMAD, combined with the epidemiological attributes of the environment and individuals, to simulate the emission of virus from infectious individuals, the spread of the emitted virus in the environment (in droplets, aerosols, and formites), and the virus dose landed on susceptible individuals at corresponding time steps to the NOMAD output. The risk identification model then identifies the risk of individuals being infected by the accumulated virus dose and interaction locations of high virus exposure based on the accumulated virus dose in the environment. The connection between virus dose and infection risk is built based on the exponential dose-response relationship developed by Nicas [1996].

This research provides an agent-based activity-specific simulation framework for SARS-CoV-2 infection risk identification, with realistic activity planning, route simulation, virus spread, and the relation between the virus dose and the infection risk. It makes up for the gap of fitting virus transmission modelling into real-life situation, the framework of which can be adapted in different scenarios to evaluate the transmission risk in other public spaces.

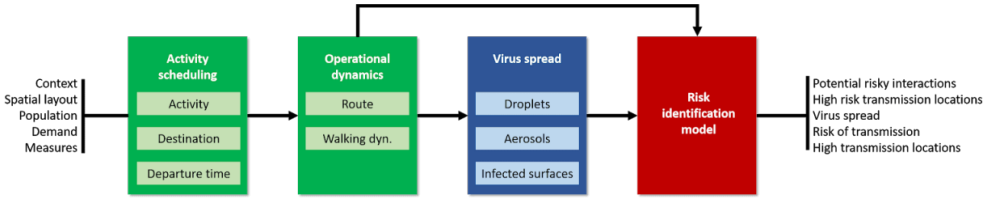


Figure 2.2: Model chain of ongoing research to identify the SARS-CoV-2 transmission risk

To conclude, these multi-route agent-based models all incorporate dose-response models to define the threshold of infection, referring to empirical dose-response curve of SARS-CoV-1 [Arav et al., 2020], influenza data [Gao et al., 2021], or assumptions of an SEIR model [Bouchnita and Jebrane, 2020]. For the consideration of individual behavior, Arav et al. [2020] generate random actions and interactions by Monte-Carlo simulations at discrete time steps; Gao et al. [2021] use predefined frequency of individual behavior from empirical studies; Bouchnita and Jebrane [2020] simulate agent movements by a social force model, incorporated with thresholds of different types of contact and transmission rates; the ongoing research models the activity scheduling, routing ans movements, fitting the most realistic pedestrian behavior to the virus transmission modelling.

Among the reviewed models, only the hybrid modeling framework [Bouchnita and Jebrane, 2020] and the activity-specific simulation framework from the ongoing research [Duives et al., 2021] include both pedestrian interaction simulation and virus spread simulation in their models, which fits the scope of this study. However, as discussed in Section 2.1.4, this study separately simulates pedestrian behavior at different activity spaces for a short time period during an event, which operates in the same way for the virus spread simulation. Therefore, the activity-specific simulation framework [Duives et al., 2021], which combines the NOMAD model with a microscopic virus spread model, will be applied in this research.

2.3 RISK ASSESSMENT

In this section, the risk assessment methods which identify the infection risks from virus spread simulation results and real-life experiments will be reviewed, as well as the risk quantification and visualization methods.

2.3.1 Risk identification

The reviewed risk identification methods include approaches used in simulations and real-life experiments.

Risk identification in simulations

Jitsuk et al. [2020] estimated the SARS-CoV-2 transmission risk of the cancelled 2020 Songkran festival in Thailand based on the estimated contact between people during the festival and the disease transmission rates obtained in Wuhan, China. Instead of simulating individual contact during the event, this study approximated an average contact frequency throughout the festival based on historical expense data and calculated the number of infections in the entire population based on the initial number of infections before the festival. Although this 'macroscopic' method provides insights on the possible impact of a national festival, it gives no indication of the source of infection risk, i.e., which facilities of the event lead to a high possibility of SARS-CoV-2 transmission, and therefore cannot assist decision making on interventions and regulations at the event to prevent SARS-CoV-2 transmission.

To identify the SARS-CoV-2 transmission risk from the accumulated virus doses on the individuals, obtained from the agent-based virus spread model [Duives et al., 2021], there are two types of methods. These methods translate the accumulated virus dose into infection risk by defining infection risk according to the pathogen dose that have been used in previous studies, namely threshold and nonthreshold models. A threshold model assumes that when the host receives a certain amount of pathogens, the infection is certain to occur, whereas when the received amount is smaller than that, the infection will certainly not occur [Nicas and Sun, 2006]. A nonthreshold model defines the infection risk as a probability by two parameters, the probability of a single organism infecting the host, denoted as α, and the expected dose that imparts a 50% chance of infecting a random individual who receives it, denoted as ID_{50} [Nicas and Sun, 2006]. They are related by the equation $ID_{50} = ln(2) \div \alpha$.

These two models are, in essence, both dose-response models. The nonthreshold method has been widely applied in infection risk identification studies [Watanabe et al., 2010; Duives et al., 2021], as it indicates a probability of infection, which implies how risky a certain amount of virus dose is. Therefore, a nonthreshold dose response model developed by Nicas and Sun [2006], applied by Duives et al. [2021] will be used in this study. This model builds up an exponential relationship between the infection risk and the exposure to the number of viral particles from different transmission routes, as the risk is influence by deposition locations (hands, lower or upper respiratory tract) and the viability of the virus [Deng et al., 2020]. This equation will be presented and further discussed in Equation 4.23 in Section 4.3.5.

Please note that, although this study separately simulates virus spread in different activity spaces, the infection risk is not derived from separate activity spaces. The purpose of this research is to derive a general risk profile of one visit to an event. As the infection risk is not linearly related to the virus dose, in this study, the infection risk of one individual from one visit to the event is determined by the accumulated virus dose after visiting a set of activity spaces. The activity spaces are

defined by the activity schedule derived from a data-driven method, as discussed in Section 2.1.1.

Risk identification in real-life experiments

Besides estimating the SARS-CoV-2 transmission risk from simulations, studies that derive the transmission scale from real-life experimental events have also been conducted.

From February 2021 to May 2021, FieldlabEvenementen [2021] have conducted 24 experimental events in the Netherlands, in cooperation with the government, universities, research groups, event organizers, and voluntary participants. The experimental events were designed carefully on the scale, facilities, interventions, and regulations to explore the actual risks of different types of events with different limitations. Within 24 hours before the events, participants were required to do a rapid test for COVID-19 and only people with negative test results are allowed in. During the events, participants were informed with expected behavioral rules, such as wearing face masks, keeping social distance, etc. Their trajectories were tracked throughout the event, from which the contact profile was derived, summarizing the contact characteristics including distance, duration, and time of occurrence. The contacts were categorized into different risk levels according to the distance and duration. After the events, the participants were again tested for COVID-19 and the number of positive cases are analyzed as an indicator for the event risk. In conclusion, the experimental events identify the transmission risk directly by the actual infection and indirectly by tracing high-risk contacts.

The advantage of real-life experiments is self-explaining. The transmission scales of certain types of events are directly revealed by the infection rate of the participation population. However, ethical concerns about the experiments exposing participants to potential high risks of infection have been raised among people including scientists [de Vrieze, 2021; RTLNieuws, 2021]. Moreover, as the method is based on tests, the accuracy of tests and the potential infections happened during commuting to and from the event location both have an impact on the results, which is difficult to measure. On the aspect of decision making support, the experimental results may have limited capabilities to offer advice on effective measures at future events, as a small number of experiments may not accurately imply the influence of different kinds of interventions on the transmission scale. A large number of variables exist in real-life experiments that are difficult to measure and control. For instance, at the same type of events in similar scale and facilities, the respiratory activities of people can be very different depending on the live atmosphere, which in turn lead to very different infection numbers. Therefore, a stable simulation tool that takes control of variables may be preferred on the aspect of decision making support. In addition, simulations do not face ethical criticism and can limit the uncertainties in the experiments, such as infections happened outside the event terrain.

2.3.2 Risk visualization

After identifying the infection probability, it is important to visualize and analyze the detected risk in order to assist the decision making of COVID-19 related regulations. The risk assessment methods being used for existing repository virus transmission studies are briefly reviewed in this section.

The most direct method is to simply illustrate the infection probability under different circumstances in number or bar charts. Gao et al. [2021] demonstrated the SARS-CoV-2 infection risks of 6 transmission routes between one infector and one infectee by probability. The contribution of each route is identified under different exposure distances. The relations of infection risk to environmental factors are derived and visualized in bar charts and line charts.

Besides using probability, the actual infection number (percentage) has also been used for risk assessment. Bouchnita and Jebrane [2020] used the infection percentage of the population plotted in line charts to illustrate the development of SARS-CoV-2 penetration since the hospitalization of the first patient. The concentration of SARS-CoV-2 in space and the status of individuals are also visualized in a heat map, where different colors of nodes represent people of different statuses, as shown in Figure 2.3.

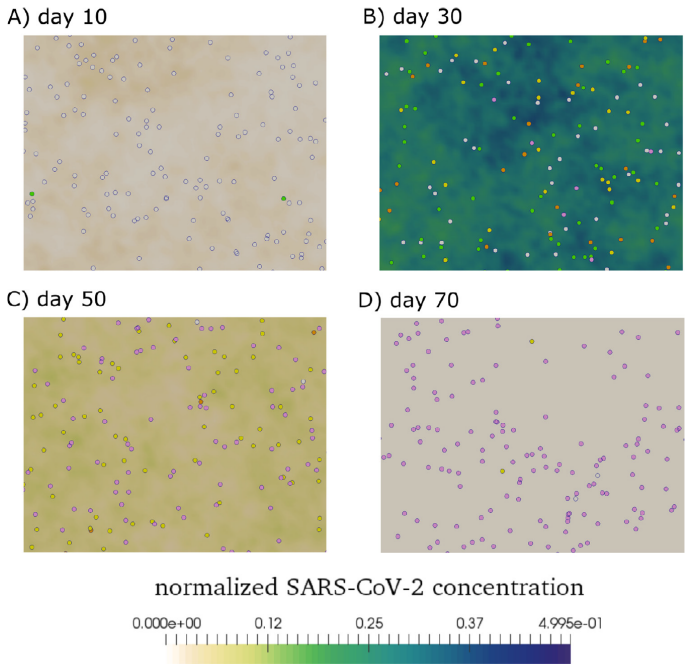


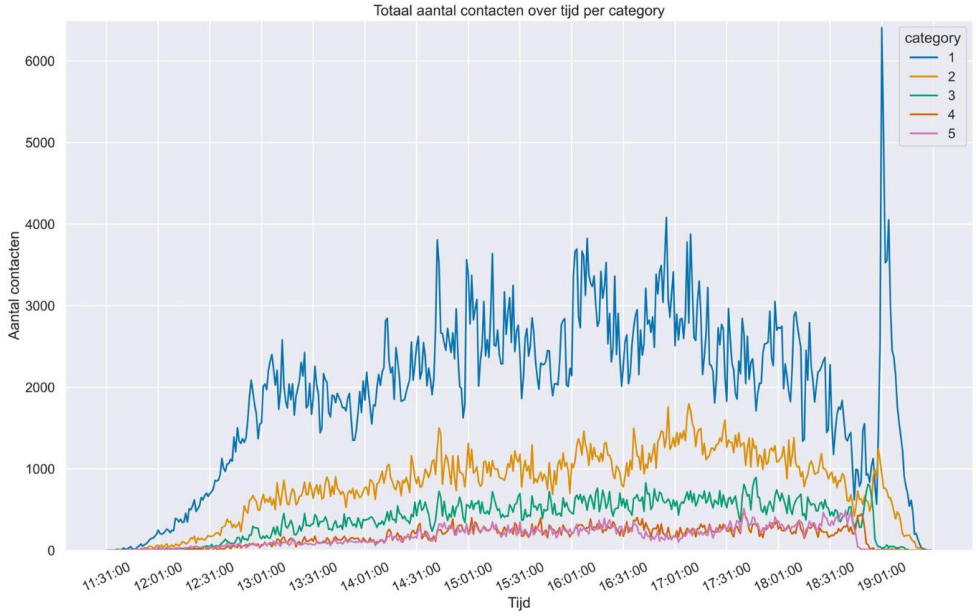
Figure 2.3: Visualization of individual's infection status and SARS-CoV-2 concentration in space (Source: Bouchnita and Jebrane [2020])

Another assessment method is based on the basic reproduction number (R_0), which is the expected number of infections directly generated by one infector in a population where all the individuals are susceptible [Fraser et al., 2009]. When the considered population is not fully uninfected, the number is defined as the effective reproduction number (R_e). Similarly, when interventions are taken, controlled reproduction number (R_c) is used. For the derivation of these R values, usually large real-life data sets or simulations with the duration from days to months are used [Nikbakht et al., 2019; Ferretti et al., 2020]. Methods to estimate R_0 include the exponential growth rate-based (EGR) method, Susceptible - Infectious - Recovered (SIR) models, etc. [You et al., 2020], which usually do not consider the details of contact between individuals. Although this study only focuses on the infection risk of individuals at a short-duration event, R_0 can be used as an indicator to evaluate the transmission scale of SARS-CoV-2 in an event under different scenarios.

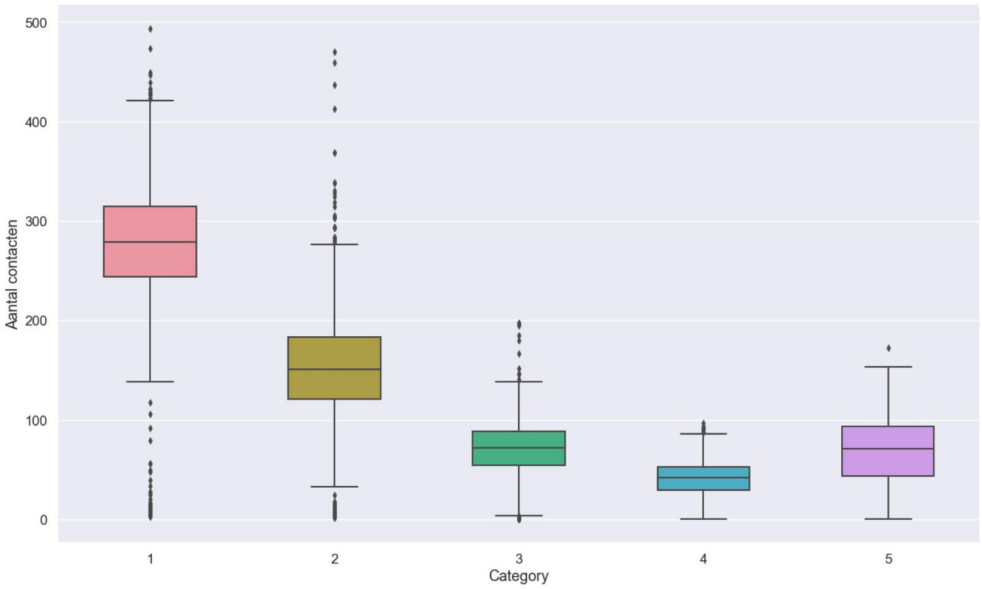
Besides direct indicators based on infection numbers and probabilities, indirect indicators have also been used to visualize infection risks. In the study of real-life experiments by FieldlabEvenementen [2021], the contacts between people are also used as a risk indicator as displayed in Figure 2.4. As epidemiological studies suggest, the possibility of airborne transmission of SARS-CoV-2 goes up when the distance between people goes down and the contact time goes up [Harrison et al., 2020; Morawska and Cao, 2020]. Therefore, categorizing the contacts by duration and distance can represent the SARS-CoV-2 transmission risk to a certain extent. In Figure 2.4, each contact category represents a risk level. The longer the duration is, the more risky the contacts become. The contact distribution of different events are compared to assess the potential infection risks and the sources of risks at these events.

However, as suggested by epidemiological studies [Harrison et al., 2020; Fears et al., 2020], the risk of SARS-CoV-2 transmission lies beyond close contact. Viral-laden droplet and aerosol particles remain for minutes and even hours on the location where the infectious individuals have stayed, depending on the ventilation rate. In this case, people can still catch virus by staying at the same location, several minutes or hours after the infectious individuals have left, without having any contact with them. Moreover, fomites transmission cannot be represented by contact tracing, as it happens when people touch the same surface. For events where touching surface is unavoidable, such as paying for drinks and using the lockers, the contribution of fomites transmission is not negligible. Therefore, contact tracing cannot accurately reflect the SARS-CoV-2 transmission risks during the entire event and will not be used in this study.

To conclude, considering the simulation setting of this study, the infection risk of individuals will be assessed by the probability of infection, which is translated from the accumulated virus dose after performing a set of activities. To evaluate the overall risks of individuals at the event, the number of infected individuals and R_0 will be the indicators to compare different event scenarios. The risk in different types of event spaces will be analyzed on the scale of virus exposure, which is visualized in bar plots.



(a) Number of different types of contacts over time



(b) Box plot of the total number of different types of contacts

Figure 2.4: Visualization of contact tracing at Dance Festival in Biddinghuizen (Category 1: contact distance ¡ 1.5m, contact time: 10sec - 1min; category 2: contact distance ¡ 1.5m, contact time: 1min - 5min; category 3: contact distance ¡ 1.5m, contact time: 5min - 10min; category 4: contact distance ¡ 1.5m, contact time: 10min - 15min; category 5: contact distance ¡ 1.5m, contact time ¿ 15min. Source: FieldlabEvenementen [2021])

3 FRAMEWORK DEVELOPMENT

In this chapter, the models and methods reviewed in Chapter 2 will be placed in a theoretical framework for the development of a SARS-CoV-2 transmission risk identification method.

3.1 FINDINGS FROM LITERATURE REVIEW

As discussed in Chapter 2, a data driven approach to derive the activity pattern, a forced-based pedestrian model, NOMAD Hoogendoorn and Bovy [2004]; Campanella [2016], a virus spread model, QVEmod [Duives et al., 2021], and a dose-response model [Duives et al., 2021; Nicas and Sun, 2006] will be adjusted and applied to develop the research method of this study.

However, the revised models together do not form a full picture of all the required steps to conduct this research. The purpose of this study is to drive a general infection risk profile of a large event, which requires combining individuals' activity schedules and virus spread situation at different event spaces.

To incorporate existing models and make computation possible, this study simulates pedestrian behavior and virus spread separately at different activity spaces of an event. As reviewed in Chapter 2, existing studies simulate only one enclosed area/public space and then derive the infection risk merely based on one visit to the simulated place, which builds up a research gap of connecting the infection risks of multiple activity spaces during a visit to a large event.

To fill the gap, this study proposes a risk quantification method which connects the processes of simulating transmitted virus doses in multiple activity areas and deriving the infection risk of individuals after visiting a number of activity spaces. The proposed method is fit into the theoretical framework and described in the following section.

3.2 THEORETICAL FRAMEWORK

Based on the above discussion, a theoretical framework of the proposed method is adapted from Figure 2.2. As is illustrated in Figure 3.1, the proposed method consists of 3 major parts, namely, the NOMAD pedestrian model, the virus spread model, and the risk quantification steps. Although the input processing steps (in dashed blue) are not part of the proposed method, they are essential for the application of the method. In the input processing steps, the input to the NOMAD pedestrian model and virus spread model is derived from the official databases, real-life data analysis, and existing researches. In addition, the activity schedule is generated from a data driven approach using real-life GPS data. The input processing steps will be thoroughly explained within each model in Chapter 4.

The first part of the method is the NOMAD model [Hoogendoorn and Bovy, 2004; Campanella, 2016], which is selected from a number of pedestrian behavior models reviewed in Section 2.1. This model was selected for its agent-based microscopic nature, continuous space and time assumption, capability to include tactical and operational factors, and realistic choice simulation. NOMAD takes the input of infrastructure settings, pedestrian movement parameters, and demand patterns, at

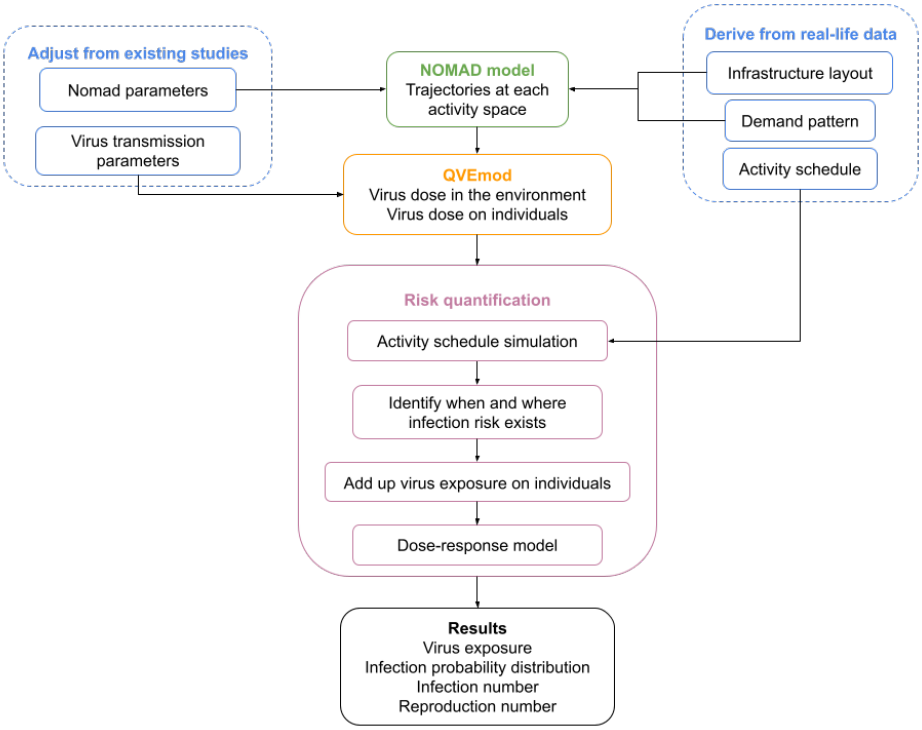


Figure 3.1: Theoretical framework

each type of activity space. As is shown in Figure 3.1, the NOMAD parameters are adapted from the existing studies to make sure the simulation of natural route choices and movements. The parameters can also be varied according to different age and gender distribution, which influences the optimal walking speed, and the 1.5 meter social distancing rule, which influences the 'forces' from other pedestrians. The output of NOMAD consist of individual trajectories within the infrastructure during the simulated period of time, which is fed to the QVEmod model for the virus spread simulation.

QVEmod [Duives et al., 2021] simulates the emissions of viruses from infectious individuals during the activities performed within the simulated space, the landing of viruses in the environment at each time step of the simulation, and eventually the viruses that get taken up by susceptible individuals. The time and space are both discrete in QVEmod. The input consists of the infrastructure settings, including definitions of touchable surfaces, movements of individuals, environmental and personal parameters of virus transmission dynamics. As illustrated in Figure 3.1, the QVEmod parameters are adjusted according to existing studies on the transmission and decay mechanics of SARS-CoV-2, which vary in different environments. For instance, personal parameters can be influenced by whether face masks are mandatory and respiratory activities. The environmental parameters can be influenced by the ventilation rate. After simultaneous simulation of virus transmission via three routes, the output is produced, which includes the accumulated virus dose in the environment and on individuals at each time step of the simulation.

After the first two steps, a general accumulated virus distribution on individuals is obtained independently in each activity space.

To derive the general infection risk of individuals visiting the entire event, the author proposed a risk assessment method which consists of four steps, namely, activity schedule simulation, identification of the location and duration of possible virus exposure, adding up accumulated virus doses, and a dose-response model [Duives et al., 2021; Nicas, 1996]. As discussed in the previous section, this quantification method is proposed to fill the gap of simulating infection risks at large events where participants visit multiple facilities.

The first step generates every visitor's activity schedule at the event, identifying where and when the infectious and susceptible individuals are located. The second step summaries the whereabouts of infectious individuals and derives the locations and duration of possible virus exposure. In the third step, the susceptible individuals' activity schedules are compared with the derived locations and duration. If a match exists, a virus exposure (obtained from the virus spread simulation) is assigned to this individual. For all the susceptible individuals, the virus exposure acquired throughout the activity schedule is added up. Finally, the accumulated virus exposure is fit into a dose-response model to calculate the general infection risk of one individual visiting this event.

From the infection risk distribution, the number of infections at this event can be calculated, as well as the reproduction number, which quantifies the general infection risk of this event. From the accumulated virus distribution of the entire event, the proportion of virus transmitted in each type of activity space can be obtained, which reflects the scale of transmission risk at facilities. Moreover, the proportion of virus transmitted via each transmission route can be derived, which reflects to what extent the route contributes to the infection profile. By analyzing the results, the facilities/routes with high risks of transmission can be identified, which can help the decision makers evaluate the infrastructure of the event, redesign the facilities, and carry out more restrictions.

The models and methods presented in this theoretical framework will be further discussed in Chapter 4.

4 METHODOLOGY

In this chapter, the method used to conduct this research is thoroughly described. 4 models of the proposed method, including the NOMAD model, the QVEmod model, contact probability model, and risk identification model, are discussed as follows.

4.1 NOMAD MODEL

As is introduced in Section 2.1, the NOMAD model, developed by Hoogendoorn and Bovy [2004], simulate pedestrian routing behavior based on the minimum walking cost (maximum utility) principle. Pedestrians are assumed to continuously balance their movement to the destination with the cost of their behavior, such as travel time and physical efforts. In the following sections, the model dynamics and case-specific components of the NOMAD model are introduced.

4.1.1 Model dynamics

This subsection describes NOMAD model dynamics according to Campanella [2016] and the ongoing project making use of NOMAD model for restaurant infection risk analysis. The fundamental equations in NOMAD are displayed as follows.

$$\frac{d}{dt}\vec{r}_p(t) = \vec{v}_p(t) \tag{4.1}$$

where:
$\vec{r}_p(t)$ is the position vector at time t.
$\vec{v}_p(t)$ is the velocity.

$$\frac{d}{dt}\vec{v}_p(t) = \vec{a}(t) \tag{4.2}$$

where:
$\vec{a}(t)$ is the acceleration.

$$\vec{a}(t) = \vec{a}_c(t) + \vec{a}_p(t) + \vec{\varepsilon}(t) \tag{4.3}$$

where:
$\vec{a}_c(t)$ is the controlled acceleration.
$\vec{a}_p(t)$ is the physical acceleration (uncontrolled).
$\vec{\varepsilon}(t)$ is the noise term.

$$\vec{a}_c(t) = \vec{a}_s(t) + \vec{a}_O(t) + \vec{a}_{pq}(t) \tag{4.4}$$

where:
$\vec{a}_s(t)$ is the path straying component.
$\vec{a}_O(t)$ is the obstacle interaction component.

$\vec{a}_{pq}(t)$ is the pedestrian interaction component.

$$\vec{a}_s(t) = \frac{(v_0(t) \cdot \vec{e}_g) - \vec{v}(t)}{\tau} \tag{4.5}$$

where:
$(v_0(t) \cdot \vec{e}_g$ is the optimal velocity (speed and direction).
τ is the constant acceleration time.

$$\vec{a}_{pq}(t) = -\vec{e}_{pq} \cdot A_0 \cdot e^{\frac{-d_{pq}}{d_i}} \tag{4.6}$$

where:
$\vec{e}_{pq}$ is the unity vector in the normal direction pointing to the other pedestrian.
A_0 is the interaction strength parameter.
d_{pq} is the anticipated distance between pedestrians.
d_i is the interaction distance parameter.

$$\vec{a}_O(t) = -\vec{e}_O \cdot A_O \sum_{o \epsilon O} \begin{cases} 1 & \text{for } 0 < d_{pO} < d_0 \\ 1 - (d_{pO} - d_0) & \text{for } d_0 < d_{pO} < 2d_0 \\ 0 & \text{for } d_{pO} > 2d_0 \end{cases} \tag{4.7}$$

where:

$\vec{e}_O$ is the unity vector in the normal direction pointing to the closest point of the obstacle.

A_O is the obstacle interaction strength, which is a balancing parameter between this components and the others.

d_{pO} is the anticipated distance to the (nearest) obstacle.
d_0 is a threshold distance (shy-away distance).

Equation 4.1 and Equation 4.2 represent the basic position dynamics of pedestrian movements in continuous space.

Equation 4.3 describes the components of the acceleration that shapes the movement of an individual, including a controlled component $\vec{a}_c(t)$, an uncontrolled physical component $\vec{a}_p(t)$, and an error term $\vec{\varepsilon}(t)$, which produces the natural fluctuations of pedestrian movements.

The components of controlled acceleration are shown in Equation 4.4, namely the path straying component $\vec{a}_s(t)$, the obstacle interaction component $\vec{a}_O(t)$, and the pedestrian interaction component$\vec{a}_{pq}(t)$.

Equation 4.5 explains the derivation of the the path straying component, which is the deviation from the optimal velocity $(v_0(t) \cdot \vec{e}_g - \vec{v}(t)$ divided by an constant acceleration time τ. The path straying component is defined for the purpose of making pedestrians try to comply to the optimal route, which aligns an optimal velocity at every position. In NOMAD, deviating from the optimal velocity is assumed to increase costs. The acceleration time determines how much the increased cost is and therefore controls the extent to which pedestrians stay close to the optimal velocity. Pedestrians with a very small acceleration time (0.0s) tend to walk closely to their optimal route with speeds around their optimal speeds. It will require very large interaction accelerations (Equation 4.6 and Equation 4.7) to make these pedestrians deviate from their routes.

The calculation of the pedestrian interaction component is illustrated in Equation 4.6. It is modelled by an exponential function that amplifies the close-by accelerations. In NOMAD, getting closer to other pedestrians is assumed to increase costs. This equation represents the interaction purpose of collision avoidance, where

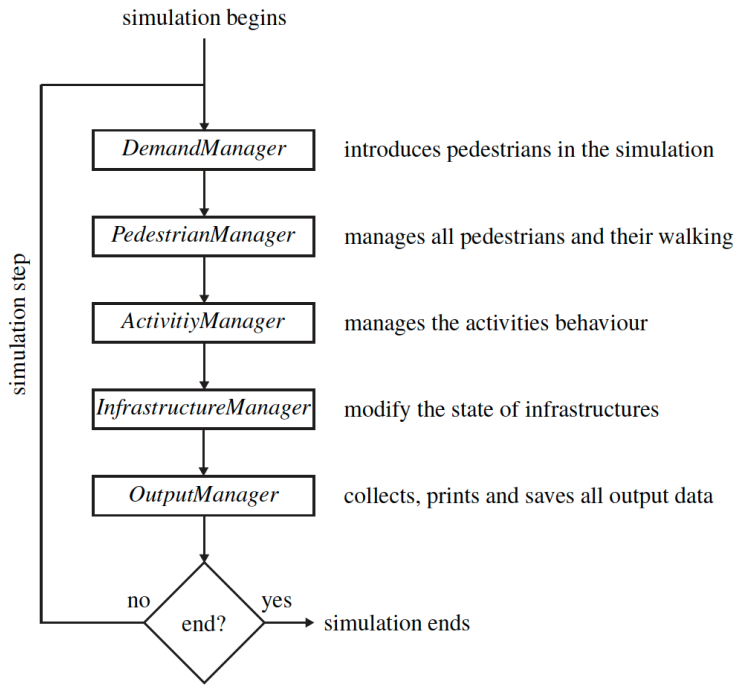


Figure 4.1: NOMAD structure (source: Campanella [2016])

the costs of proximity are inverse to the distance between pedestrians. The closer they are to each other, the more intensively pedestrians want to increase their relative distance. The increased cost is controlled by two parameters, the interaction strength parameter A_0 and the interaction distance parameter d_i. A_0 is the balancing parameter between the pedestrian interacting component and the other components. The larger A_0, the more important this component is in comparison to the others. d_i relates the distance between the pedestrians and the intensity of the interaction acceleration. Diminishing values of d_i (0.0m) diminish the distances required to generate large interaction accelerations. Diminishing d_i decreases the space between pedestrians in crowded situations.

