A Method for the Analysis of Behavioural Uncertainty in Evacuation Modelling

Enrico Ronchi[1*], Paul A. Reneke[2], Richard D. Peacock[2]

[1]Department of Fire Safety Engineering and Systems Safety, Lund University, P.O. Box 118. SE-22100 Lund Sweden. Tel +46 46 222 73 60, Fax +46 46 222 46 12

[2]National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

*corresponding author: enrico.ronchi@brand.lth.se

Abstract. Evacuation models generally include the use of distributions or probabilistic variables to simulate the variability of possible human behaviours. A single model setup of the same evacuation scenario may therefore produce a distribution of different occupant-evacuation time curves in the case of the use of a random sampling method. This creates an additional component of uncertainty caused by the impact of the number of simulated runs of the same scenario on evacuation model predictions, here named behavioural uncertainty. To date there is no universally accepted quantitative method to evaluate behavioural uncertainty and the selection of the number of runs is left to a qualitative judgement of the model user. A simple quantitative method using convergence criteria based on functional analysis is presented to address this issue. The method permits 1) the analysis of the variability of model predictions in relation to the number of runs of the same evacuation scenario, i.e. the study of behavioural uncertainty and 2) the identification of the optimal number of runs of the same scenario in relation to pre-defined acceptance criteria.

Keywords. *Evacuation modelling. Behavioural uncertainty. Human Behaviour in Fire. Functional Analysis. Convergence Criteria*

Published as: Ronchi E., Reneke P.A., Peacock., R. P. “A Method for the Analysis of Behavioural Uncertainty in Evacuation Modelling.” Fire Technology, July 9, 2013. doi:10.1007/s10694-013-0352-7.

1. Introduction

Uncertainty is divided into different components in the context of fire safety engineering and modelling [1]: model input uncertainty, measurement uncertainty, and intrinsic uncertainty.

1) Model input uncertainty is associated with the parameters obtained from experimental measurements that are used as model input, i.e. the assumptions employed to derive model input from the experiments.
2) Measurement uncertainty is associated with the experimental measurement itself, i.e., the data collection techniques employed.
3) Intrinsic uncertainty is the uncertainty associated with the physical and mathematical assumptions and methods that are intrinsic to the model formulation.

In the case of evacuation data, uncertainty includes an additional component, here named behavioural uncertainty. Behavioural uncertainty is uncertainty associated with the stochastic nature of human behaviour, i.e. human behaviour is stochastic per se [2], and a single experiment or model run may not be representative of a full range of the behaviours of the occupants. In fact, "evacuate the same building with the same people starting in the same places on consecutive days and the answers could vary significantly" [2]. There is a subsequent need for multiple experimental data-sets to understand the possible variability of occupant behaviours in each individual evacuation scenario [3]. Unfortunately, experimental data-sets on human behaviour in fire are scarce and single data-sets are often the only available reference for the study of an individual scenario. Behavioural uncertainty needs to be analysed in both experimental and modelling studies. In this context, the assessment of the variability of simulation results in relation to behavioural uncertainty is a key issue to be discussed. This is reflected in the estimation of the convergence of an individual evacuation simulation scenario towards an "average" predicted occupant evacuation time-curve. It should be noted that the term behavioural uncertainty is here introduced in the context of fire safety science, i.e. the term may have different meanings in other research fields.

Fire modellers and evacuation modellers treat uncertainty in different ways. Uncertainty is generally treated in fire models as a deterministic problem, i.e., it is studied by analysing the sensitivity of the model output in relation to the variability of the model input. This is driven by the fact that fire models are generally based on deterministic equations (e.g. [4, 5]). On the other hand, evacuation models treat uncertainty as a stochastic problem. In fact, to address the stochastic nature of human behaviour, evacuation models often employ distributions or stochastic variables to simulate people movement and behaviours [6-10] (e.g. distribution of walking speeds, distribution of pre-evacuation times, exit choice, etc.). In fact, random numbers/seeds may be employed to solve space conflict resolution, simulate exit choice, familiarity with the exit, queuing behaviour, etc. When distributions are created adopting a random sampling method, multiple occupant-evacuation time curves for the same scenario using the same model inputs are produced. Random variables may be intrinsic of the model algorithms, and model users may not have control/access to them (especially in closed-source models). This leads to the need for a study of the variability of the results associated with the random variables embedded in the models.

Therefore, evacuation modellers face the problem of selecting the appropriate number of runs to be simulated in order to be representative of the average model outcome. This problem arises both during the use of evacuation models for a fire safety design as well as during validation studies. In fact, two main questions can be asked during the simulation of evacuation scenarios that include distributions or stochastic variables: 1) Which occupant-evacuation time curve is representative of model predictions in a fire safety design? 2) Which occupant-evacuation time curve should be used as reference during the comparison with experimental data in a validation study? To date, the answers to these questions are left to a qualitative judgment by the evacuation model user. For instance, in the context of evacuation model validation, model users may select the best model prediction during the comparison with experimental data [11] or employ the model's average total evacuation time (possibly including information on the standard deviation) as representative of model predictions. The study of the average total evacuation times and their corresponding standard deviations provides insights only on the required safe escape time, rather than the whole evacuation process. There is instead a need for a method which investigates the size of the variation for the whole occupant-evacuation time curve. Nevertheless, to date, there is no universally accepted quantitative method to estimate how these average predictions may vary over the number of runs.

In addition, complex evacuation scenarios may be computationally expensive to simulate. For instance, previous research on the use of distribution curves for Monte Carlo simulations for uncertainty analysis in evacuation model predictions have demonstrated the need for a large computational effort [7]. Therefore there is a need to optimize the selection of the number of runs of the same scenario in order to be representative of occupants' "average behaviour", and provide a quantitative and computationally inexpensive measurement of the variability associated with the simulated runs (and a subsequent estimation of the behavioural uncertainty associated with an individual evacuation model setup).
A useful method for the analysis of model predictions is functional analysis. This branch of mathematics represents curves as vectors, and uses geometrical operations on the curves. Functional analysis operations are currently employed during the comparison of fire model evaluations and experimental data [12, 13] and the comparison between evacuation model results and experimental data [14]. Nevertheless, functional analysis has not been employed so far to compare evacuation model predictions against each other to analyse the uncertainty associated with the number of runs of the same evacuation scenario, i.e. behavioural uncertainty.

This paper proposes a set of convergence criteria for the analysis of the variability of evacuation model predictions of the same evacuation scenario (i.e. the same model input which includes distributions or stochastic variables) in relation to the number of runs. A procedure for the definition of the optimal number of runs - in relation to the evacuation scenario, the model in use, and the scope of the simulations - is presented. The scope of the present work is therefore to provide a quantitative method to assess the variability associated with the number of runs of the same evacuation scenario. The proposed method allows the analysis of behavioural uncertainty and the prediction of the average occupant-evacuation time curve in relation to pre-defined acceptance criteria.

A case study about the application of the method is presented. The case study is an explanatory example in which a fictitious data-set (i.e. a data-set created using a pseudo-random generator) is employed to show the convergence criteria and the evaluation procedure.

The last part of the paper discusses the benefits associated with the use of the convergence criteria and future work regarding their possible uses.

2. Method

This section presents a proposed methodology for the analysis of behavioural uncertainty. It includes the definition of five convergence criteria for the analysis of the occupant-evacuation time curves produced by evacuation models and a procedure for the assessment of the optimal number of runs in relation to pre-defined acceptance criteria.

The proposed methodology is based on the definition of a set of convergence measures that sufficiently describe the distribution of occupant-evacuation time curves. This is addressed by constructing a series for each measure and demonstrating that the measure is sufficiently close to the expected value, i.e. the series converge to the average occupant-evacuation time curve.

A series $S = \{s_{i,} \ldots, s_n\}$ converges to S_c if for any positive real value e there is an n such that $|S_c - s_n| < e$.

The series represents the evacuation time predictions of evacuation models and they are based on sample data. This will imply that the series will likely not smoothly converge, meaning that it might happen that $|S_c - s_{n+1}| > |S_c - s_n|$. In order to increase the confidence that our series have sufficiently converged, a requirement that the last b values of the series (the convergence measures) are within S_c is added. For some series we might not know the expected value S_c, i.e., the value to which the series is convergent. In those cases the last current value of the series is used as the best estimate of the value the series converges to.

2.1 Functional analysis concepts

Before discussion of convergence criteria, there is a need to introduce three concepts of functional analysis, namely the Euclidean Relative Difference (ERD), the Euclidean Projection Coefficient (EPC) and the Secant Cosine (SC). Initial applications of these concepts have been used in different research fields (e.g., mechanics [15], engineering [16], etc.), including fire science (see Peacock et al [12] and Galea et al [14]).

The single comparison of two individual points in a curve can be made by finding the norm of the difference between the two vectors representing the data. A norm represents the length of a vector. The distance between two vectors corresponds to the length of the vector resulting from the difference of the two vectors. For a generic vector $\vec{\mathbf{x}}$, the norm is represented using the symbol $||\vec{\mathbf{x}}||$. This concept can be extended to multiple dimensions. The distance between two generic multi-dimensional vectors $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ is therefore the norm of the difference of the vectors $||\vec{\mathbf{x}} - \vec{\mathbf{y}}||$. The Euclidean relative difference between two vectors can be normalized as a relative difference to the vector $\vec{\mathbf{y}}$ (see Equation 1).

$$ERD = \frac{||\vec{\mathbf{x}} - \vec{\mathbf{y}}||}{||\vec{\mathbf{y}}||} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(y_i)^2}} \quad \text{[Equation 1]}$$

The Euclidean Relative Difference (ERD) represents, therefore, the overall agreement between two curves.

