Bayesian Behavioural Model Estimation for Live Crowd Simulation

Fumiyasu Makinoshima
Fujitsu Limited
Kawasaki, Japan
f.makinoshima@fujitsu.com

Tetsuro Takahashi
Fujitsu Limited
Kawasaki, Japan
takahashi.tet@fujitsu.com

Yusuke Oishi*
Kyushu University
Fukuoka, Japan
oishi.yusuke.492@m.kyushu-u.ac.jp

ABSTRACT

The increasing availability of real-time crowd observation data enables the development of agent-based crowd simulations that incorporate real-time data feeds, i.e. *live crowd simulations*, which achieve accurate crowd forecasting for real-time interventions for improved and safer mobility. Various approaches for live crowd simulations have recently been proposed. However, existing methods cannot offer long forecasting lead times, which are crucial for planning and implementing timely interventions. To address this issue, we develop a Bayesian behavioural model estimation for live crowd simulations that sequentially estimates the underlying behavioural model assumed behind the observed crowd flows. In real crowds, although apparent behaviours change over time, an invariable rule that determines behaviour often exists behind the crowd. The developed method estimates the underlying invariable behavioural model and provides reliable long-term crowd flow forecasting. The experimental results show that the developed method can accurately forecast long-term crowd flows using aggregate observations, whereas the state-of-the-art forecasting method fails to provide reliable forecasting when the apparent behavioural tendency changes. We also demonstrate that the developed method can provide forecasting results that are sufficient to consider interventions, even in the presence of a data-model mismatch.

KEYWORDS

Agent-based crowd simulation; Data assimilation; Live simulation; Crowd flow forecasting

ACM Reference Format:
Fumiyasu Makinoshima, Tetsuro Takahashi, and Yusuke Oishi. 2024. Bayesian Behavioural Model Estimation for Live Crowd Simulation. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 8 pages.

1 INTRODUCTION

Rapid global urbanisation has led to population concentration in recent years, with over half of the world's population living in urban areas [2]. When the crowd density exceeds a certain threshold in these built environments, the crowd itself may cause injuries or fatalities without any external hazards [16], resulting in crowd disasters. Although Itaewon crowd crush in 2022 is a recent example of a major crowd disaster, a number of crowd disasters has been continuously reported worldwide for decades [11]. Due to the trend of rapid urbanisation and increasing population density, there has been a considerable rise in crowd disaster risks. This has facilitated the need for methodological development to establish effective crowd management strategies in urban environments.

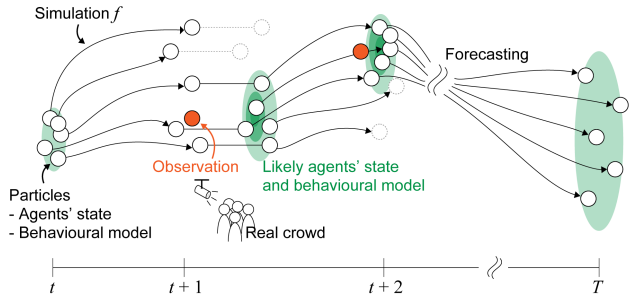

Figure 1: Schematic view of Bayesian behavioural model estimation for crowd flow forecasting.

Agent-based crowd simulation is a powerful tool for simulating crowd movements and has been used to study fundamental crowd dynamics and design environments for improved crowd mobility [18]. However, conventional agent-based simulations have been rarely combined with real-time observation data and have been limited to what-if analyses. Since real crowds often exhibit significantly different movements compared to the predefined model in what-if analyses, the conventional crowd simulations have the difficulty in providing useful information for ongoing events. Less accurate forecasting can lead to inappropriate decision-making or interventions by policy makers, which may worsen the situation.

Live simulation is the recently proposed concept that represents large-scale agent-based simulations coupled with real-time data [35]. By integrating real-time data feeds, live simulations provide short-term forecasting of an ongoing event, which is vital information for real-time interventions. In recent years, real-time crowd observation data have become readily available from various observation channels, such as cameras [33], GPS [3], Wi-Fi, and Bluetooth [32]. The number of observation channels for crowd monitoring has increased, particularly in smart cities [29]. Thus, the available data are becoming rich. Live crowd simulations based on such rich crowd observations are expected to be effective tools for preventing future crowd disasters and improving mobility in urban environments. However, despite such advanced concepts, there is a lack of effective methodological development to achieve this. Specifically, live crowd simulations that can forecast large-scale

*Work performed while employed at Fujitsu Limited.

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

microscopic crowd flows with long forecasting lead times remain a challenge. A long lead time is essential for real applications in order to consider and implement appropriate interventions in a timely manner. Such methodological developments help not only to achieve better real-time crowd management, but also to build a digital twin of cities [39], where an accurate crowd simulation is vital for evaluating urban planning and policymaking.

In this paper, we develop a Bayesian behavioural model estimation (BBME) for live crowd simulations, enabling accurate crowd forecasting with long forecasting lead times. As shown in Fig. 1, by using real-time aggregate crowd observations, the developed method can sequentially estimate the behavioural model behind the observed crowd flow, which enables highly reliable long-term crowd forecasting. Through numerical experiments, we demonstrate that the developed method can provide accurate crowd forecasting with a long lead time, which is sufficient to consider interventions to achieve improved mobility, whereas a state-of-the-art assimilation method fails to provide reliable long-term crowd forecasting.

Our contributions: (i) Present the concept and implementation of a Bayesian behavioural model estimation for live crowd simulations. (ii) Clarify the limitations of the state-of-the-art crowd forecasting method in a realistic experimental setting, wherein people's apparent behavioural tendencies change over time. (iii) Verify the performance of the developed method to provide accurate crowd forecasting with sufficient forecasting lead time in various experimental settings, including a practical scenario, i.e. a data-model mismatch scenario.

2 RELATED WORK

2.1 Crowd Simulation

Microscopic agent-based models that can consider individual-level behaviour are becoming common in simulating crowd flows. To date, various simulation models such as the force-based models [17], cellular automata-based models [4], discrete choice-based models [1] and velocity-based models [37] have been proposed. These models can simulate basic individual interactions (e.g. avoidance) and the resulting collective crowd movement [9]. Accordingly, crowd simulations have been actively used in various applications, such as evacuation safety evaluation [41] and layout design [12].