Equation 4.7 describes the calculation of the obstacle interaction component. Similar to the pedestrian interaction component, getting close to obstacles also increase the costs. The strength obstacle interaction is dependent on the distance to the obstacle d_{pO}, the interaction strength of objects in general A_O, and the direction of the nearest obstacle $\vec{e}_O$. The larger A_O, the more important the obstacle interaction is. A step-based approach is used for the definition of interaction strength, where obstacles nearby have a very large influence and obstacles outside the range of influence do not influence individuals' movement dynamics at all. Two distance thresholds, d_0 and $2d_0$ are used to govern the gradual linear decline of the obstacle avoidance force. As a result of the formulation, agents only react to obstacles when they are really close to the obstacle. This is an advantage in the case of the modelling of indoor spaces, where lots of obstacles are present.

4.1.2 NOMAD elements

The NOMAD model is structured as shown in Figure 4.1. In this section, the elements of the NOMAD model are described, namely, infrastructure setting, activity schedule (demand and activity pattern), pedestrian parameters, and output.

Infrastructure setting

In the NOMAD application to the proposed method, the walking costs of individuals are computed within the walkable area of multiple infrastructure settings, in a grid of 0.1m by 0.1m rectangular cells. NOMAD simulates all the risky activity spaces in an event, where 'high-risk contacts' are expected to happen. The 'high-risk contacts' include possible SARS-CoV-2 transmission via three identified routes, namely, droplets, aerosols, and formites. After being identified, the infrastructure layout of each risky activity space will be set up for pedestrian simulation, based on the observed (estimated) infrastructure layout of various activity spaces in large events.

5 elements compose the infrastructure setting of NOMAD, namely, walkable areas, obstacles, destinations, sources, and sinks. The walkable area defines where pedestrians can potentially walk, which is usually the area within the activity space. An obstacle occupies an area that is inaccessible to pedestrians and exerts a repulsive force on the pedestrians. When overlapping with the walkable area, the area taken by an obstacle is removed from the walkable area. Destinations are the locations that define where a pedestrian goes to perform an activity. A source is the location where pedestrians enter the model and a sink is where pedestrians leave the model. Examples are presented during the application of proposed method in Section 5.2.

Activity schedule and demand pattern

To simulate the routing behavior of individuals, the activity schedule within each activity space is predefined according to the infrastructure layout and functions. It determines the order to perform activities within the activity space. The demand pattern of each activity space is determined by infrastructure capacity estimation and the estimated demand of the entire event. In NOMAD, these two elements are set up by a demand manager, which defines how many pedestrians enter the simulation, where and when, and what their activity schedules are.

An example to derive the activity schedule and demand pattern will be given during the application of proposed method in Section 5.2.

Pedestrian parameter

As is discussed in Section 4.1.1, there are 6 essential parameters in the NOMAD model, namely, the optimal speed, $v_0(t)$,the constant acceleration time τ, the pedestrian interaction strength parameter A_0, the pedestrian interaction distance parameter d_i, the obstacle interaction strength A_O, the shy-away distance d_0.

The parameter values may vary under different pedestrian behavior assumptions. In the method application in Section 5.2, the assumptions made by this study and the selected parameter values will be presented.

Output

The output of NOMAD consists of the detailed trajectories of individuals within the simulated area. The agent ids, positions with time stamps compose the output, which will be further processed as the input to the virus spread model. An example of NOMAD output is given in the following table:

4.2 VIRUS SPREAD MODEL

The virus spread model, QVEmod, developed by the ongoing research [Duives et al., 2021], simulates the emissions of viruses from infectious individuals, the landing of viruses in space and time through the environment, and eventually the viruses that

Agent	Time[s]	X[m]	Y[m]
agentooo	34.7	0.48552994	4.50332648
agentooo	34.8	0.578413694	4.458839492
agentooo	34.9	0.671980886	4.408461202
agentooo	35	0.762384928	4.353293035
agentooo	35.1	0.849703234	4.294231629
agentooo	35.2	0.934359791	4.232008425
agentooo	35.3	1.016717364	4.167267956
agentooo	35.4	1.097182305	4.101209456
agentooo	35.5	1.17618545	4.03523484
agentooo	35.6	1.254189064	3.970662051
agentooo	35.7	1.332111589	3.91015545
agentooo	35.8	1.410100045	3.855186426
agentooo	35.9	1.486240578	3.806447296
agentooo	36	1.558026401	3.761731898

Table 4.1: NOMAD output example

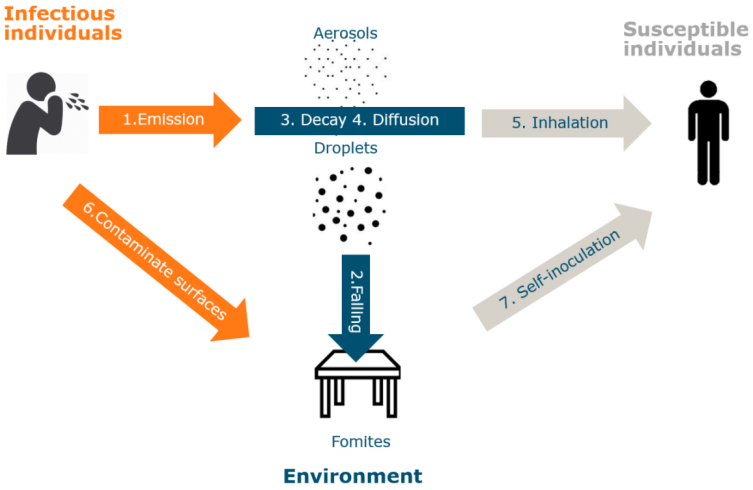


Figure 4.2: QVEmod structure (source: Duives et al. [2021])

get taken up by susceptible individuals. In this section, the structure, input and output of QVEmod are described.

4.2.1 Model structure

The QVEmod model consists of 7 parts, each representing one step of in the virus transmission process. The model structure illustrated in Figure 4.2.

As is discussed in Section 2.2, three SARS-CoV-2 transmission routes are considered, including droplets, aerosols, and fomites. In QVEmod, these three transmission routes are realized in 7 steps:

1. An infectious individual emits virus into the air in the form of virus-laden droplets and aerosols and virus-laden droplets (further referred to as aerosols and droplets, depending on the size).

2. Droplets fall onto surfaces, which creates fomites.

3. Viruses lose their infectivity at a rate that depends on its state in the environment (both in the air and on surfaces).

4. Viruses in the droplets and aerosols travel and diffuse.

5. Susceptible individuals get caught with viruses through inhaling the air with viral-laden droplets and aerosols.
6. The infectious individual contaminates surfaces by touching the surfaces of objects in the space (e.g., tables, benches, and counters).
7. Susceptible individuals get caught with viruses by touching contaminated surfaces (fomites).

Each step is simulated as a submodel, the detailed descriptions of which are provided as follows. In QVEmod, all the submodels are continuous processes and modelled in discrete time and space. The time steps of the following equations are set as small as 1 minute. The space is partitioned into 0.5m x 0.5m cells.

Step 1

Step 1 simulates the process of virus emission. Equation 4.8 describes the total volume of viruses that infectious individuals emit into the air at time t.

$$V_i(t) = V_i(t - \Delta t) + \omega\delta\sigma p_i(1 - FE_i)\Delta t \tag{4.8}$$

where:
i denotes the transmission route (aerosols or droplets).
$V_i(t)$ is the virus volume in aerosols and droplets at time t.
Δt is the time step (1 minute).
ω is the virus emission rate of a typical infectious individual (scaled to 1 per hour).
δ is an activity infectiousness scaler, which scales the heterogeneity in emission rates during different respiratory activities. It can be weighted based on the proportion of time an individual is engaged in a specific behavior.
σ is an individual infectiousness scaler, which scales the individual heterogeneity in emission rates.
p_i is the proportion of viruses emitted in the form of aerosols and droplets. $p_{aerosols}$ and $p_{droplets}$ add up to 1.
FE_i is the filter efficiency of face masks against droplets or aerosols.

As discussed in Section 2.2, infectious individuals emit viral-laden particles (droplets and aerosols) by different respiratory activities such as speaking, coughing, or sneezing, which lead to different virus emissions in the air. Equation 4.8 initially assumes that infectious individuals emit virus at a constant rate ω. The activity scaler, δ, scales the emission rate of different respiratory activities, based on the emission rate under breathing condition. Then, the individual infectiousness scaler, σ, scales different infectiousness of individuals relative to a typical emitter. In addition, the introduction of the face mask filter efficiency, FE_i, enables modeling situations in which people wear face masks.

Step 2

Step 2 simulates the process of viral-laden droplets falling onto surfaces. Equation 4.9 describes the total volume of viruses from droplets that lands on contaminated surfaces at time t.

$$V_{fomites}(t) = V_{fomites}(t - \Delta t) + V_{droplet}(t - \Delta t)u_{droplets}\Delta t \tag{4.9}$$

where:
$u_{droplets}$ is the deposition rate of viral-laden droplets.

As shown in Equation 4.9, the volume of viruses on surfaces is proportioned to the volume of viruses in droplets by $u_{droplets}$. In QVEmod, it is assumed that

high-touch surfaces can obtain viruses from droplets, while low-touch surfaces get landed with viruses only when being touched by contaminated hands. Viruses are assumed to be stationary and evenly distributed within the grid cells (0.5m x 0.5m).

Step 3

Step 3 simulates the process of viruses decaying in the air and on surfaces. Equation 4.10, Equation 4.10, and Equation 4.11 show the calculation of decayed virus volume in three transmission routes.

$$V_{aerosols}(t) = V_{aerosols}(t - \Delta t)e^{-u_{aerosols}\Delta t - ACH\Delta t} \tag{4.10}$$

$$V_{droplets}(t) = V_{droplets}(t - \Delta t)e^{-u_{droplets}\Delta t} \tag{4.11}$$

$$V_{fomites}(t) = V_{fomites}(t - \Delta t)e^{-u_{fomites}\Delta t} \tag{4.12}$$

where:

$u_{aerosols}$, $u_{droplets}$, and $u_{fomites}$ are the decay rate of viruses in aerosols, the deposition rate of viral-laden droplets, and decay rate of virus in fomites.

ACH is the indoor air change rate.

As is displayed in the formula, viruses experience an exponential decay in the environment. In aerosols, viruses lose infectivity with a constant rate while floating in the air, which is influence by the indoor air change rate (ACH). The higher ACH, the faster the viruses in the aerosol reduce. On the contrary, virus-laden droplets deposit rapidly, which justifies the inactivation of viruses in droplets in the air to be neglected. Decay in the droplet layer is therefore driven by the deposition rates. On fomites, viruses decay with a constant rate, depending on the fomite's material.

Step 4

Step 4 simulates the process of virus diffusion in the air. Equation 4.13 explains the diffusion dynamics.

$$V_{(x,y,t)} = D\frac{(V_{(x-\Delta x,y,t-\Delta t)} + V_{(x+\Delta x,y,t-\Delta t)} + V_{(x,y-\Delta y,t-\Delta t)} + V_{(x,y+\Delta y,t-\Delta t)} - 4V_{(x,y,t-\Delta t)})\Delta t}{\Delta x \Delta y} \tag{4.13}$$

where:

Δx and Δy are the length unit of the cell (both 0.5m).

D is the diffusion coefficient per cell size ($0.25m^2$).

In QVEmod, it is assumed that all particles are well mixed in the grid cells (0.5m x 0.5m) and aerosols diffuse in two directions (x and y).

Step 5

Step 5 models the process during which susceptible individuals inhale air with viral-laden droplets and aerosols. Equation 4.14 calculates the virus volume inhaled by susceptible individuals.

$$E_i(t) = V_i(t)\frac{\rho}{L}(1 - FE_i)\Delta t \tag{4.14}$$

where:

$E_i(t)$ is the recipients' exposure to the virus from route i during time step $(t - \Delta t, t)$.

ρ is the inhalation rate.

L is the cell volume.

In QVEmod, the exposure to the virus from aerosols and droplets is represented by the portion of the airborne virus recipients inhaled from the cell where they are. The inhaled virus is then deducted from the environment. As displayed in Equation 4.14, the amount of inhaled viruses is calculated as the ratio of human tidal volume, ρ, at each time step over the cell volume, L, multiplied by the total viruses from route i, $V_i(t)$, the efficiency of inhalation $(1 - FE_i)$, and the time step, Δt.

Step 6

Step 6 models the process of infectious individuals contaminate surfaces by touching them, during which the viruses on their hands are transferred to the surfaces. Equation 4.15 and Equation 4.16 calculates the virus volume on high-touch and low-touch surfaces.

$$V_{fomites-high}(t) = V_{fomites-high}(t - \Delta t) + V_{hand}(t - \Delta t))\gamma\theta\pi\Delta t \tag{4.15}$$

$$V_{fomites-low}(t) = V_{fomites-low}(t - \Delta t) + V_{hand}(t - \Delta t))\theta\pi\Delta t \tag{4.16}$$

where:

γ is the surface touching frequency.

θ is the transfer efficiency.

π is the ratio of finger pads surface relative to the contaminated area.

In QVEmod, two types of surfaces are defined. High-touch surfaces are the type of surfaces that people touch at a constant rate. Low-touch surfaces are only touched when needed. The touching behaviour of low-touch surfaces can be modelled in the NOMAD model. Hence, in Equation 4.16, the virus volume of low-touch surfaces does not include the surface touching frequency, γ.

Step 7

Step 7 models the susceptible individuals' exposure to the virus from fomites. Equation 4.17 and Equation 4.18 calculates the virus volume picked up by susceptible individuals from high-touch and low-touch surfaces.

$$E_{fomites}(t) = V_{fomites-high}(t - \Delta t)\gamma\theta\pi\Delta t \tag{4.17}$$

$$E_{fomites}(t) = V_{fomites-low}(t - \Delta t)\theta\pi\Delta t \tag{4.18}$$

where:

$E_{fomites}(t)$ is the susceptible individuals' exposure to virus from fomites during $(t - \Delta t, t)$ time step.

Viruses from surfaces transfer to hands when people touch the surfaces. Similar to hands contaminating surfaces, the virus transfer from surfaces to hands also varies over two types of surfaces, as shown in Equation 4.17 and Equation 4.18.

4.2.2 Model input

The inputs to QVEmod include surface definition, agent scripts, and related model parameters. These input elements are discussed as follows.

Surface definition

QVEmod requires detailed identification of touchable surfaces, where the process of virus landing, decaying, and being picked up by individuals is simulated. As is discussed in Equation 4.2.1, two types of touchable surfaces can be defined, namely, high-touch surfaces and low-touch surfaces. They are identified according to the infrastructure context of the simulated activity spaces, in the unit of 10cm by 10cm. When one touchable surface is larger than the unit, multiple surface units need to be identified for this one surface. A surface definition script includes the name, position, transfer efficiency, surface ratio (touchable surface/surface unit), touch frequency (for high-touch surface), and the virus decay rate on the surface. An example of one high-touch surface unit definition is illustrated in Figure 4.3.

```
{
    "name": "chair3_eetstoel:3,3",
    "x": 23,
    "y": 21,
    "transfer_efficiency": 0.23,
    "surface_ratio": 1,
    "touch_frequency": 15,
    "surface_decay_rate": 0.193
}
```

Figure 4.3: Touchable surface example

Agent scripts

The agent scripts describe the ids, movements, initial virus loads, and the respiratory activity characteristics of the agents.

The ids and movements are processed from the output of NOMAD simulation, during which the following changes are made.

First, the time stamps are aggregated in 20 seconds, instead of 0.1 seconds. Second, the positions are aggregated in 0.1m unit, and presented as relative displacements from the last time stamps. Third, with an agent leaving the simulation, the action 'leave' is assigned at the last time stamp of this agent's script, so that he/she will no longer be able to emit or pick up virus in the environment.

The initial virus loads define the amount of virus that agents bring into the simulation (every hour), which will be emitted in to the environment via droplets and aerosols. The initial contamination load at hand (of the infectious individual) is defined as a constant value [Duives et al., 2021]. Only one infectious individual is assigned in each scenario simulation.

The respiratory activity characteristics include the inhalation rate, emission rate, and proportion of viruses in the form of aerosols and droplets. Other features of agents, such as if wearing a mask, are also included in the script. An example agent script is given in Figure 4.4. The values used in this research will be presented in the method application section of the next chapter.

Model parameters

As is discussed in Section 4.2.1, QVEmod requires a large number of environmental, personal, and viral parameters in order to simulate the spread of virus in space and on individuals. Existing studies provide reference to most of the parameters. Others vary under different scenario settings. The assumptions made in this study and the corresponding parameter values will be presented during the method application in Section 5.2.

```
{
    "name": "agent002",
    "viral_load": 0,
    "contamination_load_air": 0,
    "contamination_load_droplet": 0,
    "contamination_load_surface": 0,
    "emission_rate_air": 0.252,
    "emission_rate_droplet": 0.748,
    "pick_up_air": 2.34,
    "pick_up_droplet": 2.34,
    "contamination_fraction": 1,
    "script": {
        "1": {
            "type": "enter",
            "x": 70,
            "y": 39,
            "facing": "E"
        },
        "2": {
            "type": "move",
            "x": 55,
            "y": 109,
            "facing": "E"
        },
        "3": {
            "type": "leave"
        }
    },
    "is_active": false,
    "wearing_mask": false
}
```


Figure 4.4: Agent script example

4.2.3 Model output

The output of QVEmod is exported in four tables, three of which include the virus volumes in droplets, in aerosols, and in fomites, in each grid cell in the space at each time step of the simulation. The other table, agent exposure, displays the amount of virus caught by each individual, at each position, and at each time step of the simulation. The virus accumulations obtained from this table are essential for infection risk analysis in Section 5.3.

To support the following infection risk analysis, the agent exposure data is further processed. First, at each type of activity space, an virus exposure distribution among all the susceptible agents is obtained by accumulating the virus exposure at each time step. Second, the period, during which susceptible individuals enter the simulation, are exposed to the possibility of picking up virus from the environment, denoted as t_k (k represents the activity space), is derived for all types of activity spaces. At the same time, the number of agents who are exposed to the possibility of catching the virus is also calculated.

4.2.4 Number of replications

As is introduced in Figure 4.2.2, the virus exposure results are obtained from the QVEmod simulation with processed NOMAD trajectory output. Stochastic processes exist in both models, such as the noise term in the acceleration definition in NOMAD (Equation 4.3) and choice of the infectious individual in QVEmod. These stochastic factors lead to variations in virus exposure result from the same input. To tackle the stochasticity in the simulation results, while maintaining a reasonable number of replications, the following assumption is made:

The variations in results mainly depend on the distance between people and the time they they spend in 'close contact'. When a large group of agents are considered (in the case of this study, hundreds to thousands), it is assumed that one iteration of NOMAD simulation represents an average interaction profile of a certain scenario. In this case, the stochasticity in the results can be represented by assigning a different agent (in the order of entering the simulation) as the infectious individual in each iteration of the simulation scenario.

After each iteration, a virus exposure distribution is obtained. To tackle the stochasticity in the resulting distributions, the two-step approach is applied to the obtained distributions of accumulated virus exposure from a number of iterations to determine if the current number of iterations is sufficient to make sure that the estimate of the measure of performance (MOP) is calculated within a specific confidence interval. This method is one of the two most popular methods to determine the number of simulation runs required to achieve statistically confident results, namely the two-step method and the sequential method [Truong et al., 2015]. Compared to the sequential method, the two-step method is selected as it requires less computational efforts and has produced reliable results in existing studies [Bluman, 2009; Truong et al., 2015].

In the two-step approach, a fixed number of initial runs are performed to estimate the sample mean and standard deviation of a MOP, which are then used to compute the required number of runs according to the following equation [Truong et al., 2015].

$$n = [\frac{z_{\alpha/2}S}{E}]^2 \tag{4.19}$$

where:

n is the number of required runs.

S is the sample standard deviation of the MOP from the initially conducted runs.

E is the desired margin of error (sample mean of the MOP from the initially conducted runs x desired percentage error).

$z_{\alpha/2}$ is the critical value of the normal distribution for $\alpha/2$ (desired confidence level equal to $100(1-\alpha)$).

In this study, the MOPs are the mean and the standard deviation of each distribution. The most critical value, the highest number of runs required by two MOPs, will determine the required number of runs. If the calculated n is smaller than the number of initially conducted runs, it is assumed that the current runs are sufficient for providing an acute estimate of the distribution. If the calculated n is larger than the number of initially conducted runs, more iterations need be be performed until the calculated n is smaller that the number of conducted runs. Then the iteration with the closest average exposure and standard deviation to the mean and standard deviation of all iterations will be used for further virus spread analysis.

4.3 INFECTION RISK ESTIMATION

In order to derive the general infection risk at a large event from the virus exposure at different activity spaces at the event, a number of procedures need to be conducted, as illustrate in Figure 4.5. This section presents the detailed descriptions of each step in this proposed infection risk estimation method.

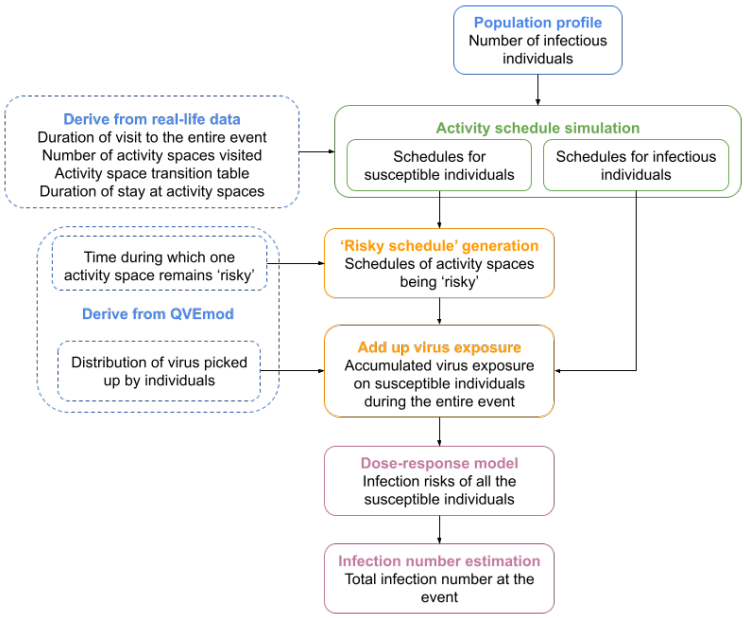


Figure 4.5: Infection risk estimation

4.3.1 Population profile estimation

First, the population profile, i.e., the number (proportion) of virus spreaders at the event, is calculated based on governmental statistics and existing research according to Equation 4.20.

$$N_{in} = p_{in} \cdot t_{in} \cdot N \tag{4.20}$$

where:

N_{in} is the number of infectious visitors in the event.

p_{in} is the average number of confirmed cases per 100,000 inhabitants per day.

t_{in} is the average duration of infectious individuals staying infectious (the average of the incubation period distribution).

N is the total number of visitors to the event.

The official COVID-19 infection statistics from the Dutch government [RIVM, 2021a] displays the historical data of the average number of confirmed cases per 100,000 inhabitants per day, denoted as p_{in}.

To calculate the number of infectious individuals in the event, denoted as N_{in}, a parameter t_{in} describing the average duration of people staying infectious is required. According to the study by Lauer et al. [2020], the average incubation period of COVID-19 (the period from getting infected with SARS-CoV-2 till symptoms develop) is 5.1 days. In this thesis, it is assumed that people who develop symptoms stay at home and do not come to the event. Therefore, the infectious individuals at the event are people who are experiencing the incubation period. Considering people being at different stages of the incubation period, the average stage of the incubation period that visitors are at is assumed to be the average of 0 to 5.1 days. Therefore, t_{in} is defined as half of the average incubation period, as 2.55 days.

4.3.2 Activity schedule simulation

After obtaining the number of infectious individuals at the event, an activity schedule simulation is conducted to obtain when and where the infectious people are throughout the entire event. Following displays the required information of activity pattern for this simulation.

1. Start and end time of the event.
2. Types and numbers of activity spaces at the event.
3. T, the duration of visit to the entire event.
4. n_a, the number of activity spaces visited during the event by one visitor.
5. The activity space transition table, displaying probabilities of visiting certain activity spaces after leaving certain activity spaces.
6. The duration of stay at each type of activity space.

To simulate the activity schedule of an infectious individual, the duration of visit to the event, T, and number of activity spaces visited, n_a, are randomly generated from the derived distributions from the activity pattern analysis. Then the order of visits to activity spaces is generated according to the transition table and the types and numbers of activity spaces at the event. Next, the time spent at each activity space, denoted as t_i is assigned to the activity spaces in the visit order. The time spent walking from one activity space to another, denoted as t_w, is assumed to be equal between different activity spaces, which is calculated as shown in Equation 4.21.

$$t_w = \frac{T - \sum_{i \epsilon n_a} t_i}{n_a - 1} \tag{4.21}$$

The starting time of the activity schedule is randomly selected between the earliest time possible and the latest time possible, given the selected T. By fitting in t_i and t_w to the order of visits, an activity schedule is generated.

To give an example, assuming at one event, there are 2 entrance gates, 10 food stands, 2 restrooms, and 2 exit gates. The event is held from 11:00 to 14:00. For one agent's schedule, T is drawn a visit duration distribution as 50 minutes; n_a is drawn as 4. According to a transition table, the probability to go to a food stand is 80% and a restroom 20%; after a food stand, the probability to go to another food stand, a restroom, or the exit gate is 50%, 20%, 30%; after a restroom, the probability to go to a food stand or the exit gate is 80%, 20%;. Assuming that a visit always starts with the entrance gate and ends with the exit gate, the agent's schedule may be generated as 'entrance 1-food stand 4-restroom 2-exit 2'. Considering the duration of the event, the entry time is selected as 12:00 for this agent. Assuming the duration of stay at all the activity spaces, t_i, is drawn as 5 minutes. Then t_w is calculated as 10 minutes. The agent's full schedule is generated as follows:

Entrance 1	12:00-12:05
Food stand 4	12:15-12:20
Restroom 2	12:30-12:35
Exit 2	12:45-13:50

4.3.3 Identification of the location and duration of possible virus exposure

From the previous step, the information of when and where the infectious individuals are staying is obtained. From the virus spread simulation results, the time from one infectious individual enters the activity space until the viruses in the environment decays completely (no one can pick up any virus from the environment), denoted as t_k (k represents the activity space), is derived.

To simulate how the virus spreads among all the event visitors, identification of the location and duration of possible virus exposure (i.e., where and when susceptible individuals are exposed to the possibility of catching the virus) at event activity

spaces is required. The location and duration are generated by fitting t_k into the activity schedules of infectious individuals (from the time when the infectious individual enters to the next t_k minutes), creating a number of time periods with exposure possibility at every activity space.

An example of the 'Risky' schedule will be given during the method application in the next chapter.

4.3.4 Accumulate virus contamination

In this step, the accumulated virus contamination on every susceptible individual is added up. First, the activity schedules of all the susceptible individuals ($N - N_{in}$) are generated in the same way as described in Section 4.3.2. Second, these activity schedules are matched with the 'risky' schedule one by one, to detect to what extent the susceptible individuals are exposed to the possibility of catching virus during the entire visit to the event. The match exists when one agent enters the activity space during its 'risky' time period. Then, out of every match of every susceptible individual, the amount of virus he/she picks up at the exact activity space, is randomly selected from the output of QVEmod, obtained by Section 4.2.3, denoted as V_{rij}. i represents the activity space and j represents the susceptible individual. At last, V_{rij} are added up together for each susceptible individual j and the accumulated virus contamination distribution throughout the entire event is created.The accumulated virus load on this susceptible individual during the entire visit to this event, V_aj is calculated in Equation 4.22.

$$V_{aj} = \sum_{i \epsilon I} V_{aij} \tag{4.22}$$

where:

I is the set of risky activity spaces visited by the susceptible individual.

4.3.5 Translation from virus dose to infection risk

The risk of being infected under exposure to a certain number of viral particles varies between different transmission routes, because of different deposition locations (hands, lower or upper respiratory tract) and the viability of the virus [Deng et al., 2020]. The relationship of infection probability and accumulated virus doses from different transmission routes is modeled by an exponential dose-response relationship in Equation 4.23, [Nicas, 1996].

$$P = \ 1 - e^{-(\frac{E_{aerosols}}{k_{aerosols}} + \frac{E_{droplets}}{k_{droplets}} + \frac{E_{fomites}}{k_{fomites}})} \tag{4.23}$$

where:

P is the probability of getting infected.

$E_{aerosols}$, $E_{droplets}$, and $E_{fomites}$ are the accumulated exposure (virus doses) from the three transmission routes.

$k_{aerosols}$, $k_{droplets}$, and $k_{fomites}$ are the route specific exposure at which individuals have 63% chance of getting infected, which is calculated according to Equation 4.24.

$$k_{route} = \frac{D_{inf}}{\phi * c_{route}} \tag{4.24}$$

where:

D_{inf} is the infection virus dose.

ϕ is the emission rate of an average individual.

c_{route} is the proportion of viral particles caught by humans that reaches the respiratory tract cells.

By fitting the virus dose distribution into the dose-response relationship, Equation 4.23, the probability distribution of all the susceptible individuals getting infected during this event is obtained. The number of infected individuals can be calculated from this distribution as well, following a random simulation method [Duives et al., 2021]. Instead of directly deriving the infection number from the average infection probability, this random simulation method is selected to make sure the high and low end of the distribution have limited contribution to the infection number calculation. It is expected that a large number of agents with very low infection probabilities will be resulting from the previous steps. To limit their impact on the estimated infection number, the selected method is used as described below:

1. Draw a random number from a uniform distribution [0,1].
2. Compare the drawn number to the individual's infection probability. This individual is assumed to be infect if the infection probability is larger than the number drawn.
3. Repeat Step 1 to 2 for all the simulated individuals.
4. Divide the total number of infections by the number of simulated individuals and obtain the possibility of being infected during this event.

4.3.6 Number of replications

As stochasticity is involved in the process of activity schedule generation and selection of virus exposure from the distribution, replications of the above steps should be conducted to make sure the ultimate output, number of infections, is representative. A two-step method is used to tackle stochasticity [Truong et al., 2015], by which the MOP is the number of infections.

5 CASE STUDY

This chapter presents two case studies conducted during this research to support the application of the proposed method and apply the proposed method.

In Section 5.1, the real-life data collected from a GPS-tracing application during a musical festival in 2019 is analyzed, from which a general profile of activity pattern is derived. The obtained activity pattern is then fed to Section 5.2 to generate activity schedules of visitors during the simulated event. The methodology developed in the previous chapter is then applied to a music festival setting, with specified scenario parameters values and assumptions.

