



DEPARTMENT OF MECHANICAL, AUTOMOTIVE
AND AERONAUTICAL ENGINEERING

Uncertainty Quantification in Microscopic Crowd Simulation Based on Polynomial Chaos Expansions

Quantifizierung von Unsicherheiten in der mikroskopischen
Personenstromsimulation basierend auf der polynomiellen Chaosentwicklung

Author:
Simon RAHN

Supervisors:
Prof. Dr.-Ing. Katina WARENDORF
Prof. Dr. Rainer FISCHER

*A thesis submitted in partial fulfilment of the requirements
for the degree of
M. Sc. in Computational Engineering*

29 July 2020

Date of Birth: 1 June 1995 in Munich | Matriculation Number: 92051514 | Study Group: 3W

Abstract

Microscopic crowd simulation, for example performed with the optimal steps model, can support the planning process of mass events if the model parameters are chosen correctly. The exact values of these parameters are often unknown, so the input suffers from uncertainties that can affect the output. To assure that the right conclusions are drawn from a simulation, modellers can employ forward uncertainty quantification methods. However, applying such methods to the optimal steps model is rather uncommon. Therefore, this thesis addresses the question which approach is suitable for estimating the uncertainty and sensitivity of the model output. It focuses on polynomial chaos expansions, which allow to derive statistical moments and Sobol' sensitivity indices for uncertainty and sensitivity analyses, respectively. The polynomial chaos expansion combined with the point collocation method and the pseudo-spectral approach is applied to a corridor scenario and the results are compared to Monte Carlo simulations. Thus, it can be shown that the point collocation method is efficient because it yields accurate results while the computational effort is low. Based on that, the method is transferred to a large scale scenario, a parade through a city center. The outcomes are interpreted to demonstrate that they are relevant to reality.

Zusammenfassung

Die mikroskopische Personenstromsimulation, zum Beispiel durchgeführt mit dem Optimal Steps Model, kann den Planungsprozess von Großveranstaltungen unterstützen, sofern die Modellparameter richtig gewählt werden. Die exakten Werte dieser Parameter sind allerdings oft unbekannt, weshalb die Eingangsgrößen mit Unsicherheiten behaftet sind und die Modellantwort beeinflussen können. Um sicherzustellen, dass die richtigen Schlüsse aus der Simulation gezogen werden, können Modellierer auf Methoden zur Quantifizierung von Unsicherheit zurückgreifen. Bisher werden solche Methoden auf das Optimal Steps Model jedoch eher selten angewandt. Deshalb geht diese Arbeit der Frage nach, welche Methode geeignet ist, um die Unsicherheit und Sensitivität des Modellergebnisses abzuschätzen. Der Fokus liegt dabei auf der polynomiellen Chaonentwicklung, mit deren Hilfe statistische Momente beziehungsweise Sobol' Sensitivitätsindizes abgeleitet werden. Die polynomielle Chaonentwicklung in Verbindung mit Kollokation nach kleinsten Quadraten beziehungsweise dem pseudo-spektralen Ansatz wird auf ein Korridor Szenario angewandt und mit Monte Carlo Simulationen verglichen. Auf diese Weise wird dargelegt, dass der Kollokationsansatz effizient ist, da genaue Ergebnisse bei geringem Rechenaufwand erzielt werden. Auf diesen Erkenntnissen aufbauend wird die Methode auf ein großes Szenario, eine Parade durch eine Innenstadt, übertragen. Die Resultate werden interpretiert, um deren Relevanz für die Realität aufzuzeigen.

Danksagung

An dieser Stelle bedanke ich mich bei allen, die mich bei der Anfertigung dieser Arbeit unterstützt haben:

Frau Prof. Dr.-Ing. Katina Warendorf, die die Betreuung von Seiten der Fakultät für Maschinenbau, Fahrzeugtechnik und Flugzeugtechnik übernommen hat. Durch das Herstellen des Kontaktes zur Pedestrian Dynamics Forschungsgruppe von Frau Prof. Dr. Gerta Köster an der Fakultät für Informatik und Mathematik wurde mir diese Arbeit erst möglich.

Ich bedanke mich bei Herrn Prof. Dr. Rainer Fischer, der mich fachlich betreut hat und jederzeit ein offenes Ohr für meine Fragen hatte. Ebenso gilt mein Dank allen Mitgliedern der Forschungsgruppe, insbesondere Marion Gödel für die wertvollen Korrekturanmerkungen und die vielen anregenden Diskussionen.

Darüber hinaus danke ich meinem Bruder für das Korrekturlesen und meinen Eltern, deren Unterstützung mir durch das ganze Studium geholfen hat.

Simon Rahn

Contents

1	Introduction	1
2	Theoretical Foundations and State of the Art	5
2.1	Modelling Pedestrian Dynamics	5
2.1.1	Microscopic Crowd Simulation	6
2.1.2	Optimal Steps Model	7
2.2	Uncertainty Quantification Methods	9
2.2.1	Uncertainty Analysis	9
2.2.2	Sensitivity Analysis	18
2.3	Uncertainty Quantification in Microscopic Crowd Simulation	22
3	Application of Uncertainty Quantification Methods	25
3.1	Adaptation of Uncertainty Quantification Methods to the Model	25
3.1.1	Object of Investigation: The Corridor Scenario	26
3.1.2	Definition of Configurations for the Methods	28
3.1.3	Verification of Methods	33
3.1.4	Rating of Methods	38
3.2	Overview of the Implemented Program Code	39
3.2.1	Embedding the Chaospy Library	41
3.2.2	Verification of the Program Code	42
4	Results and Discussion	47
4.1	Narrow Corridor Scenario	47
4.1.1	Uncertainty Analysis	48
4.1.2	Sensitivity Analysis	51
4.1.3	Discussion	55
4.2	Corridor Scenarios with Similar Topographies	56
4.2.1	Outline of Uncertainty and Sensitivity Analysis	57
4.2.2	Discussion	57
4.3	Real World Scenario: Parade Through a City Centre	58
4.3.1	Adaptations to the Method	58

4.3.2	Object of Investigation	58
4.3.3	Uncertainty Analysis	63
4.3.4	Sensitivity Analysis	64
4.3.5	Plausibility Checks	66
4.3.6	Discussion and Relevance to Reality	67
5	Conclusions and Outlook	69
	Bibliography	73
A	Appendix	I
A.1	Adapted Parameters in the Scenario Files	II
A.2	Complementary Figures for Evaluation of the Method	III
A.3	Detailed Uncertainty Quantification for Corridor Scenario with Sim- ilar Topographies	V
A.3.1	Uncertainty Analysis	V
A.3.2	Sensitivity Analysis	VII

List of Figures

2.1	Static floor field of the optimal steps model	7
2.2	Monte Carlo method with different sampling techniques	13
3.1	Corridor scenario: Topography	27
3.2	Voronoi cells	28
3.3	Convergence of the Monte Carlo method	29
3.4	Computational effort for point collocation method and pseudo-spectral approach estimated by the number of sample points	31
3.5	Qualitative comparison of output distributions for the mean Voronoi density approximated by PCE with a varying sample size	34
3.6	Qualitative comparison of output distributions for the mean Voronoi density approximated by PCE with a varying polynomial order	35
3.7	Quantitative comparison of different configurations of the UQ methods based on their errors	37
3.8	Simplified flowchart of the code	40
3.9	Convergence of sensitivity indices obtained with Jansen's approach for the Ishigami function	44
4.1	Corridor scenario: Transitions between different phases	48
4.2	Corridor scenario: Uncertainty of the Voronoi density	50
4.3	Corridor scenario: Sensitivity of the Voronoi density	54
4.4	Corridor scenario: Effect of interactions on the Voronoi density	55
4.5	Corridor scenario with different topographies: Narrow and wide versions	57
4.6	OPMoPS scenario: Topography	60
4.7	OPMoPS scenario: Normal distributions for boundary values of the uncertain parameter standard deviation of free-flow speed	62
4.8	OPMoPS Scenario: Uncertainty of the length of the parade	63
4.9	OPMoPS scenario: First few hundred frames after start of the simulation	64
4.10	OPMoPS scenario: Sensitivity indices for the length of the parade	65

A.1	Qualitative comparison of output distributions for the evacuation time approximated by PCE with a varying sample size	III
A.2	Qualitative comparison of output distributions for the evacuation time approximated by PCE with a varying polynomial degree	III
A.3	Quantitative comparison of four pre-selected configurations of the UQ methods based on their errors	IV
A.4	Corridor scenario with different topographies: Uncertainty of the Voronoi density	VI
A.5	Corridor scenario with different topographies: Sensitivity of the Voronoi density	VIII

List of Tables

2.1	Relation between common distributions of random variables and type of polynomials to be chosen for PCE	14
3.1	Corridor scenario: The uncertainty in the input parameters is described by uniform distributions.	27
3.2	Parameter set-up for the point collocation (PC) method and the pseudo-spectral (PS) approach	33
3.3	First order and total sensitivity indices for the Ishigami model function	44
4.1	Corridor scenario: Uncertainty of the mean Voronoi density and evacuation time	49
4.2	Corridor scenario: Sensitivity of the mean Voronoi density and evacuation time	52
4.3	Uncertain input parameters for the OPMoPS scenario.	61
A.1	Changes to the default values of the scenario parameters	II
A.2	Corridor scenario with different topographies: Uncertainty of the evacuation time	V
A.3	Corridor scenario with different topographies: Sensitivity of the evacuation time	VII

Acronyms

MAPE	Mean Absolute Percentage Error
MC	Monte Carlo
OPMoPS	Organised Pedestrian Movement in Public Spaces
OSM	Optimal Steps Model
PC	Point Collocation
PCE	Polynomial Chaos Expansion
PS	Pseudo-Spectral
QoI	Quantity of Interest
RMSE	Root Mean Square Error
SD	Standard Deviation
UQ	Uncertainty Quantification

1 Introduction

On 24 July 2010, 21 people died and more than 500 were injured during a stampede at the Love Parade in Duisburg, Germany. The Love Parade was a famous music festival with more than a million expected visitors. Similar crowd disasters occur each year in different parts of the world and they often shock victims and the public as well. How can such tragedies happen, despite regulations, planning guidelines and software tools that have been developed to prevent hotspot areas in advance? Helbing and Mukerji [8] investigated the incident in Duisburg from a scientific point of view and concluded that it arose from different unfortunate circumstances. Besides other proposals to avoid crowd crushes, the researchers suggest that “computer simulations with state-of-the art pedestrian software can be useful [to determine crucial areas], but model parameters must be carefully chosen” [8, p. 26].

Various models have been developed to simulate pedestrian flow on the macroscopic or microscopic scale. They are distributed either as open-source code, for example in the framework Vadere [38], or as commercial software. In either case, all of these models are typically subject to uncertainties, independent of their underlying concept, scale or licence. These uncertainties can propagate through the model and might affect the output significantly without being recognised. Thus, the user could wrongly deem simulations to be reliable and draw misleading conclusions. In the worst case, safety concepts for mass events or escape routes inside buildings could be planned based on false assumptions. To avoid such a situation, modellers should pay heed to uncertainties in their simulations.

Uncertainty quantification in pedestrian dynamics is a rather new development. In recent years, a few researchers presented different approaches to analyse the uncertainty and sensitivity of microscopic crowd simulations. For example, Monte Carlo or polynomial chaos expansions are applied in [1, 30] to determine the uncertainty and sensitivity of pedestrian models such as the optimal steps model. While the optimal steps model is widely adopted [15], none of these uncertainty quantification methods has been established as common practice in pedestrian dynamics yet. Moreover, the analysed scenarios are usually small examples and can be transferred to reality only to a limited extent. Practitioners would benefit from simulations that approximate actual scenarios more closely because they were able to identify crucial

areas or influencing factors on a more detailed level. For example, authorities could estimate better which measures are useful to ensure the safety of a parade through a city centre. The scenario for a small parade, not to be associated with a mass event as the Love Parade, is suitable as subject for a first step to apply uncertainty quantification methods to large scale scenarios because it has relatively few individual influencing factors and unknowns.

Hence two questions arise that I focus on in this thesis: Which method is appropriate to carry out forward uncertainty quantification in microscopic crowd simulations, for example performed with the optimal steps model? Once the right method is found and applied to the parade scenario, one can study the uncertainty quantification results, which leads to the second issue. How do uncertain input parameters affect the output of the model and how does the importance of those observed parameters change over time?

I adopt the following methodology to resolve these questions. The first research question is answered by comparing the results of different uncertainty quantification methods to a reference solution obtained with Monte Carlo simulations. The methods under consideration are the point collocation method and the pseudo-spectral approach. Both of them are based on the polynomial chaos expansion, which approximates the probability density functions of the model response. Measures of uncertainty and sensitivity can be derived from these distribution functions. I apply the two methods with different configurations to a small but prominent example, the corridor scenario implemented in the optimal steps model in Vadere. The simulations are evaluated in terms of accuracy and computational effort. On the basis of this rating, I select one method that is suited for further analyses.

The second research question will be satisfied by applying the chosen uncertainty quantification method to a parade scenario with two uncertain input parameters. Again, the optimal steps model is employed. In this manner, I determine the uncertainty of the output. The sensitivity analysis provides information about the importance of each parameter. It employs Sobol' indices calculated from the decomposed polynomial chaos expansion.

The thesis is structured as follows. At first, I give an overview of models for microscopic crowd simulation and uncertainty quantification methods. In this context, the optimal steps model and uncertainty quantification based on polynomial chaos expansions are presented in detail. The subsequent section connects these two fields of research, as I present the state of the art of uncertainty quantification in microscopic crowd simulation. Chapter 3 is dedicated to the adaptation of two uncertainty quantification methods to the corridor scenario. The methods under consideration are the point collocation method and the pseudo-spectral approach as part of the polynomial chaos expansion. The aim is to find a suitable configuration

for both methods and to select the one that performs best. In Chapter 4, I apply the chosen method to different scenarios. At first, the corridor scenario serves as subject of research to get acquainted with interpreting the results of the uncertainty quantification. The next step focuses on the main purpose of this thesis: I apply the polynomial chaos expansion to a large scale scenario, a parade through a city centre, to analyse the uncertainty and sensitivity of the model. Finally, these outcomes and their relevance to the real world are discussed.

2 Theoretical Foundations and State of the Art

Modelling pedestrian dynamics is becoming more and more important to engineers, safety engineers in particular. They can build on many different pedestrian locomotion models when analysing any aspect related to microscopic crowd behaviour. For instance, planning exit routes for mass events can be much easier with computer simulations than with experimental data. As interdisciplinary as this field of research is, as many different models have been developed to simulate crowd behaviour. The optimal steps model is one example, which is subject to this thesis. Therefore, its principle idea will be introduced in this chapter.

The advantage of computer models over controlled experiments is that they can be adapted to most problems simply by changing the settings and input parameters. Albeit these parameters are defined carefully, they are subject to aleatory and epistemic uncertainties [6]. These can influence the model output to a significant extent and therefore they should be accounted for, otherwise the user might draw false conclusions from the result. Uncertainty quantification (UQ) methods target this issue by estimating the uncertainty and sensitivity of the model to input parameters. Because there are various techniques that are suitable for different purposes, the second part of this chapter provides an overview of approaches relevant to the context of this thesis.

2.1 Modelling Pedestrian Dynamics

Modelling pedestrian dynamics is a method to analyse human locomotion and to establish a better understanding thereof. Most of these studies investigate crowd behaviour because computer models support planning processes in different fields of application where aggregation of pedestrians is crucial: Designing safety plans for buildings, developing crowd management strategies for mass events, or optimizing traffic flow are a few examples.

Such situations are analysed in the scope of controlled experiments, e.g. in [45].

However, the main disadvantages of controlled experiments is that they are time consuming, they require organisational effort and a lot of participants. These observations under laboratory conditions and in particular field studies depend on many different influencing factors [40]. For example, the crowd might behave differently if an experiment is repeated several times or if the participants know each other. Therefore, the measurements can vary to a considerable degree, so it is not always possible to generalise the findings to any arbitrary situation.

To solve these problems, computer models have been developed. As computing capacity rises, simulations can become computationally expensive, which allows to increase the level of detail. Kleinmeier et al. give an overview of various existing models in [15], and categorise them by their scale as follows: Macroscopic models represent an accumulation of pedestrians, i.e. a crowd, as continuum, whereas microscopic models consider all individuals and their specific attributes, e.g. desired speed, separately. This thesis focuses on a microscopic model, the optimal steps model. Application and objective of microscopic crowd simulations are presented in the following section.

2.1.1 Microscopic Crowd Simulation

The microscopic approach entails that each pedestrian's walking behaviour is simulated, which is a complex task. This issue has been addressed in many different ways and according to [15] it has lead to the development of models based on

- cellular automata [4],
- ordinary differential equations [7, 9],
- cognitive heuristics [24, 41],
- optimizing a utility, e.g. the optimal step model [25, 28, 29, 30].

Depending on the application, one approach is superior to the others. For example, in cellular automata, virtual pedestrians, often referred to as agents, move from cell to cell on a fixed grid. Rules determine the next step depending on the surrounding fields. This procedure can approximate human behaviour well, and most notably, it is fast. However, agents can neither move in arbitrary directions nor can such models represent dense crowds because compression is impossible [25].

In contrast to that, social force models are continuous in time and space. As the name suggests, they employ forces to determine the agents' movement. That is, obstacles and other agents act repulsive while the target acts attractive on the agent under consideration. This model suffers from effects such as inertia, which can lead to oscillating patterns of motion and colliding agents [16, 25].

Seitz and Köster combine the advantages of these two approaches, namely rule based, stepwise movement inspired by cellular automata and a spatially continuous

domain borrowed from social force models, and develop the optimal steps model (OSM) in [25]. An improved version thereof is implemented in Vadere, an open-source simulation framework for microscopic crowd simulation [15, 29]. Because the OSM is subject to this thesis, it is explained in detail in the next section.

2.1.2 Optimal Steps Model

The following description of the OSM refers to the implemented version in Vadere [15, 39]. Agent l walks stepwise from its source towards its target. A Nelder-Mead simplex algorithm optimizes a utility function to find the position for each step [29]. Figuratively, the utility represents a potential field, also known as floor field, in which agent l moves. A higher potential value expresses a greater geodesic distance from the target. In contrast to euclidean distances, geodesic distances include paths around obstacles and can be obtained by solving the eikonal equation:

$$\|\nabla u(x)\| f(x) = 1 \quad \text{for } x \in \Omega \subset \mathbb{R}^2 \quad (2.1)$$

$$u(x) = 0 \quad \text{if } x \in \Gamma \quad (2.2)$$

Sethian’s Fast Marching Method solves this non-linear partial differential equation [26, 27]. Figure 2.1 visualises the potential field as wave front, which starts in the target region Γ and propagates through the whole area Ω with speed f . Alternatively, one can consider the floor field as hilly terrain represented by contour lines. The agent takes always the steepest path starting from a mountain top or higher elevation down to a valley.

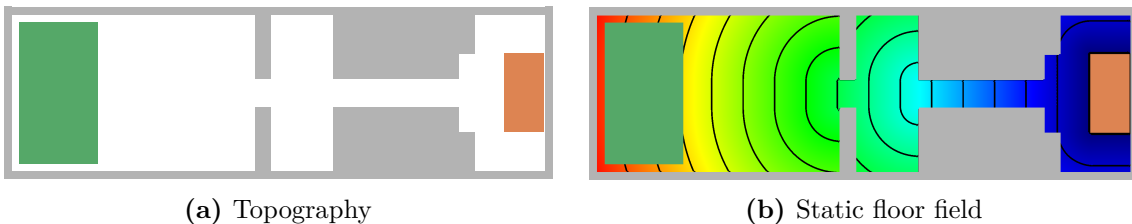


Figure 2.1: Topography (a) contains a source (green), obstacles (gray) and a target (orange). The corresponding static floor field (b) is represented by isolines, which start at the target and flow around obstacles. Low potentials are blue, whereas high potentials are red.

From a mathematical point of view, the potential field P_l is a sum of a globally defined potential P_t that depends on the distance to the target and is embodied in the solution of the eikonal equation, and local contributions to the potential inflicted by surrounding pedestrians $P_{p,l,i}$ or obstacles $P_{o,j}$. The full potential for agent l is

defined by [38]:

$$P_l(x) = P_t(x) + \sum_{i=1, i \neq l}^n P_{p,l,i}(x) + \max_{j \in \{1, \dots, m\}} P_{o,j}(x) \quad (2.3)$$

In this manner, the model takes into account that pedestrians consciously or subconsciously try to keep a certain distance from obstacles or other persons. The latter is expressed by the idea of (inter-) personal space, mathematically described in [15, 29, 31]. The potential is recalculated for each time step, which can be quite computationally expensive for large scale scenarios. For that reason, it is necessary to consider the static instead of the dynamic floor field in some cases. Then the term $P_t(x)$ in Equation 2.3 incorporates the global contributions by the target and the global penalty due to obstacles but no global effects caused by accumulated agents. Therefore, $P_t(x)$ must be calculated only once.

Agent l seeks after the lowest potential P_l within its reach. Therefore, utility P_l is optimized within a circle with radius r_l , which represents the maximum step length. r_l follows a truncated normal distribution \mathcal{N}_{tr} around the agent's mean step length s_l with variance σ_r^2 . The distribution ranges in $[-\sigma_r M, \sigma_r M]$:

$$s_l = \beta_0 + \beta_1 \cdot v_l = 0.462 \text{ m} + 0.235 \text{ s} \cdot v_l \quad (2.4)$$

$$r_l \sim \mathcal{N}_{tr}(s_l, \sigma_r^2; -\sigma_r M, \sigma_r M) \quad (2.5)$$

Constants β_0 , β_1 , σ_r and M are determined empirically in [25] for walking behaviour. Hence, the maximum step length r_l is correlated with each agent's desired velocity v_l , which is termed free-flow speed in the context of the OSM:

$$v_l \sim \mathcal{N}_{tr}(\mu_v, \sigma_v^2; v_{min}, v_{max}) \quad (2.6)$$

The free-flow speed of all agents follows a truncated normal distribution with mean μ_v and variance σ_v^2 of all agents' velocities. The random variable is bounded from below by v_{min} and from above by v_{max} to avoid unnaturally low or high speeds [25]. These values must be calibrated and they depend on the situation the user wants to simulate.

In the best case, empirical data that matches the model is available to determine parameters μ_l , σ_l , v_{min} and v_{max} . Only a few such experiments have been conducted, though. For example, Zhang et al. perform experiments in [45] to analyse pedestrian flow through corridors, bottlenecks and T junctions. For many other scenarios one has to rely on more general measurements, such as field studies. Weidmann gives an overview of previously performed observations in [40]. Based on that, he describes pedestrians' desired velocity by means of a Gaussian distribution. Depending on

the case under consideration, for example, if the modelled group is homogeneous or heterogeneous, the values for mean and standard deviation can vary. Therefore, this and other studies can be generalised to different scenarios only to some extent. As a last resort, if no data is available, one must estimate the parameters. Nevertheless, this will introduce uncertainties that have to be dealt with.

2.2 Uncertainty Quantification Methods

Before introducing the purpose of uncertainty quantification methods, I want to define the term uncertainty. There are several ways to specify uncertainty. I use the following definition, which is a widely adopted distinction and differentiates into two types [18]:

- a) **Aleatory uncertainty**, also known as variability, comprises uncertainty that is caused by the randomness of the system. As such, it is a property of the system, and it could be reduced theoretically, but only by changing the system itself [18]. For example, let a crowd of pedestrians be the subject to a modelling problem. Reducing aleatory uncertainty could imply that all pedestrians are forced to move at the same speed in a defined direction instead of allowing them to promenade around as each individual desires. To capture this systemic randomness, aleatory uncertainty is often expressed by a probability distribution (e.g. see Eq. 2.6).
- b) **Epistemic uncertainty** results from lack of knowledge. Hence it is a property of the observer and can be reduced, for example, by collecting more data [18]. Instead of estimating the value for a certain model parameter based on common sense, the modeller chooses in accordance with experimental data. However, because even the experimental information is usually incomplete, uncertainty will remain to some extent. The modeller should account for that, since any uncertainty might affect the predictive capabilities of the model.

There are several techniques to measure epistemic uncertainties, i.e. uncertainty quantification (UQ) methods. One class of methods aims at propagating uncertainties through the model and evaluating the likelihood of a specific outcome, known as uncertainty analysis. The interplay of input and output is determined by means of sensitivity analyses [12]. Some of these techniques for uncertainty and sensitivity analysis are explained in the following sections.

2.2.1 Uncertainty Analysis

Uncertainty analyses can help to evaluate the predictive capability of models and in this manner support decision-making processes. The question is: What is the

likelihood of a certain model output Y if input parameter X is not fixed but varies to some extent? The answer to this question is found by propagating uncertainties in the input parameters through the model f , expressed as

$$Y = f(X), \tag{2.7}$$

to determine prediction intervals, mean and higher statistical moments, or even the full probability density function of model response Y , also termed quantity of interest (QoI) [32].

Obviously, this requires that the researcher knows the uncertainty of input parameter X . A common approach to model uncertainty is to treat X as random variable with a certain distribution and move from the deterministic model to a stochastic one [42]. Depending on the application, the probability density function of X is obtained either experimentally, or, if that is not possible, Bayesian approaches are employed to find the input distribution by means of inverse UQ methods [32].