2.2 Model Calibration

For realistic simulations, crowd simulation models and parameters can be calibrated against real-world observations. By using crowd observations such as individual trajectories and aggregate densities, crowd simulation model parameters can be calibrated to better fit the target observations [19, 20, 40]. In addition to the simulation model parameters, behavioural models for realistic simulations can be estimated based on the observations or questionnaire surveys [10, 15, 30]. Such calibrated crowd simulations can perform well for the target crowd; however, their applicability to different crowd flows is unclear due to the variations inherent in human behaviour. Such uncertainty in human behaviour in different cases has conventionally limited the use of crowd simulations, mostly to what-if analyses. However, this has motivated the methodological development of live crowd simulations.

2.3 Live Simulation

Although this field remains largely unexplored, some studies have attempted to develop methods for live crowd simulations, and a majority are based on data assimilation techniques that have been actively developed and used in atmospheric science for weather forecasting [21]. Data assimilation for macroscopic crowds or mobility flows has been studied [6, 24, 31, 34], wherein the microscopic interactions among agents were simplified using network-based models. Such simplified simulations can involve a relatively large number of agents, but cannot consider two-dimensional microscopic crowd interactions, which is essential for providing useful information for crowd management. Some studies have attempted to apply data assimilation techniques to two-dimensional microscopic crowds [7, 8, 27, 38]. However, their results showed that crowd forecasting for large microscopic crowds is difficult due to the increased complexity associated with a large number of interacting agents. As a result, the number of agents considered in these studies was limited (e.g. tens of agents), which is insufficient for real-world problems. Recently, sequential latent parameter estimation (SLPE), which enables data assimilation for large but microscopic crowd flows, was proposed [26]. With this method, microscopic crowd flow forecasting involving thousands of agents has been presented. However, their results indicated that the method provides reliable forecasting when the behavioural tendency of the observing crowd is stationary, suggesting that the forecasting quality decreases with a long lead time, during which the observed behavioural tendency of real crowds often changes. Consequently, microscopic large-crowd forecasting with a long forecasting lead time for interventions remains challenging.

3 METHODS

3.1 Crowd Simulation Model

We employed a simplified force-based crowd simulation model [25] to simulate crowd movement. The movements of the agents are described according to the social force model [17] as follows:

$$\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = \frac{v_i^0 \mathbf{e}_i^0 - \mathbf{v}_i}{\tau_\alpha} + \sum_{i \neq j} \mathbf{F}_{ij}, \tag{1}$$

where $\mathbf{v}_i$ is the current velocity, $v_i^0 \mathbf{e}_i^0$ is the desired velocity, τ_α is the constant relaxation time, and $\mathbf{F}_{ij}$ is the interaction force between the agents. In this simplified model, the effect of obstacles is modelled by no-entry rules, referencing $1 \times 1\ m^2$ grids, rather than calculating the force from obstacles. The interaction force $\mathbf{F}_{ij}$ is calculated as follows [22]:

$$\mathbf{F}_{ij} = -\nabla_{\mathbf{r}_{ij}} (k\tau^{-2} e^{-\tau/\tau_0}), \tag{2}$$

where τ is the projected time to collision, which is calculated using linear extrapolation based on the relative velocity $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ and the relative displacements $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, and k and τ_0 are the constants. We employed this model because it was validated against various fundamental diagrams [25]; however, other models simulating realistic crowd movements, such as velocity-based models, can be used as alternatives to estimate crowd flows.

3.2 Behavioural Modelling

An agent's behaviour is modelled as a choice behaviour based on random utility theory [36]. We assume that the utility function of individual i for alternative n has the following form:

$$U_{in} = V_{in}(\mathbf{X}_{in}; \boldsymbol{\theta}) + \epsilon_{in}, \tag{3}$$

where V_{in} is the non-stochastic component of utility, which is function of the explanatory variables $\mathbf{X}_{in}$ and the parameters $\boldsymbol{\theta}$ to be estimated, and ϵ_{in} is a random term representing the unobserved components. If we assume that ϵ_{in} is independently and identically Gumbel distributed, the probability that individual i chooses alternative n can be expressed as follows [28]:

$$\pi_{in} = \frac{\exp(V_{in})}{\sum_{m \in M} \exp(V_{im})}, \tag{4}$$

where M denotes the choice set. The estimation procedure for the behavioural model is described in the following sub-section.

3.3 Bayesian Behavioural Model Estimation

Based on crowd observations, we aim to sequentially estimate a behavioural model that enables accurate long-term crowd forecasting. For this purpose, we model the transition of crowd flows and observations as a nonlinear, non-Gaussian state-space model [26]:

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k), \tag{5}$$

$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{n}_k), \tag{6}$$

where $\mathbf{x}_k$ and $\mathbf{y}_k$ are the state vector and the observation at time k, f is the agent-based crowd simulation, h is the function to transform a state $\mathbf{x}_k$ to be consistent with an observation, and $\mathbf{u}_k$ and $\mathbf{n}_k$ are the system and observation noise, respectively. Here, the state vector includes not only the agent status, such as their positions and velocities, but also the parameters $\boldsymbol{\theta}$ for their behavioural model. The stochastic factors in simulations can be considered as system noise. The observation noise was not considered in this study.

We can then sequentially estimate the crowd state, including behavioural model parameters, using a particle filter [14, 23] as we obtain the crowd observations (Algorithm 1). In this study, the weights of the particles were calculated based on the sum of the errors between the observed and simulated population maps, i.e. $\lambda_k^{(l)} \propto \exp(-\sum |\mathbf{y}_k - h(\hat{\mathbf{x}}_k)|/\sigma^2)$, where we set $\sigma = 60$. Residual systematic resampling [5] was employed as resampling algorithm. Similar to roughening [14], Gaussian noise $\epsilon_\theta \sim \mathcal{N}(0, 0.05)$ was added to the parameters $\boldsymbol{\theta}$ at each assimilation step to explore the various possible crowd states and maintain the diversity of the ensembles. The number of particles $N = 500$ was consistently employed in this study. This estimation procedure was run every 500 simulation time steps, which is equivalent to 5 s in real-time. After successive estimations, forward simulations with the estimated crowd state and behavioural model provide crowd flow forecasting.

