Specification, estimation and validation of a pedestrian walking behavior model

G. Antonini* M. Bierlaire † S. Schneider‡ Th. Robin †

July 27, 2007

Report TRANSP-OR 070727
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
transp-or.epfl.ch

*Business Optimization Group, Computer Science Dept., IBM Research GmbH, Zurich Research Laboratory, CH-8803 Ruschlikon, Switzerland, GAN@zurich.ibm.com

†Transp-OR, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, {michel.bierlaire, thomas.robin}@epfl.ch

‡Operations Research Group ROSO, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, sabina.schneider@epfl.ch

Abstract

We propose and validate a model for pedestrian walking behavior, based on discrete choice modeling. Two main behaviors are identified: *unconstrained* and *constrained*. The constrained patterns are captured by a *leader-follower* model and by a *collision avoidance* model. The spatial correlation between the alternatives is taken into account defining a cross nested logit model. The model is estimated by maximum likelihood estimation on a real data set of pedestrian trajectories, manually tracked from video sequences. The model is validated using a bi-directional flow data set, collected in controlled experimental conditions at Delft university.

1 Introduction

Pedestrian behavior modeling is an important topic in different contexts. Architects are interested in understanding how individuals move into buildings to find out optimality criteria for space design. Transport engineers face the problem of integration of transportation facilities, with particular emphasis on safety issues for pedestrians. Recent tragic events have increased the interest for automatic video surveillance systems, able to monitor pedestrian flows in public spaces, throwing alarms when abnormal behaviors occur. Special emphasis has been given to more specific evacuation scenarios, for obvious reasons. In this spirit, it is important to define mathematical models based on specific (and context-dependent) behavioral assumptions, tested by means of proper statistical methods. Data collection for pedestrian dynamics is particularly difficult and few models presented in the literature have been calibrated and validated on real data sets.

Previous methods for pedestrian behavior modeling can be classified into two main categories: *microscopic* and *macroscopic* models. In the last years much more attention has been focused on microscopic modeling, where each pedestrian is modeled as an agent, individually. Examples of microscopic models are the *social forces* model in Helbing and Molnar (1995) and Helbing et al. (2002) where the authors use Newtonian mechanics with a continuous space representation to model long-range interactions, and the multi-layer utility maximization model by Hoogendoorn et al. (2002) and Daamen (2004). Blue and Adler (2001) and Schadschneider (2002) use cellular automata models, characterized by a static discretization of the space where each cell in the grid is represented by a state variable. Another microscopic approach is based on space syntax theory where people move through spaces following criteria of space visibility and accessibility (see Penn and Turner, 2002) and minimizing angular paths (see Turner, 2001). Finally, Borgers and Timmermans (1986), Whynes et al. (1996) and Dellaert

et al. (1998) focus on destination and route choice problems on network topologies. For a general literature review on pedestrian behavior modeling we refer the interested reader to Bierlaire et al. (2003).

Leader-follower and collision avoidance behaviors have been addressed here to face with interactions between pedestrians. Existing literature has shown the occurrence of self-organizing processes in crowded environments. At certain levels of density, interactions between people give rise to lane formation. In order to model these effects formally, we took inspiration from previous car following models in transport engineering (including Newell, 1961, Herman and Rothery, 1965, Lee, 1966, Ahmed, 1999). The main idea in these models is that two vehicles are involved in a car following situation when a subject vehicle follows a leader, normally represented by the vehicle in front, reacting to its actions. In general, a sensitivity-stimulus framework is adopted. According to this framework a driver reacts to stimuli from the environment, where the stimulus is usually the leader relative speed. Different models differ in the specification of the sensitivity term. This modeling idea is extended here and adapted to the more complex case of pedestrian behavior. We want to stress the fact that in driver behavior modeling a distinction between acceleration behavior and direction change (lane change) behavior is almost natural (see Toledo, 2003 and Toledo et al., 2003), being imposed by the transport facility itself. The pedestrian case is more complex, the movements being two-dimensional on the walking plane, where acceleration and direction changes are not easily separable. The collision avoidance pattern and the constrained behaviors in general are also inspired by studies in human sciences and psychology, leading to the concept of *personal space* (see Horowitz et al., 1964, Dosey and Meisels, 1969 and Sommer, 1969). Personal space is a protective mechanism founded on the ability of the individual to perceive signals from one's physical and social environment. Its function is to create the spacing patterns that regulate distances between individuals and on which individual behaviors are based (Webb and Weber, 2003). Helbing and Molnar (1995) in their social forces model use the term "territorial effect". Several studies in psychology and sociology show how individual characteristics influence the perception of the space and interpersonal distance. Brady and Walker (1978) found for example that anxiety states are positively correlated with interpersonal distance. Similarly, Dosey and Meisels (1969) found that individuals establish greater distances in high-stress conditions. Hartnett et al. (1974) found that male and female individuals approached short individuals more closely than tall individuals. Other studies (Phillips, 1979 and Sanders, 1976) indicate that the other person's body size influences space.

2 Modeling framework

In this work we refer to the general framework for pedestrian behavior described by Hoogendoorn (2003) and Daamen (2004). Individuals make different decisions, following a hierarchical scheme: *strategical*, *tactical* and *operational*. Briefly, destinations and activities are chosen at a strategical level; the order of the activity execution, the activity area choice and route choice are performed at the tactical level, while instantaneous decisions are taken at the operational level. In this paper we address the problem of pedestrian walking behavior, naturally identified by the operational level of the hierarchy just described. We are interested in modeling the short range behavior in *normal* conditions, as a reaction to the surrounding environment and to the presence of other individuals. With the term "normal" we refer to non-evacuation and non-panic situations.

The motivations and the soundness of discrete choice methods have been addressed in our introductory work (Antonini et al., 2006). The objective of this paper is twofold. First, we aim to provide an extended disaggregate, fully estimable behavioral model, calibrated on real pedestrian trajectories manually tracked from video sequences. Second, we want to test the coherence, interpretability and generalization power of the proposed specification through a detailed validation on external data. Compared with Antonini et al. (2006), we present three important contributions: (i) we estimate the model using significantly more data representing revealed walking behavior, coming from two different sources, (ii) the model specification explicitly captures leader-follower and collision-avoidance patterns and (iii) the model is successfully validated on experimental data, not involved in the estimation process.

We illustrate in Figure 1 the operational framework. The unconstrained decisions are independent from the presence of other pedestrians and are generated by subjective and/or unobserved factors. The first of these factors is represented by the individual's destination. It is assumed to be exogenous to the model and decided at the strategical level. The second factor is represented by the tendency of people to keep their current direction, minimizing their angular displacement. Finally, unconstrained accelerations (with accelerations we mean both positive and negative speed variations) are dictated by the individual desired speed. The implementation of these ideas is made through the three unconstrained patterns indicated in Figure 1.

We assume that behavioral constraints are induced by the interactions with the other individuals in the scene. The *collision avoidance* pattern is designed to capture the effects of possible collisions on the current trajectory of the decision maker. The *leader-follower* pattern is designed to capture the tendency of people to follow another individual in a crowd, in order to benefit from the

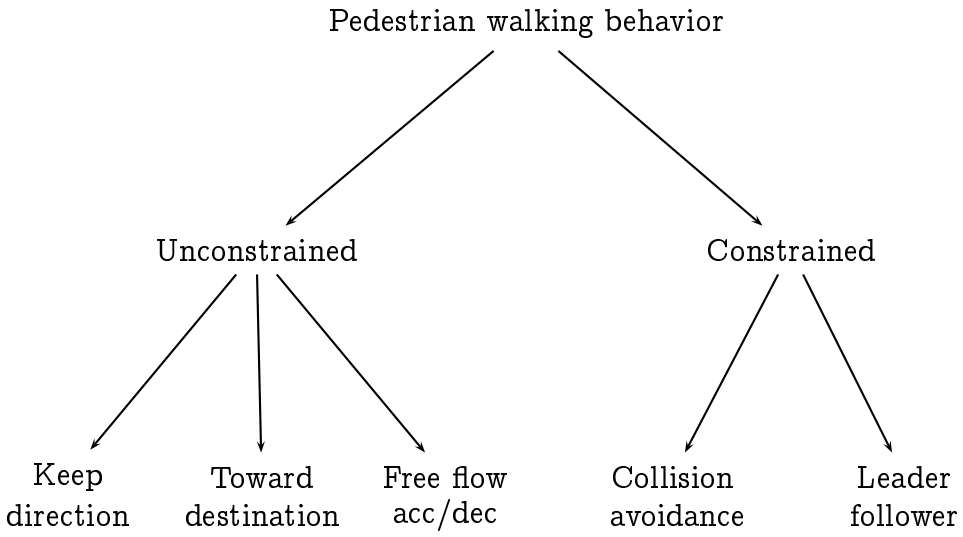


Figure 1: Conceptual framework for pedestrian walking behavior

space she is creating.

The discrete choice model introduced by Antonini et al. (2006) is extended here. The basic elements are the same and summarized below. Pedestrian movements and interactions take place on the horizontal walking plane. The spatial resolution depends on the current speed vector of the individuals. The geometrical elements of the space model are illustrated in Figure 2.

