Testing the crowd:it model for building fire evacuation scenarios

Niclas Movitz Rydhé | Division of Fire Safety Engineering | LTH | LUNDS UNIVERSITET

Testing the crowd:it model for building fire evacuation scenarios

Niclas Movitz Rydhé

Lund 2024

Title: **Testing the crowd:it model for building fire evacuation scenarios**

Author: Niclas Movitz Rydhé

Report 5716
ISRN: LUTVDG/TVBB--5716--SE

Number of pages: 79
Illustrations: 47

Keywords
Evacuation modelling, Validation, ISO 20414, crowd:it, Optimal Steps Model.

Abstract
Evacuation models can be reliably used if they are sufficiently validated for the application in question (e.g. Fire Safety Engineering). To be useful, an evacuation model should therefore be an accurate enough representation of the real-world behaviour of evacuees. The scope of this work was to determine how accurately the model crowd:it (making use of the Optimal Steps Model) provides a representation of the real-world behaviour in building fire evacuation. Validation tests were conducted in accordance with the methodology of verification and validation protocol for building fire evacuation models presented in ISO 20414. Suggested tests for assessing parts of the model representing pre-evacuation time, effect on walking speed at different densities, movement on stairwells and in flight of stairs, counter-flow, route/exit choice, and movement in bottlenecks at openings were conducted. Two global tests of full evacuations of an auditorium and a school were performed as well. There were a few tests where behaviours differed to a greater extent. In counterflow conditions, agents tended to get stuck, and agents tended to zig-zag between lanes in a flow rather than following the person in front (e.g. lane formation phenomena) as experiments showed. Ascending or descending a stairwell does not alter the movement speed of agents when stairs are modelled explicitly. They can however also be implicitly modelled similar other evacuation models. Distributions available for representing evacuees' characteristics and behaviour in crowd:it was recommended to be expanded and the log-normal distribution was added due to findings in this report. In summary, validation test results were generally in line with the experimental results and verification tests highlighted a set of model features that could be improved.

Brandteknik
Lunds tekniska högskola
Lunds universitet
Box 118
221 00 Lund

www.brand.lth.se
Telefon: 046 - 222 73 60

Division of Fire Safety Engineering
Faculty of Engineering
Lund University
P.O. Box 118
SE-221 00 Lund
Sweden

www.brand.lth.se
Telephone: +46 46 222 73 60

Preface

This thesis has been written for and with the Department of Fire Safety Engineering at Lund Institute of Technology (LTH), at Lund University and with accu:rate which develops crowd:it. It encompasses 22.5 Swedish academic credits (hp) which correlates to 600 hours of work and is the final part of my Bachelor of Science in Fire Protection Engineering.

I would like to thank my supervisor Dr. Enrico Ronchi, Senior Lecturer at Division of Fire Safety Engineering at Lund University for all the help and assistance he has provided in this work.

I would also like to thank my secondary supervisor Dr. Angelika Kneidl as well as the rest of the accu:rate team she leads for providing me with the tools for the work and assisting with any questions or queries I have had in the project.

Thank you,

Niclas Movitz Rydhé
Lund, April 2024

Summary

Evacuation models can be reliably used if they are sufficiently validated for the application in question (e.g. Fire Safety Engineering). To be useful, an evacuation model should therefore be an accurate enough representation of the real-world behaviour of evacuees. The scope of this work was to determine how accurately the model crowd:it (making use of the Optimal Steps Model) provides a representation of the real-world behaviour in building fire evacuation. This was accomplished with validation testing conducted in accordance with the methodology of verification and validation protocol for building fire evacuation models presented in ISO 20414.

A background to evacuation modelling, The Optimal Steps Model, and the model crowd:it are presented to aid the reader in understanding the testing of the model. The method section introduces the ISO 20414 protocol and methods for analysis of behavioural uncertainty in evacuation modelling and inferential statistical testing used for the analysis of test results. To get a grasp on the scope of validation work done previously in the field validation testing of the twelve most widely used evacuation models was summarized and presented.

Chapter two presents test configurations, simulation results and comparisons to experiments. Suggested tests for assessing parts of the model representing pre-evacuation time, effect on walking speed at different densities, movement on stairwells and in flight of stairs, counter-flow, route/exit choice, and movement in bottlenecks at openings were conducted. Two global tests of full evacuations of an auditorium and a school were performed as well.

Chapter three and four present discussion and conclusion. In summary, the Optimal Steps Model implementation in crowd:it generally gave results in line with experiment results for the validation scenario considered. There were a few tests where behaviours differed to a greater extent. In counterflow scenarios, agents tended to get stuck, and agents tended to zig-zag between lanes in a flow rather than following the person in front as experiments tend to show. Testing showed ascending or descending a stairwell did not alter the movement speed of agents when explicitly modelled using the Optimal Stair Model. This can accumulate large time differences in long stairwells due to evacuees slowing down due to exertion. It can however also be implicitly modelled similar other evacuation models. The selection of distributions available for representing evacuees' characteristics and behaviour in crowd:it was in some instances limiting the scenario calibration. Therefore, implementation of additional and custom distributions was recommended. The log-normal distribution was recently implemented due to the test results in this report.

Table of Contents

1 Introduction

1.1 Background

Evacuation models are used in Fire Safety Engineering (FSE) to evaluate the evacuation of people from buildings and other structures. In building fire evacuation simulation, the total evacuation time (also called required safe egress time) of the people is often compared to the time until critical conditions occur inside the building (also called available safe egress time) due to the fire. This comparison is often used to evaluate fire safety adopting performance-based design in buildings.

To draw any conclusions from these analyses the evacuation models must provide an accurate representation of evacuations in the real world. The evaluation of the relationship between model outputs and the real world is called validation. Together with verification, V&V (verification and validation) has been identified as the most crucial factor when evacuation model users choose among models (Lovreglio et. al., 2019b). Validation can in FSE be defined as the "process of determining the degree to which a calculation method is an accurate representation of the real world from the perspective of the intended uses of the calculation method" (International Organization for Standardization [ISO], 2008). This is in the context of simulation models done by comparing simulations with experimental data sets. In addition to being accurate to real world behaviour the models must also be implemented into code accurately. The evaluation of the implementation is referred to as verification. Verification can in FSE be defined as "the process of determining that a calculation method implementation accurately represents the developer's conceptual description of the calculation method and the solution to the calculation method" (ISO, 2008). In short, verification can be seen to answer the question "Is the model doing the math right?" and validation is used to answer, "Is the model doing the right math?". This concept can be extended to any type of evacuation models (e.g. agent-based models).

Several guidelines for verification and validation of evacuation models have been developed over the years. One of the most used guidelines has been developed by the International Maritime Organization (IMO) for evacuation analysis of passenger ships (IMO, 2007), and was later revised in 2016 (IMO, 2016). For German-speaking countries, a guideline for the use of pedestrian simulation tools was developed, i.e., the RiMEA (Richtlinie für Mikroskopische EntfluchtungsAnalysen, corresponding to Guidelines for microscopic analysis of evacuation). RiMEA was developed for general crowd evacuation scenarios and includes guidance on verification but not on validation (Rimea, 2016). In 2013, the American agency NIST (National Institute of Standards and Technology) developed a procedure to assess the V&V of models to be used for building fire evacuation specifically (Ronchi, Kuligowski, et al., 2013). In 2020 the International Standardization Organization (ISO) published a V&V protocol for building fire evacuation models, ISO 20414:2020. It is based on the NIST procedure and introduced several validation tests as well as a standardized reporting template.

This thesis aims at investigating V&V for a type of evacuation model using the Optimal Steps Model (OSM) for movement representation, which has not been tested before, . The OSM was developed by Seitz & Köster (2012). It represents movement by modelling each step of each pedestrian. In 2013, the evacuation model crowd:it was developed as part of Dr. Angelika Kneidl's dissertation (Kneidl, 2013) and has been continuously developed by accu:rate since then. It relies on the OSM for movement representation. crowd:it is continuously tested using RiMEA tests and verification tests in accordance with ISO 20414 have been done. Validation testing in accordance with ISO 20414 has not been performed for crowd:it or any other pedestrian software model based on the OSM.

1.2 Purpose and objective

The purpose of this thesis work was to evaluate the degree to which crowd:it, making use of the Optimal Steps Model, provides an accurate representation of the real world from the perspective of the FSE uses.

The objective was to conduct validation testing of crowd:it in accordance with ISO 20414. A literature study was done looking into pedestrian evacuation modelling in general, the OSM specifically and validation work done to the twelve most used pedestrian evacuation models (see section 1.6). Tests were done according to the ISO 20414 protocol and reported according to the ISO reporting template. The full test reports are found in Appendix A and test configurations and results are presented in chapter 2.

1.2.1 Research questions

To achieve the objective a set of research questions were identified.

- How valid is the Optimal Steps Model adopted in crowd:it for use in building fire evacuation design?
- How accurately does crowd:it provide a representation of the real world from the perspective of the intended uses of the model?

1.3 Evacuation theory and modelling, the Optimal Steps Model, and crowd:it

This section provides a background into evacuation and evacuation modelling in general, the assumptions made in the Optimal Steps Model (OSM) and in crowd:it and an introduction into the features in crowd:it and terminology used. The sections on evacuation and evacuation modelling are just brief overviews to give context for the work. More information on modelling in general is referred to other published work dealing with that specifically, such as by Ronchi and Nilsson (2016), and Ronchi (2021).

1.3.1 Evacuation

The need for fire evacuation modelling tools began with the introduction of performance-based design in the fire safety design of buildings (Nelson & Mowrer, 2002; Ronchi & Nilsson, 2016; Gwynne & Rosenbaum, 2016). In performance-based design any fire safety design can be used as long as an adequate level of safety is provided. This is an alternative to prescriptive-based approach where fire safety design follow a set of predetermined rules. Pedestrian evacuation models of differing complexity are used in FSE to evaluate performance-based design. In this context, evacuation models are applied to help determine how much time is needed to reach a place of safety, the so called Required Safe Egress Time (RSET) (ISO, 2020). This is compared to tenability assessments of the building which provide the Available Safe Egress Time (ASET) (Ronchi & Nilsson, 2016).

The study of the evacuation process is based on the academically grounded assumptions that human behaviour during evacuation generally is rational and can be predicted (Ronchi & Nilsson, 2016). There are several theoretical frameworks which are useful to qualitatively interpret the behavioural fundaments of people decision making during fire evacuations. Two reviews have discussed the most important of them which are listed below (Kobes et. al, 2010 ; Fridolf et. al., 2011). They will not be explained further in this paper due to the scope of the work.

- The role-rule model (Tong & Canter, 1984)
- The affiliate model (Sime, 1985)
- Social influence (Latané & Darley, 1970)
- Behaviour sequence model (Canter et. al., 1980)

Due to the need of quantitative results evacuation models commonly use the timeline model (Ronchi & Nilsson, 2016). It simplifies human behaviour during evacuation into consecutive steps or different phases of the evacuation process. The time of each phase can then be assessed individually and summarized to represent the RSET. Compared to more qualitative theoretical frameworks, as the ones above, the time-line model can be considered an over-simplification since it cannot represent the variations of behaviours present in an actual evacuation. However, the simplifications provide a great advantage since it enables a quantitative analysis of the evacuation process in a relatively short time as well as an easier implementation in evacuation models. The timeline model is presented in Figure 1 and the phases are described below.

Δt_{det} is the detection time. It is the time from ignition until the fire is detected.

Δt_{warn} is the alarm time. It is the time from detection to general alarm.

Δt_{pre} is the pre-evacuation time. It includes two phases, recognition time Δt_{rec} and response time Δt_{resp}. Recognition time is the time from being alarmed of the fire but before one starts to respond. Response time is the time from people recognize the need to evacuate but before movement out of the building is initiated.

Δt_{trav} is the travel time. It is the time evacuees need to walk to a safe place.

Δt_{evac} is the evacuation time. It is the sum of pre-evacuation and travel time. This is the phase evacuation models generally are useful for rather than the full timeline (Ronchi & Nilsson, 2016).

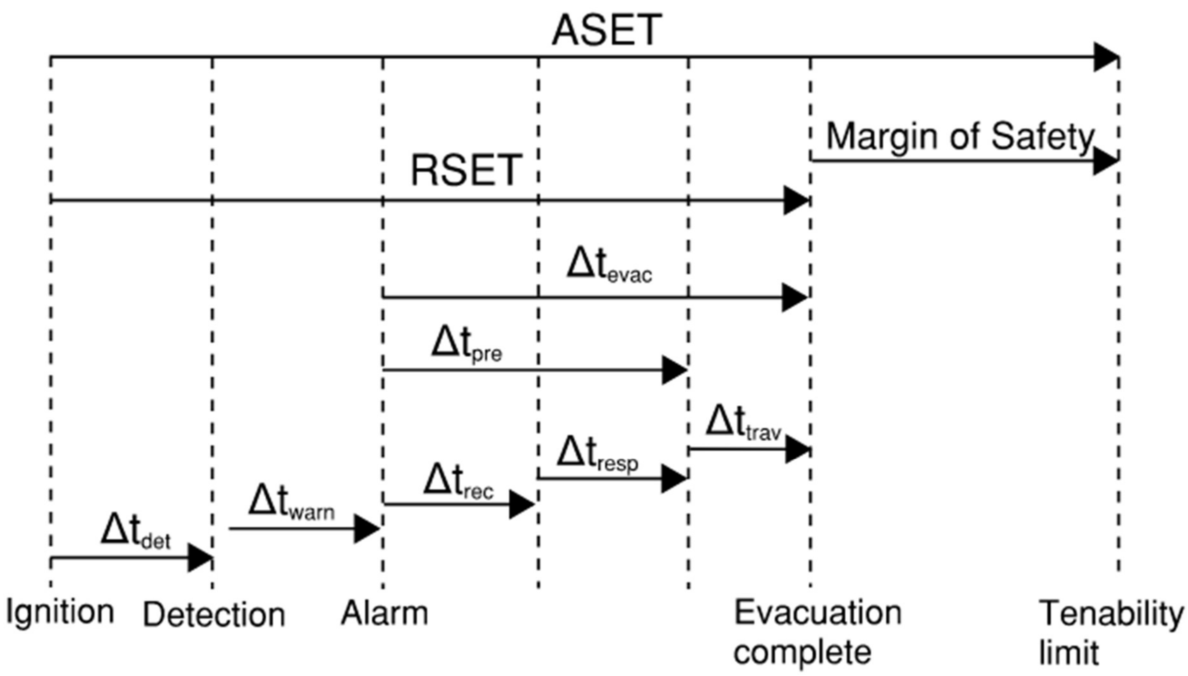

Figure 1 – Representation of the different phases in the timeline model.

Several definitions are available for the pre-evacuation phase and several terms are used to refer to the simulated times during this phase (Ronchi & Nilsson, 2016). It intends to include a variety of activities evacuees engage in which represents a delay in movement to safety. Some common and often interchangeable terms in evacuation models are pre-evacuation time, pre-movement time, pre-travel activity time, pre-travel time, delay time, waiting time and response time. Pre-evacuation time depends on different factors, such as perceived urgency, personal and cultural background, past fire experiences, training level, type of installations available in the building, emergency signage and way-finding installations. The most common method for the simulation of the pre-evacuation time in evacuation models consists of the assignment of a delay time before movement to a safe place is initiated (Ronchi & Nilsson, 2016). This is generally done using a fixed value or a pseudo-random number obtained from a distribution. Pre-evacuation times can in many scenarios be reasonably approximated using normal or log-normal distributions (Purser & Bensilum, 2001).

Travel time is sometimes also referred to as movement time or walking time (Ronchi & Nilsson, 2016). It varies depending on travel speed, route availability, usage or choice, and flow conditions and constraints. Travel speed or free-flow velocity is the maximum uncongested speed at which people move towards a safe place. It depends on the characteristics of the evacuee, such as age, gender, fitness level, and movement ability. In models it is often represented as a fixed value, or a pseudo-random number obtained from a distribution. Free-flow velocities are then modified in relation to different factors, such as crowding, the type of egress component, effects of smoke and trying to keep a group together. Route availability, usage or choice refers to the effect navigation has on evacuation time. Choices can be made based on familiarity of the exits, social factors, egress capacity, proximity, and time factors (Ronchi & Nilsson, 2016). In evacuation models routes are generally assigned deterministically by the user or they implement deterministic or probabilistic algorithms which calculate the best route based on principles such as fastest or closest exit. Flow conditions and constraints refer to the limiting factors in the architecture which adds congestions or in other ways limit the number which can pass by a certain point over a certain period (Ronchi & Nilsson, 2016).

1.3.2 Evacuation Modelling

In early history of the science of fire evacuation modelling, during the 1970's and 80's, simple calculations were developed approximating the movement of humans as the movement of fluids (Ronchi & Nilsson, 2016). These hydraulic models, sometimes referred to as macroscopic models, use the distance to exits, travel speeds and flow through building component to calculate the evacuation time and are still used today (Gwynne & Rosenbaum, 2016). Due to their inability to explicitly take into consideration the impact of human behaviour on evacuation, evacuation modelling tools evolved towards more sophisticated agent-based models, also known as microscopic models (Ronchi & Nilsson, 2016). In these models, each evacuee is represented as an autonomous agent with individual properties and agents' behaviour is often governed by a set of rules determining their decisions as well as agent-to-agent and agent-to-environment interactions.

Evacuation models also differ in their approach to represent geometry (Ronchi & Nilsson, 2016). Early on the coarse network model was mostly used. In coarse network models the geometry is represented as nodes connected with arcs. Each node represents a certain area of the geometry and is given a capacity to limit the number of agents in the area. Arcs are given flow restrictions and travel speed to represent the movement between areas. In fine network models the geometry is represented by cells in a grid-like pattern or as a fine mesh of nodes connected by arcs. Each cell or node only holds a single agent. In continuous models the geometry is represented as a continuous space with agents moving freely. Each agent is described with a position and a body size. The different methods of representing space can also be combined in hybrid models (Chooramun, 2011).

Evacuation models also differ in the way movement is modelled. Reynolds (1999) divided movement in rule-based simulation of individual agents into three hierarchies. Pathfinding is the first level of the hierarchy and refers to strategic aspects of the movement. The agent needs to have a goal of the movement and an algorithm for determining the best route. Most evacuation models assume that the agents know where they are going at the start of simulations (Ronchi & Nilsson, 2016). Local movement or steering behaviour is the next level of the hierarchy (Reynolds 1999). It refers algorithms that deal with agent-to-agent and agent-to-environment actions. Locomotion is the highest resolution of movement and refers to the way movement of the agent's body is represented. In evacuation models locomotion is if present often independent of the other hierarchies of movement.

1.3.3 The Optimal Steps Model

The Optimal Steps Model is developed by Seitz and Köster (2012). It is an agent-based model and represents geometry as a continuous space (Dietrich et al., 2014). The navigation of agents through the geometry is determined by rules of attractive and repulsing potentials (Seitz & Köster, 2012). An attractive force to the target location is represented as a navigation floor field. Two types of repulsing potentials, around each pedestrian and around each obstacle, are also aggregated into the navigation field (Von Sivers & Köster, 2015a). This ensures pedestrians do not step too close to walls or other agents. Agents move stepwise to the next position with the highest utility for them according to the navigation field. The floor field for pedestrian navigation can be illustrated as a topographic map where the repulsing potentials can be imagined as higher ground locally around agents and obstacles and the attractive potentials can be thought of as lower ground. The agents are in this illustration trying to maximize their downhill movement.

To represent the stepwise movement of humans Seitz and Köster (2012) limited the next possible position for each agent to lie on a circle around the agent. The calculations for next position were optimized by only considering a chosen number of equidistant points of the circle. To eliminate artifacts that would occur if the points always had the same position on the circle an element of randomness was implemented in choosing the points. In Figure 2 the stepwise movement and discretisation of the circle are illustrated.

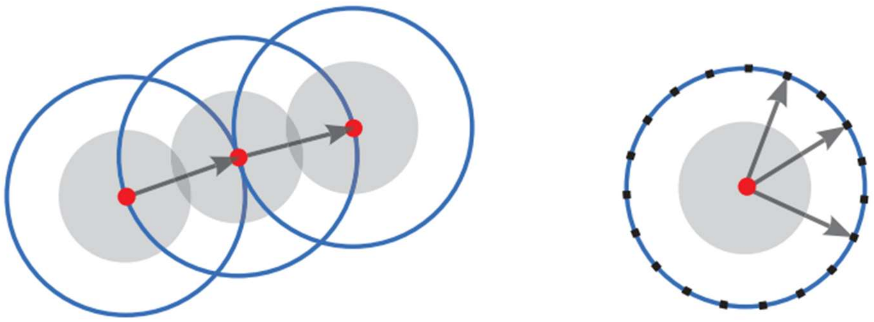

Figure 2 – Left: sequence of three time steps t = 1,2,3. The pedestrian torso is depicted with the filled inner circle. The next position has to be on the step circle around the pedestrian. Right: discretization of the circle indicated by points. The three arrows represent three possible choices for the next step. Reproduced from Seitz and Köster (2012) with permission.

To allow for different stride lengths von Sivers and Köster (2015b) replaced the optimisation on a circle around the agents with the optimisation on a disc. This is the version implemented in crowd:it. The unimpeded velocity of the agents determines the maximum stride length i.e. the radius of the disc but this made it possible for agents to take smaller steps if it was of higher utility to them according to the navigation field. Sivers and Köster (2015a) also implemented behaviour regarding personal space. The behaviour was based on the empirically tested psychological model by Hall (1966). Now agents tried to keep other agents out of their personal space, i.e. within 120 cm, and only allowed agents into the intimate space, i.e. within 45 cm, in very dense crowds. Steps are updated sequentially, and new steps can be taken when the last step is finished, i.e. the time of the step length divided by the speed. The OSM allows agents to have any shape but in crowd:it agents are represented as circles.

Many models represent stairs as an area in which velocity is decreased according to given relationships. For movement on stairs an extension of the Optimal Steps Model is developed by Köster et al. (2016) in which the area where maximum utility is searched is altered. Possible positions for agents next step are limited to intersections between the disc around the pedestrian and the next stair tread thus restricting free movement. If several positions on the next stair tread are of similar utility the closest point is chosen resulting in straighter paths (Köster et al., 2019). The possible positions for the agent's next step are illustrated in Figure 3.

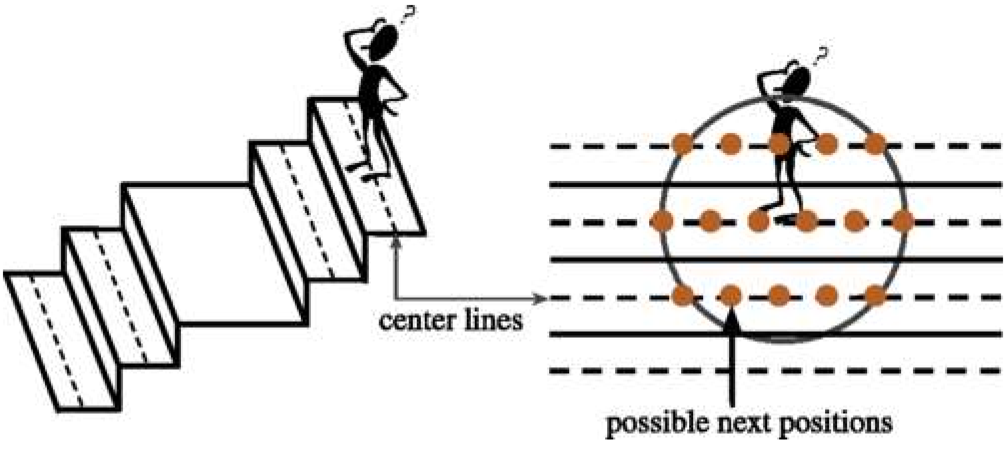

Figure 3 –Graphical illustration of the Optimal Star Model. Solid lines represent the edges of the staircase's treads and are not modelled. The dashed lines represent the centre lines of a tread on which agents are allowed to step. Reproduced from Köster et. al. (2019) with permission.

1.3.4 The crowd:it model

In this section an overview of the crowd:it software is presented. The overview is intended to give an introductory understanding of crowd:it and its capabilities and cannot be seen as exhaustive. For more detailed instructions see the crowd:it documentation (accu:rate GmbH, 2023).

1.3.4.1 Architecture

The software uses 2D floor plans for the simulation of the environment. The floor plans are imported as .dxf files, which most CAD software handles, into crowd:it as a *floor.* Several floor plans can be imported at the same time to make up several floors of a building which can later be connected to each other. Name, height, elevation, and cell discretization can be set for each floor. Height defines the height of the floor in metres and elevation specifies the position of on floor relative to the others. In the current software version floor height and elevation are only used for 3D renders of simulation results. The cell discretization determines the distance between calculated grid points for the navigational fields. It is by default set to 0.1 m. Points in between are interpolated. Objects of the floor plan are divided into the structural objects and simulation object which can be assigned a function in crowd:it.

1.3.4.2 Simulation objects

For simulation objects the following functions can be assigned:

- Origins: The source area from which agents are generated into the geometry.
- Destinations: A target area which removes agents from the simulation if it is the last object in a path.
- Scaled area: Area that affects the target speed with a factor.
- Directed scaled area: Area that affects the target speed with a factor in the positive and negative axis of a chosen direction.
- Waiting zones: Area in which agents wait.
- Stair: Rectangular area which models stairs and can connect two floors.
- Escalator: Rectangular area which models escalators which transport agents and can connect two floors.
- Elevator: Area that models an elevator and can connect several floors.
- Queueing line: A line from which a queue can form.
- Portal: A line which can act as a teleporter to connect two floors

1.3.4.2.1 Origin

An area configured as a simulation object can be set as an Origin. Origins generates a chosen number of agents into the simulation during a chosen interval. In addition to generation interval each origin is assigned personas, premovement time and agent placement. Within an origin, pedestrian of differing characteristics, or *personas,* can be generated according to a chosen percentage. Premovement time determines a time before agents are allowed to move and can be assigned according to distributions for individual agents or as a group or set times. Positioning can be set to sorted or unsorted placement within the origin and with a minimum distance between agents. Since premovement is the term used in the crowd:it model to refer to pre-travel time or pre-evacuation time, explained in section 1.3.1, it will be used when referring to assigned values to this setting in simulations.

1.3.4.2.2 Destination

An area configured as a simulation object can be assigned as a destination. If set as the last step of a path Destinations remove agents from the simulation. Destinations can also be set with intervals.

1.3.4.2.3 Scaled area and directed scaled area

An area set as simulation object can be assigned as a scaled area or a directed scaled area. Both areas affect the target speed of the agents by a chosen factor. For the directed scaled area, a direction for the acceleration or deacceleration field is chosen and different acceleration factors can be chosen for the positive and negative axis. When scaled areas are set to *only for navigation* agents will navigate through the floor as if the scaled area would affect their travel time, but their velocity will in fact not be altered.

1.3.4.2.4 Waiting zone

An area configured as a simulation object can be assigned as a waiting zone. The capacity of each waiting zone is chosen. Waiting times can be set to durations determined by a distribution or a fixed time or agents are allowed to leave at scheduled times. With the setting *meeting zone* agents continue their path if capacity is reached before the waiting time is exceeded.

1.3.4.2.5 Stair

Stairs can only be set to rectangular simulation objects. The orientation of the stair can be flipped and turned 90 degrees and stairs can be connected to another floor. The number of treads is set for each stair and can be altered to tread depths between 0.15 m and 0.40 m.

1.3.4.2.6 Escalator

Escalators can only be set to rectangular simulation objects. For the escalator you specify tread depth between 0.3 m and 0.8 m and the speed at which it transports agents. The orientation of the escalator can be adjusted in the same way as the stair.

1.3.4.2.7 Elevator

An area set as a simulation object can be set as an elevator. The elevator can connect several floors and travel time between floor can be adjusted.