5.1 DERIVE ACTIVITY PATTERN FROM GPS DATA

This section describes the methods used to collect, process, and analyze a real-life GPS data set. Initially, it was expected that this data set provides actual pedestrian interaction information that can be directly used for the dose response model. However, during this case study, it turns out that due to the incomplete data set and limited accuracy, the collected traces do not provide a full picture of event visitors' interaction characteristics. Moreover, the data collection dates back to pre-COVID-19 time, which provides no indication to pedestrian behavior under COVID-19 regulations such as social distancing. Therefore, the detailed pedestrian interaction characteristics will be fully dependent on pedestrian simulation. The purpose of this data analysis lies in deriving the activity pattern of visitors at this event, which will be used to set up the simulation model.

The case study is carried out in the following method. First, the GPS data is gathered and processed in Section 5.1.1. After being matched on the map, the pedestrians' duration of stay at different event spaces and the activity scheduling characteristics are analyzed in Section 5.1.2 and Section 5.1.3. A conclusion is drawn in Section 5.1.4.

5.1.1 Data collection and processing

The real-life data of this research was collected during the Amsterdam Open Air musical festival, which was held in Gaasperpark, Amsterdam, on June 1st and 2nd, 2019. During this two-day event, about 20000 people participated. They were offered with a mobile phone application, named Woov, to explore activity information and communicate with friends. Each time when a participant used this application, the GPS location of his or her mobile phone was collected, together with the timestamp, GPS accuracy and id of the device.

The raw data set used in this research consists of 104866 location stamps of 2688 devices. In the form of an excel sheet, 4 columns of data include the time stamp, GPS location, encrypted id, and accuracy. Each time stamp consists of date and time information, in the form of dd-mm-yyyy, hh:mm:ss. The GPS location data is coded by PostGIS, in the form of a long string with numbers and letters. The encrypted ids are short strings composed of numbers and letters. The accuracy is measured in meters, ranging from 0 to above 100 meters.

To make use of the collected data, the following data processing steps are taken, as illustrated in Figure 5.1.

1. First, the GPS location data is decoded into Well-known text (WKT) forms by PostGIS, resulting in two columns of latitude and longitude values.

2. All the location stamps are scatteredly mapped on the open street map [OpenStreetMap, 2021], which provides a detailed infrastructure background of Gaasperpark. It is observed that a part of the GPS traces is located outside the event area, which is due to the use of the Woov app before entering and after leaving the event location. Based on the location information of Gaasperpark on Google maps [**?**], a rough geographical area of the musical festival is determined. Then the data with location stamps outside the event area is filtered out, as the scope of this study is limited to the behavior at the event, excluding the commuting to and from the event.

3. Considering the large event area of this festival, it is assumed that location stamps with the accuracy smaller than or equal to 30 meters provide sufficient information of the activity areas where people stay. Therefore, the data with accuracy larger than 30 meters is filtered out.

4. The repeated data rows are deleted.

5. The encrypted ids are replaced with numbers for the convenience of further data analysis.

6. The data before 01-Jun-2019 06:00:00 and after 03-Jun-2019 06:00:00 is filtered out to make sure traces recorded before and after this event are not included. Next, the data is split into 2 two data sets, each including traces recorded in one day of the event, divided at 02-Jun-2019 06:00:00, for the convenience of further data analysis.

7. It is observed that the usage of the Woov app varies significantly among users. Some of them provide continuous data with short time intervals and long duration. However, the data of a large number of users consists of only several scattered time stamps throughout the day, from which no clue of their activity patterns can be derived. To make the best use of the limited filtered data, it is assumed that data with a time interval smaller than 15 minutes would provide sufficient information of the duration of stay at a certain activity area and the movement among different areas. Thus, for each unique id, on each day, the data rows that create a time interval larger than 15 minutes from the previous data row are deleted.

Finally, the processed data set of 01-Jun-2019 is composed of 20186 location stamps of 981 devices, the data set of 02-Jun-2019, 17266 location stamps of 916 devices.

5.1.2 Duration analysis

As is observed during data processing, the real-life data merely consists of a small part of traces of 1/10 of the visitors to this event. Although the small sample size is not sufficient for pedestrian interaction analysis, it enables deriving visitors' activity patterns, including how long they stay at a certain activity space (Section 5.1.2) and what the order of visiting activity spaces is (Section 5.1.3). By analyzing these patterns, the pedestrian characteristics as input to the detailed simulation can be determined.

To analyze people's duration of stay at different types of activity spaces, the following steps are taken.

1. First, the activity spaces are defined based on clusters of location stamp mapping on the open street map [OpenStreetMap, 2021] and an Amsterdam Open Air festival map [**?**], as shown in Figure 5.2 and Figure 5.3. It is observed

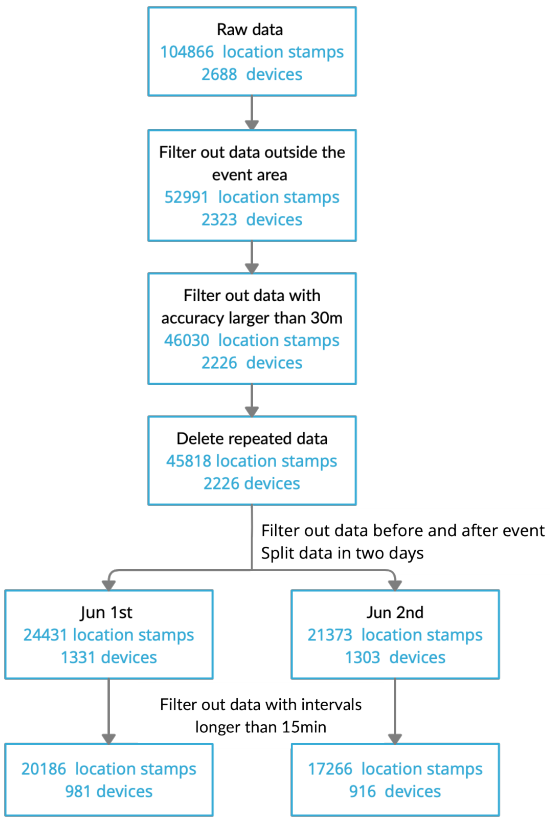


Figure 5.1: Data processing steps

that some activity spaces are not not easy to distinguish by comparing the clusters and the festival map, such as toilets and bars next to musical stages. Therefore, only those activity spaces that can be distinguished are defined, resulting in 5 musical stages, 3 bars (including food stands), 2 toilets, and 1 locker area. These spaces are defined by selecting multiple points on their edges and exporting the location information of the points.

2. After defining activity spaces, the data with location stamps that fall into each space is obtained. The same as data processing steps, the data within one activity space whose time interval is larger than 15 minutes is filtered out. It is important to be aware that the activity space definition method does not provide very accurate spatial information. People clustering around the edges or walking on the road next to these spaces might also be considered as staying in these spaces, which will result in more data of short duration being included. Therefore, it is assumed that the duration of stay shorter than 3 minutes at musical stages and shorter than 1 minute at bars, toilets, and lockers is generated by people moving in the vicinity and not preforming any activity in the activity space. The corresponding data is then deleted from the data set.

3. Next, the summation of continuous time intervals of each id within each activity space is calculated, which represents the (minimum) time that that person spent in that space. The statistics of the (minimum) duration of stay at each type of activity space is summarized in Table 5.1. Figure 5.4 shows the histogram of the (minimum) duration of stay at each type of activity space. One outlier of 01:36:49 at the activity space type of stage is too large to be included in Figure 5.4.

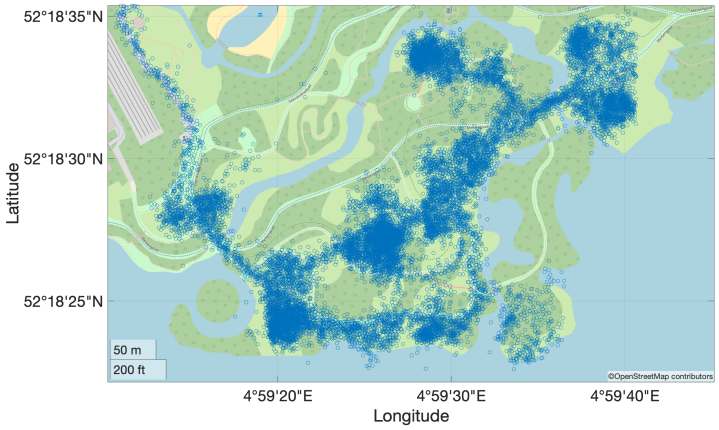


Figure 5.2: Scattered location stamps of filtered Woov data on 01-Jun-2019 on open street map (Source: OpenStreetMap [2021])

Figure 5.3: Amsterdam Open Air 2019 festival map (Source: ?)

As is shown in Table 5.1, people spent on average 11 minutes 32 seconds in front of the musical stages, 6 minutes 51 seconds at bar and food stands, 3 minutes 41 seconds at lockers and 3 minutes 21 seconds in toilets. The medians of duration is slightly shorter than the mean values for all the activity spaces. The variations exist in the duration ranges, as the standard deviations are relatively large. According to the distribution illustrated in Figure 5.4, for the activity space type of bar and toilet, the number of people staying in the area decreases nearly linearly with the increase of duration. At bars, the largest number of people are observed from 00:01:00 to 00:03:00, same for toilets. At around 00:18:00 to 00:21:00, the number drops to around 0 for bars and around 00:06:00 to 00:09:00 for toilets. For musical stages and lockers, the number remains steady for several bins and drops significantly afterwards. Most people stay for 00:01:00 to 00:06:00 at lockers and for 00:03:00 to 00:15:00 at stages.

It is important to be aware that the values shown in Table 5.1 and Figure 5.4 are calculated from the time stamps when people continuously stay in each activity space, which may not include the time when they enter and leave. Therefore, the results may slightly underestimate the actual duration of stay at each type of activity space.

5.1.3 Activity pattern analysis

As is mentioned above, not all the activity spaces can be defined to track the exact locations where people have been. However, it is still important to study the activity pattern on the aspect of the order to visit different types of event spaces based on

Activity space	Stage	Bar	Locker	Toilet
Mean	00:11:32	00:06:51	00:03:41	00:03:21
Median	00:10:06	00:05:03	00:03:15	00:02:15
Maximum	01:36:49	00:33:03	00:13:05	00:25:30
Standard deviation	00:08:09	00:05:52	00:02:30	00:03:46

Table 5.1: Duration of stay at activity spaces

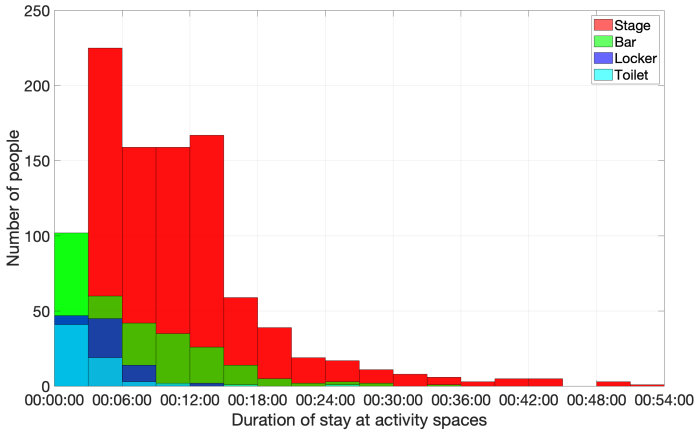


Figure 5.4: Duration of stay histogram of 4 types of activity spaces

available data. Together with the duration analysis, a full picture of visitors' activity pattern will be revealed.

To derive where visitors go after a certain type of activity space from the bulk data, the following steps are taken.

1. The metrics of data, the locations of which fall within the same type of activity space, are combined, resulting in four metrics, namely data located at musical stages, bars, toilets, and lockers, respectively including 12400, 2423, 352, and 560 rows of data.

2. For each metrix, the intervals between time stamps of the same device are calculated. For an interval larger than 15 minutes, it is assumed that this person did not stay in this activity space and went to another activity space. Considering the time it takes to reach another area, for instance, going from one stage at the west end to another stage at the east end, it is assumed that where a person is detected within the next 30 minutes (from the time stamp that has an interval larger than 15 minutes compared to the next one in the same metrix) is considered as the "next" activity space he or she visits after leaving the previous area. Therefore, for each metrix, the numbers of devices detected within 30 minutes in four metrics are calculated.

3. Then the detected device numbers of four types of activity spaces are summarized in Table 5.2, together with the percentage of each origin (previous location) and destination (next location in the next 30 minutes) pair.

As is observed from Table 5.2, the majority of people go to musical stages or bars after leaving all types of activity spaces. A small number of people are detected for visiting toilets and lockers after leaving stages and bars. Most of the derived information falls into expectation. For instance, it is observed that visitors pay more frequent visits to stages and bars than other areas. As a music festival, the main purpose of attending this event is anticipated to be enjoying the music performances. Meanwhile, people also frequently grab drinks from bars. For visits to toilets, it is expected to be fewer compared to the frequency of people switching from stage

Activity space	Stage		Bar		Toilet		Locker	
Stage	198	42%	210	44.5%	54	11.5%	10	2%
Bar	93	60%	38	25%	16	10%	7	5%
Toilet	13	37%	20	57%	2	6%	/	
Locker	17	61%	9	32%	2	7%	/	

Table 5.2: Transition matrix of observed activity pattern derived from bulk data (activity spaces in the left column represent the origins, activity spaces in the top row represent the destinations)

to stage and from stage to bar. The fewest visits are expected for the locker, as normally people only visit it when they left an important belonging at the locker.

However, some other derived information rises doubts. For instance, 25% of people leaving a bar are observed to visit a bar again and 6% of people leaving a toilet are observed to visit a toilet again, which is not likely in happen in practice. The reason that it is derived may be the long interval assumed for walking from one area to another and the imprecise definition of activity spaces. Another problem that hinders the reveal of activity patterns from Table 5.2 is the poor sample size. It is surprising that such small numbers of people moving from one activity space to another are summarized from relatively large data sets, which to some extent implies how scattered the data is.

Besides the above-mentioned problems, the following limitations of the data and analysis methods also make it difficult to make use of the bulk data to derive people's activity pattern. First, most of the data only provides short-interval location information of a short time period, for instance, within 30 minutes, which is too fragmented to be used for activity pattern analysis. Second, the limited GPS accuracy and imprecise definition of activity spaces do not allow categorizing all types of spaces. For instance, it is relatively difficult to distinguish small bars and toilets close to musical stages, compared to larger bar areas 20 meters away from other stages. Therefore, to provide an accurate activity pattern analysis, it is important to analyze individual traces with long duration and short intervals.

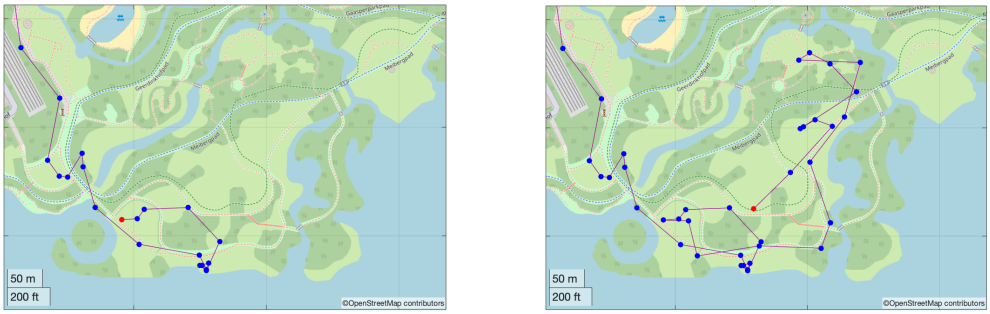
In the following steps, these traces are selected, mapped, and compared with the festival map (Figure 5.3), to derive the visiting order of different activity spaces.

1. First, the location data of the same device with more than 120 time stamps in one day of the event is selected. It is tested that all the time intervals during the GPS tracking are smaller than 15 minutes.

2. Then the processed data of 19 traces on 01-Jun-2019 and 16 traces on 02-Jun-2019 are mapped on the open street map [OpenStreetMap, 2021] to examine if all the traces cover the complete visit to the event, which is indicated by whether the traces start from entering the park and end at leaving the park through an entrance. After filtering out incomplete traces, 6 traces on 01-Jun-2019 and 11 traces on 02-Jun-2019 are used for deriving rough trajectories of individuals throughout the day.

3. For each individual, the rough trajectory is mapped time stamp by time stamp on the open street map [OpenStreetMap, 2021]. By comparing the festival map (Figure 5.3) with trajectory mapping shown in Figure 5.5, the activity pattern of each visitor is determined.

4. The results of all the traces are summarized, which compose a transition matrix of the probability of the type of activity space to be visited (destination) after visiting a certain type of activity space (origin), as shown in Table 5.3. A transition diagram visualizing the observed activity pattern is shown in Figure 5.6. The grey circles represent all types of activity spaces visited before an individual leaves the event.



(a) Trajectory at time stamp 20

(b) Trajectory at time stamp 40

Figure 5.5: Trajectory mapping

Activity space	Entrance	Locker	Stage	Bar	Toilet
Entrance	/	100%	/	/	/
Locker (not before leaving)	/	/	52%	41%	8%
Stage	/	13%	53%	23%	11%
Bar	/	11%	81%	/	9%
Toilet	/	13%	78%	9%	/
Locker (before leaving)	100%	/	/	/	/

Table 5.3: Transition matrix of observed activity pattern derived from individual traces (activity spaces in the left column represent the origins, activity spaces in the top row represent the destinations)

From the individual-based activity analysis, other characteristics are derived as follows.

1. An individual visits on from 9 to 26 (average: 17.5) activity spaces during this musical festival. The average number of visits paid to each type of activity space by one individual is listed in **??**.

Activity space	Average number of visits
Entrance/exit	2
Locker	2.6
Bar	2.8
Toilet	1.4
Music stage	8.7
Total	17.5

Table 5.4: Average number of visits at different activity spaces

2. The duration of the visit ranges from 7 hours 6 minutes to 10 hours 11 minutes, with an average of 8 hours 50 minutes.

The observed activity pattern provides insights to the actual choice behavior of people at the event and will be used as the reference to input of simulations.

5.1.4 Conclusion

From this case study, 4 important activity characteristics are derived as follows.

1. The duration of visit to the event, derived from analyzing individual trajectories, average 8 hours 50 minutes, ranging from 7 hours to 10 hours.

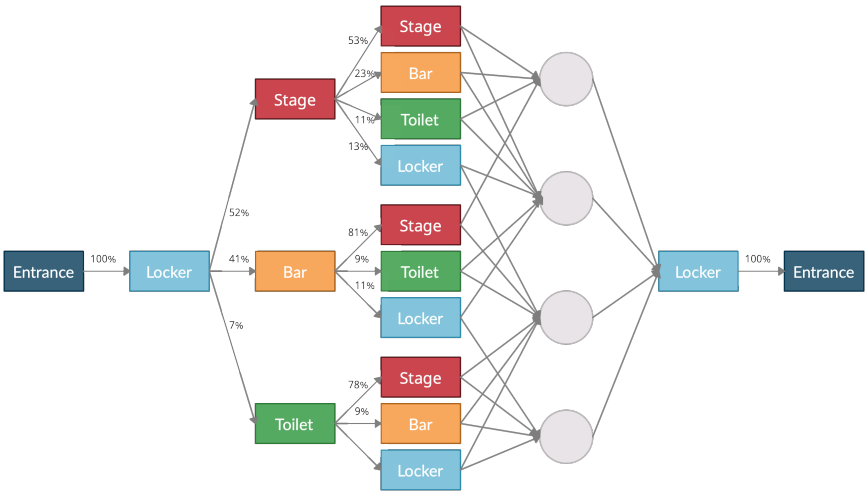


Figure 5.6: Transition diagram of observed activity pattern (grey circles represent all types of activity spaces visited before leaving the event)

2. The distribution of number of activity spaces visited during the event by one visitor, derived from analyzing individual trajectories, ranging from 9 to 26.
3. The activity space transition table, Table 5.3, displaying probabilities of visiting certain activity spaces after leaving certain activity spaces, derived from analyzing all the documented trajectories.
4. The distribution of duration of stay of individuals at each type of activity space, derived from analyzing all the documented trajectories and summarized in Table 5.1.

These characteristics will be used to generate the contact probability distribution in Section 5.2 according to the proposed method in Section 2.3.1.

5.2 METHOD APPLICATION

In this section, the developed method to identify SARS-CoV-2 transmission risk from pedestrian interactions is applied in a simulated event, reproducing the Amsterdam Open Air music festival. The simulated festival is held for one day from 12:00 to 23:00 with 10000 participants. Within the event terrain, there are 7 different types of activity spaces where the virus can be transmitted between people, namely, entrance/exit queue, locker, big bar with sitting areas, small bar without sitting areas, normal indoor toilet, portable toilet and the queue, music stages. The number of each type of activity spaces will be calculated based on the demand. The number of infectious individuals at this event will be determined according to the infection data from the government. The assumptions and inputs to the models described in Chapter 4 will be specified in this section, following the structure of the proposed methods. The purpose of this method application is to estimate the SARS-CoV-2 transmission profile at the reproduced Amsterdam Open Air music festival and analyze the potential risk at different types of facilities.

5.2.1 NOMAD model application

As introduced in Section 4.1.2, to apply the NOMAD model, definition of 4 elements is required, namely, infrastructure, activity schedule, demand pattern, and pedestrian parameters. In this subsection, these elements are defined under the simulated scenarios.

NOMAD infrastructure

In the application to the Amsterdam Open Air music festival, 5 types of risky activity spaces are defined for pedestrian simulation, namely, entrances/exits, lockers, bars, toilets, and music stages. Every activity space is independently simulated in NOMAD. These activity spaces are selected as 'high-risk contacts' are expected there, according to the high possibilities of three SARS-CoV-2 transmission routes to occur. In a music festival, a dense crowd usually accumulates in front of entrances/exits, bars, and toilets, when people form queues to get in the activity space. The largest crowds are usually gathered in front of music stages where the density can grow extremely high ($> 2P/m^2$) [Weppner and Lukowicz, 2013]. Even when under social distancing rules, people are expected to get close to each other either in the rush to reach the end of the queue or influenced by the live music atmosphere [FieldlabEvenementen, 2021], which can lead to droplets and aerosols transmission in both indoor and outdoor spaces. In places like lockers and bars, it is inevitable for visitors to touch the locker door when depositing personal belongings, touch the POS terminal when paying for a drink at a bar. Via these behaviors, virus transmission via fomites may occur. As for inside the toilets, though people are expected to disinfect themselves when washing their hands, aerosols transmission still plays an important role.

Other activity spaces within the park are not simulated in this research, including routes connecting different activity spaces and vast free areas where people can sit on the grass for a rest. These places are considered not risky, because the transmission risks in these places are assumed to be negligible for the following reasons. First, there is almost no touchable surfaces which different people would constantly touch during the visit, which nearly eliminates the possibility of fomites transmission. Second, the space these places are expected to be large enough for people to keep a relatively large social distance when walking around. With the natural outdoor ventilation, it is assumed that the possibilities of droplets and aerosols transmission are negligible.

The simulated activity spaces are described and visualized in the following paragraphs. The setup and dimensions are estimated from the music festival map Figure 5.3, an YouTube vlog shot during this this event [Huijkman, 2019], and daily life experience. For some activity spaces, different infrastructure layouts and settings are designed to explore the sensitivity of their influences on the infection risk profile. In the following figures, the grey areas within the boundaries represent walkable areas, the dark grey areas are obstacles, and the meshed areas are destinations. The measurement unit is 1 meter.

1. *Entrance/exit*

   In Figure 5.7, the queue layout at the entrance/exit is displayed. The design follows the efficient 's' shape, which is often seen at the entrance/exit of large events, such as football matches and music festivals. In a 20m by 20m area, queuing points are distributed according to the assigned distance between people. The destination placed at the top left corner is the entrance/exit gate, which is assumed to open when visitors scan the digital ticket. As a result, direct touching behavior is avoid and fomites transmission can be eliminated in this facility. To explore the influence of the queuing distance, two scenarios with 1 meter and 1.5 meter queuing distance between individuals are designed. In both scenarios, the demand will be calculated to make sure that there are always enough spots for individuals to take up during the queuing process. To guarantee one-way traffic, the entry is placed on the right bottom and exit on the left top.

2. *Lockers*

   Figure 5.8 displays the locker infrastructure layout, which is inspired from the ubiquitous locker design shown in Figure 5.9. 4 rows of lockers, each 6

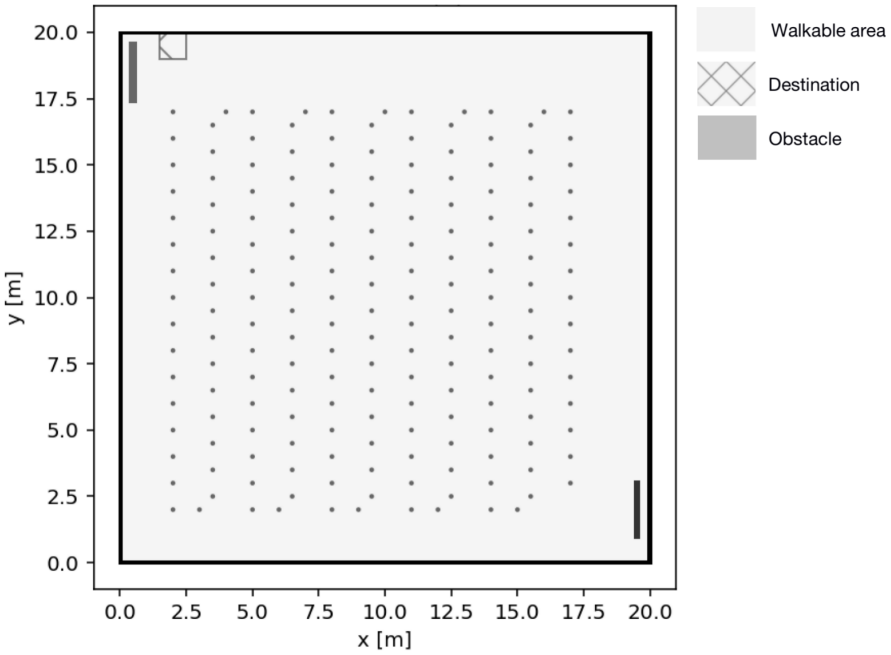


Figure 5.7: Queue layout at the entrance/exit/small bar

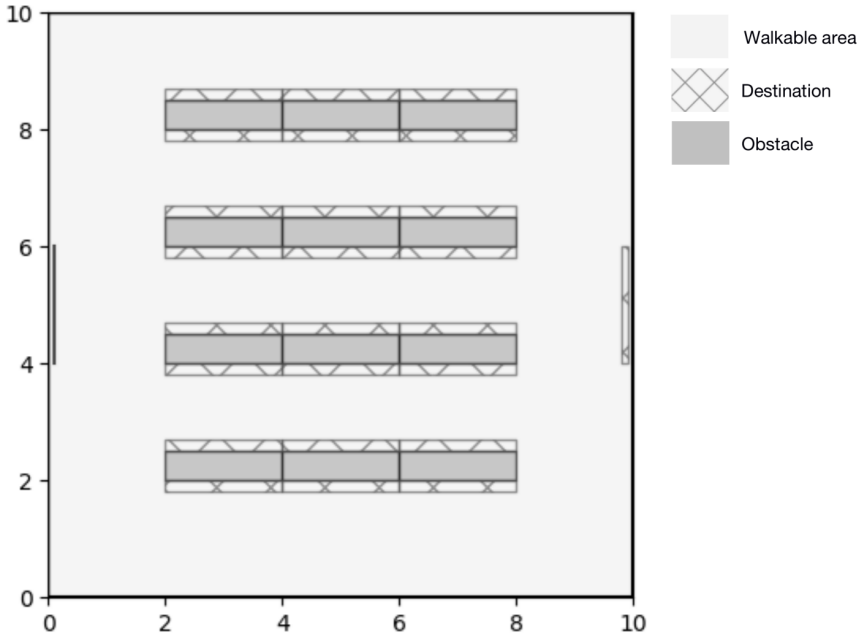


Figure 5.8: Locker infrastructure layout

Figure 5.9: A ubiquitous locker design (Source: Lockerlogic [2018])

meters long, 0.5 meters wide, are depicted as obstacles. The 0.3-meter wide and 2-meter long destinations around the lockers represent where people aim to visit during the simulation. To apply the social distancing regulation, 3 destinations on each side of the lockers are defined, each can only be taken up by one person at a time. Therefore, the maximum capacity of this locker area is 24 people at a time. To guarantee one-way traffic, the entry set is on the left middle and exit on the right middle.

3. *Bars*

   Two types of bar infrastructure are designed according to the music festival map Figure 5.3 and the YouTube vlog of Amsterdam Open Air festival 2019 by Huijkman [2019]. One is a big enclosed bar area as illustrated in **??**. The other is a small bar where no tables or benches are placed. In the big bar area, 6 bar counters lie on the right, offering drinks and snacks to visitors coming by. The 1m x 0.3m area in front of each counter is set as the visitors' destination. 14 tables and 28 benches are situated on the left and in the middle. The tables are 2 meters long and 1 meter wide, defined as obstacles. The benches around them are destinations and each bench can only be seated by one individual, in order to promote social distance. Therefore, the maximum number of people sitting in the entire bar area is 28 at a time. People are randomly assigned to sit at the table or leave directly after visiting the bar counter, the proportion of which will be calculated in the next section. One-way traffic rule is applied in this area by locating the entry at the bottom left and the exit on the top left. The design variable of the big bar infrastructure is indoor/outdoor setting.

   A small bar area only involves one bar counter. Visitors are expected to leave to other facilities after visiting the bar counter. In real-life practice, it is often observed that people form a queue in front of the bar counter, which increases the risk of transmission by droplets and aerosols. The layout of the queue is designed as the same as the queue at the entrance/exit, as displayed in Figure 5.7. In this case, the destination at the top left corner is the bar counter, where the action of paying at the POS terminal occurs. The design variables of the small bar infrastructure include indoor/outdoor setting and the queue distance. Same as the entrance/exit queue, 1 meter and 1.5 meters queuing distance are designed, which are expected to influence the scale of virus spread in the crowd.