If the probability density function of the input is known, one can use intrusive or non-intrusive approaches to propagate the uncertainty through the model. Intrusive methods require access to the underlying forward model to determine the statistics of Y , whereas non-intrusive methods consider f as black box. That is, non-intrusive methods require only input X to the model and its response Y but no information about f [3, 11, 35].

Since the terminology and categorisation of techniques to propagate uncertainties through a model differs among research communities, I distinguish the approaches according to [32] as follows:

- **Direct evaluation:** For linearly parametrised models the uncertainty of the output is computed explicitly [32]. Xiu uses the term “moment equations” [43, p. 4]. This implies that statistical moments are calculated directly from the stochastic equations of the model. Unfortunately, one often has to deal with the closure problem because deriving statistical moments usually necessitates knowledge about properties of higher moments [43].
- **Perturbation methods:** These techniques build on Taylor expansions for f . The uncertainty is derived from truncated Taylor expansions and characterized by mean, variance, or covariance. However, these measures provide no information about the shape of the probability density function of the output. First-order Taylor expansions lead to exact results for linear models, whereas the accuracy is limited for highly non-linear problems. Though higher-order terms could be included in the Taylor expansion at considerable cost, this is reasonable only if the accuracy is required [32, 42]. Besides, the input and

output uncertainties should not be greater than 10% [43].

- **Sampling based methods** such as Monte Carlo (MC) techniques: Sampling methods evaluate a model several times for different sets of input values. The results are collected and then described statistically. The advantage of sampling methods is that they are intuitive to the researcher and independent of the dimension of the parameter space. Moreover, the solution converges to the exact solution for sufficiently large samples [12]. But they converge slowly, e.g. mean value obtained with MC and N realisations converges at rate $\mathcal{O}(\frac{1}{\sqrt{N}})$. As a consequence, many realisations are required to reach sufficiently accurate results. So, these methods are feasible as long as the computational effort for evaluating the model is moderate [32].
- **Spectral representations:** These methods are based on spectral expansions. Usually, a series of orthogonal polynomials represents the random process. Therefore, they are often referred to as (generalised) polynomial chaos expansions. The spectral representations are used to calculate the uncertainties or statistics for the model response. The advantage is that, compared to MC, these approaches converge much faster and thus necessitate fewer model evaluations. They merely require a smooth relation between the random input parameters and the solution [32, 42].

For the purpose of this thesis, direct evaluation and perturbation methods are not applicable because of their aforementioned limitations: Neither can I access or alter the code of the non-linear model, nor can I employ too large samples. But the polynomial chaos expansion shows a lot of promise, since it seems to be effective and to have no severe drawbacks that would constrain the application to the optimal steps model. Besides, Monte Carlo techniques with small sample sizes might be useful as well. Therefore, I explain these two approaches in detail.

Monte Carlo Method

Uncertainty quantification with Monte Carlo methods consists of the following steps:

1. Generate a sample for input parameter X that has a normal or uniform distribution. If the input parameter follows a different distribution, one can create (pseudo-)random numbers on $[0, 1]$ and then apply the Rosenblatt transformation to map them into the domain of X .
2. Evaluate the model for each sample point. This trivial task, also called an embarrassingly parallel problem, can be executed simultaneously to exploit the full computing capacity and to speed up the process.
3. Collect the model output and visualise its distribution with histograms or use

metrics mean, standard deviation and higher moments to describe the output uncertainty.

In principle, this procedure utilizes the Law of Large Numbers: As the sample size increases, the average of Y converges to the expected value [36]. The same applies to higher-order moments but the convergence rate is smaller so more realisations are needed [12]. For a sufficiently large sample size mean and variance can be expressed as follows:

$$\mathbb{E}[Y] \approx \frac{1}{N} \sum_{i=1}^N Y_i = \bar{\mu} \quad (2.8)$$

$$\begin{aligned} V[Y] &= \mathbb{E}[(Y - \mathbb{E}[Y])^2] \\ &= \mathbb{E}[Y^2 - 2Y\mathbb{E}[Y] + \mathbb{E}[Y]^2] \\ &= \mathbb{E}[Y^2] - 2\mathbb{E}[Y]\mathbb{E}[Y] + \mathbb{E}[Y]^2 \\ &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\ &\approx \frac{1}{N} \sum_{i=1}^N Y_i^2 - \left(\frac{1}{N} \sum_{i=1}^N Y_i \right)^2 \end{aligned} \quad (2.9)$$

The efficiency of the MC method can be improved by employing Quasi Monte Carlo schemes, e.g. the Sobol sequence, or other sampling techniques such as Latin Hypercube Sampling [3]. In the following, MC with randomly drawn samples is termed MC method.

If more than one input parameter is treated as random variable, one can extend the problem to dimension d by using multivariate distributions. Then X is a random vector with mutually independent variables $\mathbf{X} = [X_1, \dots, X_d]$. For example, let

$$f(X_1, X_2) = \frac{X_1}{2} + 5^{X_2} \quad (2.10)$$

represent any arbitrary model, where input parameters X_1 and X_2 are tainted with uncertainties and thus expressed as independent random variables. Figure 2.2 shows the sample for joint distribution of X_1 and X_2 and the output uncertainty in the shape of a histogram. Alternatively, one could also calculate statistical moments from response vector Y .

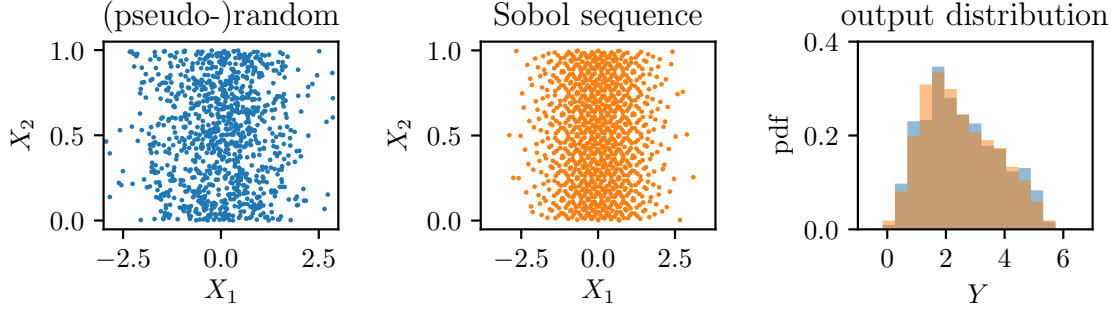


Figure 2.2: Monte Carlo method with different sampling techniques: $N = 1000$ sample points for random variables $X_1 \sim \mathcal{N}(0, 1)$ and $X_2 \sim \mathcal{U}(0, 1)$ generated with (pseudo-)random numbers and Sobol low-discrepancy sequence. The histogram shows the distribution of model response Y for both samples.

Otherwise, in case of correlated non-Gaussian, or non-uniform parameters, these variables are converted into independent normally distributed random variables with the aid of the Nataf transformation and Cholesky decomposition. However, this is beyond the scope of this thesis which is why I refer to [32] for further information.

Generalised Polynomial Chaos

Xiu [43] describes the development of the generalised polynomial chaos to explain its naming: The generalised polynomial chaos evolves from the Wiener-Hermite polynomial chaos. Norbert Wiener used Hermite polynomials to decompose Gaussian stochastic processes, but Hermite polynomials are not suited as basis to approximate non-Gaussian distributions. This problem has been solved by developing the generalised polynomial chaos expansion, which is independent of the type of polynomials [43]. The labelling can be misleading, since there is nothing chaotic about the spectral representation in the context of UQ [32]. Moreover, there are different terms and abbreviations in the literature that refer to the same idea. I use the term polynomial chaos expansion (PCE) for the spectral representation of Gaussian and non-Gaussian distributions by globally smooth orthogonal polynomials.

The PCE works similarly to Fourier series: A function is approximated but with polynomials instead of sinusoids. Let \mathbf{X} be a vector of parameters and Y the response of model f . The problem might depend on time t and space \mathbf{x} . For the sake of simplicity, I set aside the spatial component. Hence I rewrite Equation 2.7 as

$$Y = f(t, \mathbf{X}), \quad (2.11)$$

where \mathbf{X} has a known joint probability density function $\rho(\mathbf{X})$ and thus introduces uncertainties into the problem. A series of polynomials, truncated after N terms,

approximates function $f(t, \mathbf{X})$:

$$\begin{aligned} f(t, \mathbf{X}) &= \sum_{n=0}^{\infty} \hat{f}_n(t) \phi_n(\mathbf{X}) \\ &\approx \sum_{n=0}^{N-1} \hat{f}_n(t) \phi_n(\mathbf{X}) \end{aligned} \quad (2.12)$$

$\hat{f}_n(t)$ are coefficients and yet unknown. Polynomials ϕ_n are of degree n and their type depends on the distribution of the parameters $\rho(\mathbf{X})$. If Equation 2.12 approximates $f(t, \mathbf{X})$ well, one can calculate statistical properties by performing three steps [3]:

1. Construction of basis functions ϕ_n
2. Computation of coefficients $\hat{f}_n(t)$
3. Computation of statistical properties based on $\hat{f}_n(t)$

While the approximation is applicable to problems with multiple random variables if multidimensional basis functions are employed [32, 43], I refer to the univariate case for the following explanations to keep it simple.

Step 1: Construction of Basis Functions ϕ_n

According to [11], one can use different basis functions depending on the input parameter distribution as listed in Table 2.1. The choice of the type of polynomials might affect the convergence of the method [11].

Table 2.1: Relation between common distributions of random variables and type of polynomials to be chosen for PCE; An exhaustive list can be found in [43].

Distribution	Type of polynomials
Gaussian	Hermite
Gamma	Laguerre
Beta	Jacobi
Uniform	Legendre

Moreover, the polynomials ϕ_n are orthogonal. This property can be expressed

mathematically with the L^2 inner product $\langle \cdot, \cdot \rangle_\rho$ and weight $\rho(X)$:

$$\langle \phi_i(X), \phi_j(X) \rangle_\rho = \int \phi_i(X) \phi_j(X) \rho(X) dX = \gamma_i \delta_{ij} \quad (2.13)$$

$$\text{with Kronecker delta } \delta_{ij} = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}$$

$$\langle \phi_i(X), \phi_i(X) \rangle_\rho = \gamma_i \quad (2.14)$$

γ_i are normalisation constants [32, 43]. A numerically stable approach to determine the orthogonal polynomials is the three-terms recurrence relation with initial conditions $\phi_{-1}(X) = 0$ and $\phi_0(X) = 1$ [3]:

$$\phi_{n+1}(X) = (A_n X + B_n) \phi_n(X) - C_n \phi_{n-1}(X) \quad n \geq 0 \quad (2.15)$$

A_n , B_n and C_n are constants and computed on the basis of weight ρ . For example, let X be a random variable with uniform distribution $\mathcal{U}(-1, 1)$ and weight $\rho(X) = \frac{1}{2}$. According to Table 2.1, Legendre polynomials are optimal to solve this problem. Hence

$$\int_{-1}^1 \phi_i(X) \phi_j(X) \rho(X) dX = \frac{2}{2i+1} \delta_{ij}$$

and Equation 2.15 yields

$$\begin{array}{lll} \phi_{-1} = 0 & \phi_1 = X & \phi_3 = \frac{1}{2} (5X^3 - 3X) \\ \phi_0 = 1 & \phi_2 = \frac{1}{2} (3X^2 - 1) & \dots \end{array}$$

This and additional examples can be found in [36].

Note that the three-term recursion relation requires independent random variables in case of a multivariate problem. Bertran's recursion, Cholesky decomposition and modified Gram-Schmidt orthogonalization resolve this issue, since these approaches generate orthogonal polynomials also for statistically dependent random variables. More details can be found in [3].

Step 2: Computation of Expansion Coefficients $\hat{f}_n(t)$

Now that the base polynomials in Equation 2.12 are defined, one must determine the coefficients $\hat{f}_n(t)$ by means of stochastic collocation, integration schemes or the Galerkin method [3]. The latter belongs to the intrusive methods, hence it is not applicable in the context of this thesis, whereas collocation and integration might prove successful. Again, the denotation of these approaches is not standardised (compare [3, 11, 12, 32, 36, 43]): Point collocation is equivalent to stochastic collocation,

while the pseudo-spectral projection or discrete projection method is an integration scheme. I follow the notation of [3, 11] and use the terms point collocation method and pseudo-spectral approach.

Both approaches utilise of the orthogonal basis functions defined in Equation 2.13. So the truncated series of polynomials (see Equation 2.12) for the univariate case yields [43]:

$$\begin{aligned}
\langle f(t, X), \phi_m(X) \rangle_\rho &= \left\langle \sum_{n=0}^{N-1} \hat{f}_n(t) \phi_n(X), \phi_m(X) \right\rangle_\rho \\
&= \sum_{n=0}^{N-1} \hat{f}_n(t) \langle \phi_n(X), \phi_m(X) \rangle_\rho \\
&= \sum_{n=0}^{N-1} \hat{f}_n(t) \delta_{nm}
\end{aligned} \tag{2.16}$$

This can be simplified once again because the Kronecker delta $\delta_{nm} = 0$ if $n \neq m$. While for $m = n$ the indices can be exchanged and $\delta_{nm} = 1$. Hence

$$\langle f(t, X), \phi_n(X) \rangle_\rho = \hat{f}_n(t), \tag{2.17}$$

where $f(t, X)$ might not be available if the model is considered as a black box or evaluated at great computational expense. This issue is addressed by point collocation or the pseudo-spectral projection method.

Point Collocation Method

The point collocation (PC) method replaces Equation 2.17 by a system of equations [11]:

$$\begin{bmatrix} \phi_0(x_0) & \dots & \phi_N(x_1) \\ \vdots & \ddots & \\ \phi_0(x_M) & & \phi_N(x_M) \end{bmatrix} \begin{bmatrix} \hat{f}_0(t) \\ \vdots \\ \hat{f}_N(t) \end{bmatrix} = \begin{bmatrix} f(t, x_0) \\ \vdots \\ f(t, x_M) \end{bmatrix} \tag{2.18}$$

The M nodes x_m at which the model is evaluated can be chosen arbitrarily. That is, they can be defined for example as (pseudo-)random MC samples [3]. If $M = N$, one could use an interpolation approach to solve the system. But this is not recommended because interpolation is not robust in particular for multivariate problems [44]. Generally, regression methods provide better results. In case of an over-determined system, i.e. $M > N$, the method of least squares is the usual choice to minimize the error difference of approximation and exact solution [44]. Hosder, Walters, and Balch [11] define a ratio α_{PC} of actual sample size M' to minimum required number of sample points $M + 1$. The latter depends on the dimension of the parameter space d and the order of the polynomial chaos p , i.e. the expansion

in Equation 2.12 is truncated after $p + 1$ terms:

$$M + 1 = \frac{(d + p)!}{d! p!} \quad (2.19)$$

$$\alpha_{PC} = \frac{M'}{M + 1} \quad (2.20)$$

In other words, α_{PC} expresses the degree to which the system of Equations 2.18 is over-determined. It is suggested that using $M' = 2(M + 1)$ collocation points leads to better results [11].

Pseudo-Spectral Projection Method

Alternatively to collocation methods, the pseudo-spectral (PS) approach employs quadrature rules such as Gaussian quadrature: M nodes x_m and weights w_m are defined with respect to the input density ρ . Then the model must be evaluated only at the nodes to obtain the coefficients $\hat{f}_n(t)$ [32]:

$$\hat{f}_n(t) = \sum_{m=0}^{M-1} f(t, x_m) \phi_n(x_m) w_m \quad (2.21)$$

Note that $\hat{f}_n(t)$ are merely approximations of the exact solution [43]. The univariate problem can be generalised to a multivariate one by employing tensor products, but integration in higher dimensions becomes costly [3].

Step 3: Computation of Statistical Moments Based on Expansion Coefficients $\hat{f}_n(t)$

Recall that the actual objective is to calculate the statistics, e.g. mean and variance, of model response $Y = f(t, X)$. The expected value results from truncated expansion defined by Equation 2.12 and initial condition for the three-term recursion relation $\phi_0(X) = 1$ [32]:

$$\begin{aligned} \mathbb{E}[f(t, X)] &\approx \mathbb{E} \left[\sum_{n=0}^{N-1} \hat{f}_n(t) \phi_n(X) \right] \\ &= \sum_{n=0}^{N-1} \hat{f}_n(t) \mathbb{E}[1 \cdot \phi_n(X)] \\ &= \sum_{n=0}^{N-1} \hat{f}_n(t) \mathbb{E}[\phi_0(X) \cdot \phi_n(X)] \\ &= \sum_{n=0}^{N-1} \hat{f}_n(t) \delta_{0n} = \hat{f}_0(t) \end{aligned} \quad (2.22)$$

The variance exploits Equation 2.14 in the slightly different form

$$\langle \phi_n(X), \phi_n(X) \rangle_\rho = \gamma_n = \mathbb{E}[\phi_n^2(X)] \quad (2.23)$$

so that it can be simplified as

$$\begin{aligned} V[f(t, X)] &= \mathbb{E} \left[(f(t, X) - \mathbb{E}[f(t, X)])^2 \right] \\ &\approx \mathbb{E} \left[\left(\sum_{n=0}^{N-1} \hat{f}_n(t) \phi_n(X) - \hat{f}_n(t) \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_{n=1}^{N-1} \hat{f}_n(t) \phi_n(X) \right)^2 \right] \\ &= \sum_{n=1}^{N-1} \hat{f}_n^2(t) \mathbb{E} [\phi_n^2(X)] + \mathbb{E} \left[\sum_{n=1}^{N-1} \sum_{\substack{m=1 \\ m \neq n}}^{N-1} \hat{f}_n(t) \hat{f}_m(t) \underbrace{\phi_n(X) \phi_m(X)}_0 \right] \\ &= \sum_{n=1}^{N-1} \hat{f}_n^2(t) \gamma_n. \end{aligned} \quad (2.24)$$

Higher statistical moments can be obtained in a similar manner [32].

2.2.2 Sensitivity Analysis

Sensitivity analysis is perceived as the “study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input” [20, p. 45]. It often helps to identify crucial parameters and to prioritise them. Accordingly, important parameters are chosen with great care to improve the simulation, whereas non-influential ones could be fixed in order to simplify the model [21].

Commonly, the practices developed for this purpose are categorised as local or global. Local methods vary the uncertain parameter around a nominal value, hence the information about the sensitivity of the model is valid only for a restricted range of the parameter. On the contrary, global methods take into account the uncertainty of several parameters over the total parameter space at once. This allows to analyse the sensitivity on the whole interval of the uncertain parameters. Besides, global methods provide information about interaction effects among the parameters on the output [21]. The concept of global sensitivity analysis serves better for the intent of this thesis because the model under consideration has many input parameters that might interact and the Vadere users are interested in identifying these interactions.

To narrow down the broad field of global sensitivity analysis techniques, one can distinguish between regression-based and variance-based methods. Linear regression methods are useful for linear models, but they cannot analyse the sensitivity of non-

linear non-monotonic models [35]. Variance-based methods evaluate the effect of the variance in the input to a model on the variance of the output [21].

One common metric are so-called Sobol' indices [10, 33]. A slightly modified version of the Sobol' indices usually calculated from MC samples is based on spectral representations, just like uncertainty analysis with PCE. This is advantageous because information about the sensitivity can be obtained almost as by-product when conducting uncertainty analyses. Generally, sensitivity indices are a normalised measure for the sensitivity of the QoI. They quantify the influence of each input parameter that is considered uncertain on the model output. Time dependent quantities can be evaluated just as well as scalar ones by calculating the sensitivity indices for each time step [32]. This allows for tracking the importance of a certain parameter over time, which can prove beneficial if changes in the ranking of parameters are to be expected. Because of all these advantages, the Sobol' indices seem a useful tool for the analysis of crowd simulations. I introduce both PCE and MC based Sobol' indices in the next section.

Sobol' Decomposition of the PCE

The model f (see Eq. 2.7) is considered once again, now for the multivariate case:

$$Y = f(\mathbf{X}) \quad (2.25)$$

For simplicity any dependencies of space or time are neglected. Input vector $\mathbf{X} = [X_1, \dots, X_d]$ contains independent and on $[0, 1]$ uniformly distributed variables. The Sobol' decomposition works as well for other densities than uniform distributions, but the method still relies on the premise that the random variables are mutually independent. Otherwise, Nataf or Rosenblatt transformation can be applied to convert them into independent variables [32]. This is beyond the scope of the present work. Equation 2.25 can be expressed as high-dimensional model representation or Sobol' representation [32]:

$$f = f_0 + \sum_{i=1}^d f_i + \sum_i \sum_{i < j} f_{ij} + \dots + f_{1, \dots, d}, \quad (2.26)$$

where $f_i = f_i(X_i)$, $f_{ij} = f_{ij}(X_i, X_j)$. f_0 represents the mean of the model response, and f_i are effects caused by each single parameter, whereas terms with multiple indices denote contributions that originate from interactions between the input variables [32]. The variance of model response Y can be estimated as follows:

$$V[Y] = V[f] = \sum_{i=1}^d V[f_i] + \sum_i \sum_{i < j} V[f_{ij}] + \dots + V[f_{1, \dots, d}] \quad (2.27)$$

And division by the total variance $V[f]$ yields

$$1 = \sum_{i=1}^d \frac{V[f_i]}{V[f]} + \sum_i \sum_{i<j} \frac{V[f_{ij}]}{V[f]} + \dots + \frac{V[f_{1,\dots,d}]}{V[f]} \quad (2.28)$$

$$1 = \sum_{i=1}^d S_i + \sum_i \sum_{i<j} S_{ij} + \dots + S_{1,\dots,d}. \quad (2.29)$$

Equation 2.29 defines sensitivity indices as ratio of partial variances to the total variance. First order sensitivity indices S_i quantify the variance of the model response if input parameter X_i varies. One can denote this more detailed than in the equation above as $V[\mathbb{E}[Y | X_i]]$. Higher order sensitivity indices can be calculated as well, but usually their contribution is negligible. Total sensitivity indices S_{Ti} characterise the output variance if all input parameters vary except for parameter X_i . A concise expression of this is $V[\mathbb{E}[Y | X_{\sim i}]]$. Normalisation by the total variance $V[Y]$ yields

$$S_i = \frac{V[\mathbb{E}[Y | X_i]]}{V[Y]}, \quad (2.30)$$

$$S_{Ti} = 1 - \frac{V[\mathbb{E}[Y | X_{\sim i}]]}{V[Y]} = \frac{\mathbb{E}[V[Y | X_{\sim i}]]}{V[Y]}. \quad (2.31)$$

In other words, S_{Ti} evaluates the influence of parameter X_i and the interactions between X_i and all other parameters on the model response [32]. Therefore, it can be expressed as

$$S_{Ti} = S_i + \sum_i \sum_{i<j} S_{ij} + \dots + S_{1,\dots,d}. \quad (2.32)$$

Both S_i and S_{Ti} range between $[0, 1]$ due to normalisation by the total variance. Value 0 indicates that the parameter under consideration has no influence, whereas 1 denotes a significant influence.

The coefficients needed for the sensitivity indices have been calculated already in the scope of the uncertainty analysis in Section 2.2.1. Note that one must consider the multivariate case now. The expansion coefficients are selected in agreement with the dependency of the basis functions. The squares of these coefficients are summed and normalised by the total variance to obtain the sensitivity indices. Once the first, second and higher order sensitivity indices are known, the total sensitivity can be approximated by Equation 2.32, which is the sum of the sensitivity indices [35].

Monte Carlo Based Sensitivity Indices

The MC based sensitivity indices follow the same approach as the ones computed from the PCE coefficients. They only differ in the approximation of the partial and total variances in Equation 2.31 and 2.30, respectively. Several researchers have developed different estimators among which the following suggestions lead to the

most accurate results [22].

Let \mathbf{A} and \mathbf{B} be $N \times d$ matrices of independent random variables. N is the sample size and d is the number of input variables that follow a certain distribution. Matrices \mathbf{C}_i are equal to matrix \mathbf{A} except for column i , which is replaced by column i from \mathbf{B} :

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} x_{11}^A & \cdots & x_{1d}^A \\ \cdots & \cdots & \cdots \\ x_{N1}^A & \cdots & x_{Nd}^A \end{bmatrix} & \mathbf{B} &= \begin{bmatrix} x_{11}^B & \cdots & x_{1d}^B \\ \cdots & \cdots & \cdots \\ x_{N1}^B & \cdots & x_{Nd}^B \end{bmatrix} \\ \mathbf{C}_i &= \begin{bmatrix} x_{11}^A & \cdots & x_{1i}^B & \cdots & x_{1d}^A \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{N1}^A & \cdots & x_{Ni}^B & \cdots & x_{Nd}^A \end{bmatrix} \end{aligned} \quad (2.33)$$

Jansen [14] determines the partial variances with

$$V[\mathbb{E}[Y | X_i]] = V[Y] - \frac{1}{2N} \sum_{j=1}^N [f(\mathbf{B})_j - f(\mathbf{C}_i)_j]^2, \quad (2.34)$$

$$V[\mathbb{E}[Y | X_{\sim i}]] = \frac{1}{2N} \sum_{j=1}^N [f(\mathbf{A})_j - f(\mathbf{C}_i)_j]^2. \quad (2.35)$$

Saltelli et al. [22] suggest to calculate the first order sensitivities based on

$$V[\mathbb{E}[Y | X_i]] = \frac{1}{N} \sum_{j=1}^N f(\mathbf{B})_j [f(\mathbf{C}_i)_j - f(\mathbf{A})_j]. \quad (2.36)$$

The estimator for the total variance is given in [21] as

$$V[Y] = \frac{1}{N} \sum_{j=1}^N [f(\mathbf{A})_j]^2 - \left[\sum_{j=1}^N f(\mathbf{A})_j \right]^2, \quad (2.37)$$

but the result is more accurate if all independent samples, i.e. \mathbf{A} and \mathbf{B} , instead of \mathbf{A} are taken into account [19].