Two components can be considered during the comparison of two vectors, namely the distance between two vectors and the angle between the vectors.
The concept of projection coeffictient a is introduced. From a geometric point of view, the vector $a\vec{\mathbf{x}}$ is the projection of the vector $\vec{\mathbf{y}}$ onto the vector $\vec{\mathbf{x}}$ (see Fig. 1).

Fig. 1 The projection coefficient for two vectors.

a defines a factor which reduces the distance between two vectors to its minimum (see Fig. 1). The solution of the minimum problem is found and corresponds to Equation 2.

$$a = \frac{||\vec{\mathbf{y}}||}{||\vec{\mathbf{x}}||}\cos\beta \quad \text{[Equation 2]}$$

$< \vec{\mathbf{x}}, \vec{\mathbf{y}} >$ is the inner product of two vectors, i.e., the product of the length of the two vectors and the cosine of the angle between them. The inner product can be interpreted as the standard dot product; producing Equation 3.

$$< \vec{\mathbf{x}}, \vec{\mathbf{y}} > = \sum_{i=1}^{n}(x_i y_i) \quad \text{[Equation 3]}$$

The Euclidean Projection Coefficient (EPC) is found by studying the minimum problem, i.e., studying when the derivative of the function is zero (see Peacock et al [12] for the full solution of the minimum) and it corresponds to:

$$a = EPC = \frac{<\vec{\mathbf{x}}, \vec{\mathbf{y}}>}{||\vec{\mathbf{y}}||^2} = \frac{\sum_{i=1}^{n}(x_i y_i)}{\sum_{i=1}^{n} {y_i}^2} \quad \text{[Equation 4]}$$

EPC defines a factor which when multiplied by each data point of the vector $\vec{\mathbf{y}}$ reduces the distance between the vectors $\vec{\mathbf{y}}$ and $\vec{\mathbf{x}}$ to its minimum, i.e. the best possible fit of the two curves.

The concept of Secant Cosine (SC) is also introduced. It represents a measure of the differences of the shapes of two curves. This is investigated by analysing the first derivative of both curves.
For n data points, a multi-dimensional set of n-1 vectors can be defined to approximate the derivative. This produces Equation 5 [12]:

$$SC = \frac{<\vec{\mathbf{x}},\vec{\mathbf{y}}>}{||\vec{\mathbf{x}}||\,||\vec{\mathbf{y}}||} = \frac{\sum_{i=s+1}^{n} \frac{(\Delta x_{i-s})(\Delta y_{i-s})}{s^2(\Delta t_{i-1})}}{\sqrt{\sum_{i=s+1}^{n} \frac{(\Delta x_{i-s})^2}{s^2(\Delta t_{i-1})} \sum_{i=s+1}^{n} \frac{(\Delta y_{i-s})^2}{s^2(\Delta t_{i-1})}}} \quad \text{[Equation 5]}$$

Where: t is the measure of the spacing of the data, i.e. t =1 if there is a data point for each occupant;
s represents the number of data points in the interval;
n is the number of data points in the data-set.

$$\Delta x_{i-s} = x_i - x_{i-s}$$
$$\Delta y_{i-s} = y_i - y_{i-s}$$
$$\Delta t_{i-1} = t_i - t_{i-1}$$

When the Secant Cosine is equal to unity, the shapes of the two curves are identical. Depending on the value for s, the noise of the data is smoothed out. An example of the impact of different values of s on the SC is shown in Fig. 2 and 3. Fig. 2 shows two hypothetical curves (obtained by 120 values for x and y corresponding to 120 arbitrary data-points) which include noise or no noise. The comparison between the shapes of the two curves is made using different s values (from s=1 to s=60 in this example), i.e., Fig. 3 shows that the use of higher values for s reduces the impact of the noise in the comparison.

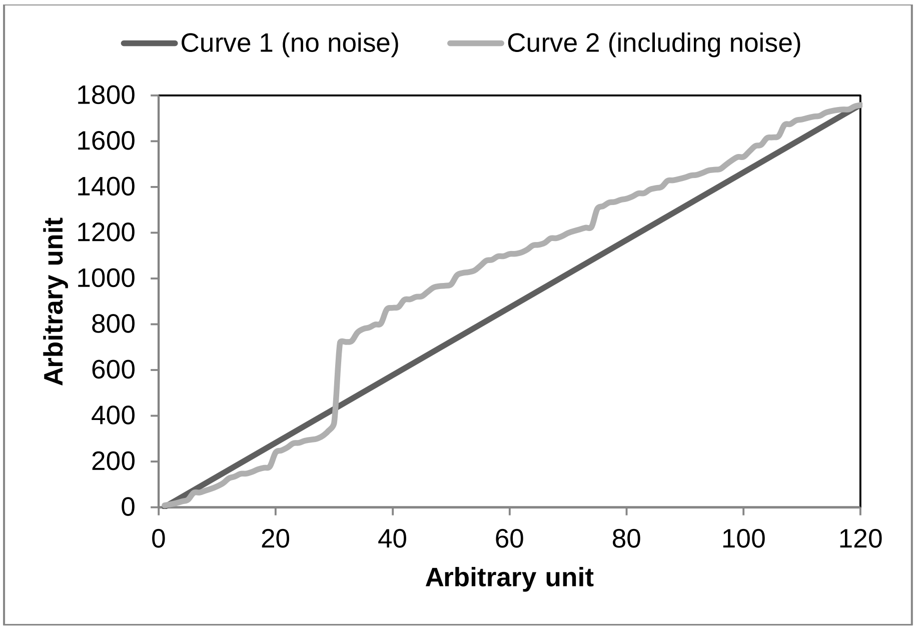

Fig. 2 Hypothetical curves including noise (grey curve) and not including noise (black curve)

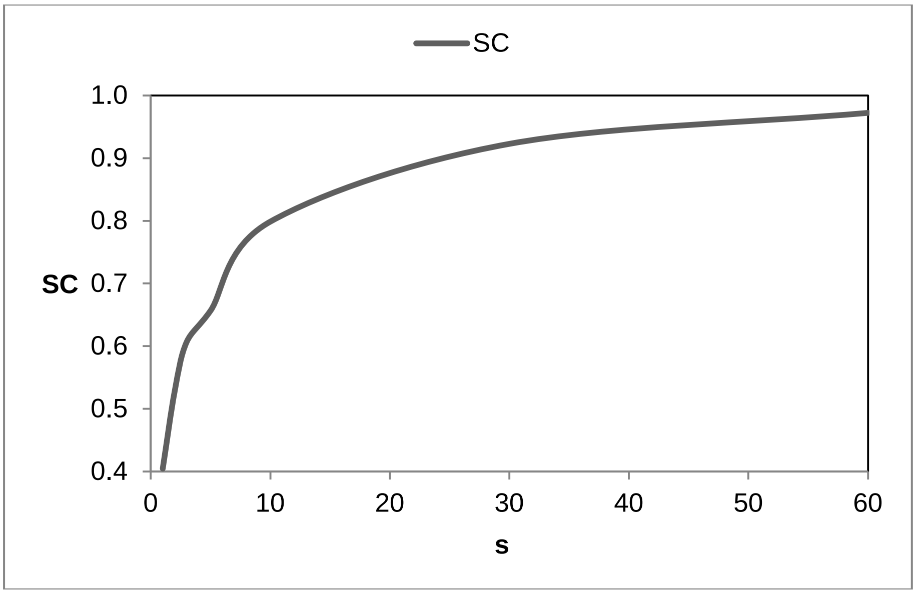

Fig. 3 Secant Cosine in relation to different s values.

Nevertheless, s should not be too large, so that the natural variations in the data are kept. An example of this issue is provided in Fig. 4, where, considering a hypothetical set of 4 data-points, different values for s generate either SC=1 for s=4 (the shape of the curves appear identical) or SC≠1 in the case of s=1 and s=2.

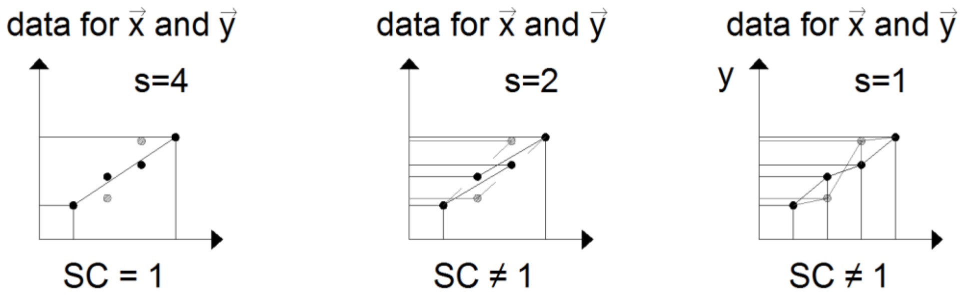

Fig. 4 Schematic representation of the use of different values for s during the calculation of the Secant Cosine.

2.2 Convergence measures

A set of variables are introduced in order to present the method of analysis of evacuation model predictions based on functional analysis and convergence criteria.

The measured experimental data are represented using vector $\vec{\mathbf{E}}$ (see Equation 6), where E_i represents the measured evacuation time for the i^{th} occupant.

$$\vec{\mathbf{E}} = (E_1, \ldots, E_n) \qquad \text{[Equation 6]}$$

For example, in the case of *i*=3 occupants, i.e., $\vec{\mathbf{E}} = (E_1, E_2, E_3)$, E_1 is the measured evacuation time corresponding to occupant 1, E_2 is the measured evacuation time corresponding to occupant 2 and E_3 is the measured evacuation time corresponding to occupant 3.