4. *Toilets*

   Two types of toilets are considered in this study. One is a normal toilet which has been originally built in the park where the festival is held. The other is

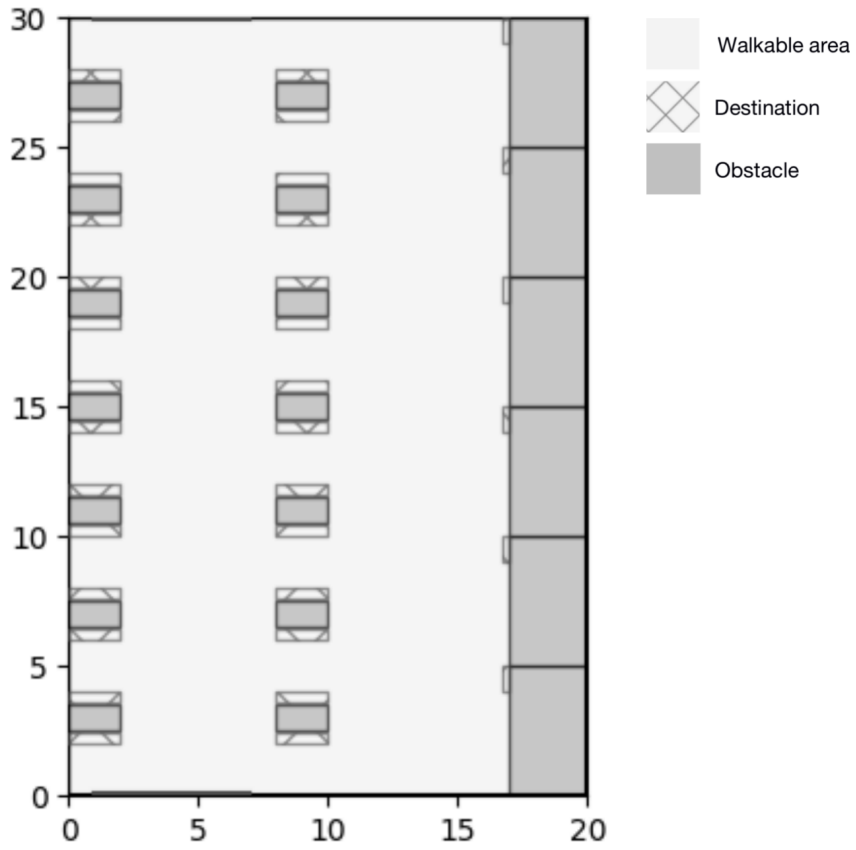


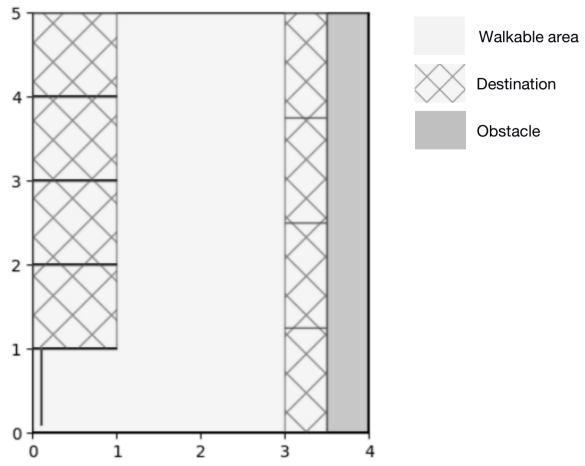
Figure 5.10: Big bar infrastructure layout

portable toilets which are brought to the event location by the organizers to satisfy the need of a large number of visitors. The normal toilet infrastructure layout is demonstrated in Figure 5.11a, inspired by everyday life experiences. Four toilets (1m x 1m) are situated on the left and 4 basins on the right. The entry and exit are through the same door, located on the bottom left corner. The normal toilet is simulated indoor. There is no other scenario at this facility.

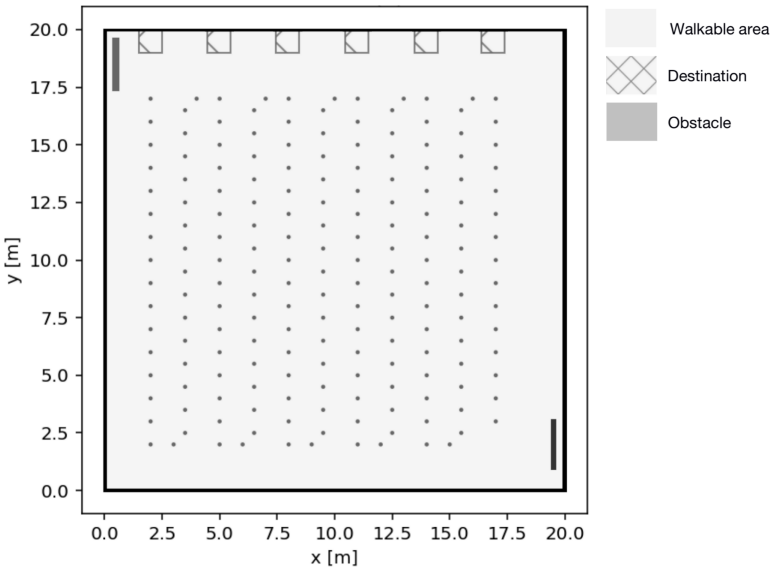
For portable toilets, as observed in everyday life experiences, they usually stand in a line in the field, where people form a queue in front. Figure 5.12 displays portable toilets lined up in a real life application, which is often observed in large temporary events. To simulate the virus spread situation in the scenario of portable toilets, it is important to take into account the risk in the queue formed in front of the toilet cubes, as people stay relatively close with each other. The portable toilets are simulated in a similar layout as the queue at entrance/exit/small bar counter, with 6 portable toilets cubes lining up on the top as the destinations of visitors. Each cube is an 1m-by-1m enclosed indoor space, where droplets and aerosols emitted by visitors remain until decay completely. Figure 5.11b shows the layout design of portable toilets and the queue, with 6 toilets aligned on the top. People wait in the outdoor queue before entering the toilets. The design variable is the queue distance, the same as the queue at entrance/exit/small bar.

5. *Music stage*

   The standing area of music stage is designed in a square area of 50m x 50m, as illustrated in Figure 5.13. The dimensions are estimated from the YouTube vlog of Amsterdam Open Air festival 2019 by Huijkman [2019]. It is assumed that the stage is on the top, entrance at the bottom left corner and exit at the



(a) Normal toilet layout



(b) Portable toilet layout

Figure 5.11: Toilet infrastructure layout

Figure 5.12: A line of portable toilets (Source: Gobbler [2020])

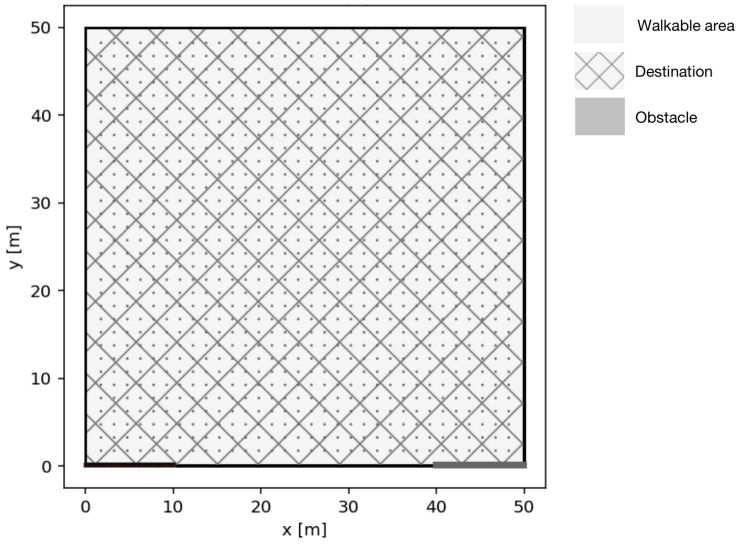


Figure 5.13: Music stage layout

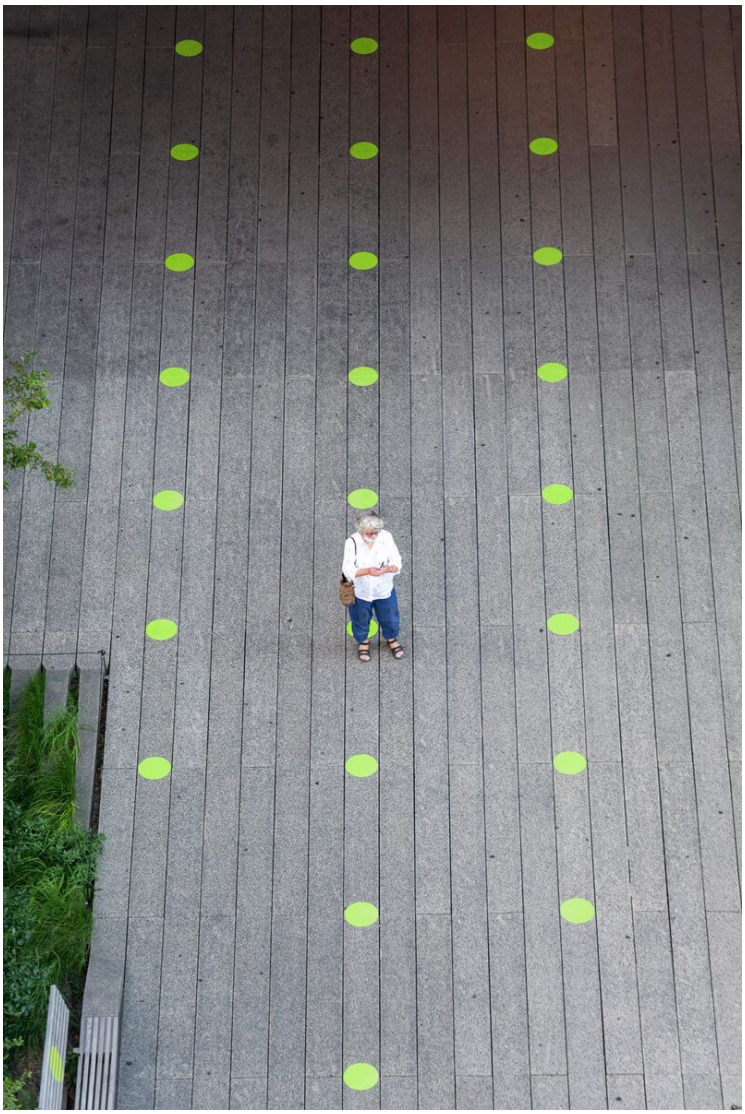


Figure 5.14: Social distancing markers (Source: Nedelcheva [2020])

bottom right corner. As is observed from real life experiences, visitors clustering in front of the music stage tend to stay around the same location during their visit, due to the crowdedness. To simulate visitors' behavior at the music stage, this study assigns uniformly distributed points as a range of available destinations to the agents. This design is inspired by the ubiquitous social distancing markers that have been widely applied since the start of the pandemic, as shown in Figure 5.14. It also allows the quantification of infection risk The distance between the points can be adjusted to simulate scenarios with different densities. To simulate visitors' choice of the destination point in reality, a 'focal point' is introduced. It is an 'attraction' point to which visitors tend to come as close as possible. During the simulation, if agents are set to follow the focal point, they always take the available destination point that is the nearest to the focal point. In this designed music stage area, the focal point is placed in the center of the top line, to mimic the visitors' behavior of clustering in the center front of the stage. In real life practice, people do not always want to go the most crowded area in front of the stage. Some would prefer staying at a random location in the field. In this study, a randomness indicator is incorporated for the choice of destination points. The indicator defines the proportion of visitors following the focal point (80% in this scenario) and the proportion of visitors randomly choosing the destination (20% in this scenario). It is assumed that there is no touchable object in the music stage area. The infrastructure related design variables include the density and indoor/outdoor setting. Four types of density are simulated, namely jam density ($5.76p/m^2$ - distance between people: 0.4m) [Duives et al., 2015], very high density ($2p/m^2$ - distance between people: 0.7m), high density ($1p/m^2$ - distance between people: 1m), ideal density ($0.4p/m^2$ - distance between people: 1.5m).

NOMAD demand pattern and activity schedule

After defining the infrastructure layout of different activity spaces, their demand patterns are derived from the infrastructure capacity estimation and the demand pattern of the entire event. As this study aims to reproduce the 2019 Amsterdam Open Air music festival, the demand considered in this research remains the same as the scale of this event, which is 10000 participants in a day. The activity schedules are derived from the infrastructure settings in the previous subsection.

The demand patterns and activity schedules of simulated activity spaces are presented as follows.

1. *Entrance/exit*

   The entire event accommodates 10000 visitors and lasts from 12:00 pm to 23:00 pm. As analyzed in Section 5.1, people spend 7 to 10 hours at the festival, indicating a scattered demand during the first few hours. Considering the simulation scale and computational limitations, this study aims to simulate an 'average' scenario for each activity space, 'averaging' the risk throughout the busy hours. It is assumed that visitors enter and leave the festival at a constant rate. Deducting the average stay of 8 hours 50 minutes from the event duration, 11 hours, the demand to the event is calculated as 10000 persons per 2 hours 10 minutes. Assuming there are 20 gates at the entrance, if all the visitors arrive at the event within the first 2 hours 10 minutes, the demand at each gate is 1 person every 15.6 seconds. In real life experiences, it often happens that the demand is bigger than the entry rate, which builds up the queue. In this study, the entry rate at the gate is set as 20 seconds per person, which is considered by the author, as a reasonable time to scan the QR code, wait till the gate open and walk through the gate. Under such an entry rate, it is calculated that the queues will resolve within the first 2 hours 47 minutes,

indicating the time spent at the queue ranges from 0 to 37 minutes. It is assumed that the same rates also apply to the exit queues.

To conclude, there are 20 entrance/exit queues, each with a demand of 38.5p/10min (1 person every 15.6 seconds) and a entry rate of 30p/10min (1 person every 20 seconds). The activity schedule is straight from the entry to the tail of the queue, then to the entrance/exit gate, and exit the facility.

2. *Lockers*

   For the locker area, the maximum capacity at a time is 24 people, as discussed in the previous subsection. According to the activity pattern analysis from Section 5.1.2, people on average spend 3'41" in the locker area, resulting in a maximum flow rate of 65p/10min. As analyzed in Section 5.1.3, people always visit the locker after entering and before leaving and usually also pay 0 to 2 visits to the locker during the entire event. Therefore, for the first 2 hours 47 minutes (total time required to let in 10000 visitors), the inflow at the locker area corresponds to the inflow at the entrance. Following the same assumption at the exit, the inflow to the locker proceeds to the exit, which makes the flow rate at the locker 10000 persons per 2 hours 10 minutes. To avoid underestimating the infection risk at the locker area, this study simulates the maximum flow rate scenario, which makes the flow rate of the entire locker area only have an influence on the number of locker facilities. To allow the higher flow rate at the exit, the number of locker facilities is calculated as 12 (10000 persons per 2 hours 10 minutes inflow divided by 65p/10min maximum flow rate). It is important to pay attention that, in practice, the flow rate at the locker area does not always remain the maximum throughout the entire event. By simulating maximum flow rate scenario, this study may overestimate the infection probability at this event.

   In conclusion, the demand at each locker facility is 65p/10min. 12 identical lockers are installed in the simulated music event to satisfy the need to deposit belongings for all the visitors.

   The activity schedule is determined as simple as 'Entry-locker destination-exit', based on the infrastructure layout.

3. *Bars*

   As summarized in Table 5.1, the average stay in the big bar area is 6'51", which includes the time spent to get a drink (or food) and the time spent sitting at the tables in the area. As the big bar infrastructure layout is set up to promote social distancing, the number of people seated is 24 at a time, which allows the flow rate of 24p/6'51" = 35p/10min. Assuming each individual spends one minute at the counter to order and pay for a drink (or food), the 6 counters in total allow the flow rate of 60p/10min. Therefore, the maximum flow rate to the big bar area is 60p/10min, while only 35 people out of the 60 are able to sit at the tables.

   Same for the small bar, maximum flow rate at a single counter is 10p/10min. The demand pattern is set as 12p/10min for the first half and 8p/10min for the second half to make sure that agents spend a period of time in the queue and be potentially exposed to the virus. To satisfy the average number of visits to the bar by on individual, 2.8 (Table 5.4), during the 11-hour event, the overall demand flow rate to all the bars is calculated as: 2.8*10000p/11h = 424p/10min. As a result, 4 big bar area (240p/10min) and 19 small bars (190p/10min) are installed in the simulated music event.

   According to the infrastructure function, the activity schedules at the big bar are determined as 'Entry-counter-table-exit' and 'Entry-counter-exit'. The latter also applies to the small bar areas.

4. *Toilets*

   The average duration of stay in toilets is 3'21", as analyzed in Table 5.1. The infrastructure setting determines the maximum number of people in the normal toilet is 4 at a time. Therefore, the maximum flow rate is derived as 4p/3'21" = 12p/10min. For the portable toilet setting displayed in Figure 5.11b, the maximum number of people is 6 and the maximum inflow is 6p/3'21" = 18p/10min. To simulate the risk in the queue, the simulated demand is set as 20p/10min for the first half and 16p/10min for the second half and guarantee the queue resolves by itself. To satisfy the need of 10000 visitors each paying 1.4 visits to the toilet (Table 5.4) during the entire event (11 hours), the overall flow rate to all the toilets is 1.4*10000/11h = 212p/10min. Therefore, 6 normal toilets (72p/10min) and 8 portable toilet queues (144p/10min) are required.

   The activity schedule is defined as 'Entry-toilet-basin-exit' at the normal toilet and 'Entry-queue-toilet-exit' at the portable toilet setting.

5. *Music stage*

   The demand pattern of a music stage is different from other facilities, as people tend to enter and leave the area around the same time as other audience to enjoy a certain show, instead of entering and leaving at a constant rate throughout the whole day. The observed average duration of stay in front of music stages is 11'32", as summarized in Table 5.1. Based on this finding, the simulation assumes that a show lasts on average 20 minutes. People enter the facility at a constant rate for 8 minutes, stay there for 11'32" and then leave. The capacity of the music stage varies according to the designed density, from 15625p ($5.76p/m^2$)to 1111p ($0.4p/m^2$). According to Table 5.4, the average number of visits to the music stage by one individual is 8.7, which makes the total number of visits to music stages 87000. As observed in real life experience, unlike other facilities which are constantly used during a day, a music stage normally does not hold music shows throughout the entire day. Instead, a certain number of performance is staged with intervals. It is important to consider that this study aims to simulate an average scenario at each activity space, under which circumstance, the difference among different facilities of the same type is ignored. For the convenience of simulation and calculation, it is assumed that 11 shows are played at each music stage at an interval of 1 hour throughout the day. Thus, when assuming each music stage accommodates 1000 people at a time, holding 11 shows throughout the day, the number of music stages to satisfy the demand is calculated as 87000/(1000*11) = 8.

   As a result, the demand at the music stage is 1000p/8min for 1 show and 11 shows during the day. The activity schedule is defined as 'Entry-destination point-exit'.

The demand patterns and activity schedules are summarized in Table 5.5.

Activity space	Nr	Demand	Max flow	Activity schedule
Entrance/exit quque	20	38.5p/10min	30p/10min	Entry-queue-exit
Locker	12	65p/10min	65p/10min	Entry-locker-exit
Big bar area	4	60p/10min	60p/10min	Entry-counter-table-exit Entry-counter-exit
Small bar + queue	19	10p/10min	12p/10min (first half) 8p/10min (second half)	Entry-queue-counter-exit
Normal toilet	6	12p/10min	12p/10min	Entry-toilet-basin-exit
Portable toilet + queue	8	18p/10min	20p/10min (first half) 16p/10min (second half)	Entry-queue-toilet-exit
Music stage	8	1000p/8min	/	Entry-destination point-exit

Table 5.5: Activity schedules in activity spaces

NOMAD parameters

Based on the parameter calibration conducted by Campanella [2016] and the NOMAD application to a restaurant scenario [Duives et al., 2021], the parameters used in this study are defined in Table 5.6. An ongoing calibration of NOMAD parameters under social distancing rules conducted by TU Delft suggested that the interaction strength parameter, A_0, and the interaction distance parameter, d_i, may increase to $5.0m/s^2$ and 0.5m in order to keep 1.5m social distance with other people. However, as agents being stuck in the corridors at the locker facility has been observed with the calibrated values, they are not used in this study.

Parameter	Value	Unit	Source
$v_0(t)$	Normal distribution: (0.9, 0.2) Min: 0.4, Max:1.4	m/s	Campanella [2016]
τ	0.5	m/s^2	Campanella [2016]
A_0	2.0	m/s^2	Campanella [2016]
d_i	0.1	m	Campanella [2016]
A_O	2.0	m/s^2	Campanella [2016]
d_0	0.3	m	Campanella [2016] Duives et al. [2021]

Table 5.6: NOMAD parameter values

5.2.2 QVEmod model application

The inputs to QVEmod include surface definition, virus transmission related model parameters and agent movement script, the last of which is processed from the output of NOMAD. This subsection discusses the assumptions and definitions of all the required inputs under the simulated scenarios.

Surface definition

QVEmod simulates virus spread in the same infrastructure setting as NOMAD. To model the transmission by fomites, QVEmod requires the definition of touchable surfaces, which constantly receive and give away virus from every touch made by individuals. The touchable surfaces at the facilities are defined as follows.

At the entrance/exit gates, it is assumed that people scan QR codes to pass through, without touching any surface. In real life, the area in front of a music stage is usually an empty ground without any items to touch. In toilets, this study assumed that people disinfect themselves by washing hands at the basin. As a result, in the simulation, transmission via fomites is assumed to be negligible at the entrance/exit gates, music stages and toilets.

At lockers, the locker doors are defined as low-touch surfaces. It is assumed that each person touches the locker door twice during every visit to the locker facility. The material is assumed to be plastic.

At bars, the POS terminals at the counters are assumed to be touched only once during one visit, as they are touched when visitors pay for their drinks. The POS terminal material is simulated as plastic. In the big bar area, tables and benches are defined as high-touch surfaces. It is assumed that people touch them at a constant rate during sitting there [Duives et al., 2021]. The tables and benches are made from wood.

QVEmod parameters

A large number of parameters are described in Section 4.2.1 to simulate detailed virus spread dynamics. Following the ongoing research by Duives et al. [2021] and

related studies, the parameter values are defined in the order of model equations displayed in Section 4.2.1.

1. Equation 4.8

   Virus mission rate, ω, represents the number of pathogens emitted per hour by a typical infectious individual while 'breathing', which is scaled to one. The 'breathing' situation is, in fact, calibrated under a dinning-at-the-restaurant situation, which involves other respiratory activities, such as talking. The emission rates for other respiratory activities than 'breathing' are calculated by multiplying ω with δ, the activity infectiousness scaler. This study takes reference from the research by Chen et al. [2021], which identified the value of δ as 25 for singing and 3.6 for talking. It indicates the number of pathogens emitted during these two respiratory activities is respectively 25 times and 3.6 times as the number of pathogens emitted during breathing. However, as the base scenario ω is not based on 100% breathing rate, for singing and talking, the infectiousness scaler, as well as other parameters related to respiratory activities, should be adjusted. The method used to calibrate the δ of the scenario of '100% breathing' to a restaurant scenario is discussed in Section 5.2.3.

   The individual infectious scaler, σ, defines the extent to which an individual emit the virus. 0 indicates a susceptible individual who does not emit the virus and 1 represents a typical infectious individual [Chen et al., 2021]. The individual infectious scaler represents the emission rate by an average infectious individual, ϕ, defined as 10^6 RNA copies per hour [Ma et al., 2020; Leung et al., 2020]. The proportion of viruses emitted in the form of aerosols and droplets, $p_{aerosols}$ and $p_{droplets}$, are defined based on three factors: 1) the threshold size distinguishing particles of aerosols and droplets, 2) the relative amounts of aerosols and droplets expelled by humans, and 3) the difference in viral copies carried by aerosol and droplet particles. In this study, the threshold size is defined as a dry size of $10\mu m$ [Duives et al., 2021]. $p_{aerosols}$ is defined as 0.252 for breathing, 0.344 for singing, and 0.434 for talking [Chen et al., 2021], which should also be adjusted based on the calibrated restaurant scenario [Duives et al., 2021]. The face mask filter efficiency FE_i may vary for different types of masks and different airborne particles (aerosols and droplets) according to studies by Pan et al. [2021]; Ueki et al. [2020]. As face masks are not considered in this study, FE_i remains 0.

2. Equation 4.9

   The deposition rate of viral-laden droplets, $u_{droplets}$, is regarded as the sedimentation rate of droplets in static air [Xie et al., 2007]. Duives et al. [2021] derived $u_{droplets}$ as 18.18 per hour from the experiment by Vuorinen et al. [2020], considering the dry size of emitted particles.

3. Equation 4.10 to Equation 4.12

   The decay rate of viral-laden aerosols, $u_{aerosols}$, refers to the study conducted by Van Doremalen et al. [2020], defined as 1.51 per hour. The decay rate of viruses transmitted by fomites, $u_{fomites}$, depends on the surface material and the environment. The values are defined based on the lab studies by Liu et al. [2021]; Van Doremalen et al. [2020]; Chin et al. [2020]. Other researches have shown UV exposure also have a great influence on the decay rate of surface virus [Nicastro et al., 2021; Raiteux et al., 2021; Carvalho et al., 2021]. However, the existing studies have concluded very different results in different scales of comparison. The study by Nicastro et al. [2021] suggested that exposure to natural UV radiation at noon can reduce the lethal time of SARS-CoV-2 to the shortest as within 2 minutes, depending on the latitude. Raiteux et al. [2021] reported that illuminance of 10 klx and 56 klx can increase the SARS-CoV-2 decay rate from 0.58% per minute to 25.5% and 106% per minute. To explore

a reliable SARS-CoV-2 decay rate in an average summer day in Amsterdam, further research needs to be done. For the purpose of demonstration, this study assumes $u_{fomites}$ to be 10 times when in an outdoor scenario.

The indoor ACH is based on recommendations by CIRES [2020], as 3 times per hour. For outdoor spaces, the ACH is calculated according to the following equation [Ratcliff, 2018]:

$$ACH = Q/Vol \tag{5.1}$$

where:

Q is the volumetric flow rate of fresh air (m^3/h).

V is the space volume (m^3).

Assuming Q equals the space volume multiplied by the wind speed at site, the equation can be interpreted as follows:

$$ACH = v_{wind} \cdot H \cdot W \cdot 1h/H \cdot W \cdot L = v_{wind} \cdot 1h/L \tag{5.2}$$

where: v_{wind} is the wind speed at site, which is assumed as the average wind speed in Amsterdam, 18.4km/h [KNMI, 2021].

H, W, L are the height, width, and length of the space.

Therefore, the ACH at the outdoor entrance/exit queue, locker, music stage, big bar, small bar queue, and portable toilet queue is calculated as 920, 1840, 368, 920, 920, and 920.

4. Equation 4.13

   No parameter involved.

5. Equation 4.14

   L, the cell volume, is calculated by 0.5m x 0.5m x 0.5m, as 125L. The inhalation rate, ρ, is defined as the ratio of human tidal volume over the cell volume per time step. The tidal volume is the volume of air that moves in or out of the lungs during each respiratory cycle, which on average equals 0.4 L for an adult [Hallett et al., 2020]. 12 respiratory cycles occur per minute [Hallett et al., 2020]. Therefore, the inhalation rate is calculated as 0.4*12*60=288L per hour. Individuals are assumed to inhale 2.304 times the cell volume (125L) each hour. The inhalation rate increases to 1.5 times while singing [Bernardi et al., 2017], which equals 432L per hour.

6. Equation 4.15 to Equation 4.18

   The surface touching frequency, γ, for tables and chairs in restaurants and bars is estimated as 0.25 touch per minute, according to the study by Lei et al. [2020]. The virus transfer efficiency from touching the surfaces, θ, is defined as 0.23 per touch [Julian et al., 2010]. The ratio of the finger pad size over the cell size, π is calculated by $10cm^2$ over $10 * 10cm^2$, as 0.1 [Gao et al., 2021].

The parameter values and sources are summarized in Table 5.7.

Infectious agent assignment

Considering the larger number of activity spaces installed in the simulated event (77 in total), the possibility of two infectious agents being in the same activity space is relatively small. To maintain a reasonable number of simulations, this study assigns one infectious agent in each simulation scenario. As discussed in Section 4.2.4, the

choice of the infectious individual represents the stochasticity in the simulation. The initial number of replications is set as 10. To make sure the simulation covers all the agents exposed to a significant amount of virus within the simulated time frame, the assigned infectious agent is randomly selected from the first 5% to 15% individuals entering the simulation.

5.2.3 Influence factors

To explore the impact of different influence factors on the transmission scale in different facilities, a number of scenarios with different factor values are designed and simulated. The factors considered in this study include different queue distance, indoor and outdoor settings (air change rate and virus decay rate on surface, $u_{fomites}$), density, and respiratory activities, as shown in Table 5.8.

As is discussed in Section 5.2.2, the activity infectiousness scaler, δ, was considered as 1 for an average respiratory activity combination in the restaurant, for the calibration of the transmission scale and other parameters such as c_{route} [Duives et al., 2021]. This average respiratory activity combination involves more than just breathing. As observed in real life experiences, people also talk a lot during a meal. Unfortunately, the scientific evidence of how long do people talk and breath at a restaurant is lacking. It is estimated by the author that during a two-person meal, people spend 40% of time having a conversation. Assuming one person stays silent when the other one is talking, it results in one average individual spending 20% of time talking and 80% of time breathing. Another assumption is made that when people have a certain respiratory activity combination, the activity infectiousness scaler of each respiratory activity linearly contributes the overall activity infectiousness scaler. In this case, the activity infectiousness scaler is calculated as $3.6 * 20\% + 1 * 80\% = 1.52$, which means a restaurant scenario with the activity infectiousness scaler of 1.52 has been calibrated with the value of 1. Therefore, for other respiratory activity combinations, the calculated activity infectiousness scaler should be corrected by $1/1.52$. For instance, for the '20%talking + 40%singing + 40%breathing' scenario, the original activity infectiousness scaler is calculated as $3.6 * 20\% + 25 * 40\% + 1 * 40\% = 11.12$. Divided by $1.52/1$, the calibrated activity infectiousness scaler is then 7.32. The same calibration process also applies to proportion of viruses emitted in the form of aerosols, $p_{aerosols}$, which should be $0.434 * 20\% + 0.252 * 80\% = 0.2884$ in the case of '20%talking + 80%breathing', but was assigned 0.252 in the restaurant case. Therefore, $p_{aerosols}$ in the scenario of '20%talking + 40%singing + 40%breathing', should be $(0.434 * 20\% + 0.344 * 40\% + 0.252 * 40\%)/(0.2884/0.252) = 0.2842$. Besides δ and $p_{aerosols}$, another parameter, inhalation rate, ρ, is also influenced by respiratory activity. Its value does not require correction, as ρ stays the same for breathing and talking. In the scenario of '20%talking + 40%singing + 40%breathing', ρ is calculated as $288L * 20\% + 432L * 40\% + 288L * 40\% = 345.6$.

Besides the variables listed in Table 5.8, following factors may also have a significant influence on the virus transmission scale.

1. Face mask

   Face masks can reduce the amount of virus emitted to and inhaled from the environment. When face masks are required, it is expected that the virus transmission scale will be limited. The extent of limitation depends on the filter efficiency of the required face mask. However, the time frame of this research is limited and the contribution of face masks has been more widely recognized and evaluated [Pan et al., 2021], compared to the simulated design variables. Therefore, face mask is not simulated as a design variable and this study assumes agents emit and inhale viral-laden aerosols and droplets without any face mask protection.

2. Willingness to comply with social distancing rules

It is expected that encouraging social distance may have an influence on people's interaction behavior [Du et al., 2021]. When people are willing to comply with the social distancing rules, the interaction strength and interaction distance may change to keep a longer distance between pedestrians, which refer to A_0 and d_i (Equation 4.6) in the NOMAD model. However, as discussed in Table 5.2.1, the designed infrastructure layouts and flow rates do not allow stable simulation with calibrated parameters. Further calibration should be done to guarantee a stable simulation in different infrastructure environments, before the impact of willingness to comply with social distancing rules can be explored.

The expected transmission routes, simulation time, and number of agents simulated in each scenario are listed in Table 5.9. The simulation time is initially set as the displayed value and can be adjusted after obtaining the results to guarantee all the agents being significantly exposed to the virus are included in the simulation.

5.3 INFECTION RISK ESTIMATION

In this section, the case-specific input to derive contact probability is introduced, including population profile estimation and activity schedule simulation.