In both cases, following Jansen or Saltelli et al., the model must be evaluated for a total number of $(2 + d)N$ sample points. Since N usually ranges between a few hundred to a few thousand, the MC based sensitivity indices becomes infeasible for computationally heavy models [21].

This section briefly introduced a selection of UQ methods. Their aim is to quantify uncertainty and sensitivity of multidimensional stochastic problems. In particular non-intrusive point collocation or integration schemes based on PCE show promise in the context of pedestrian dynamics. Compared to MC, they are usually less computationally expensive. In the next section, I give an overview of studies in

which microscopic crowd simulations have been analysed by means of UQ methods.

2.3 Uncertainty Quantification in Microscopic Crowd Simulation

The introduction of simulators for microscopic crowd behaviour in Section 2.1 emphasised that the OSM suffers from uncertainties in the input parameters. Section 2.2 about UQ methods stresses that uncertainties should be considered to improve the performance of the model and to assess the reliability of the model before drawing conclusion from simulations. In the context of pedestrian dynamics, it is crucial to assure reliable results when planning mass events or escape routes inside buildings. Safety concepts based on putative confidential assumptions can be dangerous for the people who should be guided through critical passages.

The OSM is one of the available models in the framework Vadere to describe the movement of crowds. And the question arises how its uncertainty and sensitivity can be investigated properly. So far, employing forward and inverse uncertainty quantification methods to pedestrian locomotion models is rather uncommon up to now [5]. There are a few studies where Bayesian inversion helps to infer distributions of input parameters from given output data. For example, Gödel, Fischer, and Köster [5] apply Markov chain Monte Carlo to a simple scenario modelled with the OSM in Vadere. They show that the inversion can improve the quality of input parameters and thereby provide a basis for further development of a framework for UQ methods in microscopic crowd simulation. However, Bayesian inversion methods pursue a different objective than forward propagation. As mentioned in Section 2.2.1, applying inverse methods are usually preliminary proceedings to determine input uncertainties that are required for forward uncertainty quantification.

Forward propagation methods are applied in [1, 17, 30]. Dietrich et al. [1] build a data-driven surrogate model for a relatively small scenario, the entry and exit of passengers at a train station. The main goal is to speed up the process of generating the QoI, since surrogate models usually provide a computationally less expensive approximation than direct model evaluations. Nevertheless, this approach introduces errors and it remains unclear to which extent these inaccuracies affect the subsequently performed uncertainty quantification. This final step comprises stochastic collocation whose outcomes are compared to MC simulations. I do not go into depth about constructing surrogate models because this is another field of research, but the comparison of UQ results obtained with the point collocation method and MC simulations seems interesting.

Sivers et al. [30] examine the behaviour of a model for a train evacuation scenario,

which suffers from uncertain parameters, by means of the pseudo-spectral approach with Gauss integration based on the PCE. The polynomials have degree 6, which leads to 9126 model evaluations for a three dimensional parameter space, just to give an impression of the magnitude. The results prove the predictive capability of the simulation, even though it incorporates a highly uncertain parameter because this parameter has a minor effect on the model response.

Kurtc [17] investigates the sensitivity and uncertainty of two scenarios: A bottleneck scenario, commonly rather termed corridor scenario (compare [45]), and a first attempt to simulate a protest march through a city centre. She studies the corridor scenario as a modification of the experiment in [45] with respect to the three uncertain input parameters mean of free-flow speed, standard deviation of free-flow speed and number of agents. The QoI are the evacuation time, in other words the time between ingress and egress, and the density time series in a certain area. The uncertain input parameters for the simulation of the protest march are standard deviation of free-flow speed and number of agents. The UQ is based on spectral representations where the polynomial coefficients are obtained with the stochastic collocation method. This is applied in different schemes to the model response. Since the employed OSM is not deterministic, Kurtc tries to reduce the stochastic effects in the output by considering an average of repeated model evaluations for the same sample or an average of repeated UQ applications. Thus, the reliability of the UQ results increases to some extent. A comparison between the stochastic collocation results and a reference solution obtained with MC shows that the point collocation method seems useful to determine the uncertainty and sensitivity of the OSM. The study leaves the question open in which sense one can translate the UQ results for both scenarios into reality. In particular the simulation of the protest march is only a first trial. The input parameters need to be adapted to match real situations. For example, the number of agents ranges between 10 and 20, which represents a very small group of demonstrators.

In summary, there are a few studies that address forward propagation of uncertainties in microscopic crowd simulation. Among these, it is common practice to employ polynomial chaos expansions and to compare the results to a MC simulation, which serves as the ideal expected result. The polynomial coefficients are computed either with point collocation method or pseudo-spectral approach so the question which method is efficient remains unanswered. Furthermore, it is unclear which polynomial degree is appropriate. Therefore, I target the unsettled issue regarding the choice of the method and its configuration. Especially the scenarios addressed in [17] have promise which is why I seize Kurtc's suggestion to elaborate her principle ideas. I improve the set-up of the analysed scenarios and complement the missing interpretations with respect to reality.

3 Application of Uncertainty Quantification Methods

Inspired by the previously presented works, this thesis aims at quantifying the uncertainty of different microscopic crowd scenarios modelled with the optimal steps model (OSM). The idea is to perform forward propagation based on polynomial expansions. Since the model is considered as black box, I can only use non-intrusive methods such as point collocation and integration schemes. In contrast to MC simulations, increasing the number of sample points does not necessarily lead to better results. There are several settings, with the polynomial order being an example, that must be chosen carefully. Therefore, this chapter is dedicated to the adjustment of crucial settings for both point collocation and integration methods. The best configuration can then be used to analyse the uncertainty and sensitivity of different scenarios throughout the rest of this work.

The uncertainty and the sensitivity analysis is implemented in Python 3 for which UQ toolboxes are already available. Feinberg and Langtangen give an overview of some available toolboxes in [3]. Chaospy [2] is one option and a good choice especially because of its functionality: For example, the toolbox allows to control the generation of random variables for input distributions. It provides different sampling techniques and, most importantly, point collocation and integration schemes are both implemented. Subsequent UQ studies with Vadere can build on the implementation, which is introduced in the second part of this chapter.

3.1 Adaptation of Uncertainty Quantification Methods to the Model

There are many different options to perform UQ, as shown in Chapter 2.2, but only a few of these approaches are relevant when studying models such as the OSM. I use polynomial chaos expansions whose settings can be changed in many ways. Different set-ups thereof can lead to different results. Therefore, the task is to find an appropriate configuration. On the one hand, *appropriate* implies accurate outcomes. On

the other hand, the selected method should not be too computationally demanding. In the following sections,

- the point collocation method and
- the pseudo-spectral approach

are applied to a simple corridor scenario, which is introduced in Section 3.1.1. The results are evaluated with respect to computing time and accuracy. The accuracy is measured qualitatively by qualitatively comparing the output densities to a Monte Carlo (MC) reference solution. Complementary, the metrics root mean square error (RMSE) and mean absolute percentage error (MAPE) are used to obtain a quantitative assessment. Finally, I select one method with the most efficient configuration for the evaluation of the sandbox scenario. In the subsequent chapter, this method serves to analyse further scenarios.

3.1.1 Object of Investigation: The Corridor Scenario

The above-mentioned methods are applied to a corridor scenario as defined within the framework of an experiment in [45]. I choose this scenario because it is not too computationally demanding and, since it has been analysed many times in other works, it is a familiar example. In the experiment in [45], the researchers varied the width of the entrance b_{entr} , corridor b_{cor} and exit b_{exit} . Figure 3.1 shows the version with equal widths of 180 cm, hereinafter referred to as C-180-180-180. The computer model differs marginally from the experimental setting because the scenario is adapted in such a way that it compensates unnatural behaviour introduced by artificial components, such as the source. The important parts conform with the original topography. The locomotion of the agents is described by the OSM. The model has been calibrated to the corridor experiment in [45]. This leads to adaptations of parameters which describe the speed distribution of all agents. The changes to the default values are listed in Table A.1.

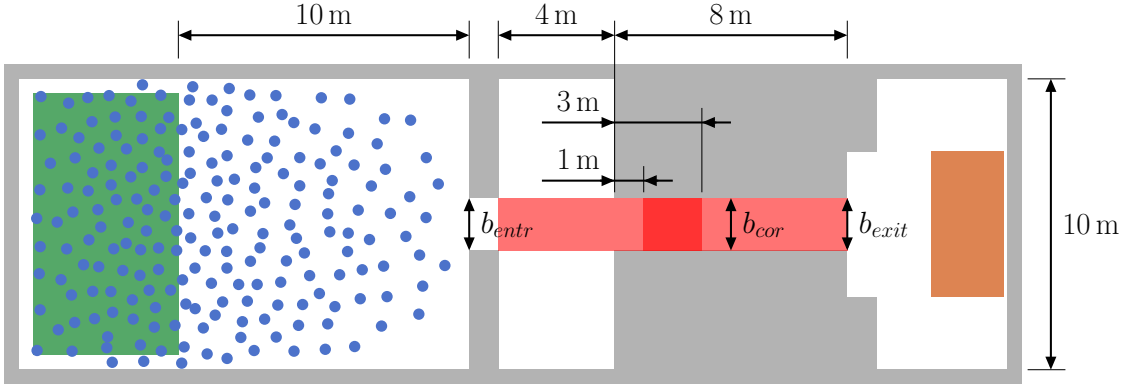


Figure 3.1: Corridor scenario modelling the experiment [45] in large part: Agents (blue) move from source to target and thereby pass the corridor. The width of the entrance, corridor and exit is defined by b_{entr} , b_{cor} and b_{exit} , respectively. The area in light red marks an auxiliary area A_{aux} for the measurement area A_D in dark red.

The calibrated values are still subject to uncertainties. This is why the parameters mean of free-flow speed μ_v , standard deviation of free-flow speed σ_v and number of agents n are considered uncertain within a defined range and it is assumed that each of them follows a uniform distribution (see Table 3.1). One should bear in mind that, theoretically, the multivariate PCE requires stochastically independent random variables. Actually, mean and standard deviation of free-flow speed are linked because the normal distribution is truncated. But I treat them as independent, acting on the assumption that their dependency is weak. Otherwise, performing UQ on models with dependent variables would become considerably complicated.

Table 3.1: Corridor scenario: The uncertainty in the input parameters is described by uniform distributions.

Parameter	Symbol	Distribution	Unit
Mean of free-flow speed	μ_v	$\mathcal{U}(1.37, 1.73)$	m s^{-1}
SD of free-flow speed	σ_v	$\mathcal{U}(0.10, 0.20)$	m s^{-1}
Number of agents	n	$\mathcal{U}(100, 200)$	

Now that the input uncertainty is specified, I can propagate it through the model by applying different methods as introduced in 3.1.2. The aim is to obtain the uncertainty of the following simulation outcomes:

- Density time series within measurement area A_D
- Mean over density time series within measurement area A_D
- Time between ingress and egress of all agents (evacuation time)

The density for both QoI a) and b) is based on Voronoi diagrams (see Figure 3.2) calculated by means of measurement areas A_D and A_{aux} . Agent l is enclosed by a Voronoi cell with area A_l . The boundary of the cell is defined by all points that are equidistant from the center of agent l and its closest neighbours [23]. Vadere calculates the density D_v inside measurement area A_D at time step t as follows [25]:

$$D_v(t) = \frac{N}{\sum_{l=1}^N |A_l|} \quad (3.1)$$

N is the number of agents whose center is located inside measurement area A_D , and $|A_l|$ the size of the corresponding Voronoi cell.

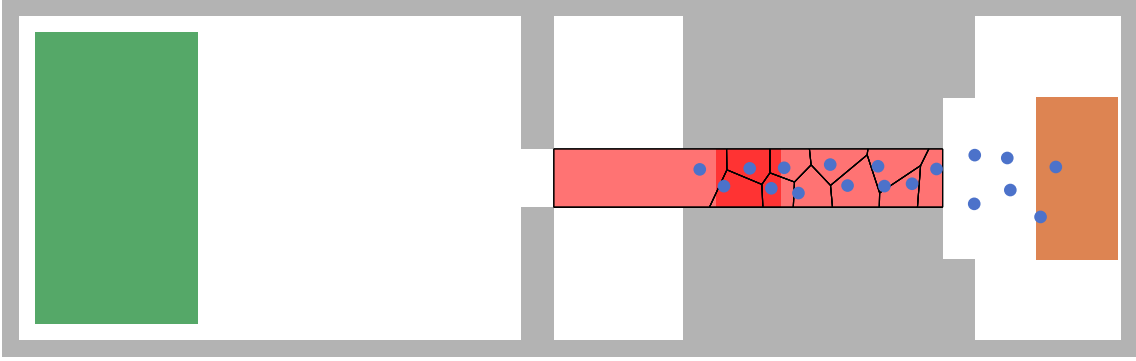


Figure 3.2: The Voronoi cells are defined by black lines that enclose each agent. In this example, the Voronoi cells are shown only inside auxiliary area A_{aux} .

3.1.2 Definition of Configurations for the Methods

The aim of this section is to define different configurations for the polynomial chaos expansion in combination with the point collocation method and the pseudo-spectral approach. Each of those methods can be adapted to the problem, for example, by defining a certain sample size or by choosing a reasonable polynomial order. The results of the adapted PCE methods are compared to a MC reference solution, which is introduced as well.

Monte Carlo Reference Solution

The reference solution is obtained by applying a Monte Carlo simulation with random sampling. It yields a good approximation if enough sample points are taken into account. An indicator for the right sample size is the convergence of the UQ output. Figure 3.3 shows the uncertainty and sensitivity for evacuation time as QoI over an increasing number of sample points. The other two quantities for mean and time series of the Voronoi density lead to qualitatively similar curves. Note that the sample size n_s for the statistical moments and the sensitivity indices differ (see Section 2.2.2). The curves for mean and standard deviation are nearly constant,

whereas some sensitivity indices change significantly if 50, 500 or 5000 sample points are used. The difference between the results for 5000 and 25000 sample points seems negligible. Therefore, the sensitivity indices calculated from 5000 sample points (or uncertainty obtained from 2000 sample points) is considered as converged and can be employed as ground truth. Now that a reliable reference solution is available, I can focus on the UQ methods that are actually of interest.

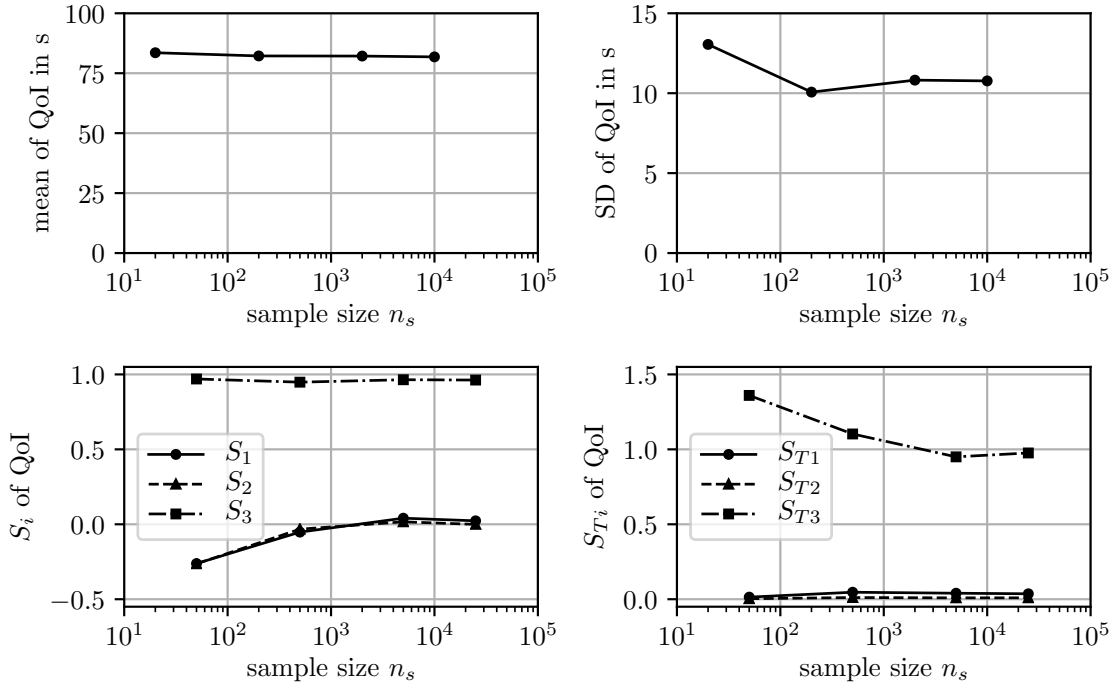


Figure 3.3: Convergence of the Monte Carlo method: Upper graphs show the UQ results for mean and standard deviation of evacuation time calculated from different sample sizes. The bottom graphs show first order (S_i) and total (S_{T_i}) sensitivity indices accordingly.

Definition of Different Sample Sizes for the PCE

I relinquish the MC simulation and return to the point collocation (PC) and pseudo-spectral (PS) approach based on the PCE. How large should the sample size be for these methods? In this paragraph, I define two key figures, α_{PC} and α_{PS} , which are used to vary the sample size systematically for the PC and PS method, respectively.

The sample size for the PC method can be chosen arbitrarily as long as the order of magnitude is correct. I want to recapitulate that the PC method leads to a system of equations (see Eq. 2.18). This system of equations can be solved by means of the least squares method to obtain the polynomial coefficients for the PCE. Let the polynomials of the PCE have degree 2 and the number of uncertain parameters is 3. Then the system of equations requires at least 10 sample points so that it can be solved with least squares (compare Eq. 2.19). However, one can also use more sample points than that, which leads to an over-determined system of

equations. The degree of how over- or under-determined the system of equations is can be specified by α_{PC} (see Eq. 2.20). For example, if 20 sample points are used instead of 10, the system is $\alpha_{PC} = \frac{20}{10} = 2$ times over-determined. In this manner, α_{PC} can be used to vary the sample size in a systematic scheme. Hosder, Walters, and Balch [11] suggest that, independent of sampling technique and polynomial degree, $\alpha_{PC} = 2$ leads to better approximations. However, it is not clear how distinct the qualitative improvements would be if this recommendation were applied to the case under consideration. Increasing α_{PC} might lead only to small differences in the result just as well as it could yield far better results, even at lower cost than it were the case for raising the polynomial order. For this reason, the PC method is analysed for $\alpha_{PC} = 1, 2, 3$. Note that this study does not entail the under-sampled case $\alpha_{PC} < 1$.

In contrast to the PC method, the sampling for the PS approach works differently. The required sample size depends on the order q of the underlying quadrature rule and number of uncertain parameters d :

$$n_s = (q + 1)^d \quad (3.2)$$

Similarly to the systematically varied factor α_{PC} , I introduce factor $\alpha_{PS} = 1, 2, 3$ and change the sample size according to the following scheme:

$$q = p \alpha_{PS} \quad (3.3)$$

where p is the polynomial degree of the PCE. For example, a problem with 3 uncertain parameters and second order polynomials requires 27, 125 and 343 sample points, respectively. The considered sample sizes for both PC and PS approach are summarised in Table 3.2.

Definition of Different Polynomial Orders for the PCE

Another setting that can affect the accuracy of the approximated distributions is the polynomial degree. It might be computationally more efficient either to enlarge the sample (via factor α_{PC} for PC or α_{PS} for PS as described above), or to increase the polynomial degree p [11]. The efficiency depends on the computational effort, which can be estimated coarsely by the number of model evaluations.

The number of sample points n_s for PC and PS approach is defined by Equation 2.19 and 3.2, respectively. Figure 3.4 visualizes the relation between α , p and n_s . The plot makes clear that the choice of p is more crucial for the PS approach, since the sample becomes large rapidly as the polynomial degree increases. As a consequence, the PS approach should be used with a low polynomial degree to avoid many model evaluations. In contrast, for the PC method the effects of p and α_{PC} are similar. So

one could think that the PC method with relatively high values for both α_{PC} and p should be acceptable in terms of computing time.

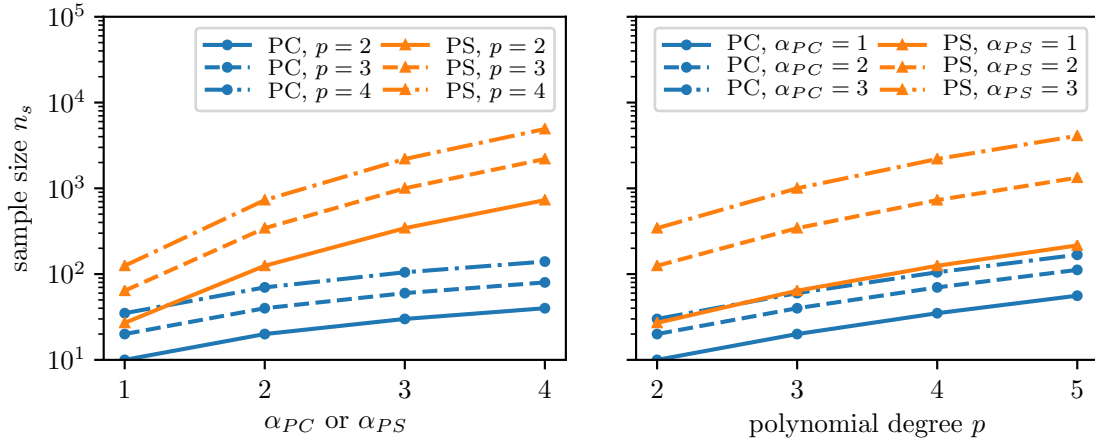


Figure 3.4: Computational effort for point collocation method and pseudo-spectral approach estimated by the number of sample points: Factors for systematic variation of the sample size α_{PC} and α_{PS} (left) are compared to the polynomial degree p of the PCE (right).

However, computations with the aid of Chaospy reveal that it takes considerably longer to compute statistics and, in particular, sensitivity indices for higher polynomial orders p . The Chaospy algorithm that calculates mean, standard deviation, first order and total sensitivity indices for a time dependent QoI sometimes runs up to 10 times longer than the code that evaluates the model. That is, the computational effort cannot be equated with the sample size. Instead, one has to record the total computing time. Moreover, a low polynomial degree is desirable to keep the runtime low.

Bearing that in mind, the following combinations for p and α seem reasonable to analyse the effect on the accuracy of the approximated distributions: The polynomial order ranges between 1 and 5. At the same time, α takes the values 1, 2, 3 but only for at maximum third order polynomials. Polynomials of degree 4 and 5 are not combined with $\alpha = 2, 3$ to avoid too heavy computations. These settings are summarised in Table 3.2.

Reducing Stochastic Behaviour of the Model Response

The third interesting aspect concerns primarily the model response but is also related to the UQ method. The OSM is not fully deterministic because it incorporates stochastic processes to improve the simulation. For example, random numbers are used to define the starting seed or to prevent agents from getting stuck while they navigate around obstacles and other agents. Therefore, simulations with identical settings of input parameters lead to slightly different results. This has an impact on

the outcome of the UQ method and its accuracy. There are two options to reduce this effect:

- a) Reduce the stochastic behaviour of the model by fixing the seed. In practice, this means that agents are spawned by the same pseudo random process for each simulation. If different UQ methods are applied to realisations that are obtained in this manner, the UQ output should be comparable. However, the results might be representative for only one specific initial condition but not for others. A different seed might yield a different response and therefore, different output uncertainties and sensitivities.
- b) Reduce the stochastic behaviour of the model response by taking the average of repeated model evaluations (in the following referred to as number of scenario runs n_r) for the same sample. Each repetition uses a different seed.

Since the computational effort increases by factor n_r , I fixed the seed and applied the UQ methods to the model response of a single run (a)) to pre-select the best three or four methods. The pre-selection is then run with the averaging (b)) to get a more general solution and to find the most suitable method based on this final evaluation. Note that, unless stated explicitly, the following sections refer to fixed initial conditions (a)).

Summary of the Parameters to be Adapted

I briefly summarise the previous paragraphs. From the considerations regarding sample size and polynomial degree arise the configurations for the point collocation method and the pseudo-spectral approach as listed in Table 3.2.

Table 3.2: Parameter set-up for the point collocation (PC) method and the pseudo-spectral (PS) approach

Method	Polynomial degree p	α_{PC} or α_{PS}	Sample size n_s
Point collocation	2	1	10
	2	2	20
	2	3	30
	3	1	20
	3	2	40
	3	3	60
	4	1	35
	4	2	70
Pseudo-spectral	5	1	112
	2	1	27
	2	2	125
	2	3	343
	3	1	64
	3	2	343
	4	1	125
	5	1	216

Each configuration is applied to the model response obtained by one single run. I pre-select the best configurations by comparing the UQ results to a Monte Carlo (MC) reference solution. The MC simulation uses 2000 sample points for the uncertainty analysis and 5000 for the sensitivity analysis. The pre-selected methods are then applied to the average of 10 repeated model evaluations with different initial conditions. The outcomes are evaluated in the next section.