1.3.4.2.8 Queueing line

A line set as a simulation object can be set as a queueing line from which a queue can be formed. Queues can be sorted or unsorted. Agents continue from the queue when the next object in simulation object in their path has free capacity. Some queuing behaviour can be adjusted. Distance between queuing agents can be adjusted. The actual distance varies up to 125% of the specified value. Max queue deviation specifies the greatest angle the queue can deviate from the perpendicular line to the queuing line. Queue width factor affects the shape of the line. It determines how many people would stand next to each other rather than behind each other.

1.3.4.2.9 Portal

A line set as a simulation object can be set as a portal which connects two floors together. The direction of portal can be flipped.

1.3.4.3 Paths

Paths are assigned to each origin and specifies which simulation objects an agent will travel to. Several paths can be assigned to an origin and a probability must then be assigned to each path.

1.3.4.3.1 PathSnippets

Several simulation objects or sets can be combined into a pathsnippet. When a pathsnippet is included in a path the agent either visit each pathsnippet element in the assigned order or in a random order. Each element in the pathsnippet can be assigned a probability of being visited. The pathsnippet can be assigned a capacity which limits the number of agents that may be assigned to the pathsnippet at the same time.

1.3.4.3.2 Sets

Several simulation objects or pathsnippets can be combined into a set. When a set is included in a path only one element of the set is included in the path selected according to specified rules.

1.3.4.4 Persona

A persona is a set of characteristics from which an agent population is modelled. For each persona, distributions of circle diameter for the representation of the torso and velocity are chosen. In the default settings, body size is set to a uniform distribution between 0.42 and 0.46 m. The default velocity settings are set to a normal distribution in accordance with Figure 4. The default settings for body size and velocity are based on the works of Weidmann (1992).

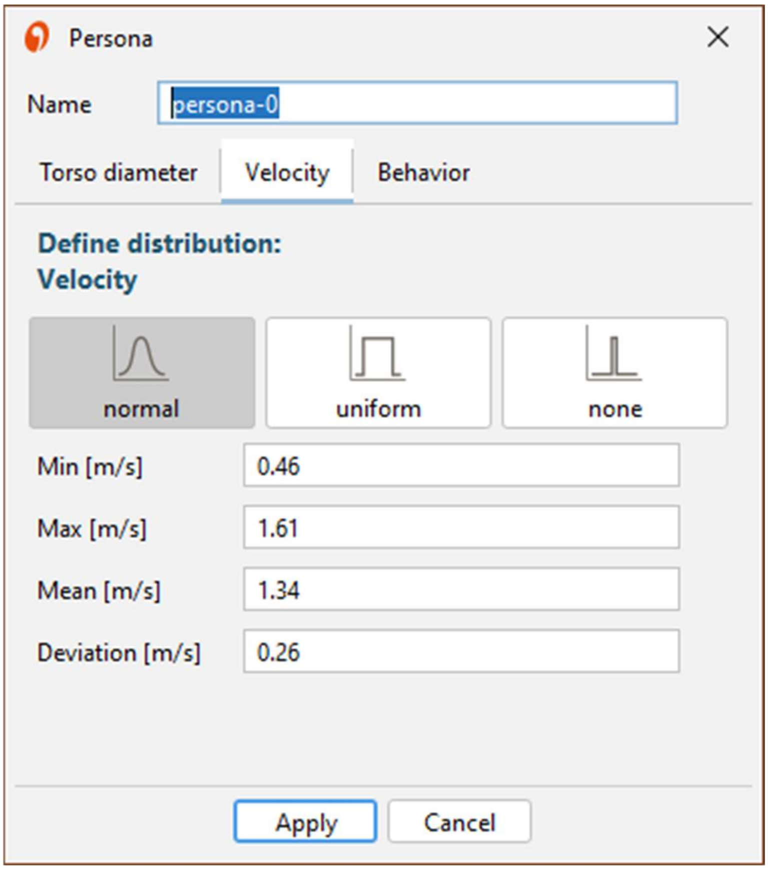

Figure 4 – screenshot of window in crowd:it for adjusting the persona settings.

In the behaviour tab the persona can be excluded from using certain types of simulation objects It also includes a social distancing feature in which the modeler can set a minimum distance between agents between 0 m and 2.5 m. Agents try to keep this distance but will shorten it for dense crowds and bottlenecks. This feature uses an additional model (Mayr & Köster, 2021) which will not be used in the ISO tests or explained further in this paper.

1.3.4.5 Distributions

Whenever a distribution can be specified, as for body size, velocity, pre-movement time and other time settings, the choice is between normal, uniform and no distribution. The available distributions can be seen in Figure 4. The normal distribution can be truncated by specifying a minimum and maximum value. In a newly released version of crowd:it the log-normal distribution was added. It was not available while tests for this report was performed and was partly implemented in response to the testing of this report.

1.3.4.6 Measurements
Data from the simulations can be analysed in many ways inside of the software. Charts can be created in the software to present data from scenario, floor, simulation objects and evaluation objects. Evaluation object are lines and areas applied on top of the finished simulation from which data can be extracted. They are explained further in the following sections. For this paper all data was exported as .csv files. Each dataset had to be manually exported from each scenario, floor or, simulation or evaluation object for each run since there was no mass export feature which would sum results from several runs. Only total scenario evacuation time can be compared across the number of statistical runs performed.

1.3.4.6.1 Tripwire
Tripwires are evaluation objects in the form of lines which measure velocity and passing time of agents traveling across the line. Tripwires can be set to be directional to only measure agents passing in one direction.

1.3.4.6.2 Rectangle and Polygon
Rectangles and polygons are evaluation objects in the form of areas which measure the number of agents in the area, density, velocity, and time spent in the area per agent as well as some other analytics. Polygons can contain corners which are directed inwards.

1.4 Method
This section describes methods used in the validation testing. ISO 20414 is described to give some understanding to the methods used for the validation testing. The method used for dealing with behavioural uncertainty in evacuation simulation and inferential statistics which was used as a tool for analysis of results are presented.

1.4.1 ISO 20414 – Verification and validation protocol for building fire evacuation models
ISO 20414 is developed to verify and validate evacuation building fire evacuation models (ISO, 2020). The document includes a list of model components and a methodology for the analysis of the model and its accuracy. It identifies four core model components and a mandatory starting list of components which needs to be validated in order to be used in building evacuation scenarios. It includes a representation of behaviours corresponding to each component. The identified core components are pre-evacuation, movement, navigation and route selection, and flow condition/constraints. The protocol suggests nine component validation tests to represent behaviours corresponding to these components. In Figure 5 the component validation tests are presented, and colour coordinated to the model core component which they test. The tests includes instructions and suggested data sets for comparison. The instructions include what scenario is to be tested, what result should be compared and the user’s actions. No geometries for the tests are provided since it should represent the conditions in the reference experimental scenarios. For each validation test examples of suitable experimental data sets for comparison are provided. All data sets used in the testing are from the provided data sets and test designs are modelled to represent the conditions in the experiments to the highest degree possible. The protocol also includes instructions and examples of suitable experimental data sets for global validation tests in which several basic components affect the evacuation process.

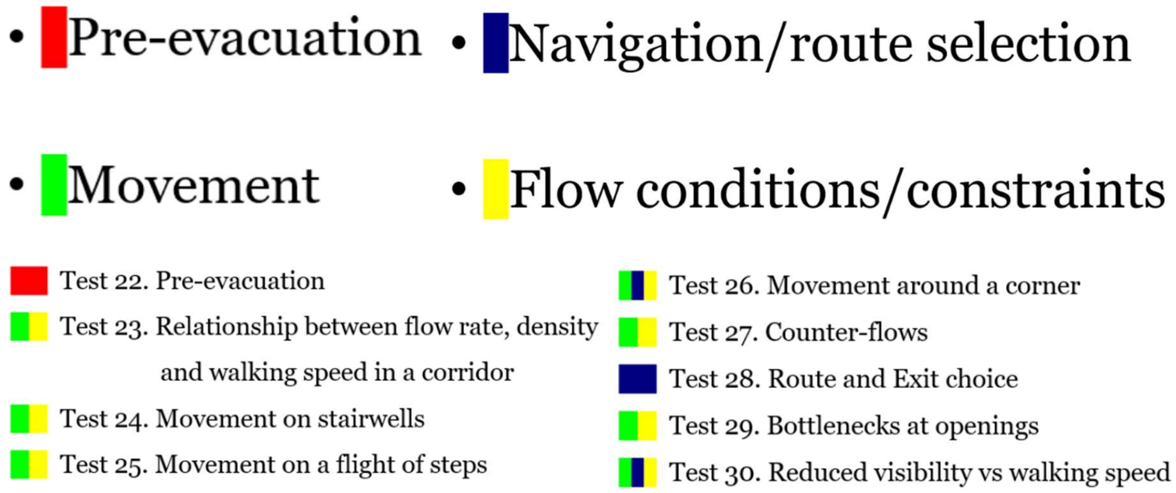

Figure 5 – Core model components colour coordinated to suggested component validation tests in ISO 20414 (2020).

For the validation tests, methods for the analysis of the results are also included. The methods of analysis are divided into basic methods and advanced methods. Method A to D includes comparisons of pre-evacuation time, arrival times, exit choice, and flows though exit/doors. Advanced methods, lettered E to M, include comparisons of arrival time curve, arrival time in different section of scenario, usage of different egress components, density at different sections, queueing time, movement paths and travelled distance, relationship between flows, walking speeds and densities, and visual inspection of occupant behaviour. All these methods consider multiple simulations and/or representative runs e.g., best vs. worst. Some of the advanced methods of analysis include concepts from a branch of mathematics called functional analysis. In functional analysis curves are represented as vectors and it is in this context used to evaluate the differences between experimental and simulated results. Euclidean Relative Difference (ERD), Euclidean Projection Coefficient (EPC) and Secant Cosine (SC) were analysed in these cases. For further explanations of these concepts the reader is referred to other published work which touch on this specifically.

1.4.2 Analysis of behavioural uncertainty in evacuation modelling

Behavioural uncertainty is the uncertainty associated to the variability of human behaviour (Smedberg et al., 2021). The full spectrum of possible evacuation behaviours during a building fire cannot be represented in a single experiment or model run (Ronchi et al., 2013). In evacuation modelling, behavioural uncertainty is often handled using a probabilistic approach (e.g., adopting pseudo-random sampling from distributions for inputs) by running multiple runs of the same scenario. Due to the law of large numbers, averaging results from a large enough number of runs would result in an average value closer to the expected value (Smedberg et al., 2021). Ronchi et al (2013) introduced a quantitative method to investigate behavioural uncertainty using convergence criteria based on functional analysis of the evacuation times. This method is generally used for determining when enough runs have been completed to address the behavioural uncertainty in such scenario. The details of the method will not be described further in this paper and interested parties are referred to the original paper describing the method by Ronchi et. al. (2013). Alternative methods exist in the literature (Grandison, 2020; Grandison et. al., 2017; Tinaburri, 2022).

The five convergence criteria proposed by Ronchi et. al. (2013) are Total Evacuation Time (TET), Standard Deviation of TET (SD), Euclidean Relative Difference (ERD),
Euclidean Projection Coefficient (EPC) and Secant Cosine (SC). When the convergence measures for the criteria are below the acceptance criteria for a given consecutive number of runs, the method assumes they are deemed representative. The impact of the selection of acceptance criteria for this method has not been investigated yet in the scientific literature, therefore the arbitrary values by Ronchi et. al. (2013) were here used. The acceptance criteria below were needed to be met at 10 consecutive runs.

$TR_{TET} = 0.5\%$

$TR_{SD} = 5.0\%$

$TR_{ERD} = 1.0\%$

$TR_{EPC} = 1.0\%$

$TR_{SC} = 1.0\%$

1.4.3 Inferential statistical testing

Statistical inference is the process of using data analysis to infer properties of an underlying distribution of probability (Upton & Cook, 2008). One inferential statistical test is the Kolmogorov-Smirnov test (KS-test). A two-sample Kolmogorov-Smirnov test is used to evaluate if two samples could come from the same distribution (Liao, 2002). A null hypothesis of both samples being of the same distribution is tested. Cumulative frequency distributions are developed for each sample and the maximum difference between the two curves indicates if there is a significant difference between the samples. The KS-test is non-parametric i.e., it does not assume a given shape for a distribution. Because of these reasons it was used in this thesis work.

1.5 Limitations

The analysis of the validation testing results relies on relevant benchmark data sets for comparison. Tests were therefore completed in relation to the data obtained. Validation cannot be considered a one-time task, but rather as a continuous effort to test model predictive capabilities. This project should therefore be considered as one step towards testing a specific model rather than a definitive attempt to validate the model.

In the test method applied in this work no benchmark pre-defined acceptance criteria are available for the tests performed. Neither can it be obtained from other sources. The lack of acceptance criteria entails limitations in the analysis of the validity of the models under consideration. The readers should therefore make their own conclusions from the results whether the models are valid for the intended uses. The answer to the validity of the models can only be given in the degree of deviation from the experimental datasets chosen for comparison.

The method used for the analysis of behavioural uncertainty in evacuation modelling suggested by Ronchi et. al. (2013) relies on acceptance criteria. Due to the lack of work looking into reliable values for acceptance criteria for this method the arbitrarily defined values by Ronchi et. al. (2013) were used. This affects the number of runs which are deemed representative of distributions used for the test configurations.

1.6 Validation summary

To get a grasp on the scope of validation work done previously it was investigated. In Table 1 publicly available documents reporting the validation of the twelve most used computer pedestrian evacuation models (Lovreglio et. al., 2019b) excluding crowd:it are summarized. The summary cannot be considered exhaustive since it is based on what validation data was readily available online. There may be more tests conducted that are not publicly retrievable.

Table 1 – Comparison of validation done to the twelve most used evacuation models.

Software [Developed by]	Validation done	Sources
Pathfinder [Thunderhead Engineering]	Fundamental diagram tests, flow rate tests, behaviour tests, test for special features such as FED calculations and walking speed reduction due to smoke as well as IMO and NIST tests. Most of it relates to verification. Some validation in fundamental diagram tests and tests for special features	(Thunderhead engineering, 2021)
	Comparison of evacuation from twin bore tunnel	(Ma et al., 2014)
	Comparison of evacuation time for full evacuation of lecture hall, sports arena, and theatre as well as specific behaviour in bottlenecks.	(Wijnhoven and Klein, 2014)
	Evacuation time and emergency exit usage compared to tunnel experiment and other models.	(Ronchi, 2013)
STEPS [Mott Macdonald]	Open test with sensitivity analysis of full building evacuation in comparison to 6-story office building with occupied basement in London with all experiment information given to modeler. Blind test for full building evacuation of 7-story building with occupied basement in Ottawa.	(Lord et. al., 2005)
	Comparison to calculations according to NFPA 130 for emergency evacuation in a mass transit station.	(Kang, 2006)
	Evacuation time and emergency exit usage compared to tunnel experiment and other models.	(Ronchi, 2013)
Massmotion [Oasys/ARUP]	Tests validating varied merging flows in stairs, flow through small opening in end of corridor and full evacuation of cinema	(Mashhadawi, 2016)
	Summary of several published validation cases and one unpublished case including stair/route/exit usage, flow rates, and total evacuation times from evacuation drills in high and medium rise buildings.	(ARUP, 2015)

VISSIM/Viswalk [PTV Group]	Four tests covering corridor with small opening, evacuation of classroom and theatre, and flow and movement time on downward stairs.	(Blomstrand Martén and Henningsson, 2014)
	Egress behaviour at bottlenecks compared to experiments for both normal and emergency conditions.	(Shi et. al., 2021)
Pedestrian dynamics [InControl]	no public validation found.	
Legion [Bentley]	1. Density and flow in large crowd entering train station through narrow opening. 2. Carriage boarding and alighting times in London metro. 3. Flow versus density in downward stairs boarding queue in London metro.	Berrou et.al., 2007)
FDS+Evac [VTT Technical Research Centre of Finland]	3 tests: 1. Specific flow through corridors. 2. Flow through staircase of an office building. 3. Evacuation from Full movement phase of evacuation from public library with dual exits. IMO verification testing.	(Korhonen, 2018)
	Comparison of evacuation from twin bore tunnel	(Ma et al., 2014)
	Evacuation time and emergency exit usage compared to tunnel experiment and other models.	(Ronchi, 2013)
Simulex [IES, Ltd.]	Tests validating Simulex have been carried out by staff at Edinburgh University, Lund University, Ove Arup (Australia) and University of Ulster. Tests were conducted on a variety of buildings, including department stores, office buildings, lecture theatres, sports stadiums, and others. (Tests demonstrated that Simulex accurately models individual movement and yields realistic results when analysing groups. The simulated flow rates correspond well with real-life evacuation flow rates, including during fire drills, in the absence of fire cues.)	(Thompson, 2018)
	Comparison of evacuation time for full evacuation of lecture hall, sports arena, and theatre as well as specific behaviour in bottlenecks.	(Wijnhoven and Klein, 2014)
	Evacuation time and emergency exit usage compared to tunnel experiment and other models.	(Ronchi, 2013)

Exodus [Fire Safety Engineering Group at the University of Greenwich]	Tests validating exit flows as well as full evacuations including hospital ward, theatre, university structure, Gothenburg fire and the pavilions at the Tsukuba world exposition.	(Galea et. al., 2017)
	Evacuation time and emergency exit usage compared to tunnel experiment and other models.	(Ronchi, 2013)
EGRESS [ESR Technology]	Comparison of evacuation time for full evacuation from Trident aircraft through door and over wing, double deck bus, two theatres and own building(ESR Technology) with blocked main exit during fire drill.	(Ketchell, 2006)
EvacuatioNZ [University of Canterbury, New Zealand]	Comparison of evacuation time for full evacuation from three high rise buildings. Effect of variance of involvement of disabled people, complexity of node configurations, and different pre-movement distributions.	(Tsai, 2007)
	Comparison of evacuation time for full evacuation of lecture hall, sports arena, and theatre as well as specific behaviour in bottlenecks.	(Wijnhoven and Klein, 2014)

2 Simulation and results

In this section test configurations, results and comparisons with experiments are presented. Tests included are based on suggested component tests in ISO 20414 (ISO, 2020) and instructions for global validation testing. For each test an experiment data set was chosen from the suggested data sets in the ISO protocol. Test configurations were based on the experiments and results from each test are evaluated in how they compare to the experiment results. Tests were documented using the suggested reporting template in ISO 20414. The full test reports with detailed descriptions of the test configurations are presented in Appendix A.

The main features listed below were tested as suggested in the ISO 20414 protocol. Test 30 was not performed due to crowd:it not having an explicit model which models the consequences of smoke on evacuation behaviour.

- Test 22. Pre-evacuation
- Test 23. Relationship between flow rate, density and walking speed in a corridor
- Test 24. Movement on stairwells
- Test 25. Movement on a flight of steps
- Test 26. Movement around a corner
- Test 27. Counter-flows
- Test 28. Route and Exit choice
- Test 29. Bottlenecks at openings
- (Test 30. Reduced visibility vs walking speed)

In addition, full evacuations of an auditorium (Global test 1) and a school (Global test 2) were modelled to tests several core components simultaneously. The experiments the global tests were based on and compared to were chosen because the geometry, population, experiment conditions and results were well described.

2.1 Test 22 Pre-evacuation

The model ability to represent pre-evacuation behaviour was tested. The ISO protocol instructs that the simulation of pre-evacuation behaviour is to be compared to the observed experimental pre-evacuation behaviour (ISO, 2020). The crowd:it model represent pre-evacuation behaviour with deterministic values or pseudo-random sampling from distributions for a delay time before the agents start to move.

Test design and experimental data for comparison were taken from experiment conducted by Gwynne and Boswell (2009). They observed an unannounced evacuation exercise from a mid-rise administrative building and collected pre-evacuation times and evacuation times. To not alert the evacuees of the exercise pre-evacuation times were recorded from alarm went off until the evacuees entered the stairway. This included a small amount of travel time when evacuees travelled from their initial location to the stairwell. Gwynne and Boswell (2009) estimated that it would have taken between 10 and 15 seconds to have travelled from the most distant part of the floor to the stair door. This study was chosen as a suitable data set for the test due to the paper presenting raw values for the recorded pre-evacuation times and providing some information about the geometry of the experiment conditions.

2.1.1 Simulation configuration

The raw values for the measured pre-evacuation times from the experiment by Gwynne and Boswell (2009) was used to create a normal distribution applied to the premovement time setting for the agent population. A normal distribution was deemed the better alternative of representing the experiment times to a fixed time or a uniform distribution. Pre-evacuation times in the experiment included the travel time to the measure point. To account for additional travel time in the experiment the premovement distributions with both 10 s and 15 s subtracted were used for simulation testing. The normal distributions based on the experiment data with different time alterations are presented in Table 2. The simulation geometry was set up to also include a maximum 10-15 s travel time.

Table 2 – Data for normal distributions for pre-evacuation time

Scenario	Data alteration	Mean (s)	Standard deviation (s)	Minimum (s)	Maximum (s)
1	-10 s	64	34	13	142
2	- 15 s	59	34	8	137

Body size and walking speed of population was set to default distributions since no other data was available for the population. Convergence of total evacuation time (TET) and pre-travel time was assessed for both configurations. For TET scenario 1 convergence criteria were met at run 20. For scenario 2 convergence criteria were met at run 23. For pre-travel times convergence criteria were met at run 19 for scenario 1 and run 26 for scenario 2. Evacuation times for each arrival were averaged across 21 runs for scenario 1 and across 26 runs for scenario 2.

The floor geometry was represented in accordance with Figure 6 to recreate dimension in experiment by Gwynne and Boswell (2009). The figure shows the walls of the building with lines in black and agent origin and destinations with orange edges. Dimensions were only given for outer walls and door sizes. Dimensions of the centre room and position of exits to the stairwells were set according to Figure 6. The origin was offset from the corners by 4 m to match experiment travel time which was estimated to be no longer than 15 s from most distant point on the floor (Gwynne & Boswell, 2009). Two identical geometries were setup next to each other only differing in the number of agents to represent the two video-recorded floors of the experiment with differing population.

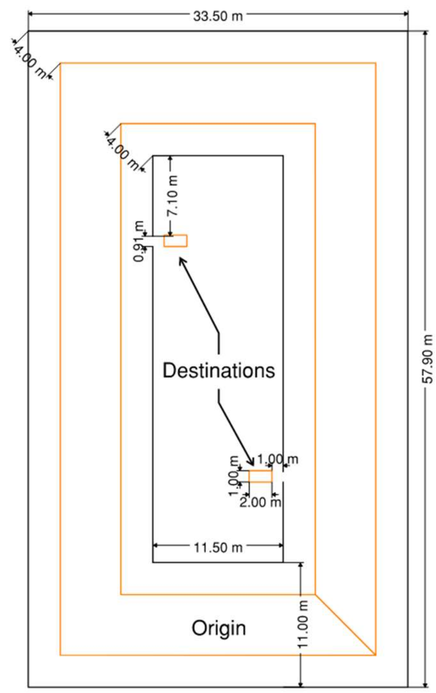

Figure 6 – Geometry of the test design in accordance with experiment by Gwynne & Boswell (2009).

80 agents were placed in the one origin zone and 52 in another within the first second to match floor 5 and 6 of the experiment. The destinations of each floor were connected as a set meaning agents travelled to the closest destination. Exit time per agent was measured.

2.1.2 Results and comparison

In Figure 7 evacuation time curves for simulations and experiment are presented representing pre-evacuation times. The simulation curves seem to follow the same shape with scenario 2 overall being about 5 seconds faster as expected as the distribution for pre-evacuation time differs 5 seconds. The experiment curve differs slightly in shape to the simulation curves. Maximum and minimum times are within 4 seconds for scenario 1 and 2 seconds for scenario 2 compared to experiment times. Experiment pre-evacuation times are faster than simulations between 10 and 100 arrivals. From 110 arrivals the simulations have faster pre-evacuation times. The difference between simulation and experiment results seems to be largest between around 60 arrivals.

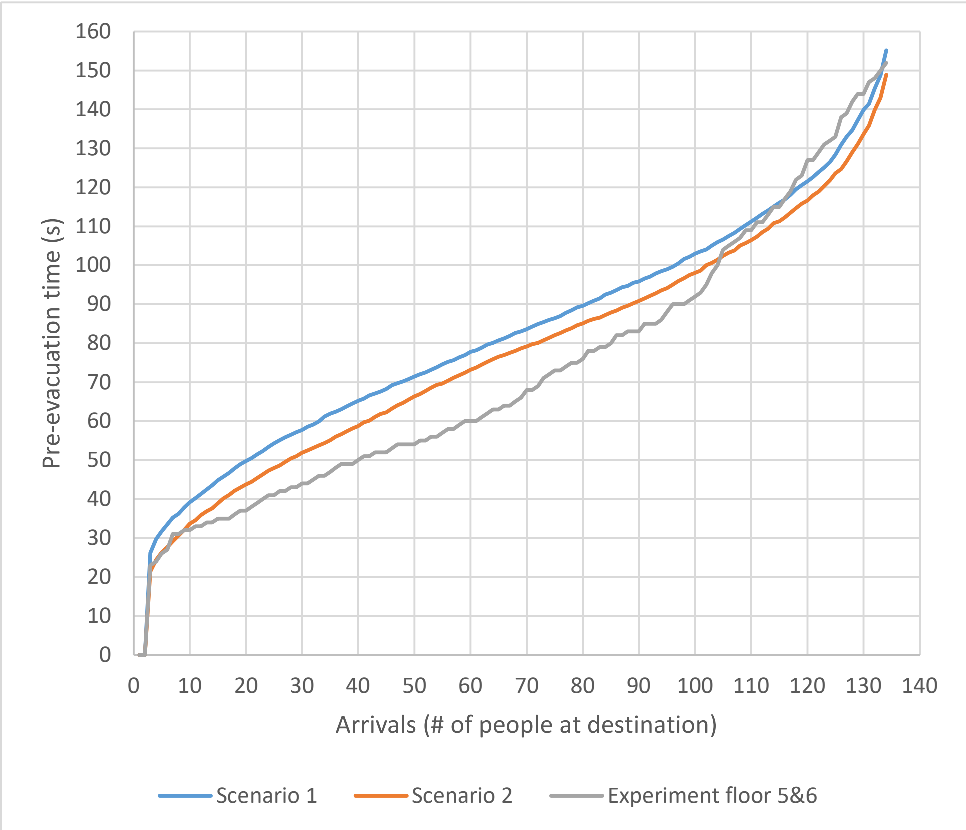

Figure 7 – Evacuation time curve representing pre-evacuation time for experiment and scenario.

When Kolmogorov-Smirnov tests (KS-test) were run with a significant level of 0.05 using the averaged time for each agent the hypothesis that scenario 1 and experiment results were from the same distributions were rejected. For scenario 2 the hypothesis was accepted. When using the data from all agents across all runs the hypothesis for both scenarios 1 and 2 were rejected.

In Table 3 Euclidean Relative Difference (ERD), the Euclidean Projection Coefficient (EPC) and the Secant Cosine (SC) are presented for averaged evacuation timed from simulation compared to the measured times from the experiment. The closer the ERD is to 0 and EPC and SC is to 1 the more similar the curves are. Both simulation evacuation curves are not too far off from 0 and 1 respectively. The values show scenario 2 being closer to the experiment values than scenario 1.

Table 3 – ERD, EPC and SC of averaged simulation evacuation times compared to experiment simulation times.

	ERD	EPC	SC
Scenario 1	0.014	1.083	0.993
Scenario 2	0.010	1.025	0.995

In Figure 8 a histogram of experiment pre-evacuation times with 15 seconds subtracted overlayed with distributions to fit the data points is presented. To fit the distributions to the experiment data the risk analysis tool @Risk by Palisade was used. Four distributions were fitted to the data, triangular, log-normal, normal, and uniform. The triangular distribution had the best fit, followed by lognormal. The uniform had the worst fit. Only normal and uniform distributions can be used in crowd:it. The difference in the evacuation time curve is probably due to the normal distribution of the pre-evacuation time not being a good fit for the experiment response time distribution. The normal distribution used in the test underrepresent pre-evacuation times shorter than 50 s and overrepresent pre-evacuation time longer than 50 s compared to experiment results.

