



Multifactor Variance Assessment for Determining the Number of Repeat Simulation Runs in Evacuation Modelling

Erik Smedberg, Department of Fire Safety Engineering, Lund University, Lund, Sweden

Michael Kinsey, Arup Fire, Shanghai, China

Enrico Ronchi , Department of Fire Safety Engineering, Lund University, Lund, Sweden*

Received: 17 April 2020/**Accepted:** 27 April 2021/**Published online:** 19 May 2021

Abstract. Evacuation models commonly employ pseudorandom sampling from distributions to represent the variability of human behaviour in the evacuation process, otherwise referred to as ‘behavioural uncertainty’. This paper presents a method based on functional analysis and inferential statistics to study the convergence of probabilistic evacuation model results to inform deciding how many repeat simulation runs are required for a given scenario. Compared to existing approaches which typically focus on measuring variance in evacuation times, the proposed method utilises multifactor variance to assess the convergence of a range of different evacuation model outputs, referred to as factors. The factors include crowd density, flowrates, occupant locations, exit usage, and queuing times. These factors were selected as they represent a range of means to assess variance in evacuation dynamics between repeat simulation runs and can be found in most evacuation models. The application of the method (along with a tool developed for its implementation) is demonstrated through two case studies. The first case study consists of an analysis of convergence in evacuation simulation results for a building including 1855 occupants. The second case study is a simple verification test aimed at demonstrating the capabilities of the method. Results from the case studies suggest that multifactor variance assessment provides a more holistic assessment of the variance in evacuation dynamics and results provided by an evacuation model compared to existing methods which adopt single factor analysis. This provides increased confidence in determining an appropriate number of repeat simulation runs to ensure key evacuation dynamics and results which may be influenced by pseudorandom sampling are represented.

Keywords: Evacuation, Human behaviour in fire, Behavioural uncertainty, Evacuation simulation, Pathfinder, Modelling

*Correspondence should be addressed to: Enrico Ronchi, E-mail: enrico.ronchi@brand.lth.se



List of Symbols

ERD	Euclidean relative difference
ERD_{convj}	The convergence measure of the ERD of two consecutive ($j - 1$ and j) aggregated vectors
EPC	Euclidean projection coefficient
EPC_{convj}	The convergence measure of the EPC of two consecutive ($j - 1$ and j) aggregated vectors
KS-test	The Kolmogorov–Smirnov test. A non-parametric statistical test
MVA	Multifactor variance assessment
OETC	Occupant evacuation time curve
Occ	Occupants
RSR	Repeat simulation runs
SC	Secant cosine
SC_{convj}	The convergence measure of the SC of two consecutive ($j - 1$ and j) aggregated vectors
SD	Standard deviation
SD of Φ	The standard deviation of a specific value in a data set of particular interest
SD of Φ_{convj}	The convergence measure of the measure SD of Φ of two consecutive ($j - 1$ and j) aggregated vectors
TET	Total evacuation time
T_{lim}	Maximum allowable assembly time
$T_{0,95}^{max50}$	Maximum 95th percentile assembly time for the last 50 simulations
$T_{0,95}^{mean50}$	Mean 95th percentile assembly time for the last 50 simulations
$T_{0,95}^{min50}$	Minimum 95th percentile assembly time for the last 50 simulations
Φ	A specific value in a data set of particular interest
Φ_{convj}	The convergence measure of the measure Φ of two consecutive ($j - 1$ and j) aggregated vectors

1. Introduction

During the evacuation process people make a variety of decisions about what they will do. Previous research has shown that human behaviour in fire depends on several factors, such as a person's own past experience/perceptions [9, 11], the environmental conditions, and social influence, etc. [8, 16]. This explains why human behaviour can be highly varied during evacuations. Therefore, distributions are often used within evacuation models to reflect this variability [1]. The uncertainty associated with the representation of variability of human behaviour within an evacuation model is often referred to as 'behavioural uncertainty' [18, 21]. Unlike other types of uncertainty which are considered within evacuation modelling, behavioural uncertainty reflects the current understanding of human behaviour in fire. Current knowledge on evacuation behaviour is limited [2], thus distributions and/or stochastic modelling is the only feasible approach to represent evacuation behaviour without requiring a user to explicitly define each individual behaviour. Both of these methods employ pseudorandom sampling whereby the model uses pseudo randomly generated numbers to perform given tasks such as sampling from a distribution or deciding if a given behaviour is adopted. The gen-

eration of these pseudo randomly generated numbers varies between repeat simulation runs (RSR). As a consequence, many evacuation models produce varying results between RSR for the same evacuation scenario which introduce variability in evacuation dynamics. The process of generating these RSR can be seen as a type of “Monte-Carlo” simulation. Consequently, an evacuation model user must decide how many RSR are required to represent the range of key evacuation dynamics which influence results and subsequent decision-making process which the evacuation modelling informs, e.g., assessing if a building design is acceptable. The analysis of how many RSR are required then relies on the law of large numbers, specifying that the average of the results should be close to the *expected value*, and will tend to become closer as more RSR are performed. In addition, it is also possible to investigate the behavioural uncertainty associated with a given number of RSR by comparing the current results to the *expected value* (the evacuation model user would have to identify a method to estimate the *expected value*).

Previous work has been done in order to suggest methods to determine the number of runs needed in relation to the acceptance criteria and scenario under consideration [4, 5, 14, 21]. However, these methods have mainly focused on considering the overall evacuation time and evacuation curves (time series data) as main output of evacuation models. Due to such methods only focusing on what time people arrive at the final exits of a simulation, they do not explicitly consider the underlying variability in evacuation dynamics which could vary significantly more than that of the exit behaviour. Indeed, it may be possible for multiple repeat runs to produce identical or very similar evacuation times/curve output but exhibit very different evacuation dynamics. To address these issues, a method has been developed which considers multifactor variance assessment (MVA) that analyses a variety of evacuation simulation outputs (hereafter referred to as factors). Such factors include not only static outputs/results such as the time people evacuate, but also factors associated with spatial assessment to provide a more complete assessment in variability of evacuation dynamics. This is deemed to provide a more comprehensive assessment of behavioural uncertainty in a variety of relevant factors which are currently used in fire safety engineering applications. An example of such issues is that current assessment of space usage (i.e. occupant density and congestion levels) are generally performed in the fire safety engineering practice by looking at the results of individual simulations (which are representative only of one possible outcome). As such, the main contribution in this work has been to include the set of key factors used in fire safety engineering practice in the convergence assessment. Due to the inclusion of space-related factors, the MVA method is aimed towards usage of microscopic evacuation models which represent people as individuals and explicitly represent the physical space they occupy.

The MVA method is based on existing approaches and further expanded to address the specific issues associated with a wide list of outputs produced by evacuation models. In addition to the evacuation times (or arrival time/evacuation time curves), the MVA method investigates behavioural uncertainty of factors such as crowd density, flowrate, spatial location, exit usage and queuing time. The use of the MVA method is here demonstrated through two separate case studies,

each with a different aim. The first case study consisting of an evacuation simulation of a seven stories building with the use of the evacuation model Pathfinder (version 2018.3.0730) [25]. The first case study has been chosen as it represents a range of evacuation dynamics and so benefits from the advantages of employing the MVA method compared the existing methods for determining the number of runs. The second simpler case study adopts a modified version of the IMO 1533 verification test 10 which represents people evacuating a series of rooms/cabins on a single level. This is aimed at demonstrating the capabilities of the method through the comparison with an estimated *expected value* represented by 10,000 RSR. The evacuation model Pathfinder within the study has been chosen since it is a microscopic agent-based simulator, thus providing the granularity required for the analysis. Any other model with such characteristics could have been used. As part of the work, the method has been implemented in a tool which is freely available to aid practicing fire engineers in conducting convergence analysis in evacuation modelling.