3.1.3 Verification of Methods

This section comprises a visual comparison of the probability density functions of the model response. In doing so, one gets a first impression of the performance of the PCE. Additionally, the metrics mean absolute percentage error (MAPE) and root mean square error (RMSE) are used to evaluate the UQ results quantitatively.

Qualitative comparisons of the PC method with the MC reference solution lead to the conclusion that polynomial expansions obtained with 10 sample points ($\alpha_{PC} = 1$) are poor approximations, while in terms of accuracy 20 ($\alpha_{PC} = 2$) or 30 ($\alpha_{PC} = 3$) sample points yield equally good results. The approximations for PS with a sample

size of 27 or 125 ($\alpha_{PS} = 1$ or 2) are acceptable. As an example, Figure 3.5 shows approximated distributions for the mean Voronoi density compared to a histogram of the MC reference simulation. The graphs for the evacuation time look qualitatively similar (see Figure A.1), so I consider only the mean Voronoi density. Independent of the number of samples used, most approximations in this plot are good. Since the actual UQ results are derived from these curves, one can expect the mean, standard deviation and sensitivity indices to be in accordance with the MC based measures. PC with 10 sample points is the sole exception because it approximates the actual probability density function poorly.

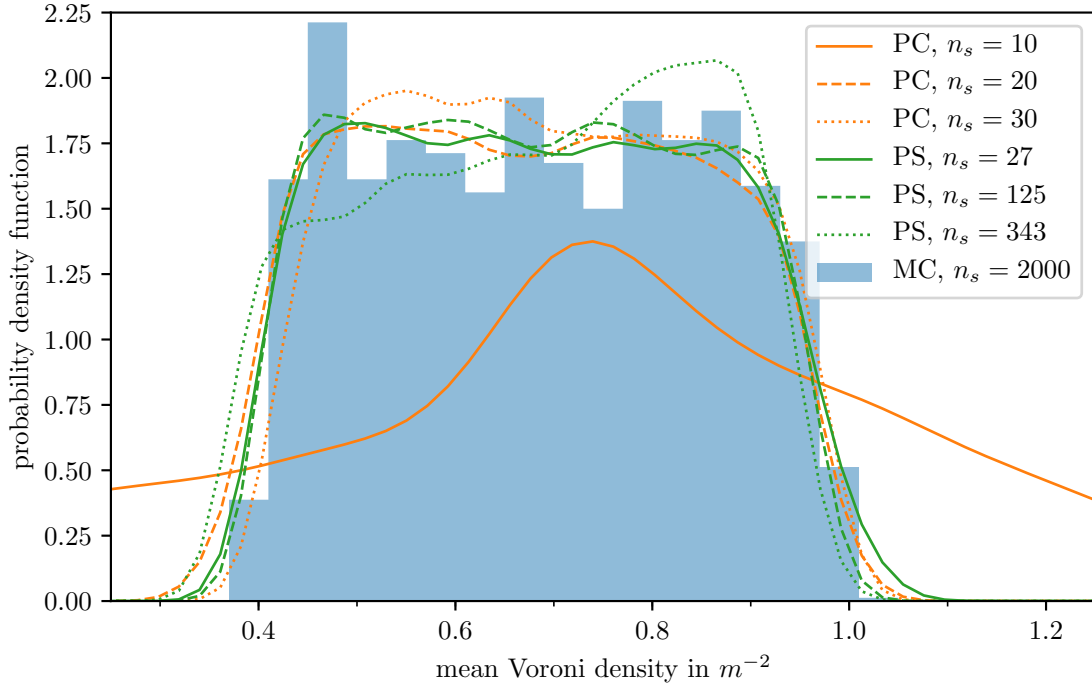


Figure 3.5: Qualitative comparison of output distributions approximated by PCE with a varying sample size: The polynomial degree is $p = 2$ for both point collocation (PC) and pseudo-spectral (PS) approach. The reference solution is obtained with Monte Carlo (MC) from 2000 sample points and represented by a histogram. The model response is the mean Voronoi density.

The influence of the polynomial degree on the accuracy of the PCE for the mean Voronoi density is shown in Figure 3.6. Again, the graphs for the evacuation time look qualitatively similar (see Figure A.2). Generally, the plot reveals that the PC method results in more stable approximations than the PS approach. As the polynomial degree is increased, the curves obtained with PC are equally good, whereas the PS approach leads to oscillating curves. The results for PC with sample sizes in accordance with $\alpha_{PC} = 1$ are not included in the graph because this approach yields unsatisfactory approximations.

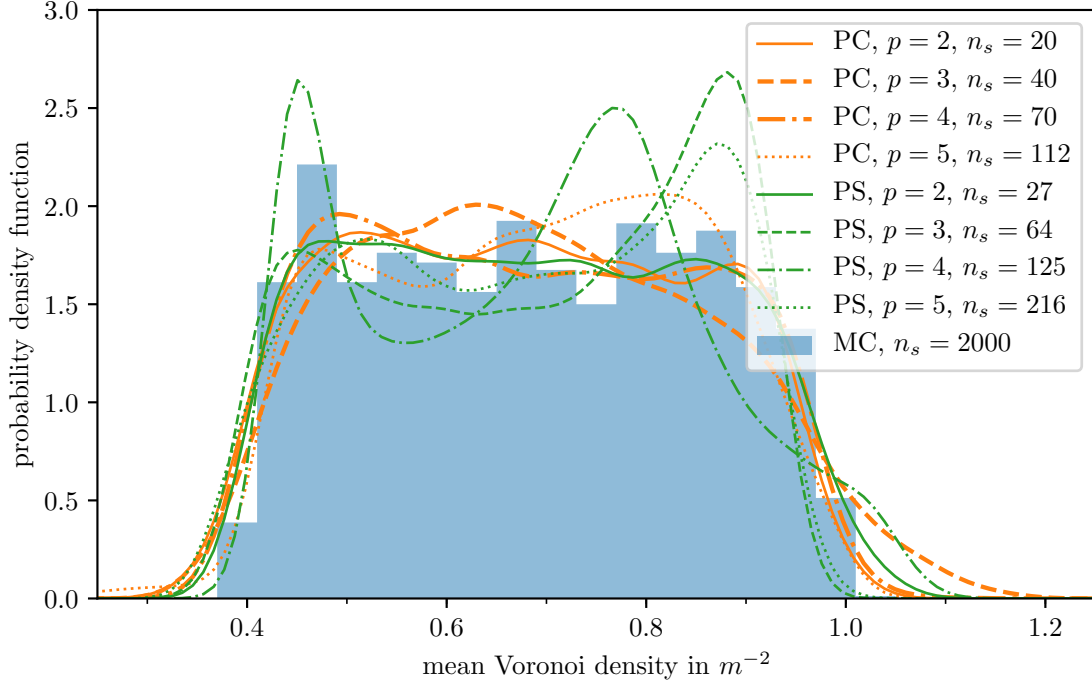


Figure 3.6: Qualitative comparison of output distributions for QoI mean Voronoi density approximated by polynomial chaos expansions with varying polynomial order p : The sample sizes n_s are in accordance with $\alpha_{PC} = 2$ in case of point collocation (PC) method and $\alpha_{PS} = 1$ for the pseudo-spectral (PS) approach. Thick lines indicate configurations of the PC method ($p = 3$ and $p = 4$) that will prove useful in the quantitative assessment. The reference solution is obtained with Monte Carlo (MC) from 2000 sample points and represented by a histogram.

Visual comparisons are useful to get a better understanding, but they cover only a part of the evaluation. Measures, such as MAPE and RMSE, capture the accuracy quantitatively and allow for reliable assessment. They are defined as follows:

$$\text{RMSE}(y) = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{ref,i} - y_i)^2}, \quad (3.4)$$

$$\text{MAPE}(y) = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_{ref,i} - y_i}{y_{ref,i}} \right| \quad (3.5)$$

where y_{ref} is the MC reference solution and y the outcome of the PC and PS method, respectively. $\text{MAPE}(y)$ is calculated for the outcome of the uncertainty analysis, while $\text{RSME}(y)$ is taken for the sensitivity analysis. That is, Equation 3.4 replaces y_{ref} and y by mean or standard deviation, while Equation 3.5 refers to first order or total sensitivity index for each parameter. If the QoI is a scalar quantity (e.g. evacuation time), $N = 1$, otherwise y_{ref} and y are time series with time steps $i = 1, \dots, N$.

I employ these two different metrics for the following reasons. The error of the uncertainty is calculated with MAPE to normalise the statistical moments of the

QoI, which have different units. Furthermore, mean and standard deviation reach values of different magnitudes. For example, the mean evacuation time might be 80 s, whereas the mean of mean Voronoi density is around 2 m^{-2} . Albeit the Voronoi density time series also reaches a few values around 0, I use the MAPE. Those few time steps where mean and standard deviation are less than a defined threshold (0.05 m^{-2}) are neglected. The sensitivity indices are dimensionless already, and they range from 0 to 1. Therefore, the RMSE is better suited.

Figure 3.7 compares the errors of the UQ results for all QoI obtained with the methods as listed in Table 3.2. Just the errors arising from the PC method with $\alpha_{PC} = 1$ are not included in the plot because their error is multiple times greater.

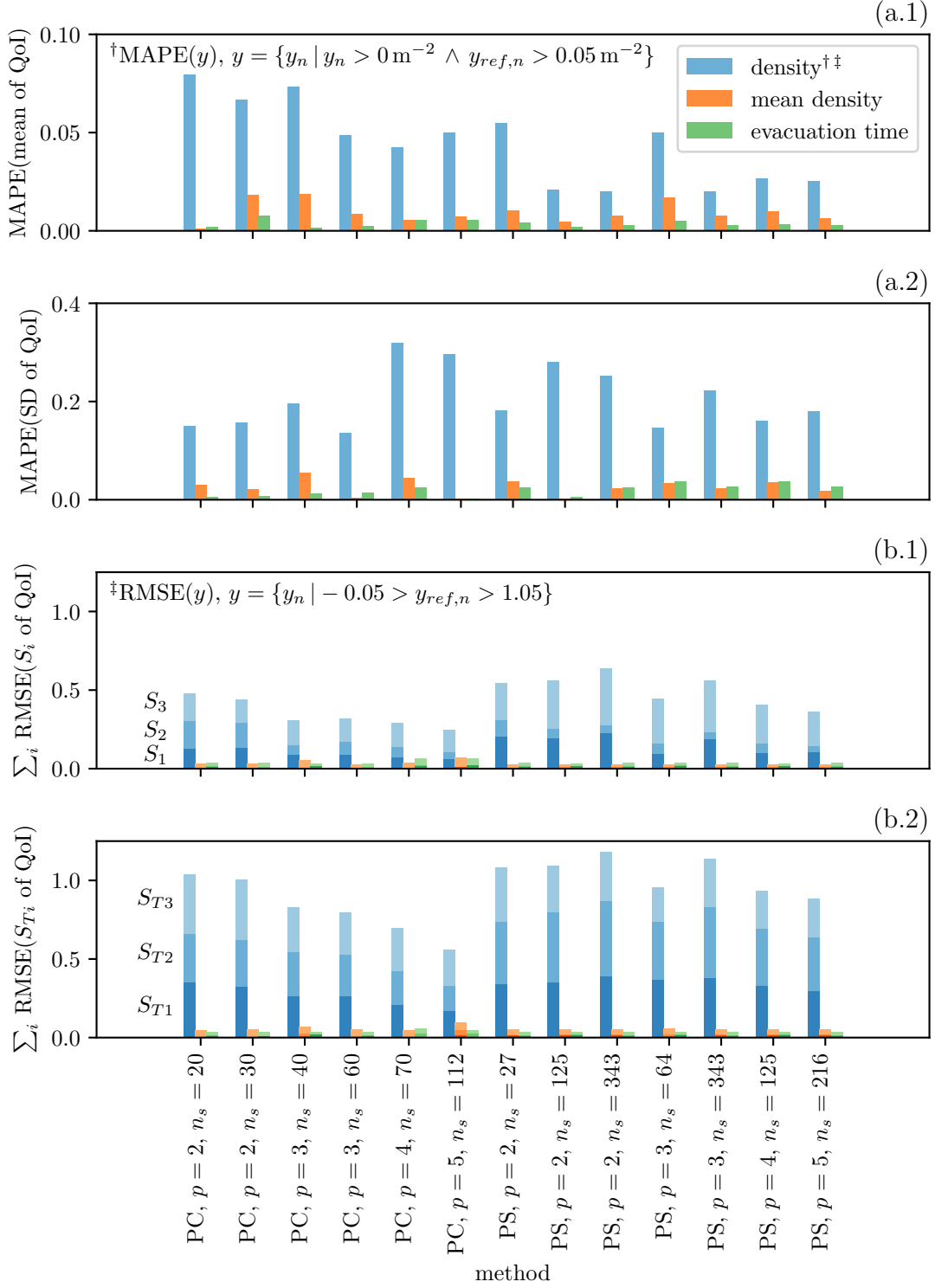


Figure 3.7: Quantitative comparison of different configurations of the UQ methods based on their errors: The UQ results for different QoI are calculated by point collocation (PC) and pseudo-spectral (PS) approach applied to the model response of a single run ($n_r = 1$). $\text{MAPE}(y)$ quantifies the error of mean (a.1) and standard deviation (a.2). $\text{RMSE}(y)$ quantifies the error of first order sensitivity indices (b.1) and total sensitivity indices (b.2). The reference values y_{ref} are derived from Monte Carlo (MC) simulations with $n_s = 2000$ sample points for the uncertainty analysis (a) and $n_s = 5000$ sample points for the sensitivity analysis (b). The estimation of the error for QoI Voronoi density time series neglects 3% of the data due to different thresholds for (a)[†] and (b)[‡].

Obviously, the uncertainty estimation for the time dependent QoI density is much more imprecise than for scalar QoI. This can be partly explained by the shape of the probability distribution for the Voronoi density time series. For a few frames at the beginning and at the end of the simulation, the PCE is not capable of approximating the actual distribution. These poor approximations adversely affect the estimates for uncertainty and sensitivity. More importantly, one must admit that the UQ results for scalar QoI are surprisingly good. Even though the PCE does not fit the exact output probability density, the UQ results derived from the polynomials are close to their true values. Figure 3.6 reveals that high order polynomials tend to oscillate around the true value. However, these deviations compensate each other by chance when mean, standard deviation and sensitivity indices are derived from the PCE. Therefore, mean, standard deviation and the sensitivity indices are close to the MC reference solution and the error is small.

In accordance with the visual comparison of PC (regression, i.e. $\alpha_{PC} > 1$) and PS approximations, none of the two approaches seem superior to one another. The influence of greater sample sizes turns out to be irrelevant in both cases. As the polynomial degree increases, the PC method yields slightly better results. The errors of different configurations for the PS approach are less distinct in relation to higher order polynomials. However, as explained above, one cannot trust only the quantitative evaluation.

Based on these considerations, I pre-select the following methods. They seem to be accurate and not too expensive in terms of computing time:

- PC: $p = 3$, $n_s = 30$
- PC: $p = 4$, $n_s = 70$
- PS: $p = 3$, $n_s = 64$
- PS: $p = 4$, $n_s = 125$

MAPE and RSME are calculated once more for these configurations. This time, the UQ methods are applied to the average model response of $n_r = 10$ repetitions in order to obtain a more reliable assessment (see Figure A.3). Compared to the previous evaluation based on one single run, the errors decline significantly. This concerns all four pre-selected configurations to the same degree so that their errors do not differ much from one another. Therefore, the best method is the most economical one. The efficiency is evaluated in the next section.

3.1.4 Rating of Methods

The final rating of the different configurations of the methods takes into account accuracy of the UQ results and computational effort. The analysis in the previous section leads to the conclusion that the error of the four pre-selected methods is

comparable. The computational effort, in other words the runtime, is primarily dominated by two factors:

Mainly, the polynomial degree p and, secondarily, the sample size n_s determines the computing time for the PCE. This is unexpected but can be explained as follows. The Chaospy algorithm requires some time to decompose the output variance sensitivity indices. Second and higher order sensitivity indices would take even longer because more coefficients must be calculated to satisfy Equation 2.29. This becomes noticeable in case of a time dependent QoI because the statistics and sensitivities are calculated for each time step. For example, PC ($p = 3$, $n_s = 40$, $\alpha_{PC} = 2$) requires less than 1 min to generate the samples and evaluate the model (preprocessing) and 21 min to calculate the PCE (post-processing), whereas PC ($p = 4$, $n_s = 70$, $\alpha_{PC} = 2$) runs 3 min and 25 min, respectively. These time measurements refer to simulations where all three QoI are considered, but the scalar ones contribute only little. One can expect the model evaluation to be more time consuming than the post-processing in case of large scenarios. In either case, these numbers suggest the user to choose low polynomial orders. In this manner, the number of sample points is kept low as well.

The PC method typically requires less sample points than the PS approach. Besides, the PC method yields more stable approximations of the probability density function of the model response. However, the user should bear in mind that the accuracy may differ, depending on the model response under consideration. For example, the time dependent Voronoi density exhibits some complicated distributions during a few time steps. Then the PCE in general is not capable of capturing their shape, no matter if one employs the PC method or the PS approach. In conclusion, the point collocation (PC) method with third order polynomials ($p = 3$) applied to a model response with 40 samples ($\alpha_{PC} = 2$) is an appropriate choice for further analyses. Its computational effort is justifiable and the method yields sufficiently accurate results. Moreover, the performance improves significantly if the method is applied to an average of repeated model evaluations.

3.2 Overview of the Implemented Program Code

The code for forward uncertainty quantification with Vadere consists of two parts: The evaluation of the model for all sample points and the uncertainty and sensitivity analysis. The flow chart in Figure 3.8 summarizes both scripts for pre- and post-processing.

At first, one has to define the scenario, its uncertainties and some other settings, such as sampling technique, number of sample points, etc. Chaospy provides an algorithm for generating the samples. The number of runs n_r specifies how many

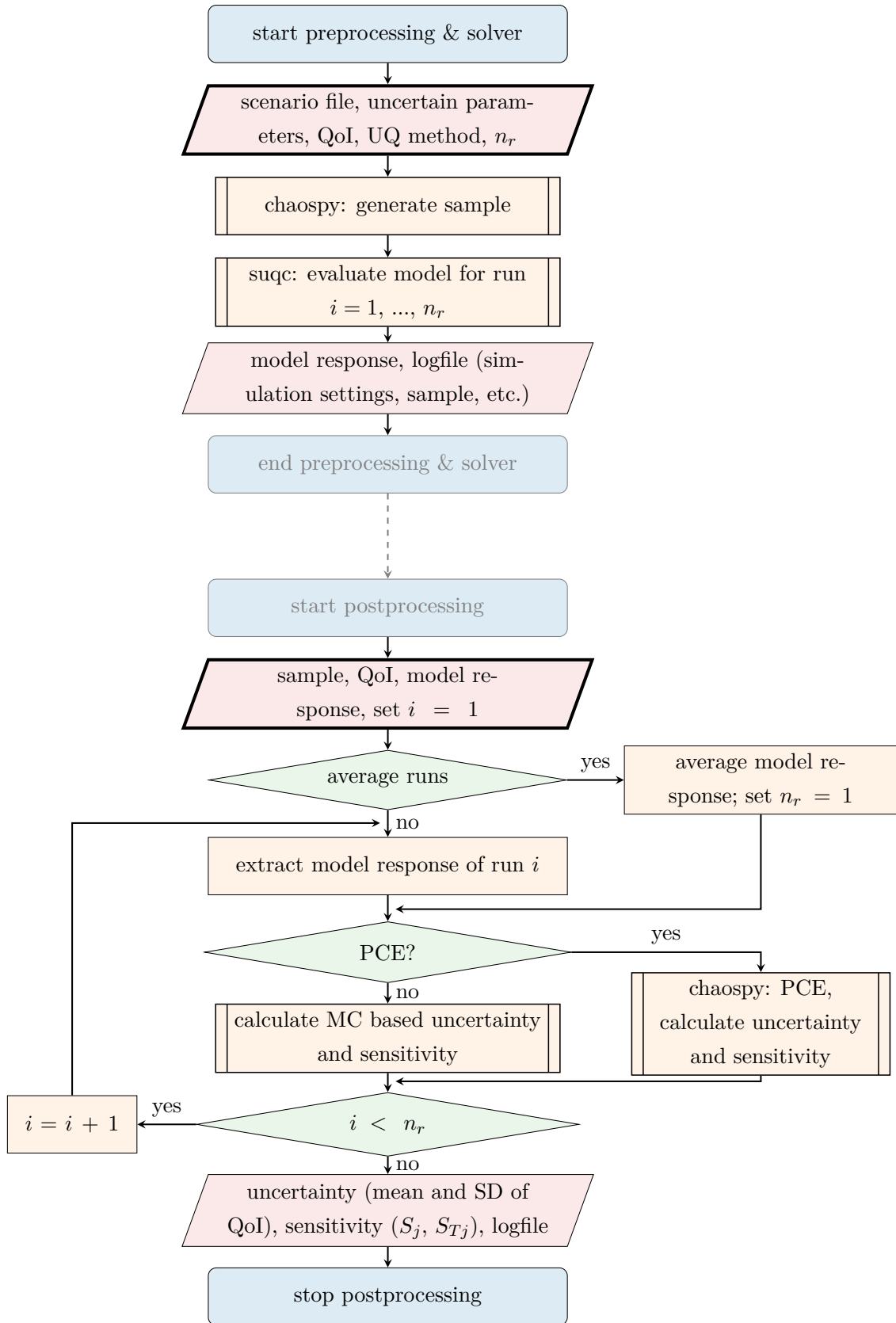


Figure 3.8: Simplified flowchart of the code: The first step after starting each process (tick outline) is important because the user must define key variables. n_r represents the number of scenario runs.

times the model is evaluated for the same sample. This can be used to reduce the effect of the stochastic behaviour of the OSM as explained above. Finally the SUQ (surrogate model and uncertainty quantification) controller [37], a tool designed for Vadere, triggers the model evaluations for each run.

The model response is processed in a second script. It converts the output data so that the results of all runs are averaged or only one single run is considered. In either case, UQ methods from the Chaospy library or a Monte Carlo based sensitivity analysis are applied to the response vector.

3.2.1 Embedding the Chaospy Library

Chaospy offers various options for PCE and Monte Carlo simulation. The most important commands and their embedding in the code are briefly explained in this section. More details can be found in [3].

This example is adapted to the corridor scenario I have discussed so far. At first, I define the distributions of the three uncertain parameters, mean and standard deviation of free-flow speed and number of agents, with continuous and discrete uniform distributions in a certain range, respectively:

```
import chaospy as cp
dist_mean_v = cp.Uniform(lower=1.37, upper=1.73)
dist_sigma_v = cp.Uniform(lower=0.10, upper=0.20)
dist_n = cp.DiscreteUniform(lower=100, upper=200)
```

The distributions are joined by

```
mdist = cp.J(dist_mean_v, dist_sigma_v, dist_n)
```

and one obtains a multivariate probability distribution.

The sampling depends on whether the user wants to apply MC, the PC or PS approach. In case of MC and PC method, the sample is generated by drawing 1000 pseudo-random numbers:

```
sample = mdist.sample(size=1000, rule="R")
```

The sampling technique is defined by the function argument `rule`. Even though the algorithm also provides different techniques, such as Latin Hypercube Sampling or the Hammersley sequence, one should be aware that the sampling technique must accord with the matrices that are constructed for the MC based Sobol' sensitivity indices (see Section 2.2.2).

The PS approach requires nodes (sample points) and their corresponding weights. The argument `order` defines the number of nodes. Chaospy offers various integration rules. Here, I apply Gauss-Legendre with `q` points because the input distribution is uniform:

```
sample, weights = cp.generate_quadrature(order=q, dist=mdist,
    rule="Gaussian")
```

Chaospy applies the three terms recursion relation to get a numerical stable PCE where `mdist` is assumed to be stochastically independent [2]. In the example, I use polynomials of third order. The least squares method or spectral projection yield an approximation of the polynomial coefficients for PC and PS approach, respectively:

```
polynomials = cp.orth_ttr(order=3, dist=mdist)
adist = cp.fit_regression(polynomials, sample, y)
adist = cp.fit_quadrature(polynomials, sample, weights, y)
```

`y` contains the response of the model.

The statistics, mean and standard deviation, and sensitivity indices are computed as follows:

```
mean = cp.E(adist, mdist)
sd = cp.Std(adist, mdist)
s1 = cp.Sens_m(adist, mdist)
st = cp.Sens_t(adist, mdist)
```

`s1` contains first order sensitivity indices and `st` the total sensitivity indices for parameter mean and standard deviation of free-flow speed and number of agents.

For MC simulations, I simply use the NumPy library to calculate mean and standard deviation of `y`. However, the MC based sensitivity indices are not computed by means of Chaospy or any other open-source library. This is why in particular these manually written functions must be checked for errors. The verification is addressed in the next section.

3.2.2 Verification of the Program Code

Code verification aims at finding and debugging errors in the program code to make sure that the software is correctly implemented. The OSM and Vadere have been verified and validated already [15]. Besides, Chaospy has been checked for proper operation and compared to other toolboxes by the developer so that one can call it state-of-the-art software [3]. I show that this external procured software is embedded and used in the right way. However, the MC based sensitivity analysis is coded manually and thus has to be tested comprehensively.