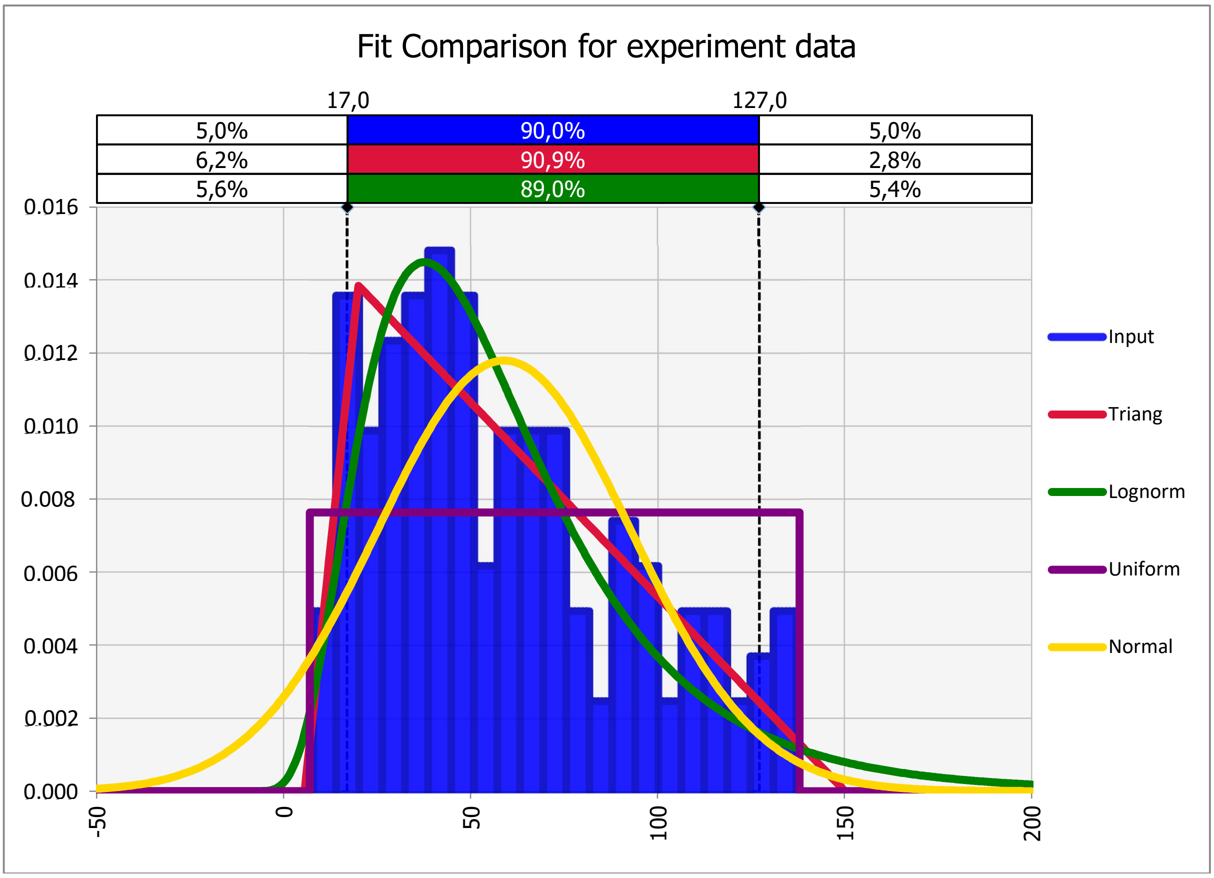

Figure 8 – Experiment pre-evacuation times overlayed with fitted distributions.

2.2 Test 23 – Relationship between walking speed, uni-directional flows, and densities

The model ability to represent the relationship between flow rate, density and walking speeds in uni-directional flow was tested. The ISO protocol instructs that the models ability to represent the expected relationship in relation to the population and the geometric layout is to be compared to the observed evacuation scenario including a uni-directional crowd flow (ISO, 2020). The crowd:it model employs the Optimal Steps Model for representation of pedestrian movement.

Test design and experimental data for comparison were taken from experiment conducted by Seyfried et al. (2007). In the experiment, university students and staff walked in a single line formation in an oval and the velocity at different densities were measured. The participants was students and staff at the Central Institute for Applied Mathematics of the Research Centre Jülich. To measure behaviour at different densities tests were performed with groups of 1, 15, 20,25,30, and 34 participants walking in the passageway. One section of the straights of the passageway was recorded and velocity and density data was extracted. The paper by Seyfried et al. (2007) also include a section in which they tried to recreate the results of the study with another evacuation model.

2.2.1 Simulation configuration

All participants in the experiment can be expected to have been able bodied, and most participants can be expected to have been between 20 and 30 years of age. The body size of the population was not specified. Therefore, it was set to default distribution for all tests. Pre-movement time was set to 0 s for all tests. No data of the experiment population's unimpeded speeds was presented. Therefore 3 different normal velocity distributions, presented in Table 4, were tested. Velocity distribution 1 was based on values used by Seyfried et al. (2007) when they were trying to recreate the experiment results in simulations using another evacuation model. Minimum and maximum values were set the same as the default distribution since only mean velocity and standard deviation was used in the study. Velocity distribution 2 is the default velocity distribution in crowd:it. Velocity distribution 3 is based on data from measurements of young adults by Gales et al. (2020).

Table 4 – Velocity distributions used in test 23.

Velocity distributions	Mean (m/s)	Std. deviation (m/s)	Minimum (m/s)	Maximum (m/s)
1	1.24	0.05	0.46	1.61
2	1.34	0.26	0.46	1.61
3	1.61	0.58	0.71	3.92

The geometry was represented in accordance with Figure 9 to recreate dimension in the experiment by Seyfried et al. (2007). Dimensions for the curves were not given and were fitted to ellipses with centre positions marked with + inside the track since this looked most like the schematics presented for the passageway in the study.

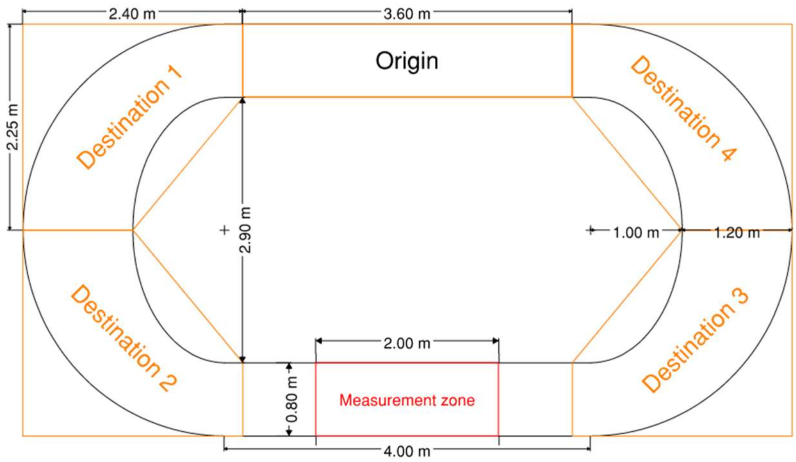

Figure 9 – Geometry of the test design in accordance with experiment (Seyfried et al., 2007)

The test conducted with the geometry presented in Figure 9 shows that faster agents were passing slower agents in the wider part of the course. In the experiment it was clear that no passing occurred. To mitigate the problem of agents passing each other in the first geometry a second geometry was produced with consistent course width of 0.8 m presented in Figure 10. All other dimensions were kept the same. The geometry in Figure 9 is referred to as geometry 1 and the geometry in Figure 10 is referred to as geometry 2. Simulations were performed with populations of 15, 20, 25, 30, and 34 to recreate the conditions in the experiment. Simulation using geometry 1 was only performed with velocity distribution 1 for comparison with the results of the altered geometry. Simulations for geometry 2 were performed using velocity distributions 1-3.

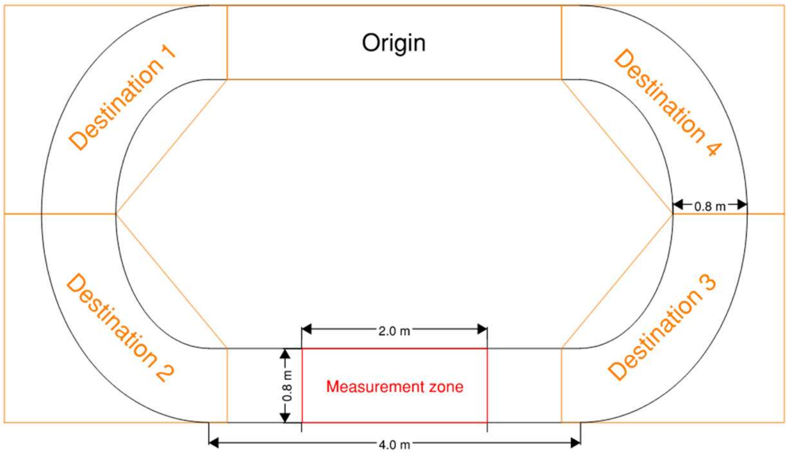

Figure 10 – Secondary test geometry with constant width

2.2.2 Results and comparison

In Figure 11, the distance needed in front of a pedestrian to reach a certain velocity from the experiment by Seyfried et al. (2007) is presented as measured points as well as a fitted linear expression. This is overlayered with averaged simulation results for each case with a fitted linear relationship with R^2 value for all test configurations. Each test configuration is identified with geometry and mean velocity in the velocity distribution used. The linear relationships are numbered y_1 to y_4 matching the descending order in the chart legend. No R^2 value was given for regression curve in the experiment report so the data points were extracted using WebPlotDigitizer and used to calculate a R^2 value. Simulation results seem to follow the same trend as experiment results although with a higher velocity at the same free distance. All tests using geometry 2 yield very similar results to one another as well as to the experiment and results for velocity distribution 1 and 2 almost overlap.

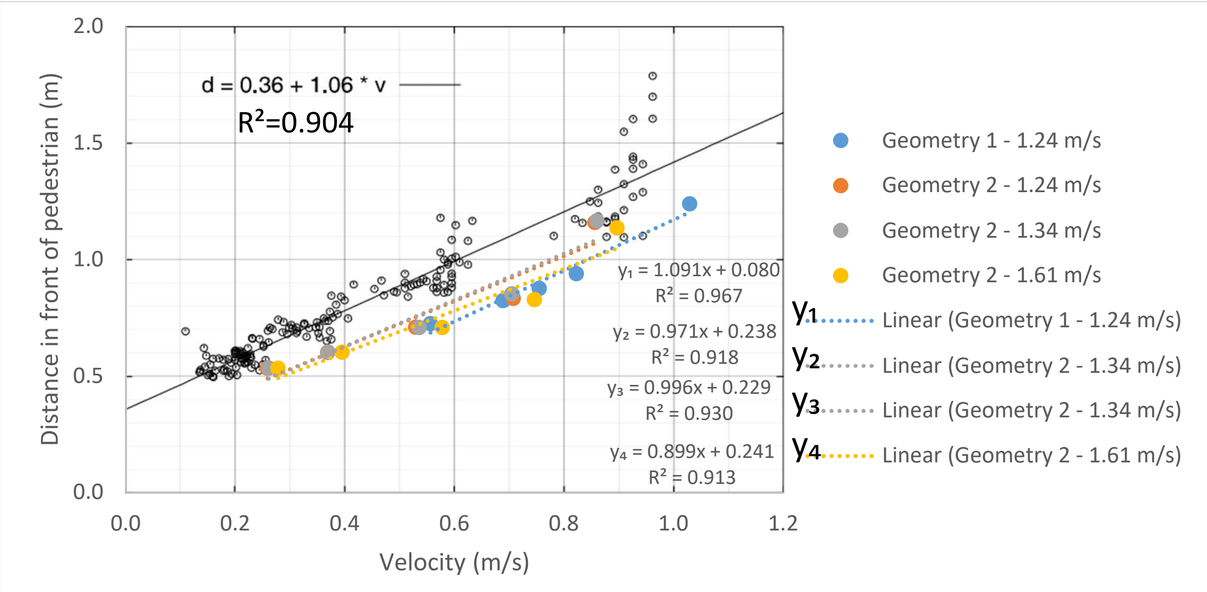

Figure 11 – Chart of mean velocity at a given free distance in front of the agent with fitted linear relationship overlayed experiment results along with a fitted linear relationship presented by Seyfried et al. (2007).

In Table 5 the differences in percentage between distance in front of agents and expected distance in front of agents are presented. Expected free distances in front of agents are calculated by applying the averaged velocities from the simulations to the fitted linear expression (d). The different test configurations are numbered in accordance with the descending order of the graph legend. Test configuration 1 represent tests with geometry 1. Test configurations 2-4 represent tests with geometry 2 and velocity distributions 1-3 respectively. For test configuration 1 the differences are 24-25% for all scenarios except for the lowest density scenario which was only 15%. For the test configurations using geometry 2 the differences were generally lower, especially for the highest and lowest density scenarios. For test configuration 4 the differences were higher than for test configurations 2 and 3 and the differences for scenarios N20 and N25 even surpassed test configuration 1.

Table 5 – Differences in percentage between mean distance in front of agents and expected distance in front of agents based on mean velocity.

		Difference to expected distance to next agent			
	Test configuration	1	2	3	4
Test population	N15	15%	9%	8%	13%
	N20	24%	25%	23%	28%
	N25	25%	23%	24%	27%
	N30	24%	20%	20%	23%
	N34	24%	16%	16%	18%

2.3 Test 24. Movement on stairwells

The model ability to represent movement behaviour on stairs was tested. Both agents ascending and descending a staircase was studied using 2 different velocity distributions for a total of 4 different configurations. The ISO protocol instructs that the simulation of stairwell evacuation movement behaviour, such as evacuation times, merging flow rations in the stairs and landing, and movement paths, is to be compared to observed experimental movement behaviour. The protocol also gives the option to compare floor emptying sequences and evaluate evacuation on escalators.

Test design and experimental data for comparison were taken from experiment conducted by Choi et al. (2013). In the experiment 30 male and 30 female participants ascended and descended the stairwell in a 50-story residential building in Korea one by one. Arrival times was recorded on each floor using cameras.

2.3.1 Simulation configuration

For the test, a staircase spanning 50 floors was set up. The geometry for the descent test case was represented in accordance with *Figure 12* to recreate dimension in the experiment (Choi et al., 2013). For the ascent test case, the destination and origin were switched. The landing halfway up the floor was modelled as part of the lower floor.

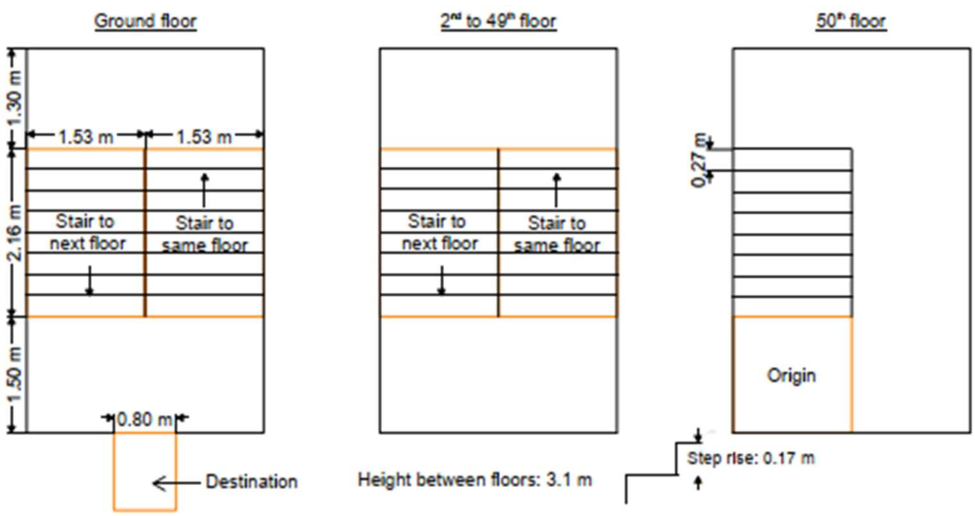

Figure 12 – Geometry of the test design in accordance with experiment (Choi et al., 2013).

Agents were set to move from origin through the stairs to the destination. A single agent was used in each run as experiments were done individually. Ascent and descent scenarios were run both assuming a walking speed of the population set to default distributions and with walking speeds calibrated in line with the recorded free walking speeds in the experiment. Body size was set to default and pre-movement time was set to 0 s for all cases.

In the experiment the population was made of 30 males and 30 females in the ages 20 to 28. The average age was 23.4 years. Most of the participants were recorded having normal weight according to BMI. 5 males were considered obese, and 7 females was considered underweight according to BMI. The free horizontal walk speed of each participant was measured. The averages were recorded to 1.44 m/s for males and 1.26 m/s for females. In Figure 13 a histogram of recorded walking speeds from the experiment with fitted distributions is presented. A normal distribution was chosen as the best fit over uniform distribution. Triangle distributions is not supported in crowd:it. The normal distribution had the mean of 1.34 m/s (same as default), standard deviation of 0.19, minimum value of 0.98 m/s, and maximum value of 1.71 m/s.

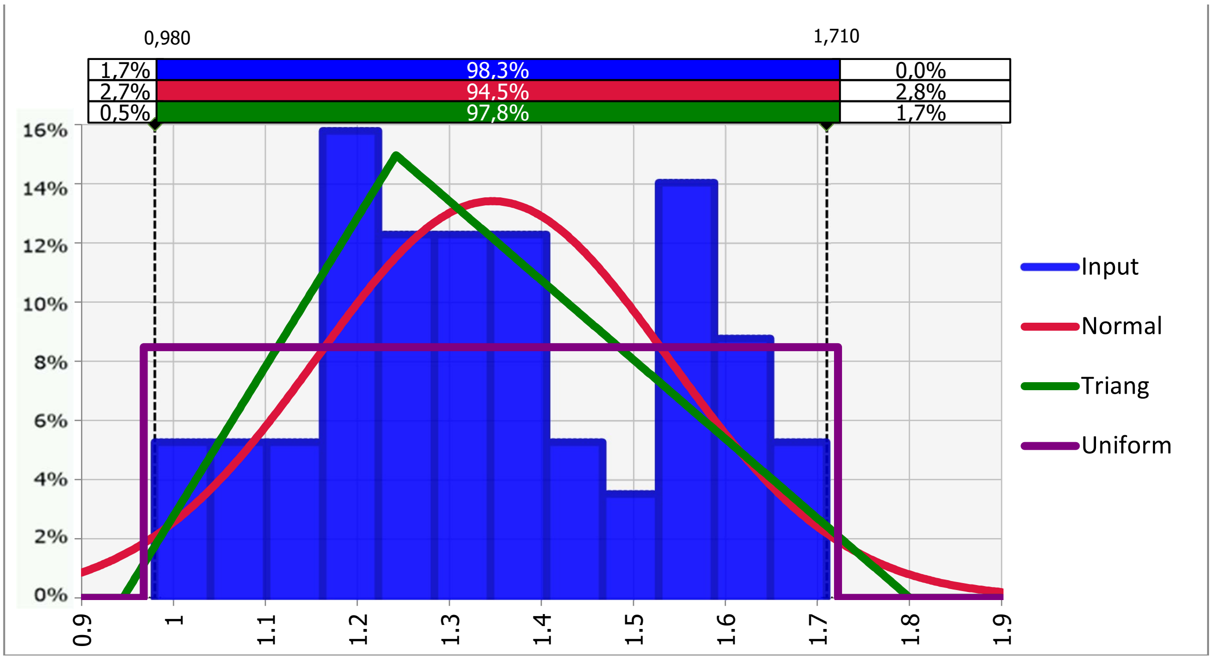

Figure 13 – Measured walking speeds from experiment overlayed with fitted distributions.

Convergence of total evacuation time was assessed for each scenario. Convergence criteria were met at 31 runs for descending runs and at 43 runs for ascending runs for scenarios with default velocity distribution. Total evacuation time for the number of runs needed for convergence was averaged. For scenarios with the calibrated velocity distribution convergence criteria were met at run 24 for both descent and ascent. Total evacuation times were averaged over 60 runs for both scenarios to match the number of data points in experiment. Since simulations were done with only 1 agent per run no secant cosine could be assessed since it analyses the derivatives of the evacuation times. This convergence criteria have therefore been omitted from this test.

2.3.2 Results and comparison

In Table 6 data from experiment in 50-story staircase by Choi et al. (2013) as well as simulation data for both ascent and descent using different velocity distributions are presented. Mean, minimum and maximum is presented for both experiment and simulation. Standard deviation is only presented for simulation result since it was not presented in the experiment report. The experimental data were given per gender. Since a single population was used in the simulation, the results of men and women were combined. Mean evacuation time from the simulations was compared to the mean evacuation time from the experiment.

Table 6 – Data for evacuation time from experiment and simulations, and comparisons between them.

		Experiment			Simulation	Difference
		Men	Women	Gender non-specific		
Descent default distr.	Average (s)	468	508	488	571	14.6%
	Std. dev. (s)	-	-	-	43	
	Min (s)	320	397	320	518	
	Max (s)	593	571	593	662	
Ascent default distr.	Average (s)	629	832	731	574	-27.3%
	Std. dev. (s)	-	-	-	47	
	Min (s)	425	596	425	520	
	Max (s)	822	955	955	708	
Descent calibrated distr.	Average (s)	468	508	488	554	11.9%
	Std. dev. (s)	-	-	-	25	
	Min (s)	320	397	320	510	
	Max (s)	593	571	593	616	
Ascent calibrated distr.	Average (s)	629	832	731	552	-32.3%
	Std. dev. (s)	-	-	-	28	
	Min (s)	425	596	425	512	
	Max (s)	822	955	955	623	

For simulation results presented in Table 6 a few aspects are worth noting. Evacuation times using the same velocity distribution was almost identical regardless of agents ascending or descending the stairs. The evacuation time for simulations using the calibrated velocity distribution was uniformly faster than simulations using the default distribution. Experimental data on the other hand showed a great difference between ascent and descent. In the experiments, the travel time as well as the spread between fastest and slowest time increases for the ascent case. This can be expected since going up the stairs rather than down exerts the participants due the effects of gravity.

For the simulated descending runs both configurations over-estimated the evacuation time. The difference for descending runs decreased for runs using the altered velocity distribution since these runs overall were faster. The difference for descending runs using the default distributions was 15% and for runs using the calibrated velocity distribution it was 12%. For ascending run both configurations under-estimated evacuation times. The difference for ascending runs increased for runs using the altered velocity distribution since these runs overall was faster. The difference for ascending runs using the default distributions was 27% and for runs using the altered velocity distribution it was 32%. The uniformity of the simulation results regardless of agents descending or ascending can be expected as the Optimal Stairs model only takes tread depth into account without considering explicitly the direction of movement. This shows the need to choose a different velocity distribution depending on if agents will be descending or ascending long stairs.

2.4 Test 25. Flight of steps

The ability of the model to represent pedestrian movement on a flight of steps was tested. The ISO protocol instructs that simulated pedestrian movement flows and behaviour along a flight of step shall be compared to the observed experimental behaviour.

Test design inspiration and experimental data for comparison were taken from experiment conducted by Burghardt et al. (2013) with further descriptions found in thesis by Burghardt (2013). Experiments were done in an external staircase and on the 1.2 metres wide stair in the upper and lower tier of the grandstand of an arena. Horizontal velocity and density was extracted from video recordings of the tests.

2.4.1 Simulation configuration

The geometry was represented in accordance with Figure 14 to recreate the dimension of the grandstand stairs in the experiment (Burghardt et al., 2013). Tread depth was set to 0.265 m. Density measurement zone was set up on top of stairs as a 2.4 m^2 rectangle. A tripwire was set up at the end of measurement zone to measure time of people passing.

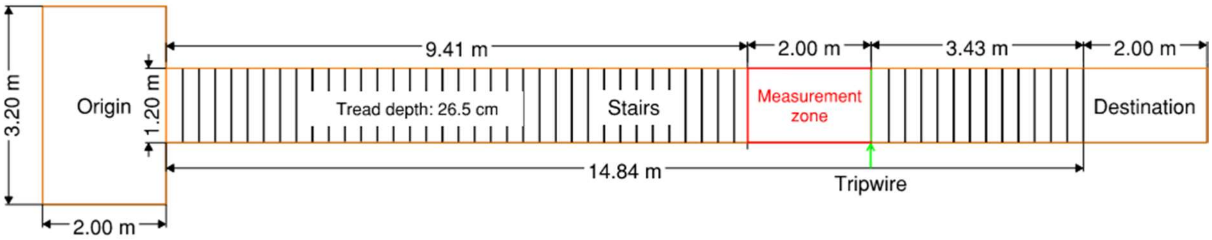

Figure 14 – Geometry of test design

Agents were set to move from origin through the stairs to the destination. Agent population was set to 300 agents to get a steady flow. Agents spawned into the origin as enough space was available. Local density was measured in the measurement zone and passing time was measured by a tripwire at the end of the measurement zone. The flow was calculated from passing times.

Experiment population was made up of 350 participants, mostly composed of students. No additional data was given for experiment population. Pre-movement time was set to 0 s. Body size of the population was set to default population for all tests since no explicit data was given. In addition to the default velocity distribution (velocity distribution 1) the tests were run with two other velocity settings using normal distributions explicit for students. The data for the distributions are presented in Table 7. Velocity distribution 2 was slower than the default distribution with the mean of 1.23 m/s and a standard deviation of 0.22 m/s (Cao et al., 2016). The minimum recorded free velocity was 0.6 m/s, and the maximum was 1.6 m/s. Velocity distribution 3 was faster than the default distribution with the mean of 1.61 m/s and a standard deviation of 0.59 m/s (Gales et al. 2020). The minimum recorded velocity was 0.71 m/s, and the maximum was 3.92 m/s.

Table 7 – Summary of data for normal velocity distributions used in test 25.

	Mean velocity (m/s)	Standard deviation (m/s)	Minimum velocity (m/s)	Maximum velocity (m/s)
Velocity distribution 1	1.34	0.26	0.46	1.61
Velocity distribution 2	1.23	0.22	0.6	1.6
Velocity distribution 3	1.61	0.59	0.71	3.92

Convergence of total evacuation time was assessed for all test configurations. For the configuration using the default velocity distribution convergence criteria were reached at 18 runs. For tests using the velocity distribution measured by Cao et al. (2016) convergence criteria were reached at 25 runs. For tests using the velocity distribution measured by Gales et al. (2020) convergence criteria were reached at 30 runs. Density and flow were averaged across the number of runs needed for convergence. Only data between 90 to 240 second was used during which a steady flow was observed.

2.4.2 Results and comparison

In Table 8 averaged data for density and specific flow for each test configuration is presented. Standard deviation and minimum and maximum values presented are for data averaged across the number of runs needed for convergence. The higher spread of the flow data in comparison to the density data can partly be attributed to the measuring method. The tripwire detects agents passing a line each second which varies more than the area measuring method used for density.

Table 8 – Time averaged density and flow data for each test configuration.

	Velocity distr. 1		Velocity distr. 2		Velocity distr. 3	
	Density(m^{-2})	Flow ($s^{-1}m^{-1}$)	Density(m^{-2})	Flow ($s^{-1}m^{-1}$)	Density(m^{-2})	Flow ($s^{-1}m^{-1}$)
Average	2.25	0.93	2.30	0.93	2.31	0.97
Std. dev.	0.07	0.13	0.05	0.10	0.06	0.09
Min	1.97	0.60	2.12	0.63	2.11	0.72
Max	2.48	1.30	2.47	1.17	2.47	1.19

In Figure 15 simulated results for each test configuration are overlayed onto experiment results. Due to crowd:it only using tread depth when simulating stairs, the simulation geometry would be the same for both upper and lower tier stairs of the stadium grandstand tested in the experiment. Thread depth was consistent and only stair riser varied between these data sets. The slope for the upper tier stairs was 35° and 27° for the lower tier stairs. The means of the simulated results are well within the spread of the values observed for the upper tier stair in the experiment. The specific flow and density in experiments are in the lower part of the values observed for the lower tier stair which overall had a higher specific flow and density. The test with velocity distribution 3 showed a slightly higher specific flow than the other tests but only by 5%.