Algorithm 1 Bayesian Behavioural Model Estimation

1: Generating initial distribution $\{\mathbf{x}_{0|0}^{(l)}\}_{l=1}^{N}$ with initial agent states and behavioural model parameters $\boldsymbol{\theta}$.
2: **for** $k = 1, \cdots, T$ **do**
3: Obtaining predictive distribution: $\{\mathbf{x}_{k|k-1}^{(l)}\}_{l=1}^{N}$ by $\mathbf{x}_{k|k-1}^{(l)} = f(\mathbf{x}_{k-1|k-1}^{(l)}, \mathbf{u}_k)$ for $l = 1, \cdots, N$
4: Calculating weights of particles: $\lambda_k^{(l)} = p(\mathbf{y}_k|\mathbf{x}_{k|k-1}^{(l)})$ for $l = 1, \cdots, N$
5: Normalising the weights of the particles: $\beta_k^{(l)} = \lambda_k^{(l)} / \sum_L \lambda_k^{(L)}$ for $l = 1, \cdots, N$
6: Resampling to obtain particles $\{\mathbf{x}_{k|k}^{(l)}\}_{l=1}^{N}$, approximating filtering distribution $p(\mathbf{x}_k|\mathbf{y}_{1:k}) \simeq 1/N \sum_l \delta(\mathbf{x}_k - \mathbf{x}_{k|k}^{(l)})$
7: **end for**

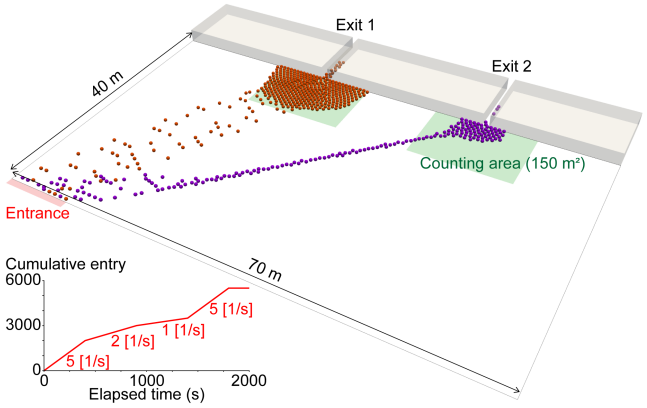

Figure 2: Simulation setup for numerical experiments. Agents enter environment based on given flow rate and choose either Exits 1 or 2. Orange and purple particles represent agents who chose Exits 1 and 2, respectively.

4 EXPERIMENTAL SETUP

4.1 Simulation Environment

We conducted various numerical experiments to verify crowd forecasting methods. Figure 2 shows the simulation environment and the experimental setup. The environment has a 10-m-wide entrance and two exits that have 1 m bottleneck width. A total of 5,500 agents entered the environment from the entrance and exited through either Exits 1 or 2. The agent entry flow rate changed from one to five people/s over the simulation time, which resulted in various crowding conditions. Similar to an existing study [8], the flow rate is assumed to be known in the experiment because such objective information can be obtained from various observations, such as those from cameras installed in connecting aisles.

Although the flow rate is known, it is still difficult to accurately forecast the crowd because we must predict the decision-making of the people, i.e. the tendency of the exit choice in this case, which is a psychological process and thus cannot be directly observed. In this environment, Exit 1 was located closer to the entrance than Exit 2, and people usually preferred Exit 1. However, some people may prefer Exit 2 to Exit 1 if Exit 1 becomes overcrowded. Such decision-making certainly exists in our daily lives, but it cannot be directly observed. For accurate forecasting, we must infer these internal decision-making rules. We aim to estimate this tendency of subjective decision-making solely from objective crowd observations. As the available observations, only crowd population maps

with the resolution of $10 \times 10\ m^2$ were considered in this study. To verify the forecasting lead time, the available observation period was consistently set to $t \leq 500\ s$, which is only the first quarter of the total simulation time. The simulated environment and behaviours are simplified; however, similar situations can often be observed in the real world, such as at train stations.

4.2 Experimental Scenario

We assume the following utility function for Exits 1 and 2 to explain crowd movements:

$$V_{in} = \theta_{in,0} + \theta_{in,1} X_{in}, \tag{7}$$

where $\theta_{in,0}$ is the alternative-specific constant only valid for Exit 1 owing to normalisation ($\theta_{i2,0} = 0$), and $\theta_{in,1}$ is the coefficient for the perceived density X_{in}, which is calculated in the counting area visualised in Fig. 2. Because the estimation of a large number of heterogeneous parameters is intractable, we assumed homogeneous parameters in the estimation ($\theta_{i1,0} = \theta_0$ and $\theta_{in,1} = \theta_1$). To prevent biased estimations, the initial behavioural model parameters were drawn from uniform distributions, i.e. $\theta_0 \sim \mathcal{U}(-5, 5)$ and $\theta_1 \sim \mathcal{U}(-5, 0)$. Although a simple behavioural model, such as a linear utility function, is often specified in discrete choice modelling as employed in this study, in principle, a more complex behavioural model can be employed in the developed framework.

To verify the forecasting methods, we considered the following scenarios with different decision-making rules:

Identical Twin Scenario. To verify the performance of the developed method, an identical twin approach [7, 8, 27, 38] was considered as the first scenario, wherein the observation data were generated from the same system used for estimation. To synthesise observation data, we ran a single forward simulation with Gaussian-distributed individual heterogeneity with a 10% standard deviation, i.e. $\theta_0 \sim \mathcal{N}(4.5, 0.45)$ and $\theta_1 \sim \mathcal{N}(-2.5, 0.25)$. Since this simulation for generating data includes heterogeneous parameters for different individuals and probabilistic choices, the model used for the estimation, which assumes homogeneous parameters, does not exactly match the synthetic data. Nevertheless, the system used in the estimation is capable of representing the behavioural tendencies of the observed data.

Data-Model Mismatch Scenario. In real applications, we cannot determine the true form of the utility function to represent observed behavioural patterns. As a result, a gap exists between real behaviours and the model used for an estimation, i.e. data-model mismatch [13]. Assuming such challenging situations in real applications, we synthesised another observation dataset using a different decision-making system, wherein agents have their own tolerance parameter γ_i, representing the acceptable degree of congestion at Exit 1. If the number of people in the counting area at Exit 1 exceeds γ_i when agent i enters the environment, the agent chooses Exit 2. Here, γ_i is drawn from a mixed distribution of two Gaussian, i.e. $\mathcal{N}(200, 50)$ and $\mathcal{N}(350, 50)$, resulting in a complex multimodal distribution that is difficult to be estimated. In this scenario, the choice of Exit 2 is caused only by the congestion at Exit 1, and the congestion at Exit 2 does not affect the choice, whereas the model assumes that the density at the exits affects the utility equally through θ_1; therefore, there is a clear data-model mismatch, in addition to the complex parameter distribution.