Verification of Simulations with the SUQ Controller

The main task of the SUQ controller is to call Vadere, execute the simulation for each sample point and return the model output. It does not visualize the simulation itself, so the user cannot be sure if the simulation was successful or if any artefacts

occurred. Checking a few dozens or even hundreds results of visually would be far too time-consuming. Therefore, I pursued the following remaining options:

- ☑ Check a few results, for example, the ones where parameter combinations are close to the parameter bounds, by means of the postvisualisation-tool (`vadere-postvis.jar`), a tool which allows to show the trajectories of the agents. This helps to identify agents that pass through walls, stuck agents, unnatural or odd trajectories.
- ☑ Compare the result of a manually started simulation in the Vadere graphical user interface (`vadere-gui.jar`) and the results of the SUQ controller. Both ways must lead to the same results.
- ☑ Check the scenario files that are created by the SUQ controller. The values for uncertain parameters in the original file should be replaced by the values that are given by the sample points.
- ☑ Introduce an auxiliary QoI. For example, one can easily check the simulation even for a few thousand sample points by recording the time between ingress and egress. The QoI is clearly arranged as vector. An evacuation time that is equal to -1 indicates that at least one agent did not reach the target in time either due to a too short simulation time or because the agent got stuck. Both cases should not occur.
- ☑ Likewise, the same principle applies to any other QoI. Screening the model response for implausible values can help to find errors in the code.

Indeed, some errors can be detected in the scenario files, but none of them concerns the application of the SUQ controller. For example, the thickness of some obstacles must be increased up to 1 m to prevent agents from walking through those walls. Unnatural trajectories are avoided by introducing intermediate targets as described in detail in Section 4.3.2. These issues are solved and one can use the SUQ controller without further ado.

Verification of MC Based Sensitivity Analysis

Since I implemented the MC based sensitivity analysis by myself, it is necessary to check the related lines of code for errors. For this purpose, I apply the Ishigami test function,

$$f(x_1, x_2, x_3) = \sin(x_1) + A \sin^2(x_2) + B x_3^4 \sin(x_1), \quad (3.6)$$

where the independent random variables x_1 , x_2 and x_3 are uniformly distributed in $]-\pi, \pi[$ and the sensitivity indices are known [13]. I follow the example in [34] and derive the sensitivity indices for $A = 7$ and $B = 0.05$ as listed in Table 3.3. These are the exact values which should be reached by the algorithm.

Table 3.3: First order and total sensitivity indices for the Ishigami model function: $D_1 = 1.9485$, $D_2 = 6.125$, $D_{13} = 0.8434$, $D = 8.9169$, $S_i = D_i/D$, $S_{12} = S_{23} = S_{123} = 0$

i	Fist order sensitivity index S_i	Total sensitivity index S_{T_i}
1	0.2185	$S_1 + S_{12} + S_{13} + S_{123} = 0.3131$
2	0.6869	$S_2 + S_{12} + S_{23} + S_{123} = 0.6869$
3	0	$S_3 + S_{13} + S_{23} + S_{123} = 0.0946$

Naturally, the MC simulation requires a large sample size to reach convergence. Moreover, since the process is random, each sample returns a different result which is why a single run might not be representative. Supposed that 10 samples ($n_r = 10$) of size n_s are sufficient, one obtains curves as in Figure 3.9 by performing the sensitivity analysis and taking the average of the 10 results. The plot shows the sensitivity indices according to [14]. They converge to the true values if the number of sample points is increased. One can conclude from this that the code is implemented correctly.

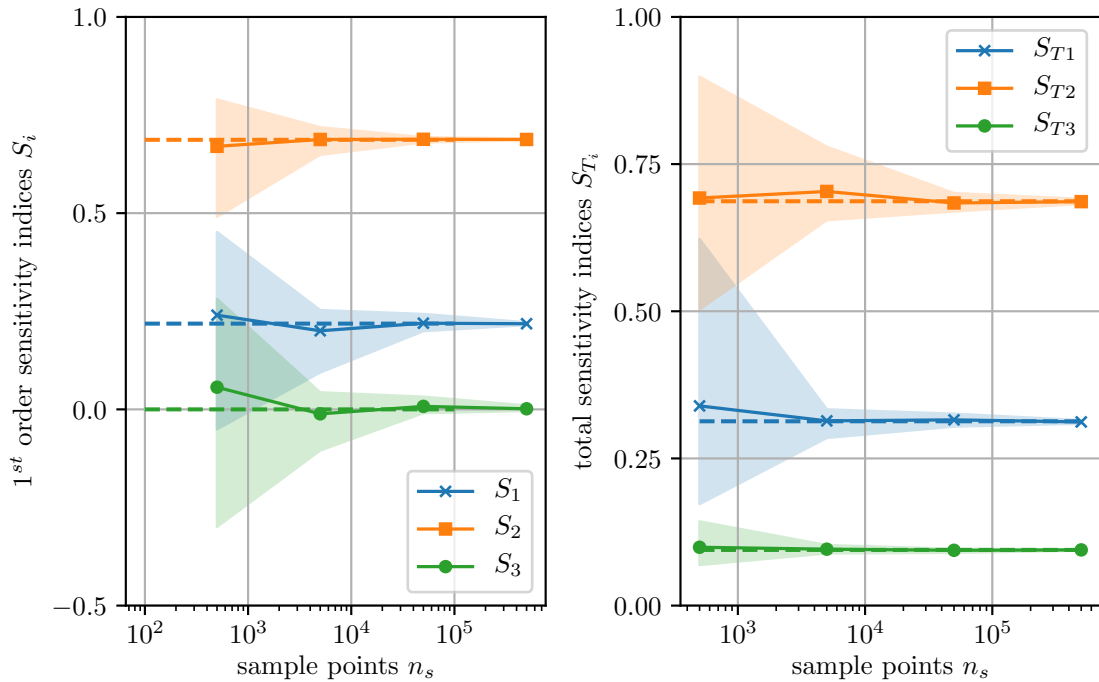


Figure 3.9: Convergence of sensitivity indices obtained with Jansen’s approach [14] for the Ishigami function: Dashed lines show the true values. Solid lines represent the mean and shaded areas represent the range between the minimum and maximum output of 10 repetitions. Each repetition uses n_s sample points.

The other approaches that are presented in Section 2.2.2 ([19, 21, 22]) approximate the sensitivity indices differently. They are implemented and tested as well by

means of the Ishigami function. A comparison of the results reveals that Jansen's approach [14] converges faster than the other methods. Furthermore, if the sample size is small, the interval of the minimum and maximum output, represented by shaded areas in Figure 3.9, is narrower than the interval resulting from the other approaches. For that reason, I use Jansen's formulas (see Eq. 2.34 and Eq. 2.35) to analyse the sensitivity of the OSM.

4 Results and Discussion

In this chapter, UQ results for two different scenarios are examined. Firstly, I discuss corridor scenarios to get acquainted with the applied method. Secondly, the method is transferred to a real world scenario, a parade through a city centre. The sections are structured similarly, as they start with an introduction of the scenario, then follow uncertainty and sensitivity analysis. For the large scale scenario, I dedicate one section to plausibility checks of the UQ results because there are no reference solutions or experimental data available to which simulation results could be compared. Each section concludes with a discussion of the UQ results.

4.1 Narrow Corridor Scenario

The narrow corridor scenario C-180-180-180 has already been introduced in Section 3.1 to adapt the UQ method to the problem. Now, I reconsider the UQ results obtained by applying the PC approach with third order polynomials and 40 sample points. Additionally, the MC solution is available with 2000 sample points for the uncertainty analysis and 5000 sample points for the sensitivity analysis. The input uncertainties (see Table 3.1) as well as the quantities mean density, density time series and evacuation time are defined in Section 3.1.1.

To give a first impression, Figure 4.1 shows the simulation for different time steps. Virtual pedestrians are represented by blue dots, which move from the source towards the target. Compare Figure 3.1 for a more detailed description of the topography. The pedestrian flow through the topography can be classified into the following five phases:

1. Time from start of the simulation until the first agent reaches the corridor;
2. Short phase in which the pedestrians can enter the corridor without hindrance and pass it with a relatively high distance between each other; this implicates a low density.
3. Agents accumulate in the waiting area in front of the entrance (area I) and 4-meter passage between entrance and corridor (area II), which implies a significant reduction in walking speed and a high density.

4. The number of accumulated agents in area II and the density inside the corridor has declined to such an extent that the agents can move relatively free again.
5. Time between the last agent has passed the measurement area A_D (dark red) and the end of the simulation;

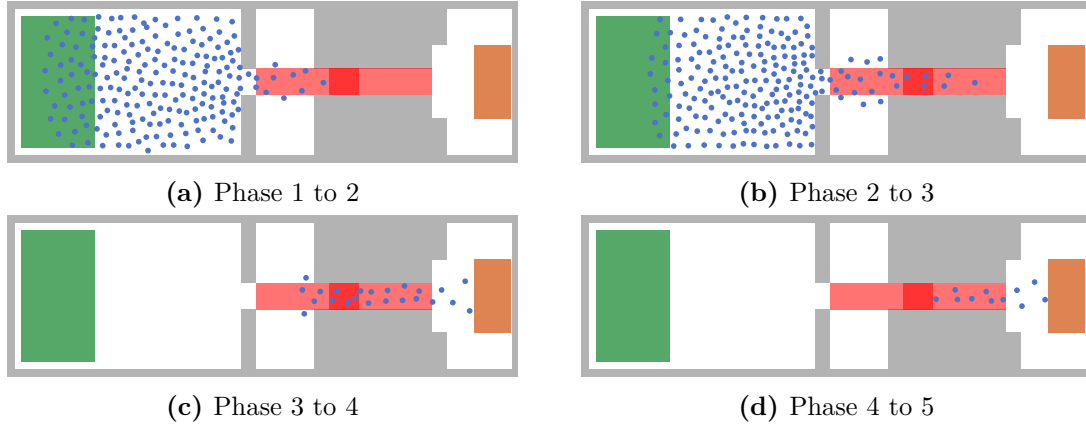


Figure 4.1: The pedestrian flow through the corridor scenario C-180-180-180 can be characterised by five phases. Figures (a) to (d) show approximately the time step during which the transitions between the phases take place.

The highest density occurs in front of a narrow passage such as the entry or the corridor during phase 3. However, the Voronoi density refers to measurement area A_D within the corridor. This definition follows the corridor experiment conducted by Zhang [45]. The five phases are now analysed with quantitative measures.

4.1.1 Uncertainty Analysis

Scalar Quantities of Interest

The quantities of interest evacuation time and mean Voronoi density are scalar and their uncertainty, quantified by mean and standard deviation, are summarized in Table 4.1. In agreement with the evaluation of the UQ method in Section 3.1, the values obtained with PC are close to the MC reference solution.

Table 4.1: The uncertainties of scalar QoI are obtained with Monte Carlo (MC) simulations and point collocation (PC) method for the corridor scenario.

	Method	
	MC ($n_s = 2000$)	PC ($p = 3, n_s = 40$)
Mean Voronoi density		
Mean	0.68 m^{-2}	0.68 m^{-2}
SD	0.17 m^{-2}	0.16 m^{-2}
Evacuation time		
Mean	81.94 s	81.91 s
SD	10.75 s	10.90 s

The UQ results can be understood as follows. The uncertainty is characterized by the standard deviation, which should be considered in relation to the mean value, since each QoI reaches different magnitudes. A larger standard deviation in the output states that at least one of the uncertain input parameters has a greater influence on the output. But the result does not state which parameter or which interaction between two or more parameters causes the uncertainty in the output. One can only guess that probably the number of agents contributes significantly to the uncertainty of the evacuation time.

The relatively great variation in the mean Voronoi density is probably more a matter of how the QoI is measured than how pedestrians pass a corridor in reality. The main reason for a seemingly high uncertain mean density might be that all time steps are taken into account when calculating the mean density. This also includes frames where all agents have reached the target already while the simulation is still running. That is, the density for those time steps is zero. These frames cannot simply be ignored because the exact duration of the simulation is unknown beforehand, and it is not possible to adapt this parameter for each sample point individually when applying UQ methods to the model. Therefore, one has to be careful when interpreting the mean and standard deviation of the mean Voronoi density. To overcome this issue, I decided to analyse the Voronoi density time series instead of its mean value.

Time Dependent Quantity of Interest

Time series can be analysed by applying the UQ method to the model output for each time step. The results for the Voronoi density are plotted in Figure 4.2. The standard deviation can be interpreted as a measure of uncertainty, whereas the mean value puts it into perspective. The plot shows both the reference solution obtained

with MC and the solution for the PC. The similarity of both curves confirms once again that PC yields good results also for time dependent quantities.

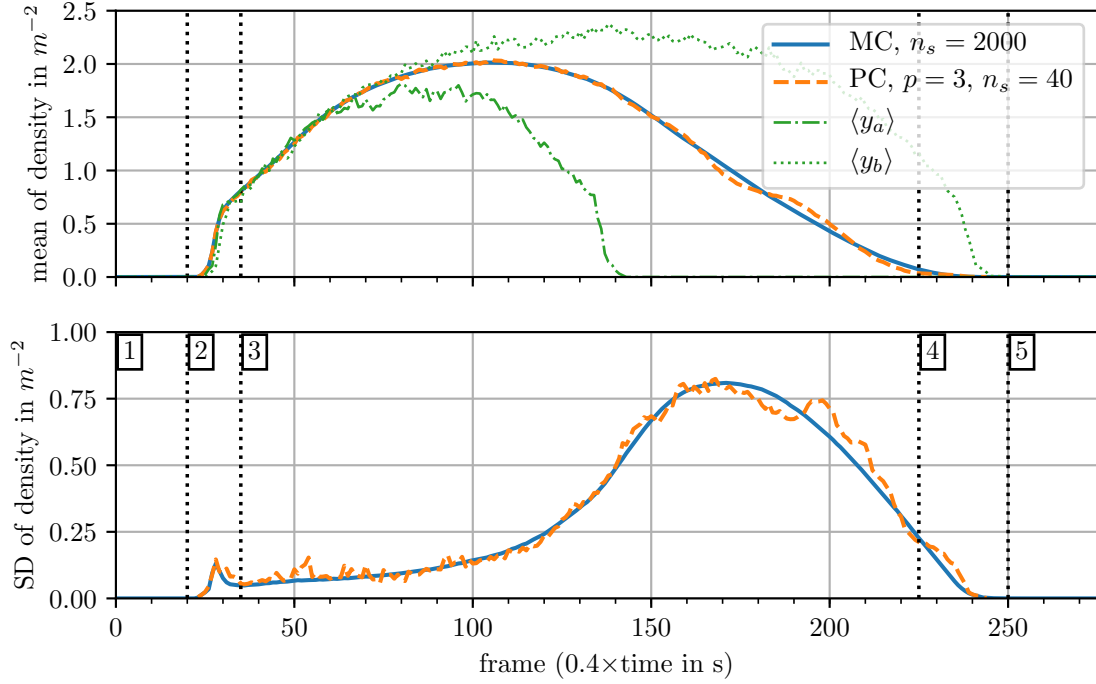


Figure 4.2: Uncertainty of QoI Voronoi density in measurement area A_D obtained by applying MC and PC method to the model response y : y is the average of $n_r = 10$ scenario runs. $\langle y_a \rangle$ and $\langle y_b \rangle$ is the averaged model response ($n_r = 10$) for parameter combination a ($\mu_v = 1.73 \frac{m}{s}$, $\sigma_v = 0.15 \frac{m}{s}$, $n = 100$) and b ($\mu_v = 1.37 \frac{m}{s}$, $\sigma_v = 0.15 \frac{m}{s}$, $n = 200$), respectively. Vertical dotted lines indicate the start of a new phase.

The mean of the density is zero as long as no agent has reached measurement area A_D . From that moment on, the mean density goes up within only a few frames (phase 2), and it continues to increase at a reduced rate during phase 3 until it reaches its maximum at around 2 agents per square meter after around 100 frames (40s). As the accumulation in front of the corridor disbands more and more, the density declines (phase 4).

Depending on the input parameters, i.e. depending on the evaluated sample point, the density time curve and particularly the part of phase 3 lasts longer. It takes more time to let a large crowd (represented by $\langle y_b \rangle$) pass a narrow corridor than a small group of people ($\langle y_a \rangle$) for a given flow rate. This also explains why the standard deviation increases until frame 170 where some simulations show a decreasing density already while the slope of other curves is still positive. The density approaches zero while the agents leave the corridor, and so does the variance of all density time series.

Besides the spawn number, there are other parameters that affect the behaviour of the QoI to some extent. For example, the small peak of the standard deviation at

approximately frame 30 (12 s), shortly after the first agent has passed the measurement area, can be traced back mainly to the input parameter mean free-flow speed and partly to standard deviation of free-flow speed. Since in every simulation the agents have a different mean free-flow speed, the first ones reach the measurement area after a different time. If one considers only a single time step near the peak, simulations with a faster mean free-flow speed yield higher density values while the ones with a slower free-flow speed have lower densities. Hence, the standard deviation of all simulations results in a local maximum. Next, I will study the strength of the individual impact of each uncertain parameter by means of sensitivity indices.

4.1.2 Sensitivity Analysis

Scalar Quantities of Interest

The sensitivity of the QoI to the input parameters is quantified by means of sensitivity indices summarized in Table 4.2. First order sensitivity indices represent the influence of a single parameter on the output, whereas the total sensitivity indices take into account the influence of a certain parameter and all interactions between this parameter and others. Again, the indices derived from the approximation (PCE) are in agreement with the ground truth (MC). This confirms the capability of the PC method. Values less than 0 and greater than 1 should not occur, and therefore indicate numerical instabilities or limitations of the method.

Table 4.2: Corridor scenario: First order sensitivity indices (S_i) and total sensitivity indices (S_{T_i}) quantify the of influence of parameters mean of free-flow speed ($i = 1$), SD of free-flow speed ($i = 2$) and number of agents ($i = 3$) on the QoI mean Voronoi density and evacuation time

	Method	
	MC ($n_s = 2000$)	PC ($p = 3, n_s = 40$)
Mean Voronoi density		
S_1	0.00	0.01
S_2	0.00	0.00
S_3	0.99	0.99
S_{T1}	0.01	0.01
S_{T2}	0.00	0.00
S_{T3}	1.00	0.99
Evacuation time		
S_1	0.02	0.03
S_2	0.00	0.00
S_3	0.97	0.97
S_{T1}	0.03	0.03
S_{T2}	0.00	0.00
S_{T3}	0.98	0.97

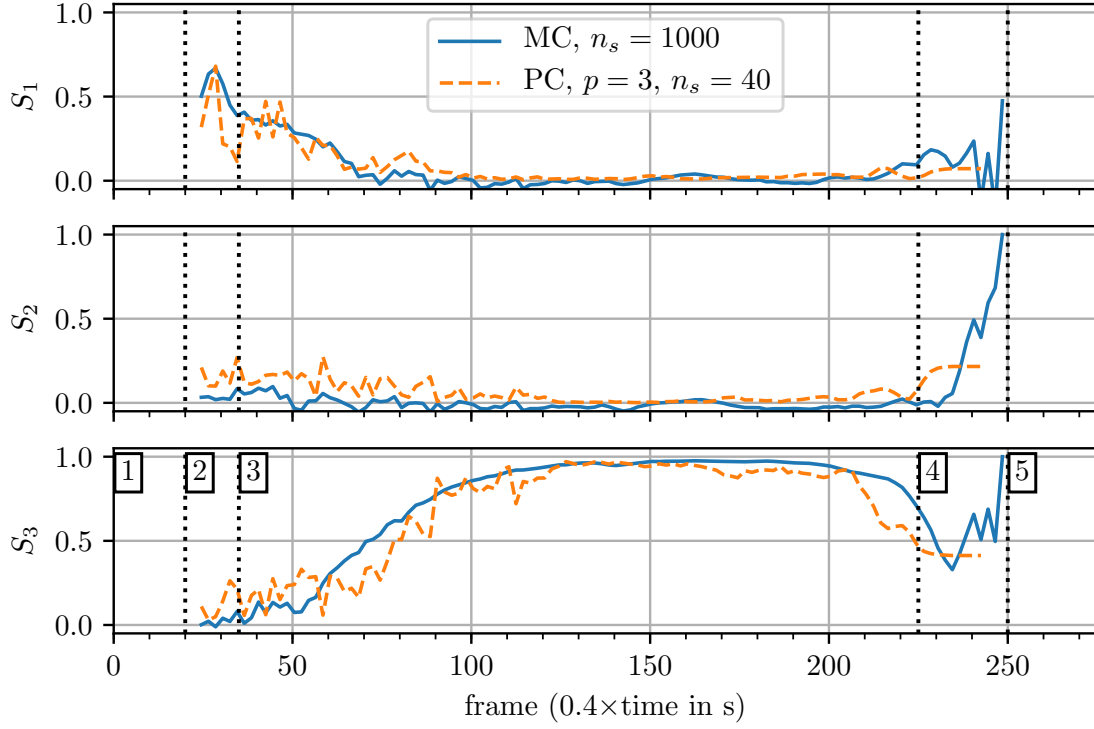
The sensitivity indices show that the number of agents is the most important parameter for both QoI, whereas the speed distribution parameters are insignificant. This is because most of the time the area of interest is packed and the agents are hindered, so they cannot reach their desired walking speed. Figure 4.1, which visualises the different phases, and Figure 4.2, which represents the uncertainty by curves for mean and standard deviation, illustrate this situation. The length of phase 3 and thus the simulation time is dominated by the number of agents.

The difference between total sensitivity indices S_{T_i} and first order sensitivity indices S_i expresses the contribution of interaction effects between the parameters to the output variance. The differences are small, so there is almost no interaction that plays a role.

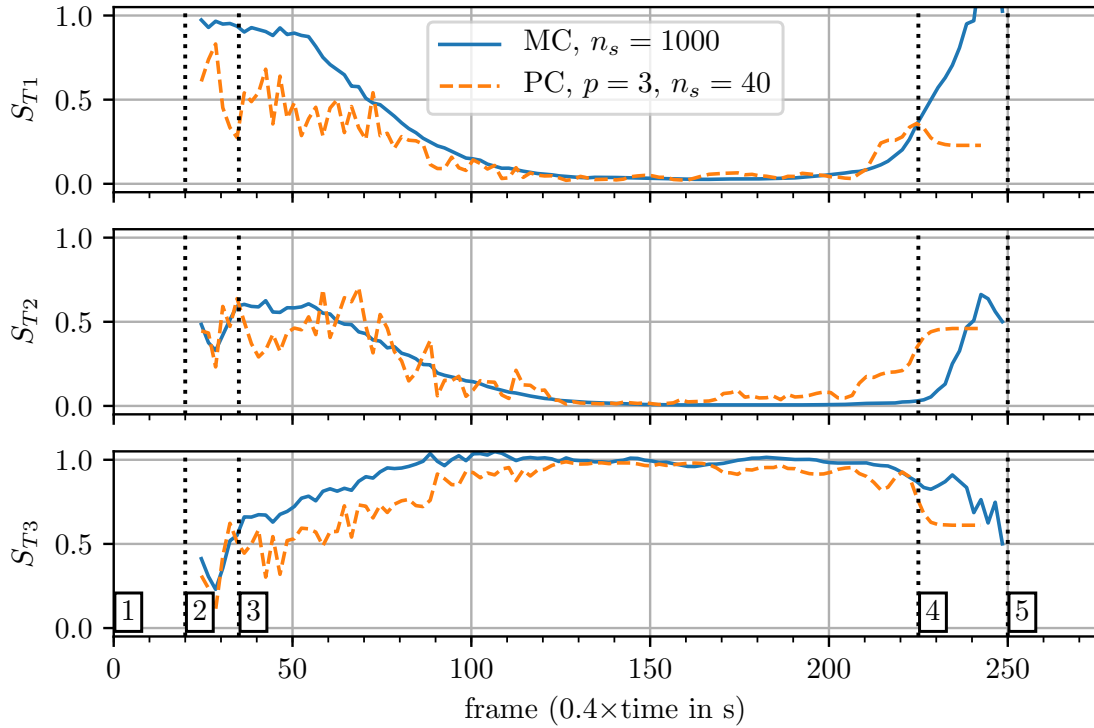
The sensitivity indices are consistent for both scalar quantities of interest. Since the information one can gain especially from the mean Voronoi density is limited, I now perform the sensitivity analysis for the Voronoi density time series.

Time Dependent Quantity of Interest

Firstly, I focus on the first order sensitivity indices of the Voronoi density time series, which express the importance of a single parameter. These are shown in Figure 4.3a. The graph reveals that the third parameter, number of agents, is the dominant one. After the first few agents have passed the measurement area, the sensitivity grows almost linearly to reach its maximum of around 0.95 for about 20 s. Then, the sensitivity decreases again because none of the parameters can influence the quantity of interest after all agents have passed the measurement area. The more agents want to pass the corridor, the longer it takes until all have reached the target. Therefore, a large spawn number causes a higher density over a longer period. Accordingly, the model and its output uncertainty stay sensitive to this parameter. The mean free-flow speed has a small impact. This variable is only important if all agents can move without being hindered considerably by other agents, which means there are no queues or accumulated agents (compare phase 2 and, to a certain degree, phase 4). The standard deviation of free-flow speed has no significant influence on the variance of the density.



(a) First order sensitivity indices S_i



(b) Total sensitivity indices S_{Ti}

Figure 4.3: Sensitivity indices obtained with MC and PC method ($n_r = 10$) quantify the influence of parameter mean free-flow speed (top, $i = 1$), SD of free-flow speed (middle, $i = 2$), number of agents (bottom, $i = 3$) on the QoI Voronoi density in measurement area A_D . Vertical dotted lines indicate approximately the start of a new phase. All curves are smoothed by taking the average of two frames.