The simulated predicted times are represented by the vector $\vec{\mathbf{m}}$ (see Equation 7), where m_i is the simulated evacuation time for the *i^th^* occupant. m_n represents the evacuation time corresponding to the last occupant out of the building.

$$\vec{\mathbf{m}} = (m_1, \dots, m_n) \quad \text{[Equation 7]}$$

Therefore, $\vec{\mathbf{m}} = (m_1, m_2, m_3)$, where m_1 is the simulated evacuation time corresponding to occupant 1, m_2 is the simulated evacuation time corresponding to occupant 2 and m_3 is the simulated evacuation time corresponding to occupant 3.

Several runs of the same scenarios are simulated. The simulated evacuation times of each occupant *i* in each j^{th} run are represented using *n* vectors $\vec{\mathbf{m}}_{\mathbf{ij}}$ (see Equation 8). Here, *q* is the total number of occupants and *n* is the total number of runs. One assumption is that occupants are ranked in accordance to their evacuation time, i.e. occupants may evacuate the building in a different order in different runs.

$$\vec{\mathbf{m}}_{\mathbf{ij}} = \left(m_{11}, \dots, m_{ij}, \dots, m_{qn}\right) \quad \text{[Equation 8]}$$

Considering nine runs of the same evacuation scenario including the same three occupants, 9 vectors $\vec{\mathbf{m}}_{\mathbf{ij}}$ are obtained where *i*=3 and j=9, i.e., $\vec{\mathbf{m}}_{\mathbf{i1}} = (m_{11},\ m_{21}, m_{31})$, $\vec{\mathbf{m}}_{\mathbf{i2}} = (m_{12},\ m_{22}, m_{32})$, …, $\vec{\mathbf{m}}_{\mathbf{i9}} = (m_{19},\ m_{29}, m_{39})$.

The next variable that is presented is associated with the calculation of the arithmetic mean of the values of the runs. The j^{th} average curve of evacuation times produced by the model considering the arithmetic mean of the values of all runs is represented using an *n* dimensional vector $\vec{\mathbf{M}}_{\boldsymbol{j}}$ (see equation 9), where $\mathrm{M}_1 = \frac{1}{n}\sum_{j=1}^{n} m_{1j}$, $\mathrm{M}_2 = \frac{1}{n}\sum_{j=1}^{n} m_{2j}$, …, $\mathrm{M}_n = \frac{1}{n}\sum_{j=1}^{n} m_{qj}$.

$$\vec{\mathbf{M}}_{\boldsymbol{j}} = (\mathrm{M}_1,\ \dots, \mathrm{M}_j, \dots, \mathrm{M}_n) \quad \text{[Equation 9]}$$

Considering the previous example, i.e. 3 occupants and 9 runs (*i*=3 and *j*=9), the average curve $\vec{\mathbf{M}}_{\mathbf{1}}$ corresponds to the values of the first run. The average curve for a sub-set of 4 runs will generate $\vec{\mathbf{M}}_{\mathbf{4}}$ which corresponds to the arithmetic means of the values up to the fourth run. In the case of all 9 runs, $\vec{\mathbf{M}}_{\mathbf{9}}$ corresponds to the arithmetic means of the values of all runs.

Figure 5 presents vector $\vec{\mathbf{M}}_{\boldsymbol{j}}$ in relation to the number of runs under consideration.

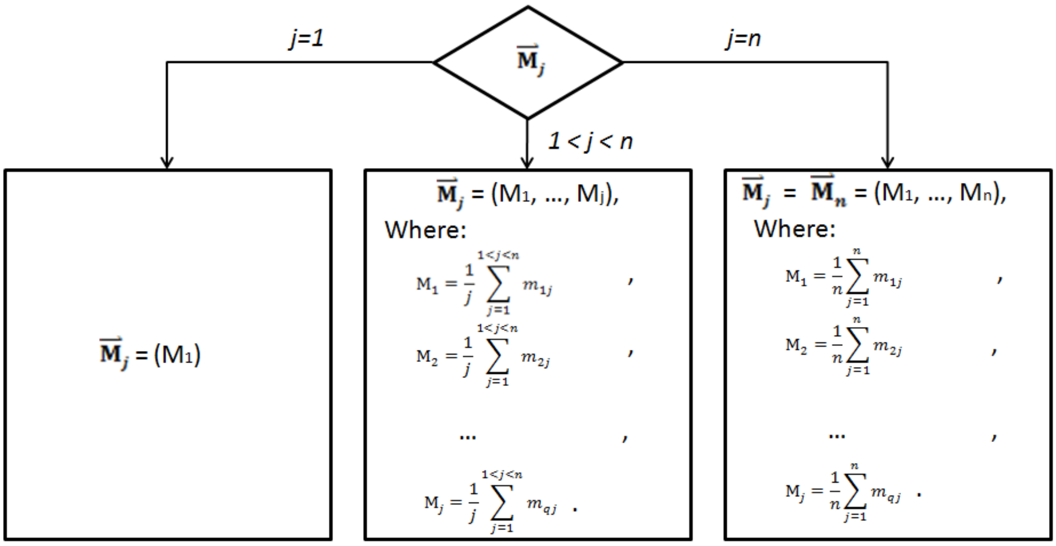

Fig. 5. Vector $\overrightarrow{\mathbf{M}}_j$ in relation to the considered number of runs.

Hence if $j=1$, $\overrightarrow{\mathbf{M}}_j$=($\mathrm{M}_1$), i.e. the average curve corresponds to the curve of the first run. If $1 < j < n$, $\overrightarrow{\mathbf{M}}_j$ becomes $\overrightarrow{\mathbf{M}}_j$=($\mathrm{M}_1$, …, M_j) where $\mathrm{M}_1 = \frac{1}{j}\sum_{j=1}^{1<j<n} m_{1j}$, $\mathrm{M}_2 = \frac{1}{j}\sum_{j=1}^{1<j<n} m_{2j}$, …, $\mathrm{M}_j = \frac{1}{j}\sum_{j=1}^{1<j<n} m_{qj}$. $\overrightarrow{\mathbf{M}}_j$ represents then the average curve corresponding to $1 < j < n$ runs. Considering 4 vectors $\overrightarrow{\boldsymbol{m}}_{ij}$ corresponding to the predicted evacuation times for three occupants in j=4 runs out of n=9 runs, $\overrightarrow{\mathbf{M}}_4$=($\mathrm{M}_1 = \frac{1}{4}\sum_{j=1}^{1<4<9} m_{1j}$,$\mathrm{M}_2 = \frac{1}{j}\sum_{j=1}^{1<4<9} m_{2j}$, $\mathrm{M}_3 = \frac{1}{4}\sum_{j=1}^{1<4<9} m_{3j}$). If $j=n$, $\overrightarrow{\mathbf{M}}_j$ becomes $\overrightarrow{\mathbf{M}}_n$=($\mathrm{M}_1$, …, M_n) where $\mathrm{M}_1 = \frac{1}{n}\sum_{j=1}^{n} m_{1j}$, $\mathrm{M}_2 = \frac{1}{n}\sum_{j=1}^{n} m_{2j}$, …, $\mathrm{M}_j = \frac{1}{n}\sum_{j=1}^{n} m_{qj}$. Thus, $\overrightarrow{\mathbf{M}}_j$ represents the average curve corresponding to all $j=n$ runs. For instance, if n=9 runs, $\overrightarrow{\mathbf{M}}_9$=($\mathrm{M}_1 = \frac{1}{9}\sum_{j=1}^{9} m_{1j}$, $\mathrm{M}_2 = \frac{1}{9}\sum_{j=1}^{9} m_{2j}$ $\mathrm{M}_3 = \frac{1}{9}\sum_{j=1}^{9} m_{3j}$).

2.2.1 Convergence measure 1: Total Evacuation Time (TET)

The vector m_n can also be called TET_j, total evacuation time (also called Required Safe Egress Time in the context of performance based design [17]), corresponding to the j^{th} run. Therefore, there are several simulated TET_j, each one corresponding to the j^{th} run for a total of n runs.

The j^{th} total evacuation times TET_i for n runs of the same scenario simulated with an evacuation model can be represented using the vector $\overrightarrow{\boldsymbol{TET}} = (TET_1, \dots, TET_n)$.

The arithmetic mean of the total evacuation times for j runs can be expressed using TET_{avj} (see Equation 10):

$$TET_{avj} = \frac{1}{j}\sum_{i=1}^{j} TET_i \qquad \text{[Equation 10]}$$

The set of all n consecutive mean total evacuation times TET_{avj} of the same scenario simulated with an evacuation model is TET_{av} =(TET_{av1}, ..., TET_{avn}). TET_{av1} is assumed to correspond to the value in run 1, TET_{av2} is the average for j=2, …, TET_{avn} the average for j= n.

Applying the law of large numbers, the consecutive mean total evacuation times TET_{avi} can be interpreted as a series converging to an expected value (the mean total evacuation time). Hence, a measure of the convergence of the series can be performed.

A measure of the convergence of two consecutive mean total evacuation times TET_{avj} (e.g. TET_{av1} and TET_{av2}) is obtained calculating TET_{convj} (see Equation 11). It is expressed (in %) as the difference of two consecutive mean total evacuation times divided by the last mean evacuation time. This convergence measure assumes that the best approximation of the expected value (the mean total evacuation time) is the last mean evacuation time.

This produces a total of $p=n-1$ TET_{convj}.

$$TET_{convj} = \left| \frac{TET_{avj}-TET_{avj-1}}{TET_{avj}} \right| \qquad \text{[Equation 11]}$$

The last TET_{convj} value, corresponding to all n runs is $TET_{convFIN}$ (see Equation 12).

$$TET_{convFIN} = \left| \frac{TET_{avp}-TET_{avp-1}}{TET_{avp}} \right| \qquad \text{[Equation 12]}$$

2.2.2 Convergence measure 2: Standard Deviation (SD) of total evacuation times

Convergence variables can also be presented in terms of the standard deviation of total evacuation times.