5 RESULTS

5.1 Identical Twin Scenario

The forecasting results with the ground truth data in the identical twin scenario is visualised in Fig. 3. Here, we also ran SLPE [26] with the same observation data and assimilation window as the state-of-the-art method for a detailed comparison. For the SLPE, we followed the implementation in the original study [26], which considered an estimation of the single exit preference parameter θ. Within the assimilation period ($t \leq 500\ s$), both the SLPE and BBME successfully estimated the crowd flows. Even with the coarse observation having the resolution of $10 \times 10\ m^2$, both the SLPE and BBME reproduced the crowd state, which was equivalent to the ground truth density map (Fig. 3 (b)). However, the SLPE failed to provide accurate crowd flow forecasting, showing less crowd density at Exit 1 and largely overestimating the crowd extent at Exit 2 (Fig. 3 (d)). In contrast, the BBME provided reliable long-term crowd flow forecasting, which was equivalent to the ground truth density map (Fig. 3 (e)).

For a quantitative evaluation, the time series of the number of people near the exits (people inside the counting area visualised in Fig. 2) are compared in Fig. 4. In this figure, the mean values and 90% credible interval (CI) of the forecasting are presented along with ground truth observations (every 5 s). As seen in the snapshots, the quality of forecasting with the SLPE started to degrade as soon as the assimilation procedure ended. As a result, congestion forecasting with the SLPE was largely diverged as the forecasting lead time increased. In forecasting with the SLPE, the relative errors of the peak population in the second crowd wave were 25.1% and 51.9% for Exits 1 and 2, respectively. With the BBME, we obtained more reliable crowd forecasting, even for long lead times. Even after the assimilation period ($t > 500\ s$), the method provided accurate estimates of actual crowd flows with a narrow CI. While ensuring sufficient lead time for interventions, the second congestion peak at the exits was accurately forecasted with the relative errors of 1.3% and 14.6% for Exits 1 and 2, respectively. These results verify that the BBME can provide reliable long-term crowd forecasting, which is sufficient for timely interventions to manage congestion.

To understand how BBME achieved these successful long-term crowd forecasts, the time series of the estimated parameters θ_0 and θ_1 were investigated with the snapshots of the ground truth (Fig. 5). Here, the mean values of the behavioural model parameters and their 90% CI are visualised as estimates. In the early phase of the assimilation period ($t \leq 50\ s$), θ_0 increased constantly, whereas θ_1 remains stable. From the crowd flow concentrated at Exit 1, the BBME estimated that Exit 1 has a higher utility than Exit 2. Since people rarely head to Exit 2, even when Exit 1 is crowded, the BBME estimated a higher θ_1, indicating a lower weight for crowd density ($t \leq 100\ s$). When Exit 1 became overcrowded and the number of people choosing Exit 2 increased, a lower θ_1 was estimated to reproduce the crowd flow towards Exit 2. In successive periods, although there were apparent behavioural changes, additional information to determine behavioural model was not included in the observations. Consequently, the estimated behavioural parameters were stable. The BBME achieved successful long-term crowd flow forecasting by estimating the invariable behavioural model behind the crowd solely from coarse observations.