Population profile estimation

In this study, N is assumed to be 10000. Taking p_{in} from April 13th 2021 [RIVM, 2021a], which is 38.7, N_{in}, according to Equation 4.20, the number of infectious people visiting the event is calculated as 10.

Event infrastructure layout

To evaluate the infection risk of an entire event according to the findings from NOMADQVEmod simulation, a 'standard' infrastructure layout (activity space combination) of a music festival is proposed, based on everyday life experience and existing information of the Amsterdam Open Air music festival 2019 [**?**Huijkman, 2019].

The number of each type of activity space has been calculated in item 5.2.1. The selected scenarios of these facilities in a 'standard' infrastructure layout are summarized in Table 5.10.

Activity schedule simulation

According to the activity pattern analysis in Section 5.1, the following information is obtained:

1. The duration of visit to the event: average 8 hours 50 minutes, ranging from 7 hours to 10 hours.
2. The number of activity spaces visited during the event by one visitor: average 17.5, ranging from 9 to 26.
3. The activity space transition table, Table 5.3, displaying probabilities of visiting certain activity spaces after leaving certain activity spaces.
4. The duration of stay of individuals at each type of activity space: summarized in Table 5.1.

Based on the information, individuals' activity schedules are generated following the method proposed in Section 4.3.2. An example of one individual's activity schedule is demonstrated in Figure 5.15.

Nr	Entry time	Exit time	Activity space	Activity space ID
1	11:52	12:00	Entrance	Nr.5 entrance
2	12:09	12:12	Locker	Nr.6 locker
3	12:51	12:53	Small bar	Nr.9 small bar
4	13:32	13:48	Stage	Nr.3 stage
5	14:27	14:29	Locker	Nr.6 locker
6	15:08	15:35	Stage	Nr.5 stage
7	16:13	16:15	Small bar	Nr.2 small bar
8	16:54	17:09	Stage	Nr.8 stage
9	17:48	18:10	Stage	Nr.7 stage
10	18:49	19:41	Portable toilet	Nr.7 portable toilet
11	20:20	20:22	Small bar	Nr.9 small bar
12	21:01	21:16	Stage	Nr.8 stage
13	21:55	22:00	Locker	Nr.6 locker
14	22:00	22:05	Entrance	Nr.5 entrance

Figure 5.15: An example of one individual's activity schedule

Stage	Nr.1	Nr.2	Nr.3	Nr.4	Nr.5	Nr.6	Nr.7	Nr.8
12:00								
12:30	12:33-12:53	12:37-12:57				12:20-13:00		
13:00			13:18-13:46		12:51-13:11			13:19-13:39
13:30	13:24-13:44	13:23-13:43						
14:00		14:11-14:31			13:45-14:05			14:07-14:27
14:30	14:24-14:44	14:34-14:54	14:39-15:09	14:47-15:07	14:32-14:52	14:21-14:41		
15:00		15:04-15:24			15:10-15:30	15:02-15:22	15:05-15:45	
15:30				15:19-15:39				
16:00		16:04-16:24				16:06-16:26		
16:30					16:19-16:39			
17:00						16:42-17:02		17:11-17:31
17:30	17:22-17:42			17:33-17:53			17:20-18:20	
18:00					17:59-18:39			
18:30		18:18-18:38						
19:00				18:46-19:06				
19:30	19:23-19:43	19:28-19:48						
20:00					20:04-20:24	20:09-20:29		19:58-20:18
20:30			20:27-20:47					
21:00				20:54-21:14				
21:30						21:39-21:59		
22:00								21:57-22:17
22:30								
23:00								

Figure 5.16: An example of one type of activity space bearing potential exposure risk for susceptible individuals

5.3.1 Identification of the location and duration of possible virus exposure

This part and the following parts of the proposed method fully rely on the output of the above discussed steps. Therefore, no further inputs or assumptions are required. Figure 5.16 shows an example of part of the output of this step.

Parameters	Value	Source
Emission rate (ω)	Scaled to 1 unit per hour (Typical infectious individual, breathing)	
Emission rate by an average infectious individual (ϕ)	10^6 RNA copies per hour (used for informing dose-response relationships)	Leung et al. [2020] Ma et al. [2020]
Activity infectiousness scaler (δ)	25 (singing relative to breathing) 3.6 (talking relative to breathing)	Chen et al. [2021]
Individual infectiousness scaler (σ)	1 (A typical infectious individual) 0 (Susceptible individual) 1000 (super-shedder)	Chen et al. [2021]
Proportion of viruses emitted in the form of aerosols ($p_{aerosols}$)	0.252 (Breathing) 0.344 (Singing) 0.434 (Talking)	Chen et al. [2021]
Transfer efficiency (θ)	0.23 per touch	Julian et al. [2010]
Ratio of finger pads size to the cell size (π)	0.1	Gao et al. [2021]
Frequency of touching surfaces (γ)	15 touch per hour	Lei et al. [2020]
Decay rate of viruses in aerosols ($\mu_{aerosols}$)	1.51 per hour	Van Doremalen et al. [2020]
Deposition rate of droplets ($\mu_{droplets}$)	18.18 per hour	Xie et al. [2007] Vuorinen et al. [2020]
Diffusion coefficient of aerosol (D)	0.05 m^2/s	Xie et al. [2007] Vuorinen et al. [2020]
Decay rate of viruses on surfaces ($\mu_{fomites}$)	Wood: 0.969 per hour, Cloth: 0.275 per hour, Plastic: 0.193 per hour, Steel: 0.180 per hour, Glass: 0.149 per hour, Paper: 1.1 per hour, Copper: 0.323 per hour, Cardboard: 0.119 per hour (10 times when exposed to the sun)	Liu et al. [2021] Van Doremalen et al. [2020] Chin et al. [2020] Nicastro et al. [2021] Raiteux et al. [2021]
Inhalation rate (ρ)	288 L per hour (breathing, talking) 432 L per hour (singing)	Hallett et al. [2020] Bernardi et al. [2017]
Volume of a cell (L)	125 L	
Infectious dose (D_{inf})	1000 RNA copies	Popa et al. [2020]
The proportion of virions reaching respiratory cells, $c_{aerosols}, c_{droplets}, c_{fomites}$	10 % (aerosols) 10 % (droplets) 1 % (fomites)	Zuo et al. [2020] Hinds [1999] Kraay et al. [2020] Adam et al. [2020]
Air change rate (ACH)	Air in a room is replaced 3 times per hour	CIRES [2020]

Table 5.7: QVEmod parameter values

Activity space	Factor	Variable value
Entrance/exit queue	Indoor/outdoor	Indoor: ACH = 3 Outdoor: ACH = 920
	Queue distance	1m 1.5m
Locker	Indoor/outdoor	Indoor: ACH = 3, $u_{fomites_plastic}$ = 0.193 Outdoor: ACH = 1840, $u_{fomites_plastic}$ = 1.93
Music stage	Indoor/outdoor	Indoor: ACH = 3 Outdoor: ACH = 368
	Density	$5.76p/m^2$ $2p/m^2$ $1p/m^2$ $0.4p/m^2$
	Respiratory activity	20%talking + 80%breathing 20%talking + 40%singing + 40%breathing
Big bar	Indoor/outdoor	Indoor: ACH = 3, $u_{fomites_plastic}$ = 0.193, $u_{fomites_wood}$ = 0.969 Outdoor: ACH = 920, $u_{fomites_plastic}$ = 1.93, $u_{fomites_wood}$ = 9.69
Small bar + queue	Indoor/outdoor	Indoor: ACH = 3, $u_{fomites_plastic}$ = 0.193 Outdoor: ACH = 920, $u_{fomites_plastic}$ = 1.93
	Queue distance	1m 1.5m
Normal toilet	Indoor	ACH = 3
Portable toilet + queue	Indoor/outdoor	Indoor: ACH = 3 Outdoor: ACH = 920
	Queue distance	1m 1.5m

Table 5.8: Selected simulation variable and their values

Activity space	Transmission route	Simulation time	Nr of agents
Entrance/exit queue	Aerosols/Droplets	2h	462
Locker	Aerosols/Droplets/Fomites	2h	780
Big bar area	Aerosols/Droplets/Fomites	2h	720
Small bar + queue	Aerosols/Droplets/Fomites	2h	300
Normal toilet	Aerosols/Droplets	2h	144
Portable toilet + queue	Aerosols/Droplets	2h	216
Music stage	Aerosols/Droplets	20min	1000

Table 5.9: Simulation attributes

Activity space	Nr	Scenario
Entrance/exit queue	20	Outdoor: ACH = 920 Queue distance: 1m
Locker	12	Indoor: ACH = 3, $u_{fomites_{plastic}}$ = 0.193
Big bar area	4	Outdoor: ACH = 920, $u_{fomites_{plastic}}$ = 1.93, $u_{fomites_{wood}}$ = 9.69
Small bar + queue	19	Outdoor: ACH = 920, $u_{fomites_{plastic}}$ = 1.93 Queue distance: 1m
Normal toilet	6	Indoor: ACH = 3
Portable toilet + queue	8	Outdoor: ACH = 920 Queue distance: 1m
Music stage	8	1/8: Indoor: ACH = 3, density: $2p/m^2$ 20%talking + 40%singing + 40%breathing 7/8: Outdoor: ACH = 368, density: $2p/m^2$ 20%talking + 40%singing + 40%breathing

Table 5.10: A 'standard' infrastructure layout of a large-scale music festival

6 RESULTS

This chapter presents the results from the method application presented Section 5.2. First, the results of virus spread simulation in different activity spaces with different design variables are visualized and processed for computing the infection risk of an entire visit to the music festival. Second, the simulated contact probability is presented. Third, the virus exposure distribution is fitted into the simulated contact probability. The accumulated virus doses then translate into the general risk profile of the entire visit to this music festival.

6.1 VIRUS SPREAD SIMULATION

In this subchapter, the results of virus spread simulation in different activity spaces are presented. The distributions are used for computing the infection risk of an entire visit to a music festival.

6.1.1 Significant threshold of accumulated virus dose

Before presenting and analyzing the simulation results, it is important to determine a significant threshold of virus exposure, which separates the virus dose that contributes to possible infections during visiting the entire event from the negligible virus dose. It is observed from the simulation results that the virus transmitted by fomites and aerosols in indoor scenarios can remain viable after several hours with very small proportion that makes very little contribution to the infection probability. To save the simulation time for QVEmod and to efficiently evaluate the accumulated virus exposure in different activity spaces throughout the entire event, a significant threshold value is introduced. Its value is determined according to the following assumption.

> ***The equivalent accumulated virus dose (Droplets + Aerosols +*** 10% ***Surface virus) one agent caught at one activity space, multiplied by the maximum number of activity spaces visited during one day (26), resulting in more than*** 0.1% ***chance of being infected is considered as a significant contribution. The amount of exposure smaller than this value is considered as negligible, and therefore will not be used for computing the infection risk of an entire visit to a music festival.***

The reason to count 10% of the surface virus is that $c_{fomites}$, the proportion of virus transmitted by fomites that reaches the respiratory tract cells, is 10% of $c_{droplets}$ and $c_{aerosols}$. The threshold infection probability of 0.1% is selected to make sure that the actual infection number is at most $0.1\% x number of visitors$ higher than the resulting infection number, the impact of which is assumed to be negligible in this study.

According to Equation 4.23 and Equation 4.24, the significant threshold value is calculated as $3.846x10^{-7}$.

6.1.2 Entrance/exit queue

In the average case of entrance/exit queue, the infectious agent is the 32nd person to enter the queue, who has the potential to infect not only the agents entering

after him/her, but also the susceptible individuals who have already been inside the space when the infectious agent enters.

The simulation result statistics of the entrance/exit queue are displayed in Table 6.1 (dur is short for duration). It is observed that the first agent to catch the virus is the first to enter the queue after the infectious individual. The amount of virus picked up by followers then decreases rapidly. Figure 6.1 plots the virus dose transmitted to the first 20 agents following the infectious agent, revealing the decay trend of transmission in different scenarios.

Statistics	1m distance		1.5m distance	
	Indoor	Outdoor	Indoor	Outdoor
Sum significant airborne exposure (10^{-4})	31	17	25	15
Maximum airborne exposure (10^{-4})	4.3	3.4	4.3	4.1
Average significant airborne exposure (10^{-4})	0.2	0.3	0.3	0.3
Number of agents with significant exposure	167	64	87	50
Entry dur of possible significant exposure (min)	43	17	23	13

Table 6.1: Virus exposure statistics at the entrance/exit queue

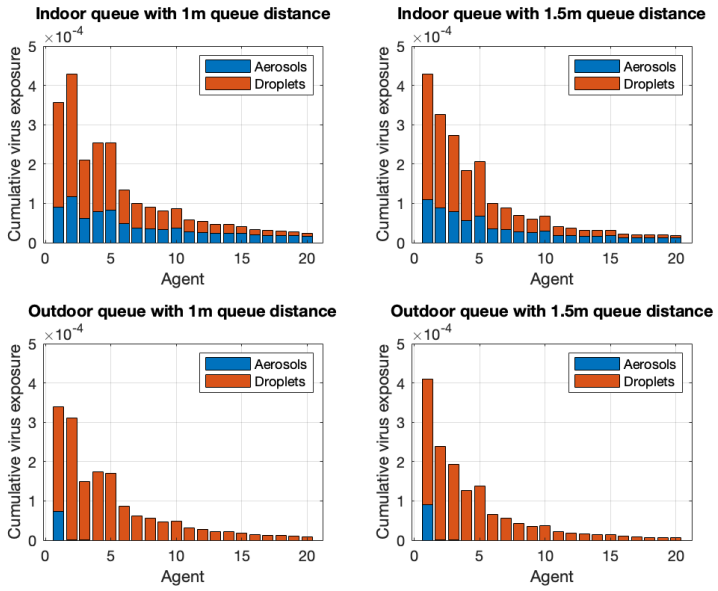


Figure 6.1: The accumulated exposure of the first 20 agents picking up virus at the entrance/exit queue

As is shown in the figure, in all scenarios, the amount of virus transmitted to susceptible individuals decreases to around 10% of the maximum exposure within 10 people. In outdoor scenarios, aerosols barely contribute to virus transmission, while in indoor scenario, they continue to transmit the virus after droplet transmission decreases drastically. In the 1.5m-queue-distance scenario, it is observed that the amount of transmitted virus decreases slightly faster compared to the 1m-queue-distance scenario. The trend is displayed more clear in Table 6.1, as all the statistical values are relatively higher in the 1m-queue-distance scenario. It can be concluded from Table 6.1 that larger queue distance and outdoor setting help reduce the scale of virus transmission in the crowd, which is within the expectation, as SARS-CoV-2 decays with distance and air change.

6.1.3 Locker

In the average locker scenario, the infectious individual is the 54th person to enter the facility, who has the potential to infect not only the agents entering after him/her, but also the susceptible individuals who have already been inside the space when the infectious agent enters.

The virus transmission trend is different in the locker facility compared to the entrance/exit queue. The amount of virus being picked up does not decrease constantly among the susceptible individuals who enter after the infectious individual. In fact, agents with a significant virus dose arise after a while, which might be the result of viral-laden aerosol and droplet particles not able to transmit further and only land on agents visiting the same and nearby lockers. Figure 6.2 illustrates the accumulated exposure on the first 60 agents who visit the locker facility following the infectious individual during the first 9 minutes. As is displayed, in both indoor and outdoor scenarios, only a small number of individuals catch a significant virus dose spread by the infectious agent. The simulation result statistics of the locker are listed in Table 6.2.

Statistics	Indoor	Outdoor
Sum significant exposure (10^{-4})	11	8.5
Maximum exposure (10^{-4})	7.7	7.4
Average significant exposure (10^{-4})	0.1	0.4
Number of agents with significant exposure	87	21
Entry dur of possible significant exposure (min)	37	13

Table 6.2: Virus exposure statistics at the locker

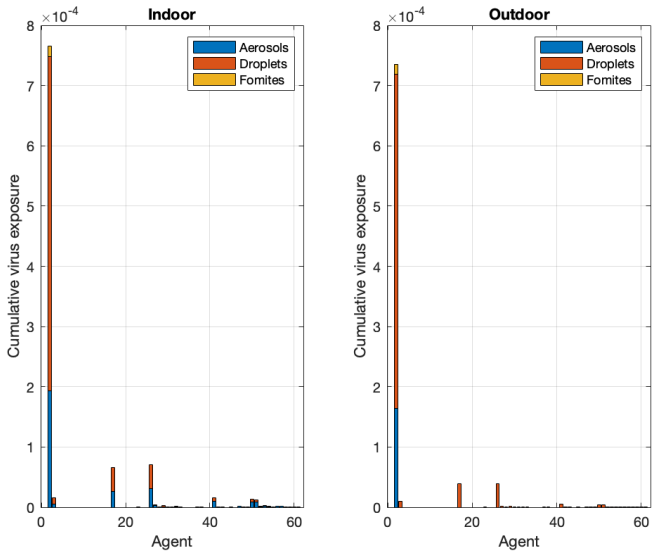


Figure 6.2: The accumulated exposure of the first 60 agents following the infectious agents at the locker

As is discussed in Chapter 5, the viruses caught via different routes contribute differently to the infection risks. Viruses transmitted by fomites are 1/10 times as effective as viruses transmitted by droplets and aerosols. In Figure 6.2, the exposure from three routes are demonstrated in proportion. It is clear that fomites transmission barely contributes to the virus spread in the locker. It corresponds to the expectation that fomites play a limited role in the transmission process. It is also observed that only one agent throughout the distribution catches a relative large

amount of virus, the dose of which consists the majority of the virus picked up by all the following agents.

Concluding from Table 6.2, virus remains longer and spread to more people (in an extremely low volume) in the indoor scenario. The outdoor scenario significantly reduces the virus transmission scale on the aspect of the total number of agents being exposed. However, the highest exposure and the sum of exposures are similar in both situations, indicating the exceeding number of agents with significant exposure, in fact, catch only a limited amount of virus.

6.1.4 Big bar area

During the simulation, it is discovered that in the indoor scenario, the amount of virus caught by agents does not decrease to below the significant threshold in the initially set simulation time, 2 hours. Therefore, the simulation is rerun for a longer period, up to 5 hours. The range of assigned infectious individual remains the same as the initial calculation in the 2-hours time frame, i.e., between the first 5% to 15% people entering the facility.

In an average scenario, the infectious individual is the 44th person to enter the facility. The results show that the exposure risk becomes negligible after 4.5 hours and 1.5 hours in the indoor and outdoor scenario. The statistics are summarized in Table 6.3. The accumulated virus distribution is similar to the locker scenario, where the exposure depends both on when the agents enter and where the agents stay in the facility. Figure 6.3 displays the accumulated virus dose of the first 90 agents that enter the facility after the infectious agents.

Statistics	Indoor	Outdoor
Sum significant exposure (10^{-4})	48	24
Maximum exposure (10^{-4})	8.5	8.2
Average significant exposure (10^{-4})	0.2	0.3
Number of agents with significant exposure	208	90
Entry dur of possible significant exposure (min)	273	91

Table 6.3: Virus exposure statistics at the big bar area

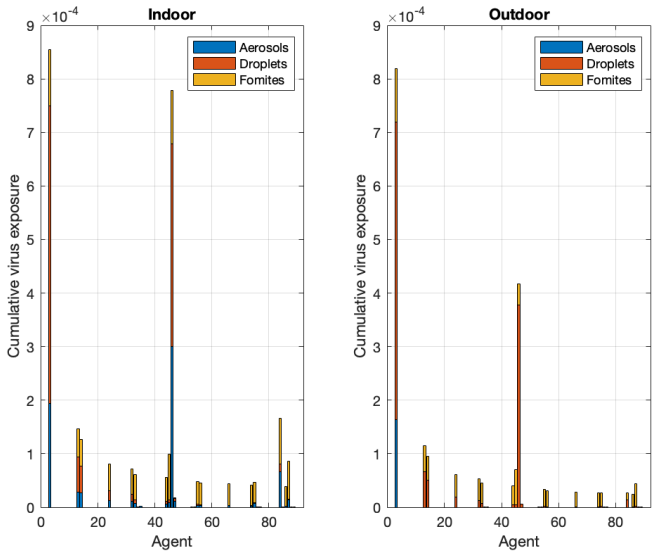


Figure 6.3: The accumulated exposure of the first 90 agents following the infectious agents at the big bar area

As is illustrated in Figure 6.3, in both indoor and outdoor scenarios, among all the agents entering after the infectious individual, only a small number of agents catch a significant amount of virus. The largest exposure happens immediately after the infectious individual, which is expected to be picked up by the individual who sits at the same table as the infectious individual. Three routes all contribute the virus transmission. In the indoor scenario, aerosol transmission plays an important part for the 3 highest virus exposures, while it only contributes the 1 highest virus exposure in the outdoor scenario. Same as aerosols, virus transmitted by fomites also decreases faster in the outdoor scenario, due to the high decay rate under the solar radiance.

It is observed in Table 6.3 that in the indoor scenario, the total amount of virus transmitted is twice as high as the outdoor scenario, while the maximum exposure remains similar. Due to the low air change rate and surface decay rate, more agents are exposed to the possibilities of catching virus in the indoor scenario. In the first 4.5 hours after the infectious individual has entered the facility, 208 out of 1638 visitors are exposed to a significant amount of virus in the indoor scenario. While in the outdoor scenario, the proportion is 90 out of 547 and the duration is shortened to 1.5 hours.

6.1.5 Small bar

It occurred during the simulation that in the indoor scenario, the accumulated virus dose of agents does not decrease to below the significant threshold within the initially set time period, 2 hours. The simulation is then rerun for 4 hours to include all the agents catching a significant amount of virus.

In the average case of the small bar and queue scenario, the infectious agent is the 18th person to enter the queue, who has the potential to infect not only the agents entering after him/her, but also the susceptible individuals who have already been inside the space when the infectious agent enters.

The simulation result statistics of the small bar are displayed in Table 6.4. The same as the trend observed in the entrance/exit queue, it is the first agent entering the queue after the infectious individual catches the largest amount of virus. Then the accumulated virus dose on the followers decreases rapidly. What's worth noticing is that the agent who enters before infectious individual also catches the virus, but the amount is too far from being significant. This might be because of that the agents who stand in front of the infectious individual never occupy the position where the infectious agent has stayed and emitted virus in the environment. Therefore, the virus volume around the people who stand in front of the infectious agent never becomes large enough to be picked up. For the convenience to observe, Figure 6.1 only plots the virus dose transmitted to the first 20 agents following the infectious agent, revealing the decay trend of transmission in different scenarios.

Statistics	1m distance		1.5m distance	
	Indoor	Outdoor	Indoor	Outdoor
Sum significant exposure (10^{-4})	49	27	46	26
Maximum exposure (10^{-4})	14	11	13	10
Average significant exposure (10^{-4})	0.3	0.3	0.2	0.3
Number of agents with significant exposure	189	89	189	89
Entry dur of possible significant exposure (min)	186	74	186	74

Table 6.4: Virus exposure statistics at the small bar + queue

It is clear from Figure 6.4 that fomites transmission plays a very limited role in the transmission process for the first 20 agents. As expected, virus transmitted by aerosols contributes more in the indoor scenario while rapidly decreasing to a negligible level after being picked up by 1 or 2 agents. However, compared to the simulation results at the entrance/exit queue, the value of the largest is almost 10

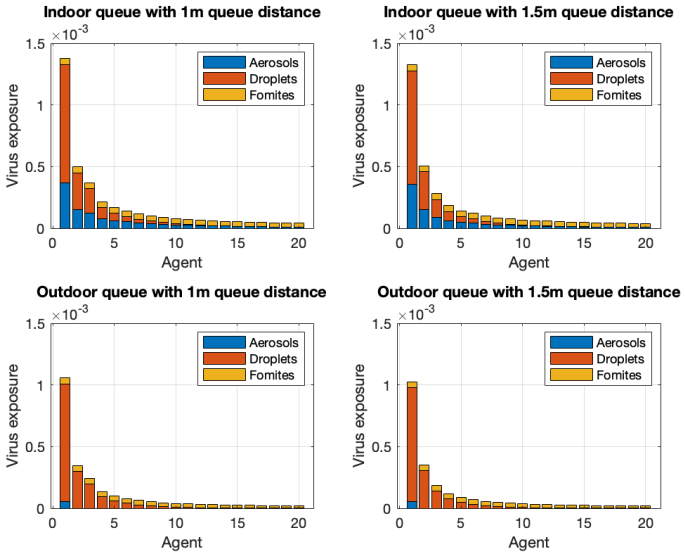


Figure 6.4: The accumulated exposure of the first 20 agents following the infectious agents at the small bar

times as high and the overall decreasing rate is also much higher. This is expected to be caused by the relatively long time it takes to move up one point further (60s instead of 20s) in the small bar queue. As a result, people spend more time in the location with a high concentration of virus in the environment, where the infectious agent emits viral-laden particles.

Table 6.4 confirms the findings in Figure 6.4, that the sum and maximum of significant exposure are higher compared to the entrance/exit queue, which results in more agents during a longer period of entry being exposed to the virus. It is also observed that an outdoor setting helps reduce the number of agents being exposed to a significant amount of virus to less than a half compared to the indoor setting, mainly thanks to the UV radiance. As is displayed in Figure 6.4, after the first 10 to 15 agents, the virus transmitted by fomites begins to dominate the accumulated virus exposure. Without the UV radiance, the time it takes for fomites virus to decrease to an insignificant level is more than doubled in the indoor scenario. Surprisingly, as is demonstrated in Table 6.4, increasing the queue distance from 1m to 1.5m only slightly reduces the total virus being picked up by agents, the degree of which is smaller than what is discovered at the entrance/exit queue. This might be the result of fomites playing a major role in the virus transmission and not influenced by the queue distance.

To conclude, a longer queue distance and an outdoor setting reduces the virus transmission scale in the small bar, the latter being more efficient compared to the queue distance increasing from 1m to 1.5m.

6.1.6 Music stage

In the average case of the music stage scenario, the infectious agent is the 98th person to enter the facility, who has the potential to infect not only the agents entering after him/her, but also the susceptible individuals who have already been inside the space when the infectious agent enters.

During the simulation, it is observed that whether and how much a person pick up virus does not depend on the time to enter the facility, but depends on the chosen destination point. According to the results, even under the 'safest' scenario, agents entering both before and after the infectious agent are exposed to the possibility of picking up the virus. As people stay in the same destination point for a relatively

long period of time (more than 10 minutes), it is expected that the virus transmission mainly occurs to people standing in the vicinity. Figure 6.5 shows the 20 highest accumulated exposure in 16 scenarios on the log scale, due to huge differences between the values in different scenarios. The axes are not specified in the graph for the clarity of display, the meanings of which are explained in the caption. The statics are summarized in Table 6.5 and Table 6.6.

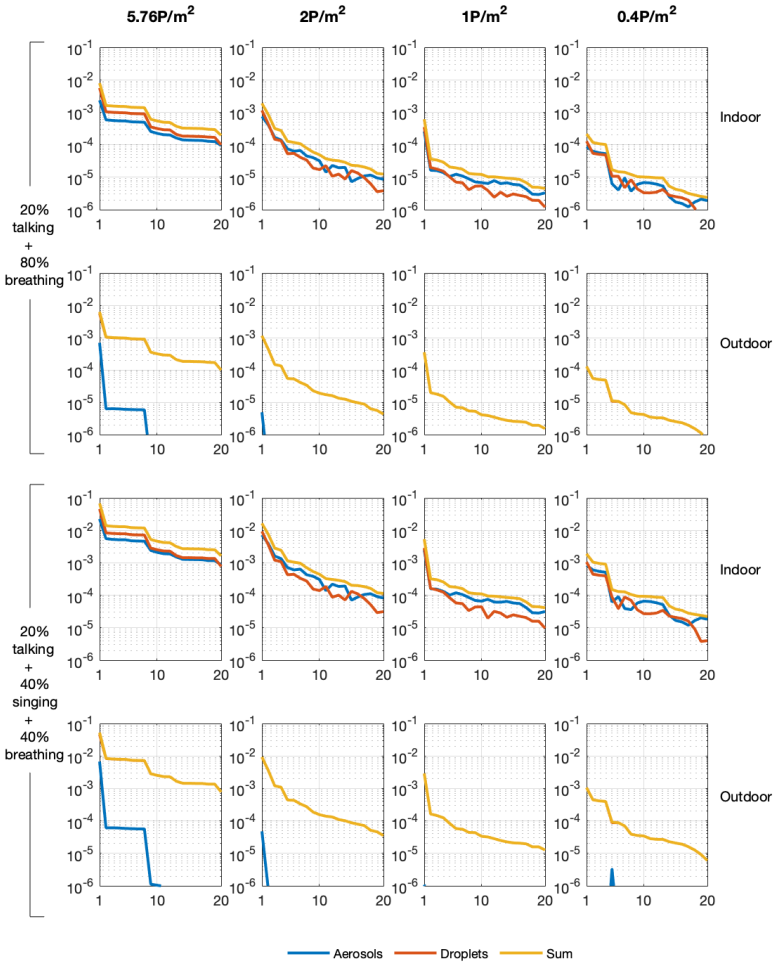


Figure 6.5: The accumulated exposure of the 20 agents with the highest virus dose at the music stage (Y axis - virus exposure, X axis - agent; when virus exposure by droplets is the same as the total exposure, the orange line is covered by the yellow line)

Statistics	Indoor								Outdoor							
	$5.76p/m^2$		$2p/m^2$		$1p/m^2$		$0.4p/m^2$		$5.76p/m^2$		$2p/m^2$		$1p/m^2$		$0.4p/m^2$	
		pct		pct		pct		pct		pct		pct		pct		pct
Sum significant exposure (10^{-4})	260	100%	42	16.2%	9	3.6%	7	2.6%	171	100%	22	13%	4.8	2.8%	3.5	2.1%
Maximum exposure (10^{-4})	79	100%	19	23.8%	6	7.7%	2	2.7%	63	100%	11	18.2%	3.5	5.6%	1.3	2.1%
Average significant exposure (10^{-4})	1.6	100%	0.7	44.1%	0.2	12.6%	0.2	13%	1.2	100%	0.5	37.2%	0.1	11%	0.1	11.3%
Number of agents with significant exposure	164	100%	60	36.8%	46	28.2%	33	20.2%	137	100%	48	35%	35	25.5%	25	18.2%
Entry dur of possible significant exposure (min)	8	100%	8	100%	8	100%	8	100%	8	100%	8	100%	8	100%	8	100%

Table 6.5: Virus exposure statistics at the music stage (20%talking + 80%breathing): percentages are calculated based on the statistics of the $5.76p/m^2$-density scenarios

Statistics	Indoor								Outdoor							
	$5.76p/m^2$		$2p/m^2$		$1p/m^2$		$0.4p/m^2$		$5.76p/m^2$		$2p/m^2$		$1p/m^2$		$0.4p/m^2$	
		pct		pct		pct		pct		pct		pct		pct		pct
Sum significant exposure (10^{-4})	2226	100%	374	16.8%	82	3.7%	61	2.7%	1391	100%	182	13.1%	40	2.9%	29	2.1%
Maximum exposure (10^{-4})	679	100%	164	24.1%	54	8%	19	2.8%	523	100%	93	17.8%	29	5.5%	11	2.1%
Average significant exposure (10^{-4})	10	100%	4	40%	1	10%	1	10%	7.5	100%	2.5	33.3%	0.7	9.3%	0.7	9.3%
Number of agents with significant exposure	219	100%	93	42.5%	76	34.7%	61	27.9%	186	100%	73	39.2%	53	28.5%	43	23.1%
Entry dur of possible significant exposure (min)	8	100%	8	100%	8	100%	8	100%	8	100%	8	100%	8	100%	8	100%

Table 6.6: Virus exposure statistics at the music stage (20%talking + 40%singing + 40%breathing): percentages are calculated based on the statistics of the $5.76p/m^2$-density scenarios

As is illustrated in Figure 6.5, Table 6.5, and Table 6.6, indoor/outdoor setting, density, and respiratory activity, these three factors all have significant influences on the transmission scale.