The total sensitivity indices in Figure 4.3b, which account for all contributions to the variance of the response due to first order effects and interactions between the parameters, show slightly different curves. The total sensitivity of parameters mean and standard deviation of free-flow speed follow almost the same course as the regarding first order sensitivity indices. The only difference is that the standard deviation of free-flow speed is a little less relevant, especially during phase 2 and 4. However, throughout phase 2 to 4, the number of agents makes a major contribution to the output variance which leads to a total sensitivity index of almost 1 (if MC simulation is considered) for all time steps.

Note that during phase 1 and 5 the model response is 0. For this reason, the sensitivity indices are not considered. Moreover, the PC method tends to underestimate the total influence of parameter mean free-flow speed.

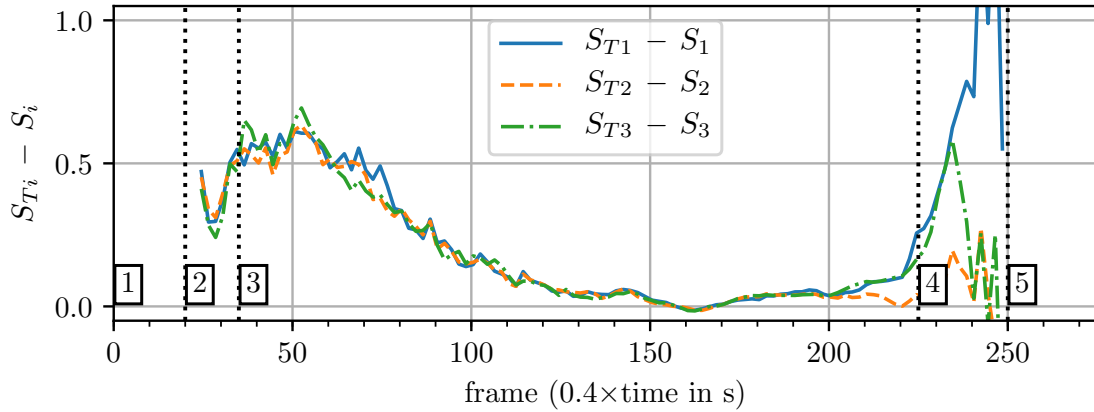


Figure 4.4: Interactions between the parameters ($S_{ij}, S_{ijk}, 1 \leq i < j < k$) affect the Voronoi density time series. The sensitivity indices are calculated from MC simulations with 5000 sample points ($n_r = 10$). Vertical dotted lines indicate the start of a new phase. All curves are smoothed by taking the average of two frames.

If S_{T_i} is subtracted from S_i , the interaction among the input parameters on the output variance is isolated. As one can derive from Figure 4.4, interactions are significant within phase 2, at the beginning of phase 3 and in phase 4. Higher order interactions could be calculated with greater computational effort. Because the agents accumulate in phase 3, there are little to none interaction effects because the free-flow speed parameters do not come into play. The fact that interactions occur, underlines the necessity to perform a global sensitivity analysis, for example, with Sobol' sensitivity indices. Local methods could not capture these effects so easily.

4.1.3 Discussion

Major results found in the uncertainty and sensitivity analysis are:

- The Voronoi density time series is superior to its mean value because it captures more information and is not biased.
- The simulation can be classified into five phases. The longest phase 3 is characterised by an accumulation of agents in front of and inside the corridor. Therefore, the agents cannot walk at free-flow speed.
- As a consequence, the free-flow parameters in the model have only minor effects.
- Independent of the QoI, the number of agents is the most important parameter because it is responsible for the accumulation in phase 3.

Analysing uncertainty and sensitivity in theory leaves the following question unanswered: What do the UQ results mean to practitioners, e.g. a civil engineer who is involved in planning the exit routes of a building? Of course, these findings cannot simply be transferred to cases of emergency because pedestrians would probably behave different in such a situation. Nevertheless, the scenario matches for example the daily rush hour in a railway building.

As expected, the main statement is that for reaching a fast evacuation of the waiting area in front of a relatively narrow corridor one has to make sure that only few persons want to pass it. This can be reached by planning more exit doors, so that the crowd can spread out and use different ones. An interesting example is the Colosseum in Rome, Italy. Although it could hold several ten thousand spectators, five minutes were enough to evacuate the building due to its efficient architecture [8]. However, the planners' hands are tied sometimes. Not always do they have influence on the amount of people who must pass the corridor or the number of exits to be built. The only option left might be to change the design of the corridor itself to some extent. The next section sheds light on the influence of the topography so as to find out whether the uncertainty and the sensitivity of the model response is different if the corridor is wider. Moreover, I study what happens if there is some narrow passage like a bottleneck at the end of the corridor.

4.2 Corridor Scenarios with Similar Topographies

In this section, I further analyse the corridor scenario but with modified topographies. QoI are evacuation time and Voronoi density time series as in the previous section (definitions can be found in Section 3.1.1). The uncertainty in the input parameters stay the same (see Table 3.1). So far, the PC method has returned reliable results which is why I now apply only the PC approach and save a many hours of computing time by leaving MC simulations behind. This method is applied to the average of 10 scenario runs. The topography is varied as shown in Figure 4.5.

The notation follows [45], that is hyphenated numbers denote the width of entrance b_{entr} , corridor b_{corr} and exit b_{exit} in centimetres, respectively (C- b_{entr} - b_{corr} - b_{exit} , see Section 3.1.1 for more details). Corridor is used as an umbrella term for the corridor with a narrow exit, also referred to as bottleneck, and corridors with a straight exit alike.

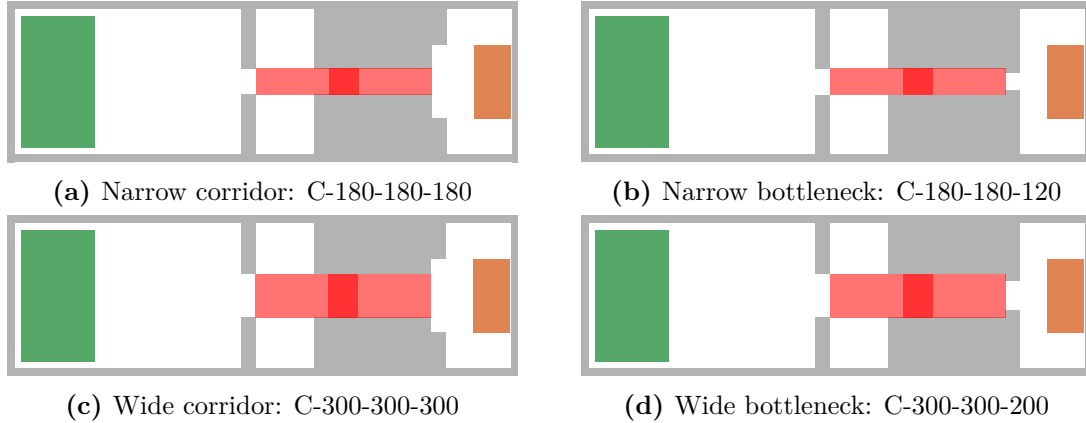


Figure 4.5: Corridor scenario with different topographies

4.2.1 Outline of Uncertainty and Sensitivity Analysis

The uncertainty and sensitivity analysis of a corridor scenario with different topographies can be summed up as follows. Mainly the parameter number of agents affects both uncertainties in the evacuation time and Voronoi density, regardless of the topography under consideration. However, this might change if the corridor were wide enough to let pass all agents at a time or at least in such a way that phase 2 and 4 were significantly longer than phase 3. As a result, the agents would not accumulate and the speed distribution parameters could become more important. This goes hand in hand with the previous analysis of the single corridor scenario in Section 4.1. Nevertheless, the uncertainty and sensitivity of the model does not change much as the topography is varied. A more detailed analysis is appended (see Section A.3).

4.2.2 Discussion

What can a practitioner infer from these uncertainty and sensitivity studies? As expected, a wider passage allows faster evacuations and lower densities inside the corridor at the same time. But it might be interesting that if the width of the exit is reduced, the evacuation time and the density increase not as much as if the whole passage were narrowed by the same ratio. Therefore, if I only take into account the evacuation time, a bottleneck is to be favoured marginally over a corridor that is as wide as the narrowest part of the bottleneck. But, depending on each

individual's perception, it might be better to avoid higher densities as they occur in the bottleneck. To state more about this aspect one should analyse the area where the density reaches the highest values.

4.3 Real World Scenario: Parade Through a City Centre

So far, I have considered rather academic examples. The aim is now to transfer the methods I got acquainted with in the previous sections to real world problems. Various models are being developed to simulate mass events. One aim is to take the right measures in advance based on these simulations. However, the application of UQ methods is relatively new in this field of study. Therefore, it suggests itself to analyse the uncertainty and sensitivity of a parade or protest march, for instance. The following example of a parade through a city centre is rather small, but it provides a basis for further studies of organized pedestrian movement in public spaces (OPMoPS) simulated with Vadere.

4.3.1 Adaptations to the Method

Both uncertainty and sensitivity analysis are performed by means of the PC method in accordance with the evaluation of the configurations of the UQ methods in Section 3.1.3. Therefore, the following analysis acts on the assumption that the PC method leads to correct results, although the object of investigation and quantities of interest are different now. It is not feasible to compare the UQ results to MC simulations, since this parade scenario is computationally too demanding to be evaluated several hundred or thousand times. This is why I supplement sanity checks for the UQ results.

The parameter space is only two dimensional, as will be explained in Section 4.3.2. The polynomial coefficients are determined on the basis of $n_s = 20$ sample points ($\alpha_{PC} = 2$). Because of the relatively high computational effort, the model is evaluated only three instead of ten times for each sample point ($n_r = 3$). The PC method is applied to the average of the output y to reduce the stochastic behaviour of the model. Moreover, a dynamic floor field is no longer feasible which is why a static floor field is used (see Table A.1).

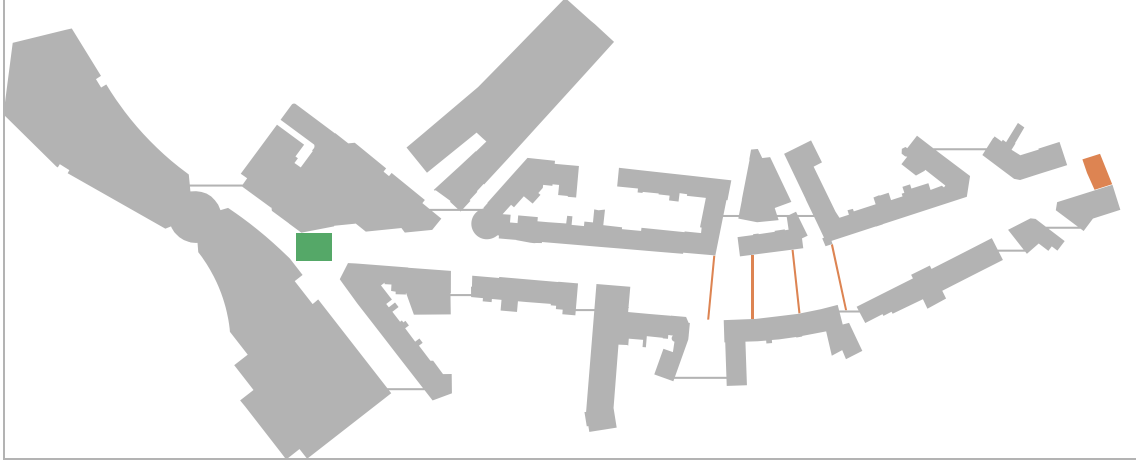
4.3.2 Object of Investigation

Topography

Objective of the uncertainty and sensitivity analysis is a fictitious, but analogue to a typical, parade through Richard Wagner Straße in Kaiserslautern, Germany. The

topography, which extends over an area of $228 \text{ m} \times 559 \text{ m}$, is shown in Figure 4.6a. I modify the original map as follows to improve the performance of the simulation:

- Four intermediate targets are introduced so that the agents follow realistic trajectories. Otherwise, they would not use the whole width of the street but try to take the shortest path as soon as entering the bend. This is a consequence of using a static instead of a dynamic floor field. However, intermediate targets in combination with a static floor field are functional and more efficient than the dynamic floor field.
- On closer inspection, it appears that the final target is a polygon. The slightly bent finish line makes sure that the agents approach it along realistic trajectories.
- The side streets are blocked by connecting the buildings with thin walls (grey lines). This is needed to calculate the length of the parade l_p as QoI by means of each pedestrian's potential towards the final target. Without those road blocks, the potential defined by the geodesic distance between agent and target might be biased due to short cuts the agents could possibly take. Besides, the blocked roads prevent agents from taking unrealistic routes. This is necessary because up to now, the simulator does not distinguish between the main road and side streets.



(a) Computer model



(b) Satellite image¹

Figure 4.6: The topography for the OPMoPS scenario is the northern part of Richard Wagner Straße (rotated 90° ccw.). The computerised version (a) of the real map (b) is slightly simplified. Agents start from the source (green rectangle), pass four intermediate targets (orange lines) before they reach the final target (orange polygon).

Quantities of Interest

As mentioned above, the QoI is the length of the parade l_p defined by the geodesic distance between agents in the rear and in the front of the parade. It is important to use the geodesic distance instead of the euclidean distance to assure that the length is measured correctly even for curved paths. The simulator provides the potential for each agent, which can be translated into the geodesic distance between agent and the final target. Thus the difference between the minimum and maximum potential of all agents

$$l_p = f_c \Delta\phi = f_c (\phi_{max} - \phi_{min}) \quad (4.1)$$

¹Richard Wagner Straße, Kaiserslautern, 49°26'35"N and 7°45'56"S. GeoBasis-DE/BKG and Google Earth. 2009. Retrieved 24 June 2020.

approximates the length of the parade l_p in m with conversion factor $f_c = 1$. The solution becomes more representative by taking the average of the 10 greatest and 10 smallest values of potential ϕ instead of ϕ_{max} and ϕ_{min} , respectively. This is reasonable because in particular the front of the parade is not so dense due to different free-flow speeds. Thus, one would scarcely consider the first agent to be part of a parade.

Since the QoI length of the parade is rather unfamiliar, a second QoI is introduced for checking purposes: The duration of the parade t_p . This is equivalent to the evacuation time for the previously analysed scenarios, so it has been used in a different context already. Moreover, it is more intuitive than the length of the parade which makes it easier to estimate its minimum, mean and maximum so that one can use it for sanity checks.

Uncertain Parameters

Uncertain input parameters are standard deviation of free-flow speed and the number of agents as summarized in Table 4.3. In accordance with previous analyses and calibrations of the OPMoPS scenario (compare demo files in [39]), mean free-flow speed is fixed to $\mu_v = 0.6 \text{ m s}^{-1}$ and not considered uncertain because its influence on the output variance is negligible.

Table 4.3: OPMoPS scenario: The uncertainty in the input parameters is described by uniform distributions.

Parameter	Symbol	Distribution	Unit
SD of free-flow speed	σ_v	$\mathcal{U}(0.05, 0.10)$	m s^{-1}
Number of agents	n	$\mathcal{U}(400, 1200)$	

The interval for parameter standard deviation of free-flow speed follows the variation given in [40]: $\sigma_v = 0.19 \cdot \mu_v \approx 0.1 \text{ m s}^{-1}$. Compared to other scenarios, this value appears to be quite low, but it is reasonable because the speed among participants of a parade is more homogeneous than among pedestrians observed under ordinary conditions. Furthermore, the standard deviation of free-flow speed is assumed to reach only 0.05 m s^{-1} (compare demo files in [39]). That is why an uncertainty in the range of $[0.05, 0.15] \text{ m s}^{-1}$ seems adequate. However, this choice is misleading because effectively it does not result in a sufficiently large variation of the speed distribution but in an almost constant effective standard deviation of the agents' free-flow velocity $\sigma_{v,i} \approx 0.3 \text{ m s}^{-1}$.

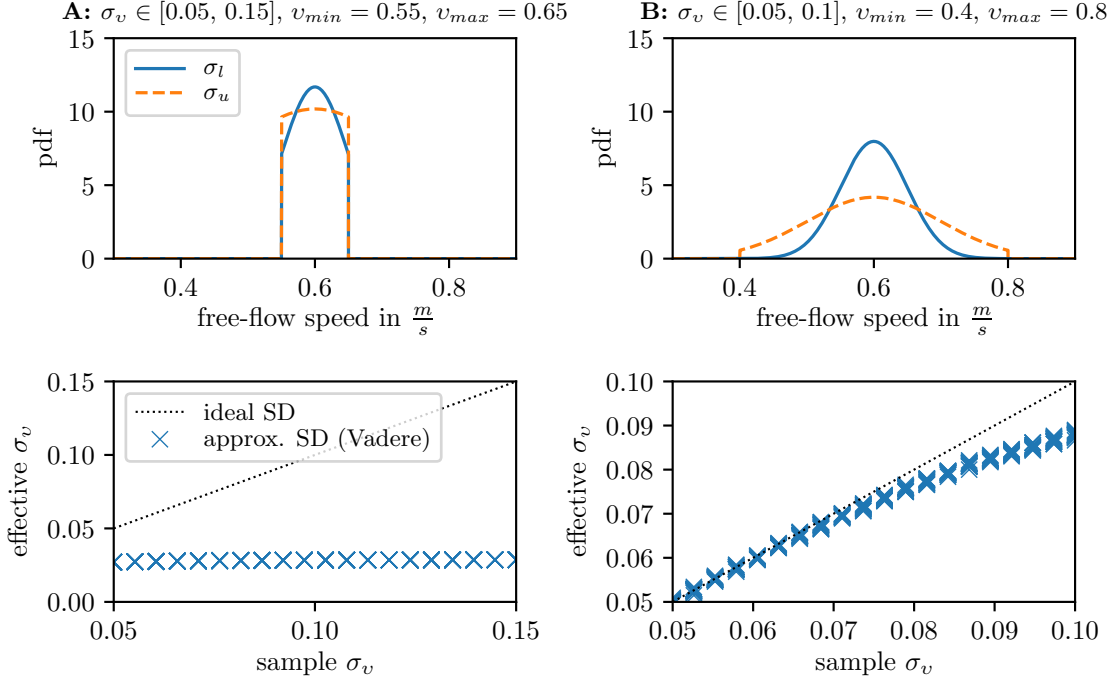


Figure 4.7: Normal distributions for boundary values of the uncertain parameter standard deviation of free-flow speed $\sigma_v = [\sigma_l, \sigma_u]$: Distributions are truncated at v_{min} and v_{max} , which leads to a different effective σ_v when generating speed distributions from 20 equidistant sample points in Vadere. $\mu_v = 0.6 \frac{m}{s}$ for both cases A (left) and B (right). σ , v_{min} , and v_{max} are given in $\frac{m}{s}$.

This becomes more clear by visualising the probability density functions for the truncated speed distributions in Figure 4.7. Case A (left) shows that the speed distributions within the range of uncertainty barely vary. Therefore, it would make almost no difference if the parameter were fixed or not. Although Vadere draws a speed distribution for each realisation, it does not matter which standard deviation of free-flow speed is given initially. The resulting standard deviation of the speed values assigned to each agent is the same for each sample point, as shown in the bottom graph. From a sensitivity analysis based on parameter set A, one could easily draw the false conclusion that input parameter standard deviation of free-flow speed does not contribute to the variance of the output and thus is not important. This can be prevented by changing the values v_{min} and v_{max} at which the speed distribution is truncated. Instead of $v_{min} = 0.55 \text{ m s}^{-1}$ and $v_{max} = 0.65 \text{ m s}^{-1}$ (case A), as proposed in [38], the speed distribution in case B is truncated at 0.4 m s^{-1} and 0.8 m s^{-1} , respectively. Moreover, the uncertainty of input parameter σ_v is defined in the range $[0.05, 0.10] \text{ m s}^{-1}$. From that follows an effective variation of the free-flow speed between 0.05 m s^{-1} and approximately 0.09 m s^{-1} as shown in the bottom right graph.

The second uncertain input parameter is the number of agents, which ranges between 400 and 1200 pedestrians. This represents a rather small parade, since the

number of participants of a parade can also shoot up to 1 million people or more [8]. If large crowds are considered, also the topography, and therefore, the distance between source and target, has to extend over a greater area. Otherwise, it might happen that some agents are still being spawned while others have reached the target already. As a consequence, the length of the parade is given by the constant distance between source and target. Obviously, such simulations are meaningless.

4.3.3 Uncertainty Analysis

The uncertainty of the length of the parade is quantified by mean and standard deviation in Figure 4.8. The curves are only relevant in between the vertical dotted lines, i.e. between frame 100 and 1200. I neglect all frames before frame 100 to avoid artefacts from the source. Figure 4.9a and 4.9b show the difference between the unrealistic starting position in the first frame and a more realistic one after 100 frames. Analogously, frame 1200 to 1400 are omitted because in some simulations the agents are still about to reach the target, whereas in others the first few agents have reached it already. In the latter case, the model response is meaningless and thus distort the statistics.

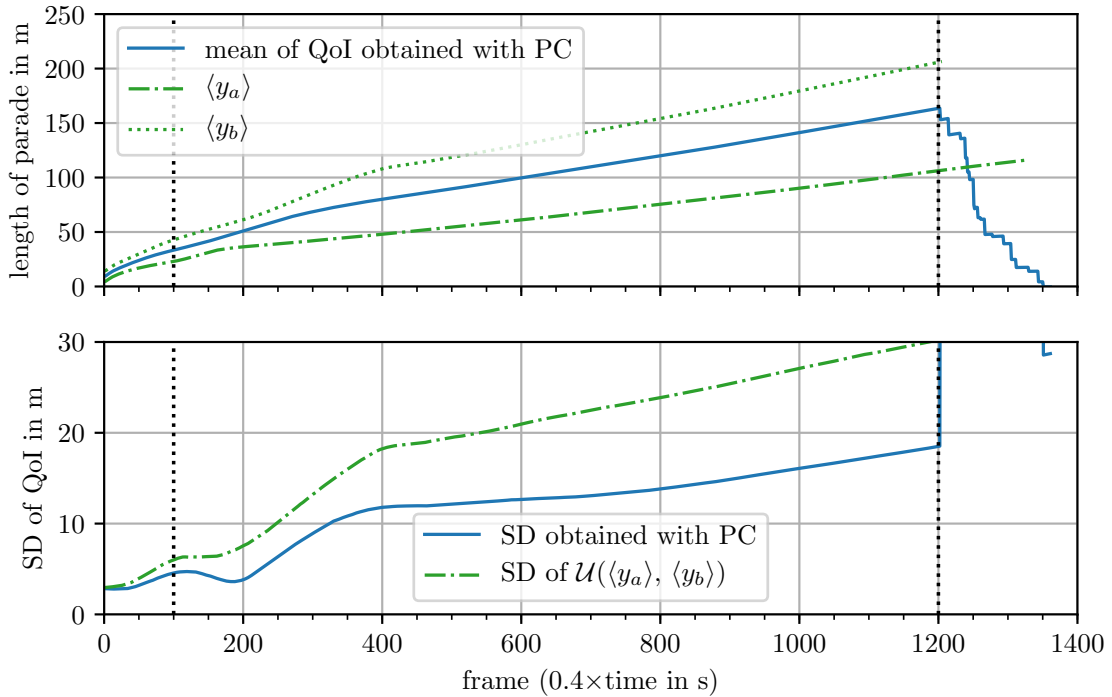


Figure 4.8: Uncertainty of QoI length of the parade: The PC method with third order polynomials ($p = 3$) is applied to the model response y for $n_s = 20$ samples. y is the average of $n_r = 3$ scenario runs. $\langle y_a \rangle$ and $\langle y_b \rangle$ is the averaged model response ($n_r = 3$) for parameter combination a ($\sigma_v = 0.05 \frac{m}{s}$, $n = 400$) and b ($\sigma_v = 0.10 \frac{m}{s}$, $n = 1200$), respectively. The combinations a and b represent two boundaries (vertices) of the input parameter space. The results are only valid between the two vertical dotted lines.

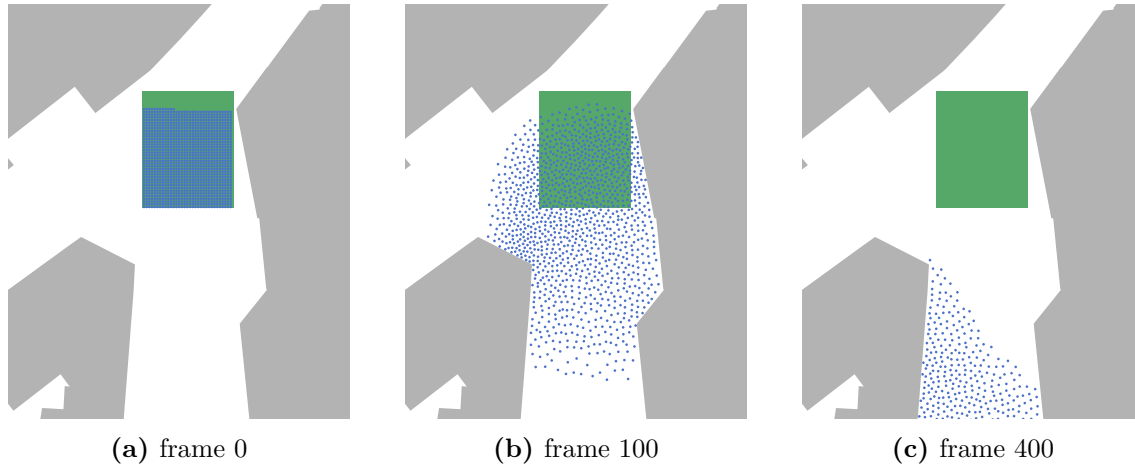


Figure 4.9: First few hundred frames after start of the simulation of the OPMoPS scenario with $n = 1200$ agents.