This paper initially presents an overview of random sampling variables commonly found in evacuation models. Next follows an overview of the current methods available to determine the number of RSR in evacuation models. This overview provides context and basis of comparison with the proposed MVA method. The issue of data formatting, in particular when addressing the use of functional analysis for data series made of varying data points is also addressed. A description of the MVA method is then provided along with two case studies of its application. A discussion on the importance of addressing behavioural uncertainty in fire safety engineering practice is then provided.

Table 1
Random Sampling Variables

Variable	Description
Starting location	The initial starting location of occupants in a simulation
Occupant characteristics	The individual properties of an occupant (e.g. body size, comfort distance)
Movement speed	The assigned movement speed which occupants move whilst traversing the geometry
Pre-evacuation time	The delay time of occupants from the start of a simulation to when they start evacuating
Weights for route selection/collision avoidance algorithm	The definition of weights associated with route selection/collision avoidance algorithms which use a cost-based (or similar) function to make routes more/less attractive and represent pedestrian navigation
Exit attractiveness/availability	The likelihood of an occupant being aware of a given (available) exit
Influenced by toxic species	The likelihood a person will be affected by toxic species from a fire

2. Random Sampling Variables

Pseudorandom sampling from distributions/ranges can be used for a range of variables within evacuation modelling [12, 13, 20]. Table 1 presents a non-exhaustive list of potential variables which use random sampling within evacuation modelling. It is variation in such variables between RSR which cause differences in output from an evacuation model: the greater the variation in these variables the higher the likelihood of more RSR being required to achieve convergence in results. It should be noted that not all of these variables may be represented using pseudorandom sampling in an evacuation model, i.e., it is possible to represent many of them using static or fixed values whereby there would be no variation between RSR. In addition, where a model does use random sampling for a given variable, the underlying algorithm implemented within an evacuation model will influence the level of variation between RSR [17] which may give rise to variation between different evacuation models of the same scenario, i.e., some evacuation models might exhibit greater variation in evacuation modelling results. It should also be noted that such variables may be represented in varying degrees within a given evacuation modelling application, e.g., some may not be represented at all, which will influence the number of RSR required to achieve convergence in results.

3. Review of Current Methods

In order to highlight what is novel in the proposed MVA method, this section presents a review and comparison of methods, both at a conceptual level, and by reviewing previously developed methods.

Kinsey [10] has categorized possible methods to choose the number of RSR within evacuation modelling into four types:

1. *Brute force*: By simulating all possible permutations of the stochastic variables, it is ensured that the complete range of results has been obtained.
2. *Fixed number*: Setting a fixed number of RSR which is said to represent the potential outcomes of human behaviour sufficiently, as is the case in the recommendation to run the model 500 times in the IMO guidelines [6].
3. *Qualitative visual assessment*: Visually assessing the differences between runs to assess the differences in results then decide if more runs are needed.
4. *Dynamic assessment of variance in an output variable/series*: Utilizing the results from the simulation runs to assess whether or not convergence has been met with the use of quantitative methods.

The level of sophistication in the above methods varies with the latter being considered the most sophisticated as it considers a feedback process whereby output from results influences the number of RSR. It should be noted that all methods described above have their advantages and disadvantages. A brute force method will probably yield a theoretically infinitely large number of runs required, making it impractical in engineering practice. Utilizing a fixed number of runs

Table 2
Key Characteristics of Different Methods Available

References	[21]	[14]	[6]	[5]	[4]
Output(s) analysed	Data series: Evacuation time/occupants evacuated	Data series: Evacuation time/occupants evacuated	Assembly time (= Evacuation time)	Data point: Evacuation time/occupants evacuated	Data series: Evacuation time/occupants evacuated
Number of data points (per simulation run) used in the analysis	The number of occupants in the evacuation simulation	The number of occupants in the evacuation simulation	One (TET)	One (TET)	The number of occupants in the evacuation simulation
Measures included in the analysis	ERD EPC SC TET SD of TET No	ERD EPC SC TET SD of TET No	TET	TET	ERD EPC SC TET SD of TET NO
Reliant on a set ASET	No	No	Yes	Yes	NO
Method(s) to detect convergence in results	$M_{convg} = \left \frac{M_j - M_{j-1}}{M_j} \right $ (M = measure)	$M_{convg} = \left \frac{M_j - M_{j-1}}{M_j} \right $ (M = measure) KS-test	$\left T_{lim} - T_{0,95}^{mean50} \right \geq T_{0,95}^{max50} - T_{0,95}^{min50}$	Comparing the CI of the 95 th percentile simulated RSET with the set ASET	Comparing the width of the CI's of the measures with a specified tolerance

fails to evaluate if the number is too small or too large and assess convergence. A visual assessment might not be reliable for many RSR and may be associated with differences among different model users. The advantages and disadvantages of the last method will be addressed in the discussion section, as this is the conceptual model used in the proposed MVA method.

A number of methods have been proposed to determine the number of RSR of evacuation modelling results by assessing the variance in results [4, 5, 14, 21] (see Table 2). The method by Ronchi, Reneke and Peacock [21] is described as a method for quantitatively analysing the variance in modelling results in evacuation models. It is based on functional analysis concepts, which were previously applied in the context of a comparison between fire simulation data and experiments [19]. The method uses different operators (Euclidean relative difference (ERD), Euclidean projection coefficient (EPC) and the Secant cosine (SC)) to compare aggregated occupant-evacuation time curves. These operators are considered together with two additional outputs to assess variance between evacuation times (looking at aggregated normalized arithmetic mean and standard deviation of total evacuation time runs) [21]. The method then analyses the measure of the relative difference between j and $j - 1$ runs to detect convergence in results. The method by Lovreglio et al. [14] is an extension of the above mentioned method, complemented with inferential statistical testing (i.e. KS-test) to assess the same factors considered by Ronchi et al. [21] for the assessment of the variance between aggregated repeat runs.

The IMO guidelines [6] proposes two methods for determining the number of RSR: (1) To run at least 500 repeat runs randomizing the input variables, or (2) To use an “appropriate method” to determine the number of repeat runs through demonstrating *convergence of results*. The IMO guidelines provide one example of such method for determining convergence [6]. The first method implies that results may be acceptable according to the guidelines, even though they may not be convergent, after 500 repeated runs.

The method proposed by Grandison et al. [5] uses the 95th percentile of the total evacuation time (TET) for determining convergence. This means that only one data point per simulation run is used in the analysis, in contrast to using time series data. As such, the method does not consider the complete occupant/evacuation time curve data, thus convergence is likely to mainly rely on the total evacuation time. This method simplifies the analysis of data compared to time series data analysis methods previously discussed and is deemed suitable for RSET/ASET assessment. Nevertheless, it does not capture the variability of evacuation time curves, which may in turn be very important in the assessment of the whole evacuation process. This issue can be very important in certain scenarios, e.g., when studying phased evacuation procedures, assessment of the evolution of congestion at exit doors, etc.

The most recently proposed method [4] is also an extension of the first mentioned method. The method complements the method by Ronchi et al. [21] by introducing confidence intervals (CIs). The method is recognized to be more computationally costly than its predecessor, but provides a more standard interpretation of convergence through its use of well recognized statistical methods.

A limitation of the above-mentioned methods is that only the evacuation time and the occupant-evacuation time curves have been addressed in the behavioural uncertainty analysis. However, it may be possible for RSR to have comparable evacuation time curves, but variability in the underlying evacuation dynamics between different repeat runs. To address this issue, the MVA method is here proposed. The method takes into account several factors that may provide a more accurate assessment of convergence of repeat evacuation simulation results.