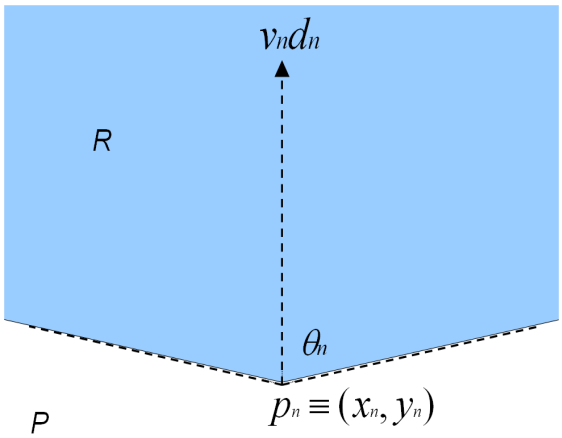


Figure 2: The basic geometrical elements of the space structure

The current position of the decision maker n is p_n, her current speed $v_n \in \mathbb{R}$, her current direction is $d_n \in \mathbb{R}^2$ (normalized, so that $\|d_n\| = 1$) and her visual

angle is θ_n. The region of interest is situated in front of the pedestrian, ideally overlapping with her visual field, and is schematically represented by the shaded area in Figure 2. An adaptive discretization is obtained assuming three speed regimes, where the individual can accelerate up to two times the speed and decelerate up to half time the speed or can maintain the current speed. These hypothesis seem to be coherent with real pedestrian movements. Therefore, for a given time t, the next position will lie into one of the zones, as depicted in Figure 3 (left). A choice between 11 radial directions is allowed, as illustrated in Figure 3 (right).

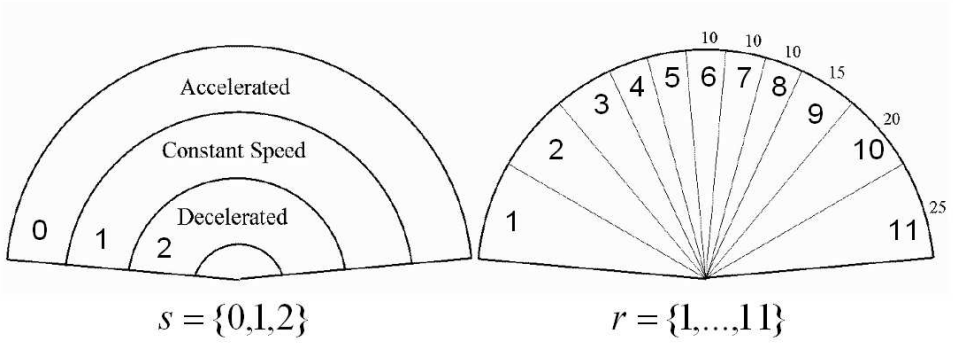


Figure 3: The spatial discretization is generated assuming three speed regimes and 11 radial directions. The external numbers in the right-hand figure represent the angular amplitudes of the radial cones, in degrees. The internal numbers (r) refer to the enumeration of directions while s in the left-hand figure represents the indexes used for speed regimes

A choice set of 33 alternatives is generated where each alternative corresponds to a speed regime v and a radial direction d. They are numbered using $na = 11s + r$ where na is the index of the alternative, s and r are, respectively, the speed regime and the direction indices, as reported in Figure 3. Each alternative is identified by the physical center of the corresponding cell in the spatial discretization c_{vd}, that is

$$c_{vd} = p_n + vtd,$$

where t is the time step. The choice set varies with direction and speed therefore the distance between an alternative's center and other pedestrians will vary with the speed of the decision maker. As a consequence, differences in individual speeds are naturally mapped into differences in their relative interactions.

3 Behavioral patterns

Individuals walk on a 2D plane and we model two kind of behavior: changes in direction and changes in speed, i.e. accelerations. This specification is important to perform walking behavior analysis, and hypotheses have to be made about the unobserved factors in the model and how they are related to the observed data (see Figure 1). Five behavioral patterns are defined. In a discrete choice context, they have to be considered as competitive terms entering the utility functions of each alternative, as reported in Equation 1. The utilities describe the space around the decision maker and under the rational behavior assumption the individual chooses that location (alternative) with the maximum utility. In the following, we discuss the different patterns and the associated assumptions in more details.

3.1 Unconstrained patterns

The unconstrained patterns are identified by those behaviors that are independent from the presence of other pedestrians. We assume that three factors influence the individual behavior.

- **Toward destination** The first factor is represented by the choice of the final destination which can be a specific area where the individual wants to perform an activity in her schedule. To be coherent with the general framework introduced in Section 1, we assume that the destination choice is performed at the strategical level in the hierarchical decision process. Such a higher level choice is naturally reflected on the short term behavior as the tendency of individuals to choose, for the next step, a spatial location that minimize both the angular displacement and the distance to the destination.

- **Keep direction** The second factor influencing the unconstrained behavior is represented by the tendency of people to avoid frequent changings in direction. People choose their next position in order to minimize the angular displacement from their current movement direction. In addition to the behavioral motivation of this factor, it also plays a smoothing role in the model, avoiding drastic changes of direction from one time period to the next.

- **Free flow acceleration** In free flow conditions the behavior of the individual is driven by her desired speed. The acceleration is then a function of the difference between current speed and desired speed. However, this factor is an unobserved individual characteristic and it cannot be introduced

explicitly in the model. As a consequence, we assume that the attractiveness of an individual for an acceleration is dependent on her current speed value. Increasing speed values correspond to decreasing attractiveness for further accelerations. A similar idea is applied to decelerations (see Antonini et al., 2006).

3.2 Constrained patterns

Constrained behaviors are induced by the presence of other individuals in the scene and capture the pedestrian-pedestrian interactions. We identify the following patterns:

- **Leader-follower** We assume that the decision maker is influenced by leaders. In our spatial representation 11 radial cones partition the space (see Figure 3). In each of these directions a possible leader can be identified among a set of *potential leaders*. A potential leader is an individual which is inside a certain region of interest, *not so far* from the decision maker and with a moving direction *close enough* to the direction of the radial cone where she is. Among the set of potential leaders for each radial direction, one of them is selected as leader for that direction (the closest to the decision maker). Once identified, the leader induces an attractive interaction on the decision maker. Similarly to car following models, a leader acceleration corresponds to a decision maker acceleration.

- **Collision avoidance** This pattern captures the effects of possible collisions on the decision maker trajectory. For each direction in the choice set, a collider is identified among a set of *potential colliders*. Another individual is selected as a potential collider if she is inside a certain region of interest, *not so far* from the decision maker and walking against the decision maker herself. The collider for a radial direction is chosen from the set of potential colliders for that direction as the individual whose walking direction forms the larger angle with the decision maker walking direction. This pattern is associated with repulsive interactions in the obvious sense that pedestrians change their current direction to avoid collisions with other individuals.

4 The model

Following the framework proposed in Figure 1 we report here the systematic utility as perceived by individual n for the alternative identified by the speed

regime v and direction d:

$$
\begin{aligned}
V_{vdn} = \beta_{dir} dir_{dn} \quad &+ \Big\} && \textit{keep direction} \\
\beta_{ddist} ddist_{vdn} \quad &+ \Big\} && \textit{toward destination} \\
\beta_{ddir} ddir_{dn} \quad &+ \Big\} && \\
\beta_{acc} I_{v,acc} (v_n / v_{max})^{\lambda_{acc}} \quad &+ \Big\} && \textit{free flow acceleration} \\
\beta_{dec} I_{v,dec} (v_n / v_{max})^{\lambda_{dec}} \quad &+ \Big\} && \\
I_{v,acc} I^L_{acc} \alpha^L_{acc} D_L^{\rho^L_{acc}} \Delta v_L^{\gamma^L_{acc}} \Delta\theta_L^{\delta^L_{acc}} \quad &+ \Big\} && \textit{leader-follower} \\
I_{v,dec} I^L_{dec} \alpha^L_{dec} D_L^{\rho^L_{dec}} \Delta v_L^{\gamma^L_{dec}} \Delta\theta_L^{\delta^L_{dec}} \quad &+ \Big\} && \\
I_{d,d_n} I_C \alpha_C e^{-\rho_C D_C} \Delta v_C^{\gamma_C} \Delta\theta_C^{\delta_C} \quad &\Big\} && \textit{collision avoidance}
\end{aligned}
\tag{1}
$$

where all the β parameters as well as λ_{acc}, λ_{dec}, α^L_{acc}, ρ^L_{acc}, γ^L_{acc}, δ^L_{acc}, α^L_{dec}, ρ^L_{dec}, γ^L_{dec}, δ^L_{dec}, α_C, ρ_C, γ_C, δ_C are unknown and have to be estimated. Note that this specification is the result of an intensive modeling process, where many different specifications have been tested. We explain in the following the different terms of the utilities.

- **Keep direction** This behavior is captured by the term

  $$\beta_{dir} dir_{dn}$$

  where the variable dir_{dn} is defined as the angle in degrees between direction d and direction d_n, corresponding to the central cone, as shown in Figure 4. For the β_{dir} parameter we expect a negative sign.