For now, I focus only on the curves for mean and standard deviation obtained with PC in Figure 4.8. I get back to $\langle y_a \rangle, \langle y_b \rangle$ and standard deviation of $\mathcal{U}(\langle y_a \rangle, \langle y_b \rangle)$ in Section 4.3.5 for checking purposes. In the beginning, approximately between frame 100 and 300, the slope of the mean curve is slightly greater than in the subsequent period. The reason is that initially the crowd is spread over the area of the T junction, but the street where the parade starts is narrower. Agents at the rear have to wait while the ones in the front start marching down Richard Wagner Straße already. Due to this difference in speed, the parade stretches faster. This phase is shorter for simulations with fewer agents, and for simulations with many agents this phase is longer. Frame 200 and 400 mark the time step at which all agents at the rear start walking in simulations with few and many agents, respectively (see Figure 4.9c). For this reason the uncertainty, represented by the standard deviation of l_p , slightly decreases until frame 200. From that moment on, it follows a steep upward trajectory until frame 400. As soon as all agents move approximately at mean free-flow speed, both mean and standard deviation of the QoI rise steadily until frame 1200. One could guess that the length of the parade is less determined by the number of agents but more and more by the variation of the pedestrians' desired walking speed, i.e. by parameter standard deviation of free-flow speed. The sensitivity analysis in the next section will break down the influence of each parameter.

4.3.4 Sensitivity Analysis

The sensitivity analysis determines how parameters standard deviation of free-flow speed and number of agents influence the length of the parade l_p . Figure 4.10 shows the first order and total sensitivity indices. As explained in the previous section, only the period between the vertical dotted lines is relevant.

Because the curves of first order and total sensitivity indices are very similar, there is little interaction between the parameters that would affect the variance of the output. Furthermore, the comparison of both plots can be regarded as sanity check. The sum of all indices for two dimensional problems must equal 1 (see Eq. 2.29):

$$S_1 + S_2 + S_{12} = 1 \quad (4.2)$$

$$S_2 + S_{12} = S_{T2} \quad (4.3)$$

S_1 and S_{T2} are known, so the left hand side of Equation 4.2 can be calculated with the aid of Equation 4.3. The condition holds true for all time steps, as plotted in the top graph of Figure 4.10. This also becomes clear by comparing the mirrored shape of S_1 and S_{T2} or S_2 and S_{T1} .

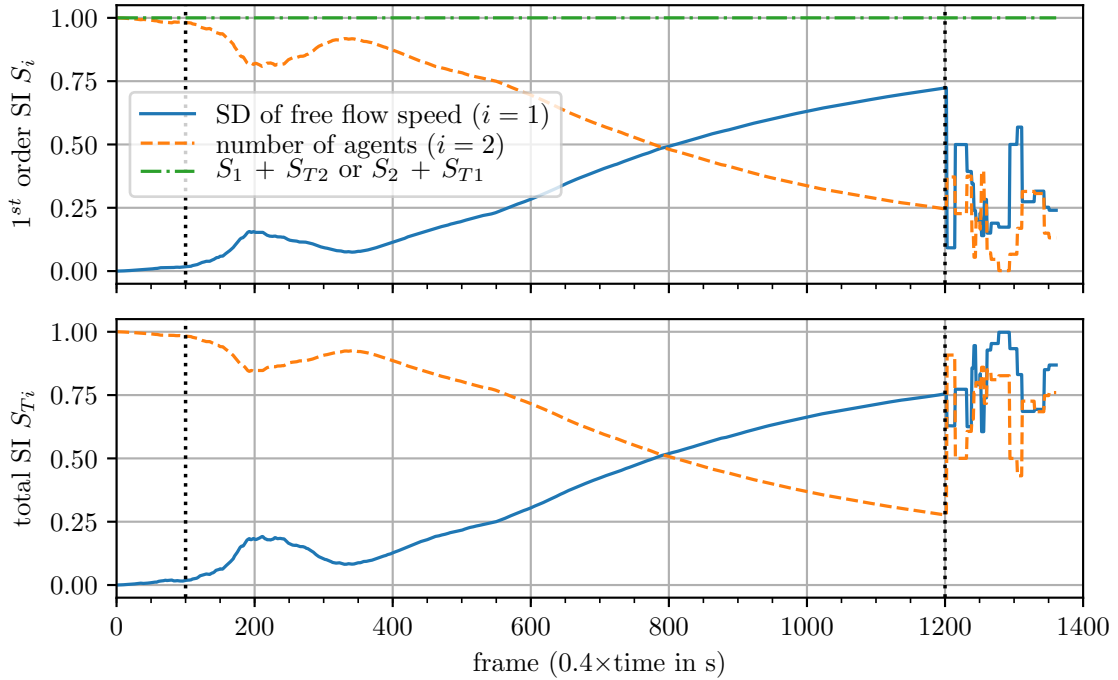


Figure 4.10: Sensitivity of the OPMoPS scenario: First order sensitivity indices (top) and total sensitivity indices (bottom) are obtained with PC method ($p = 3$, $n_s = 20$, $n_r = 3$). They quantify the influence of uncertain input parameters standard deviation of free-flow speed ($i = 1$) and number of agents ($i = 2$) on the QoI length of the parade l_p with respect to time. The results are only valid between the two vertical dotted lines.

Right after the agents have been spawned and start marching down the street, solely the number of agents affects the QoI. As soon as they are set in motion, also the speed parameter contributes increasingly to the output variance. However, the effect of parameter standard deviation of free-flow speed diminishes again between frame 200 and circa 350. This is linked to the topography and confirms the hypothesis

stated within the scope of the uncertainty analysis in Section 4.3.3: The time until all agents move at similar speed, which has a major effect on the output uncertainty, depends on how many agents are about to enter the relatively narrow street.

As time passes, the influence of parameter standard deviation of free-flow speed S_1 rises, whereas the importance of the number of agents S_2 declines steadily. The intersection point marks the time step where both input parameters have the same impact on the variance of the output. Generally, this interchange of the sensitivity indices would occur also in similar scenarios, but the moment when the relation turns is rather specific. Also Kurtc [17] showed this change in the importance of the two uncertain parameters in her report. Since she analysed a quite small group of agents, the intersection point is at the very beginning of the simulation.

In case of an infinitely long route, the time between ingress and egress t_p would increase because each agent has a different speed. In contrast, the number of agents would not affect the output variance for $t \rightarrow \infty$. How fast S_1 converges to 1 and S_2 to 0 depends mainly on how different the minimum and maximum speed among the agents is. This becomes obvious if the equation that describes the length of the parade (see Eq. 4.1) is rewritten, acting on the assumption that the first (last) agent is the fastest (slowest). QoI length of the parade l_p is now expressed by the difference between the minimum and maximum potential with respect to time:

$$\begin{aligned}\Delta\phi(t) &= \phi_n(t) - \phi_1(t) \\ &= \left(\phi_{n,t_0} - \int v_n(t) dt\right) - \left(\phi_{1,t_0} - \int v_1(t) dt\right) \\ &= \Delta\phi_{t_0} + \int v_1(t) - v_n(t) dt,\end{aligned}\tag{4.4}$$

where ϕ_{t_0} is the initial potential and v the agents' speed. Index 1 and n denote the first and last agent, respectively. The term $\int v_1(t) - v_n(t) dt$ depends on parameter standard deviation of free-flow speed and on the boundaries at which the speed distribution is truncated. This term becomes more important over time, while the influence of $\Delta\phi_{t_0}$, which equals the initial length of the parade l_{p,t_0} , vanishes because it is not time dependent. $\Delta\phi_{t_0}$ is determined by the width of the parade, density and number of agents. Hence, if the magnitude of parameter number of agents were greater, the intersection point in Figure 4.10 would be shifted towards the right. It remains unanswered whether the corresponding time step would range within the bounds of the considered time interval or not.

4.3.5 Plausibility Checks

The time between ingress and egress t_p as QoI is used for sanity checks because it can easily be compared to manually computed quantities. The route is approximately

400 m long, thus an average agent needs $\frac{400 \text{ m}}{0.6 \text{ m s}^{-1}} \approx 670 \text{ s}$. In fact, mean of QoI is $t_p = 1011 \text{ s}$. Obviously, there is a substantial difference, but one has to keep in mind that the last agent does not start walking before frame 300 (120 s) on average. Additionally, it cannot move at desired speed due to slower agents ahead which is why a value between minimum speed 0.4 m s^{-1} and mean speed 0.6 m s^{-1} , e.g. 0.45 m s^{-1} , is more realistic. If these considerations are taken into account, the UQ results bear comparison with the manually estimated duration of the parade:

$$\frac{400 \text{ m}}{0.45 \text{ m s}^{-1}} \approx 1011 \text{ s} - 120 \text{ s} \approx 890 \text{ s} \quad (4.5)$$

Similarly to this approach, also the length of the parade can be calculated manually. The maximum length should not be greater than

$$\begin{aligned} l_p &= \frac{\Delta t_1}{v_1} - \frac{\Delta t_n}{v_n} \\ &= \frac{(1200 - 100) \cdot 0.4 \text{ s}}{0.8 \text{ m s}^{-1}} - \frac{(1200 - 300) \cdot 0.4 \text{ s}}{0.4 \text{ m s}^{-1}} = 224 \text{ m}, \end{aligned} \quad (4.6)$$

where Δt is the period during which the agent moves at free-flow speed v . Index 1 and n denote the first and last agent, respectively. The mean of the length of the parade l_p in Figure 4.8 is about 160 m and therefore meets the condition. However, this is a coarse estimate and represents only the worst case.

A second way to test the UQ results for plausibility is to compare them to the model response resulting from certain sample points. Figure 4.8 shows the response for combinations of the input parameters that are expected to yield the shortest and longest parade. The mean value is in the range of both model responses at any time step. Therefore, the PC method presumably returns correct mean values. Assumed that the model responses for different sample points follow a uniform distribution within the bounds of $\langle y_a \rangle$ and $\langle y_b \rangle$, which errs on the side of conservatism, the standard deviation should lead to greater values than the ones obtained with the PC method. Thus, the comparison of both curves in the bottom graph in Figure 4.8 reveals that the UQ method returns reasonable results.

Probably the best method to validate the simulation is to compare the output to experimental data. However, this kind of data is not available. At least, I take a look at the real topography in the next section to better estimate the relevance of the simulation to reality.

4.3.6 Discussion and Relevance to Reality

In contrast to the computer model, there are many obstacles such as trees, parking lanes, etc. in Richard Wagner Straße (compare Figure 4.6b). Depending on the

pedestrian's constitution, even road kerbs might have an effect on some individuals and their movement. If the parade targets a certain group that is affected by such barriers, it might be that not only a few but the whole crowd behaves differently. Many other influencing factors are imaginable. So even though the model seems quite specific, it rather represents an average scenario. Generally, the insights gained in the previous sections can be transferred to other situations and help to evaluate or develop new crowd-management strategies.

The topography can influence the length of the parade significantly. For instance, the T junction where the source is placed and the parade starts, or more in general a bottleneck, can result in both stretching and compressing the parade. Both extreme cases and, with respect to the pedestrians' safety, in particular high densities should be avoided. With this in mind, it makes sense to guide a parade by means of physical barriers. This can allow for better flow, and, as a side effect, confrontations with spectators are reduced. Such physical barriers could be included in the model to find out whether fencing is necessary or not. Additionally, spectators could be placed either inside or outside designated viewing areas, but this raises the question how the spectators' behaviour should be modelled.

Secondly, the variation of the desired speed among the participants of a parade determines the length of the parade in the long run. To keep the parade compact, one has to make sure that the standard deviation of the pedestrians' speed is as small as possible. Usually, this is achieved by floats which limit or control the speed. Also broad banners have a similar effect because they reduce permeability in the direction of motion. They can be placed in front, in the middle, or behind the pedestrians. Which position is best depends on the speed distribution. In the present case, it would be more efficient to place a float in front of the parade because the density is lower in the front part and some fast agents outrun the others. Another option that might support a steady pace is playing music at a certain tempo. This idea is inspired by military parades or music play lists for runners created by commercial streaming services. If the stride length is constant and equal among the participants, and if the listeners are in strict time, it is easier to move forward at constant speed.

If these measures are successful, the remaining important parameter is the number of pedestrians. This is not always easy to control, although event managers typically have to apply for a permit, which defines a maximum number of participants. In combination with this parameter, there are other aspects that could also affect the length of a parade. Recently, as the concept of social distancing becomes more important, participants in any mass event are obliged to keep a certain distance to each other. This leads to lower densities and the crowd covers a larger area. Further studies could examine how the extension of the personal space influences the variance of the length of the parade.

5 Conclusions and Outlook

This thesis aimed to quantify the uncertainty and sensitivity of microscopic crowd simulations performed with the optimal steps model in Vadere [15, 39]. The optimal steps model suffers from uncertain input parameters, which affect the output. To quantify the influence of some parameters, these were treated as random variables with defined density functions and thus propagated through the model. The resulting distribution of the output can be approximated by polynomial chaos expansions from which I calculated the statistical measures mean and standard deviation to characterise the uncertainty in the model response. The sensitivity of the output was defined by sensitivity indices derived from the Sobol' decomposition of the polynomial chaos expansion. These analyses were carried out with the Chaospy toolbox [3], a state-of-the-art Python library for uncertainty quantification.

I chose a specific set-up of the uncertainty quantification method based on a precedent comparison of polynomial chaos expansions computed with the point collocation method and the pseudo-spectral approach. Those were varied in terms of sample size and polynomial degree to apply different configurations to a small example scenario in which agents have to pass a corridor.

A quantitative comparison to a reference solution obtained with Monte Carlo showed that both the point collocation method and pseudo-spectral approach with third or fourth order polynomials yield similarly good results. Independent of the method and its configuration, the uncertainty quantification results are more reliable when calculated from an averaged model response. That is, the model is evaluated repeatedly for exactly the same sample and the output is averaged.

In the final rating of the two approaches with different set-ups, I opted for the point collocation method with polynomials of degree 3 because the resulting approximations to the output distribution are more stable than the ones generated with the pseudo-spectral approach. Besides, the point collocation method with third order polynomials is not as computationally demanding as the other configurations. This is due to two reasons: Firstly, a lower polynomial order leads to a shorter runtime when determining the coefficients of the polynomial chaos expansion. Secondly, the point collocation method requires fewer sample points and hence fewer model evaluations. However, one must define the sample size carefully. If the number of

sample points is at least two times greater than required, the series of polynomials approximates the actual distribution of the model response considerably better. As a consequence, also the uncertainty quantification results closely match the ground truth obtained with Monte Carlo.

In general, the comparison to Monte Carlo simulations serves as proof of concept, since it showed that the polynomial chaos expansion is capable of estimating the uncertainty and sensitivity of the corridor scenario sufficiently accurate.

Therefore, these findings were transferred to a large scale scenario, a parade through a city centre. I assumed the parameters standard deviation of free-flow speed and number of agents to follow uniform distributions and analysed their effect on the length of the parade over simulation time. Most interestingly, the sensitivity indices revealed that the influence of the two uncertain input parameters interchange with time. At the beginning, the number of agents determines the length of the parade, while the standard deviation of free-flow speed is of little importance. As the parade moves towards the target, the number of agents contributes barely to the output variance, whereas the standard deviation of free-flow speed becomes the dominating factor. The moment when the sensitivity of the model response is equal for both input parameters depends on the input uncertainties. Smaller crowds or a greater standard deviation of the free-flow speed shifts the intersection towards an earlier point in time. More agents or less variation in the speed lead to the opposite effect. The uncertainty of the length of the parade accumulates over simulation time and therefore increases steadily unless the topography has a great impact, for example due to bottlenecks. Although the scenario, input parameters and the quantity of interest differs from the precedent analysis of the corridor scenario, the point collocation method with polynomials of third order is still applicable and yields reasonable results. I confirmed the reliability of the uncertainty quantification results by various plausibility checks for lack of experimental data.

Finally, I discussed the significance of the uncertainty quantification results to reality. According to expectations, any measure that reduces the variation of all pedestrians' velocities is effective to control the length of the parade, whereas the number of participants is less important. To achieve that the participants of a parade walk at the same speed, one could use physical barriers such as banners, floats, or other pacemakers in front, in between or at the rear of the parade. Alternatively, playing music aloud could have a positive, i.e. unifying, impact on the pedestrians' desired speed. These are rather general interpretations, which might apply also to similar parades, even though the topography of the scenario is quite specific. However, to draw more individual conclusions, the level of detail must be increased. For example, this concerns the topography, which should include obstacles and more individual properties of the participants of the parade as well. The

important parameter standard deviation of free-flow speed might highly depend on the participants' characteristics.

Further studies could address this issue by performing inverse uncertainty quantification to improve the quality of the input uncertainties before using forward techniques. Alternatively, one could try to enhance the probability density functions of the input parameters based on experimental data. Defining the input uncertainties is a delicate task. They must be realistic, but they should also lead to a sufficiently large variation in the simulation to obtain meaningful uncertainty quantification results. Especially the speed distribution parameters are crucial because they are processed in the optimal steps model in a manner that is not obvious to the common user. Changing the speed distribution parameters requires some knowledge about the internal algorithms.

As a next step, one could identify more parameters that potentially play a crucial role. For example, as the concept of social distancing emerges, it would be interesting to vary the extent of the pedestrians' personal space. By analysing additional parameters, one might detect interaction effects between parameters. Note that the polynomial chaos expansion requires stochastically independent input parameters. Otherwise, one must use techniques to transform dependent variables into independent ones.

From a methodological point of view, the application of the polynomial chaos expansion and the Monte Carlo simulations could be improved by employing more efficient sampling techniques, such as Latin Hyper Cube Sampling or Importance Sampling. One can further reduce the computational effort when generating the model response by the aid of surrogate models. Faster techniques would allow to model more complicated scenarios, for example larger crowds and topographies.

Identifying other crucial and more realistic scenarios is surely among the next tasks. Up to now, the optimal steps model in *Vadere* is not able to model mass events such as the Love Parade in all its complexity. However, that sort of all-encompassing simulation seems unlikely in order to derive the right crowd management strategies. As demonstrated in this thesis, less complicated scenarios can be helpful to understand the importance of certain parameters already. Therefore, this thesis and the developed code provide a foundation for further studies and the application of uncertainty quantification methods to the optimal steps model.

Bibliography

- [1] Felix Dietrich et al. “Fast and flexible uncertainty quantification through a data-driven surrogate model”. In: *International Journal for Uncertainty Quantification* 8 (2 2018), pp. 175–192. DOI: 10.1615/Int.J.UncertaintyQuantification.2018021975.
- [2] Jonathan Feinberg. *Chaospy - Toolbox for performing uncertainty quantification*. online. <https://github.com/jonathf/chaospy>. 2016.
- [3] Jonathan Feinberg and Hans Petter Langtangen. “Chaospy: An open source tool for designing methods of uncertainty quantification”. In: *Journal of Computational Science* Volume 11 (2015), pp. 46–57. DOI: 10.1016/j.jocs.2015.08.008.
- [4] Peter G. Gipps and Bertil S. Marksjö. “A micro-simulation model for pedestrian flows”. In: *Mathematics and Computers in Simulation* 27.2–3 (1985), pp. 95–105. DOI: 10.1016/0378-4754(85)90027-8.
- [5] Marion Gödel, Rainer Fischer, and Gerta Köster. “Applying Bayesian inversion with Markov Chain Monte Carlo to Pedestrian Dynamics”. In: *UNCECOMP 2019, 3rd ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering*. 2019. DOI: 10.7712/120219.6322.18561.
- [6] Marion Gödel, Rainer Fischer, and Gerta Köster. “Sensitivity Analysis for Microscopic Crowd Simulation”. In: *Methods and Applications of Uncertainty Quantification in Engineering and Science* 13 (7 2020): *Methods and Applications of Uncertainty Quantification in Engineering and Science*. DOI: 10.3390/a13070162.
- [7] Dirk Helbing and Péter Molnár. “Social Force Model for pedestrian dynamics”. In: *Physical Review E* 51.5 (1995), pp. 4282–4286. DOI: 10.1103/PhysRevE.51.4282.
- [8] Dirk Helbing and Pratik Mukerji. “Crowd disasters as systemic failures: analysis of the Love Parade disaster”. In: *EPJ Data Science* 1.7 (2012), pp. 1–40. DOI: 10.1140/epjds7.

- [9] K. Hirai and K. Tarui. “A simulation of the behavior of a crowd in panic”. In: *Proc. of the 1975 International Conference on Cybernetics and Society*. 1975, p. 409.
- [10] Toshimitsu Homma and Andrea Saltelli. “Importance measures in global sensitivity analysis of nonlinear models”. In: *Reliability Engineering & System Safety* 52.1 (1996), pp. 1–17. ISSN: 0951-8320. DOI: 10.1016/0951-8320(96)00002-6.
- [11] Serhat Hosder, Robert W. Walters, and Michael Balch. “Efficient Sampling for Non-Intrusive Polynomial Chaos Applications with Multiple Uncertain Input Variables”. In: *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. 2007. DOI: 10.2514/6.2007-1939.
- [12] Gianluca Iaccarino. “Quantification of Uncertainty in Flow Simulations Using Probabilistic Methods”. In: *VKI Lecture Series*. Sept. 8 – 12. 2008.
- [13] T. Ishigami and T. Homma. “An importance quantification technique in uncertainty analysis for computer models”. In: *[1990] Proceedings. First International Symposium on Uncertainty Modeling and Analysis*. 1990, pp. 398–403. DOI: 10.1109/ISUMA.1990.151285.
- [14] Michiel J.W. Jansen. “Analysis of variance designs for model output”. In: *Computer Physics Communications* 117.1 (1999), pp. 35–43. DOI: 10.1016/S0010-4655(98)00154-4.
- [15] Benedikt Kleinmeier et al. “Vadere: An Open-Source Simulation Framework to Promote Interdisciplinary Understanding”. In: *Collective Dynamics* 4 (2019). DOI: 10.17815/CD.2019.21.
- [16] Gerta Köster, Franz Treml, and Marion Gödel. “Avoiding numerical pitfalls in social force models”. In: *Physical Review E* 87.6 (2013), p. 063305. DOI: 10.1103/PhysRevE.87.063305.
- [17] Valentina Kurtc. *Uncertainty quantification and sensitivity analysis in application to pedestrian dynamics*. Tech. rep. Peter the Great St. Petersburg Polytechnic University, 2019.
- [18] William L. Oberkampf and Christopher J. Roy. *Verification and Validation in Scientific Computing*. Cambridge: Cambridge University Press, 2010.
- [19] Andrea Saltelli. “Making best use of model evaluations to compute sensitivity indices”. In: *Computer Physics Communications* 145.2 (2002), pp. 280–297. ISSN: 0010-4655. DOI: 10.1016/S0010-4655(02)00280-1.

- [20] Andrea Saltelli et al. *Sensitivity Analysis In Practice*. John Wiley & Sons, Ltd., 2004. ISBN: 978-0-470-87093-8.
- [21] Andrea Saltelli et al. *Global Sensitivity Analysis. The Primer*. John Wiley & Sons, Ltd., 2008, pp. 1–292. DOI: 10.1002/9780470725184.
- [22] Andrea Saltelli et al. “Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index”. In: *Computer Physics Communications* 181.2 (2010), pp. 259–270. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2009.09.018.
- [23] Andreas Schadschneider and Armin Seyfried. “Empirical results for pedestrian dynamics and their implications for modeling”. In: *Networks and Heterogeneous Media* 6.3 (2011), pp. 545–560. DOI: 10.3934/nhm.2011.6.545.
- [24] Michael J. Seitz, Nikolai W. F. Bode, and Gerta Köster. “How cognitive heuristics can explain social interactions in spatial movement”. In: *Journal of the Royal Society Interface* 13.121 (2016), p. 20160439. DOI: 10.1098/rsif.2016.0439.
- [25] Michael J. Seitz and Gerta Köster. “Natural discretization of pedestrian movement in continuous space”. In: *Physical Review E* 86.4 (2012), p. 046108. DOI: 10.1103/PhysRevE.86.046108.
- [26] J. A. Sethian. “A fast marching level set method for monotonically advancing fronts”. In: *Proceedings of the National Academy of Sciences* 93.4 (1996), pp. 1591–1595. DOI: 10.1073/pnas.93.4.1591.
- [27] J. A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge: Cambridge University Press, 1999.
- [28] Isabella von Sivers and Gerta Köster. “How Stride Adaptation in Pedestrian Models Improves Navigation”. In: *arXiv* 1401.7838.v1 (2014). URL: <http://arxiv.org/abs/1401.7838v1> (visited on 07/21/2020).
- [29] Isabella von Sivers and Gerta Köster. “Dynamic Stride Length Adaptation According to Utility And Personal Space”. In: *Transportation Research Part B: Methodological* 74 (2015), pp. 104–117. DOI: 10.1016/j.trb.2015.01.009.
- [30] Isabella von Sivers et al. “Modelling social identification and helping in evacuation simulation”. In: *Safety Science* 89 (2016), pp. 288–300. ISSN: 0925-7535. DOI: 10.1016/j.ssci.2016.07.001.