The j^{th} standard deviation SD_j for n runs of the total evacuation time of the same scenario simulated with an evacuation model can be represented by the vector $\overrightarrow{\boldsymbol{SD}} = (SD_1, \ldots, SD_n)$.

Also in this case, the application of the law of large numbers permits the interpretation of the consecutive standard deviations of total evacuation times SD_j as a series convergent to an expected value (the mean standard deviations of total evacuation times). Therefore, a measure of the convergence of the series is possible.

A measure of the convergence of two consecutive standard deviations SD_j (e.g. SD_1 and SD_2) is obtained by calculating SD_{convj}. It is expressed (in %) as the difference of two consecutive standard deviations divided by the last standard deviation (see Equation 13). This produces a total of $p=n-1$ SD_{convj}. This convergence measure assumes that the best approximation of the expected value (the mean standard deviation of total evacuation times) is the last standard deviation of total evacuation times.

$$SD_{convj} = |\frac{SD_j - SD_{avj-1}}{SD_j}| \quad \text{[Equation 13]}$$

The last SD_{convj} value, corresponding to all *n* runs, is $SD_{convFIN}$ (see Equation 14).

$$SD_{convFIN} = |\frac{SD_{avp} - SD_{avp-1}}{SD_{avp}}| \quad \text{[Equation 14]}$$

2.2.3 Convergence measure 3: Euclidean Relative Difference (ERD)

A set of Euclidean Relative Differences (ERD) can be calculated, each one corresponding to two consecutive pairs of vectors $\overrightarrow{\mathbf{M}_j}$ representing the progressive average occupant-evacuation time curves.

A vector $\overrightarrow{\boldsymbol{ERD}} = (ERD_1, \dots, ERD_p)$ is made of *p* consecutive ERD_j where *p=j-1*, corresponding to average *j* runs of the same scenario simulated with an evacuation model. For instance, in the case of *j=4* runs, $\overrightarrow{\boldsymbol{ERD}} = (ERD_1, ERD_2, ERD_3)$ where ERD_1 is calculated from the comparison between M_1 and M_2, ERD_2 is calculated from the comparison between M_2 and M_3 and ERD_3 is calculated from the comparison between M_3 and M_4. M_1 represents the curve from run 1, M_2 represents the average curve generated by the arithmetic means of the individual occupant evacuation times for run 1 and run 2, M_3 represents the average curve generated by the arithmetic means of the individual occupant evacuation times for run 1, run 2 and run 3. M_4 represents the average curve generated by the arithmetic means of the individual occupant evacuation times for run 1, run 2, run 3 and run 4.

The consecutive ERD_j can be interpreted as a series convergent to the expected value equal to 0 (the case of two curves identical in magnitude). Hence, a measure of the convergence of the series is possible. A measure of the convergence of two consecutive Euclidean Relative Differences ERD_j, corresponding to two consecutive average curves $\overrightarrow{M_j}$ can be obtained calculating ERD_{convj} (see Equation 15). It is expressed as the absolute value of the difference of two consecutive Euclidean Relative Differences ERD_j and ERD_{j-1}.

$$ERD_{convj} = |ERD_j - ERD_{j-1}| \quad \text{[Equation 15]}$$

The last ERD_{convj} value, corresponding to the differences between the latest average curves is $ERD_{convFIN}$ (see Equation 16).

$$ERD_{convFIN} = |ERD_p - ERD_{p-1}| \quad \text{[Equation 16]}$$

Calculation of ERD_{convj} permits estimation of the impact of the number of runs on the overall differences between consecutive average curves. $ERD_{convFIN}$ represents therefore a tool to understand the *behavioural uncertainty* associated with multiple runs of an individual evacuation scenario.

2.2.4 Convergence measure 4: Euclidean Projection Coefficient (EPC)

The same type of convergence measures can be produced for the Euclidean Projection Coefficient (EPC).

The consecutive EPC_j can be interpreted as a series convergent to the expected value equal to 1 (the best possible agreement between two consecutive EPC_j). Hence, a measure of the convergence of the series can be performed. This results in Equation 17 and 18.

$$EPC_{conv\,j} = |EPC_j - EPC_{j-1}| \quad \text{[Equation 17]}$$

$$EPC_{convFIN} = |EPC_p - EPC_{p-1}| \quad \text{[Equation 18]}$$

$EPC_{conv\,j}$ permits the estimation of the impact of the number of runs on the possible agreement between two consecutive average curves. $EPC_{convFIN}$ is therefore another indicator of the *behavioural uncertainty* associated with multiple runs of an individual evacuation scenario.

2.2.5 Convergence measure 5: Secant Cosine (SC)

Convergence measures can be developed for the Secant Cosine (SC). The consecutive SC_j can be interpreted as a series convergent to the expected value equal to 1 (the case of two identical shapes of consecutive curves). Hence, a measure of the convergence of the series can be performed and it is presented in Equation 19 and 20.

$$SC_{conv\,j} = |SC_j - SC_{j-1}| \quad \text{[Equation 19]}$$

$$SC_{convFIN} = |SC_p - SC_{p-1}| \quad \text{[Equation 20]}$$

$SC_{conv\,j}$ allows understanding of the impact of the number of runs on the possible differences between the shapes of two consecutive average curves. $SC_{convFIN}$ represents therefore a variable to understand the behavioural uncertainty associated with the average shape of the simulated curves, given a certain number of runs *n* of the same evacuation scenario.

2.3 The evaluation method

Five variables have been presented in the previous section, namely $TET_{convFIN}$, $SD_{convFIN}$, $ERD_{convFIN}$, $EPC_{convFIN}$, and $SC_{convFIN}$. Those variables represent the basis for a novel evaluation method. The proposed method addresses two key aspects of evacuation modelling:

1) The analysis of behavioural uncertainty of an individual evacuation scenario.
2) The identification of the optimal number of runs to produce a stable evacuation curve of the same scenario in relation to the evacuation scenario and the model in use.

An iterative method is suggested for the evaluation of evacuation model results. The method is based on five steps (see Fig. 6).

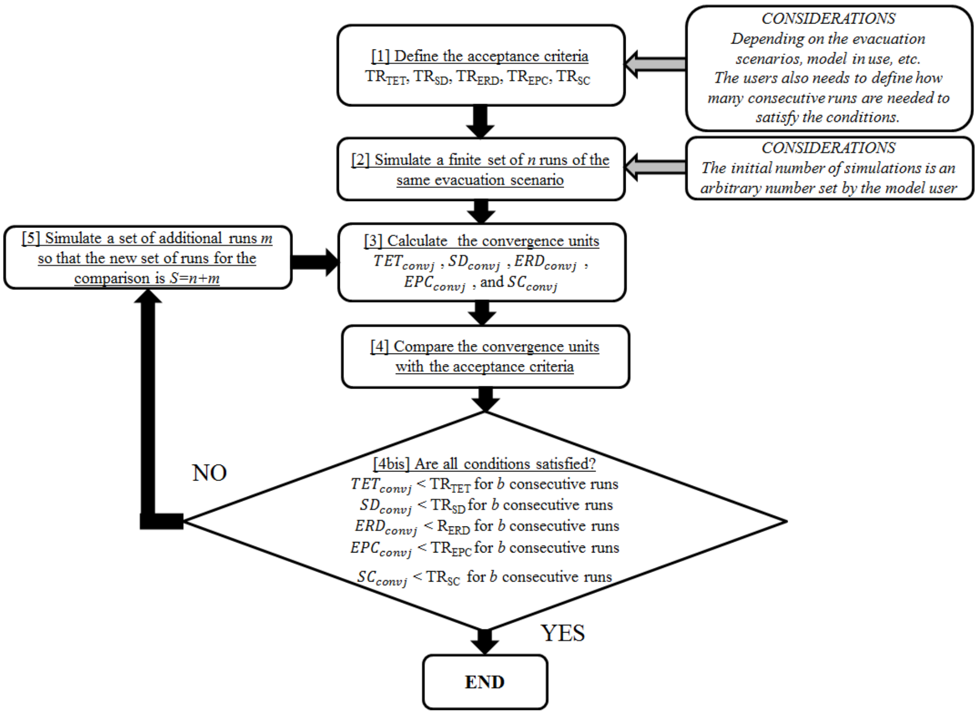

Fig. 6 Schematic flow chart of the proposed evaluation method.

Step 1. Define the acceptance criteria. (see (1) in Fig. 6)

The first step of the method consists of the identification of the acceptable thresholds to be achieved, i.e. the accepted behavioural uncertainty associated with the average curve obtained by multiple runs of the same scenario. The aim is to obtain an evacuation curve that is sufficiently stable given the scope of the analysis. For example, in the case of the use of evacuation modelling in the context of performance based design, the identification of these acceptable thresholds can be based on the estimated uncertainty during the calculation of the ASET (Available Safe Escape Time) produced using a fire model. This approach permits a joint analysis of the uncertainty associated with both the fire and evacuation simulations. Five thresholds (corresponding to the five convergence measures) are identified, namely TR_{TET}, TR_{SD}, TR_{ERD}, TR_{EPC}, TR_{SC}. It should be noted that there is an additional acceptance criteria that needs to be assessed, i.e., a finite number of consecutive runs *b* for which the acceptable thresholds must not be crossed. This needs to be assessed in order to verify that the convergence measures are stable under certain thresholds over a pre-defined number of runs. This requirement is based on the assumptions described in Section 2. The larger is b, the higher is the confidence that can be put on the fulfilment of the acceptance criteria.

The identification of the acceptance criteria may depend on several factors such as the evacuation scenario, the model in use, etc. The selection of the acceptance criteria - which may or may not include all convergence measures - may be identified by the evacuation modeller itself or from a third party.