4. Multifactor Variance Analysis (MVA) Method

The MVA method includes factors associated with spatial results (rather than only time-related results), thus providing a more comprehensive assessment of variance in evacuation dynamics. The factors included in the proposed method are: crowd density, flowrate, queuing time, spatial location and exit usage. This selection was made performing a review of the outputs produced by the most used evacuation models [15] and prioritizing those deemed of key importance within evacuation dynamics for fire safety engineering applications. Note that this list of factors could be tailored for specific evacuation models or application domains depending on the underlying evacuation dynamics and the extent each factor occurs. For example, an evacuation scenario which does not involve large amounts of queuing or large numbers of occupants simultaneously moving may derive little benefit from assessing variation in flow rates as a factor between RSR.

The MVA method makes use of the assumption adopted in the method developed by Ronchi et al. [21], i.e., factors can be represented as vectors to calculate convergence. Each vector represents the results from one simulation run. In other words, the MVA method describe each factor as a multi-dimensional vector for which functional analysis operators can be calculated.

4.1. Factors

The factors included in this work (crowd density, flowrate, queuing time, spatial location and exit usage) are described, along with the unit of measures considered in their analysis (see Table 3).

Crowd Density (here referred to as local density) is an important factor to consider in evacuation dynamics since high densities may lead to issues such as crowd crush, congestions and other comfort and safety risks. In the context of pedestrian dynamics, local density can be calculated in several manners, but it is often measured in terms of the number of occupants per unit of floor area [3]. A more general definition of density in the evacuation modelling context relates to the number of simulated occupants in a 2D space, called a referenced area or space. Density is a concept derived from an analogy with fluid dynamics in which there are situations with a seemingly infinite number of particles [23]. In the context of evacuation, the number of occupants is discrete, which in turns makes the density concept more difficult to apply.

The issue that arises when the number of occupants is discrete, or when the reference area is small, is that large fluctuations may occur when occupants pass in

Table 3
Description of Factors Considered in the MVA Method

Factor	Description	Unit of measurement
Crowd density	The number of occupants occupying a reference area at a certain time	Occupants/m ²
Spatial location	The number of occupants occupying a room or area at a certain time	Number of occupants
Flowrate	The number of occupants passing a door over a set time	Occupants/s
Queuing time	The amount of time spent moving at a velocity lower than a certain threshold for the duration of the evacuation	s
Exit usage	The cumulative number of occupants which has used an exit	Number of occupants

and out of the reference area. These fluctuations may be treated by averaging over space and/or time, with the cost of lower resolution [23]. Another solution may be the use of Voronoi diagrams which is often referred as an appropriate method to maintain high resolution [23]. The MVA method does not depend on the use of Voronoi diagrams but the analysis of density based on this approach may yield more accurate results since the scatter is limited. In the case studies presented in this paper, Voronoi diagrams are used for the density estimations.

The MVA method includes the *spatial location* of the occupants since it provides insights into how the building is used in the evacuation. In large scenarios, to compare the exact locations of individual occupants would be too computational expensive. It could also be argued that an extremely high level of detail may not be necessary in fire safety engineering applications, while the location of the “mass” of occupants would be sufficient. This issue could be considered by counting the number of occupants in specific spaces, e.g., rooms or hallways (i.e., global density). This factor shows a high degree of similarity with the measurement of crowd density, with the difference that this would not be divided by the reference area and that the reference area of interest is generally larger than the reference area when measuring crowd density. In the MVA method, *spatial location* refers to the number of occupants in a specific area or room at a specific time.

Flowrate was included because it provides essential information about the evacuation process: at what rate occupants are evacuating, thus providing information about evacuation efficiency. Optimizing flowrates also enables the possibility of achieving lower evacuation times and higher safety levels. *Flowrate* is, just as density, a concept derived from fluid dynamics. In the context of pedestrian dynamics, it is often measured as the number of occupants which have passed through a doorframe or similar during a specific time interval. It is possible also in this case to apply Voronoi diagrams to maintain a high level of resolution while minimizing scatter. Similar to crowd density, the MVA method does not rely on the use of Voronoi diagrams for measuring flowrate.

Queuing time is a measure of the delay in the evacuation time. In addition, long queuing times may cause distress for the occupants, thus potentially affecting comfort and safety. Fruin [3] gives this definition of queuing: “*Queuing may be broadly defined as any form of pedestrian waiting that requires standing in a relatively stationary position for some period of time*”. It is important to note that queuing does not necessarily mean that the occupants stand in line. As occupants move towards an exit, they may try to maximize their chances of exiting yet allowing occupants in front to exit first. This phenomenon is called *dislocable queue* and has been observed in experiments [27]. Queuing may be measured in different ways: the congestion near the exits, the number of occupants in a queue, etc. Nevertheless, since queuing delays the evacuation, it is important to assess to which extent the *queuing time* is related to the evacuation process efficiency. Queuing causes occupants to slow down, which means that it is possible to measure the total time during the evacuation in which an evacuee moves at a slower pace (velocity) than desired, which is also the definition used in the MVA method.

The inclusion of the factor *exit usage* allows the increase of the level of understanding of the evacuation process, as it can be used to include a key underlying behaviour in the analysis. By considering the *exit usage*, it is possible to optimize the design of a building (i.e. placement and width of exits). For the purpose of the proposed method, *exit usage* will be defined as exit-specific occupant-evacuation-time-curves (OETCs), i.e. which exit is used by each occupant. It is noted that *exit usage* and *flowrate* describe the same behaviour or phenomena for the same exit to a certain extent. However, the concept of functional analysis, which the proposed method is based upon, is simply a way of calculating differences between two data sets by appearance. This means that the results from such a calculation may differ, even though the behaviour or phenomena is the same.

4.2. Methodology

The MVA method consists of a number of different steps in order to calculate and assess convergence. The flowchart in Fig. 1 is a modified version of the method flowchart proposed by Ronchi et al. [21], designed to fit the purpose of this work. The MVA method consists of an iterative process where acceptance criteria are specified and compared to the results from an arbitrarily defined number of repeat runs. If the acceptance criteria are not met, additional runs are required. The additional runs are then added to the existing batch of repeat runs. The results from this larger batch of repeat runs are then compared to the acceptance criteria again. This iteration is continued until the acceptance criteria are fulfilled and the analysis is concluded.

4.3. Mathematical Representation

The description of each individual factor in the form of multi-dimensional vectors can be found in the report associated with this paper [22]. In this paper, we only include the general mathematical description of the vectors. The MVA method relies on the representation of factors as vectors since it is based on a combination of functional analysis and inferential statistics. Factors are first defined as multi-

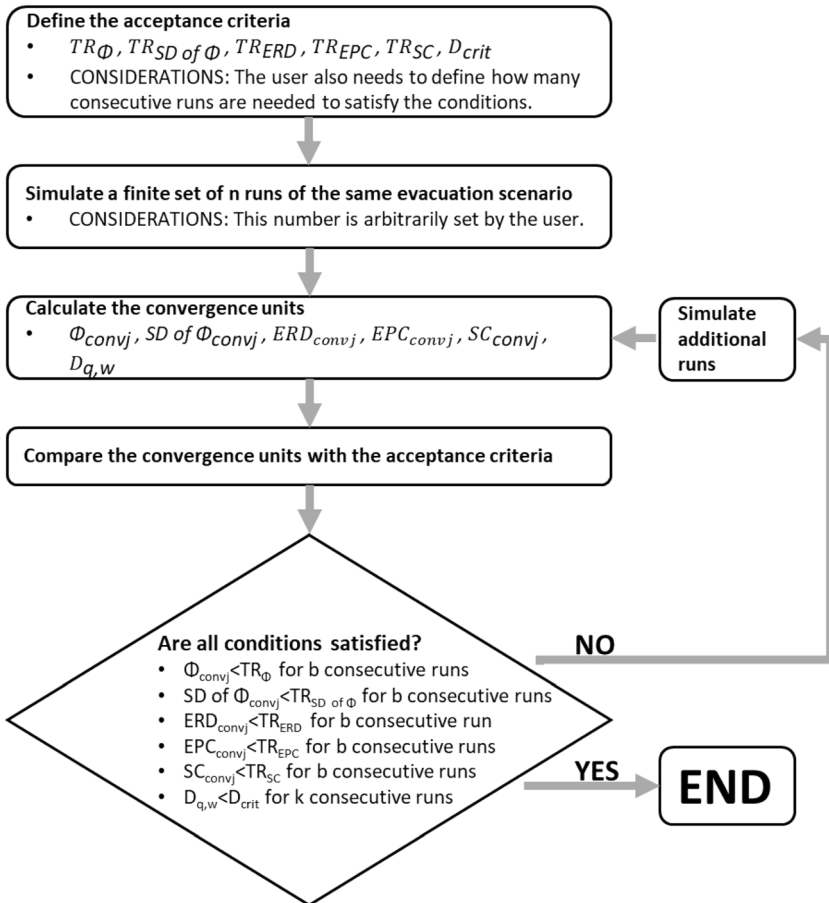


Figure 1. Schematic representation of the evaluation process.