- **Toward destination** This behavior is captured by the term

  $$\beta_{ddist} ddist_{vdn} + \beta_{ddir} ddir_{dn}$$

  where the variable $ddist_{vdn}$ is defined as the distance (in meters) between the destination and the center of the alternative C_{vdn}, while $ddir_{dn}$ is defined as the angle in degrees between the destination and the alternative's direction d, as shown in Figure 4. We expect a negative sign for both the β_{ddir} and β_{ddist} parameters.

- **Free flow acceleration** We define two parameters for the free flow acceleration (deceleration) terms, $\tilde{\beta}_{acc}$ and $\tilde{\beta}_{dec}$:

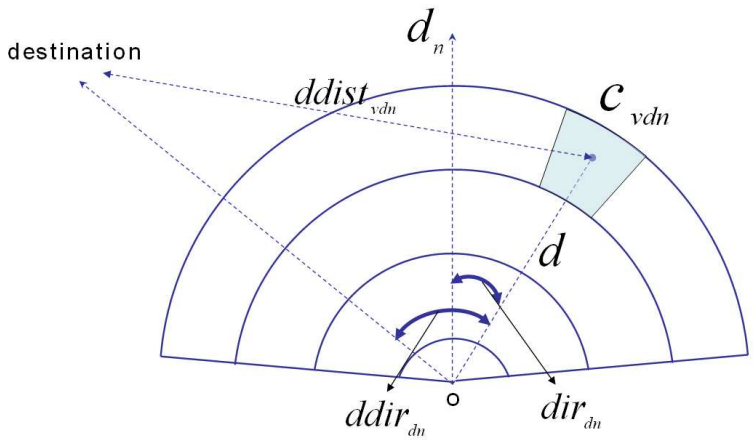


Figure 4: The elements capturing the *keep direction* and *toward destination* behaviors

$$\begin{aligned}\tilde{\beta}_{acc} &= I_{v,acc}\beta_{acc}(v_n/v_{max})^{\lambda_{acc}},\\ \tilde{\beta}_{dec} &= I_{v,dec}\beta_{dec}(v_n/v_{max})^{\lambda_{dec}}\end{aligned}$$

The attribute $I_{v,acc}$ is 1 if $v = v_{acc}$, that is, if the alternative corresponds to an acceleration and 0 otherwise. $I_{v,dec}$ is similarly defined. The two parameters are non-linear functions of the current speed of the decision-maker v_n. β_{acc} is the value of the parameter associated with $v_n = v_{max}$ and λ_{acc} is the elasticity of the parameter with respect to speed. v_{max} represents the maximum value of the observed speed module. We expect negative signs for the β_{acc}, β_{dec} and λ_{dec} parameters, while a positive sign is expected for λ_{acc}.

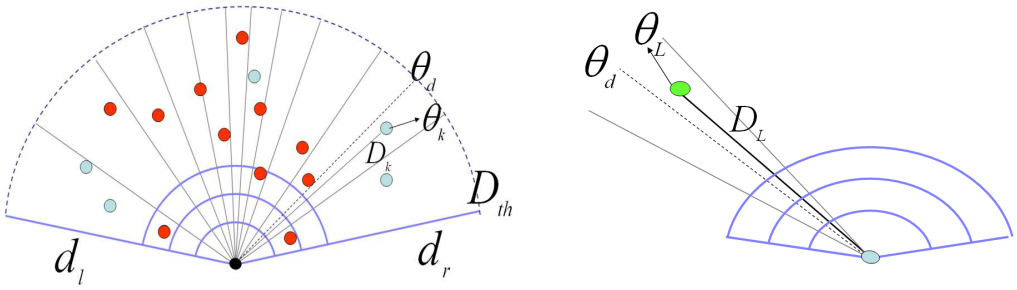
- **Leader-follower** The leader-follower model captures the attractive interactions among pedestrians and is given by the following terms

$$I_{v,acc}I^L_{acc}\alpha^L_{acc}D^{\rho^L_{acc}}_L\Delta v^{\gamma^L_{acc}}_L\Delta\theta^{\delta^L_{acc}}_L + I_{v,dec}I^L_{dec}\alpha^L_{dec}D^{\rho^L_{dec}}_L\Delta v^{\gamma^L_{dec}}_L\Delta\theta^{\delta^L_{dec}}_L.$$

It is described by a *sensitivity/stimulus* framework. The leader for each direction is chosen considering several *potential leaders*, as shown in Figure 5(a). An individual k is defined as a potential leader based on the following indicator function:

$$I^k_g = \begin{cases} 1, & \text{if } d_l \le d_k \le d_r \text{ (is in the cone),}\\ & \text{and } 0 < D_k \le D_{th} \text{ (not too far),}\\ & \text{and } 0 < |\Delta\theta_k| \le \Delta\theta_{th} \text{ (walking in almost the same direction),}\\ 0, & \text{otherwise,}\end{cases}$$

where d_l and d_r represent the bounding left and right directions of the choice set (defining the region of interest) while d_k is the direction identifying the pedestrian k position. D_k is the distance between pedestrian k and the decision maker, $\Delta\theta_k = \theta_k - \theta_d$ is the difference between the movement direction of pedestrian k (θ_k) and the angle characterizing direction d, i.e. the direction identifying the radial cone where individual k lies (θ_d). The two thresholds D_{th} and $\Delta\theta_{th}$ are fixed at the values $D_{th} = 5D_{max}$, where D_{max} is the radius of the choice set, and $\Delta\theta_{th} = 10$ degrees. We assume an implicit *leader choice* process, executed by the decision maker herself and modeled choosing as leader for each direction the potential leader at the minimum distance $D_L = \min_{k\in K}(D_k)$, illustrated in Figure 5(a) by the darker circles. Finally, the indicator functions $I_{v,acc}$ and $I_{v,dec}$ discriminate between accelerated and decelerated alternatives, as for the free flow acceleration model.



(a) (b)

Figure 5: Figure 5(a) illustrates how many potential leaders are considered for each direction and how only the nearest one is chosen as leader for a specific direction (darker circles). Figure 5(b) shows the leader's movement direction, θ_L, the direction of the radial cone where the leader lies, θ_d, and her distance from the decision maker, D_L, used in the definitions of both the sensitivity and the stimulus terms

For a given leader, the sensitivity is described by

$$\text{sensitivity} = f(D_L) = \alpha_g^L D_L^{\rho_g^L} \tag{2}$$

where D_L represents the distance between the decision maker and the leader. The parameters α_g^L and ρ_g^L have to be estimated and $g = \{acc, dec\}$ indicates when the leader is accelerating with respect to the decision maker. Both α_{acc}^L and α_{dec}^L are expected to be positive while a negative sign is expected for ρ_{acc}^L and ρ_{dec}^L.

The decision maker reacts to stimuli coming from the chosen leader. We model the stimulus as a function of the leader's relative speed Δv_L and the leader's relative direction $\Delta\theta_L$ as follows:

$$\text{stimulus} = g(\Delta v_L, \Delta\theta_L) = \Delta v_L^{\gamma_g^L} \Delta\theta_L^{\delta_g^L} \tag{3}$$

with $\Delta v_L = |v_L - v_n|$, where v_L and v_n are the leader's speed module and the decision maker's speed module, respectively. The variable $\Delta\theta_L = \theta_L - \theta_d$, where θ_L represents the leader's movement direction and θ_d is the angle characterizing direction d, as shown in Figure 5(b). Positive signs are expected for both the γ_{acc}^L and γ_{dec}^L parameters, while we expect a negative sign for both the δ_{acc}^L and δ_{dec}^L. A leader acceleration induces a decision maker's acceleration. A substantially different movement direction in the leader reduces the influence of the latter on the decision maker.

- **Collision avoidance** The collision avoidance model captures the repulsive interactions among pedestrians and is given by the following term

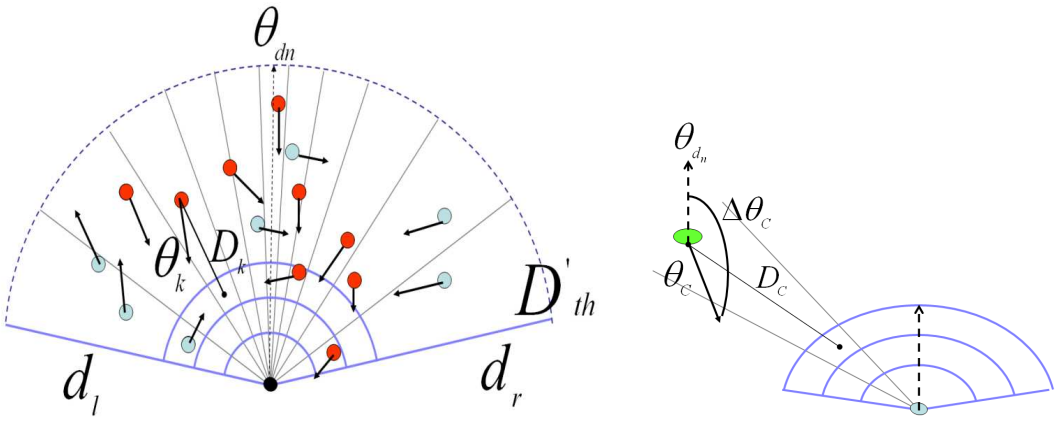
$$I_{d,d_n} I_C \alpha_C e^{-\rho_C D_C} \Delta v_C^{\gamma_C} \Delta\theta_C^{\delta_C}.$$