Step 2. Simulate a finite set of n runs of the same evacuation scenario (see (2) in Fig. 6)

Evacuation model users select an arbitrary initial number of simulations of an individual evacuation scenario, i.e., the same model input is used. n vectors $\vec{\boldsymbol{m}}_{\boldsymbol{ij}} = (m_{11}, \dots, m_{ij}, \dots, m_{qn})$ corresponding to the simulated evacuation times of each occupant i^{th} in each j^{th} run are obtained. The occupant-evacuation time curves are produced ranking the occupants in relation to their evacuation time.

The vector corresponding to the consecutive average curves $\vec{\mathbf{M}}$=($\mathrm{M}_1, \dots, \mathrm{M}_n$) is also generated.

In order to optimize the iterative process, the selection of the initial arbitrary number of runs may be based on a qualitative evaluation made by the evacuation modeller of the variability of the predicted outcome given the model input of the scenario under consideration. Nevertheless, this judgment - which is the current qualitative method adopted by evacuation modellers to estimate the optimal number of runs - is not mandatory, since the proposed method permits a quantitative study of the impact of the number of runs on the occupant-evacuation time curve produced by the model.

Step 3. Calculate the convergence measures (see (3) in Fig. 6)

The convergence measures presented in the previous sections are calculated for all runs, i.e., TET_{convj} , SD_{convj} , ERD_{convj} , EPC_{convj} , and SC_{convj}.
In order to perform the calculation of the secant cosines for all runs, model users need also to identify a finite set of values for s, needed for the calculation of SC_{convj}. As described in Section 2.1, the choice of the values for s relies on the dataset under consideration. SC_{convj} are calculated for all runs for as many s values as chosen by the model user.

Step 4-4bis. Compare the convergence measures with the acceptance criteria (see (4-4bis) in Fig. 6)

The model user compares the calculated convergence measures against the acceptable thresholds defined during Step 1. This produces five tests that need to be accomplished:

TEST 1:

$TET_{convj} < \mathrm{TR_{TET}}$ for b consecutive number of runs [Equation 21]

TEST 2:

$SD_{convj} < \mathrm{TR_{SD}}$ for b consecutive number of runs [Equation 22]

TEST 3:

$ERD_{convj} < \mathrm{TR_{ERD}}$ for b consecutive number of runs [Equation 23]

TEST 4:

$EPC_{convj} < \mathrm{TR_{EPC}}$ for b consecutive number of runs [Equation 24]

TEST 5:

$$SC_{convj} < \text{TR}_{\text{SC}} \text{ for } b \text{ consecutive number of runs} \quad \text{[Equation 25]}$$

It should be noted that the criteria need to be satisfied for a pre-defined finite number of consecutive b runs (as defined during Step 1). The values corresponding to the j^{th} run where the conditions are verified for b consecutive runs represent $TET_{convFIN}$, $SD_{convFIN}$, $ERD_{convFIN}$, $EPC_{convFIN}$, and $SC_{convFIN}$.

If the five conditions are all satisfied for a pre-defined number of consecutive runs, the curves generated by *n* number of runs meet the acceptance criteria, i.e. the average curve is estimated given an accepted behavioural uncertainty associated with the number of runs (based on the acceptance criteria). If one or more of the conditions are not satisfied, the model user needs to proceed with Step 5.

Step 5. Simulate a set of additional simulations *m* so that the new set of runs for the comparison is *S*=*n*+*m*. (see (5) in Fig. 6)

The model user sets an arbitrary number of additional simulations to be run. The definition of the additional runs can be set in accordance with a qualitative analysis of the failed tests. A new set of $S=n+m$ $\vec{\boldsymbol{S}}_{\boldsymbol{ij}}$ vectors $\vec{\boldsymbol{S}}_{\boldsymbol{ij}} = (S_{11}, \dots, S_{ij}, \dots, S_{q\,S})$ corresponding to the average simulated evacuation times of each occupant i^{th} in each j^{th} run are obtained. The same methodology of Step 2 is adopted to produce the occupant-evacuation time curves, i.e., the occupants are ranked in relation to their evacuation time. The model user can now re-start the procedure starting from Step 3.

3. Case study

An application of the method presented in Section 2 is described to provide an example of the concepts. Given the explanatory scope of the example, data used in this section are fictitious, i.e., they do not correspond to real data. This choice has been driven by the current lack of repeated experimental data, i.e. the method has been applied to study simulation results. Data are created in order to be representative of the results obtained with an evacuation model for a hypothetical evacuation scenario. A fictitious set of numbers is produced using Wichman & Hill's [18] pseudo-random generator. The pseudo-random numbers are used as input to produce lognormal-distributed values. This choice was made in order to be representative of a hypothetical evacuation scenario which is influenced by pre-evacuation times (which generally follow a log-normal distribution [17]).The fictitious data are then used to create fictitious individual evacuation times calculated by progressively summing the values obtained (in order to be representative of a hypothetical real case study where evacuation ranges approximately between 1100 s and 1900 s). For example, if the first pseudo-random generated number is 12.41 and the second pseudo-random generated number is 18.18 s, the evacuation time of the first occupant out would correspond to 12.41 s and the evacuation time of the second occupant out would be 12.41 s + 18.18 s = 30.59 s. The procedure is repeated for all 120 occupants (See Table 1).

An example of one possible curve is provided in Fig. 7. The assumed population consists of 120 occupants. The evaluation of the number of runs to be simulated is the unknown variable of this example.

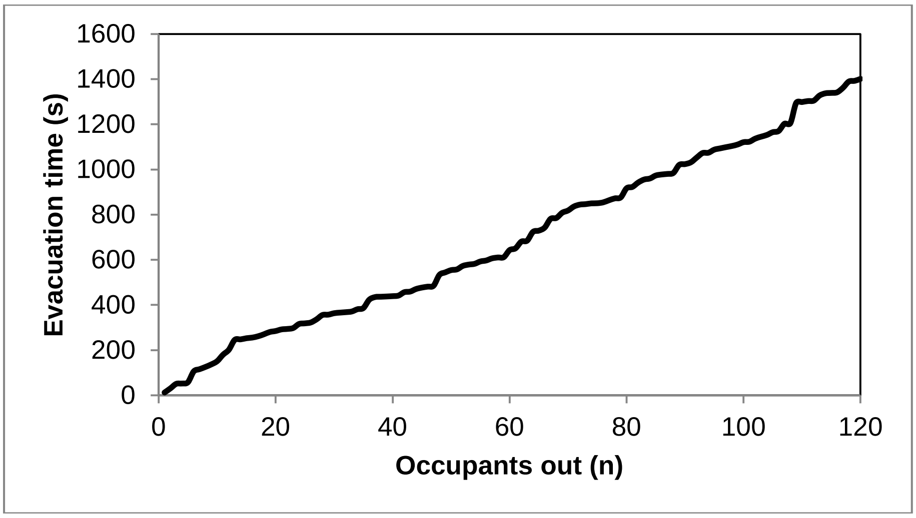

Fig. 7 Fictitious data representing one possible curve of evacuation times.

Table 1. Example of fictitious data representing one possible occupant-evacuation curve.

Occupants out	**Pseudo-random generated number**	**Evacuation time (s)**
1	12.41	12.41
2	18.18	30.59
3	20.22	50.81
4	8.43	59.24
…	…	…
120	…	1401.09

The steps of the evaluation method are applied as follows.

Step 1. Define the acceptance criteria.

This step deals with the definition of the five acceptable thresholds TR_{TET}, TR_{SD}, TR_{ERD}, TR_{EPC}, TR_{SC} about the impact of the number of runs on the predicted outcome of the evacuation model for the same evacuation scenario (see Equation 26-30). The number of consecutive runs $b=10$ for which the acceptance thresholds needs to be accomplished is also defined.

$$\mathrm{TR_{TET}} = 0.5\% \quad \text{[Equation 26]}$$
$$\mathrm{TR_{SD}} = 5\% \quad \text{[Equation 27]}$$
$$\mathrm{TR_{ERD}} = 1\% \quad \text{[Equation 28]}$$
$$\mathrm{TR_{EPC}} = 1\% \quad \text{[Equation 29]}$$
$$\mathrm{TR_{SC}} = 1\% \quad \text{[Equation 30]}$$

For instance, this means that the acceptance criteria are satisfied if $TET_{convj} < \mathrm{TR_{TET}}$ for 10 consecutive runs, $SD_{convj} < \mathrm{TR_{SD}}$ for 10 consecutive runs, etc.

It should be noted that the acceptance criteria have been selected with the only purpose of showing the procedure, i.e., they do not represent recommended values for use in real engineering analyses. Nevertheless, those criteria represent possible values in the context of fire safety engineering in relation to all types of uncertainty associated with modelling results. In fact, the authors argue that thresholds below 5 % would permit the assessment of the required safe egress time with a reasonable degree of accuracy. The definition of the criteria would be dependent on several factors, such as the type of evacuation scenario, data under consideration, the scope of the analysis, etc.

Step 2. Run a finite set of n runs of the same evacuation scenario

An arbitrary initial number of simulations of the same scenario is set to 35. *n=35* vectors of 120 dimensions $\overrightarrow{\boldsymbol{m}}_{\boldsymbol{ij}} = \left(m_{1\,1}, \dots, m_{i\,j}, \dots, m_{120\,35}\right)$ corresponding to the simulated evacuation times of each occupant i^{th} (for a total of 120 occupants) in each j^{th} run are obtained (for a total of 35 runs).

In the present example, the 35 fictitious curves have been generated using the method described at the beginning of Section 3. They result in the 35 curves showed in Fig. 8. The curves presented in Fig. 7 are representative of a set of repeated results of an evacuation model in the case of a hypothetical evacuation scenario for lognormal distribution of evacuation times [19]. It should be noted that the shape of the evacuation curves may be different than the example provided here (e.g. s-shaped occupant-evacuation curves). The method is based on convergence measures which are independent on the shape of the curves, thus it may be applicable for any type of curve.