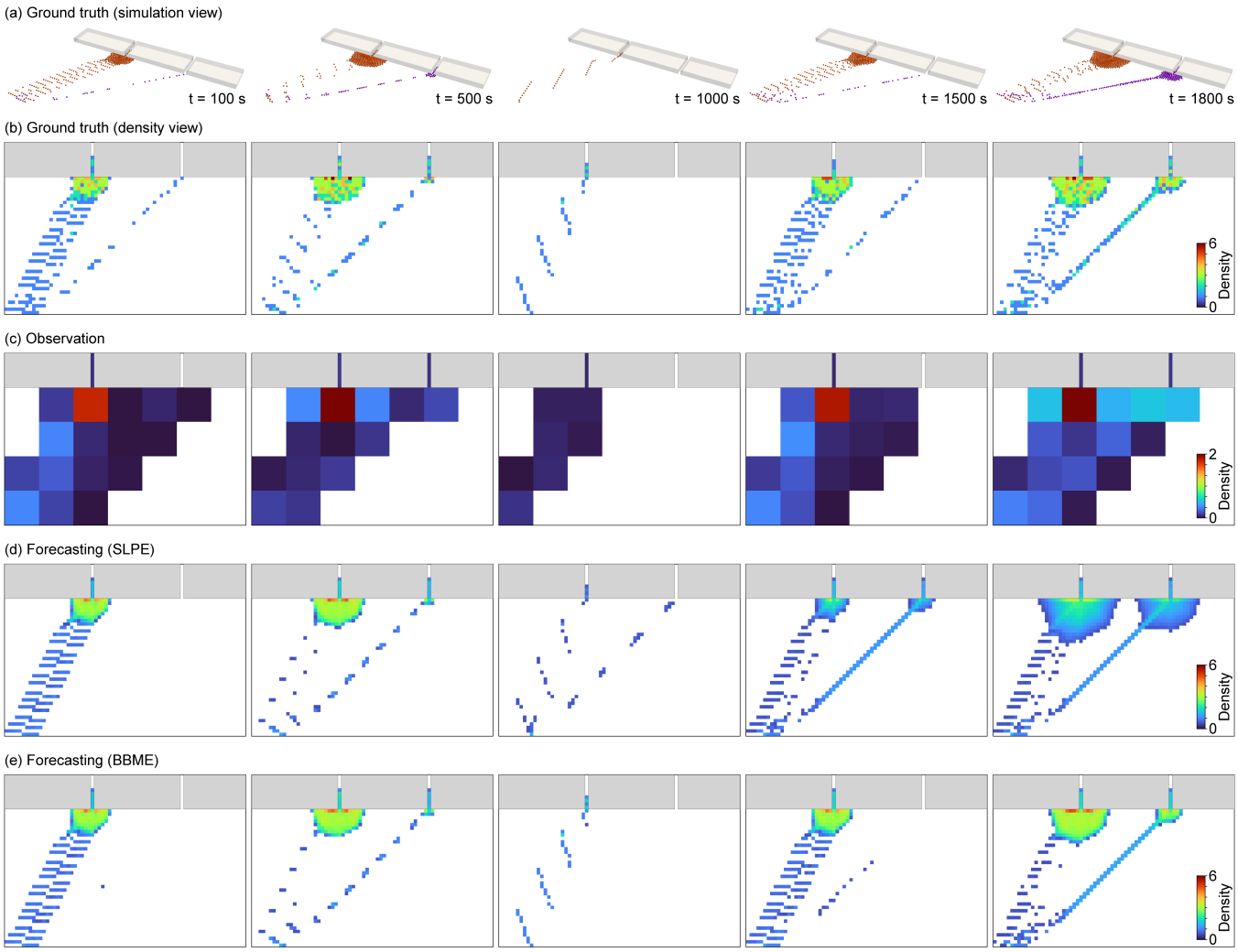

Figure 3: Forecasting results (t = 100, 500, 1000, 1500 and 1800 s) in the identical twin scenario. (a) Ground truth (simulation view). (b) Ground truth (density). (c) Observation (density). (d) Forecasting results with SLPE. (e) Forecasting results with BBME.

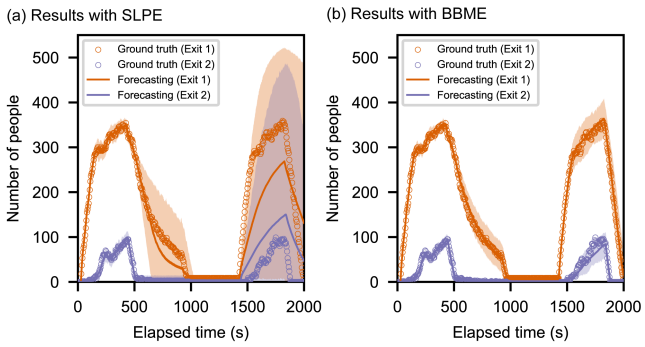

Figure 4: Forecasted exit congestion in the identical twin scenario. Coloured range represents 90% CI of the forecasting.

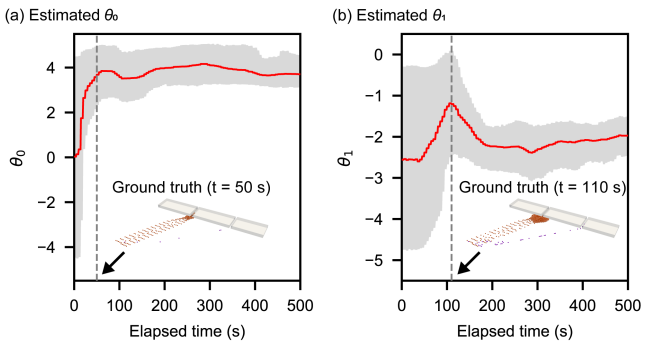

Figure 5: Estimated behavioural model parameters (red line) and the 90% CI (grey range) in the idential twin scenario.

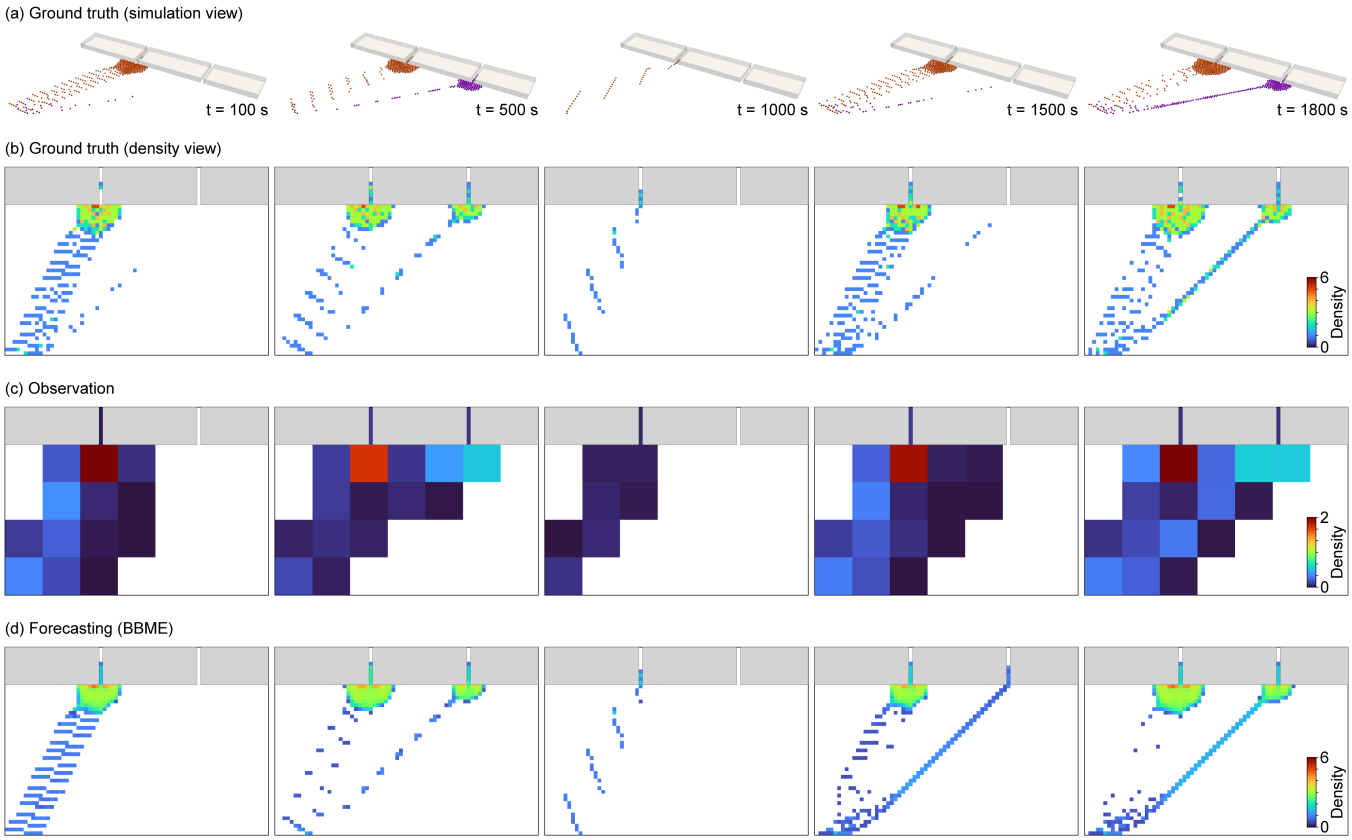

Figure 6: Forecasting results (t = 100, 500, 1000, 1500 and 1800 s) in the data-model mismatch scenario. (a) Ground truth (simulation view). (b) Ground truth (density). (c) Observation (density). (d) Forecasting results with BBME.