The collider for each direction is chosen considering several *potential colliders*, as shown in Figure 6(a). An individual k is defined as a potential collider based on the following indicator function:

$$I_C^k = \begin{cases} 1, & \text{if } d_l \le d_k \le d_r \ \text{ (is in the cone),} \\ & \text{and } 0 < D_k \le D'_{th} \ \text{ (not too far),} \\ & \text{and } \frac{\pi}{2} \le |\Delta\theta_k| \le \pi \ \text{ (walking in the other direction),} \\ 0, & \text{otherwise,} \end{cases}$$

where d_l, d_r and d_k are the same as those defined for the leader-follower model. D_k is now the distance between individual k and the center of the alternative, $\Delta\theta_k = \theta_k - \theta_{d_n}$ is the difference between the movement direction of pedestrian k, θ_k, and the movement direction of the decision maker, θ_{d_n}. The value of the distance threshold is now fixed to $D'_{th} = 10 D_{max}$. We use a larger value compared to the leader-follower model, assuming the collision avoidance behavior being a longer range interaction, happening also at a lower density level. We assume an implicit *collider choice* process, which is deterministic and decision-maker specific. Among the set of K_d potential colliders for direction d, a collider is chosen in each cone as that individual having $\Delta\theta_C = \max_{k \in K_d} |\Delta\theta_k|$. The related indicator function is I_C. Finally, the collision avoidance term is included in the

utility functions of all the alternatives, with the exception of the central ones which are used as references. So, the indicator function I_{d,d_n} is equal to 1 for those alternatives that are not in the current direction ($d \neq d_n$), 0 otherwise.



(a) (b)

Figure 6: Figure 6(a) shows many potential colliders taken into account for each direction. Figure 6(b) shows the collider and decision maker movement directions, θ_C and θ_{dn} respectively. D_C represents here the distance of the collider with the center of the alternative

We apply a similar sensitivity/stimulus framework, where the sensitivity function is defined as

$$\text{sensitivity} = f(D_C) = \alpha_C e^{-\rho_C D_C} \tag{4}$$

where the parameters α_C and ρ_C, that have to be estimated, are expected to have both a negative sign and D_C is the distance between the collider position and the center of the alternative, as shown in Figure 6(b). We choose the exponential to keep the same functional form as that used in Antonini et al. (2006). The decision maker reacts to stimuli coming from the collider. We model the stimulus as a function of two variables:

$$\text{stimulus} = f(\Delta v_C, \Delta\theta_C) = \Delta v_C^{\gamma_C} \Delta\theta_C^{\delta_C} \tag{5}$$

with $\Delta\theta_C = \theta_C - \theta_{d_n}$, where θ_C is the collider movement direction and θ_{d_n} is the decision maker movement direction, and $\Delta v_C = v_C + v_n$, where v_C is the collider's speed module and v_n is the decision maker's speed module. The parameters γ_C and δ_C have to be estimated and a positive sign is expected for both of them. Individuals walking against the decision maker at higher speeds and in more frontal directions (higher $\Delta\theta_C$) generate stronger reactions, weighted by the sensitivity function.

We use the cross nested logit (CNL) specification used in Antonini et al. (2006). Such a model allows flexible correlation structures in the choice set, keeping a closed form solution. The CNL being a Multivariate Extreme Value model (MEV, see McFadden, 1978), the probability of choosing alternative i within the choice set C is:

$$P(i|C) = \frac{y_i \frac{\partial G}{\partial y_i}(y_1, ..., y_J)}{\mu G(y_1, ..., y_J)} \tag{6}$$

where J is the number of alternatives in C, $y_j = e^{V_j}$ with V_j the systematic part of the utility described by (1) and G is the following generating function:

$$G(y_1, ..., y_J) = \sum_{m=1}^{M} \left(\sum_{j \in C} (\alpha_{jm}^{1/\mu} y_j)^{\mu_m} \right)^{\frac{\mu}{\mu_m}} \tag{7}$$

where M is the number of nests, $\alpha_{jm} \geq 0, \forall j, m$, $\sum_{m=1}^{M} \alpha_{jm} > 0, \forall j$, $\mu > 0$, $\mu_m > 0, \forall m$ and $\mu \leq \mu_m, \forall m$. This formulation leads to the following expression for the choice probability formula, using $y_i = e^{V_i}$:

$$P(i|C) = \sum_{m=1}^{M} \frac{\left(\sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} y_j^{\mu_m} \right)^{\frac{\mu}{\mu_m}}}{\sum_{n=1}^{M} \left(\sum_{j \in C} \alpha_{jn}^{\mu_n/\mu} y_j^{\mu_n} \right)^{\frac{\mu}{\mu_n}}} \frac{\alpha_{im}^{\mu_m/\mu} y_i^{\mu_m}}{\sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} y_j^{\mu_m}} \tag{8}$$

We assume a correlation structure depending on the speed and direction and we identify five nests: *accelerated*, *constant speed*, *decelerated*, *central* and *not central*. This correlation structure is illustrated in Figure 7. Given the lack of any a priori information, we fix the degrees of membership to the different nests (α_{jm}) to the constant value 0.5. The parameter μ is normalized to 1, and the nest parameters μ_m are estimated.

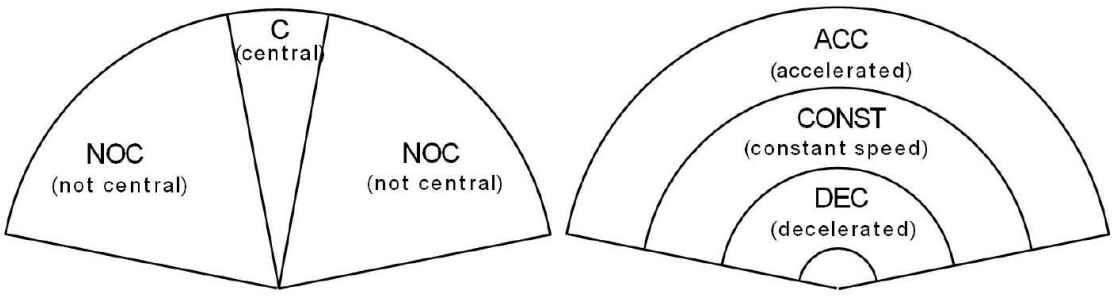


Figure 7: **left:** Nesting based on direction **right:** Nesting based on speed

(a) Japanese scenario (b) Swiss scenario

Figure 8: Images from the two scenarios used to collect the data set

5 Data

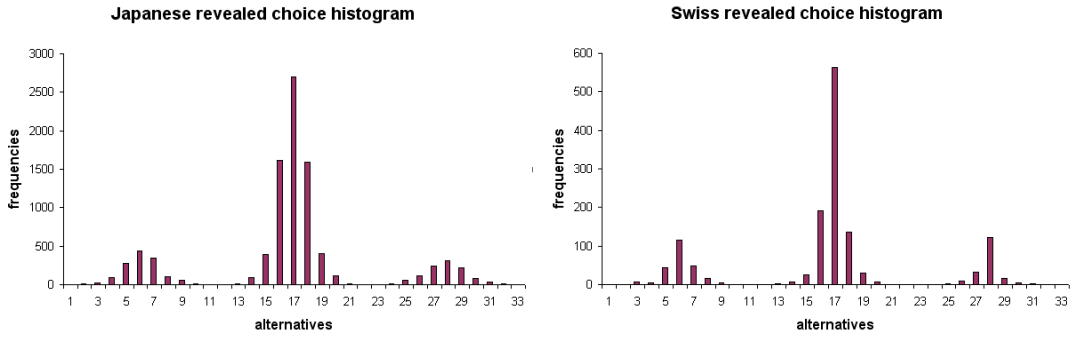
The data set used to estimate the model consists of pedestrian trajectories manually tracked from video sequences. We have pooled together two different data sets, collected separately in Switzerland and Japan.

The Swiss data set This part of the data set consists of 36 pedestrian trajectories, manually tracked from a digital video sequence. The scene has been recorded out of the Flon metro station in Lausanne, in 2002, for a total of 1675 observed positions. Each position refers to a reference system on the walking plane, after a calibration of the camera. For a detailed description of this first data set we refer the reader to Antonini et al. (2006).

The Japanese data set This data set has been collected in Sendai, Japan, on August 2000 (see Teknomo et al., 2000, Teknomo, 2002). The video sequence has been recorded from the 6th floor of the JTB parking building (around 19 meters height), situated at a large pedestrian crossing point. Two main pedestrian flows cross the street, giving rise to a large number of interactions. In this context, 190 pedestrian trajectories have been manually tracked, with a time step of 1 second, for a total number of 10200 position observations. The collected data contains the pedestrian identifier, the time step and the image coordinates. The mapping between the image plane and the walking plane is approximated by a 2D-affine transformation, whose parameters are calibrated by linear regression. The reference system on the walking plane has the origin arbitrarily placed on

the bottom left corner of the zebra crossing. The x axis represents the width of the crossing while the y axis is the crossing length.