dimensional vectors consisting of several data points (each of them representing a dimension). The data points may be either occupants or time steps, depending on what is applicable for the specific factor.

Consider a simulation consisting of q number of data points. The vector that describes the generic factor x would then be denoted as in Eq. 1:

$$\vec{x}_i = (x_1, x_2, \dots, x_{q-1}, x_q) \tag{1}$$

where i denotes a specific data point, x_1 corresponds to the first data point of x , x_2 to the second data point of x and so on. If we were to simulate n runs of the same scenario, n vectors \vec{x}_{ij} would be obtained, where n is the total number of runs and j denotes a specific run (see Eq. 2).

$$\vec{x}_{ij} = (x_{1j}, x_{2j}, \dots, x_{(q-1)j}, x_{qj}) \tag{2}$$

The next step is to present a variable which is associated with the arithmetic mean of the values of the runs. This means that the factors represent the arithmetic mean of the previous runs and not only the values for the specific run. If the total number of data points is still denoted q , and a specific run is denoted j , then the j th average curve, \vec{X}_j , is described by the following vector, presented in Eq. 3:

$$\vec{X}_j = (X_1, X_2, \dots, X_{q-1}, X_q) \tag{3}$$

where

$$X_1 = \frac{1}{j} \sum_{j=1}^{1 \leq j \leq n} x_{1j}, X_2 = \frac{1}{j} \sum_{j=1}^{1 \leq j \leq n} x_{2j}, \dots, X_q = \frac{1}{j} \sum_{j=1}^{1 \leq j \leq n} x_{qj}$$

For example, if $j = 1$, then \vec{X}_j would correspond to the values of the first run, i.e. \vec{x}_{i1} . If $j = 4$, then \vec{X}_j would correspond to the arithmetic mean of the values of four consecutive runs, i.e. the arithmetic mean of \vec{x}_{i1} , \vec{x}_{i2} , \vec{x}_{i3} and \vec{x}_{i4} .

The method proposed includes the use of six different convergence measures, five of which are included in the original method proposed by Ronchi et al. [21]: Φ_{convj} , SD of Φ_{convj} , ERD_{convj}, EPC_{convj}, SC_{convj}. Note that in the method by Ronchi et al. [21], Φ was denoted TET since this was the only factor included. Φ represents a specific value (the largest or most interesting) of each factor. The definition of Φ for each factor is presented in Table 4. Note that there is one value of Φ for each factor and RSR.

The additional convergence measure is the application of a non-parametric statistical test, the Kolmogorov–Smirnov (KS) test. This test has previously been used for the study of convergence evacuation model results, see Lovreglio et al. [14]. The KS-test was chosen in this application since it is a non-parametric test,

Table 4
The Definition of Φ for the Different Factors

Factor	Definition of Φ
Crowd density	The maximum crowd density measurement
Spatial location	The maximum number of occupants occupying a room or area at the same time
Flowrate	The maximum flowrate measured for the specified exit
Queuing time	The maximum queuing time measured for an occupant
Exit usage	The maximum number of occupants using the specified exit

i.e. no assumptions are made on the distribution type. This is deemed appropriate given the fact that the methodology should be applicable for different sample sizes.

The KS-test is used here to determine if two samples come from the same underlying distribution. In this case, the test will be applied on the aggregated values of two consecutive runs, i.e. between the curves \bar{X}_j and \bar{X}_{j-1} . This is achieved using the significance value α , and a number of consecutive runs the test needs to be passed, denoted k .

To conduct the test, the values need to be presented in the form of a cumulative distribution function, named $F_{j,q}(X)$, where q is the number of data points and j is the number of the run. To use TET as an example, the y axis on the graph of the function would represent the percentage of the occupants that have evacuated at that time (the x-axis value).

After the calculation of the convergence measures has been performed, they need to be compared to the user specified acceptance criteria. This includes both the absolute value of the acceptance criteria of the measures (e.g. 1% between \bar{X}_j and \bar{X}_{j-1}) as well as the arbitrarily set number of runs which the acceptance criteria must be fulfilled (e.g. 10 runs). Note that the proposed method does not provide any guidance on how to set these criteria, which will be discussed in Sect. 6.

4.4. Data Formatting

In the original method proposed by Ronchi et al. [21] the analysis was based on the number of occupants, which meant that the number of data points were the same between different runs. This could be done since the factor (evacuation time) is connected to the individual occupants, i.e. each occupant evacuates at a certain time. For some of the factors included in this work, the number of data-points have been identified based on a constant time interval, δt . This is because not all factors can be ascribed to individual occupants. Crowd density (at a specific place) for example is an emergent property which is the outcome of collective behaviour and can vary between runs. To solve this issue, the data set is divided into data points using time intervals as a delimiter. It could also be argued that the time aspect is important when measuring crowd density (or other factors) since it could provide useful insights into the evacuation process. The definition of the time step size is done prior to introducing the data sets to the MVA method. The method has currently been tested on a time step size of 1 s, as this is the default time step in the Pathfinder version used in the case studies presented in Sect. 5.

A possible issue when making the separation based on the constant time interval δt is that the number of time steps, q , will vary when the TET does, i.e. the number of time steps is smaller when the TET is shorter and vice versa. Different solutions can be adopted to solve this issue. A solution for this could be to let the simulation run with the most time steps, i.e. the simulation with the largest TET, set the number of time steps for all simulation runs. The rest of the simulation

runs are then filled out with time steps with a set value of zero or similar. However, this approach might yield unrealistic differences when comparing the curves since the factors do not necessarily approach the set value (i.e. zero or similar) when the simulation is completed. Doing the opposite, i.e. letting the simulation run with the smallest amount of time steps decide the number of time steps to include might yield more realistic results, but it will also mean that important data might be disregarded. Alternatively, the use of the average number of data points corresponds to a combination of the two above methods, thus carrying both their limitations.

The fourth option is to manipulate the data sets so that they all contain the same number of data points. This would mean that the division into data points would not be done with the use of a constant time interval (e.g. 1 s) but instead based upon how much of the simulation has been completed, i.e. a relative time difference, (e.g. 0.1%). This would mean that two data points may not represent the value at the exact same time in the simulation but rather how much of the evacuation process that is completed. As long as the TET does not vary significantly between runs, this approach is deemed useful to make the different curves comparable. In the current implementation of the MVA method, this option is implemented by modifying the number of time steps to be the same as the simulation run with the largest number of time steps through the use of linear interpolation.

4.5. Factor Averaging

The amount of variation in results is expected to be high depending on the assumptions used by the model to generate the factors (e.g., models make use of different approaches to estimate densities, flows, etc.). To compensate for this issue, where possible, a moving average approach is adopted to smooth vectors, thus making the method less sensitive to localised peaks caused by differences in calculation methods. This means that if the moving average interval is defined to be ± 10 s, then the value at the central data point is an average of the 10 previous, the central and the 10 sequent data points. Near the ends of the data sets, where data to conduct the moving average calculations does not exist, the data is cropped out.