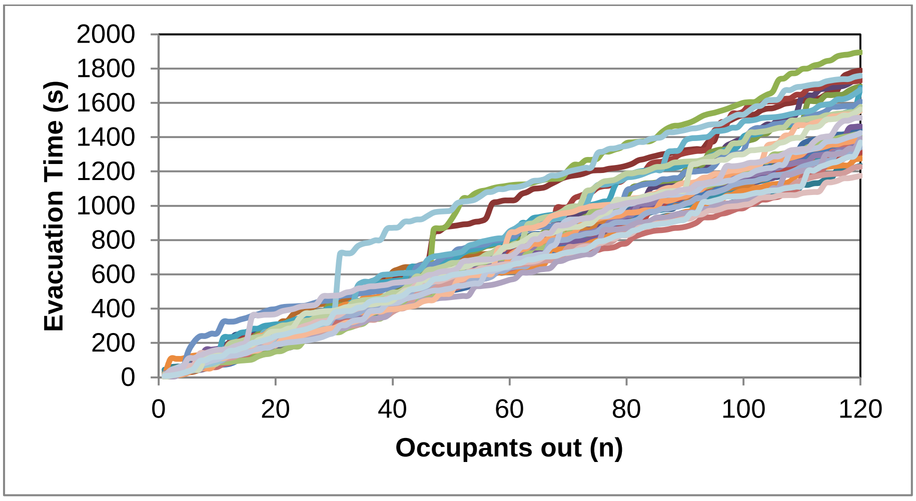

Fig. 8 Fictitious data representing 35 runs of the same hypothetical evacuation scenario.

The vector corresponding to the consecutive average curves $\overrightarrow{\mathbf{M}}$=($M_1$, ..., M_{35}) is also generated.

Step 3. Calculate the convergence measures

The convergence measures presented in the previous sections are calculated for all 35 runs, i.e., TET_{convj} , SD_{convj} , ERD_{convj} , EPC_{convj} , and SC_{convj} in accordance to Equation 11, Equation 13, Equation 15, Equation 17, and Equation 19, respectively. In this example a single value for *s* in Equation 19 has been used, namely *s=4*. Results are presented in Table 2.

Table 2. Results corresponding to 35 runs of the same evacuation scenario (expressed in %).

Run (n)	**TETconvj (%)**	**SDconvj (%)**	**ERDjconv (%)**	**EPCjconv (%)**	**SCjconv (%)**
1	/	/	/	/	/
2	12.163	/	/	/	/
3	3.852	8.294	10.800	18.674	10.337
4	3.369	0.886	1.555	5.446	0.439
5	4.691	16.110	1.059	4.962	1.238
6	2.024	3.639	2.399	3.022	0.176
7	0.405	7.698	0.368	0.052	0.045
8	2.054	2.709	1.321	2.658	0.115
9	1.323	1.400	1.403	1.086	0.365
10	1.238	0.875	0.055	1.900	0.165
11	0.582	3.723	0.158	0.088	0.024
12	1.139	0.358	0.169	0.187	0.011
13	0.626	2.473	0.122	0.138	0.008
14	0.820	0.772	0.081	0.068	0.004
15	1.901	11.480	1.761	2.673	0.392
16	0.140	3.194	1.883	2.505	0.396
17	0.758	0.312	0.235	0.969	0.042
18	0.861	0.779	0.134	1.393	0.064
19	0.403	1.859	0.166	1.248	0.062
20	0.453	1.400	0.254	1.355	0.062
21	0.234	2.124	0.578	0.696	0.003
22	0.420	1.159	0.028	0.092	0.010
23	0.569	0.080	0.443	0.980	0.032
24	0.298	1.473	0.514	0.893	0.029
25	0.393	0.819	0.170	0.206	0.012
26	0.663	1.576	0.206	0.210	0.002
27	0.267	1.297	0.138	1.033	0.016
28	0.198	1.480	0.084	0.955	0.006
29	0.666	2.348	0.666	1.631	0.095
30	0.166	1.442	0.958	1.043	0.079
31	0.129	1.481	0.031	0.231	0.020
32	0.652	2.721	0.323	0.359	0.012
33	0.184	1.193	0.382	0.646	0.013
34	0.075	1.441	0.044	0.075	0.001
35	0.207	0.956	0.181	0.424	0.012

Step 4 - 4bis. Compare the convergence measures with the acceptance criteria

Results for 35 runs are compared with the acceptance criteria defined in Step 1 (see also Equation 21-25). Table 3 shows the results of the tests in relation to the number of runs. When the box shows “FAILED”, it means that the test is failed. When the test is passed, the box is left blank. After 10 consecutive runs (given the acceptance criteria defined in Step 1), when the test is passed, the box shows “OK”, which means that the acceptance criteria have been met.

Table 3. Summary of the results of the tests in Step 4.

Run	TEST 1	TEST 2	TEST 3	TEST 4	TEST 5
1	/	/	/	/	/
2	FAILED	FAILED	/	/	/
3	FAILED	FAILED	FAILED	FAILED	FAILED
4	FAILED		FAILED	FAILED	
5	FAILED	FAILED	FAILED	FAILED	FAILED
6	FAILED		FAILED	FAILED	
7		FAILED			
8	FAILED		FAILED	FAILED	
9	FAILED		FAILED	FAILED	
10	FAILED			FAILED	
11	FAILED				
12	FAILED				
13	FAILED				
14	FAILED				
15	FAILED	FAILED	FAILED	FAILED	**OK**
16			FAILED	FAILED	
17	FAILED				
18	FAILED			FAILED	
19				FAILED	
20				FAILED	
21					
22					
23	FAILED				
24					
25		**OK**			
26	FAILED		**OK**		
27				FAILED	
28					
29	FAILED			FAILED	
30				FAILED	
31					
32	FAILED				
33					
34					
35					

In this example, Test 1 failed, Test 2 is passed after 25 runs, Test 3 is accomplished after 26 runs, Test 4 is failed, and Test 5 is accomplished after 15 runs. This means that our predicted curve meet the acceptance criteria with regards of the standard deviation of the total evacuation time, the Euclidean Relative Difference and the Secant Cosine. Nevertheless, there are two criteria that have not been met (Total Evacuation Time and Euclidean Projection Coefficient). It is therefore necessary to proceed with Step 5 by conducting additional runs.

Step 5. Simulate a set of additional simulations *m* so that the new set of runs for the comparison is *S=n+m.*

Another set of runs m=35 of the same scenario – corresponding to additional 35 occupant-evacuation time curves – are considered for a total of *S=n+m=35+35=70* runs. In this example, additional fictitious data are produced using the same method as the first 35 curves. A new set of *S=n+m* $\vec{\boldsymbol{S}}_{\boldsymbol{ij}}$ vectors $\vec{\boldsymbol{S}}_{\boldsymbol{ij}} = \left(S_{11}, \dots, S_{ij}, \dots, S_{q\,S}\right)$ corresponding to the average simulated evacuation times of each of the 120 occupant i^{th} in each of the 70 j^{th} run *S* are produced.

The evaluation method is now repeated for S=70 runs, starting from Step 3, called here Step 3.2.

Step 3.2 Calculate the convergence measures

The failing convergence measures are calculated for *S=70* runs, i.e., $TET_{convFIN}$, and $EPC_{convFIN}$ for our case study.

Step 4.2-4.2bis. Compare the convergence measures with the acceptance criteria
Results for *S=70* runs are compared again with the acceptance criteria defined in Step 1. Table 4 shows the results of the tests that were previously failing in relation to the number of runs.

Table 4. Results of Test 1 and Test 4 for 70 runs.

Run	TEST 1	TEST 4	Run	TEST 1	TEST 4	Run	TEST 1	TEST 4
1			24			47		
2	FAILED		25			48		
3	FAILED	FAILED	26	FAILED		49		
4	FAILED	FAILED	27		FAILED	50		
5	FAILED	FAILED	28			51	FAILED	
6	FAILED	FAILED	29	FAILED	FAILED	52		
7			30		FAILED	53		
8	FAILED	FAILED	31			54		
9	FAILED	FAILED	32	FAILED		55		
10	FAILED	FAILED	33			56		
11	FAILED		34			57		
12	FAILED		35			58		
13	FAILED		36			59		
14	FAILED		37			60		
15	FAILED	FAILED	38			61	**OK**	
16		FAILED	39			62		
17	FAILED		40	FAILED	**OK**	63		
18	FAILED	FAILED	41			64		
19		FAILED	42	FAILED		65		
20		FAILED	43			66		
21			44			67		
22			45			68		
23	FAILED		46			69		
						70		

Table 4 shows that Test 4 is accomplished after 40 runs. An example of the number of runs required to accomplish different criteria for $\mathrm{TR_{TET}}$ (where $TET_{convj} < \mathrm{TR_{TET}}$ for 10 consecutive runs) for the fictitious data-set under consideration is shown in Fig. 9. The grey dashed vertical line refers to the acceptance criteria $\mathrm{TR_{TET}}$ = 0.5% which has been selected for the analysis of the total evacuation time in Step 1. Test 1 is passed after 61 runs if the convergence criteria are $TET_{convj} < 0.5\,\%$ for 10 consecutive runs. This means that our predicted curve now meet all acceptance criteria.