Two frames from the two video sequences are reported in Figure 8. In Figure 9 we report the frequency of the revealed choices as observed in the two data sets. The three peaks in the distributions arise on the central alternatives (6, 17, 28), as expected.



(a) Japanese data set

(b) Swiss data set

Figure 9: Revealed choices histograms

We report in Figure 11 two examples of trajectories and in Figure 12(a) and Figure 12(b) the related speed-time graphs. In Figure 10 we report the speed histogram and in Table 1 the speed statistics.

Data post-processing The original Swiss data set has been post-processed in order to generate the input data for the estimation process. At each step, the observed choice made by the current decision maker has been measured 3 steps ahead in time, i.e. 0.9 seconds. As a consequence, the last four positions of each trajectory are not used. Moreover, in both the data sets those observations corresponding to a static pedestrian ($v_n = 0$) and those corresponding to an observed choice out of the choice set have been discarded.

When the two data sets are pooled together, we obtain a total of 10783 observations. Their repartition across the nests defined in Figure 7 is detailed in Table 2.

6 Estimation results

We report in Table 3 the estimation results. The parameters have been estimated using the Biogeme package (Bierlaire, 2003, biogeme.epfl.ch). It is a freeware package for the estimation of a wide range of random utility models.

Mean	0.668
Standard Error	0.00355
Median	0.580
Mode	0
Standard Deviation	0.358
Minimum	0
Maximum	3.940

Table 1: Speed statistics

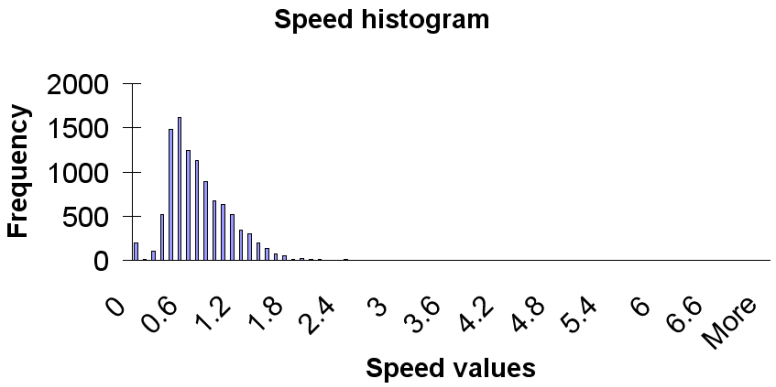


Figure 10: Speed histogram

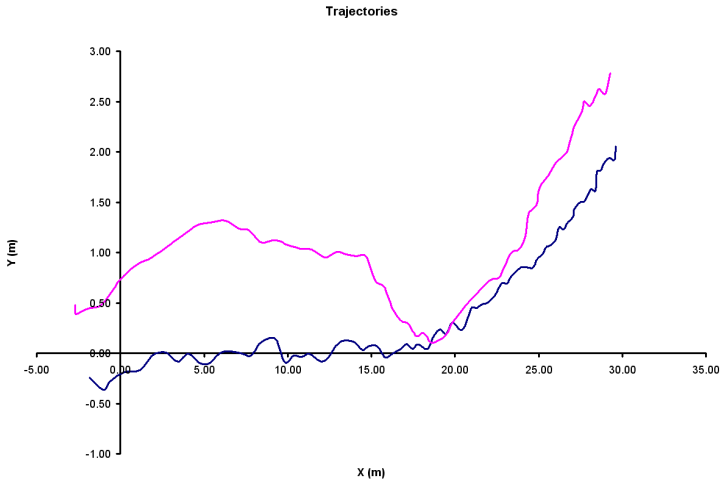
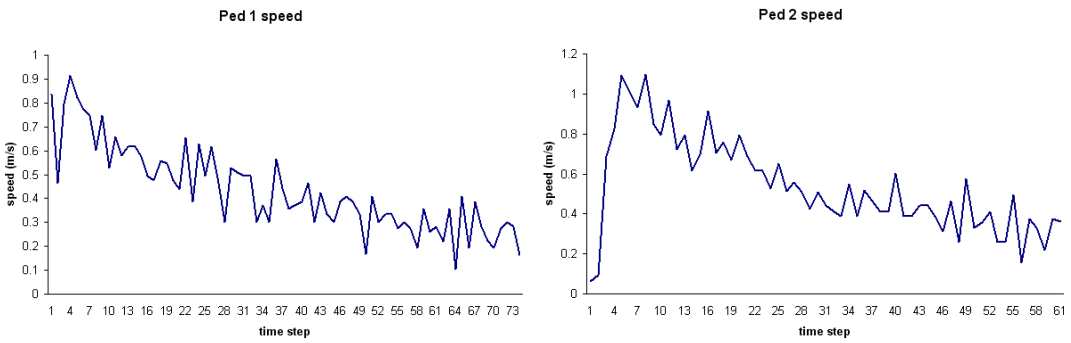


Figure 11: Examples of two manually tracked trajectories



(a) (b)

Figure 12: Speed-time graphs for the trajectories in Figure 11

Nest	# steps	% of total
acceleration	1609	14.92%
constant speed	7894	73.21%
deceleration	1280	11.87%
central	4257	39.48%
not central	6526	60.52%

Table 2: Number of chosen steps in each nest for the real data set

We first shortly comment the results for those parameters related to the unconstrained models (toward destination, keep direction and free flow acceleration). This part of the model specification is similar to that presented in Antonini et al. (2006). The *toward destination* coefficients β_{ddir} and β_{ddist} have been estimated significantly different from zero. The assumption that destination distance and direction capture two different effects is supported by the data, being related to the 2D nature of the pedestrian movements. Their signs are negative, as expected, reflecting the tendency of individuals to move directly towards their final destination, through the shortest path. The destination being exogenous to the model, we interpret this behavior as the short range projection of higher level decisions, made at the tactical level, such as (intermediate) destination choice and/or activity area choice. The *keep direction* parameter, β_{dir}, is significant and has a negative sign, as expected. It captures the tendency of people to minimize the angular displacement along their trajectories. Finally, 3 out of 4 of the *free flow acceleration* parameters, namely β_{acc}, β_{dec} and λ_{acc} have been estimated significantly different from zero. The negative signs for β_{acc} and β_{dec} indicate the tendency of pedestrians to perceive variations in speed as a disutility, both positive and negative. A positive value for the acceleration elasticity λ_{acc} indicates that the attractiveness of an acceleration reduces with

Variable name	Coefficient estimate	t test 0	t test 1
β_{ddir}	-0.075	-11.81	
β_{ddist}	-0.661	-4.06	
β_{dir}	-0.044	-5.61	
β_{acc}	-4.06	-14.86	
β_{dec}	-2.9	-18.30	
λ_{acc}	0.746	18.00	
$\alpha^{\text{L}}_{\text{acc}}$	4.91	3.27	
$\rho^{\text{L}}_{\text{acc}}$	-0.890	-3.78	
$\gamma^{\text{L}}_{\text{acc}}$	0.824	9.18	
$\alpha^{\text{L}}_{\text{dec}}$	3.96	6.53	
$\rho^{\text{L}}_{\text{dec}}$	-0.767	-7.18	
$\gamma^{\text{L}}_{\text{dec}}$	0.431	8.25	
$\delta^{\text{L}}_{\text{dec}}$	-0.0843	-1.31	
α_{C}	-0.0059	-3.86	
ρ_{C}	-0.603	2.40	
γ_{C}	0.287	5.14	
μ_{const}	1.4	11.39	3.26
$\mu_{\text{not\_central}}$	1.04	7.05	0.29
μ_{scale}	0.591	-	-210.31
Sample size = 10783			
Number of estimated parameters = 21			
Init log-likelihood = -26270.8			
Final log-likelihood = -22652.0			
Likelihood ratio test = 30101.6			
$\bar{\rho}^2 = 0.399$			

Table 3: **CNL** estimation results for the pooled data set

increases in speed, as expected. We now comment on the constrained models' parameters. For the *leader-follower* behavior we note that in the case of an accelerating leader, 3 out of 4 parameters have been estimated significantly different from zero. The positive value for the $\alpha^{\text{L}}_{\text{acc}}$ multiplicative coefficient indicates that when a leader is present (or several potential leaders are present, so that the closest to the decision maker is considered), a leader's acceleration

induces a corresponding acceleration on the decision maker. The negative sign for the distance exponential coefficient, $\rho^{\mathrm{L}}_{\mathrm{acc}}$, indicates that the influence of the leader on the decision maker acceleration behavior reduces when their relative distance increases, as expected. The positive sign for the speed exponential coefficient, $\gamma^{\mathrm{L}}_{\mathrm{acc}}$, shows that the utility of an acceleration increases with higher values of the relative leader speed, as expected. The same interpretation is given for the parameters corresponding to a decelerating leader. In this case we keep in the model the exponential coefficient related to the direction, $\delta^{\mathrm{L}}_{\mathrm{dec}}$, with t-test statistics equal to 1.31. Its negative sign is coherent with the leader-follower behavior. It reflects the fact that in those cases where the leader's relative direction is higher, the influence of the leader on the decision maker is lower, resulting in a lower utility value for the decelerated alternatives. The same parameter in the accelerating case, $\delta^{\mathrm{L}}_{\mathrm{acc}}$, is not significant and it has been removed from the model. For the estimation of the *collision avoidance* parameters, we fix the exponential coefficient related to the collider relative direction, δ_{C}, equal to 1 for numerical convenience. The other three free parameters have been estimated significantly different from zero. The multiplicative coefficient α_{C} is negative, as expected. It indicates that those directions more likely to lead to a collision have a lower utility with respect to the central (current) direction. The latter is taken as the reference one for normalization purposes. The exponential coefficient related to the distance between the collider and the alternative, ρ_{C}, has a negative sign. It shows that a more distant collider has a less negative impact on the alternative utility. Finally, the exponential coefficient related to the relative speed, γ_{C}, is positive, as expected. It captures the fact that faster colliders have a more negative impact on the utilities than slower individuals. The correlation structure is captured by the cross nested specification. Three nest parameters have been fixed to 1 while two are left free in the model, capturing the correlation between the constant speed and the not central alternatives. The nest parameter $\mu_{\mathrm{not\_central}}$ is not significantly different from 1. However, we decided to keep it in the model to avoid potential misspecification. Finally, the scale factor (μ_{scale}) for the Swiss data captures the variance ratio of the associated error term between the two data sets. The scale is less than 1, so that the variance of the error term for the Japanese data set is (significantly) lower than the variance of the error term for the Swiss observations.