For the indoor/outdoor setting, the same trend identified in other activity spaces is observed in Figure 6.5, that the virus transmitted by aerosols decays rapidly in an outdoor setting, as the viral-laden aerosol particles are removed as a much higher frequency in an outdoor scenario. Only under the maximum density scenario, the agent with the highest exposure, who stood next to the infectious agent for more than 10 minutes, catches a relatively large amount of virus via aerosols. When the density decreases, aerosol transmission's role becomes negligible. This finding reveals that when spending equal or more than 11 minutes within a very short distance (equal or shorter than 0.4m) to an average infectious individual, a high air change rate (equal or higher than 368) does not eliminate the risk of aerosol transmission. What makes this finding different from the previously discussed activity space is that agents do not move during staying at the music stage, which eliminates the possibility of following agents staying at the same location as where the infectious individual has stayed. In a queue/locker room/bar, an infectious individual stays in one location for a short period of time, emits virus in the environment, and then leave. The next individual then comes in immediately, takes up this location and picks up the virus emitted in the form of aerosols. In these scenarios, there is always at least one individual picking up viral-laden aerosol particles on the same scale as droplet particles. However, in the music stage simulation, this only happens when the density is as high as $5.76p/m^2$ (distance between people 0.4m), which indicates that people may pick up less virus transmitted by aerosols when staying in the vicinity of the infectious individual, compared to when they stay in the same location after the infectious individual has just left. In the latter case, people may still catch a relatively large amount of virus without being close to the infectious individual. According to Table 6.5 and Table 6.6, the total amount of virus picked up by individuals decreases by around 35% (high density) to 50% (low density) in the outdoor scenarios compared to the indoor scenarios. The maximum exposure, average exposure, and the number of people catching a significant amount of virus are around 20% smaller in the outdoor scenarios.

For the influence of densities, it is observed in Figure 6.5 that with the increase of density, the transmission scale increases as well. Table 6.5 and Table 6.6 clearly display the differences in transmission scale under different densities. It appears that the sum of significant virus exposure shrinks remarkably to 16.2%-16.8%, 3.6%-3.7%, and 2.6%-2.7% in the indoor scenarios, when the density changes from $5.76p/m^2$ (0.4m distance between people), to $2p/m^2$ (0.7m distance between people), $1p/m^2$ (1m distance between people), and $0.4p/m^2$ (1.5m distance between people), while the proportions are even lower (13%-13.1%, 2.8%-2.9%, and 2.1%) in the outdoor scenarios. Other indicators, except the entry duration of possible significant exposure, also display a similar trend, though less significant. This finding indicates that density has a huge influence on the transmission scale, especially when it becomes larger than $1p/m^2$, as the transmission scale does not increase by several times when the density increase from $0.4p/m^2$ to $1p/m^2$, while it does from $1p/m^2$ to higher. It was expected that a higher density triggers a higher transmission scale. However, this simulation results also identify that the influence of increasing density is much higher in scenarios of density larger than $1p/m^2$ compared to scenarios of density smaller than $0.4p/m^2$. In addition, the influence of density is also bigger in outdoor scenarios compared to indoor scenarios, which indicates that the density (distance between people) has a more profound influence on droplet transmission compared to aerosol transmission.

For the influence of respiratory activity, it is clearly shown in Figure 6.5 that the scenarios under 20%talking + 40%singing + 40%breathing respiratory activity combination pose a much bigger threat to susceptible agents, as the 20 highest accumulated virus exposures are around 1 log scale higher in the scenarios of 20%talking

+ 40%singing + 40%breathing. As is displayed in Table 6.5 and Table 6.6, the total and highest virus exposures are about 7.5 to 8 times higher, while the average significant virus exposure is around 5 to 4 times higher. However, the change of respiratory activity does not have a huge impact on the number of people picking up a significant amount of virus, as it only goes up by 34% (high density) to 85% (low density), the increase of which is relatively small compared to other indicators. The relatively small increase may be the result of the limited transmission distance of viral-laden aerosol/droplet particles, which could not accumulate after a certain distance with a density that is large enough to transmit a significant virus dose to susceptible individuals.

When replacing 40% time spent on singing with breathing, the amount of virus picked up by individuals goes down by more than 85% in a 20-minute-duration 1000-participant music concert. The impact is even stronger in outdoor music stages, which indicates singing affects droplet transmission more than aerosol transmission. The higher increase proportion of accumulated virus dose under outdoor scenarios is out of expectation. It is assumed that airborne viral-laden particles decay faster with a higher air change rate, especially for aerosols. The compositions of aerosol and droplet particles emitted by an individuals during breathing and singing are different and a bigger share of aerosols is emitted while singing. Therefore, it is expected that, when replacing breathing with singing, the proportion of aerosols goes up and a smaller share of viral-laden droplets should lead to a smaller increase of accumulated virus dose. However, the results have proven that the activity infectiousness factor has a bigger impact on the spread of virus via droplets, making it increase by a larger proportion even when aerosols take a bigger proportion of emissions.

In a nutshell, at the music stage, keeping the density low, increasing air change rate, and avoiding singing can effectively lower the virus transmission scale. When the density is lower than $1p/m^2$, increasing the distance between people still limits the transmission scale, but is less efficient compared to when the density is higher than $1p/m^2$. In an outdoor scenario, the risk of picking up the virus via droplets might be higher when the agent takes up the position where the infectious individual has stayed compared to staying in the vicinity of the infectious individual.

6.1.7 Normal toilet

In the average normal toilet scenario, the infectious individual is the 9th person to enter the facility, who has the potential to infect not only the agents entering after him/her, but also the susceptible individuals who have already been inside the space when the infectious agent enters.

Similar to the locker and big bar facility, the amount of virus picked up by a susceptible individual does not merely depend on the entry time but also depends on the location where the agent stays. Figure 6.6 displays the virus exposures of the first 30 agents following the infectious agents at the normal toilet. The statistical indicators are summarized in Table 6.7.

Statistics	Normal toilet
Sum significant exposure (10^{-4})	9.4
Maximum exposure (10^{-4})	3.1
Average significant exposure (10^{-4})	0.2
Number of agents with significant exposure	52
Entry dur of possible significant exposure (min)	47

Table 6.7: Virus exposure statistics at the normal toilet

As is illustrated in Figure 6.6, the second agent to enter the facility after the infectious individual begins to pick up the virus and the fifth agent catches the largest amount of virus (1/3 of the sum of all the significant exposures). Then the

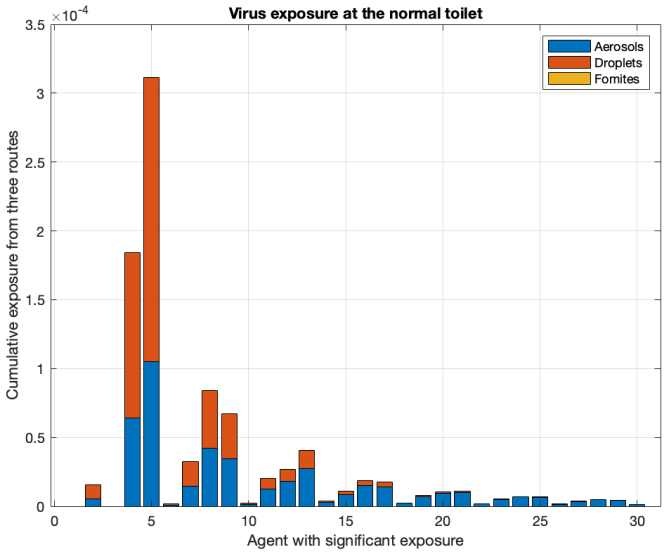


Figure 6.6: The accumulated exposure of the first 30 agents following the infectious agents at the normal toilet

virus exposure peaks in a group of 3 among every 4 agents while declining. This phenomenon occurs might be because of the the following behavior pattern. First, the infectious agent takes 1 out the 4 toilet cubes and 1 out of the 4 washing basins. Then the following agents who take the same toilet cube/washing basin and the closest toilet cubes/washing basins pick up a relatively large amount of virus, as they stay in the same location as and in the vicinity of where the infectious agent has stayed. Droplet transmission declines rapidly to a negligible level within 20 agents (17 minutes), while viral-laden aerosol particles continue to infect people with a very small dose until the 56th agent (47 minutes). According to Table 6.7, the virus transmission scale in the normal toilet is very limited (similar to the locker facility), compared to other activity spaces such as entrance/exit queues and small bars due to the limited time people spend in the facility and limited close contact (distance between people mostly larger than 1m).

In short, the virus transmission scale in the normal toilet is relatively small but the duration of possible significant exposure is as long as 47 minutes.

6.1.8 Portable toilet

In the average normal toilet scenario, the infectious individual is the 23th person to enter the facility, who has the potential to infect not only the agents entering after him/her, but also the susceptible individuals who have already been inside the space when the infectious agent enters.

As the exposure risk at the portable toilet facility exists both in the queue and in the toilet cube, the virus transmission displays a different trend compared to other queuing scenarios. Although, same as other queuing scenarios, agents entering the queue after the infectious individual start to pick up a significant amount of virus, the first agent following the infectious agent is not longer the most exposed person. Figure 6.7 illustrates the accumulated exposure of the first 20 agents following the infectious agent at the portable toilet. The statistics are summarized in Table 6.8.

As is shown in Figure 6.7, the agent with the highest exposure is the sixth to enter after the infectious individual, who stayed at the same toilet cube. Comparing the virus obtained in the cube with the virus obtained in the queue, it is clear that staying in the same cube (for 3 minutes 22 seconds) as the infectious agent immediately after he/she left (after being there for 3 minutes 22 seconds) poses a

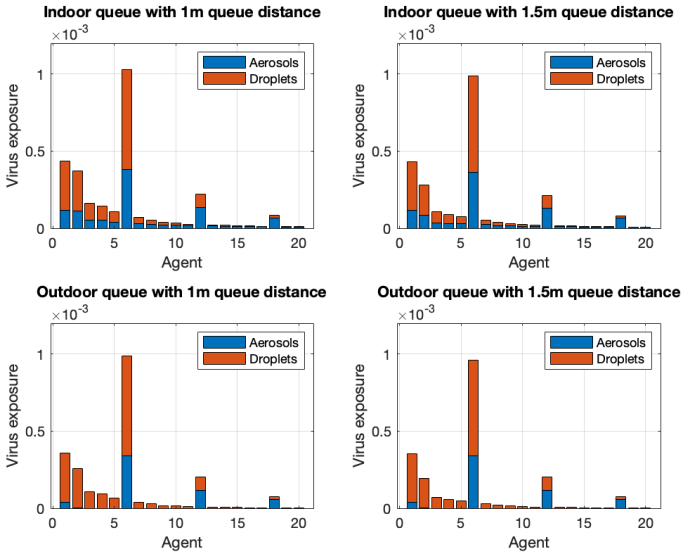


Figure 6.7: The accumulated exposure of the first 20 agents following the infectious agent at the portable toilet

Statistics	1m distance		1.5m distance	
	Indoor	Outdoor	Indoor	Outdoor
Sum significant exposure (10^{-4})	31	24	27	21
Maximum exposure (10^{-4})	10	9.9	9.9	9.6
Average significant exposure (10^{-4})	0.4	0.7	0.4	0.7
Number of agents with significant exposure	71	32	68	30
Entry dur of possible significant exposure (min)	42	33	46	33

Table 6.8: Virus exposure statistics at the portable toilet + queue

much greater risk of catching a large amount of virus comparing to staying 1m to 1.5m away from the infectious individual in the queue. It is calculated that the first agent to enter after the infectious individual spends approximately 3 minutes and 20 seconds in the queue. This finding corresponds to the finding at the music stage, which concludes that agents may be exposed to greater amount of virus via aerosols when staying at where the infectious agent has stayed compared to standing close to the infectious agents. As the toilet cube is simulated as an indoor cube, we can tell from here if a similar trend is displayed in an outdoor scenario or how much does the trend change in an outdoor scenario.

The remaining virus in the cube keeps declining as time goes by. The second person to enter the infectious agent's cube catches much less virus than the first one. According to Table 6.8, in total, around 30 (outdoor, 1.5m queue distance) to 71 people (indoor, 1m queue distance) pick up a significant amount of virus in different scenarios.

In general, outdoor setting and longer queue distance help with reducing the total amount of exposure and the number of people catching virus. The contribution of changing queue distance from 1m to 1.5m is very limited, while changing indoor setting to outdoor reduced the number of agents catching virus to half.

6.1.9 Conclusion

After analyzing the virus spread simulation results of 7 different activity spaces, it can be concluded that, in general, an outdoor setting, longer queue distance,

smaller density, and calmer respiratory activities (breathing, talking) help reduce the transmission scale of SARS-CoV-2.

An outdoor setting significantly limits aerosol and fomite transmission and increases the decay rate of viral-laden droplet particles by a high air change rate and a higher virus decay rate on the surface.

Increasing the queue distance from 1m to 1.5m does not have a huge impact on the transmission scale, as the total amount of virus picked up by susceptible agents and the number of agents with significant amount of virus are only slightly lower in the 1.5m queue distance scenario.

However, as observed in the music stage simulation results, when the distance between people is smaller than 1m, increasing the distance (0.4m to 0.7m, 0.7m to 1m) remarkably lowers the transmission scale. When the density decreases from the jam density ($5.76p/m^2$ - 0.4m distance) to very high density ($2p/m^2$ - 0.7m distance), high density ($1p/m^2$ - 1m distance) and optimal density ($0.4p/m^2$ - 1.5m distance), the total amount of virus caught by agents declines to around 13% - 16%, 3% - 4%, 2% - 3%, and the number of agents with significant amount of virus is reduced to about 35% - 42%, 28% - 35%, 20% - 28%. Lowering density has a bigger influence on the outdoor scenario, which indicates that droplet transmission is effected more by the density.

For respiratory activities, it is discovered that increasing the percentage of singing significantly rises the transmission scale, with the total virus exposures about 7.5 to 8 times higher under the scenario of 20%talking + 40%singing + 40%breathing, compared to 20%talking + 80%breathing, as the infectious scaler and inhalation rate under singing are much higher. The impact on number of people getting infected is much smaller, which indicates that people on average are exposed to a greater amount of virus.

For the transmission routes, droplets play a major role in outdoor scenarios, as aerosols decay rapidly with a high air change rate. In indoor scenarios, they both contribute to SARS-CoV-2 transmission, while the viral-laden aerosol particles remain in the environment for much longer than droplet particles. Fomites contribute very little to the virus transmission, as the amount of virus picked up by susceptible agents is much less compared to airborne transmission. However, when not exposed to UV rays, virus on the surface decay slowly and can be picked up by people with a very small dose even after 3 hours (in the small bar case).

For transmission pattern, it is derived from the observed pattern at the music stage and the portable toilet that staying in the same location where the infectious individual has been and just left might pose a greater risk of infection compared to standing next to the infectious individual (with 1m distance) for the same duration of time. This finding is based on the observed pattern at the music stage where aerosol transmission does not play any role in an outdoor setting when the distance between people is longer than 0.7m, while in other queuing scenarios, where people continuously take up the infectious individual's position, aerosol transmission always contributes to the highest exposure. In the portable toilet scenario, it is observed that the first agent who enters the same cube as the infectious individual catches more virus than the person who stays next to the infectious individual in the queue for approximately the same duration of time. However, this finding is derived from the activity-space-specific simulations, where the demand pattern, other variables may all play a role on the virus spread scale. Further research should be done to evaluate under which circumstances this finding is valid and how this can contribute to crowd management during this COVID-19 pandemic.

6.2 RISK ESTIMATION

In order to translate the above simulation results of each type of activity space into the infection risk of an entire event, a 'standard' music festival activity space com-

Infection probability	Number of agents	Percentage
0 - 1%	4582	91.04%
1 - 5%	404	8.03%
5 - 10%	18	0.36%
10 - 20%	24	0.48%
20 - 30%	5	0.10%
30 - 100%	0	0
Total	5033	100%

Table 6.9: Infection probability of the 'standard' scenario

bination has been proposed by the author, in Section 5.3. Based on the combination and activity schedule generation, the accumulated virus dose distribution will be obtained and then the infection risk distribution according to the dose-response model. The total number of infected agents and the reproduction number will be calculated based on the results.

6.2.1 'Standard' scenario

After simulating the activity schedules and matching virus exposure with agents under the 'standard' music festival activity space combination, the accumulated virus exposure and infection risk distribution are obtained.

In total, 5033 agents catch a significant amount of virus, the 200 of which with the highest exposure and infection probabilities are plotted in Figure 6.8, for the convenience and clarity of display. Figure 6.9 illustrates the distribution of total amount of accumulated virus in different activity spaces and via different routes. Table 6.9 summarizes the contribution proportion of each transmission route and the number of agents and percentages of certain infection probability ranges.

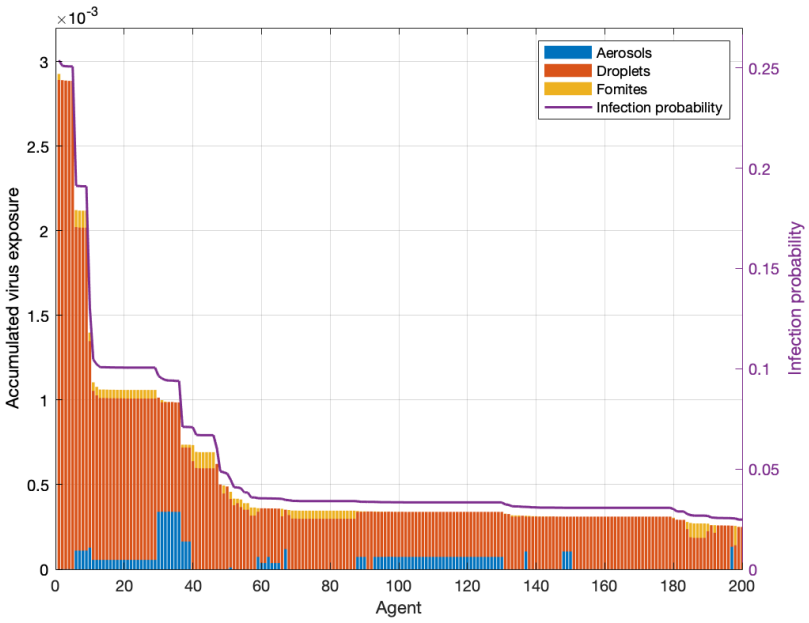
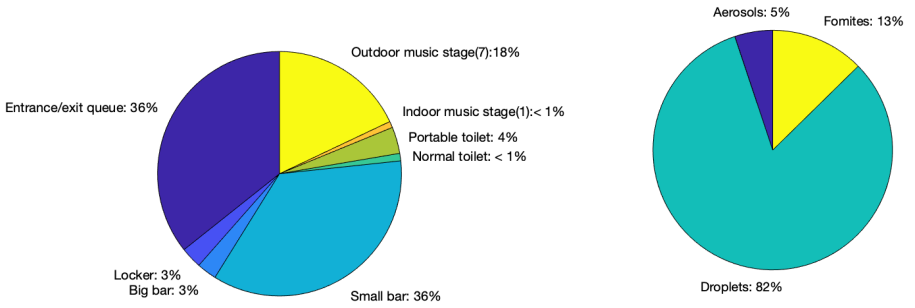


Figure 6.8: The 200 highest accumulated exposure and infection probability of agents at the 10000-people music festival

As displayed in Figure 6.8, under the 'standard' scenario, a very small number of agents (under 10) pick up a virus dose of more than $1.5x10^{-3}$ and have an infection probability of more than 15%. Under 50 agents are exposed to more than $0.5x10^{-3}$ virus and have an infection probability of more than 5%. The infection probability



(a) Distribution of virus picked up in different activity spaces

(b) Distribution of virus transmitted via different routes

Figure 6.9: Virus distribution under 'standard' scenario

curve then declines slowly after 5%, to which the majority of people are exposed. As summarized in Table 6.9, more than 90% of the agents have a infection probability smaller than 1% and only 0.94% of the exposed population, in total 29 people, has a probability of more than 5% to get infected. Out of these 29 people, only 5 have a relatively high infection probability, above 20%. Considering the large scale of this simulated event, it can be concluded that for most visitors, the possibility of getting infected in the event is as low as below 1% and only a very small number of people have a higher infection probability, the highest being 25%.

The stacked bar plot of virus exposure in Figure 6.8 displays several plateaus, where the volume and composition of the exposures are almost the same. This is the result of drawing virus exposure from one distribution in each type of activity space. For instance, the accumulated virus distribution is the same in 20 entrance/exit queues. When people are exposed to the virus in entrance/exit queues, they are assigned one value from the same distribution no matter which entrance/exit queue they are in. In total, there are 8 distributions for 77 activity spaces and 10000 agents in the simulation. When agents have the same combination of activity space types where they are exposed to the virus, it can happen that they draw the same values from the virus distributions, especially when the number of potential risky activity spaces is small. As a result, there are small clusters of virus exposures of the same volume and composition.

Figure 6.8 also illustrates that droplet transmission plays a major role in the 'standard' scenario among the agents with relatively large exposures (the highest 200), which is within the expectation, as most of the facilities are set outdoor where aerosols decay rapidly. This is confirmed in Figure 6.9b that droplets transmission takes up 81% of the total amount of virus caught by susceptible agents. Surprisingly, fomites transmission contributes to as much as 14% of the total amount of virus caught by agents, which is more than twice as the contribution of aerosols transmission. Observing from Figure 6.8, the virus transmitted by fomites takes up a small part of accumulated exposure for almost all 80 agents with the highest exposures. This indicates that a large number of agents are exposed to very small amounts of virus transmitted by fomites. These agents make up the majority of people with very low virus exposure, which leads to an infection probability smaller than 1%.

Figure 6.9a displays the distribution of the total amount of transmitted virus in all types of activity spaces, which summarizes the actual virus transmission risk (possibility x impact) of each facility. It is observed that small bars and entrance/exit queues are the most risk-prone facilities, as they each transmit 36% of the virus. Considering the average number of visits people pay to these two activity spaces, 2.8 (including big bars) and 2 (entrance/exit), which are relatively small, it is without doubt that these two facilities are the most 'dangerous' places regarding the amount

of virus possibly caught by visitors at each visit. 8 music stages take up 19% of the total amount of transmitted virus, making them the second most risk-prone facilities. The average number of visits paid to the music stages is 7.8. Dividing the proportions of virus transmitted in them by the average number of visits, it is discovered that one visit to the entrance/exit queue or the smaller is much more risky than one visit to the music stage. More virus is transmitted in an average small bar compared to an average big bar, which indicates the risk of virus transmission is higher in the queues of the small bars than the sitting areas in big bars. Locker and toilet facilities are the least risk-prone facilities. It can be concluded that activity spaces where people form queues (entrance/exit queues and small bars) or stand close to each other for a relatively long period of time (more than 10 minutes in music stages) may be in general more risky. Other facilities where people do not have close interactions and stay for a short period of time (around 3 minutes at lockers and normal toilets) may be less risky.

Following the method proposed in Section 4.3.5, it is calculated that the number of infection during the entire event is about 23.3, which makes the infection percentage 0.23%. As the initial number of infectious agents are set as 10, the basic reproduction number, R_0, of this event is then calculated as 2.33. With a R_0 of 2.33, this 'standard' scenario music festival would surely worsen the infection scale in the population.

It can be concluded that in the 'standard' event scenario, the general infection probability is relatively low, as more than 90% people has an infection probability smaller than 1% and the percentage of people getting infected is as low as 0.23%. However, considering the small number of initial infectious agents and the scale of the event, which accommodates 10000 visitors, the R_0 is 2.3, which leads to a larger scale of infection in the population. The most risk-prone facilities are the music stages, entrance/exit queues and the bar areas. Lockers and toilets are relatively less risky.

6.2.2 'Group' scenario

For the convenience of simulation and computation, this study assumes that all the simulated agents attend the event alone and make their own decisions despite other people. They plan their own activity schedules independently and every other individual poses the same attractiveness to them. However, in practice, people usually go to large events, such as music festivals and sports events, in a small group, who stay close to each other for most of the time during the event.

To make up for the lack of group consideration during the simulation, this subsection aims to reveal the possible group effect purely based on the existing virus spread simulation results of activity spaces.

This method assumes that the infectious individuals travel in a group of 4 and always stay with the rest of the team during the entire event. In the entrance/exit queues, lockers, big bar areas, small bars, normal toilets, and portable toilets, it is assumed that the 3 other people in the group are the first 3 to enter the facility after the infectious individual. In the music stage, it is assumed that the 3 are the 3 agents who stay the closest to the infectious individual and therefore catch the highest exposures. For 10 infectious agents, the 30 group mates of them are assumed to have the same activity schedules throughout the day and each is randomly assigned with one out of the 3 exposure values in each facility. The exposures taken by the group mates are deducted from the possible exposures to the rest of the population. Then the entire procedure of activity schedule generation and adding up exposures are done for the rest 9960 agents. Taking virus exposure from the group mates and the general population, the resulting number of significant exposure is 5077, almost as the same as the 'standard' scenario. The resulting statistics are summarized in Figure 6.10, Figure 6.11, and Table 6.10.

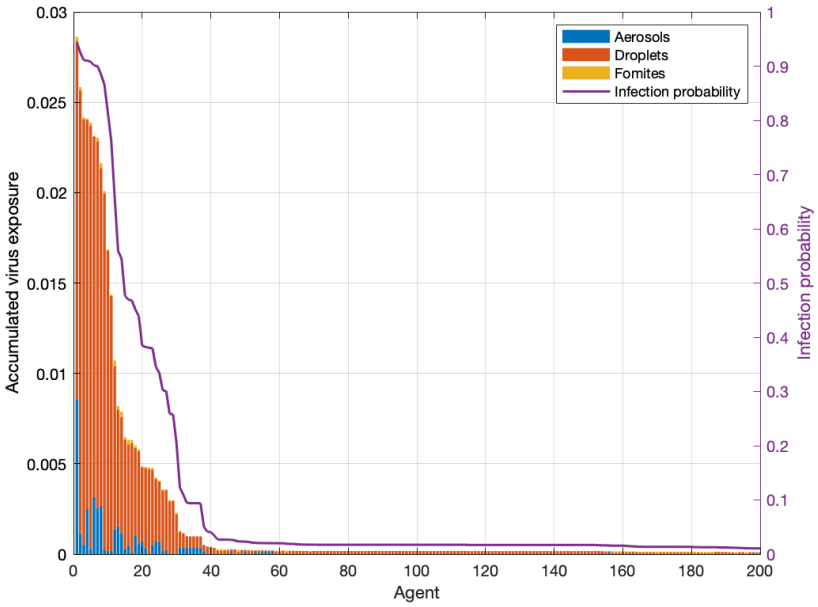
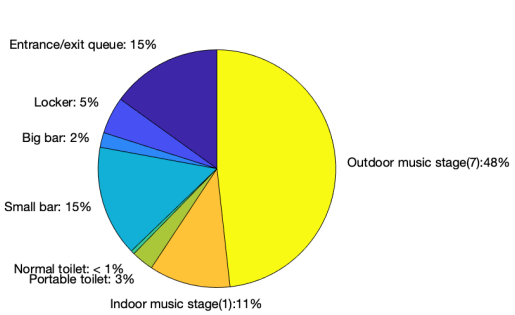
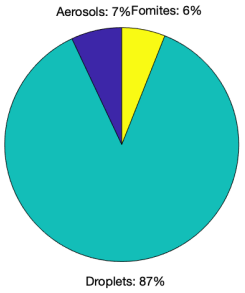


Figure 6.10: The 200 highest accumulated exposure and infection probability of agents at the 10000-people music festival with 'group' behavior



(a) Distribution of virus picked up in different activity spaces



(b) Distribution of virus transmitted via different routes

Figure 6.11: Virus distribution under 'group' scenario

For the convenience and clarity of display, Figure 6.10 demonstrates the 200 highest virus exposure and infection probability. As is illustrated, when adding up the virus load on people who stay the closest to the infectious agents, the accumulated virus distribution displays a much higher peak (almost 10 times as high as the 'standard' scenario) and a relatively large number of agents with high exposures and infection probabilities. The phenomenon of plateaus cannot be clearly observed in Figure 6.10, which is because of the highest exposures are mainly from the group mates of 10 infectious agents, among whom the activity schedules are not the same. For the rest of the population, who get assigned virus doses by random draw, plateaus certainly still exist. Observing from Figure 6.10, the composition of the accumulated virus distribution is similar to the 'standard' scenario, as droplet transmission takes up the majority of the virus transmitted and aerosols and fomites play a very limited role. As displayed in Figure 6.11b, virus transmitted by fomites composes a much smaller part in this 'group' scenario, which is corresponding to the finding of the highest exposures being caught by the group mates of infectious agents. The characteristic of fomites transmission is that the amount of the transmitted virus remains low and decreases slowly. On the contrary, droplet transmission

Infection probability	Number of agents	Percentage
0 - 1%	4842	95.37%
1 - 5%	199	3.92%
5 - 10%	6	0.12%
10 - 20%	0	0%
20 - 30%	3	0.06%
30 - 40%	8	0.16%
40 - 50%	5	0.14%
50 - 60%	2	0.04%
60 - 70%	1	0.02%
70 - 80%	1	0.02%
80 - 90%	3	0.06%
90 - 100%	7	0.14%
Total	5077	100%

Table 6.10: Infection probability of the 'group' scenario

contributes the most to the nearest agents and decreases rapidly as the distance grows bigger. Therefore, the group mates of the infected agents are exposed to a much higher volume of airborne viral-laden particles compared to the rest of the visitors, but the amount of virus they pick up via fomites transmission is not significantly different from the rest. As a result of the remarkable increase of droplet transmission, the contribution of fomites transmission is proportionally smaller.

For the virus distribution in different types of activity spaces, it is observed from Figure 6.11a that the percentage of virus picked up at music stages has increased to 59% and the proportions of virus transmitted in the entrance/exit queues and the bar areas become smaller. Lockers and toilets remain the least risk-prone facilities. The increase of virus proportion at the music stage follows a similar principle as the increase of droplet transmission. As discussed in Section 6.1.6, the highest exposures at the music stages are significantly higher than the highest exposures at other activity spaces. As group mates of the infectious individuals are exposed to the highest amount of virus at all kinds of activity spaces, the virus volumes accumulated at the music stages are proportionally higher than the rest of the facilities, which results in a higher proportion of virus transmitted in music stages. In general, the distribution stays similar as the 'standard' scenario, with the music stages, entrance/exit queues and the small bars remaining the most risk-prone facilities, with an increase of risk at the music stages.