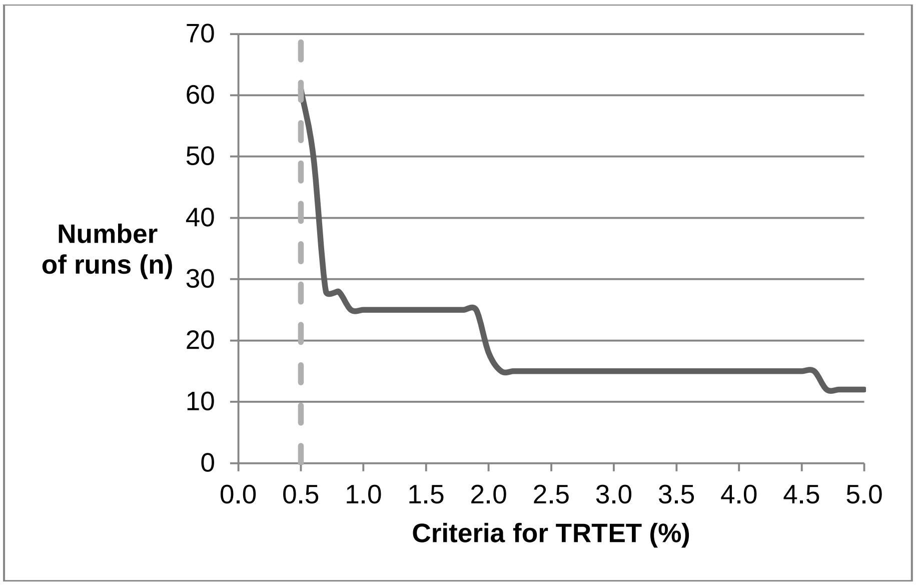

Fig. 9 Number of required runs in relation to different criteria for TRTET

4. DISCUSSION

The analysis of the trend of the convergence measures is useful to obtain general information on the type of data-set under consideration. For example, it is possible to assess behavioural uncertainty and therefore estimate the impact of the use of stochastic variables/distributions on evacuation model results.

An example from the data of the case study in Section 3.1 is presented in Fig. 10 where TET_{convj} and SD_{convj} are shown as well as Fig. 11 where ERD_{convj} , EPC_{convj} , and SC_{convj} are shown (convergence measures are calculated for a total of 140 consecutive average number of runs, i.e., 70 additional runs have been analysed).

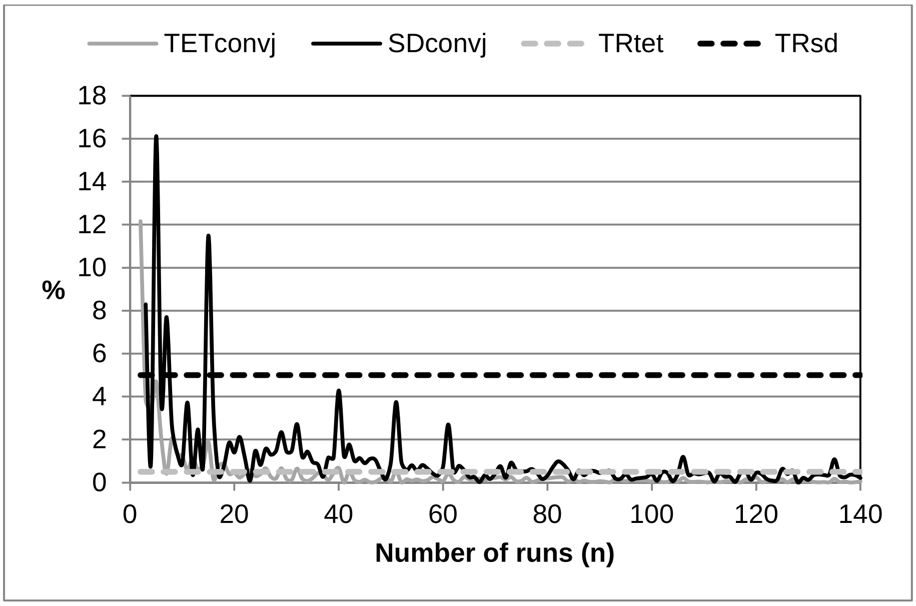

Fig. 10 TET_{convj} , SD_{convj} in relation to the consecutive average number of runs (expressed in %).

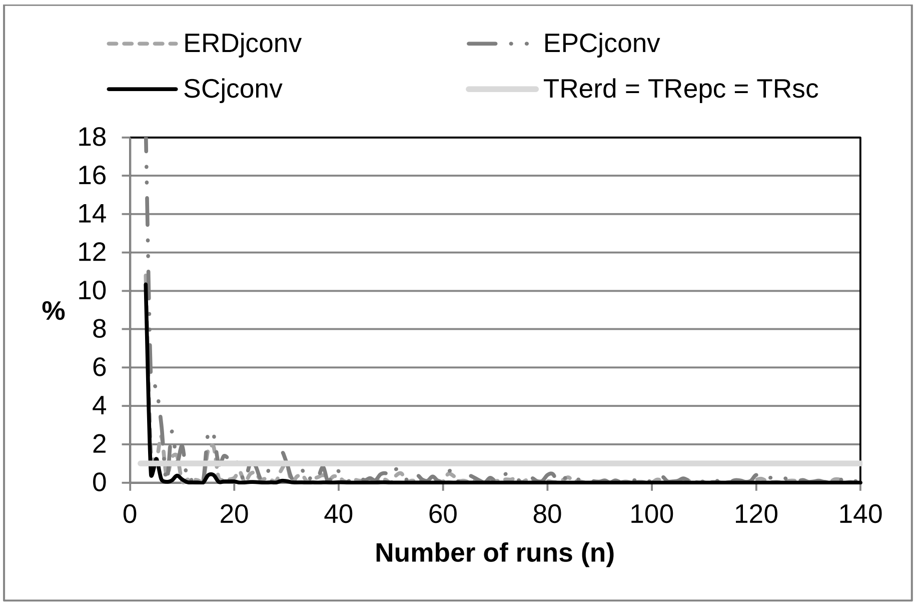

Fig. 11 ERD_{convj} , EPC_{convj} , and SC_{convj} *in relation to the consecutive average number of runs (expressed in %).*

Fig. 10 and Fig. 11 show that the standard deviation of the evacuation time SD_{convj} of the last occupant is the slowest converging variable. Together with TET_{convj}, those variables are useful to understand the variability of the total evacuation time in relation to the number of runs. An estimation of uncertainty (including behavioural uncertainty) associated with the total evacuation time is a key aspect of fire safety engineering analysis since it represents the RSET (Required Safe Egress Time) [19], the time needed by all occupants to perform a safe evacuation.

The analysis of the convergence of ERD_{convj} , EPC_{convj} , and SC_{convj} is also a significant contribution to the understanding of behavioural uncertainty, since it permits the analysis of the variability of the predicted occupant-evacuation time curves in relation to the number of runs. In the example provided here, the convergence measures are below 2.5% after 17 runs, thus permitting the estimation of the average occupant-evacuation time curve with an admitted 2.5% variability in a relatively small number of runs.

The simulation of additional 70 runs (for a total of 140 runs in Fig. 10 and Fig. 11) shows that, as expected, results continue to converge and the effect of behavioural uncertainty on average occupant evacuation time is progressively reduced. Nevertheless, if the acceptance criteria include the requirement of being below the thresholds for a sufficient number of consecutive runs (i.e. a critical number that the model user should select in relation to the scenario under consideration in order to verify the stability of the convergence), the simulation of additional runs does not provide any additional benefits to the modeller. The selection of the number of runs is optimised in relation to the pre-defined acceptance criteria and there is no need to simulate additional runs.

A statistical estimation of the uncertainty associated with the use of the convergence measures can be performed in relation to the number of runs. This includes the study of the uncertainty of the sample average total evacuation times and the sample standard deviations.

Assume that each total evacuation time in the vector $\overrightarrow{\boldsymbol{TET}}$ is a sum of random variables corresponding to the inter-temporal times between each occupant. Applying the central limit theorem, the series corresponding to the vector $\overrightarrow{\boldsymbol{TET}}$ consists of normally distributed values $TET_j \sim N(\mu, \sigma^2)$, where μ is the true mean value and σ^2 is the true variance. The sample variance is:

$$s^2 = \frac{\sum_{j=1}^{n}\left(TET_j - TET_{avj}\right)^2}{n-1} \qquad \text{[Equation 31]}$$

Where n is the number of runs. Applying Cochran's theorem, $s^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}$, a chi-squared distribution with n-1 degrees of freedom. Then, the variance of the sample variance, $Var(s^2)$, corresponds to:

$$Var(s^2) = Var\left(\frac{\sigma^2}{n-1}\chi^2_{n-1}\right) = \left(\frac{\sigma^2}{n-1}\right)^2 Var(\chi^2_{n-1}) =$$
$$= \left(\frac{\sigma^2}{n-1}\right)^2 2(n-1) = \frac{2\sigma^4}{n-1} \qquad \text{[Equation 32]}$$

The sample standard deviation *s* is distributed as a chi distribution with n-1 degrees of freedom, i.e., $s \sim \frac{\sigma}{\sqrt{n-1}}\chi_{n-1}$. Hence the variance of the standard deviations of the sample data corresponds to:

$$Var(s) = Var\left(\frac{\sigma}{\sqrt{n-1}}\chi_{n-1}\right) = \frac{\sigma^2}{n-1} Var(\chi_{n-1}) =$$
$$= \frac{\sigma^2}{n-1}\left[n-1-2\left\{\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right\}^2\right] \qquad \text{[Equation 33]}$$

Where $\Gamma(n)$ is a gamma function. Hence, it is possible to estimate the relative standard deviation (the relative difference between the use of sample standard deviations and the standard deviations corresponding to the true distribution):

$$relative\ Std(s) = \sqrt{\frac{\left[n-1-2\left\{\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right\}^2\right]}{n-1}} \qquad \text{[Equation 34]}$$

This information permits an estimation of the uncertainty associated with the use of the estimate standard deviations SD_j employed in the evaluation method in relation to the number of runs under consideration.