We conclude this section underlying the fact that it seems natural that individual characteristics such as age, sex, weight, height (among others) influence the spatial perception, interpersonal distance and human-human interactions. However, given the available data (trajectories) it is not possible to take into account such characteristics.

7 Model validation

The validation procedure consists in applying two models on two data sets. In addition to the model presented in Section 4, we consider also a simple model, where the utility of each alternative is represented only by an alternative specific constant (ASC). This ASC model perfectly reproduces the observed shares in the sample, with 31 parameters. Indeed, there are 33 alternatives, minus one which is never chosen, minus one constant normalized to 0. The two data sets are the Swiss-Japanese data set described in Section 5, and a data set collected in the Netherlands, which is described below.

7.1 Swiss-Japanese data set

We first apply our model with the parameters described in Table 3 on the Swiss-Japanese data set, using the Biosim package (Bierlaire, 2003). For each observation n, we obtain a probability distribution $P_n(i)$ over the choice set.

Figure 13 represents the histogram of the probability value $P_n(i_n^*)$ assigned by the model to the chosen alternative i_n^* of each observation n, along with the hazard value $1/33$ (where 33 is the number of alternatives). We consider observations below this threshold as outliers. We observe that there are 12.7% of them.

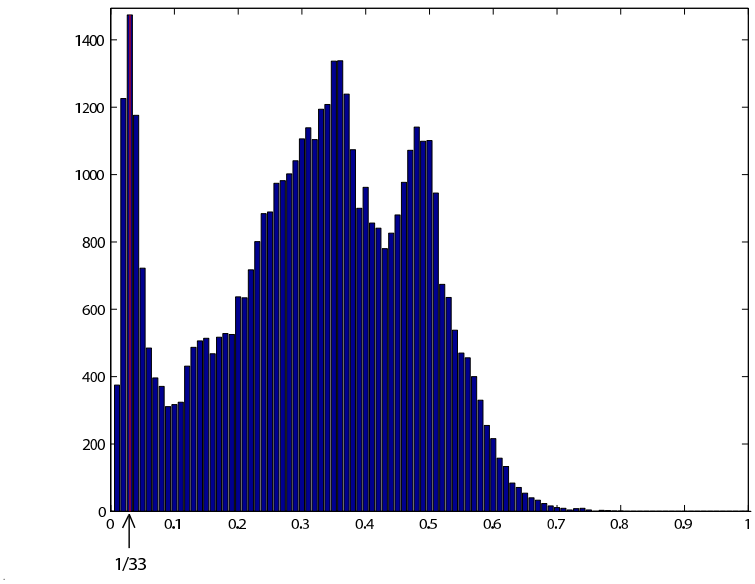


Figure 13: Predicted probabilities of the Swiss-Japanese data

We compare also the estimated model with the ASC model. We observe that our model improves the loglikelihood from -25018.22 up to -22652.0, with less

parameters (21 instead of 31). The number of outliers with the ASC model is 13.7%

The top part of Figure 14 reports, for each i, $\sum_n P_n(i)$, and the bottom part reports $\sum_n y_{in}$, where y_{in} is 1 if alternative i is selected for observation n, 0 otherwise. As expected, the two histograms are similar, indicating no major specification error.

This is confirmed when alternatives are aggregated together, by directions (see Table 4) and by speed regimes (see Table 5). For a group Γ of alternatives, the quantities

$$\begin{aligned} M_\Gamma &= \sum_n \sum_{i\in\Gamma} P_n(i), \\ R_\Gamma &= \sum_n \sum_{i\in\Gamma} y_{in}, \end{aligned}$$

and

$$(M_\Gamma - R_\Gamma)/R_\Gamma$$

are reported in columns 3, 4 and 5, respectively, of these tables.

The relative errors showed in Table 4 and Table 5 are low, except for groups of alternatives with few observations.

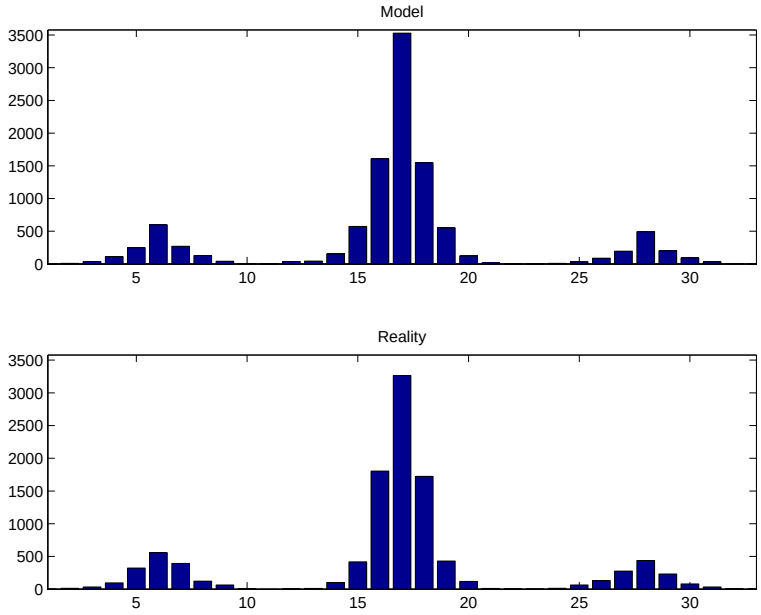


Figure 14: Predicted and observed shares for the Swiss-Japanese data set

We do not discuss the application of the ASC model on this data set as, by design, it reproduces the shares. Actually, we obtain relative errors in the range of 0.0%–0.5%, due to accumulation of rounding errors.

Although the above analysis indicates good specification and performance of the model, it is not sufficient to fully validate it. Consequently, we perform now

Cone	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	$6, 17, 28$	4572.95	4257	7.42%
Left	$3-5, 14-16, 25-27$	3075.46	3245	−5.22%
Right	$7-9, 18-20, 29-31$	3035.96	3197	−5.04%
Extreme left	$1, 2, 12, 13, 23, 24$	70.75	49	44.39%
Extreme right	$10, 11, 21, 22, 32, 33$	27.88	35	−20.34%

Table 4: Predicted (M_Γ) and observed (R_Γ) shares for alternatives grouped by directions with the Swiss-Japanese data set.

Area	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	$1-11$	1579.06	1609	−1.86%
constant speed	$12-22$	7924.63	7894	0.39%
deceleration	$23-33$	1279.30	1280	−0.05%

Table 5: Predicted and observed shares for alternatives grouped by speed regime with the Swiss-Japanese data set.

the same analysis on a validation data set, not involved in the estimation of the model.

7.2 Dutch data set

This data set has been collected at Delft University, in the period 2000-2001 (Daamen and Hoogendoorn, 2003, Daamen, 2004). Volunteer pedestrians are called to perform specific walking tasks in a controlled experimental setup, in order to create specific pedestrian motion patterns such as one-directional flow, bi-directional flow, walking through narrow and wide bottlenecks and crossing flows. A first set of *experimental* variables (free speed, walking direction, density, bottlenecks) are modified during the experiments while a second group of *context* variables are pedestrian-specific.

For the purpose of our validation procedure we use the subset of the Dutch data set corresponding to a bi-directional flow. This situation is the experimental version of the Swiss-Japanese data set, which corresponds to a walkway. The subset includes 724 subjects for 47471 observed positions, collected by means of pedestrian tracking techniques on video sequences, at a frequency of 1Hz. The data format includes a pedestrian identifier, the time step and the x-y coordinates. In Figure 15 we report a typical picture illustrating the experimental scenario. The repartition of the observations across nests defined in Figure 7 is detailed in Table 6. We note the very low number of decelerations.