5.2 Data-Model Mismatch Scenario

Figure 6 shows the forecasting results with ground truth data in the data-model mismatch scenario. As observed in the identical twin scenario, the BBME reproduced the density map that is equivalent to the ground truth from coarse observations within the assimilation period ($t \leq 500$ s) in the data-model mismatch scenario. Even in the presence of the data-model mismatch, the BBME provided a good prediction of the timing of crowd flow moving towards Exit 2 ($t \sim 1500$ s) after the assimilation period. The resulting extent of the congestion of the second peak of the crowd flow forecasted by the BBME also agreed well with the ground truth data ($t = 1800$ s). Overall, the BBME could provide qualitatively sufficient information to consider timely interventions for managing the forecasted congestion.

Figure 7 shows the forecasted exit congestion in the data-model mismatch scenario. Similar to the results in the identical twin scenario, the BBME followed the observations and successfully reproduced the time series of congestion during the assimilation periods ($t \leq 500$ s). However, some discrepancies between the ground truth and forecasting were observed in successive periods in the data-model mismatch scenario. Although the BBME provided good estimates of the congestion at Exit 2 for the entire simulation time, it failed to provide an accurate forecast of the exit congestion, especially for Exit 1. After the assimilation period ($t > 500$ s), the gap between the ground truth and forecasting for Exit 1 became apparent, and the second congestion peak at Exit 1 was underestimated. The observed errors for Exit 1 indicate a limitation of the developed method.

This performance degradation can be explained by the mismatch between the observation data and the model used for the estimation. In the behavioural model assumed in the BBME, the crowd density at Exits 1 and 2 affects the utility of the exits with shared weight θ_1. In contrast, the synthesised choice behaivour in the data-model mismatch scenario is controlled solely by the tolerance parameter γ_i for Exit 1, and the congestion at Exit 2 does not affect the choice behaviour. Thus, to reproduce the observed behavioural patterns (both congestion at exits and crowd flow towards Exit 2) with the assumed behavioural model, the relative exit-specific constant θ_0 had to be lowered to decrease the utility for Exit 1. To verify this explanation, behavioural model parameters estimated in the data-model mismatch scenario were investigated, as shown in Fig. 8. As discussed, although θ_0 was stable in the identical twin scenario (Fig. 5), θ_0 in the data-model mismatch scenario decreased in the latter phase of the assimilation period ($t = 350$ to 500 s) when Exits 1 and 2 were crowded. This estimated lower θ_0 led to the

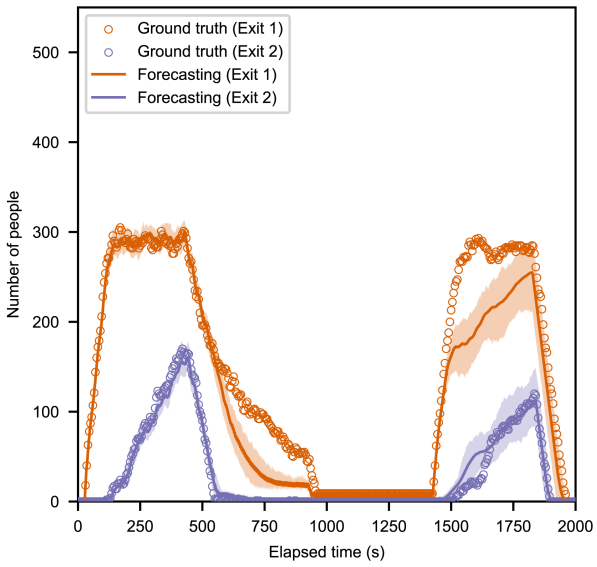

Figure 7: Forecasted exit congestion in the data-model mismatch scenario. Coloured range represents 90% CI of the forecasting.

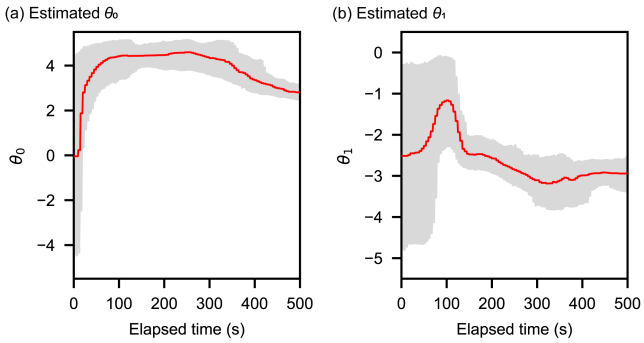

Figure 8: Estimated behavioural model parameters (red line) and the 90% CI (grey range) in the data-model mismatch scenario.

observed underestimation of the time series of congestion at Exit 1, as shown in Fig. 7. The data-model mismatch caused degradation of the forecasting performance; however, the second peak of the congestion was forecasted with reasonable errors, i.e. the relative errors of 13.2% and 6.2% for Exits 1 and 2, respectively, which is still sufficient for planning effective interventions (e.g. guidance) to mitigate exit congestions.

6 DISCUSSION

We developed the Bayesian behavioural model estimation for live crowd simulation, which enables long forecasting lead times for large crowd flows. We further verified the method in both identical twin and data-model mismatch scenarios. As a state-of-the-art method for large crowd forecasting, the performance of the SLPE [26] was also verified using the same numerical settings for comparison. Even with coarse observations, both methods could reproduce fine density maps, which agreed successfully with the ground truth data. In the successive period without observations, while the SLPE failed to provide reliable long-term crowd forecasting, the BBME offered good estimates of the future crowd state, which was sufficient to plan effective timely interventions for better crowd management.