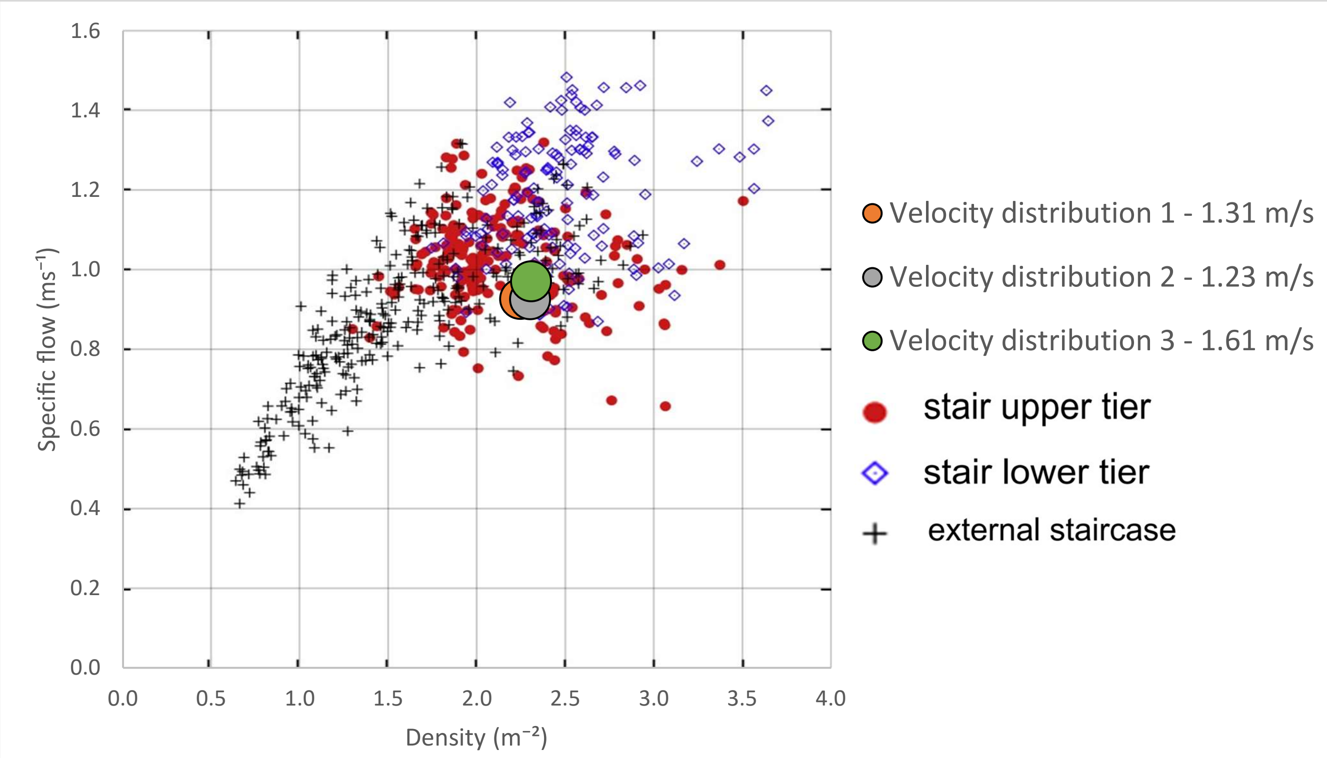

Figure 15 – Specific flow over density for averaged data from simulations per test configuration overlayed with data from experiment (Burghardt et al., 2013). Experiment data presented in red, blue, and black. Averaged data from simulations presented in orange, grey, and green.

2.5 Test 26. Movement around a corner

The ability of the evacuation model to represent evacuation movement around a corner was tested. The ISO protocol instructs that navigation and movement of simulated occupants is to be compared to an evacuation scenario in which unidirectional evacuation movement around a corner has been observed.

Test design and experimental data for comparison were taken from experiments conducted by Nilsson and Petersson (2008). In the experiment participants were recorded transversing a 90-degree corner in groups of 5 and 75. A 1.6 metres wide L-shaped corridor was built for the experiment. Three cameras were placed above the three sections of a corner aimed straight down. From the recorded tests rate of usage and usage time of each 0.1 m by 0.1 m area was extracted. Participants in the experiment were university students or personnel with the reported mean age of 24 years with a standard deviation of 5.4 years. Mean weight and length were reported as 74 kg and 180 cm with the standard deviations of 11 kg and 9.4 cm. 71% reported being men. Data was only collected from 70 of the 75 participants.

2.5.1 Simulation configuration

The geometry was represented in accordance with Figure 16 to recreate dimension in the experiment by Nilsson and Petersson (2008). Starting area was limited to 1 meter depth for test variation with 5 people to keep group together at start.

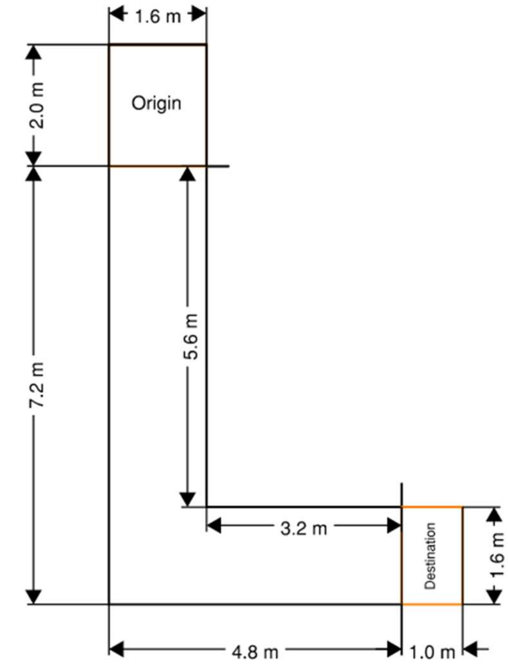

Figure 16 – Geometry of test design for test 26.

Agents were set to travel from the origin to the destination through the L-shaped corridor. Tests were performed with both population of 5 and 75 uniformly distributed people in the origin area within the first second or until population size was reached. A heat map was configured showing the number of agents passing each tile. Tiles in the heat maps in crowd:it are restricted to squares with minimum side of 0.5 m. To track the whole width of the corridor tile size was set to squares with the side of 0.533 m.

Body size of simulation population was set to default distribution for all test since no explicit data was given for the experiment population. Pre-movement time was set to 0 s.

In addition to the default velocity distribution (velocity distribution 1) the tests were run with two other velocity settings using normal distributions explicit for students. The data for the distributions are presented in Table 9. Velocity distribution 2 was slower than the default distribution with the mean of 1.23 m/s and a standard deviation of 0.22 m/s (Cao et al., 2016). The minimum recorded free velocity was 0.6 m/s, and the maximum was 1.6 m/s. Velocity distribution 3 was faster than the default distribution with the mean of 1.61 m/s and a standard deviation of 0.59 m/s (Gales et al. 2020). The minimum recorded velocity was 0.71 m/s, and the maximum was 3.92 m/s.

Table 9 – Summary of data for normal velocity distributions used in test 25.

	Mean velocity (m/s)	Standard deviation (m/s)	Minimum velocity (m/s)	Maximum velocity (m/s)
Velocity distribution 1	1.34	0.26	0.46	1.61
Velocity distribution 2	1.23	0.22	0.6	1.6
Velocity distribution 3	1.61	0.59	0.71	3.92

Convergence of total evacuation time was assessed. Runs needed until each convergence criteria were met for each scenario are presented in Table 10. Heat map values were averaged across the number of runs needed for convergence criteria to be met.

Table 10 – Number of runs needed for each scenario until convergence criteria were met.

	Velocity distribution 1	Velocity distribution 2	Velocity distribution 3
5 agents	63 runs	61 runs	71 runs
75 agents	26 runs	38 runs	25 runs

2.5.2 Results and comparison

Data of number of agents stepping into each square of the corner was collected from simulations and averaged across the number of runs needed to meet convergence criteria. The data was normalized by the number of agents in the simulation to give rate of usage and presented in a heat map. Due to minimum possible tile size in the software being squares with side length of 0.5 m, simulation results are very coarse compared to experiment results which is given per 10 cm squares. Heat maps in crowd:it only registers agents step which can lead to agents skipping tiles. In the experiment the full path of participants was examined.

In Figure 17 rate of usage from experiments with 5 participants are presented. The authors note that most participants take a line close to the inner wall and the rest concentrate to a second line 0.3 to 0.7 meter further out (Nilsson & Petersson, 2008). The shorter distance between the lines is observed in the corner and the distance is farther before and after the corner.

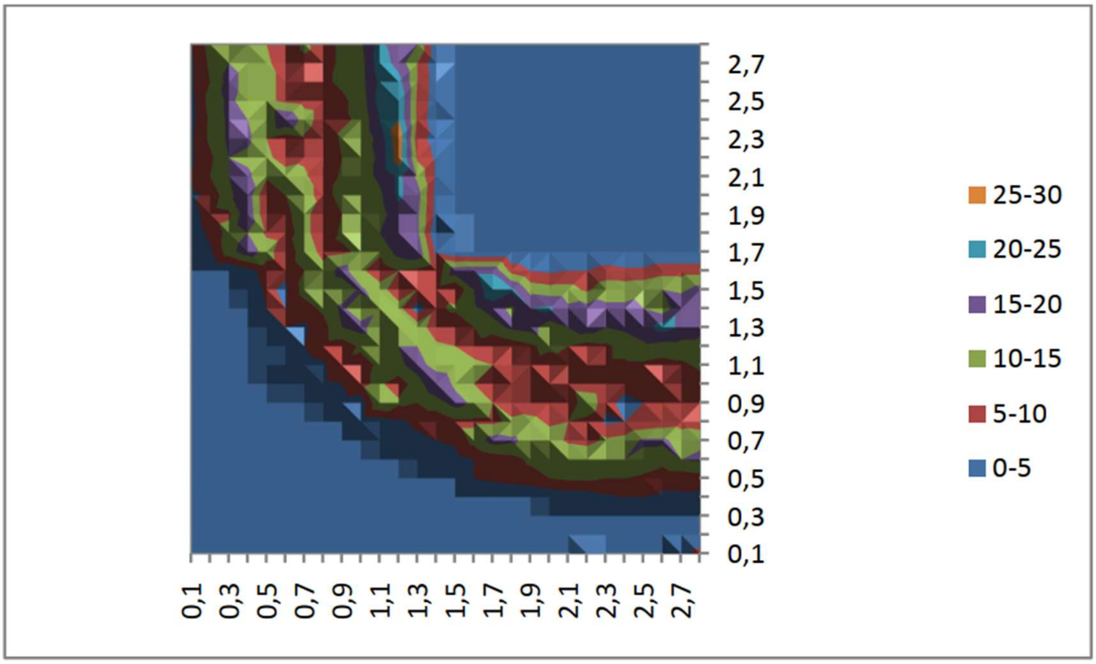

Figure 17 – Rate of usage for experiment with 5 participants. Reproduced from paper by Nilsson and Petersson (2008).

In Figure 18 rate of usage for simulations with 5 agents with velocity distribution 1 is presented in percentage as well as colour graded from red to green according to the value of each cell. Red is coded to the for the lowest value and green for the highest value. The value of each tile is averaged across the number of runs needed to meet convergence criteria. Due to the averaging across runs and to agents sometimes stepping past tiles each row does not necessarily sum to 100%. Each axis presents distance in meters. Due to Nilsson and Petersson (2008) presenting cell values as point values on a surface and the size of simulation heat map tiles the axes differ slightly between charts.

Y							
3.2							
2.7	9	71	7	0	0	0	
2.1	8	69	10	0	0	0	
1.6	7	70	15	0	0	0	
1.1	1	56	45	11	6	5	
0.5	0	12	55	77	75	82	
0.0	0	0	0	2	3	4	
X	0.0	0.5	1.1	1.6	2.1	2.7	3.2

Figure 18 – Rate of usage in percentage for simulations with 5 agents with velocity distribution 1.

In Figure 19 rate of usage for simulations with 5 agents with velocity distribution 2 is presented. Results are very similar to results from test with the default velocity distribution.

Y							
3.2							
2.7	8	73	8	0	0	0	
2.1	8	71	11	0	0	0	
1.6	9	65	16	0	0	0	
1.1	3	59	39	12	7	5	
0.5	0	15	61	79	79	80	
0.0	0	0	0	2	3	2	
X	0.0	0.5	1.1	1.6	2.1	2.7	3.2

Figure 19 – Rate of usage in percentage for simulations with 5 agents with velocity distribution 2.

In Figure 20 rate of usage for simulations with 5 agents with velocity distribution 3 is presented. Values are overall lower than results from test with default velocity. This is probably due to agents travelling further in each step.

Y							
3.2							
2.7	5	58	7	0	0	0	
2.1	4	62	5	0	0	0	
1.6	5	58	10	0	0	0	
1.1	1	40	42	7	4	2	
0.5	0	8	41	66	67	65	
0.0	0	0	0	1	1	1	
X	0.0	0.5	1.1	1.6	2.1	2.7	3.2

Figure 20 – Rate of usage in percentage for simulations with 5 agents with velocity distribution 3.

Even though simulation results are coarse a few aspects can be noted when comparing to experiment results. The outside corner is not used in neither the experiment or the simulations. In the simulations agents kept to the middle and no evidence of walking in dual parallel lines can be seen. Agents walk close to the wall in the corner similar to experiment participants.

In Figure 21 rate of usage for experiments with 75 agents is presented. The authors note that the tendency for lane development is still present for tests with 75 participants. Most participants tended to take the outside line. The distance between the outer and inner lane increased compared to tests in the experiment with 5 participants. They estimate it to 0.5 to 0.9 m with the shorter distance being observed in the corner as in the tests with 5 participants. They also note the presence of a middle lane, but it was not as prominent as the two other lanes.

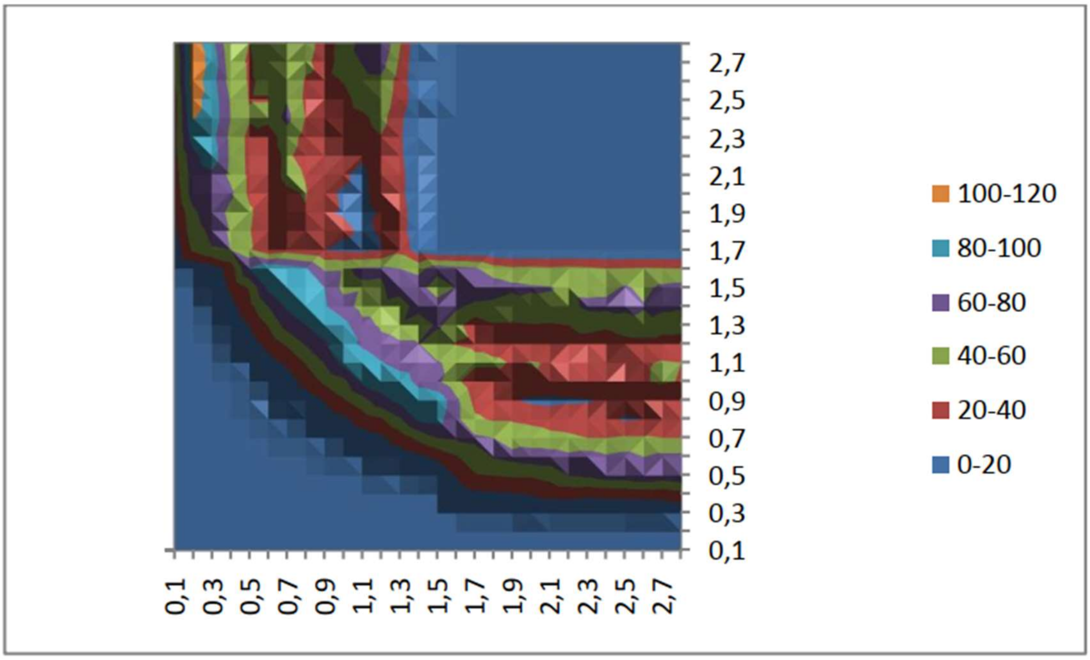

Figure 21 – Rate of usage for experiment with 75 participants at normal walking pace. Reproduced from paper by Nilsson and Petersson (2008).

In 10 additional tests the 75 participants were asked to simulate crowding behaviour by walking closer together. Rate of usage from these tests are presented in Figure 22. During these tests three lanes with similar rate of usage were observed.

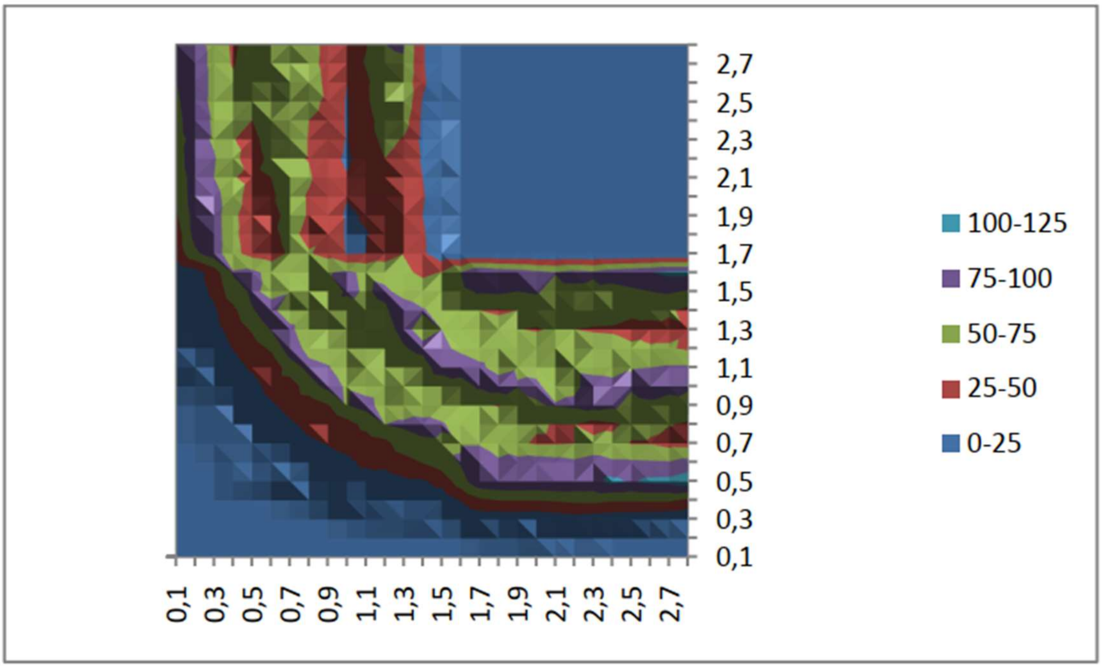

Figure 22 – Rate of usage for experiment with 75 participants at hurried walking pace. Reproduced from paper by Nilsson and Petersson (2008).

In Figure 23 rate of usage for simulations with 75 agents set to velocity distribution 1 is presented in percentage as well as colour graded from red to green according to the value of each cell, red for the lowest value and green for the highest value. Each axis presents distance in meters. Due to the averaging across runs and to agents sometimes stepping past tiles or sidestepping into adjacent tiles each row does not necessarily sum to 100%. Due to Nilsson and Petersson (2008) presenting cell values for rate of usage as point values on a surface and the size of simulation heat map tiles the axes differ slightly between charts.

Y							
3.2							
2.7	25	50	25	0	0	0	
2.1	23	47	29	0	0	0	
1.6	22	42	36	0	0	0	
1.1	17	38	46	30	26	26	
0.5	10	37	45	52	52	52	
0.0	1	7	17	21	21	21	
X	0.0	0.5	1.1	1.6	2.1	2.7	3.2

Figure 23 – Rate of usage per cell in percentage for simulations with 75 agents set to velocity distribution 1.

In Figure 24 rate of usage for simulations with 75 agents set to velocity distribution 2 is presented. Results are very similar to results from test with the default velocity distribution but show a slightly higher use of the sides of the corridor.

Y						
3.2						
2.7	26	50	27	0	0	0
2.1	23	46	31	0	0	0
1.6	22	43	37	0	0	0
1.1	18	39	45	32	28	26
0.5	10	38	46	52	52	52
0.0	1	7	17	21	22	22
X	0.0	0.5	1.1	1.6	2.1	2.7

3.2

Figure 24 – Rate of usage per cell in percentage for simulations with 75 agents set to velocity distribution 2.

In Figure 25 rate of usage for simulations with 75 agents set to velocity distribution 3 is presented. Values are overall lower than results from test with default velocity. This is probably due to agents travelling further in each step.

Y						
3.2						
2.7	21	40	21	0	0	0
2.1	21	38	24	0	0	0
1.6	21	35	29	0	0	0
1.1	14	34	39	25	21	21
0.5	10	31	40	45	43	45
0.0	1	6	13	20	19	20
X	0.0	0.5	1.1	1.6	2.1	2.7

3.2

Figure 25 – Rate of usage per cell in percentage for simulations with 75 agents set to velocity distribution 3.

Rate of usage for simulations with 75 agents are more evenly distributed across the width of the corridor than in tests with 5 agents. Although the same tendency to keep to the middle can still be observed. The tendency to cut the corner is still present but usage of the rest of the corner increased compared to tests with 5 agents.

In Figure 26 screenshots from 3 simulations of 75 agents set to velocity distribution 1 are presented with agents' paths traced. No clear formations of lane can be seen but some areas seem to be used more than others in agreement with the heatmap. Agent do seem to zigzag or move sideways across the corridor to a great extent. This is not in line with the tendency to follow the person in front Nilsson and Petersson (2008) observed.

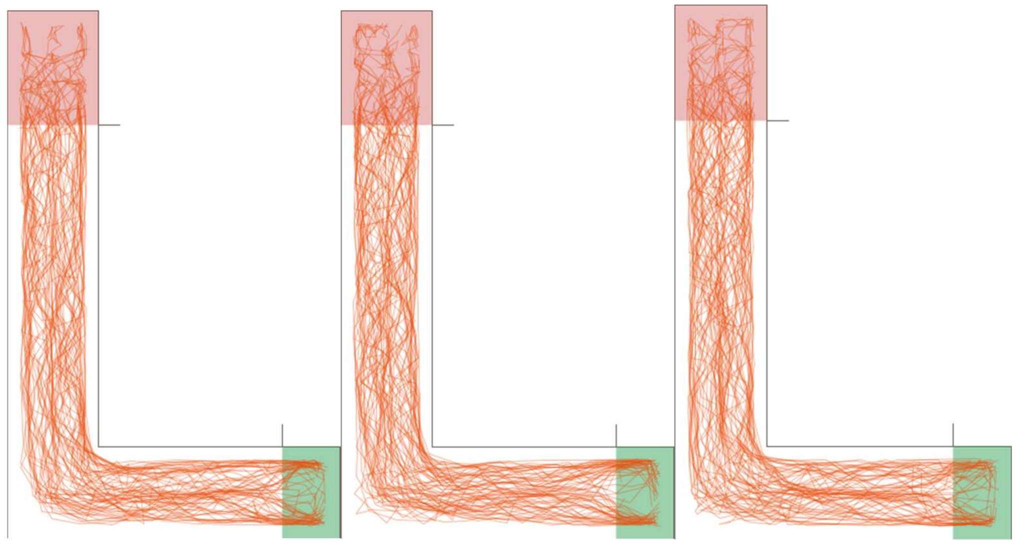

Figure 26 – Screenshots from end of 3 simulations with 75 agents with agents' paths traced.

The behaviour of agents in the simulation is to be expected of the Optimal Steps Model. Agents have no social behaviours and simply move to the next open position with the best value. Therefore, agents walk in the middle away from the walls when possible. This was seen most strongly in the configuration with 5 agents. In the corner, the shorter distance to the destination poses a greater value to the agents than wall proximity diminishes the value. This allows the agents to keep close to the inside corner. After passing the corner the agents revert back to the middle of the corridor to keep away from the wall which exerts a repulsing potential on the agents. In the configuration with 75 agents the proximity to other agents somewhat moves agents away from the middle, as expected. The crowding limits the agents from walking at their desired velocity and agents therefore zigzag across the corridor rather than maintaining the queue-like behaviour which was observed in the experiments.

2.6 Test 27. Counter-flows

The ability of evacuation model to represent bi-directional pedestrian counter-flows was tested. The ISO protocol instructs that simulated flows and movement behaviour is to be compared to behaviour in an evacuation scenario which includes bi-directional pedestrian counter-flows (ISO, 2020). The counter-flow behaviour can be for horizontal egress components, as corridors, and for vertical egress components, such as stairs.

Test design and experimental data for comparison were taken from experiment conducted by Kretz et al. (2006). In the experiment 67 participants walked through a 34 metres long corridor in varying groups and with varying ratio of counter-flow. The amount of counter-flow tested, given in a ratio of the total participants in each test, was 0, 0.1, 0.34, and 0.5. Cameras were mounted at 3 points 5 metres apart in the corridor and passing times, walking speed, and specific flux was presented. The participants began to walk from 5 metres away from the first measuring line at a given signal. Kretz et al. (2006) describe the participants of the experiment as mostly students of Duisburg-Essen University and in their twenties. 34 were female and 33 were male. No other information was given about the participants.

2.6.1 Simulation configuration

The geometry was represented in accordance with Figure 27 to recreate the layout in the experiment (Kretz et al., 2006)

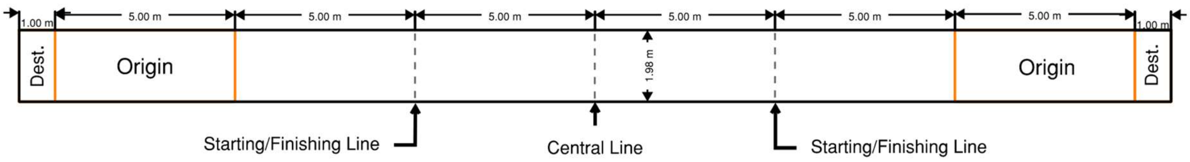

Figure 27 – Geometry of the test design in accordance with the counterflow experiment (Kretz et al., 2006).

In the experimental corridor the width broadened by 40 cm at 0.98 m above the floor making it possible for the participants to somewhat lean out above the low wall as represented in Figure 28. Since crowd:it represents geometry in 2D this was not possible to represent, and width was set to 1.98 m.

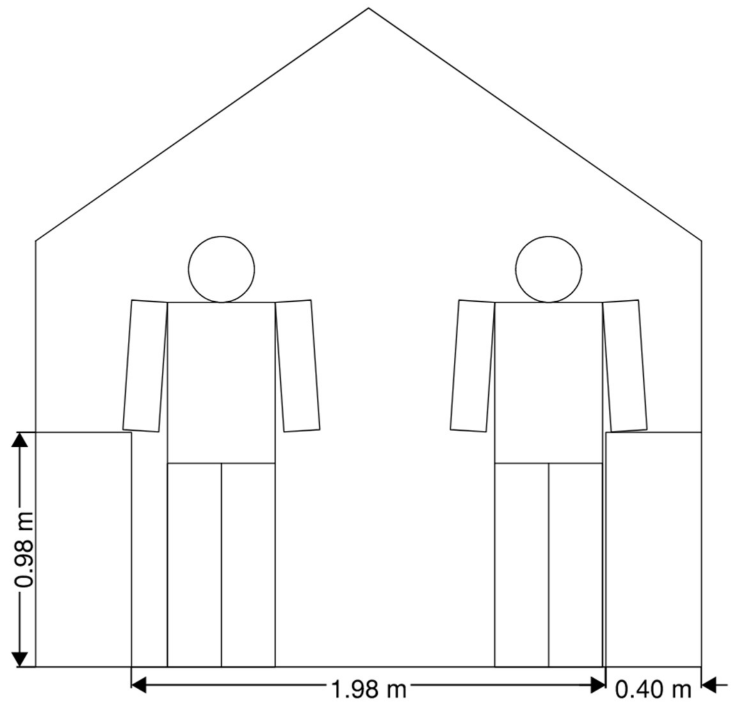

Figure 28 – Cross section view of experiment corridor.