Figure 15: A representative frame from the video sequences used for data collection

Nest	# steps	% of total
acceleration	5273	11.12%
constant speed	42147	88.78%
deceleration	51	0.12%
central	22132	46.62%
not central	25339	53.38%

Table 6: Number of chosen steps in each nest for Dutch data

We apply our model with the parameters described in Table 3 on the Dutch data set, using the Biosim package. For each observation n, we obtain a probability distribution $P_n(i)$ over the choice set.

Figure 16 represents the histogram of the probability value $P_n(i^*_n)$ assigned by the model to the chosen alternative i^*_n of each observation n, along with the hazard value 1/33 (where 33 is the number of alternatives) illustrating a purely random model with equal probability. Again, we consider observations below this threshold as outliers. We observe that there are 6.56% of them. This is less than for the data set used for parameters estimation. The shape of the curve, as well as the low number of outliers are signs of a good performance of the model.

Applying the estimated model to the Dutch data set, we obtain a loglikelihood of -52676.78. When the ASC model is applied, that is the model replicat-

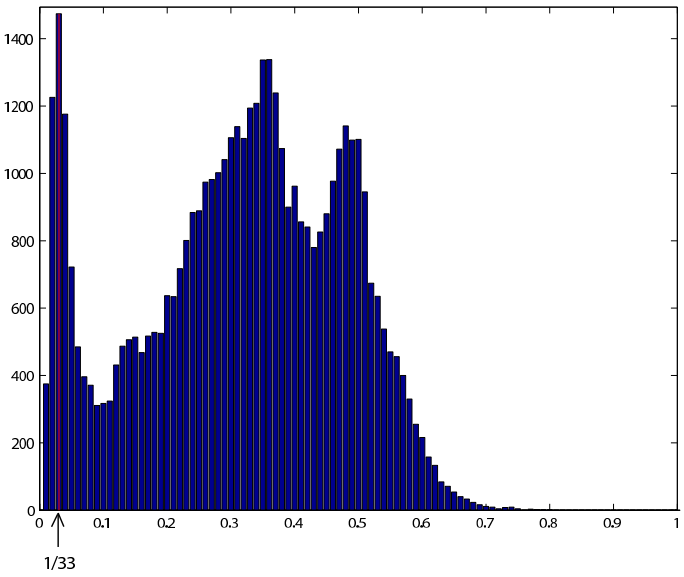


Figure 16: Predicted probabilities for the Dutch data

ing the shares of the Swiss-Japanese data set, the loglikelihood deteriorates to -85565.72. This clearly illustrates the superior forecasting power of our model compared to the simple one.

The top part of Figure 17 reports the predicted probabilities, that is, for each i, $\sum_n P_n(i)$, and the bottom part the observed shares, that is $\sum_n y_{in}$, where y_{in} is 1 if alternative i is selected for observation n, 0 otherwise. We observe some discrepancies between the two histograms. In particular, the model predicts more decelerations (alternatives 22 to 33) and less accelerations (alternatives 1 to 11) compared to reality.

In order to obtain a more robust validation, that is less sensitive to the space discretization, we aggregate alternatives together. By doing so, we decrease the impact of small errors, where predicting neighboring cells is more valid than predicting other cells.

Cone	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	$6, 17, 28$	22032.21	22132	-0.45%
Left	$3-5, 14-16, 25-27$	12566.31	12939	-2.88%
Right	$7-9, 18-20, 29-31$	12659.29	12379	2.26%
Extreme left	$1, 2, 12, 13, 23, 24$	93.99	14	571.35%
Extreme right	$10, 11, 21, 22, 32, 33$	119.20	7	1602.88%

Table 7: Predicted (M_Γ) and observed (R_Γ) shares for alternatives grouped by directions with the Dutch data set.

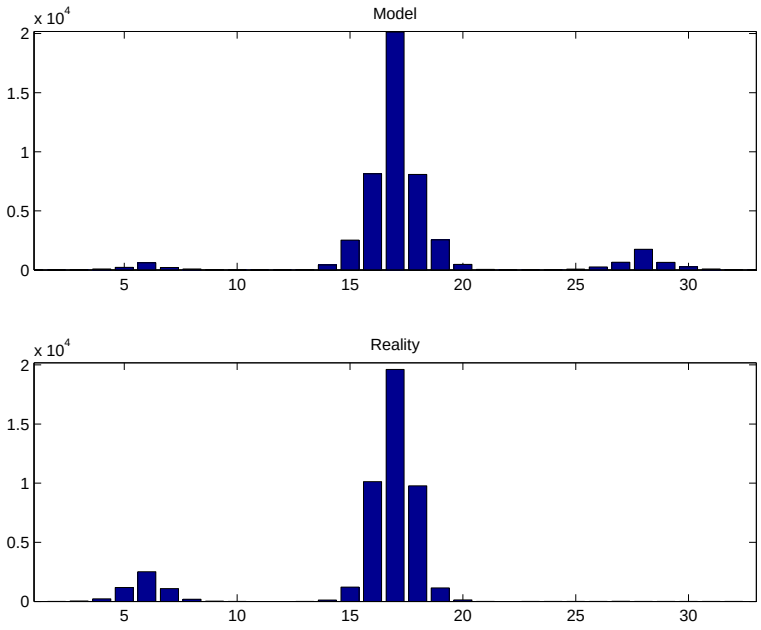


Figure 17: Choice histogram predicted by the model against the revealed choices in the Dutch data set

Area	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	$1-11$	2668.69	5273	-49.39%
constant speed	$12-22$	39292.86	42147	-6.77%
deceleration	$23-33$	5509.46	51	10702.87%

Table 8: Predicted (M_Γ) and observed (R_Γ) shares for alternatives grouped by speed regime with the Dutch data set.

Tables 7 and 8 show that the model predicts well the direction and the constant speed. We confirm the previous observation, that decelerations are over predicted, and accelerations under predicted.

There are two explanations for this phenomenon. First, the Dutch data set was collected in controlled experimental conditions, which may have introduced a bias in pedestrian behavior, depending on the exact instructions they have received. This assumption is supported by the absence of decelerations in the data set. Second, the Dutch pedestrians walk faster than the Japanese, as reported in Table 9 and in Figure 18. In this case, the model can predict a deceleration because of the higher speed value. A similar reasoning holds for accelerations.

The speed distribution is quite different for the Swiss-Japanese data set and the Dutch data set as shown in Figure 18. Indeed, the Dutch distribution seems to be Gaussian with high mean speed, which characterize the experimental conditions; while the Swiss-Japanese distribution is left-centered, relevant for a real

situation where are more interactions (higher density of population). In this case, high speeds are rare events.

Data Set	Mean speed [m/s]
Dutch (experimental)	1.27
Japanese (real)	0.69
Swiss (real)	1.46

Table 9: Average pedestrian speed in the data sets

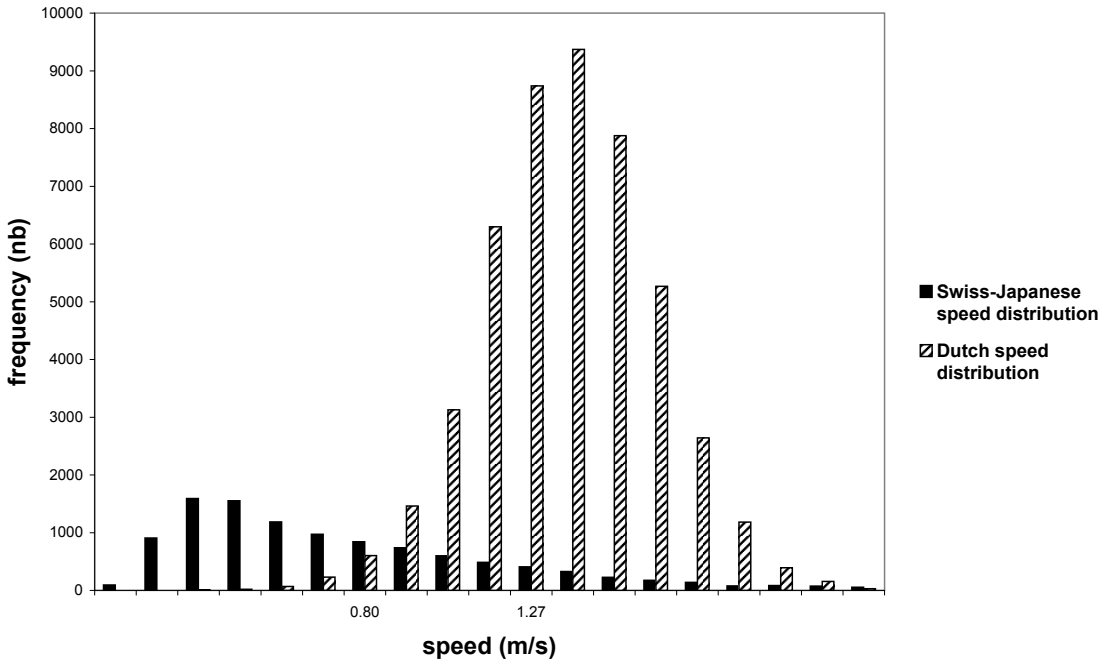


Figure 18: Distribution of speed in the two data sets

We have applied the simple ASC model in the Dutch data set. As expected, it is less powerful for prediction than the developed model (see Tables 10 and 11).