The behavioural model assumed in the BBME is the key for reliable long-term crowd forecasting. In the identical twin scenario, the SLPE estimated the exit preference parameter to reproduce the apparent distributed crowd flow towards Exits 1 and 2. Consequently, the forward simulation with the estimated parameter results in the misestimation of crowd flows towards Exit 2, even when Exit 1 is not crowded. The results demonstrate that forecasting with the SLPE is effective when the apparent behavioural tendency does not change, although such a condition would rarely be observed in real applications. In contrast, the BBME estimated the invariable underlying behavioural model and successfully provided reliable long-term forecasting. The combined use of the discrete choice model and agent-based crowd simulation has been actively studied in existing studies [10, 15, 30]. However, in previous studies, the individual behavioural data for the model estimation were obtained by identifying and tracking individuals, and the estimation of the model was conducted offline using the individual observation data. In contrast to these approaches, the BBME effectively estimates the behavioural model in an online fashion using privacy-aware coarse observations that do not locate individuals in grids. This characteristic supports the applicability of the BBME in real-world applications.

An additional advantage of the BBME is the flexibility in the design of the utility function, which is assumed in the estimation. A linear utility function is often assumed in discrete choice modelling for the convenience of parameter estimation. Although this study similarly employed a simple form of the utility function because the primary purpose of this study was to verify the methods, the BBME can consider a more complex form of the utility function in principle. In addition, because the BBME does not require individual observation, it can incorporate various forms of observation to estimate the utility function. For example, the number of people passing a certain aisle might be considered as an observation for estimating route choice models to reproduce and forecast crowd flows. Future work should investigate the applicability of the BBME to different data or model settings.

A persisting challenge in real-world applications is the specification of the decision-making model assumed for the estimation. While the potential factors affecting decision-making are clear in some conditions, such as exit or route choice problems, these factors are not usually obvious in some real-world applications. The experimental results showed acceptable forecasting performance for interventions, even in the data-model mismatch scenario. However, the results also suggest that the forecasting performance would be degraded if the assumed behavioural model was severely biased or mis-specified. Since reliable parameter estimation would be difficult in the developed framework based on a particle filter when the number of parameters to be estimated becomes large, it is important to consider effective explanatory variables affecting decision-making from many potential variables. The development

of a less-biased data-driven design for the decision-making model, which broadens the applicability of the forecasting method, is an interesting challenge for future studies.

ACKNOWLEDGMENTS

The computational resources of the AI Bridging Cloud Infrastructure (ABCI) provided by the National Institute of Advanced Industrial Science and Technology (AIST) was used.

REFERENCES

[1] Gianluca Antonini, Michel Bierlaire, and Mats Weber. 2006. Discrete choice models of pedestrian walking behavior. *Transportation Research Part B: Methodological* 40, 8 (2006), 667–687.

[2] World Bank. 2022. Urban population (% of total population). (2022). Available at https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS.

[3] Ulf Blanke, Gerhard Tröster, Tobias Franke, and Paul Lukowicz. 2014. Capturing crowd dynamics at large scale events using participatory GPS-localization. In *2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*. 1–7.

[4] Victor J. Blue and Jeffrey L. Adler. 1998. Emergent Fundamental Pedestrian Flows from Cellular Automata Microsimulation. *Transportation Research Record* 1644, 1 (1998), 29–36.

[5] Miodrag Bolic, Petar M Djuric, and Sangjin Hong. 2003. New resampling algorithms for particle filters. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).*, Vol. 2. II–589.

[6] Mingfei Cai, Yanbo Pang, Takehiro Kashiyama, and Yoshihide Sekimoto. 2021. Simulating Human Mobility with Agent-Based Modeling and Particle Filter Following Mobile Spatial Statistics. In *Proceedings of the 29th International Conference on Advances in Geographic Information Systems* (Beijing, China) *(SIGSPATIAL '21)*. Association for Computing Machinery, New York, NY, USA, 411–414.

[7] Robert Clay, Le-Minh Kieu, Jonathan A. Ward, Alison Heppenstall, and Nick Malleson. 2020. Towards Real-Time Crowd Simulation Under Uncertainty Using an Agent-Based Model and an Unscented Kalman Filter. In *Advances in Practical Applications of Agents, Multi-Agent Systems, and Trustworthiness. The PAAMS Collection*, Yves Demazeau, Tom Holvoet, Juan M. Corchado, and Stefania Costantini (Eds.). Springer International Publishing, Cham, 68–79.

[8] Robert Clay, Jonathan A. Ward, Patricia Ternes, Le-Minh Kieu, and Nick Malleson. 2021. Real-time agent-based crowd simulation with the Reversible Jump Unscented Kalman Filter. *Simulation Modelling Practice and Theory* 113 (2021), 102386.

[9] Dorine C. Duives, Winnie Daamen, and Serge P. Hoogendoorn. 2013. State-of-the-art crowd motion simulation models. *Transportation Research Part C: Emerging Technologies* 37 (2013), 193–209.

[10] Dorine C. Duives and Hani S. Mahmassani. 2012. Exit Choice Decisions during Pedestrian Evacuations of Buildings. *Transportation Research Record* 2316, 1 (2012), 84–94.

[11] Claudio Feliciani, Alessandro Corbetta, Milad Haghani, and Katsuhiro Nishinari. 2023. Trends in crowd accidents based on an analysis of press reports. *Safety Science* 164 (2023), 106174.

[12] Tian Feng, Lap-Fai Yu, Sai-Kit Yeung, KangKang Yin, and Kun Zhou. 2016. Crowd-Driven Mid-Scale Layout Design. *ACM Trans. Graph.* 35, 4, Article 132 (jul 2016), 14 pages.

[13] Thomas Gaskin, Grigorios A. Pavliotis, and Mark Girolami. 2023. Neural parameter calibration for large-scale multiagent models. *Proceedings of the National Academy of Sciences* 120, 7 (2023), e2216415120.

[14] Neil J Gordon, David J Salmond, and Adrian FM Smith. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE proceedings F (radar and signal processing)*, Vol. 140. IET, 107–113.

[15] Milad Haghani and Majid Sarvi. 2019. Laboratory experimentation and simulation of discrete direction choices: Investigating hypothetical bias, decision-rule effect and external validity based on aggregate prediction measures. *Transportation Research Part A: Policy and Practice* 130 (2019), 134–157.