Table 6.10 summarizes the distribution of infection probability. Compared to the 'standard' scenario, the proportion of extremely low infection probability ($0-1\%$) is higher and infection probability between 5 to 20% is almost negligible. On the other edge, there are 10 agents with extremely high infection probability ($80-100\%$) and the 30 group mates of the infectious individuals all have an infection probability higher than 20%, which means the visitors who are not in the same group as the infectious individuals all have an infection probability lower than 10%. This indicates that if infectious individuals stay in a group throughout the entire event, except their group mates, the possibility of anyone from rest of the event visitors being infected is very low, while their group mates are very likely to be infected.

It is calculated that the number of infections during the entire event is about 26.8, which makes the infection percentage 0.27% and R_0 2.68. To be more specific, it is calculated that 17.3 out of 26.8 infections are derived from the group mates of the infectious individuals, which makes the rest of the population in this event exposed to the infection possibility of as low as 9.5/9960.

It can be concluded that when considering group behavior, the transmission scale of this event is expanded with the R_0 increased by 0.35. However, 65% of the infections happen to the group mates of the infectious individuals. The infection proba-

bility of the rest of the population is 59% lower compared to the 'standard' scenario. The risk at the music stages, entrance/exit queues and the small bars remains the highest, with an increase of virus proportion transmitted at the music stages and a slight decrease at the entrance/exit queues and the small bars. Fomites transmission plays a smaller role as the virus transmitted by droplets is significantly higher among the group mates of the infectious individuals compared to the rest of the population, while it does not become much higher via fomites.

Deriving from the current finding, it can be expected that when the group grows bigger, the number of total infections may increase, but the majority of infections will stay in the groups of the infectious individuals and the rest of the population will have a even smaller infection probability.

7 SENSITIVITY ANALYSIS

To explore how the SARS-CoV-2 transmission scale of the event changes under different combinations of activity space scenarios, a sensitivity analysis is conducted. This analysis quantifies the variations in the impacts of infrastructure settings and interventions, including queue distance, indoor/outdoor setting, density, and respiratory activities, the findings of which may help assist the policy makers with major decisions regarding COVID-19 related regulations concerning large events.

This chapter presents the design of 10 event scenarios in Section 7.1. The simulation results are summarized and discussed in Section 7.2.

7.1 SCENARIO DESIGN

10 scenarios are designed to explore to which extent the three major variables, namely, queue distance, indoor/outdoor setting, density, and respiratory activities, influence the infection scale at large events. The design of the scenario combinations is illustrated in Figure 7.1. More detailed scenario settings are presented in Appendix A, Table A.1, **??**, Table A.3.

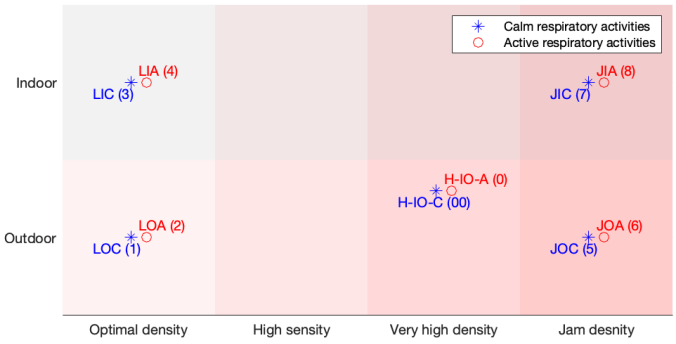


Figure 7.1: Design of event scenarios

As displayed in Figure 7.1, there are 5 groups of event scenarios, each representing a different indoor/outdoor setting and density level. Every group has two scenarios, representing a calm respiratory activity combination (20%talking + 80%breathing) and an active respiratory activity combination (20%talking + 40%singing + 40%breathing).

The scenarios are labeled by their characteristics. The labels each consists of three letter, representing the scenario design on the aspect of density, indoor/outdoor setting, and respiratory activity. For density, 'J', 'H', 'L' represent the jam density ($5.76p/m^2$), the very high density ($2p/m^2$), and the optimal density (low density: $0.4p/m^2$). The indoor/outdoor settings are represented by 'I' and 'O' and respiratory activity by 'A' for active and 'C' for calm.

Scenario 0 is the 'standard' scenario analyzed in Chapter 6. As it includes one indoor music stage and 7 outdoor stages, it is labeled by H-IO-A. Scenario 00 is the calm version of the 'standard' scenario, labeled by H-IO-C. Scenario 1&2 represent the optimal infrastructure combination, i.e., the most strict intervention, labeled by

LOC and LOA. All the event facilities are placed outdoor and the distance between individuals is kept as 1.5m. Scenario 3&4 are designed for optimal distance (1.5m) but indoor situations, labeled by LIC and LIA. Comparing with LOC and LOA, the difference of indoor and outdoor setting under low density and two types of respiratory activity combinations will be revealed. Scenario 5&6 are on the other side of the graph, involving the utmost jam density in outdoor facilities, labeled by JOC and JOA. Scenario 7&8 are expected to be the worst case scenarios, labeled by JIC and JIA, as the highest density and indoor settings come together.

7.2 RESULT COMPARISON

The resulting statistics under the 10 designed scenarios are summarized in the following figures. Figure 7.2 displays the number of infections generated in every scenario. Figure 7.3 illustrates the infection probability curves of all scenarios, of which the highest 100 values are included for the convenience and clarity of display.

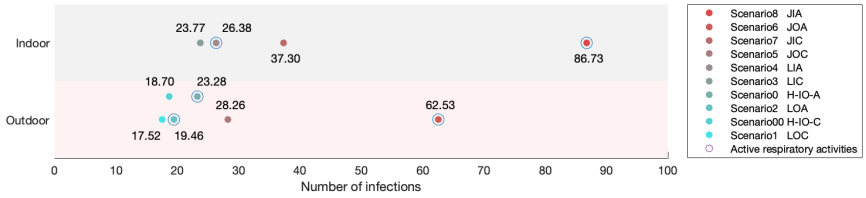


Figure 7.2: Number of infections in different scenarios

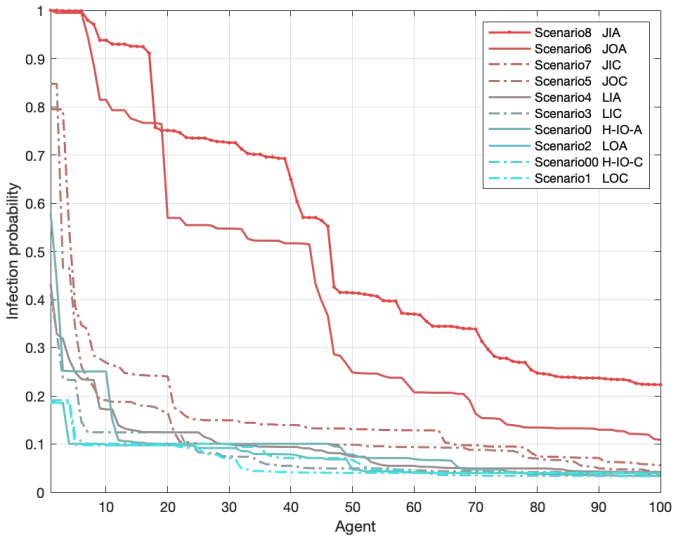


Figure 7.3: The 100 highest infection probability in different scenarios (legend ranking based on the number of infections)

As is shown in Figure 7.2, scenario 8 (JIA) has the largest number of infections, 86.73, followed by scenario 6 (JOA), 62.53, scenario 7 (JIC), 37.3, and scenario 5 (JOC), 28.26. The difference between scenario 5 (JOC) and other scenarios are much smaller. Scenario 3 (LIC) and 4 (LIA) have infections numbers slightly higher than scenario 00 (H-IO-C) and 0 (H-IO-A). Scenario 1 (LOC) is as expected the least risky scenario considering the total infection number while scenario 2 (LOA) has an infection number that is only slightly higher than scenario 00 (H-IO-C).

Looking at scenarios in the group of calm respiratory activities, the expected trend of high density and indoor setting leading to more infections is discovered. The same trend is displayed in the group of active respiratory activities. However, the increase of density does not force a linear increase in the number of infections. In an outdoor setting, when the density rises from $0.4p/m^2$ (scenario 1 (LOC) and 2 (LOA)) to $2p/m^2$ (scenario 00 (H-IO-C) and 0 (H-IO-A)), even though one of the music stages is set indoor in scenario 00 (H-IO-C) and 0 (H-IO-A), the number of infections only increases by 1.18 (6.7%) and 3.82 (19.7%) under calm and active respiratory activities. When the density goes up to $5.76p/m^2$ (scenario 5 (JOC) and 6 (JOA)), the number of infections increases by 10.70 (61.3%) and 43.08 (221.4%) compared to scenario 1 (LOC) and 2 (LOA) ($0.4p/m^2$), and 9.52 (51.2%) and 39.26 (168.6%) compared to scenario (H-IO-C) and 0 (H-IO-A) ($2p/m^2$). It can be concluded that when the density goes above $2p/m^2$, increasing density has an much bigger impact on the virus transmission scale compared to when the density is below $2p/m^2$. For the influence of respiratory activity combinations, it is observed that when more singing is involved, the impact of increasing density is also much higher than when agents are more silent and mostly just breath.

In Figure 7.2, comparing the line of outdoor scenarios and the line of indoor scenarios, it is observed that when the event is held in completely indoor facilities, the number of infections goes up by 35.7% (scenario 1 (LOC) to 3 (LIC)), 35.6% (scenario 2 (LOA) to 4 (LIA)), 32% (scenario 5 (JOC) to 7 (JIC)), 38.7% (scenario 6 (JOA) to 8 (JIA)), which indicates a stable impact under two different density settings and two different respiratory activity combinations.

Figure 7.3 displays the trend in the infection probability among the agents with the 100 highest infection probabilities. A clear split is observed between scenario 6 (JOA) and 8 (JIA) and other scenarios. The most noticeable difference is the shape of the curves. For scenario 6 (JOA) and 8 (JIA), the infection probability declines gradually from 100%, while for other scenarios, there is an immediate drop of infection probability to around 10% to 30% within the 10 agents with highest infection probabilities. In general, scenarios with the highest density (scenario 5, 6, 7, 8, i.e., JOC, JOA, JIC, JIA) display higher infection probabilities than scenarios with lower density. Among the lower density group, scenario 1 (LOC) and 00 (H-IO-C) with calm respiratory activities, and scenario 2 (LOA) with the lowest density and active respiratory activities have the lowest infection probabilities, with the maximum reaching below 20%. The in between scenarios have a higher maximum infection probability, with a sudden decline within 10 to 20 agents, after which the infection probability begins to steadily slowly decreasing around 15% to 3%.

Comparing Figure 7.2 and Figure 7.3, it is expected that the difference of infection probability curves between scenario 6 (JOA) and 8 (JIA) and other scenarios lead to the huge difference in infection numbers. As in other scenarios, it is relatively rare for agents to reach an infection probability higher than 15%, while for scenario 6 (JOA) and 8 (JIA), there is no sudden decrease in the infection probability , which indicates much more people are exposed to a high infection probability, making the total infection number higher than other scenarios.

Figure 7.4 demonstrates the distribution of total amount of virus transmitted throughout the entire event in different activity spaces.

It is observed that in all scenarios, the virus transmitted in entrance/exit queues, small bars and music stages makes up for the majority of the total virus exposure, which corresponds to the findings from Chapter 6. The change of the proportion of virus transmitted in the music stages is significant. In scenario 00, 1, and 3 (i.e., H-IO-C, LOC, and LIC), where agents follow calm respiratory activities, music stages only contributes around 1% of the total amount of virus, while the proportion goes up to around 10% to 19% in scenario 0, 2, 4 (i.e., H-IO-A, LOA, and LIA), which have the same infrastructure setting as scenario 00, 1, and 3 (H-IO-C, LOC, and LIC)), only with active respiratory activities. When the density increase to $5.76p/m^2$, the music stage takes a much share in the total amount of transmitted virus, especially

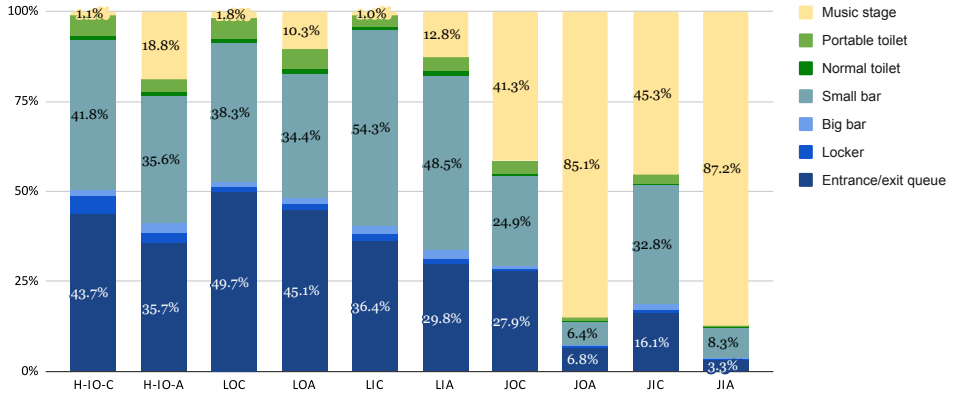


Figure 7.4: Distribution of transmitted virus in different activity spaces

under active respiratory activities, where the proportion of virus transmitted in music stages grows from above 40% (scenario 5 (JOC) and 7 (JIC)) to more than 80% (scenario 6 (JOA) and 8 (JIA)). With the increase of the music stage's share, the proportion taken by entrance/exit queues and small bars goes down accordingly. This does not mean that the virus transmitted in these facilities becomes less. On the contrary, the infection risk at the entrance/exit queues and small bars are relatively stable in either indoor or outdoor setting, as increasing queue distance from 1m to 1.5m does not make a great difference in the virus transmission scale within the facilities. It can be concluded that the increase of density from $2p/m^2$ to $5.76p/m^2$ significantly enlarges the virus transmitted in the music stages. Under scenarios with active respiratory activity, music stages also have a bigger contribution to the total amount of virus, which poses a even bigger impact when the density reaches the maximum.

Figure 7.5 demonstrate the distribution of total amount of virus transmitted throughout the entire event via different routes.

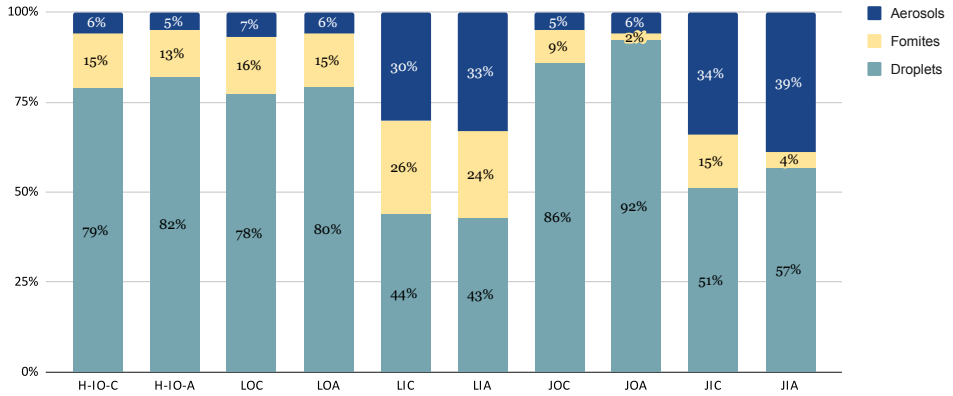


Figure 7.5: Distribution of transmitted virus via different routes

As is displayed, for scenarios with most facilities outdoor, droplets transmission takes up most of the transmitted virus and the contribution of aerosols is steadily around 6%. For scenarios with completely indoor settings, aerosols take up more than 30% of the transmitted virus and the contribution of fomites also increases from around 15% (scenario 1 & 2, LOC & LOA) to around 25% (scenario 3 & 4, LIC & LIA) and from 9% and 2% (scenario 5 & 6, JOC & JOA) to 15% and 4% (scenario 7 & 8, JIC & JIA). When the density is high, droplets play a more important role compared to low-density scenarios and active respiratory activities make the proportion of droplets transmission even bigger. It can be concluded that indoor

scenarios increase the amount of virus transmitted by aerosols and fomites and the increase of density and active respiratory activities enlarge the contribution of droplets transmission.

8 DISCUSSIONS

This chapter discusses the implications and limitations of the results and sensitivity analysis, as well as the limitations of the proposed method and the method application.

8.1 RESULT INTERPRETATIONS

As is summarized in Chapter 6, the general trend at all the activity spaces is identified that the transmission of SARS-CoV-2 is limited when the facility is located outdoor, the queue distance is increased (from 1m to 1.5m), the density is lowered (from $5.76p/m^2$ to $0.4p/m^2$), and the respiratory activities are calmer (from 20%talking + 40%singing + 40%breathing to 20%talking + 80%breathing).

Among these four variables, the impact of increasing the queue distance from 1m to 1.5m is the smallest, the reason of which can be derived from the findings at the music stage, where people keep 1.5m distance under $0.4p/m^2$ density and 1m distance under $1p/m^2$ density.

It is observed that the virus transmission scale varies much more significantly when increasing the density from $1p/m^2$ to $2p/m^2$ and $5.76p/m^2$ compared to when increasing the density from $0.4p/m^2$ to $1p/m^2$, which indicates the capability of SARS-CoV-2 to transmit via airborne routes decays to a turning point at around 1m distance. Before reaching this point, interventions that increase the distance between people have a strong impact on infection prevention, while after reaching this point, the impact is much smaller when the distance becomes longer. This finding implies that great risks lie under the type of events where fixed seating is not required and visitors tend to stand close to each other (within 1m distance), which is commonly observed at music festivals.

Moving indoor events to an outdoor environment also has a positive influence on limiting the spread of SARS-CoV-2, halving the total amount of virus transmitted in the facility, as the viral-laden particles in aerosols and fomites decay much faster in an outdoor and UV-exposed environment. Therefore, events held at indoor facilities can reduce the SARS-CoV-2 transmission risk by introducing outdoor venues.

Performing calm respiratory activities can also significantly reduce the virus transmission scale. When replacing 40% time spent on singing with breathing, the amount of virus picked up by individuals goes down by more than 85% in a 20-minute-duration 1000-participant music concert. The impact is even stronger in outdoor music stages, which indicates singing affects droplet transmission more than aerosol transmission. As a result, discouraging singing at events may significantly reduce the transmission scale of SARS-CoV-2.

It is also discovered that under the same distance between people, when staying for the same duration of time in the same facility, the queues, where people move to take up the place of the person in front of them, may be more risky than the music stages, where people stand close to each other but do not move. It is because the virus diffused in the environment is of the highest concentration at the location where the infectious individual stays. When moving in a queue, people who stand in the queue after the infectious individuals constantly move into the space of the highest virus concentration. As a result, there is a great probability

that they catch a relatively large amount of virus. This finding justifies the reason why contact tracing is not used in this study, that is the risk of SARS-CoV-2 transmission lies beyond close contact. People can pick up virus by staying at the same location, several minutes or hours after the infectious individuals have left, without having any contact with them. Fomites transmission also contributes to the infection probability, which cannot be represented by contact tracing. Therefore, contact tracing would not provide an accurate reflection of the SARS-CoV-2 transmission risk at the simulated event. Moreover, this finding suggests that compared to limiting the distance between people, it may be more effective to limit the scenarios where visitors take up other visitors' positions, such as forming long queues.

The sensitivity analysis indicates a large variability in the actual number of infections based on the context in which the pedestrian crowd resides. The infection numbers range from 17.52 in the optimal scenario to 86.73 in the least expected scenario, out of 10000 participants in the event, which is 0.18% to 0.87% of the total susceptible population. This indicates that the context has a very large impact on the number of infections, and as such the risk of virus spread at events. It also concludes the impact of the infrastructure settings and respiratory activities on the transmission scale and the variation in the virus proportion transmitted in different types of activity spaces. For instance, the share of virus transmitted at music stages increases from 1% under scenarios with low density and calm respiratory activities (LOC, LIC, H-IO-C) to 85% under the jam density and active respiratory activities (JOA and JIA), which indicates their great impact on the transmission scale of the entire event.

8.2 RESULT IMPLICATIONS

It is difficult to validate the case study results with real-life data, due to the following reasons caused by the nature of this research.

1. A large number of parameters are involved in the simulation process, some of which are based on assumptions and are difficult to measure in real life. For instance, respiratory activities and number of initial infectious visitor have a great influence on the transmission scale but are difficult to observe in real life. Validating the simulation results of inaccurately estimated parameters with the real-life experimental results may lead to inaccurate implications.
2. Organizing experimental large events to expose people in a potentially high-risk environment for infection can be controversial on the ethical aspect, which makes it difficult to obtain the real-life virus transmission data.

Therefore, instead of validation, the results and findings are critically compared to the Dutch government policies and recommendations and some of the real-life infection data obtained from the experimental events.

8.2.1 Implications on policies

In general, the advice from the OMT, which will be discussed in the next paragraph, corresponds to the findings of this research. OMT is a group of specialists and experts with different backgrounds and knowledge about COVID-19, convened by the Dutch National Institute for Public Health and the Environment (RIVM) to discuss how to control the COVID-19 outbreak based on the latest information on the transmission scale in the population, their professional expertise and the available scientific literature.

In their advisory reports, OMT has given advice on keeping social distance, strictly limiting the scale of indoor activities/events, limiting the scale of outdoor activities/events (bigger scale than indoor), limiting the duration of festivals, applying

fixed seating with low densities, requiring negative test results, and discouraging festival attendees from singing [RIVM, 2021b; NLtimes, 2021]. The findings of this research justify their suggestions. This study shows that these measures limit initial infectious individuals, limit exposure time, and help reduce the infection probability of event participants.

Not all the proposed measures have been implemented due to difficulties implementing and/or enforcing the measures. For instance, as is discovered by this research and suggest by OMT, singing significantly rises the scale of virus transmission, it is not feasible and not enforceable at large events [NLtimes, 2021]. With active respiratory activities, the transmission scale of SARS-CoV-2 can increase significantly. With longer intervals allowed for tests, the possibility of people getting infected in the time frame between the test and walking into the event terrain goes up, which results in a higher probability of infectious people spreading the virus at the event.

To better assist the policy makers to adopt feasible measures that also guarantee the safety of the population, it is suggested that when adopting new measures in the future, their risks and uncertainties need to be assessed under consideration of the actual expected behaviour of the crowd. Models, such as the one developed in this thesis, can support policy makers in quantifying the impact of their policies and identifying the risk of super spreading events.

8.2.2 Implications on real-life infection data

FieldlabEvenementen [2021] has conducted 4 types of experiments during February and March, 2021, including indoor events with static visitors, indoor events with moving visitors, outdoor events with moving visitors, and outdoor events with freely moving public. Interventions, such as compulsory social distancing, have been applied to all events, except the last category, which features all kinds of festivals. On the 20th and 21st of March, a 9-hour dance festival and a 9-hour rock festival were held in Walibi, Biddinghuizen, each with less than 2000 visitors. With all visitors tested negative within 48 hours before the festival, the infection probabilities of these two festivals are 0.072% (dance) and 0.032% (rock) per hour. Dividing the simulated infection probabilities of this study by 11 hours of the event duration, the probability in outdoor scenarios ranges from 0.016% to 0.057% per hour. The maximum simulated probability lies in between the experimental infection probabilities, which indicates a large variation in the real-life virus transmission scale and a possible underestimation under the least expected simulated scenario, with the density reaching the jam level and people spending 40% of time singing.

After a two-day outdoor festival held in Utrecht in July 2021 that attracted approximately 20000 people, 1050 people were tested positive of COVID-19 Ellyatt [2021]. Assuming 10000 people attended the festival on each day, from which 525 people were infected, the infection scale of this event is much higher than the simulated worst case scenario for outdoor facilities.

Following are various potential reasons for the underestimation of this research, which in themselves also hold important lessons pertaining our ability to model their behaviour and quantify the impact of measures that limit virus spread.

1. Different event types and infrastructure settings may lead to a different virus transmission scale. In the simulated event, there are in total 77 facilities, including 20 entrance/exit queues, 12 lockers, 4 big bars with seating area, 19 small bars with queues, 6 normal toilets, 8 portable toilets, and 8 music stages to guarantee that the pedestrian flow is well managed, in which case people spend limited time in queues and all kinds of facilities. While in real-life events, there might be limited number of facilities due to the limitations of the event location, leading to longer time spent in the queue, smaller-distance

headways between people (higher densities), which can significantly increase the infection risks.

2. In the simulated event, the time that visitors are exposed to infection risk might be underestimated. When simulating the activity schedule of individuals, the time spent at each facilities is strictly based on the activity pattern obtained from a real-life festival, which may result in underestimation of the activity duration. In addition, it is assumed that the time spent in between facilities are evenly distributed, during which there is no risk of exposure. In real-life event, people may spend more time in facilities and still be exposed to infection risk in between facilities, which leads to higher infection risks.

3. In this research, the smallest queue distance is 1m, as it is assumed that in the simulated event, visitors, to some extent, stick to the social distancing regulation. Therefore, in case of the queue, the transmission risk may be underestimated due to the assumed long distances, as the number of people within the most dangerous sphere of the influence of the infectious individual is underestimated.

4. The heterogeneity in the emission rate during real-life events can lead to a bigger transmission scale. In the simulated event, it is assumed that everyone performs the same respiratory activities within an activity space. While in real-life events, people may have very different respiratory activities. For instance, as is frequently observed at festivals, at the music stages, there may be a group of very passionate music lovers who would loudly sing along during a performance. The respiratory activities of this group can be much more active than the rest of the visitors, which may lead to a bigger virus transmission scale in this group.

Moreover, limitations exist in both real-life experiments and the proposed simulated method and therefore, direct comparisons have very limited implications. For instance, the number of infectious individuals in the events remains unknown as people participating in the events all have been tested negative. As the number of infectious individuals is one of the most important influence factors of the event infection probability, it directly impacts the amount of virus spread in the environment.

In a nutshell, although this method application may underestimate the virus spread scale at the simulated event, it is capable of showing when and where the major risks occur. By comparing the results of different scenarios, it also gives indications on crowd management measures and interventions that can help reduce the virus transmission scale.

The limitations of the proposed method and its application are discussed in the following section.

8.3 METHOD LIMITATIONS

This section discusses the major limitations of the proposed method and its application, including the lack of consideration of group behavior and the heavy dependency on detailed virus transmission related parameters and activity pattern assumptions.

8.3.1 Group behavior

In this study, pedestrians are considered as independent individuals, whose behavior results from the joint force of all agents and obstacles in the vicinity. No bonds are formed within a group of people, as all the agents are considered the same to

the moving pedestrian. However, in real-life observation, people, especially at large events, usually move in a cluster with a small group of people, the distance among whom is usually kept relatively close [Qiu and Hu, 2010].

Group behavior is different to capture and model, yet it forms an important part of crowd dynamics and may have a huge influence on the virus spread simulation. People in a group usually cohere to each other during the entire event. They usually have close physical contact, which increases the risk of virus transmission by fomites. The time they spent in close contact is also much longer than random encounters at different activity spaces, which leads to higher risks of droplets and aerosols transmission.

As is demonstrated in the results of the simplified 'group' scenario, if one infectious person stays close to other members throughout the event, people in the group usually have a high exposure. On the other hand, people in other groups may have a lower infection probability, as it is less likely for them to have close contact with infectious people in their own groups.

The purpose of the simulated 'group' scenario is mainly to illustrate possible outcomes if infectious individuals stay close to their group members. In fact, modeling group behavior is more sophisticated than the demonstration. The lack of consideration of the group behavior of susceptible individuals may lead to an overestimation of infection risk, as the social forces between groups are expected to be bigger than the forces between individual pedestrians. On the other hand, due to the group behavior, people would tend to stay close to their group members even in routes connecting major activity spaces and vast open ground where they are able to keep social distance, ignoring which may lead to an underestimation of infection risk within the group, as people stay in close contact throughout the entire event.

To conclude, for a more realistic simulation of the infection risk at events, group behavior must be considered in the process of activity schedule generation and pedestrian route choice and movement modeling.

8.3.2 Parameter and activity pattern requirement

As is introduced in Chapter 5, this proposed method is heavily dependent on detailed virus transmission related parameters. On the one hand, it gives solid theoretical support to the results. On the other hand, varying the parameter values based on different assumptions of a scenario can give very different results, which might be potentially misleading for the risk evaluation of the simulated event.

The same problem applies to activity patterns. For instance, as observed in the case study application results, facilities with queues are the most risky among all types of activity spaces. One of the decisive factors of transmission scale in queues is the queuing time, which determines how long people spend in the queue transmitting and picking up virus. However, as no data is available for how long people spend on queuing to get into a facility, this study assumes a whole process of queue building in a designed facility, during which the first half is spent building the queue, the second half is used to resolve the queue. To make sure the simulation time covers all possible virus transmission, only agents who enter the facility as the first 5% to 15% of the population are assumed to be infectious, which does not give an indication of how long an average visitor spend in the queue. Under this assumption, it can be expected that if the infectious individual enters when the queue has accumulated longer and people would stay longer in the queue, the transmission scale will be even bigger.

Therefore, making sure the parameters are validated under the simulated scenarios and acquiring detailed real life activity pattern data is of great importance for the proposed method to give reliable results.

8.4 METHOD IMPLICATIONS

This proposed method provides a quantifying approach to evaluate the infection risk of SARS-CoV-2 in large events. Despite its limitations discussed above, it has generated promising results in a large music festival simulation.

Compared to existing studies on the transmission of SARS-CoV-2, this study has achieved a breakthrough in simulating the infection risk of visiting large-scale events, where multiple activity spaces all potentially contribute to the virus transmission at the same time.

In the case study application, the same phenomena as the existing studies have been identified, including longer distance between people, outdoor scenarios with high air change rates and UV exposure, low density, and calm respiratory activities leading to a lower infection risk, vice versa. Yet, it sheds new light on the understanding of infection risk in different environment. For instance, by comparing simulation results from queues and music stages, it is discovered that taking up an infectious individual's position might be more risky compared to standing next to the infectious individual. It also quantifies the influence of queue distance, indoor/outdoor setting, density, and respiratory activities on the transmission scale of SARS-CoV-2, providing a reference to COVID-19 crowd management at large events.

The main purpose of the proposed method is to quantify the infection risk of SARS-CoV-2 in large events, by combining activity scheduling, route choice and movement modeling, and virus spread simulation. Through the case study application, it has been proven to be capable of revealing the general infection risk and identifying risk-prone areas. Therefore, this method can be used as a risk evaluation tool for big events, comparing the potential infection risk under different scenarios (infrastructure design, physical intervention) and spotting where the most risk lies. According to the results, crowd management approaches and interventions can be used to reduce the infection risk at the simulated event, making events feasible and safe again.

9 CONCLUSIONS

This chapter presents the conclusions, including the answer to the main research question and recommendations for future research.

9.1 ANSWER TO MAIN RESEARCH QUESTION

The research objective of this study is to develop a SARS-CoV-2 transmission risk analysis method at events by modelling crowd interactions at different types of event spaces and quantifying the SARS-CoV-2 transmission risks in the process. The main research question that will be answered in this section is:

How to model SARS-CoV-2 transmission risks based on pedestrian behavior and virus spread simulation at large events?

The proposed method provides a method to connect activity scheduling, pedestrian route choice and movement modeling, virus spread modeling, and infection risk identification to determine the SARS-CoV-2 transmission risks at large event by the process shown on Figure 3.1.