- [31] Isabella Katharina Maximiliana von Sivers. “Modellierung sozialpsychologischer Faktoren in Personenstromsimulationen - Interpersonale Distanz und soziale Identitäten”. PhD thesis. Technische Universität München, 2016. URL: <https://mediatum.ub.tum.de/doc/1303742/1303742.pdf> (visited on 07/21/2020).
- [32] Ralph C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*. Computational Science and Engineering. Society for Industrial and Applied Mathematics, 2014. ISBN: 978-1-611973-21-1.
- [33] I.M. Sobol’. “Sensitivity estimates for nonlinear mathematical models”. In: *Mathematical Modelling and Computational Experiment* 1 (1993), pp. 407–414.
- [34] I.M. Sobol’ and Yu.L. Levitan. “On the use of variance reducing multipliers in Monte Carlo computations of a global sensitivity index”. In: *Computer Physics Communications* 117.1 (1999), pp. 52–61. DOI: 10.1016/S0010-4655(98)00156-8.
- [35] Bruno Sudret. “Global sensitivity analysis using polynomial chaos expansions”. In: *Reliability Engineering & System Safety* 93.7 (2008). Bayesian Networks in Dependability, pp. 964–979. ISSN: 0951-8320. DOI: <https://doi.org/10.1016/j.ress.2007.04.002>.
- [36] T.J. Sullivan. *Introduction to Uncertainty Quantification*. 1st ed. Springer International Publishing, 2015. DOI: 10.1007/978-3-319-23395-6.
- [37] Vadere team. *suq-controller*. <https://gitlab.lrz.de/vadere/vadere/commit/eed76a68e3f4272c5755eff8c798cbb0c8393164>. 2020.
- [38] Vadere team. *Vadere: Open Source Framework for Pedestrian Simulation*. 2020. URL: <http://www.vadere.org/> (visited on 07/21/2020).
- [39] Vadere team. *Vadere: Open Source Framework for Pedestrian Simulation*. <https://gitlab.lrz.de/vadere/vadere/commit/02dfe88621369b025c67ef9fe60b1869591aae76>. 2020.
- [40] Ulrich Weidmann. *Transporttechnik der Fussgänger*. 2nd. Vol. 90. Schriftenreihe des IVT. Zürich: Institut für Verkehrsplanung, Transporttechnik, Strassen- und Eisenbahnbau (IVT) ETH, 1992. DOI: 10.3929/ethz-a-000687810.
- [41] Y. Xiao et al. “Investigation of Voronoi diagram based direction choices using uni- and bi-directional trajectory data”. In: *Physical Review E* 97.5 (2018). DOI: 10.1103/PhysRevE.97.052127.

- [42] Dongbin Xiu. “Fast numerical methods for stochastic computations: a review”. In: *Communications in computational physics* 5.2-4 (2009), pp. 242–272. DOI: 10.1.1.148.5499.
- [43] Dongbin Xiu. *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton, NJ: Princeton University Press, 2010. ISBN: 978-0-691-14212-8.
- [44] Dongbin Xiu. “Stochastic Collocation Methods: A Survey”. In: *Handbook of uncertainty quantification*. Ed. by R. Ghanem, H. Owhadi, and D. Higdon. Vol. 45. 2017, pp. 699–716. DOI: 10.1007/978-3-319-12385-1_26.
- [45] J. Zhang et al. “Transitions in pedestrian fundamental diagrams of straight corridors and T-junctions”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2011.06 (2011), P06004. DOI: 10.1088/1742-5468/2011/06/P06004.

A Appendix

A.1 Adapted Parameters in the Scenario Files

Table A.1: Changes to the default values of the scenario parameters

Parameter	Default scenario	Corridor scenarios	OPMoPS scenario
release	1.11	1.11	1.12
cacheType	NO_CACHE	NO_CACHE	BIN_CACHE
finishTime	500	150 (for C-180-180- b_{exit}) 100 (for C-300-300- b_{exit})	1500
useFixedSeed	true	true (for $n_r = 1$) false (for $n_r = 10$)	false
cognition	SimpleCognitionModel	CooperativeCognitionModel	CooperativeCognitionModel
speedDistributionMean	1.34	$\mathcal{U}(1.37, 1.73)$	0.6
speedDistributionStandardDeviation	0.26	$\mathcal{U}(0.10, 0.20)$	$\mathcal{U}(0.05, 0.10)$
minimumSpeed	0.5	1.37	0.40
maximumSpeed	2.2	1.73	0.80
attributesCar	{[...]}	null	null

A.2 Complementary Figures for Evaluation of the Method

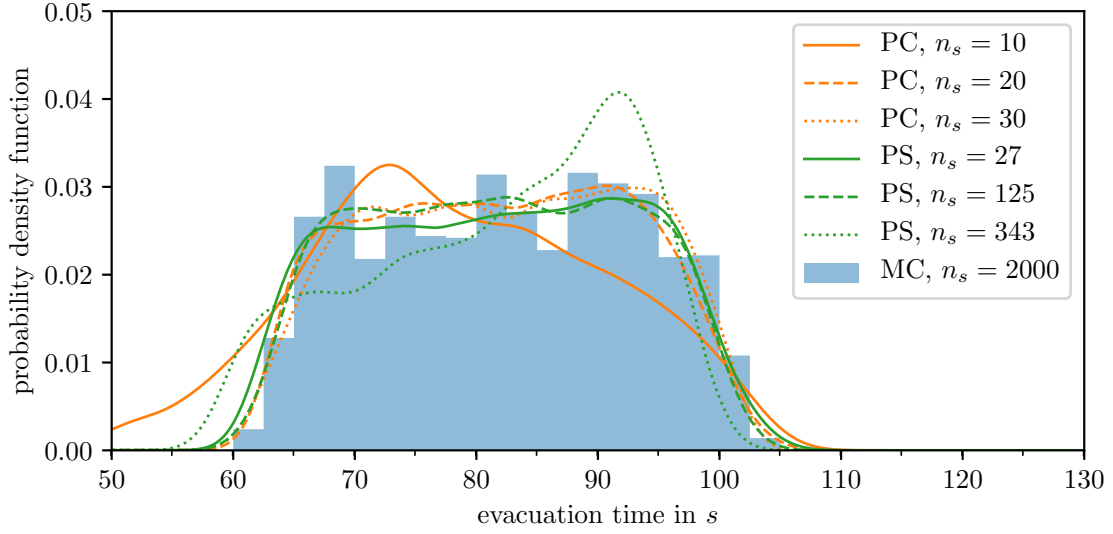


Figure A.1: Qualitative comparison of output distributions approximated by PCE with a varying sample size: The polynomial degree is $p = 2$ for both point collocation (PC) method and pseudo-spectral (PS) approach. The reference solution is obtained with Monte Carlo (MC) from 2000 sample points and represented by a histogram. The model response is the evacuation time.

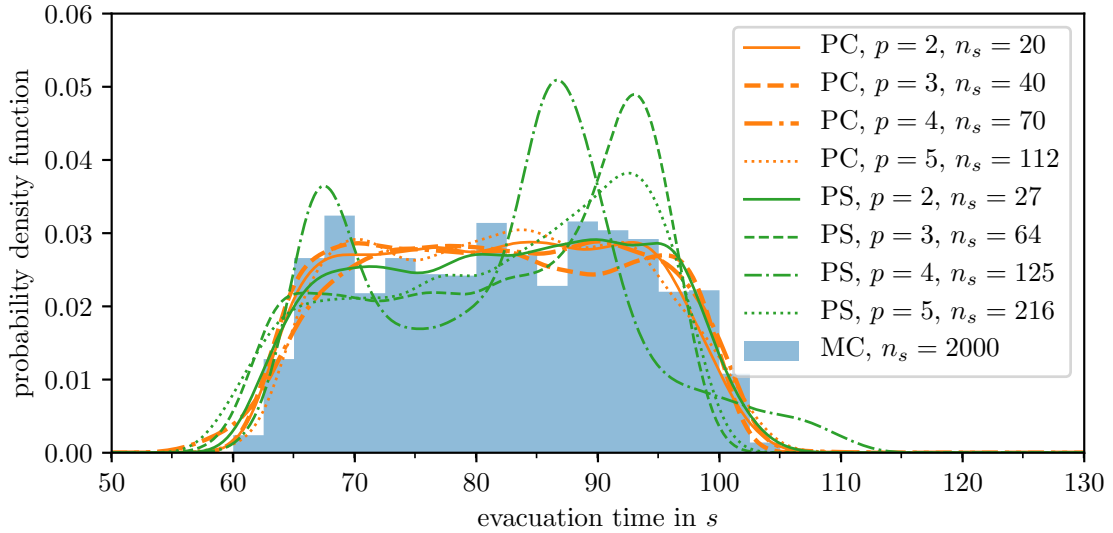


Figure A.2: Qualitative comparison of output distributions approximated by polynomial chaos expansions with varying polynomial order p : The sample sizes n_s are in accordance with $\alpha_{PC} = 2$ in case of point collocation (PC) method and $\alpha_{PS} = 1$ for the pseudo-spectral (PS) approach. Thick lines indicate configurations of the PC method ($p = 3$ and $p = 4$) that will prove useful in the quantitative assessment. The reference solution is obtained with Monte Carlo (MC) from 2000 sample points and represented by a histogram. The model response is the evacuation time.

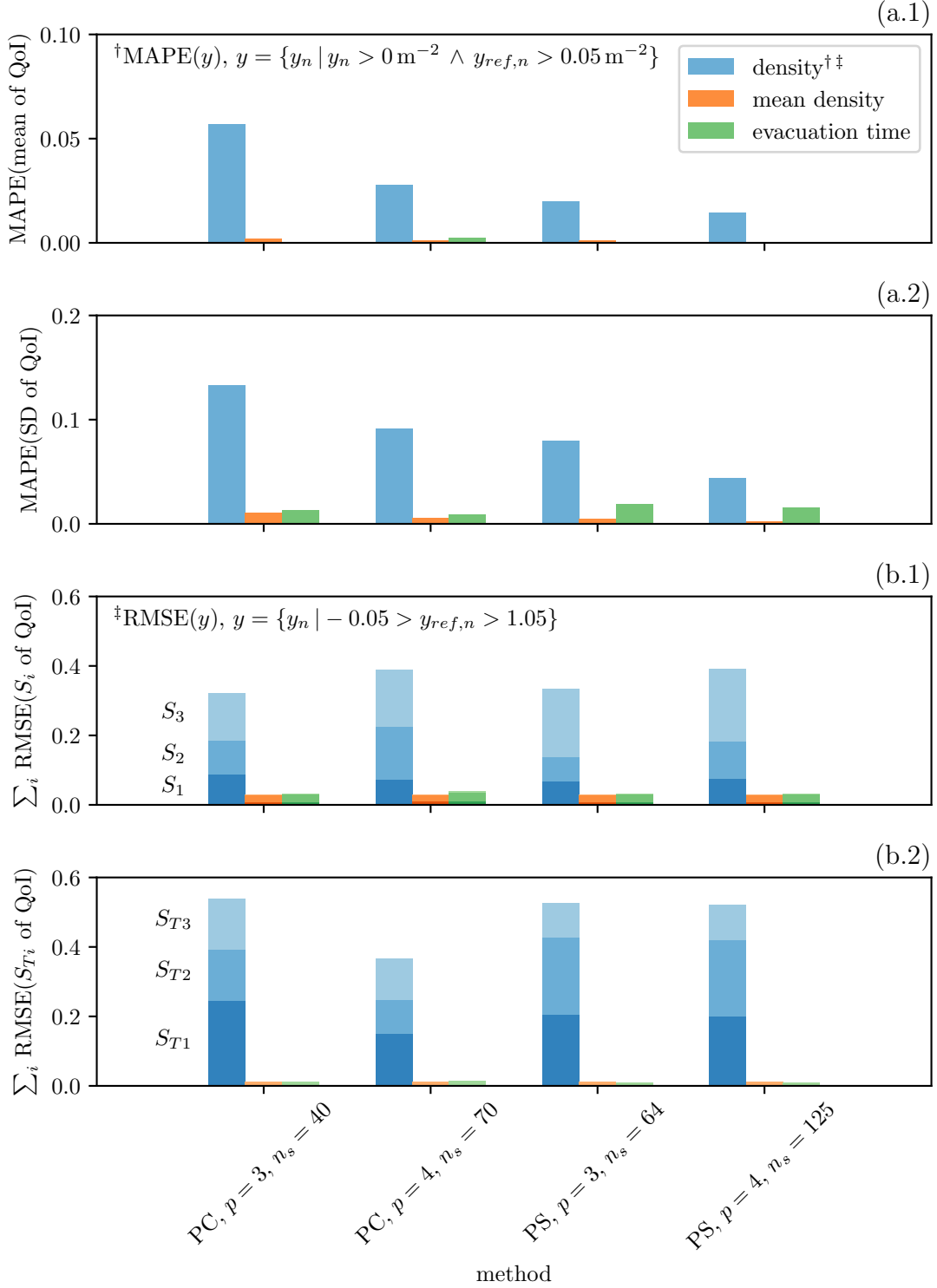


Figure A.3: Quantitative comparison of of four pre-selected configurations of the UQ methods based on their errors: The UQ results for different QoI are calculated by point collocation (PC) and pseudo-spectral (PS) approach applied to the model response of 10 runs ($n_r = 10$). $\text{MAPE}(y)$ quantifies the error of mean (a.1) and standard deviation (a.2). $\text{RMSE}(y)$ quantifies the error of first order sensitivity indices (b.1) and total sensitivity indices (b.2). The reference values y_{ref} are derived from Monte Carlo (MC) simulations with $n_s = 2000$ sample points for the uncertainty analysis (a) and $n_s = 5000$ sample points for the sensitivity analysis (b). The estimation of the error for QoI Voronoi density time series neglects 3% of the data due to different thresholds for (a)[†] and (b)[‡].

A.3 Detailed Uncertainty Quantification for Corridor Scenario with Similar Topographies

This section analyses the uncertainty and sensitivity for the corridor scenarios with similar topographies more in detail. The scenarios are defined in Section 4.2.

A.3.1 Uncertainty Analysis

Scalar Quantity of Interest: Evacuation Time

The scalar values for mean and standard deviation of the evacuation time for all scenarios are summarized in Table A.2. As expected, mean values of evacuation time are greater if the corridor or bottleneck in total is narrowed by about one third. In both cases, corridor and bottleneck, this leads to approximately 1.4 times longer mean evacuation times. Whereas, if the bottleneck is compared to the corridor scenario, mean evacuation time is only a little higher. Reducing only b_{exit} by $1/3$, i.e. transforming the corridor into a bottleneck scenario, yields a 1.1 times longer mean evacuation time.

The reason might be that, in case of a bottleneck, the average flow is reduced. At the same time, many agents in front of a narrow passage lead to a higher densities. This can be translated into the oppression that the agents feel in this area. Videos generated with the post-visualisation tool in Vadere support this statement. They show that the highest density occurs right in front of the bottleneck and the agents pass faster through the constricted egress. Thus the average flux through the narrowing is higher than through a straight exit.

Table A.2: The uncertainty of evacuation time for the corridor scenario with different topographies is represented by mean and standard deviation. The values are given in seconds.

Scenario	Mean	SD
C-180-180-180 (narrow corridor)	81.91	10.90
C-180-180-120 (narrow bottleneck)	90.07	11.94
C-300-300-300 (wide corridor)	60.33	7.00
C-300-300-200 (wide bottleneck)	65.82	7.53

The standard deviation of evacuation time is less for the wider corridor (C-300-300-300) in comparison to the narrow one (C-180-180-180) because the accumulation of agents in the waiting area is not as large. The standard deviation of evacuation time is greater for the bottleneck scenario than for the corridor scenario. In this

case, the accumulation in the waiting area is more significant, which mainly depends on the number of agents. In general, any change in the topography that leads to less accumulation of agents and thus to slowed down agents has an influence on the uncertainty in the output: The evacuation time does not depend so much on the uncertainty in the number of agents but more on speed distribution parameters.

Time dependent Quantity of Interest: Voronoi density

The Voronoi density time series for the four scenarios reveal that the five phases that have been identified previously occur again for similar topographies because the curves are similar. Figure A.4 shows mean and standard deviation of Voronoi density for each scenario. As expected, the graphs for all four scenarios do not differ much. Their course and the co-domain are similar, while only the maximum density reaches about 1.3 higher values for C-300-300-300 versus C-180-180-180 and C-300-300-200 versus C-180-180-120, respectively. Mainly, the curves vary in the extension with regard to time.

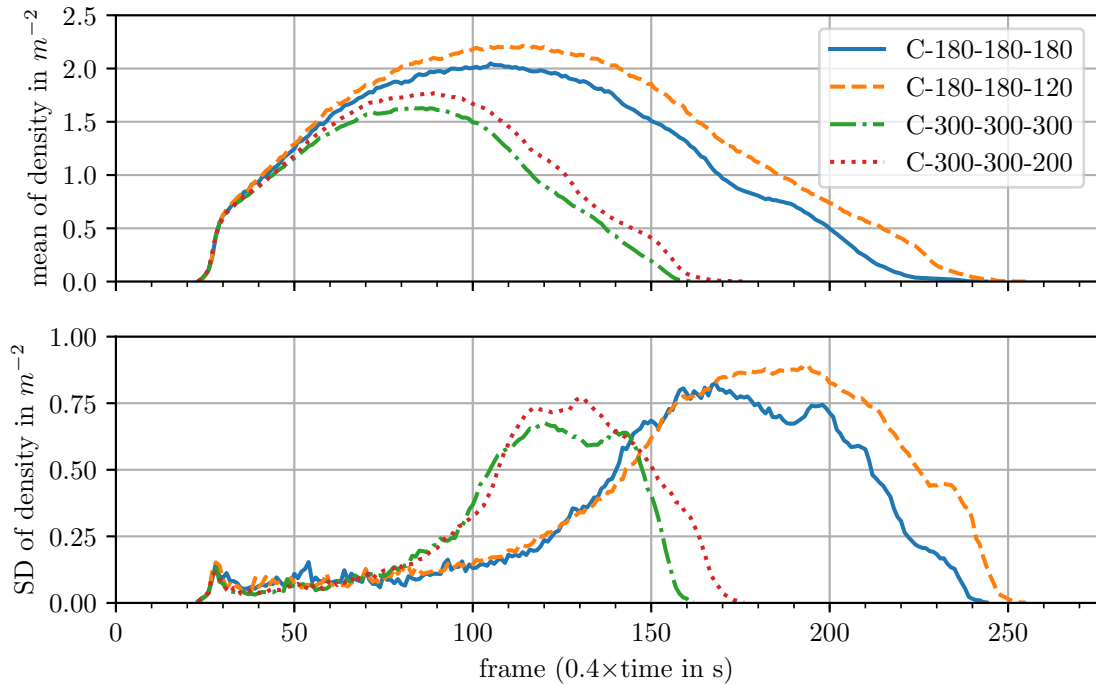


Figure A.4: Uncertainty of QoI Voronoi density in measurement area A_D obtained by applying PC method to the model response y : y is the average of $n_r = 10$ scenario runs.

The standard deviation is greater for narrow scenario (C-180-180-180, C-180-180-120) than for its wide complementary scenario (C-300-300-300, C-300-300-200). Furthermore, the standard deviation tends to be greater if the exit is narrow instead of straight. The causes for such curves for mean and standard deviation are the same as for QoI evacuation time. Parameter number of agents is mainly responsible for

the output uncertainties. This can be demonstrated again by means of sensitivity indices.

A.3.2 Sensitivity Analysis

Scalar Quantity of Interest: Evacuation Time

The influence of the input parameters on the QoI evacuation time is quantified by the total sensitivity indices as shown in Table A.3. First order sensitivity indices are not listed because they yield the same values. That means, there are no significant interactions between the parameters.

Table A.3: Total sensitivity indices (S_{Ti}) quantify the influence of parameters mean of free-flow speed ($i = 1$), SD of free-flow speed ($i = 2$) and number of agents ($i = 3$) on the QoI evacuation time for the corridor scenario with different topographies.

Scenario	S_{T1}	S_{T2}	S_{T3}
C-180-180-180 (narrow corridor)	0.03	0.00	0.97
C-180-180-120 (narrow bottleneck)	0.03	0.00	0.97
C-300-300-300 (wide corridor)	0.04	0.00	0.96
C-300-300-200 (wide bottleneck)	0.04	0.00	0.96

As expected, the speed distribution parameters have no significant influence on the output variance. The all-dominant parameter is the number of agents. The minor difference between the narrow (C-180-180-180, C-180-180-120) and the wide (C-300-300-300, C-300-300-200) scenarios might be traced back to the fact that the agents pass the broader scenario faster and thus phase 3 is shorter. However, this change is so faint that it could also result from stochastic behaviour of the model or the limited precision of the UQ method.

Time dependent Quantity of Interest: Voronoi density

In principle, the same applies to the time series of both 1st order and total sensitivity indices for the Voronoi density time series. Occasionally, interactions are more relevant than for the evacuation time, though. The first order sensitivity indices have been analysed for the single corridor scenario (see Section 4.1.2) and the conclusions drawn apply also here: Parameter mean of free-flow speed has a minor effect as long as the agents can move freely. Standard deviation of free-flow speed is of no importance. And the number of agents is important especially during phase 3. The modified topography does not change the results for the first order sensitivities apart from a dilation along the time axis. The effect of modified topographies can be explained better by considering the total sensitivity indices shown in Figure A.5.

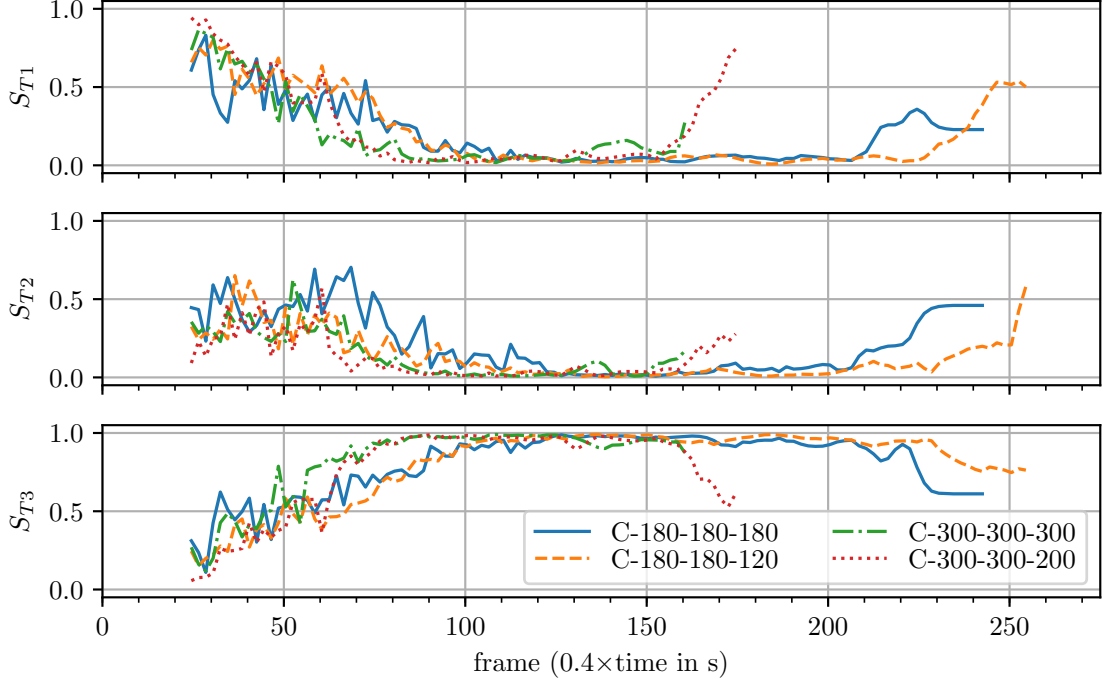


Figure A.5: Total sensitivity indices obtained with PC method applied to the average of $n_r = 10$ repeated model evaluations. They quantify the influence of parameter mean free-flow speed (top, S_{T1}), SD of free-flow speed (middle, S_{T2}), number of agents (bottom, S_{T3}) and their interactions on the QoI Voronoi density in measurement area A_D . All curves are smoothed by taking the average of two frames.

As mentioned within the framework of the uncertainty analysis, the curves for narrow scenarios and the ones with a narrow exit are stretched. The differences between the four topographies are not so plain during phase 2 and 3, since the curves fluctuate too much. A coarser smoothing function straightens things to some extent, and yet, one cannot make out large deviations between the four curves. Besides, interesting information about phase 4 gets lost if too many frames are averaged. Nevertheless, it turns out that parameter mean and standard deviation of free-flow speed are more relevant for bottlenecks than for straight corridors. As soon as the waiting area is empty, there is no more pressure from behind that would force the agents to move closer together. Hence, they can move more freely, and parameter mean and standard deviation of free-flow speed become more important.

Statutory Declaration

I herewith declare that I have composed the present thesis myself and without use of any other than the cited sources and aids. Sentences or parts of sentences quoted literally are marked as such; other references with regard to the statement and scope are indicated by full details of the publications concerned. The thesis in the same or similar form has not been submitted to any examination body and has not been published. This thesis was not yet, even in part, used in another examination or as a course performance.

Munich, 29 July 2020

Simon Rahn