4.6. Tool Implementation

This section describes how the MVA method was implemented in a spreadsheet tool which can be downloaded for free [22]. The MVA tool was developed using Visual Basic for Applications (VBA) and was made to read output files from the Pathfinder evacuation model version 2018.3.0730 [25] though could be adapted for other evacuation models with machine readable output files. In order to use the developed tool with other simulation software, an additional piece of code may be required to alter the format of the output data to be used as factors in the method. For example, the tool has been used in a validation study performed with the software FDS + Evac (version 2.5.2) [26].

The MVA tool reads the output data from the simulator and places them into vectors. The tool then proceeds in calculating the convergence measures described. For these calculations to be conducted, the user needs to define acceptance criteria and insert them as input into the tool. The user can also select the approach to be used for the varying number of data points discussed in Sect. 4.4. The user is also provided with the option to calculate and use a moving average in the convergence assessment to minimize the scatter in the data.

The output from the tool includes the descriptive statistics from the simulation runs analysed. The user is also presented with a graph containing the simulation results from all RSR, as well as a graph containing the aggregated runs. The results of the convergence assessment measures are also presented in a table format (i.e., individually for all factors).

5. Case Studies

Two separate case studies were conducted in order to demonstrate the functionalities/capabilities and limitations of the MVA method. The input/output values are purely exemplary in scope and no conclusion of the fire safety of the building under consideration should be drawn. The first case study (Case study 1) is used to demonstrate the MVA method on a realistic and comprehensive case. The second case study is used to investigate the performance and predictive capabilities on a simpler case (Case study 2), but where the results are also compared to the *expected value* (i.e. the value to which the sample is deemed to converge to), here represented by 10,000 simulation runs. In both cases, the simulation software Pathfinder version 2018.3.0730 [25] was used. The Pathfinder “Steering Mode” was used in both cases.

5.1. Case Study 1

The building used in the first case study is a hypothetical university building, consisting of seven stories, with two of them located below ground level. The building was equipped with three staircases between the floors. A total of five exits was present, all located on the ground floor. Two of them are to be regarded as main exits (see Fig. 2). The total occupancy is 1855 occupants distributed between the floors for the purpose of the case study. An overview of the geometry can be seen in Fig. 2.

A practical example is used as a case study rather than adopting a mathematical fictitious case, as done in previous studies [4, 21]. This is deemed appropriate given the fact that the factors included are of varying character (rather than only referring to evacuation times). The benefits of displaying the methodology using values obtained mathematically with a pseudorandom number generator (generalizability, lower computational cost, etc.) are however recognized by the authors.

The different factors to be analysed were measured at various points in the building, denoted here as points of interest (see Fig. 2). It should be noted that the factors which are not related to space (evacuation time and queuing time) do not require a point of interest to be measured. The selection of points of interest

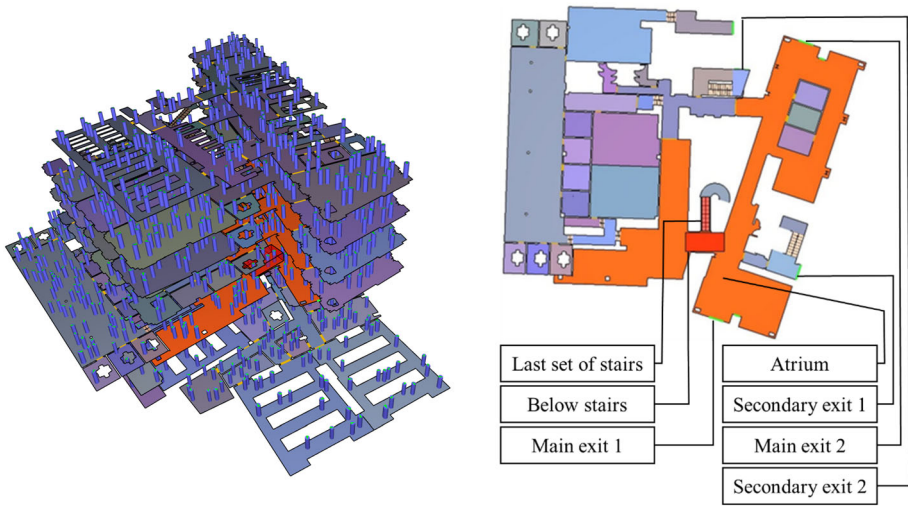


Figure 2. Screenshot from Pathfinder of the school building considered in the case study. Left: Overview of the model. Right: Ground floor with points of interest marked.

should be evaluated based on the specific case under consideration. To address this issue, the user may run a pilot study identifying potentially interesting areas (e.g. where congestion is present, etc.). A summary over the points of interest chosen can be found in Table 5. In order to minimize the scatter in results a moving average approach was utilized for Crowd density (± 15 s), Flowrate (± 30 s) and Spatial location (± 5 s). This was chosen after visually analysing the results from a pilot test.

**Table 5
Points of Interest Chosen for the Case Study**

Point of interest	Description	Factor
Main exit 1	One of the two main exits. Width = 2.0 m	Flowrate
Main exit 2	One of the two main exits. Width = 2.0 m	Exit usage
Secondary exit 1	Emergency exit adjacent to one of the staircases. Width = 2.0 m	Flowrate
Secondary exit 2	Emergency exit adjacent to one of the staircases. Width = 2.0 m	Exit usage
Last set of stairs	Density measurement region at the last set of stairs in the central staircase	Crowd density
Below stairs	Density measurement region at the bottom of the central staircase	Crowd density
Atrium	The main escape route on the ground floor	Spatial location

In order to obtain different results from the repeat runs of the simulation model, distributions/variables need to be included in the model to represent the variable nature of human behaviour. The inputs regarding occupant characteristics (i.e., walking speed and pre-evacuation time) were introduced as distributions. The horizontal walking speed follows a truncated normal distribution with the mean value of 1.5 m/s, standard deviation of 0.5 m/s, minimum value of 0.5 m/s and maximum value of 2.0 m/s. The pre-evacuation time is represented through a truncated log-normal distribution with $\mu = 4.5$ and $\sigma = 1.0$. This results in a median value of 90 s. The distribution was truncated at 5 and 300 s. No more user-defined inputs were made apart from these, i.e. the default values present in Pathfinder were utilized.

For each individual run, occupant characteristics have been randomly sampled and assigned to the occupants. Similarly, for each run, the initial occupant location within the domain has been randomized. Due to the algorithms present in Pathfinder (the locally quickest path algorithm [24]), route choice will also vary between runs as a result of occupant characteristics and positioning.

In order to conduct the analysis, acceptance criteria need to be defined and presented to the tool. The acceptance criteria used in the case study are presented in Table 6. These have been arbitrarily set in this example to show the application of the MVA method, i.e., they are purely for demonstration purpose.

A total of 80 simulation runs was arbitrarily chosen as the starting number of repeated simulations. This assumption is here arbitrary to show how a possible application of the MVA method would work. Table 7 provides the user with a description of the results from the 80 simulation runs used in the case study, as well as some description about its variation.

The results presented in Table 8 contain information about whether or not convergence was detected for the specific factor and, if so, at what run. Table 8 shows the results from the case study when utilizing the option to normalize the number of data points in the analysis.