Agent origins and destinations were set up in a way that the groups were determined to cross each other. The left origin was connected to the right destination and vice versa. Tripwires were placed at lines 5 meters from each origin to record passing times for each agent group in accordance with the experiment. Tests were performed with different populations without counter-flow and with a third of the agents walking in the opposite direction.

Pre-movement time was set to 0 s. The body size of the population was set to default distribution. Both the default velocity distribution as well as a faster velocity distribution specific to student were used. The faster velocity distribution measured by Gales et al. (2020) is a normal distribution with the mean of 1.61 m/s and the Table 11.

Table 11 – Summary of data for normal velocity distributions used in test 27.

	Mean velocity (m/s)	Standard deviation (m/s)	Minimum velocity (m/s)	Maximum velocity (m/s)
Velocity distribution 1	1.34	0.26	0.46	1.61
Velocity distribution 2	1.61	0.59	0.71	3.92

Simulations were run with no counterflow for both 35 and 67 agents. Counter-flow tests were run for population of 67 and 51 with a counter-flow ratio of 0.34, i.e. 44 versus 22 and 34 versus 17. Convergence of total evacuation time was assessed for each setup with differing population. Number of runs needed for convergence criteria for each test configuration is presented in Table 12. Convergence was not met for cases with counterflow at 100 runs and was not assessed further because of the unpredictable behaviour that occurred at counterflow which is discussed further in results. For the test configurations without counter-flow passing times were averaged over the number of runs needed for convergence to be met. For test configurations with counter-flow passing times were averaged over 60 runs.

Table 12 – Number of runs needed to meet convergence criteria for each test configuration.

Agents in flow vs. counterflow	67 vs. 0	33 vs. 0	44 vs. 22	34 vs. 17
Velocity distribution 1	40 runs	33 runs	No conv. at 100 runs	No conv. at 100 runs
Velocity distribution 2	40 runs	37 runs	No conv. at 100 runs	No conv. at 100 runs

2.6.2 Results and comparison

The passing times was measured in the simulation at the starting and finishing line. Averaged passing times at no counterflow are presented in Table 13 along with standard deviation and maximum and minimum passing times. Passing time is here defined as the time from the first until the last person passes the chosen line.

Table 13 – Passing times for simulations at no counterflow.

	Group size	67		35	
	Position	Starting line	Finishing line	Starting line	Finishing line
Velocity distribution 1	Average (s)	30	36	17	23
	Std. dev. (s)	1	2	1	3
	Min. (s)	28	32	7	15
	Max. (s)	34	42	22	33
Velocity distribution 2	Average (s)	26	34	15	22
	Std. dev. (s)	1	2	1	3
	Min. (s)	24	30	6	13
	Max. (s)	29	37	19	29

In Figure 29, the simulation results for tests with no counter-flow are overlayed with experimental results for tests with no counter-flow (Kretz et al., 2006). The sample of simulated results is far smaller than the results from the experiment, but simulation results seem to follow a similar trend as the experiment result. Simulation passing times are slower than experiment times except for starting line times using velocity distribution 2. In experiment the difference between passing times at the different positions was merely seconds indicating a small variance of speed across the population. In the simulations the difference in passing times between the starting and finishing line are much greater. This indicates a greater variance in speed across the population.

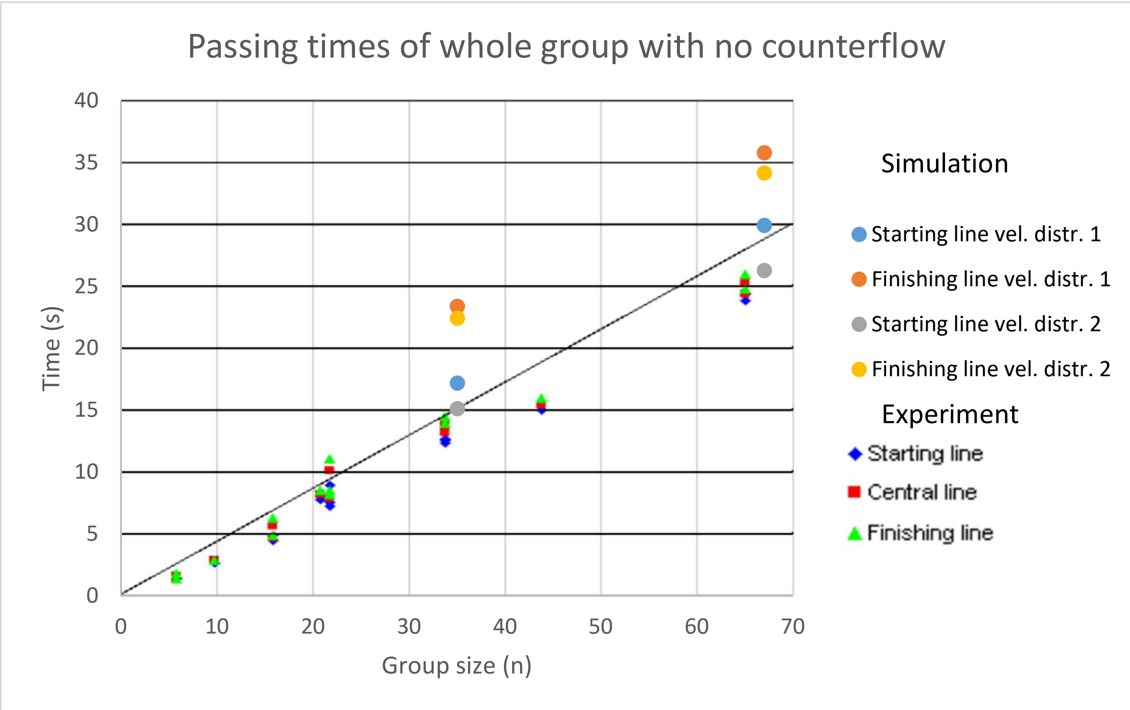

Figure 29 – Passing times at no counterflow for both simulation and experiment (Kretz et al., 2006). Passing time is measured as time from first person passes a line until the last person passes the line.

Passing times for tests with one third of the population in counterflow are presented in Table 14 along with standard deviation across averaged runs and maximum and minimum passing times. Passing times generally rose significantly from tests without counterflow. This is due to agents getting stuck in the middle when the two groups collide. The problem is worse for the tests with a larger population(44 vs. 22) than for the tests with a smaller population(34 vs. 17). Judging by the minimum passing times of the finishing line at least one run with the population of 34 versus 17 agents seems to have resolved without major standstills. This was especially true for simulations using velocity distribution 2.

Table 14 – Passing times at 34% counterflow for simulations.

		Majority groups				Minority groups			
	Group size	44		34		22		17	
	Position	Starting line	Finishing line	Starting line	Finishing line	Starting line	Finishing line	Starting line	Finishing line
Velocity distribution 1	Average (s)	246	693	61	181	30	689	11	176
	Std. dev. (s)	270	312	71	97	70	312	2	97
	Min. (s)	33	202	17	53	10	198	8	39
	Max. (s)	1144	1608	391	452	408	1602	20	449
Velocity distribution 2	Average (s)	283	567	55	131	21	561	12	125
	Std. dev. (s)	288	306	53	78	25	308	3	79
	Min. (s)	20	191	16	33	11	168	7	23
	Max. (s)	1095	1403	251	389	185	1401	19	386

In Figure 30 mean passing times for simulations with 34% counter-flow are overlayed experiment passing times for tests with 34% counter-flow. Most simulation mean passing times vastly exceeds the data measured in the experiment and are therefore not presented in the graph. Only three mean passing times of the starting line from the simulations are within 25 seconds. With basis in the data of simulations with no counter flow, one would expect simulation passing times for tests with counterflow to be slightly higher and with a slightly larger spread between times for starting finishing line across the board.

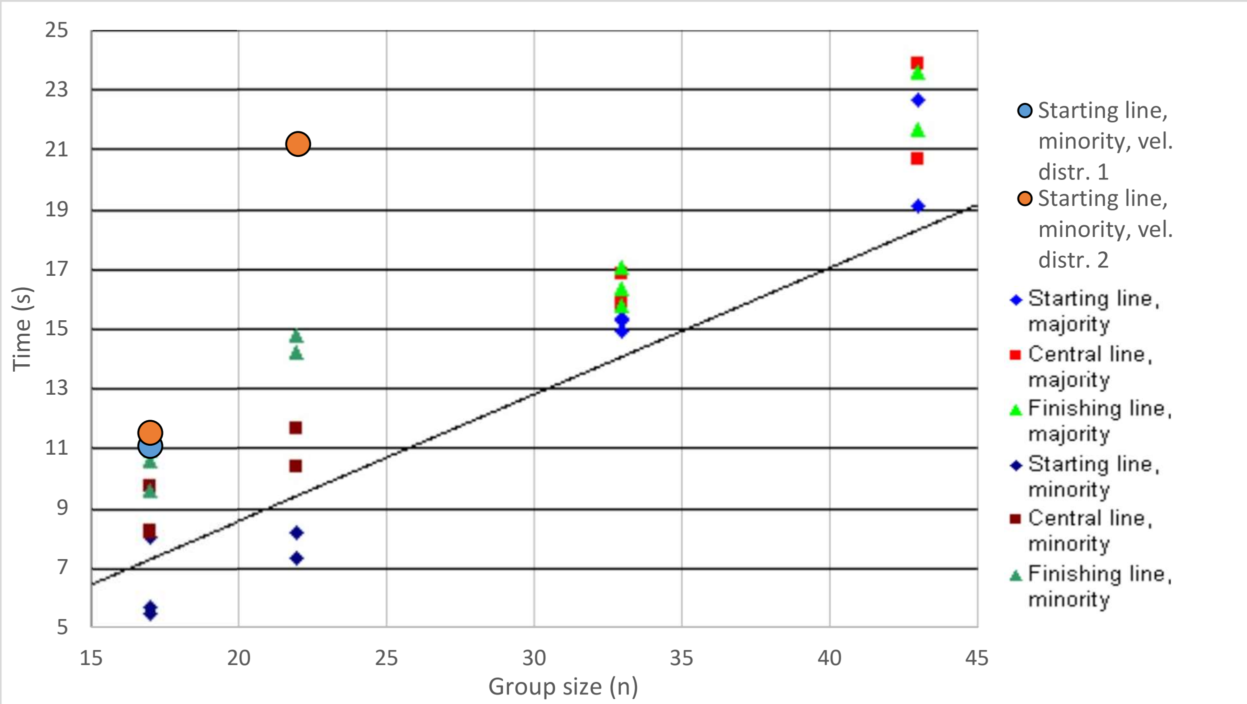

Figure 30 – Passing times 34% counterflow from experiment (Kretz et al., 2006).

In the simulations a complete standstill can be observed for most of the counterflow cases with agents slowly moving through the crowd until the population is small enough to resolve. In Figure 31 two snapshots of the simulation can be seen with 2,5 minutes difference. Between snapshots 10 agents have made it to their destination. Roughly a minute later total evacuation was achieved. In the experiment lane formation was observed (Kretz et al., 2006). This was not present in the simulations. The standstill in the simulations is probably due to the selfish nature of the Optimal Steps Model in which no crowd behaviour is modelled, and agents only want to optimize their own next step.

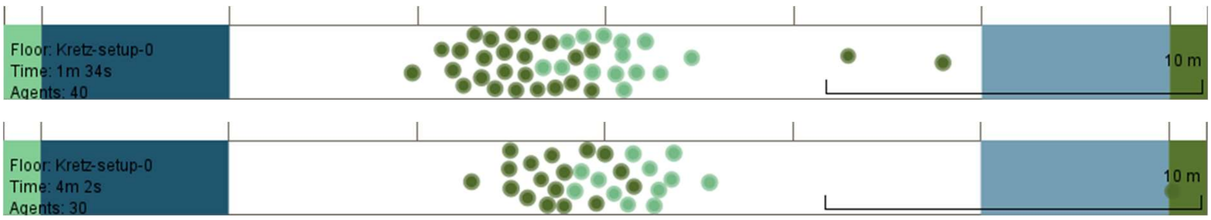

Figure 31 – Screenshots of 9th run of case with 34 vs. 17 agents using velocity distribution 1 at roughly 1,5 (upper) and 4 minutes (lower). Agents are in the colour of their destination.

2.7 Test 28. Route/exit choice

The ability of evacuation models to represent route/exit choice was tested. The ISO protocol instructs that simulated pedestrian route/exit choice and behaviours is to be compared to observed experimental behaviour in a scenario in which an individual or a crowd has to select the rout/exit for evacuation in relation to the availability to different way-finding installations. Exit choice can in crowd:it only be explicitly modelled with a deterministic approach.

Test design and experimental data for comparison were taken from experiments conducted by Nilsson et al. (2005).They performed three evacuation experiments which investigated how the design of emergency exits affects evacuation exit choice.

2.7.1 Simulation configuration

The geometry was represented as shown in Figure 32 to recreate the layout in the experiment by Nilsson et al. (2005). The experiment corridor was 3 m wide but due to shelves on the wall, the effective width was reduced to 2.6 m. For experiment 1 there was an exit at each end of the corridor. The first starting position was at equal distance from both exits. For the second starting position the distance to exit 1 was twice the distance to the second exit. For experiment 2 the starting position was at the end of the corridor. There was an exit at the other end of the corridor as well as an alternative exit halfway through the corridor. All doorways were 0.8 m wide but since flow was of no concern in this test this had no impact on the result. All origins were 1x1 m^2.

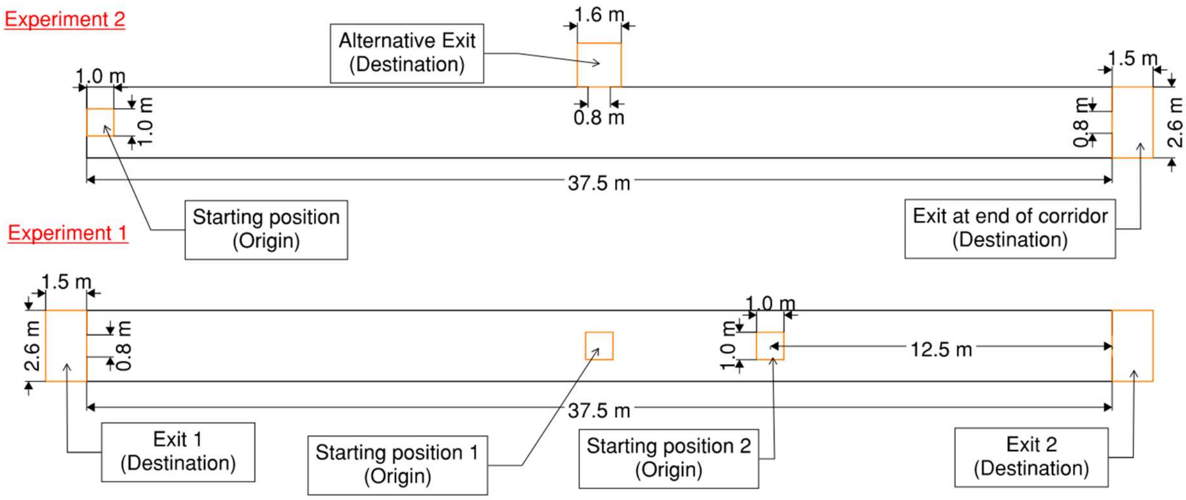

Figure 32 – Geometry of the test design in accordance with experiments by Nilsson et al. (2005).

Agents were set to walk from origin to the destinations. The destinations were added to a set with a relative threshold on the number of agents allowed for each destination. The thresholds for each exit were matched to results from experiments by Nilsson et al. (2005). 3 tests were done for each starting positions of experiment 1 and 2 test were done for experiment 2 with varied distributions for each exit. In the experiments by Nilsson et al. (2005) the experiments were done individually. To ease the workload, simulations were run with multiple agents at the same time since agents do not influence other agents' exit choice as would be the matter in the experiments. Simulations for each case were run with same the number of agents reported for each scenario by Nilsson et al. (2005).

Since no speed or flow was measured agent characteristics had no impact on result. Body size and walking speed of the population was therefore set to default distributions. Pre-movement time was set to 0 s. The thresholds for each exit were matched to results from experiments by Nilsson et al. (2005).

Since time or flow had no impact on the results the convergence of total evacuation time was not assessed for this test. Each scenario was run 10 times to ensure consistency and in fact it gave identical results.

2.7.2 Results and comparison

In the experiment by Nilsson et al. (2005) the way-guiding system was varied at the exits in accordance with Table 15 giving results also presented in Table 15. The 8 scenarios were simulated with matching start positions. The threshold of each exit was set to match the result of each scenario. Across the 10 runs for each scenario the number of participants who walked to each exit were identical to experiment results.

Table 15 – Results from experiment by Nilsson et al. (2005).

Experiment	Scenario	Start position	Way-guidance system at exit 1		Daylight at exit 2	Number (proportion) of participants who walked to	
			Flashing lights	Strobe light		Exit 1	Exit 2
Experiment 1	1	1	No	No	No	8 (50 %)	8 (50 %)
	2	1	Green	No	No	9 (75 %)	3 (25 %)
	3	1	No	Yes	Yes	11 (65 %)	6 (35 %)
	4	2	No	No	No	12 (75 %)	4 (25 %)
	5	2	No	No	No	11 (58 %)	8 (42 %)
	6	2	Green	No	No	7 (88 %)	1 (12 %)

Experiment	Scenario	Strobe light at alternative exit	Number (proportion) of participants who walked to	
			Alternative exit	Exit at the end of the corridor
Experiment 2	7	No	6 (55 %)	5 (45 %)
	8	Green	13 (93 %)	1 (7 %)

2.8 Test 29. Bottlenecks at openings

The ability of evacuation models to represent bottlenecks at openings was tested. The ISO protocol instructs that simulated pedestrian movement flows, congestion levels and behaviour at opening and in the area close to the opening is to be compared to the observed experimental behaviour in an evacuation scenario in which a crowd has to pass through an opening. The behaviour at openings is in crowd:it is emergent due to the way the Optimal Steps Model is implemented.

Test design and experimental data for comparison were taken from experiment conducted by Nicolas et al. (2017). It examined pedestrian flows through a narrow doorway. It was a controlled experiment conducted in the gymnasium of Centro Atómico Barloche in Argentina. The 80 participants were instructed to pass through the 0,72 m wide doorway and then circle around and join the crowd again from the back. Some participants were asked to behave more selfishly and were instructed to elbow their way through the crowd, with mild contact but no violence. All others were asked to avoid any contact and try to keep their distance. The experiment was also conducted with different general instructions. The first series of the experiments, referred to as the experiment with placid walkers, participants were only instructed to head for the door. In the second series of experiments, referred to as hurried walkers, participants were instructed to hurry a bit more but without running, pushing, or hitting others. Two cameras were placed above the doorway from which density locally around the doorway and the flow through the doorway were extracted.

2.8.1 Simulation configuration

The geometry was represented as shown in Figure 33 to recreate the layout in the experiment by Nicolas et al. (2017). Since the pedestrians in the experiment circled round to the start after going through the door the destination was extended to the sides to simulate the movement to the sides after exiting. Density was measured in in a rectangular zone (0.5m depth×0.84m width) in front of the door to match the density zone given in the experiment as 0.42 m^2. A tripwire was setup in the door from which the number of agents passed each second was gathered.

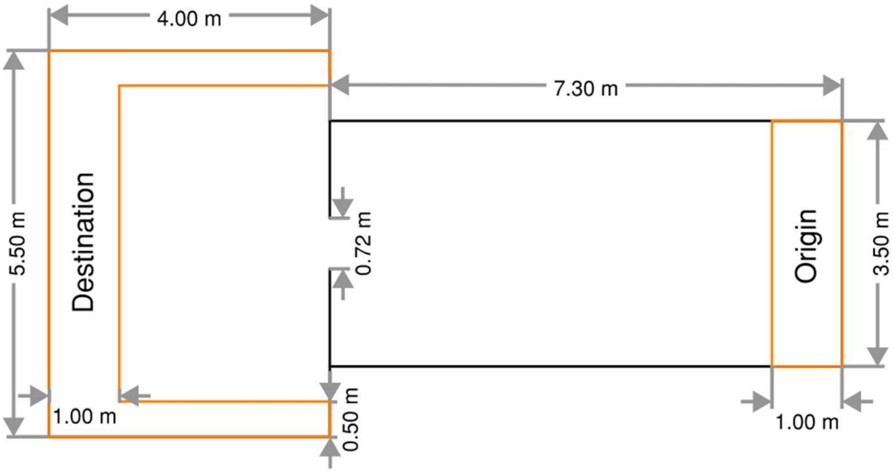

Figure 33 – Geometry of the test design in accordance with experiment by Nicolas et al. (2017).

250 agents were set to move from origin to destination through a 0.72 m wide door opening. Agents spawn in as room is given in the origin area to simulate the recircling of pedestrians in the experiment (Nicolas et al., 2017).

Body size of population and walking speed of population was set to default distributions. Pre-movement time was set to 0 s. In the experiment the population consisted of 80 participants aged 20 to 55 with a woman to men ratio of 1 to 3. Participants were reported to be students and researchers. No additional data was given for the population and therefore default values were used.

Convergence of total evacuation time was assessed and reached at 31 runs. Flow through the door and local density in the area in front of the door was averaged for the 31 runs. Only data between 30 and 150 seconds was used for the results to ensure steady flow.

2.8.2 Results and comparison

The number of agents that passed the tripwire each second was given from the software. This data was used to calculate flow by comparing values for each passing second. Local density was given from rectangular area in front of the door. Maximum values for each second were given from the software. Average values were calculated from data points between 30 and 150 s across 31 runs. Values for measured local density and flow from experiment was 3.90 m^{-2} and 1.36 s^{-1} and are presented alongside measured values in the experiment in Table 16 (Nicolas et al., 2017).

C_s refer to the amount of the population who in the reference experiment was asked to behave selfishly. C_s^* refers to the effective amount behaving selfishly since they tended to circle round more effectively. Nicolas et al. (2017) also tested scenarios where the population were asked to "Head for the door more hurriedly". In the experiment a pre-factor $\alpha \doteq 1$ was introduced in the definition of the density unit to reflect the uncertainty of the camera placement.

Table 16 – Flow and density data from simulation and experiment (Nicolas et al., 2017)

Sim.				Density (m^{-2})	Flow rate (s^{-1})	Flow rate [(sm^{-1})]
				3.90	1.36	1.80
Experiment		C_s	C_s^*	Density (αm^{-2})	Flow rate (s^{-1})	Flow rate [$(sm)^{-1}$]
	Placid walk	0%	0%	2.69	1.01	1.40
		30%	45%	4.09	1.35	1.88
		30%	47%	4.94	1.41	1.96
		60%	71%	6.04	1.71	2.38
	Hurried walk	0%	0%	3.70	1.26	1.75
		10%	18%	4.49	1.39	1.93
		60%	71%	7.63	2.20	3.06
		90%	92%	8.26	2.36	3.28
		100%	100%	8.98	2.41	3.35

In the experiment far higher densities and flows were recorded for scenarios with higher percentage of people with selfish behaviour. Especially in scenarios where people were asked to hurry. The higher flow rates can be attributed to several participants in the experiment by Nicolas et al. (2017) passing through the doorway simultaneously as can be seen in Figure 34. Since the doorway was smaller than the size of two agent torso diameters, two agents would not be able to fit in the doorway at the same time during the simulations. The high density in the experiment can be attributed to the compression of the bodies at higher density which cannot be represented in this model. It is also a specific condition driven by the instruction given to the participants.

Figure 34 – A randomly selected video frame from the experiment taken from Nicolas et al., (2017).

Although higher and lower values for flow and density was recorded the averages from the simulation can be considered not too dissimilar from the measured results as can be seen in Figure 35. When flow and density from the simulation are compared to flow and density of scenario with next higher density (placid walk, $C_S^* = 45\%$) the differences are 1.1% and 4.9% for flow and density respectively. Compared to the next lower density (hurried walk, $C_S^* = 0\%$) the differences are 7.7% for flow and 5.1% for density.

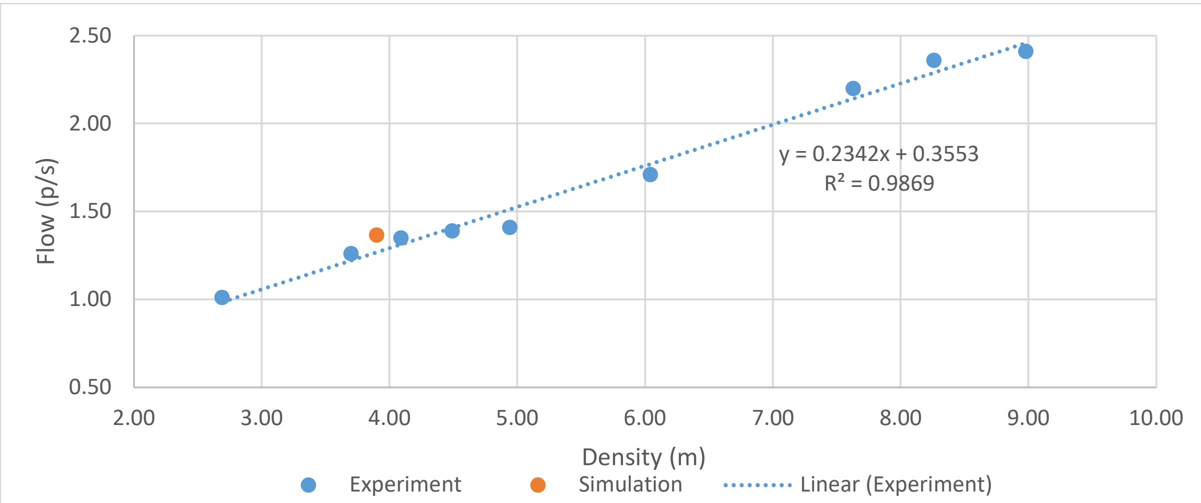

Figure 35 – Chart of flow dependence of density for measured and simulated result with a linear trendline fitted to the measured data points.

In Figure 36 screenshots from the simulation are presented in which the crowding in front of the door can be seen and how an agent passes through it from one second to the next.

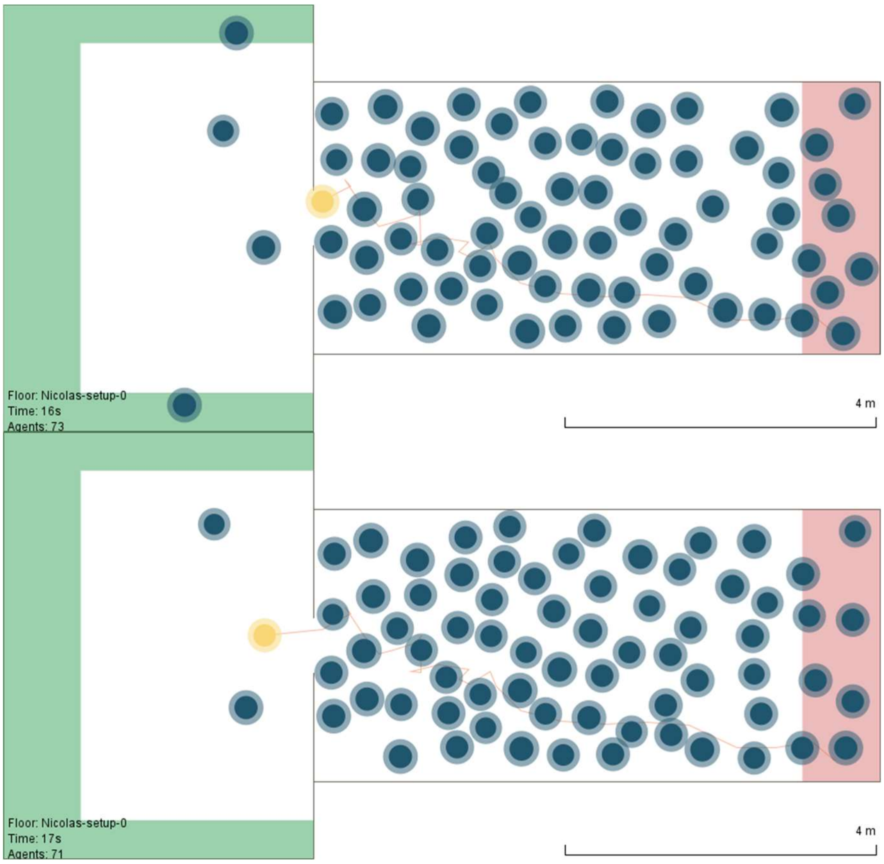

Figure 36 – Screenshots from simulation of 16 and 17 seconds from the top with the same agent highlighted.