[16] Dirk Helbing, Lubos Buzna, Anders Johansson, and Torsten Werner. 2005. Self-Organized Pedestrian Crowd Dynamics: Experiments, Simulations, and Design Solutions. *Transportation Science* 39, 1 (2005), 1–24.

[17] Dirk Helbing and Péter Molnár. 1995. Social force model for pedestrian dynamics. *Phys. Rev. E* 51 (May 1995), 4282–4286. Issue 5.

[18] Dirk Helbing, Péter Molnár, Illés J Farkas, and Kai Bolay. 2001. Self-Organizing Pedestrian Movement. *Environment and Planning B: Planning and Design* 28, 3 (2001), 361–383.

[19] Serge P. Hoogendoorn, Winnie. Daamen, and Ramon Landman. 2007. Microscopic calibration and validation of pedestrian models — Cross-comparison of models using experimental data. In *Pedestrian and Evacuation Dynamics 2005*, Nathalie Waldau, Peter Gattermann, Hermann Knoflacher, and Michael Schreckenberg (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 253–265.

[20] Anders Johansson, Dirk Helbing, and Pradyumn K. Shukla. 2007. Specification of the Social Force Pedestrian Model by Evolutionary Adjustment to Video Tracking Data. *Advances in Complex Systems* 10, supp02 (2007), 271–288.

[21] Eugenia Kalnay. 2002. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press.

[22] Ioannis Karamouzas, Brian Skinner, and Stephen J. Guy. 2014. Universal Power Law Governing Pedestrian Interactions. *Phys. Rev. Lett.* 113 (Dec 2014), 238701. Issue 23.

[23] Genshiro Kitagawa. 1996. Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models. *Journal of Computational and Graphical Statistics* 5, 1 (1996), 1–25.

[24] Jordan Lueck, Jason H. Rife, Samarth Swarup, and Nasim Uddin. 2019. Who goes there? Using an agent-based simulation for tracking population movement. In *2019 Winter Simulation Conference (WSC)*. 227–238.

[25] Fumiyasu Makinoshima, Fumihiko Imamura, and Yoshi Abe. 2018. Enhancing a tsunami evacuation simulation for a multi-scenario analysis using parallel computing. *Simulation Modelling Practice and Theory* 83 (2018), 36–50.

[26] Fumiyasu Makinoshima and Yusuke Oishi. 2022. Crowd flow forecasting via agent-based simulations with sequential latent parameter estimation from aggregate observation. *Scientific Reports* 12, 1 (July 2022), 11168.

[27] Nicolas Malleson, Kevin Minors, Le-Minh Kieu, Jonathan A. Ward, Andrew West, and Alison Heppenstall. 2020. Simulating Crowds in Real Time with Agent-Based Modelling and a Particle Filter. *Journal of Artificial Societies and Social Simulation* 23, 3 (2020), 3.

[28] Daniel McFadden. 1973. Conditional logit analysis of qualitative choice behavior. In *Economic theory and mathematical economics*, Paul Zarembka (Ed.). Academic Press, New York, Chapter 4, 105–142.

[29] Attila M. Nagy and Vilmos Simon. 2018. Survey on traffic prediction in smart cities. *Pervasive and Mobile Computing* 50 (2018), 148–163.

[30] Ryo Nishida, Masaki Onishi, and Koichi Hashimoto. 2023. Crowd Simulation Incorporating a Route Choice Model and Similarity Evaluation Using Real Large-Scale Data. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems* (London, United Kingdom) *(AAMAS '23)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2751–2753.

[31] Jason H. Rife, Samarth Swarup, and Nasim Uddin. 2019. A Behavior-Based Population Tracker Can Parse Aggregate Measurements to Differentiate Agents. In *2019 IEEE International Symposium on Technologies for Homeland Security (HST)*. 1–5.

[32] Lorenz Schauer, Martin Werner, and Philipp Marcus. 2014. Estimating Crowd Densities and Pedestrian Flows Using Wi-Fi and Bluetooth. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (London, United Kingdom) *(MOBIQUITOUS '14)*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Brussels, BEL, 171–177.

[33] Julio Cezar Silveira Jacques Junior, Soraia Raupp Musse, and Claudio Rosito Jung. 2010. Crowd Analysis Using Computer Vision Techniques. *IEEE Signal Processing Magazine* 27, 5 (2010), 66–77.

[34] Akihito Sudo, Takehiro Kashiyama, Takahiro Yabe, Hiroshi Kanasugi, Xuan Song, Tomoyuki Higuchi, Shin'ya Nakano, Masaya Saito, and Yoshihide Sekimoto. 2016. Particle Filter for Real-Time Human Mobility Prediction Following Unprecedented Disaster. In *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (Burlingame, California) *(SIGSPATIAL '16)*. Association for Computing Machinery, New York, NY, USA, Article 5, 10 pages.

[35] Samarth Swarup and Henning S. Mortveit. 2020. Live Simulations. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems* (Auckland, New Zealand) *(AAMAS '20)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1721–1725.

[36] Louis L Thurstone. 1927. A law of comparative judgment. *Psychol. Rev.* 34, 4 (July 1927), 273–286.

[37] Jur van den Berg, Ming Lin, and Dinesh Manocha. 2008. Reciprocal Velocity Obstacles for real-time multi-agent navigation. In *2008 IEEE International Conference on Robotics and Automation*. 1928–1935.

[38] Minghao Wang and Xiaolin Hu. 2015. Data assimilation in agent based simulation of smart environments using particle filters. *Simulation Modelling Practice and Theory* 56 (2015), 36–54.

[39] Gary White, Anna Zink, Lara Codeca, and Siobhan Clarke. 2021. A digital twin smart city for citizen feedback. *Cities* 110 (2021), 103064.

[40] David Wolinski, Stephen J. Guy, Anne-Hélène Olivier, Ming Lin, Dinesh Manocha, and Julien Pettré. 2014. Parameter estimation and comparative evaluation of crowd simulations. *Computer Graphics Forum* 33, 2 (2014), 303–312.

[41] Xiaoping Zheng, Tingkuan Zhong, and Mengting Liu. 2009. Modeling crowd evacuation of a building based on seven methodological approaches. *Building and Environment* 44, 3 (2009), 437–445.