The row titled “All (Max)” summarizes all convergence units for the factors studied. The highest value in this row determines when convergence has been met for all factors and convergence units. The results show that convergence has been met at the 58th run for all factors and convergence units studied. This implies that the user could have started with a lower number of runs and iteratively incremented this number until all criteria are met. The results are only presented for

Table 6
Acceptance Criteria Chosen for the Case Study

Factor/criteria	TR $_{\phi}$ (%)	TR $_{SD}$ of ϕ (%)	TR $_{ERD}$ (%)	TR $_{EPC}$ (%)	TR $_{SC}$ (%)	b	α (%)	k
Evacuation time	0.1	1.0	0.1	1.0	0.1	10	5	5
Queuing time	1.0	1.0	1.0	1.5	0.5	10	5	5
Crowd density	1.0	1.0	1.0	1.5	0.5	10	5	5
Flowrate	0.5	1.0	0.5	1.0	0.5	10	5	5
Spatial location	0.5	1.0	0.5	1.0	0.5	10	5	5
Used exit	0.5	1.0	0.5	1.0	0.5	10	5	5

Table 7
A Description of the Results from the 80 Simulation Runs Used in the Case Study

	Max	Min	Standard deviation	Average
Last set of stairs, Max density [occ/m ²]	3.04	1.43	0.48	2.17
Below stairs, Max density [occ/m ²]	2.62	0.72	0.51	1.58
Main exit 1, Max flowrate [occ/s]	2.61	1.95	0.13	2.39
Secondary exit 1, Max flowrate [occ/s]	1.75	1.07	0.13	1.43
TET [s]	860	578	43	637
Max queuing time [s]	589	293	44	364
Atrium, Max occupancy [occ]	181	123	11	151
Main exit 2, Max usage [occ]	380	287	20	347
Secondary exit 2, Max usage [occ]	419	328	20	382

Table 8
Results Provided by the MVA for the Case Study Under Consideration. The table reports if the acceptance criteria are met and if so after how many runs

	Has convergence been met (Yes/No)? At what run?																	
	Last set of stair		Below stairs		Main Exit 1		Secondary Exit 1		TET		Queuing time		Atrium		Main Exit 2		Secondary Exit 2	
φ	Y	45	Y	58	Y	28	Y	28	Y	44	Y	30	Y	38	Y	42	Y	28
SD	Y	30	Y	30	Y	17	Y	19	Y	19	Y	30	Y	18	Y	21	Y	18
ERD	Y	19	Y	42	Y	34	Y	17	Y	43	Y	33	Y	35	Y	31	Y	23
EPC	Y	46	Y	47	Y	19	Y	18	Y	13	Y	34	Y	29	Y	31	Y	20
SC	Y	28	Y	19	Y	15	Y	15	Y	12	Y	12	Y	15	Y	15	Y	13
KS-test	Y	15	Y	14	Y	13	Y	15	Y	6	Y	6	Y	8	Y	9	Y	14
All (max)	Y	46	Y	58	Y	34	Y	28	Y	44	Y	34	Y	38	Y	42	Y	28
All (min)	-	15	-	14	-	13	-	15	-	6	-	6	-	8	-	9	-	13
All (range)	-	31	-	44	-	21	-	13	-	38	-	28	-	30	-	33	-	15

the option to normalize the number of data points since this was shown to be the most effective [22].

5.2. Case Study 2

For the more simple case study, the IMO 1533 verification test 10 [6] was adopted which represents an evacuation scenario comprising 12 rooms/cabins connected via a corridor with a main exit and a secondary exit (see Fig. 3). Unlike in the IMO guidelines which suggest assigning an instant pre-evacuation time to all peo-

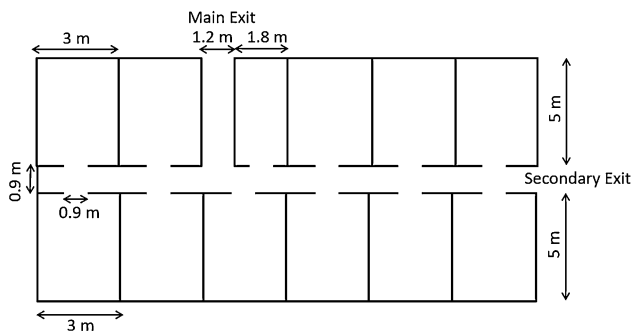


Figure 3. IMO 1533 verification test 10 [6].

ple, the pre-evacuation times were assigned from a normal distribution with mean (μ) = 15 s, standard deviation (σ) = 10 s, maximum = 30 s and minimum = 0 s. This is intended to introduce added variance to the case study and can be considered more realistic than all people responding at the same time. People were also assigned a walking speed from the same distribution as in Case study 1. The default exit choice algorithm of the evacuation model was used. The purpose of the IMO test is to demonstrate that agents use their assigned exit, however, only the geometric layout of the test is here used as a base to demonstrate the MVA method. The total number of people prescribed in IMO test is 24, however, this has been changed to 50 for the case study to increase levels of congestion along with associated contraflow. The 50 people were randomly positioned between the rooms at the start of the simulation. This was made as it was likely to generate larger variations in the results.

All factors included in the method were analysed in this case study. The number of people which used each exit and the exit flowrates were measured. Density was measured at areas in front of each exit. Spatial location was measured in the corridor connecting the rooms/cabins with the exits. As with Case study 1, occupant characteristics and occupant locations were randomly sampled between simulation runs.

A total of 10,000 RSR was conducted. The choice of this number of runs is arbitrary, but it was made to obtain a large sample of results which was likely to be convergent. This was done to represent the hypothetical *expected value*, which the results from a lower number of runs could then be compared to. This also means that acceptance criteria were not defined beforehand. Instead, the acceptance criteria are here seen as the results of the analysis for different number of runs. Given the scope of the analysis, the KS-test was not performed for this case study.

To demonstrate the trend of convergence of the results, a graph representing the change in average TET is shown in Fig. 4. The figure shows that the average TET broadly converges as the number of RSR increases, indicating that 10,000 RSR was a reasonable estimate of the *expected value* in this case.

Table 10
Results for Crowd Density at the Main Exit from Analysing 10,000 RSR

	Number of RSRs								
	25	50	100	250	500	1000	2500	5000	10,000
Average Max Density [occ/m ²]	0.99	0.97	0.94	0.91	0.91	0.92	0.93	0.93	0.93
Difference [occ/m ²]	0.06	0.04	0.01	- 0.02	- 0.02	- 0.01	0	0	0
Difference [%]	6.43	4.01	0.77	- 1.63	- 1.95	- 0.61	0.09	0.19	0
TR _Φ [%]	1.55	1.2	0.36	0.14	0.06	0.04	0.02	0.01	0
TR _{SD of Φ} [%]	0.7	0.63	0.19	0.07	0.03	0.02	0.01	0.01	0
TR _{ERD} [%]	1.36	0.72	0.25	0.05	0.1	0.04	0.01	0.01	0
TR _{EPC} [%]	1.9	1.16	0.78	0.2	0.11	0.06	0.02	0.01	0
TR _{SC} [%]	0.06	0.02	0	0	0	0	0	0	0

different number of RSR. It should be noted that the minimum number of consecutive runs for which the criteria needs to be fulfilled (b) was set to 10. The difference between the *expected value*, i.e. 10,000 runs and the current value at a given number of RSR, are presented for the value Φ as a demonstration of the adherence to convergence (no data is reported in the last column of Tables 9 and 10 for the differences and acceptance criteria as they would mean comparing the case of 10,000 runs against itself). This is denoted as “Difference” in the table. The results are presented in Tables 9 and 10.

As it can be seen in Tables 9 and 10, the values of the acceptance criteria needed to detect convergence decreases as the number of runs are increased. In some instances, the criteria increase temporarily when increasing the number of runs. This could indicate that the value of b in use, i.e. 10, might need to be increased to avoid this phenomenon.

6. Discussion

This work demonstrates the ability of the MVA method to analyse convergence for factors previously not included in this type of analysis. The analysed factors can represent a range of different metrics of the evacuation dynamics within the simulations. It was also shown that the method was able to detect convergence in cases where different numbers of data points between RSR were present.