2.9 Test 30. Reduced visibility vs walking speed

The ISO protocol instructs that simulated walking speed in relation to the visibility conditions are to be compared to an evacuation scenario in which a crowd has walked in a smoke-filled environment. There is no model in crowd:it that explicitly models the impact of smoke on agents and its impact on evacuation. Scaled velocity areas can be used to decrease the velocities of agents in an area which could be used to explicitly model the decreased walking speed in smoke. In the scaled area all agents' desired speeds are scaled by a factor. It is not possible to change the speed factor during simulation to simulate a change in visibility condition over time. Due to crowd:it not having an explicit model which models the consequences of smoke on evacuation behaviour no test was done.

2.10 Global validation test 1 – Theatre

During global validation full evacuation scenarios is to be tested. The test is replicating a full evacuation of an auditorium. Test design and experimental data for comparison were taken from experiment conducted by Imanishi and Sano (2018). Evacuation times and flow through exits are compared.

2.10.1 Simulation configuration

The floor geometry was represented in accordance with Figure 37 to recreate dimension in experiment by Imanishi and Sano (2018). The width of each seat as well as distance between the armrests to the back of the seat in front was given to 0.6 m. The inner width of the exits was all given to be 1.65 m. The minimum width of the middle aisles was given to 0.95 m and 0.75 m for the side aisles. All other measurements were approximated from schematic figures in the experiment paper.

Due to crowd:it only being able to represent stairs in rectangular shapes the geometry of the steps in connection to the front exits were simplified. The front stairs were modelled with the same width along the length of the stairs rather than the lower threads being wider as they are represented in the floor plans in the paper by Imanishi and Sano (2018). Some outer groups of seats not used in the experiments were removed to simplify the geometric representation. Each row of seats was at different elevation and the floor was stepped. For the simulations, the floor was represented as flat since the stepped floor was not possible to represent in crowd:it. The stair function in crowd:it was not able to be used for this purpose since the maximum tread depth is set to 0.40 m. This is because agents must be able to reach the next step of the stair in one step due to the design of the Optimal Stair Model. The stair function is also not able to represent angled treads and can only be set to rectangular simulation objects. Due to the rise of each step being small and the thread depth being approximately 1.2 m the walking behaviour is unlikely to resemble that to a stair. Though it seems likely that the stepped floor would have some impact on the velocity of the participants and perhaps the rhythm of the walk.   A directional scaled area could be used which scales the velocity in the positive and negative axis of a chosen direction. For example, walking up the auditorium could be scaled by a factor below 1 and walking down the steps could be kept to a factor 1 or a factor below or above 1. It is unclear though what factors to choose for the scaling since limited data are available for the calibration of this type of people movement. Therefore, no scaled areas were used.

The exits were named according to experiment. The agent origins were set to the seats in accordance with reported seating positions. In the first drill both doors of exit 1D were closed. In the third drill the outer doors of exits 1C and 1D were closed. In the first drill two small ladders to the stage were reported to be limiting the access to the walking space between the stage and the front row. No exact measurements were used. The ladders were not present in the subsequent drills and therefore not modelled since only drill 3 were modelled.

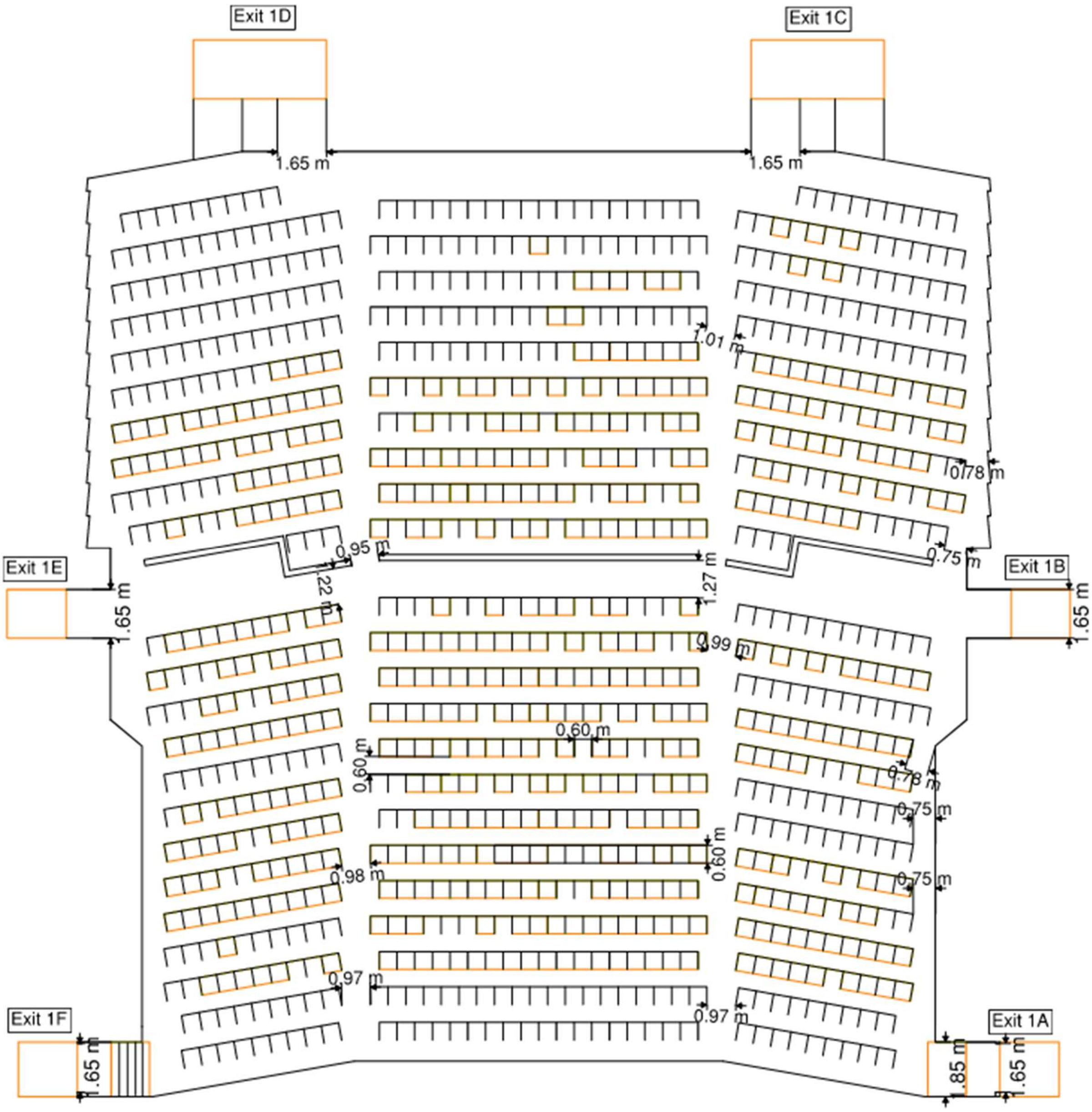

Figure 37 – Geometry of the test design in accordance with drill 3 of the experiment (Imanishi and Sano, 2018).

Participants agreed to participate in the drill upon booking tickets to the event. The drill scenario was that an earthquake occurred which triggered a fire on stage. The earthquake which had made no serious harm to the building was presented. The audience were instructed to stay in their seats, protect their head and await further instructions. After 5 minutes the fire broke out occupants were ordered to evacuate the theatre. Smoke effect and red-orange light was used to simulate the fire at the right-hand side of the stage.

Facility staff were present in the auditorium to control the flow during the drills, but the methods varied between drills. Three types of staff were present during the drills. During all three drills one staff member was always on stage making the announcements using a megaphone. Exit-door staff stood by each exit door and called out the location of the available exit. They waved LED traffic wands and called out in natural voices "Here is an exit". Aisle staff were distributed in the aisles and directed evacuees to closest exit. They had some choice to move about and stood in the way of evacuees to ensure the closest exit was used. In the second drill aisle staff was not present. In the third drill only silent exit staff at Exit 1F were present. Otherwise, evacuees were not directed. The one-stage announcer also referred from informing about uncrowded exits which he had done during the previous drills.

The number of people participating in each drill and their abilities are presented in Table 17. The people from the disability care centres allowed the other occupants to go first before evacuating, assisted by their own care centre staff. They were seated in the front part of the rear side blocks with. Wheelchair users were seated in the wheelchair area in front of the rear side block close to exit 1E. Wheelchair users evacuated with the help of their attendants. In the first drill the wheelchair users evacuated at the same time as the other attendants. In the second drill they evacuated after all others.

Table 17 – Number of participants per drill and their abilities.

	1st drill	2nd drill	3rd drill
Able-bodied	349	449	456
Wheelchair users	2	2	0
From disability care centres	47	89	20
Total	398	540	476

The simulation tests were run using the data from drill 3 since this drill had the least flow control and no wheelchair users reducing the need to model a third population. Agents were assigned to the seats of the theatre in accordance with reported seating position. A few people did not participate in the evacuation drills and stayed in their seats. These have not been modelled. The agents were assigned to a destination in accordance with chosen exits reported for each participant in the report.

The drill was conducted in Tokyo, Japan. The average age in the first drill was reported to be 59.4 years, 58.6 years in the second and 60.8 years in the third drill. Age and gender of participants were collected using questionnaires and presented in the report by Imanishi and Sano (2018). Distribution of age and gender per drill is presented in Figure 38. The data points were extracted using WebPlotDigitizer from the original report. The recovery rate for the questionnaires were 67.6% in the first drill, 67.3% in the second drill and 63.4% in the third drill. The percentages in Figure 38 are in relation to total number of participants per drill. The distributions do therefore not add up to 100% per drill.

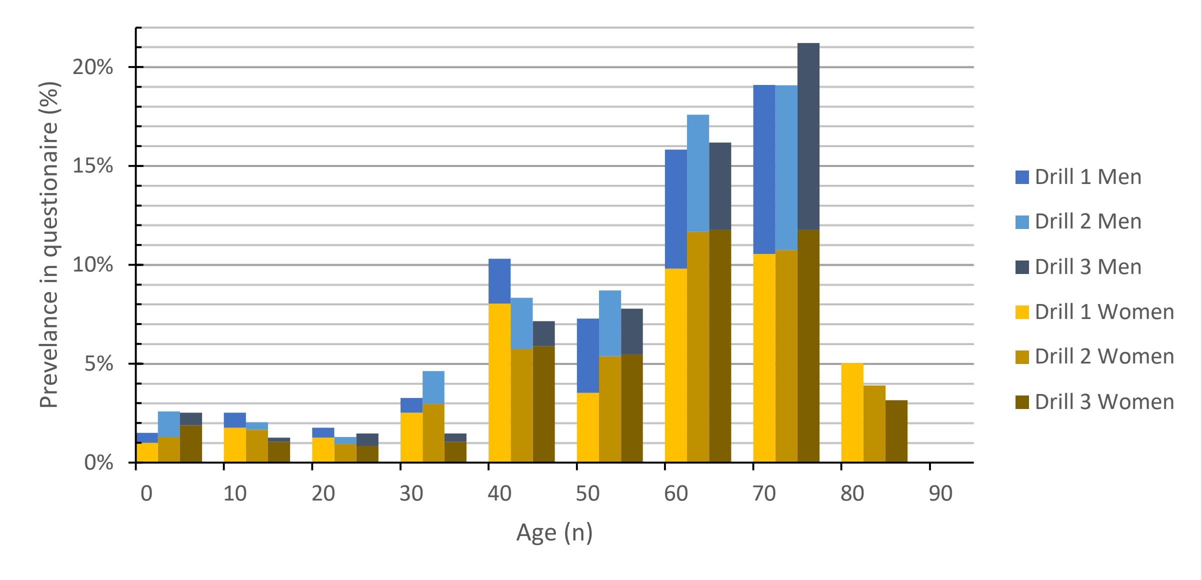

Figure 38 – Age distribution in percentage for each drill based on questionnaires.

No unimpeded speeds were presented for the populations and therefore other data was used for calibrating the unimpeded agent velocities. For the general population the default velocity distribution was used since all ages were present even though ages 40 to 80 are overrepresented. To assess the results dependency of mean velocity the tests were also run with an altered velocity distribution in which the mean velocity was set to 1 m/s. For the agents representing participants from the disability care centres a normal velocity distribution with data from Boyce et. al. (1999). Data used for agents' velocity distributions are presented in Table 18.

Table 18 – Data used for normal velocity distributions for each population.

	Mean (m/s)	Standard deviation (m/s)	Minimum (m/s)	Maximum (m/s)
Able-bodied (Default velocity)	1.34	0.26	0.46	1.61
Able-bodied (Altered velocity)	1	0.26	0.46	1.61
From disability care centres	0.78	0.34	0.21	1.4

Participation in the drills was a requisite for attending the event and participants in the evacuation drills were prepared for the evacuation cue. Imanishi and Sano (2018) reported that no pre-travel time was observed in the experiment. No pre-movement time was therefore set for the general population in the simulations. Since the participants from disability care centres were reported letting other occupants go before them a pre-movement distribution was set for them. It was calibrated for each drill to let the general population clear the middle aisles. In the third drill scenario the pre-movement time was set to 40 s.

Mean shoulder breadth of 449 mm for males and 402 mm for females were collected from a Japanese database (Research Institute of Human Engineering for Quality Life, 1994). According to Lin et al. (2004) the Japanese anthropometric database includes 178 anthropometric items from more than 34 000 people with age ranging from 7 to 90 years old. Agents' torso diameter was set to a uniform distribution with a value range of 0.40 m to 0.45 m.

Convergence of total evacuation time was assessed for each configuration. For drill 3 using the default velocity distribution for the able-bodied population convergence criteria were met at run 51. For drill 3 using the altered velocity distribution for the able-bodied population convergence criteria were met at run 45. All results were averaged over the number of runs needed to meet convergence criteria.

2.10.2 Results and comparison

Results and comparison to the reference experiment are sectioned by cases. Case 1 is simulations where the able-bodied population are configured with the default velocity distribution. Case 2 is simulations where the able-bodied population are configured with the altered velocity distribution.

2.10.2.1 Case 1

In Figure 39 the simulated evacuation time curves of each exit using the default velocity distribution for the able-bodied population is overlayed to the measured evacuation time curves from drill 3. Simulated evacuation times were faster than real-world results by about 10-20 s and all curves were steeper during steady flow for the simulation than the experiment indicating higher flow values.

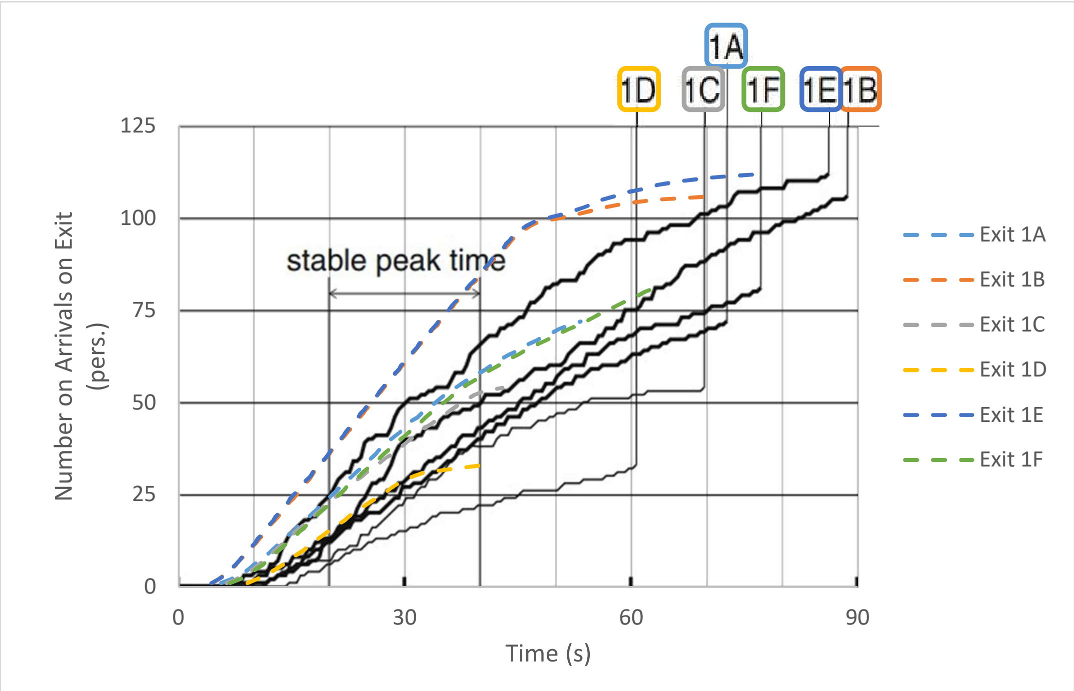

Figure 39 – Evacuation time curve for each exit for simulation using default velocity for the able-bodied population overlayed experiment results shown as black lines.

In Table 19 total evacuation times per exit are compared for experiment and simulation using the default velocity distribution for the able-bodied population. TET for simulation is 10-39% faster. Averaged over all exits and without the back exits the simulation is 25% and 19% faster.

Table 19 – Comparison of total evacuation times per exit for experiment and simulation using the default velocity distribution for the able-bodied population.

Exit	Experiment total evacuation time (s)	Simulation total evacuation time (s)	Difference (s)	Difference (%)
1E	86	77	9	10.4%
1B	89	71	18	20.2%
1F	77	63	14	18.2%
1A	73	53	20	27.0%
1C	70	43	27	38.7%
1D	61	40	21	33.7%
Average				
1A/1B/1E/1F			15	19.0%
All			18	24.7%

In Table 20 simulation and experiment specific flows for both total evacuation and peak flow are presented. The default velocity distribution was used for the able-bodied population for this data set. The specific flow was calculated at each exit by dividing the number of agents which passed in the time measured by the width of the exit and the measurement time. Peak specific flow is for experiment data collected in the stable peak time (20-40 s) noted in Figure 39. Simulation peak specific flow is measured between 15 and 30 s. Simulation flow was higher at all exits. The difference is simulation results compared to experiment results.

Table 20 – Specific flow for drill 3 from simulation using default velocity for the able-bodied population and experiment for both total evacuation and peak flow.

Exit	Total specific flow (pers./m/s)		Difference (%)	Peak specific flow (pers./m/s)		Difference (%)
	Simulation	Experiment		Simulation	Experiment	
1A	0.82	0.60	37%	1.17	0.85	37%
1B	0.90	0.73	24%	1.51	1.09	39%
1C	0.76	0.47	62%	1.01	0.94	7%
1D	0.49	0.33	50%	0.86	0.48	78%
1E	0.88	0.79	12%	1.50	1.24	21%
1F	0.78	0.64	22%	1.14	0.91	25%
Average						
1A/1B/1E/1F	0.85	0.69	23%	1.33	1.02	30%
All	0.77	0.59	30%	1.20	0.92	30%

In Figure 40 a screenshot from a simulation using the default velocity distribution for the general population is presented. Agents are coloured by their destination. Crowding can be seen at the exits in the middle, 1B and 1E. No crowding can be seen at the back exits, 1C and 1D. In the front exits agents crowd in the seating rows and along the side wall rather than in front of the exits. The stairs in front of the exits seem to limit the flow as well as width of the side aisles rather than the width of the exits. The front rows empty into the stairs before the agents in the side aisles get around the corner onto the stairs. This behaviour can be seen as representative of all runs.

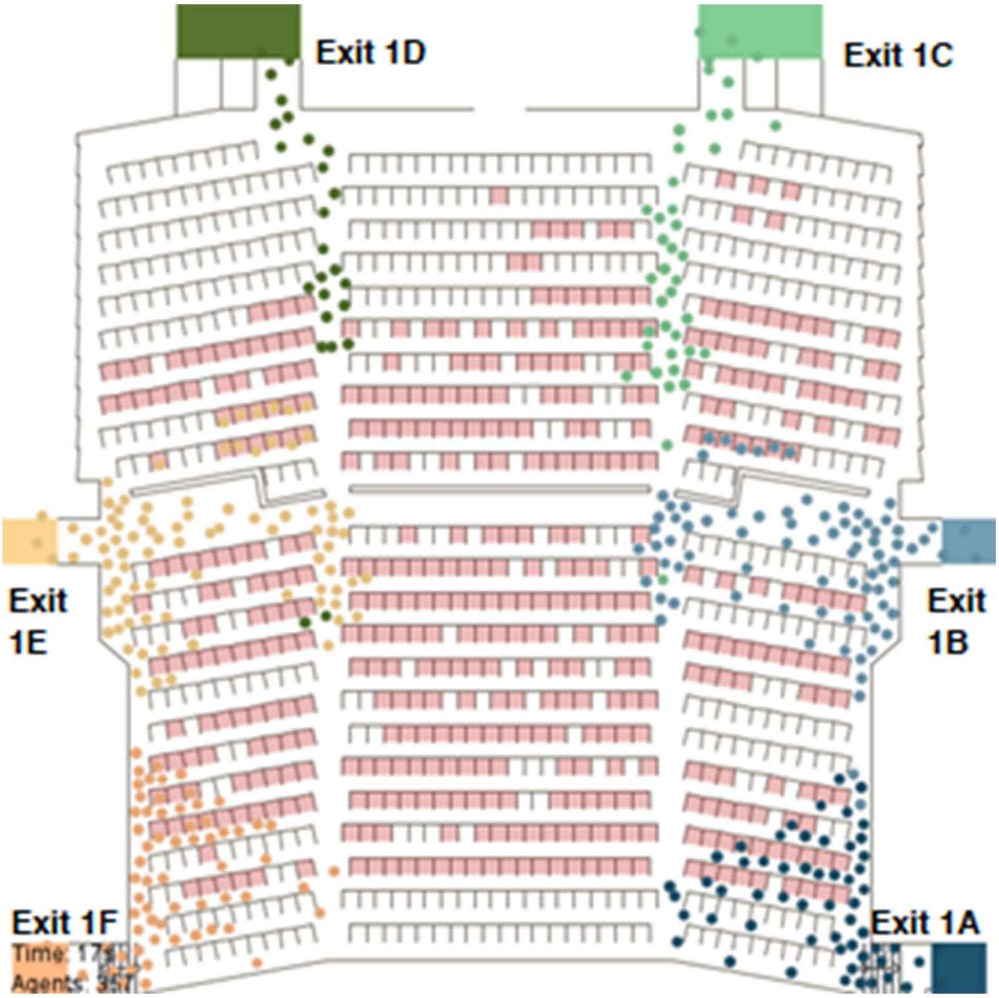

Figure 40 – Screenshot from 17 seconds into a simulation using the default velocity distribution for the general population. Agents are coloured by their destination.

2.10.2.2 Case 2

In Figure 41 the simulated evacuation time curves of each exit using the altered velocity distribution for the able-bodied population is overlayed the measured evacuation time curves from drill 3. Simulation evacuation time curves became more similar to experimental curves when mean velocity was lowered. For Exit 1F the two curves overlap very well. Total evacuation times for each exit are generally within 10 seconds of the experiment times. This indicated that the results are highly dependent on agent velocity.

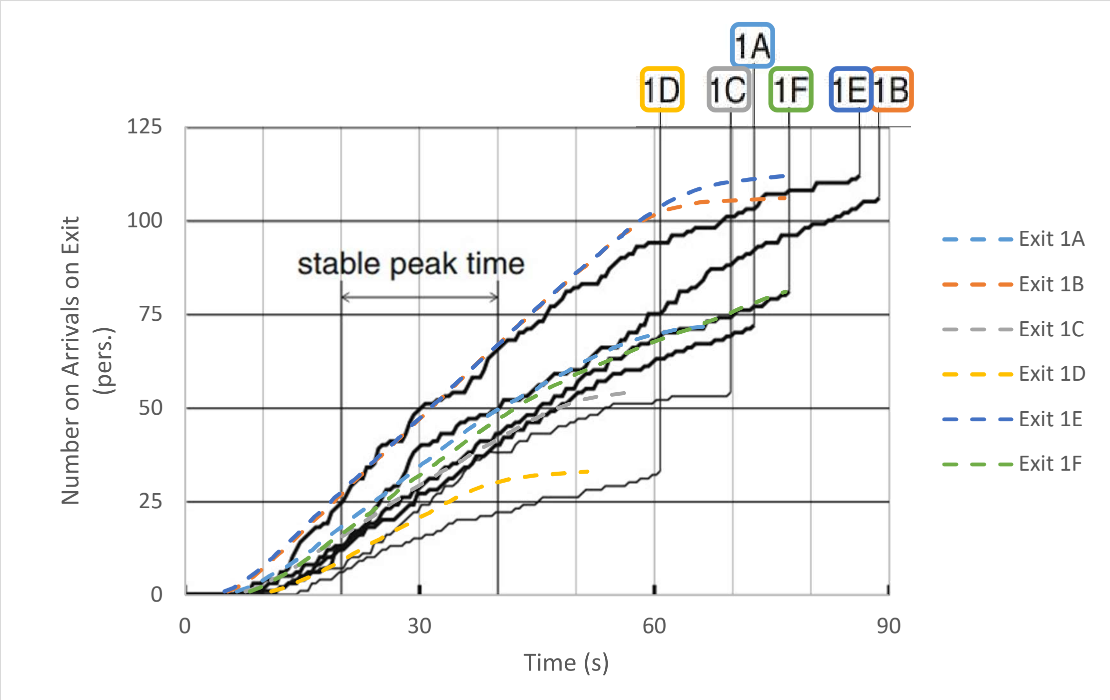

Figure 41 – Evacuation time curve for each exit for simulation using altered velocity for the able-bodied population overlayed experiment results shown as black lines.

In Table 21 total evacuation times per exit are compared for experiment and simulation using the altered velocity distribution for the able-bodied population. TET for simulation is 0-20% longer. Averaged over all exits and without the back exits the simulation is 11% and 8% longer.

Table 21 – Comparison of total evacuation times per exit for experiment and simulation using the altered velocity distribution for the able-bodied population.

Exit	Experiment total evacuation time (s)	Simulation total evacuation time (s)	Difference (s)	Difference (%)
1E	86	77	9	10.4%
1B	89	77	12	13.9%
1F	77	77	0	0.3%
1A	73	68	5	7.2%
1C	70	56	14	19.6%
1D	61	51	10	15.7%
Average				
1A/1B/1E/1F			7	7.9%
All			8	11.2%

In Table 22 simulation and experiment specific flows for both total evacuation and peak flow are presented. The altered velocity distribution was used for the able-bodied population for this data set. Peak specific flow is for both simulation and experiment data collected in the stable peak time (20-40 s) noted in Figure 41. The difference between simulation and experiment flow decreased when a lower mean velocity was used. Simulation flow was still generally higher at most exits than experiment flow but during peak flow the simulation flow was lower than experiment flow at exit 1C and 1E. Overall was the simulation flow 12% higher than experiment flow and during the stable peak time simulation flow was only 4% higher than experiment flow.