For the sake of completeness, an ASC model has been calibrated on the Dutch data set, in the same way than for the Swiss-Japanese. Our model estimated on the Swiss-Japanese data is better than the ASC model estimated on the Dutch data, when applied on the Dutch data set, both for log-likelihood (-52676.78 against -77871.06) and prediction (6.76 %, percentage of bad observations against 11.45 %). We have summarized the various loglikelihood values in Table 12, where each column corresponds to a model, and each row to a data set.

In summary, we observe that our model, estimated on the Swiss-Japanese data, performs very well in reproducing the Dutch data set in terms of directions

Cone	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	$6, 17, 28$	18639.80	22132	-15.78%
Left	$3-5, 14-16, 25-27$	14335.35	12939	10.79%
Right	$7-9, 18-20, 29-31$	14101.00	12379	13.91%
Extreme left	$1, 2, 12, 13, 23, 24$	230.38	14	1545.55%
Extreme right	$10, 11, 21, 22, 32, 33$	164.48	7	2249.65%

Table 10: Predicted (M_Γ) using the ASC model and observed (R_Γ) shares for alternatives grouped by directions with the Dutch data set.

Area	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	$1-11$	7275.62	5273	37.98%
constant speed	$12-22$	34378.45	42147	-18.43%
deceleration	$23-33$	5816.93	51	11305.75%

Table 11: Predicted (M_Γ) using the ASC model and observed (R_Γ) shares for alternatives grouped by speed regime with the Dutch data set.

and constant speed. The model does not perform well in forecasting accelerations and decelerations.

8 Conclusions

In this paper we propose a discrete choice model for pedestrian walking behavior. The short range walking behavior of individuals is modeled, identifying two main patterns: constrained and unconstrained. The constraints are generated by the interactions with other individuals. We describe interactions in terms of a leader-follower and a collision avoidance models. These models capture self-organizing effects which are characteristic of crowd behavior, such as lane formation. Inspiration for the mathematical form of these patterns is taken from driver behaviors in transportation science, and ideas such as the car following model and lane changing models have been reviewed and re-adapted to the more

Data set	Our model	ASC model based on Swiss-Japanese data	ASC model based on Dutch data
Swiss-Japanese	-22652.00	-25018.22	—
Dutch	-52676.78	-85565.72	-77871.06

Table 12: Loglikelihood of each model applied on the two data sets

complex pedestrian case. The difficulties to collect pedestrian data as well as the limited information conveyed by pure dynamic data sets limit the possibilities in the model specification step. Important individual effects cannot be captured without the support of socio-economic characteristics. Recent development of pedestrian laboratories, where the set up of controlled experimental conditions is possible, represents an important step in this direction. We use experimental data in a two step validation procedure. First, the model is validated on the same data set used for estimation in order to check for possible specification errors. Second, the model is run on a new data set collected at Delft University under controlled experimental conditions. The proposed validation procedure underline a good stability of the model and a good generalization performance. Few observations are badly predicted, mostly concentrated at the extreme of the choice set. The estimated coefficients are significant and their sign is consistent with our behavioral assumptions. Differently from other previous models, we can quantify the influence of the relative kinematic characteristics of leaders and colliders on the decision maker behavior. Moreover, such quantitative analysis has been performed using real world pedestrian data.

Future developments will focus in analyzing more and improving the acceleration and deceleration patterns. In particular, we plan to investigate the use of an adaptive resolution of the choice set, as well as incorporating in the model some physical characteristics of the pedestrians or of their ethnic group, such as average height and average speed.

Acknowledgments

We are very grateful to Kardi Teknomo, Serge Hoogendoorn and Winnie Daamen, who provided us with the data sets.

References

Ahmed, K. I. (1999). *Modeling drivers' acceleration and lane changing behaviors.*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Antonini, G., Bierlaire, M. and Weber, M. (2006). Discrete choice models of pedestrian walking behavior, *Transportation Research Part B: Methodological* **40**(8): 667–687.

Bierlaire, M. (2003). BIOGEME: a free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland. www.strc.ch.

Bierlaire, M., Antonini, G. and Weber, M. (2003). Behavioral dynamics for pedestrians, *in* K. Axhausen (ed.), *Moving through nets: the physical and social dimensions of travel*, Elsevier.

Blue, V. J. and Adler, J. L. (2001). Cellular automata microsimulation for modeling bi-directional pedestrian walkways, *Transportation Research Part B* **35**(3): 293–312.

Borgers, A. and Timmermans, H. (1986). A model of pedestrian route choice and demand for retail facilities within inner-city shopping areas, *Geographical analysis* **18**(2): 115–128.

Brady, A. T. and Walker, M. B. (1978). Interpersonal distance as a function of situationally induced anxiety, *British Journal of Social and Clinical Psychology* **17**: 127–133.

Daamen, W. (2004). *Modelling Passenger Flows in Public Transport Facilities*, PhD thesis, Delft University of Technology, The Netherlands.

Daamen, W. and Hoogendoorn, S. P. (2003). Experimental research of pedestrian walking behavior, *Transportation Research Record* **1828**: 20–30.

Dellaert, B. G., Arentze, T. A., Bierlaire, M., Borgers, A. W. and Timmermans, H. J. (1998). Investigating consumers' tendency to combine multiple shopping purposes and destinations, *Journal of Marketing Research* **35**(2): 177–188.

Dosey, M. A. and Meisels, M. (1969). Personal space and self-protection, *Journal of Personality and Social Psychology* **11**(2): 93–97.

Hartnett, J. J., Bailey, K. G. and Hartley, C. S. (1974). Body height, position, and sex as determinants of personal space, *Journal of Psychology* **87**: 129–136.

Helbing, D., Farkas, I., Molnar, P. and Vicsek, T. (2002). Simulation of pedestrian crowds in normal and evacuation simulations, *in* M. Schreckenberg and S. Sharma (eds), *Pedestrian and Evacuation Dynamics*, Springer, pp. 21–58.

Helbing, D. and Molnar, P. (1995). Social force model for pedestrian dynamics, *Physical review E* **51**(5): 4282–4286.

Herman, R. and Rothery, R. W. (1965). Car following and steady-state flow, *Proceedings on 2nd international symposium on the theory of traffic flow*, pp. 1–11.

Hoogendoorn, S. (2003). Pedestrian travel behavior modeling, *in* K. Axhausen (ed.), *Moving through nets: the physical and social dimensions of travel*, Elsevier.

Hoogendoorn, S., Bovy, P. and W.Daamen (2002). Microscopic pedestrian wayfinding and dynamics modelling, *in* M. Schreckenberg and S. Sharma (eds), *Pedestrian and Evacuation Dynamics*, Springer, pp. 123–155.

Horowitz, J. J., Duff, D. F. and Stratton, L. O. (1964). Body buffer zone: Exploration of personal space, *Archives of General Psychiatry* **11**(6): 651–656.

Lee, G. (1966). A generalization of linear car following theory, *Operations Research* **14**: 595–606.

McFadden, D. (1978). Modelling the choice of residential location, *in* A. Karlquist *et al.* (ed.), *Spatial interaction theory and residential location*, North-Holland, Amsterdam, pp. 75–96.

Newell, G. (1961). Nonlinear effects in the dynamics of car following, *Operations Research* **9**: 209–229.

Penn, A. and Turner, A. (2002). Space syntax based agent simulation, *in* M. Schreckenberg and S. Sharma (eds), *Pedestrian and Evacuation Dynamics*, Springer, pp. 99–114.

Phillips, J. R. (1979). An exploration of perception of body boundary, personal space, and body size in elderly persons, *Perceptual and Motor Skills* **48**: 299–308.

Sanders, J. L. (1976). Relationship of personal space to body image boundary definiteness, *Journal of Research in Personality* **10**: 478–481.

Schadschneider, A. (2002). Cellular automaton approach to pedestrian dynamics — Theory, *in* M. Schreckenberg and S. Sharma (eds), *Pedestrian and Evacuation Dynamics*, Springer, pp. 75–86.

Sommer, R. (1969). *Personal Space: The behavioral bases of design*, Prentice Hall, Englewood Cliffs, NJ.

Teknomo, K. (2002). *Microscopic Pedestrian Flow Characteristics: Development of an Image Processing Data Collection and Simulation Model*, PhD thesis, Tohoku University, Japan, Sendai.

Teknomo, K., Takeyama, Y. and Inamura, H. (2000). Review on microscopic pedestrian simulation model, *Proceedings Japan Society of Civil Engineering Conference*, Morioka, Japan.

Toledo, T. (2003). *Integrated Driving Behavior Modeling.*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Toledo, T., Koutsopoulos, H. N. and Ben-Akiva, M. E. (2003). Modeling integrated lane-changing behavior, *Transportation Research Record* **1857**: 30–38.

Turner, A. (2001). Angular analysis, *In Proceedings 3rd International Symposium on Space Syntax*, pp. 30.1–30.11.

Webb, J. D. and Weber, M. J. (2003). Influence of sensor abilities on the interpersonal distance of the elderly, *Environment and behavior* **35**(5): 695–711.

Whynes, D., Reedand, G. and Newbold, P. (1996). General practitioners' choice of referral destination: A probit analysis, *Managerial and Decision Economics* **17**(6): 587.