The evacuation modelling Case study 1 required a total of 58 RSR to be conducted to reach the convergence based on the acceptance criteria in use. This number was due to the ‘below stairs’ crowd density assessment requiring 58 RSR. The factor which required the second most RSR was the ‘last stair’ crowd density assessment (requiring 46 runs). It is highlighted that users should be cautious in comparing the required number of RSR between factors due to inherent charac-

teristics of the functional analysis concept used, as well as the arbitrarily defined acceptance criteria. Nevertheless, the crowd density factor was the one who reached convergence last despite relatively non strict acceptance criteria (see Table 6). This together with a larger standard deviation (approximately 20% and 30% of the mean) suggest that crowd density is one of the factors showcasing a higher degree of variation in this case study. The results obtained can be used to interpret the scenarios. For instance, in case study 1, the variation in the crowd density factor is due to small differences in contributing factors such as movement speeds, arrival rates, pre-evacuation times, occupant starting location, all of which are defined according to random sampling. It should be noted that for Case study 1, the TET varied at most 35% from the average. As stated previously, the option to normalize the number of data points can be performed when there is not a significant change in TET. The results from the case study showed that the method was able to detect convergence despite this seemingly large variation in TET. This could be seen as guidance for future users of the method when evaluating the possibility to normalize the number of data points. In Case Study 1, a decision was made to start with 80 RSR, which was later proven to be larger than needed since convergence was detected after 58 RSR for all factors studied. However, as shown in Fig. 1, it is also possible to start with a smaller number of runs and iterate forward with batches of extra RSR until convergence is reached.

The evacuation modelling Case study 2 included a total of 10,000 RSR to estimate the expected value and then presented an assessment of variance at different numbers of RSR. After 25 RSR the average TET did not vary by more than one second. Similarly, the average maximum crowd density at the main exit did not vary by more than 0.1 occ/m^2 after 25 runs. This reflects the simplicity of the scenario being a small geometry with small numbers of people causing a reduction in variability in evacuation dynamics between RSR. This exemplifies that the users should evaluate the trade-offs between the complexity of the evacuation scenario and the required number of RSR. In addition, the evacuation model user would be required to represent the range of evacuation dynamics which are impacted by pseudorandom sampling. As with Case Study 1 the crowd density factor has the widest variation between RSR and the lowest level of variation between RSR were in the TET. Case Study 2 could be used to provide some guidance when selecting appropriate acceptance criteria. By comparing the difference from the *expected value* (i.e. 10,000 RSR) and the values of the acceptance criteria needed to detect convergence, a user can determine what to use. Nevertheless, it should be noted that this is only one case study, and another case study with different inputs would produce different results. Users of the method should also be aware that there is no use in comparing acceptance criteria for different factors or application domains. The concept of functional analysis simply compares curves by shape, and curves that look in a given manner (e.g. OETC or Flowrate) may be more or less easily detected as similar. One factor converging faster than another (given the same acceptance criteria) does not mean that any conclusions can be drawn when comparing different factors. In this application, functional analysis is simply used as a method to calculate similarities between time series data.

7. Limitations and Further Work

It should be highlighted that the MVA method relies on the selection of appropriate acceptance criteria (those have been arbitrarily chosen in the first case study). Further work is required to define what acceptance criteria could reliably be used or a method which allows an evacuation modelling user to calculate them. Considering the level of variation will be heavily influenced by the number and type of pseudorandom sampling method used by a model, it is contended that it be advantageous to use the range/distribution of these defined factors to inform the selection of suitable acceptance criteria.

Due to increased sensitivity in variation of certain factors (e.g. crowd density) it may be advantageous to adopt a smoothing process in the calculation of such factors whereby it is measured in larger time intervals in order to make the factor less sensitive to small differences in densities between RSR. Further work would be required to determine what time interval would be suitable in order to reflect significant differences in evacuation dynamics in sufficient granularity for a given area. In addition, careful selection is clearly required regarding which and how many areas within a geometry should be included for assessing crowd density. Such a process may benefit from a user initially running a model to identify key area of congestion to inform which areas are considered in the assessment.

By comparison to the MVA, if only the TET factor was considered when determining the number of RSR as adopted in past studies [21], the convergence for the case study would have been reached in 44 RSR, representing a 24.1% decrease compared to the number of RSR being required using the MVA given the chosen acceptance criteria. This is expected as the larger the number of factors considered in an assessment, the greater the likelihood of one of the factors taking longer to converge. The user is therefore required to balance the number and type of factors assessed required in relation to their need to suitably capture any key differences in evacuation dynamics in a given application. This may vary between different evacuation modelling applications so it is expected there may not necessarily be a 'one size fits all' list of factors to consider for all evacuation modelling applications.

The factors included in this paper are exemplary and should not be seen as an exhaustive representation of the evacuation process. The MVA method proposed in this paper could be applied to even more factors (e.g. mean walking speed, elevator usage, etc.) if necessary, since all factors included in this work were analysed efficiently and that they represent the variety of factors possible to measure in an evacuation scenario.

The proposed MVA method may also be used in validating simulation software against real world experiments (similarly to what has been done by Lovreglio et al. [14]). The factors implemented in the method could be measured during an evacuation trial rather than coming from an evacuation simulator. The benefits of the application of the MVA method are linked to the opportunity to assess how many repeated experiments are needed in a given condition to identify convergence of observations. From a modelling perspective, this would lead to a more

rigid validation procedure as more factors would be included and the behavioural uncertainty would be evaluated in the experimental data.

A limitation of the MVA method that was discovered through the case studies is that it is not efficient in analysing convergence when there is limited or no change in the data. This typically occurs at the start or end of the calculation for the following factors: flowrate, crowd density, exit usage and spatial location. This issue relates to the calculation of ERD. Even where there is a very limited change in the data, there may still be a constant (but small) difference between two aggregated data sets. This difference is then picked up by the ERD. A possible solution could be to crop the data set so that only the parts that is subject to regular change in the factor values would be included. This was however not tested in this work.

An alternative application of the proposed MVA method could be based on a bottom-up perspective, meaning that the user determines (visually and/or by descriptive statistics) when the results have converged enough for the given application of interest. This means that the user would present the acceptance criteria that would detect convergence at that point. This removes the daunting task of choosing appropriate acceptance criteria. If this approach were to be used by many users and for many different cases, a database of acceptance criteria could be developed, informing the community on what criteria are deemed appropriate for given scenarios.

The study of convergence is motivated by the concept of behavioural uncertainty. Therefore, it is important to note that the MVA method, along with previous methods, analyse the uncertainty implemented in the model, i.e. it is not a method to analyse behavioural uncertainty per se. Despite advances in understanding about human behaviour in fire, there is clearly a large amount of uncertainty regarding the subject matter and why variation in behaviour occurs [7]. As understanding regarding how people behave in fire progresses, evacuation models will extend in complexity and accuracy.

The MVA method uses a simple calculation of convergence as presented in the method developed by Ronchi et al. [21]. The work by Grandison [4] provides a novel extension to the method by Ronchi et al. [21] by using the concept of CI. Future work could attempt applying Grandison's [4] approach to multiple factors, in a similar fashion to what has been performed in this paper with the MVA method. The main limitation of the MVA method is the selection of suitable acceptance criteria. To date, there is little/no guidance on how these acceptance criteria should be set. It is beyond the scope of the paper and the topic for future work to investigate the methods for determining associated acceptance criteria. However, preferably, a statistical method would be applied in determining these criteria, possibly with the use of empirical data. This has been addressed by Grandison [4] but further work is needed in order to merge the methods. The focus of this paper is to highlight the importance of considering a range of factors for assessing variance in results between repeat simulation runs and thereby facilitate determining an appropriate number of repeat simulation runs.