Table 22 – Specific flow for drill 3 from simulation using altered velocity for the able-bodied population and experiment for both total evacuation and peak flow.

Exit	Total specific flow (pers./m/s)		Difference (%)	Peak specific flow (pers./m/s)		Difference (%)
	Simulation	Experiment		Simulation	Experiment	
1A	0.64	0.60	7%	0.95	0.85	12%
1B	0.84	0.73	15%	1.21	1.09	11%
1C	0.58	0.47	24%	0.79	0.94	-16%
1D	0.39	0.33	18%	0.64	0.48	33%
1E	0.88	0.79	11%	1.19	1.24	-4%
1F	0.64	0.64	0%	0.93	0.91	3%
Average						
1A/1B/1E/1F	0.75	0.69	9%	1.07	1.02	5%
All	0.66	0.59	12%	0.95	0.92	4%

Similar flows can be achieved with calibrated velocity distributions, but no certain conclusions can be drawn from this data set due to the uncertainty introduced in the assumptions both regarding the geometry and the agent velocities.

2.11 Global validation test 2 – School

The test is replicating a full evacuation of a primary and secondary school in Spain. Test design and experimental data for comparison were taken from experiment conducted by Cuesta and Gwynne (2016). Evacuation times as well as stair velocities are compared to data from an observed unannounced evacuation drill.

2.11.1 Simulation configuration

The school consists of two buildings. Architectural diagrams were acquired from Cuesta and Gwynne (2016) and adapted for use in crowd:it. The floor geometry was represented in accordance with Figure 42**.** The small building is a two-storey structure with a single point of egress (Exit E). The main building is a four-storey structure with three exit points. One exit was only used by one class and did not interact with the other evacuees. They were not observed and therefore excluded from the study. Only width was specifically given for the stair geometry. Thread depth was set 0.3 m for all stairs in accordance with the architectural diagrams.

Figure 42 – Geometry of the test design in accordance with experiment (Cuesta and Gwynne, 2016).

In Figure 43 tripwire placement is presented. For Stair 1 the velocity was measured at the lower part of the staircase leading into the ground floor lobby. For Stair 2 the velocity was measured at the lower part of the staircase in the basement.

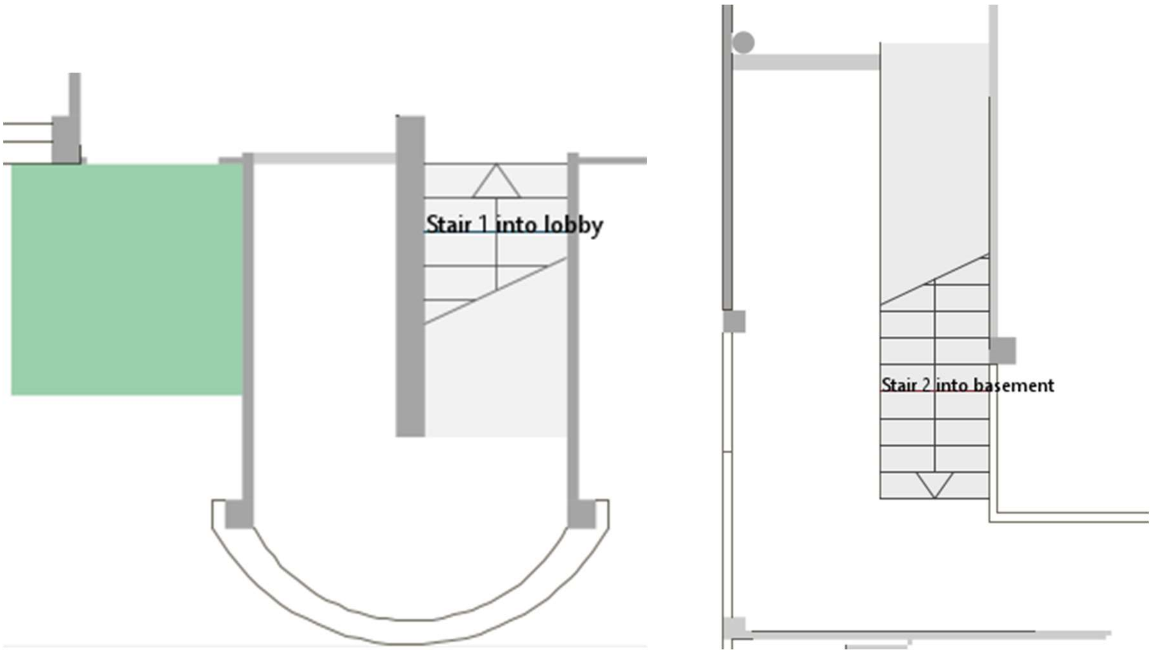

Figure 43 – Screenshots of tripwire placement used to measure horizontal stair velocities in simulations.

The trials were part of the routine evacuation drills conducted each year at Altamira School in Camargo Spain. Five trials were conducted. The drills were unannounced except for the first drill where staff knew the day of the drill but not the time. Precise conditions differed between trials, but geometry and the general evacuation plan was consistent.

The classes evacuated sequentially to reduce congestion. Each classroom was to evacuate according to a fixed evacuation plan which directed the classrooms to the nearest stair and exit. The lower floors were to evacuate first with no prioritization for classes on the same floor. Teachers independently phased the movement of the classes on the same floor to reduce congestion. Trial E3 was conducted on the eve of Christmas holidays and the classes were involved in Christmas activities as opposed to routine lessons during the other trials. This had a great impact on pre-travel times as well as resulted in a breach of the evacuation route strategy for two classrooms. 34 children from classrooms C8 and C9 along with a teacher evacuated through Exit D instead of Exit F.

For the simulation a trial specific approach was used. Population distribution, pre-travel times and route use was replicated from the selected trial according to the measured data made available by the authors.

Horizontal movement was measured in two locations during the trials. In the lobby area of the ground floor of the main building the travel speeds of students from classroom C4 and C5 (age 6-8) along with teachers were observed walking towards Exit D. On the second floor of the main building students from classroom C10–13 (age 12–16) along with teachers were observed moving towards Stair 1. Mean, standard deviation, minimum and maximum horizontal velocity measured in the trials by Cuesta and Gwynne (2016) divided by age are presented in Table 23. The horizontal velocity of the pre-school children in the small building was not measured. The small building was therefore excluded from the simulations. Agent velocities was set to normal velocity distributions with data from Table 23. The populations of classrooms C4–9 were set to normal distributions based on the data of the primary school students. The populations of classrooms C10–13 were set to normal distributions based on data of the secondary school students. Teachers and other personnel were set to a normal velocity distribution based on the data of the adults.

Table 23 –Velocity distributions based on measured horizontal velocity divided by age along with sample size.

	Primary	Secondary	Adult
Mean (m/s)	1.58	1.21	1.47
SD (m/s)	0.44	0.28	0.40
Min (m/s)	0.98	0.70	0.88
Max (m/s)	3.13	2.15	2.32
N	159	214	20

In Figure 44, Figure 45 and Figure 46 the velocity data sets are presented in histograms along with fitted distributions. The velocities are presented in m/s and the percentages are in relation to the whole data sets. No distribution is a perfect fit and since only uniform and normal distributions are currently available in crowd:it the normal distributions were chosen. A lognormal distribution would better represent the primary school data set while for the other data sets the best fit is not as clear.

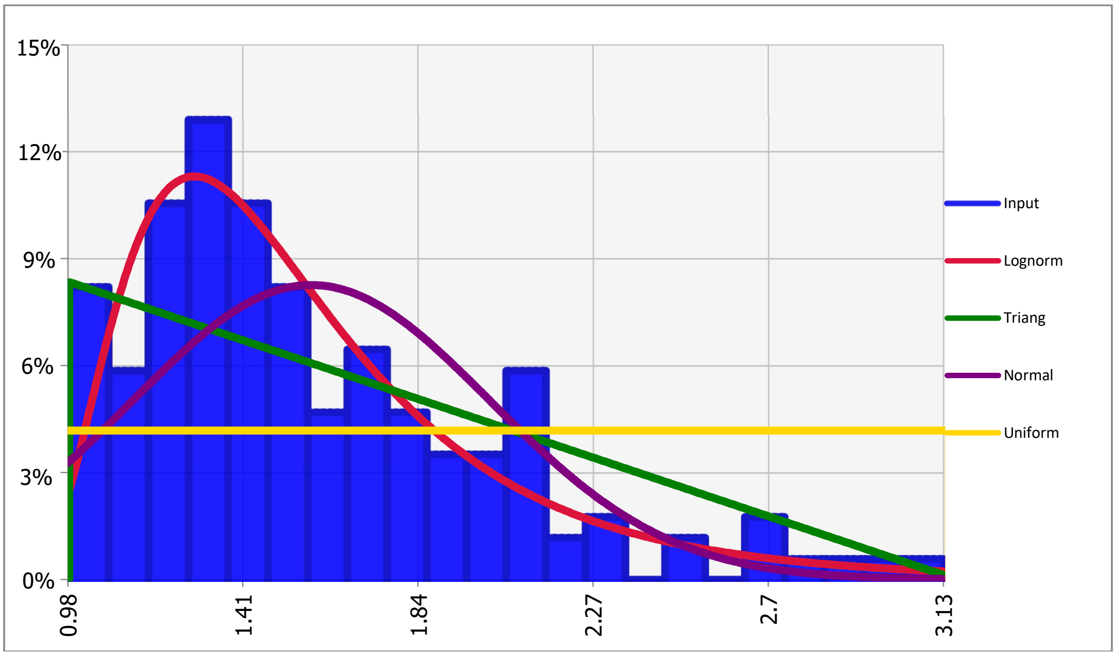

Figure 44 – Histogram of horizontal velocity measured in primary school students along with fitted distributions.

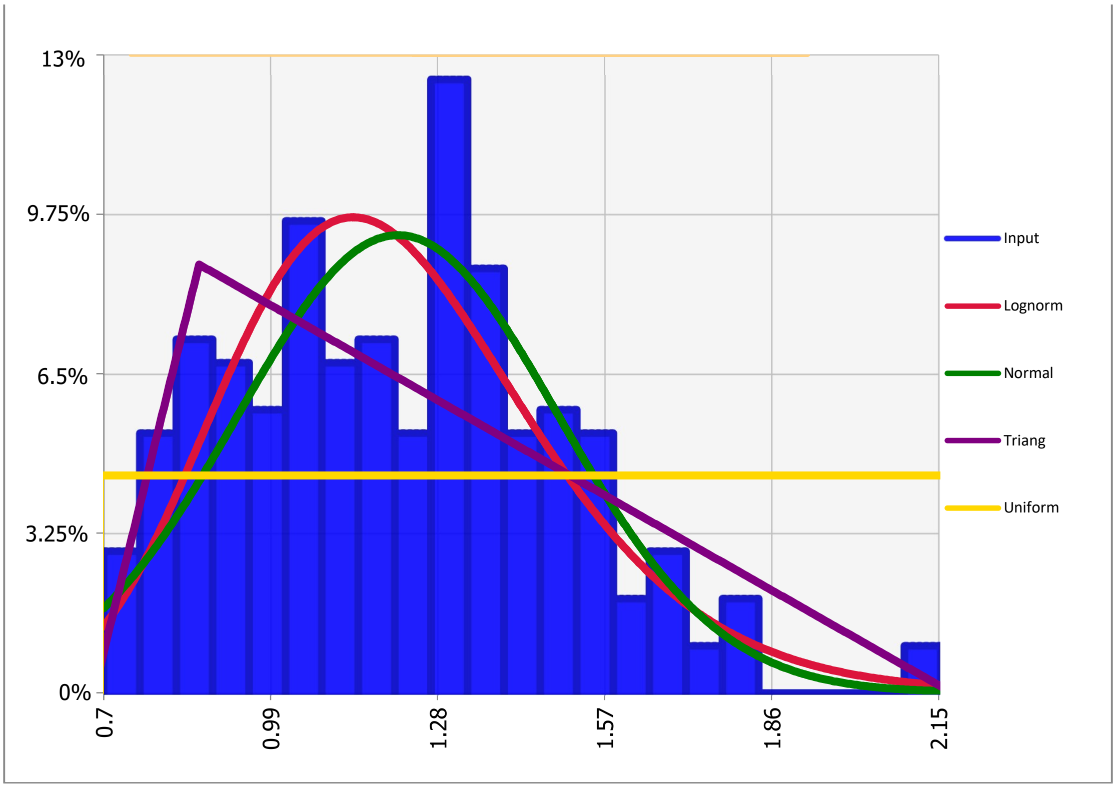

Figure 45 – Histogram of horizontal velocity measured in secondary school students along with fitted distributions.

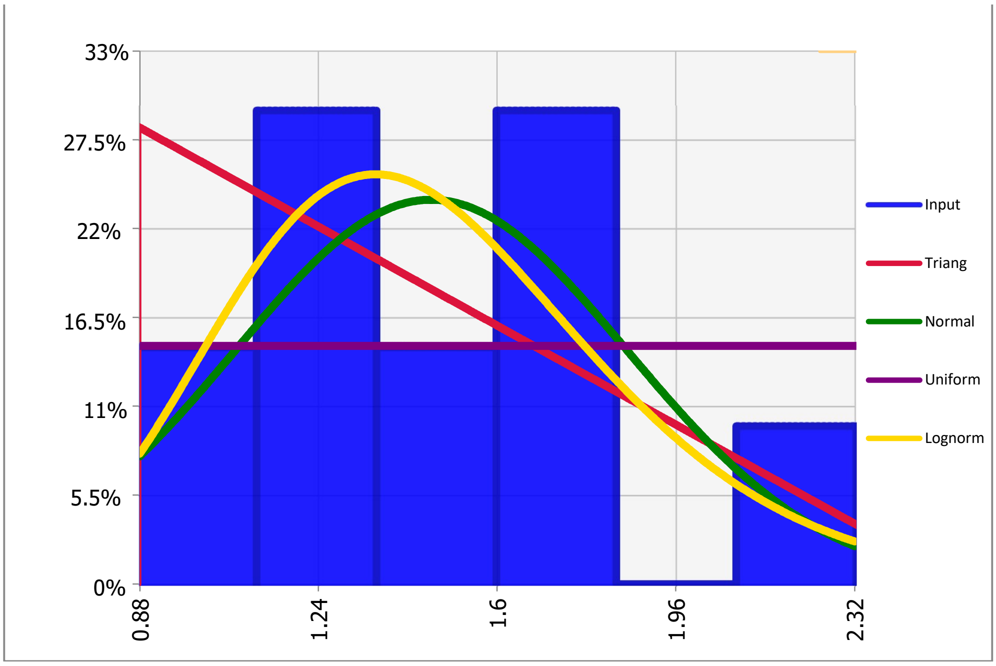

Figure 46 – Histogram of horizontal velocity measured in adults along with fitted distributions.

The torso diameter of each population of each classroom is represented as uniform distributions with values presented in Table 24. The diameters are approximations based on mean waist and upper arm circumferences by age and sex measured in American youths published by Fryar et al. (2021). For ease of calculation the torso diameter was calculated with the assumption that the circumferences were of perfect circles. The mean waist diameter and the mean arm diameter times 2 were added up to the torso diameter for each age and sex. Since the school classes are a mix of ages a span of the lowest and highest mean value in the ages and gender of each class were used. Approximated torso diameters start at 30 cm for age 6 and increase up to age 14 where values were very similar to the default values in crowd:it. The default values were used for teachers and the older students since they were similar. The use of American data of other body measurement is due to the lack of accessible data from the relevant population. These approximations are crude but are deemed to be more accurate than setting the torso diameter of the whole population to default values or guessing torso diameters.

In anthropometric measurements of Portuguese workers, the mean bi-deltoid shoulder breadth was 406.9 mm for women and 463.3 mm for men (Filho et al., 2023). 343 subjects participating in the sample, 169 were male (49.3%), and 174 were female (50.7%). The average male age was 36.2 ± 9.9 years, and the average female age was 42.7 ± 11.0 years. 98.5% of the participants were Portuguese workers, and 1.5% were foreign workers. 40% of foreign workers consisted of Brazilians followed by Angolans (20.0%), Nepalese (20.0%), and Venezuelans (20.0%). This data corroborates the default torso diameter for adults in a population close to the one in the test. Anthropometric data for Portuguese adults were used since no data from Spain could be attained and the data was deemed applicable due to geographic proximity and genetic origin (Bycroft et al., 2019).

Table 24 – Range of torso diameter of each class as well as adults.

Class (age)	Torso diameter (cm)
C4(6–7)	30-32
C5(7–8)	32-34
C6(8–9)	33-35
C7(9–10)	35-37
C8(10–11)	37-39
C9(11–12)	39-41
C10(12–13)	39-43
C11(13–14)	42-44
C12(14-15) C13(15-16) Adults	42-46

Due to lack of local data on torso diameter the height and weight of children used for the calculation of torso diameter were compared to Spanish children between the ages of 6 to 10. The data for the American children are from the same study as the waist and arm diameters were used for the torso approximations by Fryar et al. (2021). The data for Spanish children are from a study of Csilla et al. (2023) comparing anthropometric parameters of children in six European countries. 20 832 children were included in the study (48.7% boys) from which 3422 were from Spain (51.2% boys). The data and comparisons are presented in Table 25. The differences in mean height and weight of the children are within ±5%. The comparison indicates that the US data seems to be not too dissimilar to be applicable to situation.

Table 25 – Comparison of height and weight of children of ages 6 to 10 from the United States and Spain.

Origin of data		Spain	US		Spain	US	
Sex	Age	Mean weight (kg)	Mean weight (kg)	Difference (%)	Mean height (cm)	Mean height (cm)	Difference (%)
Male	6	24.7	23.9	-4%	121.9	118.8	-3%
	7	27.6	27.7	0%	127.4	126.1	-1%
	8	31.0	31.6	2%	133.0	132.1	-1%
	9	34.0	35.4	4%	137.7	135.9	-1%
Female	6	24.6	23.7	-4%	121.2	118.8	-2%
	7	27.2	27.3	0%	126.1	124.1	-2%
	8	30.6	31.3	2%	131.7	129.8	-1%
	9	34.0	34.9	3%	136.9	136.9	0%

For the trials the pre-travel time of each classroom was measured. For trial E3 the pre-travel times are presented in Table 26. Cuesta and Gwynne (2016) divided the pre-travel time into preparation time and holding time. Preparation time is the time spent by children and teacher to form a queue in the classroom and be ready to start their evacuation movement. Holding time is the time that the class is held queuing at the door until the teacher deemed the evacuation path to be cleared. The class then left the classroom together. The pre-movement time of the agents in each classroom are set to the total measured pre-travel time in Table 26. The agents were spawned in a queuing behaviour at the door to mimic the measured behaviour rather than spread in the classrooms. During trial E3 there was one secretary in the office which participated in the drill. Since no pre-travel time was presented for the office in the data sets, they are omitted from the simulation as well as the experiment result.

Table 26 – Pre-travel time of each classroom split by preparation and hold time for trial E3.

	TRIAL 3		
Classroom	Prep time (s)	Hold time (s)	Total (s)
C4	43	14	57
C5	40	0	40
C6	10	27	37
C7	19	39	58
C8	31	42	73
C9	68	0	68
C10-11	59	55	114
C12	57	96	153
C13	30	136	166

Routes used in trial E3 are presented in Table 27. The secretary in the office is omitted from this data. For the simulations students in classrooms C8 and C9 were split equally between classrooms and the teacher was placed in classrooms C8 since all of classroom C8 went the wrong way and classroom C9 had no hold time. The students of classroom C9 that went to exit D are presumed to have been in the front since they evacuated without teacher guidance. People tend to use the exit they used to enter rather than exits seldom used as Exit F in the basement. 27 students were placed in classroom C8 with a teacher and 7 students were placed in classroom C9 and routed to exit D.

Table 27 – Routes used during trial E3.

Evacuation route	Pupils	Adults	Total
C-5-Exit D	18	1	19
C-4-Exit D	22	2	24
C8-9-Stair 2-Exit D	34	1	35
C-10-Stair 1-Exit D	29	1	30
C-11-Stair 1-Exit D	24	1	25
C-12-Stair 1-Exit D	21	1	22
C-13-Stair 1-Exit D	20	1	21
Total – Exit D	**168**	**8**	**176**
C-6-Stair 2.2-Exit F	25	2	27
C-7-Stair 2.2-Exit F	22	1	23
C-9-Stair 2.1-Stair 2.2-Exit F	20	0	20
Total – Exit F	**67**	**3**	**70**
Total – Exit D and Exit F	**235**	**11**	**246**

Convergence of total evacuation time was assessed. For trial E3 convergence criteria were met at run 32.

2.11.2 Results and comparison

In Figure 47 Total evacuation times for both experiment and simulations are presented. Simulation TET are averaged across 32 runs. Experiment TET is divided by origin classroom and exit used. Simulation evacuation times are overall similar to the evacuation drill. For the downstairs classroom TETs are very similar while the simulations lack somewhat behind for the upper classrooms. In total the simulation evacuation took 15 seconds or 7.6% longer than Trial E3. For use in fire engineering purposes an over-estimation of total evacuation times is far less hazardous than an under-estimation.

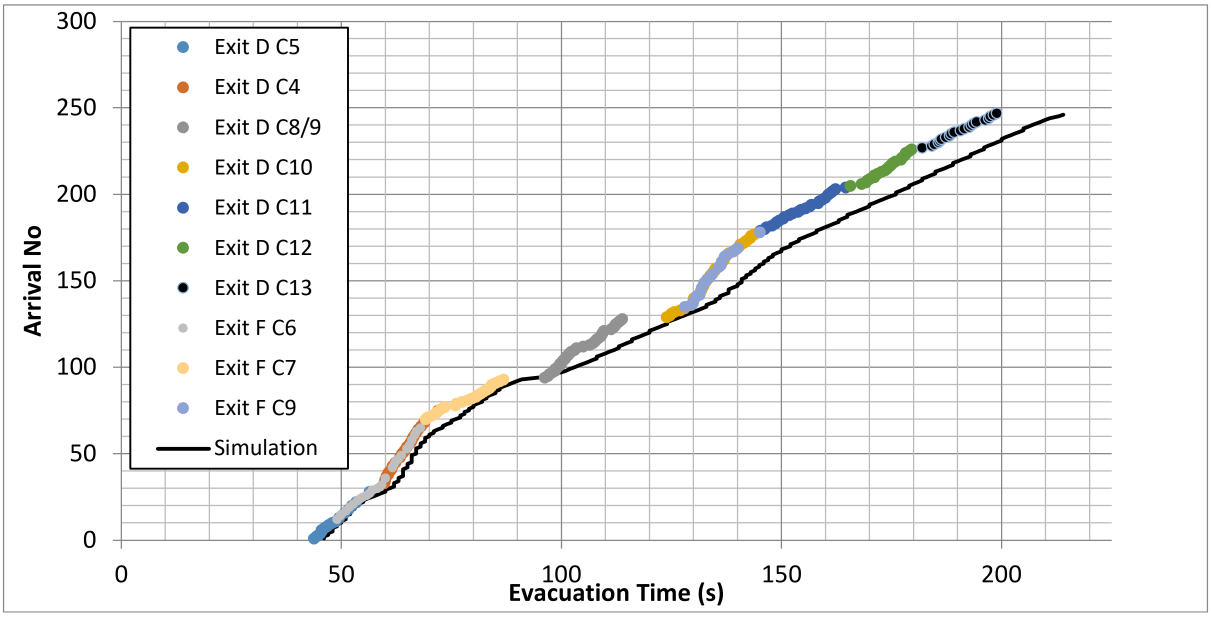

Figure 47 – Chart of averaged TET for simulation overlayed experiment TET divided by classroom and exit.

Kolmogorov-Smirnov tests (KS-test) were run both using averaged total evacuation times over 32 runs and all values across 32 runs. To assess the sensitivity of the KS-tests several significance levels were assessed. The hypothesis tested is that the simulation data set is of the same distribution as the experiment results. The hypothesis is rejected at a significance level of 0.05 when using all values over 32 runs. The hypothesis is accepted at a significance level of 0.05 when using averaged total evacuation time over 32 runs. This indicates that the simulation results can be considered to be from the same distribution but there is great variance between runs.

In the trials stair velocity was measured in both Stair 1 and Stair 2 with use of cameras. The velocity was measured along the incline of the stairs. Since crowd:it represents structures in the plane with no vertical movement the stair data need to be processed before comparison. No specific data was given for stair heights in the paper or architectural diagrams therefore a range of stair angles was assessed.

Horizontal stair velocities from the simulations were used together with trigonometry to give the simulation stair velocities for angles 30°, 35° and 40°. In Table 28 mean, standard deviation, minimum and maximum of horizontal velocities from Stair 1 and 2, calculated angled velocities and measured stair velocities in Trial E3 are presented. The stair velocities in the simulations are significantly lower than measured stair velocities. This could explain the longer evacuation times of the second-floor classrooms. The lower stair velocities of the simulation could in some part be due to differing geometry after the adaptation of the architectural diagrams for crowd:it. It could also be due to the Optimal Stair Model adopted for the simulation of the pedestrian movement on stairs and differing behaviours of school children and adults. The observed mean stair velocities for only adults in Trial E3 were slightly lower at 0.93 and 1.00 m/s.

Table 28 – Measured stair velocity in Trial E3 compared to simulation stair velocity.

		Horizontal stair velocity in simulation	Velocity at angle (°)			Stair velocity in Trial E3
			30	35	40	
Stair 1	Mean (m/s)	0.51	0.59	0.62	0.67	0.98
	SD (m/s)	0.06	-	-	-	0.18
	Min (m/s)	0.23	0.27	0.29	0.31	0.70
	Max (m/s)	0.82	0.94	1.00	1.06	1.79
Stair 2	Mean (m/s)	0.57	0.66	0.70	0.74	1.07
	SD (m/s)	0.07	-	-	-	0.15
	Min (m/s)	0.29	0.33	0.35	0.38	0.80
	Max (m/s)	0.87	1.01	1.06	1.14	1.46

3 Discussion

As mentioned in section 1.5, due to lack of acceptance criteria the validity of the models can only be given in the degree of deviation from the experimental datasets chosen for comparison. It is up to the users to determine if the uncertainty of a specific component in the model and in the model as a whole is acceptable for the specific intended uses.

The testing in this report cannot be considered exhaustive in relation to the scenarios which can arise in building fire evacuation. This report can only be used as an initial step towards validation for the specific scenarios which have been modelled. There are several components or parts of components which have not been tested such as the use of scaled areas, mentioned in section 1.3.4.2.3, for representing stairs or to explicit model lanes in counter-flow, measure flow in 90°- and 180°-degree corners, performance of escalators, and the queueing model for bottlenecks. This can also serve as a list of suggestions for additional testing.

Assessment of results in this report is done in comparison with observed evacuation scenarios. This hinges on the assumption of observed evacuation scenarios accurately describing behaviour in actual evacuations. The observed evacuation scenarios are in some instances unannounced or semi-announced evacuation drills in which measurement is done as discreetly as possible to not alert the experiment population and affect the results. Several of the observed evacuation scenarios used in this work are conducted in controlled environment where the evacuation is done at the command of the experimenters, often with instructions how to behave. One is not necessarily better than the other, but they come with their own set of uncertainties. Due to these uncertainties the experiment cannot be considered accurate description of reality or at least not all-encompassing. Real world behaviour in fire building evacuation scenarios varies with culture, age, gender, prior experiences, the behaviour of other people, the perceived threat level, etc. (Gwynne, 2010; Fahy & Proulx, 2001; Kuligowski, 2016). The ability of the models to produce results which may compare well to a specific experiment may not necessarily mean it will accurately describe reality well in other instances and vice versa. To assess the reliability of the results in this report the specific components tested in this report should also be validated against other observed evacuation scenarios.

In test 22 the model ability to represent pre-evacuation behaviour was tested. The crowd:it model represent pre-evacuation behaviour with deterministic values or pseudo-random sampling from distributions for a delay time before the agents start to move. How well the pre-evacuation behaviour is represented is therefore contingent on how well the pre-evacuation times of the population can be represented by the available distributions in crowd:it.