It is also possible to perform an estimation of the uncertainty associated with the use of the estimate variance of the sample data s^2 when calculating the average sample Total Evacuation Time TET_{avj}. In fact, the average of the Total Evacuation Time TET_{avj} is distributed as $\frac{s}{\sqrt{n}}t_{n-1} + \mu$, where t_{n-1} is a Student *t* random variable with *n-1* degrees of freedom.

Therefore the variance of the sample average TET_{avj} corresponds to:

$$Var(TET_{avj}) = \frac{s^2}{n}\frac{n-1}{n-3}$$ [Equation 35]

And the uncertainty of the TET_{avj} is:

$$SD(TET_{avj}) = s\sqrt{\frac{n-1}{n(n-3)}}$$ [Equation 36]

It is therefore possible to estimate the uncertainty associated with the number of runs given the use of the sample TET_{avj}.

To date, behavioural uncertainty is generally treated only in a qualitative manner (performing a qualitative evaluation of the number of runs to be simulated). It is argued that the present work would encourage evacuation model users to perform a quantitative treatment of this type of uncertainty given the simplicity of the method proposed.

The benefits obtained from the use of the method apply to design studies as well as model validation. The proposed method permits an estimation of the convergence of the simulated occupant-evacuation curve towards the average curve, thus increasing the understanding of the model predictions. This is reflected in a better understanding of the variability of RSET and the possible estimation of the margin of safety of a specific design in relation to behavioural uncertainty.

From a model validation perspective, to date, two antithetical approaches may be used to present model comparison with experimental data, namely 1) the use of the best model estimation for the occupant-evacuation time curve, or 2) the average occupant-evacuation time curve. The method presented in this section increases the usability of the second approach, since it allows a thorough quantitative understanding of the average curves produced by evacuation models. Future work based on the presented method is therefore a definition of an evacuation model validation protocol which uses the convergence measures to assess the differences between model predictions and experimental data by taking into account behavioural uncertainty.

A possible application of the method presented in this paper may be its use for the comparison of model predictions produced by different evacuation models. It would be in fact possible to quantify the impact of the stochastic variables and assumptions used by different evacuation models given the same evacuation scenario.

5. LIMITATIONS

A set of limitations of the proposed method can be identified both in terms of its assumptions as well as its applicability.

The first limitation of the method is that it uses the concepts of convergence in mean and the central limit theorem rather than a statistical estimation of the expected values. Hence, the choice of the requirement for the finite number of consecutive runs *b* for which the acceptable thresholds must not be crossed should be carefully evaluated by the modeller in relation to the data under consideration. This limitation is tempered by the simplicity of the proposed method, i.e., it can be applied by evacuation modellers to analyse behavioural uncertainty without a complex inferential statistical treatment of the data, which may require time and user expertise.

Another limitation of the method is associated with the assumptions that evacuation curves can be identical between model runs even in the case of different behaviours occurring, i.e., the arrival rates to the exits are the same but they refer to different occupants or different exits.

With regards of the method applicability, multiple data-sets of a single evacuation scenario are rarely available in the literature. This makes it difficult to study the impact of behavioural uncertainty on experimental data. Given the current stage of experimental evacuation research, the proposed method is mainly applicable to the analysis of behavioural uncertainty in simulation results. Once additional experimental data on individual scenarios will be available, researchers will be able to use the same concepts introduced in this paper for the analysis of behavioural uncertainty in experimental data.

Without multiple experimental data, a single experiment often represents the only reference on that specific evacuation scenario, but it is not clear whether it represents average behaviour or it is a tail of the curve. In fact, the assessment of experimental and evacuation model results may also include the analysis of the tails of the distribution rather than the analysis of the peaks (i.e. average values). Nevertheless, the authors argue that the study of the average model predictions together with the variability of results around the average is deemed to be a useful method to analyse *behavioural uncertainty*. The research community of human behaviour in fire is aware of the lack of experimental data and the need to fill this gap with data collection efforts [2]. In recent years, significant data collection efforts have been carried out (e.g. several projects were performed for different aspects/conditions of the evacuation process using several tool to aid the collection and quality of data [20]). Therefore, it can be argued that considering a long-term perspective, it will be possible to assess behavioural uncertainty also for experimental data (thus making the method proposed in the paper applicable also for that issue).

The method is presented using a case study based on pseudo-random generated numbers. Future work can be based on the analysis of the results of an evacuation model from a real world case study.

6. CONCLUSIONS

This paper introduces a simple methodology to analyse the variability of evacuation model predictions of an individual evacuation scenario in relation to the number of test cases. The method allows increasing the understanding of the uncertainty in evacuation model results, which depend on the stochastic nature of human behaviour, here called *behavioural uncertainty*.

This paper presents a step forward towards a more accurate interpretation of evacuation model results, because it introduces a novel method to perform quantitative estimation of the convergence of evacuation model predictions. In fact, the use of functional analysis operators permitted the analysis of the entire occupant-evacuation time curves, rather than a study focused on the average and standard deviations of evacuation times only.

This paper represents the starting point for a better quantitative interpretation of evacuation behaviour. In fact, future applications are associated with the interpretation of behavioural experimental data as well as the development of a standard model validation protocol.

ACKNOWLEDGMENTS

The authors wish to acknowledge Erica Kuligowski, Craig Weinschenk, Thomas Cleary, Anthony Hamins, and Rita Fahy for a comprehensive review of the paper before publication. Enrico Ronchi thanks Daniel Nilsson for the fruitful discussions which have led to the idea of this paper. Finally, the authors wish to thank Blaza Toman for her valuable suggestions for the assessment of the uncertainty associated with the proposed criteria.

REFERENCES

[1] Hamins A, McGrattan K (2007) Verification and Validation of Selected Fire Models for Nuclear Power Plant Applications. Final Report NUREG-1824. U.S. Nuclear Regulatory Commission.
[2] Averill JD (2011) Five Grand Challenges in Pedestrian and Evacuation Dynamics. In: Proceedings of the 5th International Pedestrian and Evacuation Dynamics Conference pp 1-11.
[3] Kuligowski E (2011). Predicting Human Behavior During Fires. Fire Technology 49: 101–120.
[4] McGrattan K, Hostikka S, Floyd J, Baum H, Rehm R, Mell W, McDermott R (2007) Fire Dynamics Simulator (Version 5) Technical Reference Guide. NIST Special Publication 1018-5.
[5] Jones WW, Peacock RD, Forney G, Reneke PA (2009) CFAST – Consolidated Model of Fire Growth and Smoke Transport (Version 6). Technical Reference Guide. NIST Special Publication 1026.
[6] Kuligowski ED, Peacock RD, Hoskins BL (2010) A Review of Building Evacuation Models. 2nd edition. National Institute of Standards and Technology pp 36. Fire Research Division. Technical Note 1680.
[7] Lord J, Meacham B, Moore A, Fahy R, Proulx G (2005) Guide for evaluating the predictive capabilities of computer egress models. National Institute of Standards and Technology. NIST GCR 06-886 pp 107.
[8] Ronchi E, Kinsey M (2011) Evacuation models of the future. Insights from an online survey on user‘s experiences and needs. In: Proceedings of the Advanced Research Workshop: Evacuation and Human Behaviour in Emergency Situations, pp 145-155.
[9] Capote JA, Alvear D, Abreu O, Cuesta A, Alonso V (2012). A Stochastic Approach for Simulating Human Behaviour During Evacuation Process in Passenger Trains. Fire Technology 48: 911–925.

[10] Ronchi E, Nilsson D, Gwynne SMV (2012). Modelling the Impact of Emergency Exit Signs in Tunnels. Fire Technology 48: 961–988.
[11] Galea ER, Filippidis, L, Deere S, Brown R, Nicholls, Hifi Y, Besnard B (2012) IMO Information paper – The SAFEGUARD Validation Data-Set and Recommendations to IMO to Update MSC/Circ. 1238. Safeguard Passenger Evacuation Seminar, 30 Nov 2012, London UK.
[12] Peacock RD, Reneke PA, Davis WD, Jones WW (1999) Quantifying Fire Model Evaluation Using Functional Analysis, Fire Saf J 33: 167-184. doi: 10.1016/S0379-7112(99)00029-6
[13] International Standards Organization (2008) ISO/FDIS 16730. Fire Safety Engineering – Assessment, verification and validation of calculation methods. ISO/TC92/SC4.
[14] Galea ER, Deere S, Brown R, Filippidis L (2013) An Evacuation Validation Data Set for Large Passenger Ships. In Proceedings of the Pedestrian and Evacuation Dynamics 2012 Conference (In Press).
[15] Prugovecki E (1981) Quantum mechanics in Hilbert space, 2d ed. ed, Pure and applied mathematics, a series of monographs and textbooks. Academic Press, New York.
[16] Bains R (1992). Applied functional analysis and variational methods in engineering. Engineering Analysis with Boundary Elements 9: 365–366.
[17] Purser DA, Bensilum M (2001) Quantification of behaviour for engineering design standards and escape time calculations. Safety Science 38(2):157-182.
[18] Wichmann BA, Hill ID (1982) Algorithm AS 183: An Efficient and Portable Pseudo-Random Number Generator, Appl Statistics 31:188-190.
[19] Ronchi E, Nilsson D (2013) Modelling total evacuation strategies for high-rise buildings. Building Simulation. In Press. doi: 10.1007/s12273-013-0132-9
[19] Gwynne SMV, Kuligowski ED, Nilsson D, (2012) Representing evacuation behavior in engineering terms. Journal of Fire Protection Engineering 22(2):133-150
[20] Gwynne SMV(2013) Improving the Collection and Use of Human Egress Data. Fire Technology 41(1):83-99.