To make use of the proposed method, a few factors need to identified beforehand. First, after determining the type of event to evaluate and the general demand, the infrastructure types and activity spaces should be identified as the input information to the NOMAD model. Then the activity pattern data of such an event needs to be collected and analyzed, from which the activity schedules and demand pattern at activity spaces can be derived. The former will be used for NOMAD simulation and the latter for identification of virus locations and matching up the virus exposure on susceptible individuals.

With the above mentioned information, the proposed method determine the SARS-CoV-2 transmission risks in the following steps.

The first part of the proposed method is a tactical and operational level pedestrian model, NOMAD, which simulates pedestrian route choice and movement with the input of infrastructure layout, social force parameters, and demand pattern. With NOMAD, a number of activity spaces with potential SARS-CoV-2 transmission possibility can be simulated under different infrastructure layouts and physical interventions, such as 1.5m or 1m queue distance. NOMAD generates pedestrian trajectories with the accuracy of 0.1m at each 0.1-second time step.

Then the output of NOMAD is transformed into agent scripts which include the movement of pedestrians at the accuracy of 1m at each 20-second time step, the respiratory characteristics of the agents, whether wearing a face mask, the initial viral load, and other virus transmission related parameters. The second part of the proposed method is QVEmod, the model that simulates the virus spread in different activity spaces via 7 processes that the virus goes through, including emission, falling onto surfaces, decay, diffusion, inhalation, contaminating surfaces by touching, and being picked up from surfaces by touching. QVEmod generates the accumulated virus loads on agents via three routes, namely droplets, aerosols, and fomites.

The third part of the method include e smaller steps of activity scheduling, possible exposure time and location identification, adding up virus obtained from different activity spaces, and infection risk estimation. The first step generates every visitor's activity schedule at the event, identifying where and when the infectious

and susceptible individuals are located. The second step summaries the whereabouts of infectious individuals and derives the locations and duration of possible virus exposure. In the third step, the susceptible individuals' activity schedules are compared with the derived locations and duration. If a match exists, a virus exposure (obtained from the virus spread simulation) is assigned to this individual. For all the susceptible individuals, the virus exposure acquired throughout the activity schedule is added up. Finally, the accumulated virus exposure is fit into a dose-response model to calculate the general infection risk of visiting this event.

In the case study application, the proposed method has shown its capability of revealing the general infection risk and identifying risk-prone areas in an event. This method can be used as a risk evaluation tool for big events, comparing the potential infection risk under different scenarios (infrastructure design, physical intervention) and spotting where the most risk lies. As a result, the method can further assist decision making on crowd management approaches and interventions to be used to reduce the infection risk at the simulated event.

9.2 RECOMMENDATIONS

Recommendations for future research are given in this section, which are based on the method implications and limitations discussed in Chapter 8.

9.2.1 Recommendations on method development

As is discussed in Section 8.3, the major limitations of the proposed method lie in the lack of consideration of group behavior and the heavy dependency on validated parameters and detailed activity patterns derived from real life data.

Group behavior

It is recommended that group behavior to be taken into consideration during activity scheduling, where the activity schedules within a group vary from each other within a certain threshold. For instance, it can be applied that within each activity space, the difference of entry and exit time of different group members should not be more than 2 minutes.

The coherence of group members should also be considered outside the major activity spaces, as people in the same group usually tend to stay close to each other throughout the entire visit to the event. This can be achieved by simulating routes connecting different activity spaces as one of the potential risky activity spaces, where the movements of individuals are modeled and virus spread can be simulated.

To implement group behavior in the NOMAD model, a smaller social force can be applied within a group and a larger social force between different groups to make sure the group members cohere to each other and maybe keep a longer distance to people from other groups.

More literature research or surveys should be conducted to identify the average group size in different types of events, to assist the activity scheduling and NOMAD simulation.

Dependency on validated parameters and detailed activity pattern information

Dependency on accurate parameters and detailed activity pattern information lies in the nature of the proposed method, making it difficult to be validated and applied in difference scenarios.

However, it is the solid theoretical support and real-life data that makes the simulation results reliable and helpful for decision making. Therefore, for future re-

search, it is recommended that for each type of event that this method is used for, the virus spread related parameters are validated under the specific scenario. The activity pattern information should be obtained from a same type of event and analyzed thoroughly to support activity scheduling and derive input to NOMAD simulations.

9.2.2 Recommendations on method application

In the case study application of the proposed method, due to the limited time frame, incomplete data sets, and inaccurate parameter values, the result implications have been limited. For future applications, the following recommendations are identified:

1. The GPS trajectory data should be matched with specified activity space GPS locations. In this research, the exact locations of activity spaces has not been available. The trajectories matching with the activity spaces mainly depends on a simplified festival map and the cluster of trajectory points, which adds up the inaccuracy of identifying the activity schedules and the time people spend in different types of activity spaces. If the data allows, specified activity space GPS locations would help make the derived activity patterns more accurate.

2. The sample size of the GPS trajectory data should be as big as possible, to better capture the full picture of activity patterns at the events.

3. The time stamps of the GPS trajectory data should be as continuous as possible. In this study, data with intervals longer than 15 minutes was filtered out to balance the continuity and the size of the data. However, 15-minute interval may not be enough to capture short activities such as going to the toilet. It is recommended to use GPS tracking device that obtain continuous trajectories instead of only tracking the location when people use the application on the phone.

4. The time step of the QVEmod should be shorter if possible. With a 20-second time step, the movements of agents in small activity spaces might not be captured by QVEmod, which may lead to an underestimation of virus transmission scale in the facility.

5. QVEmod parameters, such as respiratory activity characteristics, need to validated under the simulated scenarios.

6. The choice of the influence factors can be more diverse. Due to time limitation, whether wearing face mask and the willingness to comply to social distance have not be taken into account as a design variable, yet they have an huge influence on the virus transmission. If more influence factors are considered, the risk evaluation may see huge variations in the resulting infection numbers and may assist crowd management in a different direction.

7. The activity schedule simulation can be more realistic if the distance between activity spaces are considered, instead of averaging the interval between visiting two activity spaces during the entire event.

A SCENARIO DESIGN

Activity space	Nr	Scenario 00	Scenario 0	Scenario 1	Scenario 2
Entrance/exit queue	20	Outdoor: ACH = 920 Queue distance: 1m	Outdoor: ACH = 920 Queue distance: 1m	Outdoor: ACH = 920 **Queue distance: 1.5m**	Outdoor: ACH = 920 **Queue distance: 1.5m**
Locker	12	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93
Big bar area	4	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93, $u_{fomites_{wood}}$ = 9.69	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93, $u_{fomites_{wood}}$ = 9.69	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93, $u_{fomites_{wood}}$ = 9.69	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93, $u_{fomites_{wood}}$ = 9.69
Small bar + queue	19	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93 Queue distance: 1m	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93 Queue distance: 1m	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93 **Queue distance: 1.5m**	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93 **Queue distance: 1.5m**
Normal toilet	6	Indoor: ACH = 3	Indoor: ACH = 3	Indoor: ACH = 3	Indoor: ACH = 3
Portable toilet + queue	8	Outdoor: ACH = 920 Queue distance: 1m	Outdoor: ACH = 920 Queue distance: 1m	Outdoor: ACH = 920 **Queue distance: 1.5m**	Outdoor: ACH = 920 **Queue distance: 1.5m**
Music stage	8	1/8: Indoor: ACH = 3 7/8: Outdoor: ACH = 368 density: $2p/m^2$ 20%talking + 80%breathing	1/8: Indoor: ACH = 3 7/8:Outdoor: ACH = 368 density: $2p/m^2$ **20%talking + 40%singing + 40%breathing**	8/8: Outdoor: ACH = 368 **density:** $0.4p/m^2$ 20%talking + 80%breathing	8/8: Outdoor: ACH = 368 **density:** $0.4p/m^2$ **20%talking + 40%singing + 40%breathing**

Table A.1: Design of event scenario 00 - 2

Activity space	Nr	Scenario 3	Scenario 4	Scenario 5	Scenario 6
Entrance/exit queue	20	Indoor: ACH = 3 Queue distance: 1.5m	Indoor: ACH = 3 Queue distance: 1.5m	Outdoor: ACH = 920 **Queue distance: 1m**	Outdoor: ACH = 920 **Queue distance: 1m**
Locker	12	Indoor: ACH = 3 $u_{fomites_{plastic}}$ = 0.193	Indoor: ACH = 3 $u_{fomites_{plastic}}$ = 0.193	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93
Big bar area	4	Indoor: ACH = 3 $u_{fomites_{plastic}}$ = 0.193, $u_{fomites_{wood}}$ = 0.969	Indoor: ACH = 3 $u_{fomites_{plastic}}$ = 0.193, $u_{fomites_{wood}}$ = 0.969	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93, $u_{fomites_{wood}}$ = 9.69	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93, $u_{fomites_{wood}}$ = 9.69
Small bar + queue	19	Indoor: ACH = 3 $u_{fomites_{plastic}}$ = 0.193 Queue distance: 1.5m	Indoor: ACH = 3 $u_{fomites_{plastic}}$ = 0.193 Queue distance: 1.5m	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93 **Queue distance: 1m**	Outdoor: ACH = 920 $u_{fomites_{plastic}}$ = 1.93 **Queue distance: 1m**
Normal toilet	6	Indoor: ACH = 3	Indoor: ACH = 3	Indoor: ACH = 3	Indoor: ACH = 3
Portable toilet + queue	8	Indoor: ACH = 3 Queue distance: 1.5m	Indoor: ACH = 3 Queue distance: 1.5m	Outdoor: ACH = 920 **Queue distance: 1m**	Outdoor: ACH = 920 **Queue distance: 1m**
Music stage	8	8/8: Indoor: ACH = 3 density: $0.4p/m^2$ 20%talking + 80%breathing	8/8: Indoor: ACH = 3 density: $0.4p/m^2$ **20%talking + 40%singing + 40%breathing**	8/8: Outdoor: ACH = 368 **density:** $5.76p/m^2$ 20%talking + 80%breathing	8/8: Outdoor: ACH = 368 **density:** $5.76p/m^2$ **20%talking + 40%singing + 40%breathing**

Table A.2: Design of event scenario 3 - 6

Activity space	Nr	Scenario 7	Scenario 8
Entrance/exit queue	20	Indoor: ACH = 3 Queue distance: 1m	Indoor: ACH = 3 Queue distance: 1m
Locker	12	Indoor: ACH = 3 $u_{fomites_plastic}$ = 0.193	Indoor: ACH = 3 $u_{fomites_plastic}$ = 0.193
Big bar area	4	Indoor: ACH = 3 $u_{fomites_plastic}$ = 0.193, $u_{fomites_wood}$ = 0.969	Indoor: ACH = 3 $u_{fomites_plastic}$ = 0.193, $u_{fomites_wood}$ = 0.969
Small bar + queue	19	Indoor: ACH = 3 $u_{fomites_plastic}$ = 0.193 Queue distance: 1m	Indoor: ACH = 3 $u_{fomites_plastic}$ = 0.193 Queue distance: 1m
Normal toilet	6	Indoor: ACH = 3	Indoor: ACH = 3
Portable toilet + queue	8	Indoor: ACH = 3 Queue distance: 1m	Indoor: ACH = 3 Queue distance: 1m
Music stage	8	8/8: Indoor: ACH = 3 density: $5.76 p/m^2$ 20%talking + 80%breathing	8/8: Indoor: ACH = 3 density: $5.76 p/m^2$ **20%talking + 40%singing + 40%breathing**

Table A.3: Design of event scenario 7 - 8

BIBLIOGRAPHY

Adam, D. C., Wu, P., Wong, J. Y., Lau, E. H., Tsang, T. K., Cauchemez, S., Leung, G. M., and Cowling, B. J. (2020). Clustering and superspreading potential of sars-cov-2 infections in hong kong. *Nature Medicine*, 26(11):1714–1719.

Aloulou, F. (2018). The application of discrete choice models in transport. *Statistics: Growing Data Sets and Growing Demand for Statistics*, page 85.

AMT, A. (2020). Information on entry restrictions and quarantine regulations in germany.

Antonini, G., Bierlaire, M., and Weber, M. (2006). Discrete choice models of pedestrian walking behavior. *Transportation Research Part B: Methodological*, 40(8):667–687.

Arav, Y., Klausner, Z., and Fattal, E. (2020). Understanding the indoor presymptomatic transmission mechanism of covid-19. *medRxiv*.

Barros, J., Araujo, M., and Rossetti, R. J. (2015). Short-term real-time traffic prediction methods: A survey. In *2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 132–139. IEEE.

Bernardi, N. F., Snow, S., Peretz, I., Perez, H. O., Sabet-Kassouf, N., and Lehmann, A. (2017). Cardiorespiratory optimization during improvised singing and toning. *Scientific reports*, 7(1):1–8.

Blue, V. J. and Adler, J. L. (2001). Cellular automata microsimulation for modeling bidirectional pedestrian walkways. *Transportation Research Part B: Methodological*, 35(3):293–312.

Bluman, A. G. (2009). *Elementary statistics: A step by step approach*. McGraw-Hill Higher Education New York;.

Bouchnita, A. and Jebrane, A. (2020). A hybrid multi-scale model of covid-19 transmission dynamics to assess the potential of non-pharmaceutical interventions. *Chaos, Solitons & Fractals*, 138:109941.

Campanella, M. C. (2016). Microscopic modelling of walking behaviour.

Carvalho, F. R., Henriques, D. V., Correia, O., and Schmalwieser, A. W. (2021). Potential of solar uv radiation for inactivation of coronaviridae family estimated from satellite data. *Photochemistry and Photobiology*, 97(1):213.

Chen, P. Z., Bobrovitz, N., Premji, Z., Koopmans, M., Fisman, D. N., and Gu, F. X. (2021). Heterogeneity in transmissibility and shedding sars-cov-2 via droplets and aerosols. *Elife*, 10:e65774.

Chin, A., Chu, J., Perera, M., Hui, K., Yen, H.-L., Chan, M., Peiris, M., and Poon, L. (2020). Stability of sars-cov-2 in different environmental conditions. *MedRxiv*.

CIRES (2020). Covid-19 airborne transmission tool available.

Crump, K., Hoel, D., Langley, C., and Peto, R. (1976). Fundamental carcinogenic processes and their implications for low dose risk assessment. *Cancer research*, 36(9 Part 1):2973–2979.

Daamen, W. (2002). Simped: a pedestrian simulation tool for large pedestrian areas. In *Conference Proceedings EuroSIW*, pages 24–26. Simulation Interoperability Standards Organization Orlando, Fla.

Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation research part B: methodological*, 28(4):269–287.

de Dios Ortúzar, J. and Willumsen, L. G. (2011). *Modelling transport*. John wiley & sons.

de Vrieze, A. (2021). Fieldlab presenteert eindresultaat: evenementen onder voorwaarden mogelijk.

Dekker, T., Hess, S., Arentze, T., and Chorus, C. (2014). Incorporating needs-satisfaction in a discrete choice model of leisure activities. *Journal of Transport Geography*, 38:66–74.

Deng, W., Bao, L., Gao, H., Xiang, Z., Qu, Y., Song, Z., Gong, S., Liu, J., Liu, J., Yu, P., et al. (2020). Ocular conjunctival inoculation of sars-cov-2 can cause mild covid-19 in rhesus macaques. *Nature communications*, 11(1):1–7.

Desyllas, J., Duxbury, E., Ward, J., and Smith, A. (2003). Pedestrian demand modelling of large cities: an applied example from london.

Du, B., Zhang, C., Shen, J., and Zheng, Z. (2021). A dynamic sensitivity model for unidirectional pedestrian flow with overtaking behaviour and its application on social distancing's impact during covid-19. *IEEE Transactions on Intelligent Transportation Systems*.

Duives, D., Chang, Y., Sparnaaij, M., Wouda, B., Boschma, D., Liu, Y., Yuan, Y., Daamen, W., de Jong, M., Teberg, C., Schachtschneider, K., Sikkema, R., van Veen, L., and ten Bosch, Q. (2021). The multi-dimensional challenges of controlling sars-cov-2 transmission in indoor spaces: Insights from the linkage of a microscopic pedestrian simulation and virus transmission models.

Duives, D. C., Daamen, W., and Hoogendoorn, S. P. (2015). Quantification of the level of crowdedness for pedestrian movements. *Physica A: Statistical Mechanics and its Applications*, 427:162–180.

Ellyatt, H. (2021). Music festival in the netherlands leads to over 1,000 covid infections.

Fears, A. C., Klimstra, W. B., Duprex, P., Hartman, A., Weaver, S. C., Plante, K. S., Mirchandani, D., Plante, J. A., Aguilar, P. V., Fernández, D., et al. (2020). Persistence of severe acute respiratory syndrome coronavirus 2 in aerosol suspensions. *Emerging infectious diseases*, 26(9):2168.

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., and Fraser, C. (2020). Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing. *Science*, 368(6491).

FieldlabEvenementen (2021). Adviesaanvraag type iv – versie 1.0.

Fraser, C., Donnelly, C. A., Cauchemez, S., Hanage, W. P., Van Kerkhove, M. D., Hollingsworth, T. D., Griffin, J., Baggaley, R. F., Jenkins, H. E., Lyons, E. J., et al. (2009). Pandemic potential of a strain of influenza a (h1n1): early findings. *science*, 324(5934):1557–1561.

Gao, C. X., Li, Y., Wei, J., Cotton, S., Hamilton, M., Wang, L., and Cowling, B. J. (2021). Multi-route respiratory infection: when a transmission route may dominate. *Science of the Total Environment*, 752:141856.

Gobbler, G. (2020). Portable toilet.

Hall, R. (2012). *Handbook of transportation science*, volume 23. Springer Science & Business Media.

Hallett, S., Toro, F., and Ashurst, J. V. (2020). Physiology, tidal volume. *StatPearls [Internet]*.

Hankin, B. and Wright, R. A. (1958). Passenger flow in subways. *Journal of the Operational Research Society*, 9(2):81–88.

Hänseler, F. S., Lam, W. H., Bierlaire, M., Lederrey, G., and Nikolić, M. (2017). A dynamic network loading model for anisotropic and congested pedestrian flows. *Transportation Research Part B: Methodological*, 95:149–168.

Harrison, A. G., Lin, T., and Wang, P. (2020). Mechanisms of sars-cov-2 transmission and pathogenesis. *Trends in immunology*.

Helbing, D. and Molnar, P. (1995). Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282.

Hinds, W. C. (1999). *Aerosol technology: properties, behavior, and measurement of airborne particles*. John Wiley & Sons.

Hoogendoorn, S. P. and Bovy, P. H. (2004). Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2):169–190.

Huang, H.-J. and Lam, W. (2002). Quasi-continuous dynamic equilibrium assignment with departure time choice in congested unidirectional pedestrian networks. *Journal of the Operational Research Society*, 53(1):97–107.

Huijkman, I. (2019). Amsterdam open air 2019 aoa vlog.

Iltanen, S. (2012). Cellular automata in urban spatial modelling. In *Agent-based models of geographical systems*, pages 69–84. Springer.

Jitsuk, N. C., Suttirat, P., Modchang, C., et al. (2020). Effect of the songkran festival on covid-19 transmission in thailand. *Asian Pacific Journal of Tropical Medicine*, 13(7):331.

Johansson, A. and Kretz, T. (2012). Applied pedestrian modeling. In *Agent-based models of geographical systems*, pages 451–462. Springer.

Julian, T., Leckie, J., and Boehm, A. (2010). Virus transfer between fingerpads and fomites. *Journal of applied microbiology*, 109(6):1868–1874.

Kaparounaki, C. K., Patsali, M. E., Mousa, D.-P. V., Papadopoulou, E. V., Papadopoulou, K. K., and Fountoulakis, K. N. (2020). University students' mental health amidst the covid-19 quarantine in greece. *Psychiatry research*, 290:113111.

Karamouzas, I. and Overmars, M. (2010). A velocity-based approach for simulating human collision avoidance. In *International Conference on Intelligent Virtual Agents*, pages 180–186. Springer.

Karlström, A., Waddell, P., and Fox, D. (2009). Scaling up the microeconomic dynamic discrete choice model of activity-based scheduling. In *Proc. Eur. Transport Conf*. Citeseer.

KNMI (2021). Klimaatviewer.

Kraay, A. N. M., Hayashi, M. A., Berendes, D. M., Sobolik, J. S., Leon, J. S., and Lopman, B. A. (2020). Risk of fomite-mediated transmission of sars-cov-2 in child daycares, schools, and offices: a modeling study. *MedRxiv*.

Lai, Y. and Kontokosta, C. E. (2018). Quantifying place: Analyzing the drivers of pedestrian activity in dense urban environments. *Landscape and Urban Planning*, 180:166–178.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., Reich, N. G., and Lessler, J. (2020). The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, 172(9):577–582.

Lei, H., Xiao, S., Cowling, B. J., and Li, Y. (2020). Hand hygiene and surface cleaning should be paired for prevention of fomite transmission. *Indoor air*, 30(1):49–59.

Leung, N. H., Chu, D. K., Shiu, E. Y., Chan, K.-H., McDevitt, J. J., Hau, B. J., Yen, H.-L., Li, Y., Ip, D. K., Peiris, J. M., et al. (2020). Respiratory virus shedding in exhaled breath and efficacy of face masks. *Nature medicine*, 26(5):676–680.

Li, Y., Yu, R., Shahabi, C., and Liu, Y. (2017). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926*.

Liu, Y., Li, T., Deng, Y., Liu, S., Zhang, D., Li, H., Wang, X., Jia, L., Han, J., Bei, Z., et al. (2021). Stability of sars-cov-2 on environmental surfaces and in human excreta. *Journal of Hospital Infection*, 107:105–107.

Lockerlogic (2018). Charging storage lockers @ dancefestopia 2018.

Løvås, G. G. (1994). Modeling and simulation of pedestrian traffic flow. *Transportation research Part B: methodological*, 28(6):429–443.

Lue, G. and Miller, E. J. (2019). Estimating a toronto pedestrian route choice model using smartphone gps data. *Travel behaviour and society*, 14:34–42.

Ma, J., Qi, X., Chen, H., Li, X., Zhang, Z., Wang, H., Sun, L., Zhang, L., Guo, J., Morawska, L., et al. (2020). Coronavirus disease 2019 patients in earlier stages exhaled millions of severe acute respiratory syndrome coronavirus 2 per hour. *Clinical Infectious Diseases*.

Maital, S. (2020). The global economic impact of covid-19: A summary of research.

McKibbin, W. and Fernando, R. (2020). The economic impact of covid-19. *Economics in the Time of COVID-19*, 45.

Meyer-König, T., Klüpfel, H., and Schreckenberg, M. (2002). Assessment and analysis of evacuation processes on passenger ships by microscopic simulation. *Schreckenberg and Sharma [2]*, pages 297–302.

Morawska, L. and Cao, J. (2020). Airborne transmission of sars-cov-2: The world should face the reality. *Environment international*, 139:105730.

Nedelcheva, K. (2020). Dotted social distancing markers.

Nicas, M. (1996). An analytical framework for relating dose, risk, and incidence: an application to occupational tuberculosis infection. *Risk Analysis*, 16(4):527–538.

Nicas, M. and Sun, G. (2006). An integrated model of infection risk in a health-care environment. *Risk Analysis*, 26(4):1085–1096.

Nicastro, F., Sironi, G., Antonello, E., Bianco, A., Biasin, M., Brucato, J. R., Ermolli, I., Pareschi, G., Salvati, M., Tozzi, P., et al. (2021). Solar uv-b/a radiation is highly effective in inactivating sars-cov-2. *Scientific Reports*, 11(1):1–11.

Nikbakht, R., Baneshi, M. R., Bahrampour, A., and Hosseinnataj, A. (2019). Comparison of methods to estimate basic reproduction number (ro) of influenza, using canada 2009 and 2017-18 a (h1n1) data. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences*, 24.

NLtimes (2021). New covid rules for festivals less strict than experts advised.

OpenStreetMap (2021). Openstreetmap.

Pan, J., Harb, C., Leng, W., and Marr, L. C. (2021). Inward and outward effectiveness of cloth masks, a surgical mask, and a face shield. *Aerosol Science and Technology*, pages 1–16.

Paris, S., Pettré, J., and Donikian, S. (2007). Pedestrian reactive navigation for crowd simulation: a predictive approach. In *Computer Graphics Forum*, volume 26, pages 665–674. Wiley Online Library.

Parisi, D. R., Gilman, M., and Moldovan, H. (2009). A modification of the social force model can reproduce experimental data of pedestrian flows in normal conditions. *Physica A: Statistical Mechanics and its Applications*, 388(17):3600–3608.

Popa, A., Genger, J.-W., Nicholson, M. D., Penz, T., Schmid, D., Aberle, S. W., Agerer, B., Lercher, A., Endler, L., Colaço, H., et al. (2020). Genomic epidemiology of superspreading events in austria reveals mutational dynamics and transmission properties of sars-cov-2. *Science translational medicine*, 12(573).

Qiu, F. and Hu, X. (2010). Modeling group structures in pedestrian crowd simulation. *Simulation Modelling Practice and Theory*, 18(2):190–205.

Rahman, K., Ghani, N. A., Kamil, A. A., Mustafa, A., and Chowdhury, M. A. K. (2013). Modelling pedestrian travel time and the design of facilities: A queuing approach. *PloS one*, 8(5):e63503.

Raiteux, J., Eschlimann, M., Marangon, A., Rogée, S., Dadvisard, M., Taysse, L., and Larigauderie, G. (2021). Inactivation of sars-cov-2 by simulated sunlight on contaminated surfaces. *Microbiology Spectrum*, 9(1):e00333–21.

Ratcliff, M. A. (2018). Determination of laboratory airflow rates.

RIVM (2020). Generic framework for coronavirus measures.

RIVM (2021a). Confirmed cases: Coronavirus dashboard.

RIVM (2021b). Outbreak management team (omt).

Robin, T., Antonini, G., Bierlaire, M., and Cruz, J. (2009). Specification, estimation and validation of a pedestrian walking behavior model. *Transportation Research Part B: Methodological*, 43(1):36–56.

Rossi, R., Socci, V., Talevi, D., Mensi, S., Niolu, C., Pacitti, F., Di Marco, A., Rossi, A., Siracusano, A., and Di Lorenzo, G. (2020). Covid-19 pandemic and lockdown measures impact on mental health among the general population in italy. *Frontiers in psychiatry*, 11:790.

RTLNieuws (2021). Wat was het doel van een testevenement met 10.000 mensen? 'methode onduidelijk'.

Statista (2021). Number of festivals in the netherlands 2020.

Stubenschrott, M., Kogler, C., Matyus, T., and Seer, S. (2014). A dynamic pedestrian route choice model validated in a high density subway station. *Transportation Research Procedia*, 2:376–384.

Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U., Li, B., Cai, J., Xu, B., Yang, Q., et al. (2020). An investigation of transmission control measures during the first 50 days of the covid-19 epidemic in china. *Science*, 368(6491):638–642.

Ton, D., van den Heuvel, J., Daamen, W., and Hoogendoorn, S. (2015). Route and activity location choice behaviour of departing passengers in train stations. In *hEART (European Association for Research in Transportation) 2015 Conference. Copenhagen, Denmark*, pages 9–11.

Treuille, A., Cooper, S., and Popović, Z. (2006). Continuum crowds. *ACM Transactions on Graphics (TOG)*, 25(3):1160–1168.

Truong, L. T., Sarvi, M., Currie, G., and Garoni, T. M. (2015). How many simulation runs are required to achieve statistically confident results: a case study of simulation-based surrogate safety measures. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 274–278. IEEE.

Ueki, H., Furusawa, Y., Iwatsuki-Horimoto, K., Imai, M., Kabata, H., Nishimura, H., and Kawaoka, Y. (2020). Effectiveness of face masks in preventing airborne transmission of sars-cov-2. *MSphere*, 5(5).

University, A. J. H. (2021). Covid-19 dashboard by the center for systems science and engineering (csse) at johns hopkins university (jhu).

Van Den Berg, J., Patil, S., Sewall, J., Manocha, D., and Lin, M. (2008). Interactive navigation of multiple agents in crowded environments. In *Proceedings of the 2008 symposium on Interactive 3D graphics and games*, pages 139–147.

Van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., Tamin, A., Harcourt, J. L., Thornburg, N. J., Gerber, S. I., et al. (2020). Aerosol and surface stability of sars-cov-2 as compared with sars-cov-1. *New England journal of medicine*, 382(16):1564–1567.

Västberg, O. B., Karlström, A., Jonsson, D., and Sundberg, M. (2020). A dynamic discrete choice activity-based travel demand model. *Transportation science*, 54(1):21–41.

Vuorinen, V., Aarnio, M., Alava, M., Alopaeus, V., Atanasova, N., Auvinen, M., Balasubramanian, N., Bordbar, H., Erästö, P., Grande, R., et al. (2020). Modelling aerosol transport and virus exposure with numerical simulations in relation to sars-cov-2 transmission by inhalation indoors. *Safety Science*, 130:104866.

Wang, X., Liono, J., McIntosh, W., and Salim, F. D. (2017). Predicting the city foot traffic with pedestrian sensor data. In *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pages 1–10.

Watanabe, T., Bartrand, T. A., Weir, M. H., Omura, T., and Haas, C. N. (2010). Development of a dose-response model for sars coronavirus. *Risk Analysis: An International Journal*, 30(7):1129–1138.

Watts Jr, J. M. (1987). Computer models for evacuation analysis. *Fire Safety Journal*, 12(3):237–245.

Wells, W. F. et al. (1955). Airborne contagion and air hygiene. an ecological study of droplet infections. *Airborne Contagion and Air Hygiene. An Ecological Study of Droplet Infections.*

Weppner, J. and Lukowicz, P. (2013). Bluetooth based collaborative crowd density estimation with mobile phones. In *2013 IEEE international conference on pervasive computing and communications (PerCom)*, pages 193–200. IEEE.

Xie, X., Li, Y., Chwang, A., Ho, P., and Seto, W. (2007). How far droplets can move in indoor environments–revisiting the wells evaporation–falling curve. *Indoor air*, 17(3):211–225.

Xiong, Q., Xu, M., Li, J., Liu, Y., Zhang, J., Xu, Y., and Dong, W. (2021). Clinical sequelae of covid-19 survivors in wuhan, china: a single-centre longitudinal study. *Clinical Microbiology and Infection*, 27(1):89–95.

You, C., Deng, Y., Hu, W., Sun, J., Lin, Q., Zhou, F., Pang, C. H., Zhang, Y., Chen, Z., and Zhou, X.-H. (2020). Estimation of the time-varying reproduction number of covid-19 outbreak in china. *International Journal of Hygiene and Environmental Health*, 228:113555.

Yuhaski, S. J. and Smith, J. M. (1989). Modeling circulation systems in buildings using state dependent queueing models. *Queueing Systems*, 4(4):319–338.

Zanlungo, F., Ikeda, T., and Kanda, T. (2011). Social force model with explicit collision prediction. *EPL (Europhysics Letters)*, 93(6):68005.

Zhang, S. and Lin, Z. (2020). Dilution-based evaluation of airborne infection risk-thorough expansion of wells-riley model. *medRxiv*.

Zuo, Y. Y., Uspal, W. E., and Wei, T. (2020). Airborne transmission of covid-19: Aerosol dispersion, lung deposition, and virus-receptor interactions. *ACS nano*, 14(12):16502–16524.