In this and a few other tests it became apparent that some of the underlying behavioural distributions could not be represented in crowd:it during testing. Only uniform and normal distributions were implemented. For pre-travel time the gamma, lognormal, loglogistic or Weibull distributions are commonly suggested for building fire evacuation simulation (Lovreglio et. al., 2019a). The implementation of more distributions and custom distributions is therefore suggested to improve the usability of the model for building fire evacuation purposes. Since the testing phase for this report was concluded and partly due to sharing the results of tests in this report the log-normal distribution has been implemented in crowd:it version 2.26. For fire engineering uses distributions are generally used to represent behaviour and characteristics of agents since the exact values are not commonly known However, for occasions when full data sets are available, such as validation work, the ability to use custom values would produce more reliable results. Unimpeded walking speeds for fire engineering purposes are generally given in means, range of means or standard deviation (Gwynne & Boyce, 2016).

In test 23 the model ability to represent the relationship between flow rate, density and walking speeds in uni-directional flow was tested. The crowd:it model employs the Optimal Steps Model for representation of pedestrian movement. The relationship between flow rate, density and walking speeds are implicitly modelled. In the simulated tests velocities decreased with rising densities similar to the compared experiment. However, the model tests using the velocity distribution which gave the most similar results to experiment showed 8-24% higher velocities compared to experiment velocities at the same densities. Model tests were performed with three different velocity distributions since no free-flow velocities were reported for the experiment population. Results depended on the chosen velocity distribution for the test population. Due to time constraints the dependence of body size as not examined. Additional testing with differing velocity distributions and body size could show results closer to compared experimental results.

In test 24 which assessed movement on a stairwell it became apparent that ascending and descending stair movement does not affect agents' velocity on the stairs when stairs are explicitly modelled using the current crowd:it implementation of the Optimal Stair model. For short stairs this may not be noticeable but in ascending evacuation scenarios with long stairs, e.g. metro station, the physical exertion is noticeable (Ronchi et. al., 2015). The effect might also be less apparent if the stairwell is full, and velocities are lowered due to higher densities. More testing is needed to confirm such a statement. Test 25 and global test 2 both features descending stair movement. Test 25 gave mean flows in the lower part of recorded flows in the experiment. In the experiment chosen for comparison flows varied depending on the stair riser, and therefore also stair angle, which is not considered in the Optimal Stair Model. In the global test, evacuation times seem to lag behind the experiment evacuation times especially for the classrooms using the stairs due to stair velocities being lower than experiment stair velocities. Modelled descending stairs can be expected to be slower than real life and ascending stairs long stairs can be expected to be faster than real life when the stairs are explicit modelled. Stairs can also be modelled implicitly with the use of scaled areas as is the convention in evacuation models. Fatigue can by this method be modelled with a higher deacceleration factor for the scaled areas which represent stairs further up. Tests were only performed using the implicit model for stairs.

In test 26 the ability of the evacuation model to represent evacuation movement around a corner was tested. The crowd:it model employs the Optimal Steps Model for representation of pedestrian movement. The movement around a corner is implicitly modelled. The movement around a corner can only be discussed qualitatively, due to limitations in the available experimental data-sets, and limitations of the heat map functionality in crowd:it. In the modelled test agents tended to walk close to the inside corner and the outside of the corner was not used similarly to the experiment. However, in the experiment lane formation was observed which was not evident in the modelled tests. Agents also tended to zig-zag across the corridor rather than follow the person in front as in the experiment. There was no evidence that the selfish behaviour of the OSM resulted in a different flow through the corner.

In test 27, which assessed the model ability to represent counter-flow, agents tended to get stuck trying to pass each other. This is a common problem in continuous models (Ronchi & Nilsson, 2016). No lane formation was observed in the simulation in contrast with experiments. This is expected since no interactions between agents are modelled in the Optimal Steps Model other than the repulsing potentials around other agents which affect the navigational fields. These repulsing potentials upholds the agent's personal space when possible but makes no effort to conform to social rules or behaviours such as strict lane formation or tendencies to follow the evacuation. For counter-flow conditions, evacuation times will be skewed by agents getting stuck. Counter-flow is important in fire scenarios since firefighters may need to get into the building while people are evacuating. For areas where sever counter-flow can be expected, lanes can explicitly be modelled with scaled areas used only for navigation added to agents' paths. It is suggested that further testing evaluate the performance of this solution.

In test 28 the ability of evacuation models to represent route/exit choice was tested. Exit choice can in crowd:it only be explicitly modelled with a deterministic approach. Tests therefore showed identical results to the experiment chosen for comparison.

In test 29 the ability of evacuation models to represent bottlenecks at openings was tested. The behaviour at openings is in crowd:it is emergent due to the way the Optimal Steps Model is implemented. Mean flow and density from modelled tests show a difference of 1% and 5% for flow and density respectively compared to experiment scenario with the closest density. In the experiment chosen for comparison several scenarios were tested where the population were given different instructions on how hurried or selfish to behave. The experiment results showed densities up to nearly 9 people/m². These densities cannot usually be achieved in due to the use of rigid bodies. Since agents' body size is represented as circles in this model and the body diameter was set to almost 0.5 m a higher density than the recorded mean density of 3.9 m^{-2} in the simulation cannot be expected. The limitation in representing extremely high-density conditions is a known problem in several evacuation models (Pelechano & Malkawi, 2008; Cilenti, 2019).

Test 30 was not performed since the impact of smoke on evacuees cannot be modelled explicitly in crowd:it.

In global validation test 1 a full evacuation of an auditorium in Japan was modelled. Evacuation times and flow through exits were compared. Since no free velocity data for the experiment population was presented, simulations were run with the default velocity distribution for the able-bodied population and a velocity distribution with a lower mean velocity to evaluate the results dependence on agent velocity. The geometry of the simulations was also simplified from the experiment geometry. Most notably the floor was represented as flat while in reality it was stepped in 1,2 meters treads. Due to lack of data of the velocity effects of the stepped floor it was modelled as flat. For the first case with the default velocity distribution for the able-bodied population the mean TET from the simulations were 19% shorter than mean experiment TET. The mean peak flow through the exits in the simulations were 30% higher than experiment equivalents. For the second case with the slower velocity distribution for the able-bodied population mean TET from the simulations were 8% shorter than mean experiment TET. The mean peak flow through the exits in the simulations were 5% higher. Similar flows and evacuation times can be achieved with calibrated velocity distributions, but due to the uncertainty introduced in the assumptions both regarding the geometry and the agent velocities certain conclusions cannot be drawn from this data set.

In global validation test 2 a full unannounced evacuation of a primary and secondary school in Spain was modelled. Evacuation times as well as stair velocities from the simulations were compared to data from the observed unannounced evacuation drill. Geometry, pre-evacuation times and velocities was given for the building and the population of the observed drill. In total the modelled evacuation took 7.6% longer than the observed evacuation trial which was represented. The stair velocities in the simulations were significantly lower than measured stair velocities. This could explain the longer evacuation times of the second-floor classrooms. The lower stair velocities of the simulation could in some part be due to differing geometry after the adaptation of the architectural diagrams for crowd:it. It could also be due to the Optimal Stair Model adopted for crowd:it for pedestrian movement on stairs and differing behaviours of school children and adults.

4 Conclusion

Validation testing shows that crowd:it, and its implementation of The Optimal Steps Model, is generally able to reasonably represent pedestrian movement if the input is carefully calibrated. This is shown in the results for test 22, 23, 29 and the two global tests which are not too dissimilar to the experiment they model. It is worth noting that the testing in this report cannot be considered exhaustive and there are several components or parts of components which have not been tested. The assessment of validity of the performance of the model is done in comparison with observed evacuation scenarios. Due to these uncertainties the experiment cannot be considered full description of reality or at least not all-encompassing. The reliability of the results in this report and the validity of the model would benefit from validation testing of the same components against other observed evacuation scenarios.

There were a few tests where behaviours differed to a greater extent and users should take under special consideration if used. In counterflow scenarios, agents tended to get stuck, and agents tend to zig-zag between lanes in a flow rather than following the person in front as experiments tend to show. Ascending or descending a stairwell does not alter the movement speed of agents when stairs are implicitly modelled using the Optimal Stair Model. This can accumulate large time differences in long stairwells. A distinguishing between ascending and descending behaviour in the Optimal Stair Model might give more realistic results and make it more reliable compared to the alternative of implicitly model stairs using scaled areas.

The selection of distributions available for representing evacuees' characteristics and behaviour in crowd:it was in some instances limiting the scenario calibration. Therefore, implementation of additional distributions and custom datasets is recommended in addition to the log-normal distribution already implemented in response to the testing in this report.

Suggested future research in the field could focus on the development of benchmark acceptance criteria for validation testing, validation of the model components tested in this work against different experimental datasets, or validation of model components or parts of components which have not been tested in this work.

5 References

accu:rate GmbH. (2023). crowd:it Documentation. Retrieved from https://www.accu-rate.de/wp-content/uploads/2023/12/crowdit_Documentation_en-2.25.0-1.pdf

ARUP. (2015). The Verification and Validation of MassMotion for Evacuation Modelling. Retrieved from https://www.oasys-software.com/wp-content/uploads/2017/11/The-Verification-and-Validation-of-MassMotion-for-Evacuation-Modelling-Report.pdf

Berrou, J.L., Beecham, J., Quaglia, P., Kagarlis, M.A., Gerodimos, A. (2007). Calibration and validation of the Legion simulation model using empirical data. In Waldau, N., Gattermann, P., Knoflacher, H., Schreckenberg, M. (eds) *Pedestrian and Evacuation Dynamics*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-47064-9_15

Blomstrand Martén, J., Henningsson, J. (2014). Verification and Validation of Viswalk for Building Evacuation Modelling. Department of Fire Safety Engineering Lund University, Sweden.

Boyce, K.E., Shields, T.J. & Silcock, G.W.H. (1999) Toward the Characterization of Building Occupancies for Fire Safety Engineering: Capabilities of Disabled People Moving Horizontally and on an Incline. *Fire Technology 35*, 51–67 (1999). https://doi.org/10.1023/A:1015339216366

Burghardt, S. (2013). Dynamik von Personenströmen in Sportstadien. Forschungszentrum Jülich.

Burghardt, S., Seyfried, A., & Klingsch, W. (2013). Performance of stairs – Fundamental diagram and topographical measurements. *Transportation Research Part C-emerging Technologies, 37*, 268–278. https://doi.org/10.1016/j.trc.2013.05.002

Bycroft, C., Fernández–Rozadilla, C., Ruíz-Ponte, C., Quintela, I., Carracedo, Á., Donnelly, P., & Myers, S. (2019). Patterns of genetic differentiation and the footprints of historical migrations in the Iberian Peninsula. *Nature Communications, 10*(1). https://doi.org/10.1038/s41467-018-08272-w

Canter, D., Breaux, J., & Sime, J. (1980). Domestic, Multiple Occupancy And Hospital Fires. In D. Canter, *Fires and Human Behavior*, 117– 136. John Wiley & Sons Ltd.

Cao, S., Zhang, J., Salden, D., & Ma, J. (2016). Fundamental Diagrams of Single-File Pedestrian Flow for Different Age Groups. In Lecture Notes in Computer Science. Springer Science+Business Media. https://doi.org/10.1007/978-3-319-41000-5_15

Choi, J., Galea, E. R., & Hong, W. (2013). Individual Stair Ascent and Descent Walk Speeds Measured in a Korean High-Rise Building. *Fire Technology, 50*(2), 267–295. https://doi.org/10.1007/s10694-013-0371-4

Chooramun, N., Lawrence, P. J., & Galea, E. R. (2017). Evacuation simulation using Hybrid Space Discretisation and Application to Large Underground Rail Tunnel Station*. Physical Sciences Reviews, 2*(9). https://doi.org/10.1515/psr-2017-0001

Cilenti, I. (2019). Crowd evacuation in high-density scenarios. LUTVDG/TVBB.

Csilla, S., Szöllösi, G. J., Ilyés, I., Cardon, G., Latomme, J., Iotova, V., Bazdarska, Y., Lindström, J., Wikström, K., Herrmann, S. M., Schwarz, P. E. H., Karaglani, E., Manios, Y., Makrilakis, K., Aznar, L. a. M., González-Gil, E. M., & Rurik, I. (2023). Differences in anthropometric parameters of children in six European countries. *Children (Basel), 10*(6), 983. https://doi.org/10.3390/children10060983

Cuesta, A., & Gwynne, S. (2016). The collection and compilation of school evacuation data for model use. *Safety Science, 84*, 24–36. https://doi.org/10.1016/j.ssci.2015.11.003

Dietrich, F., Köster, G., Seitz, M., & Von Sivers, I. (2014). Bridging the gap: From cellular automata to differential equation models for pedestrian dynamics. *Journal of Computational Science, 5*(5), 841–846. https://doi.org/10.1016/j.jocs.2014.06.005

Fahy, R. & Proulx, G. (2001). Toward creating a database on delay times to start evacuation and walking speeds for use in evacuation modeling. In 2nd International Symposium on Human Behaviour in Fire, Boston, MA., U.S.A., pp. 175-183

Filho, P. C. M., Da Silva, L., De Mattos, D. L., Pombeiro, A., Castellucci, I., Colim, A., Carneiro, P., & Arezes, P. (2023). Establishing an anthropometric database: A case for the Portuguese working population. *International Journal of Industrial Ergonomics, 97*, 103473. https://doi.org/10.1016/j.ergon.2023.103473

Fridolf, K., Nilsson, D., & Frantzich, H. (2011). Fire evacuation in underground transportation systems: A review of accidents and Empirical research. *Fire Technology, 49*(2), 451–475. https://doi.org/10.1007/s10694-011-0217-x

Fryar, C. D., Carroll, M. D., Gu, Q., Afful, J. B. A., & Ogden, C. L. (2021). Anthropometric Reference Data for Children and Adults: United States, 2015-2018. National Center for Health Statistics. *Vital Health Stats, 3*(46), 1–44. https://pubmed.ncbi.nlm.nih.gov/33541517/

Galea et. Al. (2017). BuildingEXODUS v 6.3 Theory manual. Fire Safety Engineering Group, University of Greenwich. London.

Gales, J., Ferri, J., Harun, G., Jeanneret, C., Young, T., Kinsey, M., Wong, W., Stock, J., Chen, L., Thompson, P., Frantzich, H., Arias, S. & Friholm, J. (2020) Anthropomorphic Data and Movement Speeds. SFPE Scientific and Educational Foundation, Inc.

Grandison, A. (2020). Determining confidence intervals, and convergence, for parameters in stochastic evacuation models. *Fire Technology, 56*(5), 2137–2177. https://doi.org/10.1007/s10694-020-00968-0

Grandison, A., Deere, S., Lawrence, P. J., & Galea, E. R. (2017). The use of confidence intervals to determine convergence of the total evacuation time for stochastic evacuation models. Ocean Engineering, 146, 234–245. https://doi.org/10.1016/j.oceaneng.2017.09.047

Gwynne, S. (2010). Conventions in the Collection and Use of Human Performance Data. NIST GCR 10-928. National Institute of Standards and Technology

Gwynne, S., & Boswell, D. (2009). Pre-evacuation Data Collected from a Mid-rise Evacuation Exercise. *Journal of Fire Protection Engineering*. https://doi.org/10.1177/1042391508095093

Gwynne, S., & Boyce, K. (2016). Engineering data. In: Hurley, M.J., et al. *SFPE Handbook of Fire Protection Engineering*. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-2565-0_64

Gwynne, S.M.V., Rosenbaum, E.R. (2016). Employing the Hydraulic Model in Assessing Emergency Movement. In: Hurley, M.J., et al. *SFPE Handbook of Fire Protection Engineering*. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-2565-0_59

Hall, E. T. (1966). The Hidden Dimension. Garden City, NY: Doubleday.

Helbing, D., & Molnár, P. (1995). Social force model for pedestrian dynamics. *Physical Review, 51*(5), 4282–4286. https://doi.org/10.1103/physreve.51.4282

Imanishi, M., & Sano, T. (2018). Route choice and flow rate in theatre Evacuation Drills: Analysis of Walking Trajectory Data-Set. *Fire Technology, 55*(2), 569–593. https://doi.org/10.1007/s10694-018-0783-2

International Maritime Organization. (2007). Guidelines for evacuation analysis for new and existing passenger ships. MSC. 1/Circ. 1238.

International Maritime Organization. (2016). Revised Guidelines on Evacuation Analysis for New and Existing Passenger Ships. MSC. 1/Circ. 1533

International Organization for Standardization. (2008). Fire Safety Engineering – Assessment, verification and validation of calculation methods (16730:2008).

International Organization for Standardization. (2020). Fire safety engineering – Verification and validation protocol for building fire evacuation models (20414:2020).

Kang, Kai. (2006). Application of NFPA 130 for emergency evacuation in a mass transit station. The Free Library. Retrieved from https://www.thefreelibrary.com/Application+of+NFPA+130+for+emergency+evacuation+in+a+mass+transit...-a0156720163

Ketchell, N. (2006). A Technical Summary Of The EGRESS Code. ESR Technology, Warrington. Retrieved from https://www.esrtechnology.com/images/egresspage/Egress-Technical-Summary.pdf

Kneidl, A. (2013). Methoden zur Abbildung menschlichen Navigationsverhaltens bei der Modellierung von Fußgängerströmen. Fakultät für Bauingenieur- und Vermessungswesen.

Kobes, M., Helsloot, I., De Vries, B., & Post, J. (2010). Building safety and human behaviour in fire: A literature review. *Fire Safety Journal, 45*(1), 1–11. https://doi.org/10.1016/j.firesaf.2009.08.005

Korhonen, T. (2018). Fire Dynamics Simulator with Evacuation: FDS+Evac, Technical Reference and User's Guide (FDS 6.6.0, Evac 2.5.2, DRAFT),VTT Technical Research Centre of Finland. Retrieved from http://virtual.vtt.fi/virtual/proj6/fdsevac/documents/FDS+EVAC_Guide.pdf

Kretz, T., Grünebohm, A., Kaufman, M., Mazur, F., & Schreckenberg, M. (2006). Experimental study of pedestrian counterflow in a corridor. *Journal of Statistical Mechanics: Theory and Experiment, 2006*(10), P10001. https://doi.org/10.1088/1742-5468/2006/10/p10001

Kuligowski, E.D. (2016). Human Behaviour in Fire. In: Hurley, M.J., et al. *SFPE Handbook of Fire Protection Engineering*. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-2565-0_58

Köster, G., Lehmberg, D., & Dietrich, F. (2016). Is Slowing Down Enough to Model Movement on Stairs? In *Traffic and Granular Flow '15*. Springer, Cham. https://doi.org/10.1007/978-3-319-33482-0_5

Köster, G., Lehmberg, D., & Kneidl, A. (2019). Walking on stairs: Experiment and model. *Physical Review, 100*(2). https://doi.org/10.1103/physreve.100.022310

Latané, B., & Darley, J. M. (1970). The unresponsive bystander : why doesn't he help? Appleton-Century Crofts. http://ci.nii.ac.jp/ncid/BA37928544

Liao, T. F. (2002). Statistical group comparison. In *Wiley series in probability and statistics*. https://doi.org/10.1002/9781118204214

Lin, Y., Wang, M. J., & Wang, E. M. (2004). The comparisons of anthropometric characteristics among four peoples in East Asia. *Applied Ergonomics, 35*(2), 173–178. https://doi.org/10.1016/j.apergo.2004.01.00

Lord, J., Meacham, B., Moore, A., Fahy, R., & Proulx, G. (2005). Guide for evaluating the predictive capabilities of computer egress models. Nist Gcr, 6, 886.

Lovreglio, R., Kuligowski, E. D., Gwynne, S., & Boyce, K. (2019a). A pre-evacuation database for use in egress simulations. *Fire Safety Journal, 105*, 107–128. https://doi.org/10.1016/j.firesaf.2018.12.009

Lovreglio, R., Ronchi, E., & Kinsey, M. (2019b). An online survey of pedestrian evacuation model usage and users. *Fire Technology, 56*(3), 1133–1153. https://doi.org/10.1007/s10694-019-00923-8

Ma, N., Song, W., Qi, X., Lü, W., & Cao, S. (2014). Simulation of evacuation in a twin bore tunnel: analysis of evacuation time and egress selection. *Procedia Engineering, 71*, 333–342. https://doi.org/10.1016/j.proeng.2014.04.048

Mashhadawi, M. (2016). MassMotion Evacuation Model Validation. Department of Fire Safety Engineering Lund University, Sweden.

Mayr, C. M., & Köster, G. (2021). Social Distancing with the Optimal Steps Model. *Collective Dynamics, 6*. https://doi.org/10.17815/cd.2021.116

Nelson, H.E. & Mowrer, F.W. (2002). Emergency Movement. In: Philip J. DiNenno (ed). SFPE handbook of fire protection engineering, third edition. Society of Fire Protection Engineers.

Nicolas, A., Bouzat, S., & Kuperman, M. N. (2017). Pedestrian flows through a narrow doorway: Effect of individual behaviours on the global flow and microscopic dynamics. *Transportation Research Part B-Methodological, 99*, 30–43. https://doi.org/10.1016/j.trb.2017.01.008

Nilsson, D., Frantzich, H., & Saunders, W. (2005). Coloured Flashing Lights To Mark Emergency Exits - Experiences From Evacuation Experiments. *Fire Safety Science, 8*, 569–579. https://doi.org/10.3801/iafss.fss.8-569

Nilsson, J., & Petersson, R. T. (2008). Utvärdering av videoanalysmetoder för utrymning med tillämpning på hörn. LUTVDG/TVBB--5256--SE; (2008).

Pelechano, N., & Malkawi, A. (2008). Evacuation simulation models: Challenges in modeling high rise building evacuation with cellular automata approaches. *Automation in Construction, 17*(4), 377–385. https://doi.org/10.1016/j.autcon.2007.06.005

Purser, D., & Bensilum, M. (2001). Quantification of behaviour for engineering design standards and escape time calculations. *Safety Science, 38*(2), 157–182. https://doi.org/10.1016/s0925-7535(00)00066-7

Research Institute of Human Engineering for Quality Life, 1994. Japanese body size data. Human Engineering for Quality Life, Japan (in Japanese).

Reynolds, C.W., 1999. Steering Behaviors For Autonomous Characters. Presented at *the Game developers conference*, pp. 763–782.

Rimea. (2022). Richtlinie für Mikroskopische Entfluchtungsanalysen. Retrieved from https://rimea.de/

Ronchi, E. (2013). Testing the predictive capabilities of evacuation models for tunnel fire safety analysis. *Safety Science, 59*, 141–153. https://doi.org/10.1016/j.ssci.2013.05.008

Ronchi, E. (2021). Developing and validating evacuation models for fire safety engineering. *Fire Safety Journal, 120*, 103020. https://doi.org/10.1016/j.firesaf.2020.103020

Ronchi, E., Kuligowski, E. D., Reneke, P. A., Peacock, R. D., & Nilsson, D. (2013). The process of verification and validation of building fire evacuation models. https://doi.org/10.6028/nist.tn.1822

Ronchi, E. & Nilsson, D., (2016). Basic Concepts and Modelling Methods, in: Cuesta, A., Abreu, O., Alvear, D. (Eds.), Evacuation Modeling Trends. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-20708-7

Ronchi, E., Norén, J., Delin, M., Kuklane, K., Halder, A., Arias, S., & Fridolf, K. (2015). Ascending evacuation in long stairways: Physical exertion, walking speed and behaviour. (TVBB-3192; Vol. 3192). Department of Fire Safety Engineering and Systems Safety, Lund University.

Ronchi, E., Reneke, P. A., & Peacock, R. D. (2013). A method for the analysis of behavioural uncertainty in evacuation modelling. *Fire Technology, 50*(6), 1545–1571. https://doi.org/10.1007/s10694-013-0352-7

Seitz, M. J. (2016). Simulating pedestrian dynamics: Towards natural locomotion and psychological decision making. https://mediatum.ub.tum.de/1293050

Seitz, M., & Köster, G. (2012). Natural discretization of pedestrian movement in continuous space. *Physical Review E, 86*(4). https://doi.org/10.1103/physreve.86.046108

Seyfried, A., Steffen, B., Klingsch, W., Lippert, T., & Boltes, M. (2007). The Fundamental Diagram of Pedestrian Movement Revisited — Empirical Results and Modelling. Springer EBooks, 305–314. https://doi.org/10.1007/978-3-540-47641-2_26

SFPE Europe Digital Issue 19 - ANTHROPOMETRIC DATA AND MOVEMENT SPEEDS. (n.d.). Higher Logic, LLC. https://www.sfpe.org/publications/periodicals/sfpeeuropedigital/sfpeeurope19/europeissue19feature3

Shi, X., Xue, S., Feliciani, C., Shiwakoti, N., Lin, J., Li, D., & Ye, Z. (2021). Verifying the applicability of a pedestrian simulation model to reproduce the effect of exit design on egress flow under normal and emergency conditions. *Physica D: Nonlinear Phenomena, 562*, 125347. https://doi.org/10.1016/j.physa.2020.125347

Sime, J. D. (1985). Movement toward the Familiar: Person and Place Affiliation in a Fire Entrapment Setting. *Environment and Behavior*, 17(6), 697–724. https://doi.org/10.1177/0013916585176003

Smedberg, E., Kinsey, M., & Ronchi, E. (2021). Multifactor variance assessment for determining the number of repeat simulation runs in evacuation modelling. *Fire Technology, 57*(5), 2615–2641. https://doi.org/10.1007/s10694-021-01134-w

Thompson, P. (2018). How to validate solutions from Simulex. Procedings at Fire and Evacuation Modeling Technical Conference (FEMTC) Gaithersburg, Maryland, October 1-3, 2018. Retrieved from https://files.thunderheadeng.com/femtc/2018_d1-09-thompson-paper.pdf

Thunderhead Engineering. (2021). Pathfinder Verification and Validation. Retrieved from https://support.thunderheadeng.com/docs/pathfinder/2021-4/verification-validation/

Tinaburri, A. (2022). Principles for Monte Carlo agent-based evacuation simulations including occupants who need assistance. From RSET to RiSET. *Fire Safety Journal, 127, 103510*. https://doi.org/10.1016/j.firesaf.2021.103510

Tong, D. & Canter, D. (1985). The decision to evacuate: a study of the motivations which contribute to evacuation in the event of fire. *Fire Safety Journal, 9*(3), 257–265. https://doi.org/10.1016/0379-7112(85)90036-0

Tsai, W. (2007). Validation of EvacuatioNZ Model for High-Rise Building Analysis. Department of Civil Engineering, University of Canterbury. New Zealand: Christchurch. http://dx.doi.org/10.26021/2305

Upton, G., & Cook, I. (2008). A Dictionary of Statistics. Oxford University Press, USA.

von Sivers, I., & Köster, G. (2015a). Dynamic stride length adaptation according to utility and personal space. *Transportation Research Part B-methodological, 74*, 104–117. https://doi.org/10.1016/j.trb.2015.01.009

von Sivers, I., Köster, G. (2015b). Realistic Stride Length Adaptation in the Optimal Steps Model. In: Chraibi, M., Boltes, M., Schadschneider, A., Seyfried, A. (eds) *Traffic and Granular Flow '13*. Springer, Cham. https://doi-org.ludwig.lub.lu.se/10.1007/978-3-319-10629-8_20

Weidmann, U. (1992). Transporttechnik der Fussgänger. Strasse Und Verkehr, 78(3), 161–169. https://doi.org/10.3929/ethz-a-000687810

Wijnhoven, P., Klein, M. (2014). Validation Of Evacuation Models Based On Video Footage Of People Leaving A Room. DGMR, Arnhem, Netherlands. Retrieved from https://files.thunderheadeng.com/femtc/2014_d2-16-wijnhoven-paper.pdf

Appendix A – Test reports

In this section full test reports for all tests are published. Due to copyright reasons this section can’t be published and is redacted in public available versions of this document.