Lastly, the associated evacuation modelling results from RSR will be used to inform a given decision within the building design process. It is therefore of

importance to appreciate that whilst variations may occur in the results of an evacuation model, consideration must also be given to the consequence of those variations for informing a decision within the wider building design process. Indeed, it should be highlighted that there may be circumstances where results vary in evacuation modelling output between RSR; however, this may not impact the wider decision-making process for a given building design process.

8. Conclusion

The proposed MVA method is designed to allow an analysis of evacuation model output convergence for a wide range of factors (not only for evacuation time related factors, as currently done by existing methodologies). This enables an analysis which is more comprehensive and at a greater level of detail than existing methods. This ensures that the underlying behaviours that govern TET also have converged, i.e. the problem that different behaviours may produce the same TET has been addressed.

By conducting this analysis, the user increases the likelihood that the possible types human behaviour represented in the model (influenced by pseudorandom sampling) have been simulated and that the range of results therefore represents the range of results which may take place in real life. When comparing evacuation model results to the ASET, this implies that the building under consideration is assessed taking the variability in human behaviour into account.

It is important to note that this type of analysis is dependent on the inputs defined by the user (e.g., distributions). The proposed MVA method does not analyse behavioural uncertainty per se but only the effect of the distributions and algorithms implemented in the model which is supposed to represent behavioural uncertainty.

A tool which implements the MVA method is freely released to fire safety engineers. This is deemed to increase the number of evacuation model users performing this type of analysis which may require a higher work load if conducted manually.

Acknowledgements

Enrico Ronchi wishes to acknowledge Thunderhead Engineering for providing the academic license of the Pathfinder software to perform this study, and Bryan Klein for the useful discussions on the topic of convergence of evacuation models.

Funding

Open access funding provided by Lund University.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alvear D, Abreu O, Cuesta A, Alonso V (2014) A new method for assessing the application of deterministic or stochastic modelling approach in evacuation scenarios. *Fire Saf J* 65:11–18. <https://doi.org/10.1016/j.firesaf.2014.02.005>
2. Averill JD (2011) Five grand challenges in pedestrian and evacuation dynamics. In: Peacock RD, Kuligowski ED, Averill JD (eds) *Pedestrian and evacuation dynamics* Springer US, Boston, pp 1–11. [10.1007/978-1-4419-9725-8_1](https://doi.org/10.1007/978-1-4419-9725-8_1)
3. Fruin JJ (1987) *Pedestrian planning and design* (revised edition). Elevator World, Inc, Mobile
4. Grandison A (2020) Determining confidence intervals, and convergence, for parameters in stochastic evacuation models. *Fire Technol* . <https://doi.org/10.1007/s10694-020-00968-0>
5. Grandison A, Deere S, Lawrence P, Galea ER (2017) The use of confidence intervals to determine convergence of the total evacuation time for stochastic evacuation models. *Ocean Eng* 146:234–245. <https://doi.org/10.1016/j.oceaneng.2017.09.047>
6. IMO (2016) MSC.1/Circ.1533. Revised guidelines on evacuation analysis for new and existing passenger ships
7. Kinateder M, Ronchi E (2019) Letter to the Editor: burning down the silos in a multi-disciplinary field. towards unified quality criteria in human behaviour in fire. *Fire Technol* 55:1931–1935. <https://doi.org/10.1007/s10694-019-00867-z>
8. Kinateder M, Ronchi E, Gromer D, Müller M, Jost M, Nehfischer M, Mühlberger A, Pauli P (2014) Social influence on route choice in a virtual reality tunnel fire. *Transport Res Part F Traffic Psychol Behav* 26:116–125. <https://doi.org/10.1016/j.trf.2014.06.003>
9. Kinateder MT, Kuligowski ED, Reneke PA, Peacock RD (2015) Risk perception in fire evacuation behavior revisited: definitions, related concepts, and empirical evidence. *Fire Sci Rev* 4:1. <https://doi.org/10.1186/s40038-014-0005-z>
10. Kinsey MJ (2016). *Evacuation/pedestrian modelling: methods for determining the number of repeat simulations*. London
11. Kuligowski E (2016) Human behaviour in fire. In: *SFPE handbook of fire protection engineering*. Springer, New York, pp 2070–2114
12. Kuligowski ED, Peacock RD, Hoskins BL (2010) A review of building evacuation models, 2nd edn. NIST Technical Note 1680. National Institute of Standards and Technology

13. Lord J, Meacham B, Moore A, Fahy R, Proulx G (2005) Guide for evaluating the predictive capabilities of computer egress models NIST GCR 06-886. National Institute of Standards and Technology
14. Lovreglio R, Ronchi E, Borri D (2014) The validation of evacuation simulation models through the analysis of behavioural uncertainty. *Reliab Eng Syst Saf* 131:166–174. <https://doi.org/10.1016/j.res.2014.07.007>
15. Lovreglio R, Ronchi E, Kinsey MJ (2019) An online survey of pedestrian evacuation model usage and users. Submitted for publication
16. Lovreglio R, Ronchi E, Nilsson D (2016) An Evacuation Decision Model based on perceived risk, social influence and behavioural uncertainty. *Simul Model Pract Theory* 66:226–242. <https://doi.org/10.1016/j.simpat.2016.03.006>
17. Lovreglio R, Spearpoint M, Girault M (2019) The impact of sampling methods on evacuation model convergence and egress time. *Reliab Eng Syst Saf* 185:24–34. <https://doi.org/10.1016/j.res.2018.12.015>
18. Machado Tavares R, Ronchi E (2015) Uncertainties in evacuation modelling: current flaws and future improvements. In: Boyce K (ed) Downing College, Cambridge, UK. Interscience, Downing College, Cambridge, UK, pp 185–196. ISBN 978-0-9933933-0-3
19. Peacock RD, Reneke PA, Davis WD, Jones WW (1999) Quantifying fire model evaluation using functional analysis. *Fire Saf J* 33:167–184. [https://doi.org/10.1016/S0379-7112\(99\)00029-6](https://doi.org/10.1016/S0379-7112(99)00029-6)
20. Ronchi E, Kinsey MJ (2011) Evacuation models of the future. Insights from an online survey on users experiences and needs. In: Capote J, Alvear D (eds) Proceeding of the advanced research workshop: “evacuation and human behaviour in emergency situations” Universidad de Cantabria, Santander, pp 145–155
21. Ronchi E, Reneke PA, Peacock RD (2014) A method for the analysis of behavioural uncertainty in evacuation modelling. *Fire Technol* 50:1545–1571. <https://doi.org/10.1007/s10694-013-0352-7>
22. Smedberg E (2019) The analysis of results of stochastic evacuation models. Report 5587. Department of Fire Safety Engineering, Lund University, Sweden
23. Steffen B, Seyfried A (2010) Methods for measuring pedestrian density, flow, speed and direction with minimal scatter. *Phys A* 389:1902–1910. <https://doi.org/10.1016/j.physa.2009.12.015>
24. Thornton C, O’Konski R, Klein B, Hardeman B, Swenson D (2014) New wayfinding techniques in pathfinder and supporting research. In: Weidmann U, Kirsch U, Schreckenberg M (eds) Pedestrian and evacuation dynamics Springer International Publishing, Cham, pp 1315–1322. [10.1007/978-3-319-02447-9_108](https://doi.org/10.1007/978-3-319-02447-9_108)
25. Thunderhead Engineering (2018) Pathfinder—Technical Reference
26. Wu Y (2019) Testing the International Standards Organization Verification and Validation protocol for evacuation simulations-An application to the FDS+ Evac model. LUTVDG/TVBB
27. Zhang J, Song W, Xu X (2008) Experiment and multi-grid modeling of evacuation from a classroom. *Phys A* 387:5901–5909. <https://doi.org/10.1016/j.physa.2008.06.030>