Probabilistic speed-density relationship for pedestrians based on data driven space and time representation

Marija Nikolić
Michel Bierlaire
Bilal Farooq

Transport and Mobility Laboratory, EPFL

April 2014

STRC

14th Swiss Transport Research Conference

Monte Verità / Ascona, May 14-16, 2014

Transport and Mobility Laboratory, EPFL

Probabilistic speed-density relationship for pedestrians based on data driven space and time representation

Marija Nikolić, Michel Bierlaire
Transport and Mobility Laboratory
School of Architecture, Civil and
Environmental Engineering
École Polytechnique Fédérale de Lausanne
phone: +41 21 693 24 08
fax: +41 21 693 80 60
{marija.nikolic,michel.bierlaire}@epfl.ch

Bilal Farooq
Géologique et des Mines
Département des génies civil
Polytechnique Montréal, Canada
phone: +1 514 340-4711 ext. 4802
fax: +1 514 340-3981
bilal.farooq@polymtl.ca

April 2014

Abstract

This paper proposes a mathematical framework that provides the detailed characterization of the pedestrian flow. It is specifically designed to address the heterogeneity of pedestrian population which is to be reflected through the pedestrian flow indicators. The key components of the presented work are: (i) data driven space discretization framework based on the Voronoi tessellations that allow pedestrian-oriented definition of density indicator; (ii) statistical and data driven approach to time aggregation, allowing for the pedestrian oriented definition of speed indicator; (iii) probabilistic model for speed-density relationship, so as to capture the empirically observed heterogeneity among pedestrians. The estimation and validation of the proposed model are performed on the basis of a pedestrian tracking input. Data is collected in a Lausanne railway station where a large-scale network of cameras has been installed to automatically locate and track thousands of pedestrians. Additionally, the performance provided by this methodology is compared with the well-accepted models published in the literature against empirical data with the aim at improving research on the pedestrian flow characterization.

Keywords

pedestrian flow indicators, fundamental diagram, time aggregation, space discretization, Voronoi diagram, individual trajectories

1 Introduction

Understanding the pedestrian flow is essential to many areas ranging from urban planning, traffic forecasting, designing public spaces (airports, stadiums, train stations, etc.), the optimization of the processes within public facilities to the evacuation studies. To answer the questions with respect to pedestrian flow, there are mainly two different approaches that exist in the literature: (i) empirical pedestrian data analysis and (ii) mathematical modeling approaches aiming at a realistic representation of the phenomena typical for pedestrian flow and movement behavior. Originally, data on pedestrian behavior and movement were based on surveys and counting whereas in recent years modern technology has allowed pedestrian tracking (GPS, Wi-Fi, Bluetooth, video-based tracking) that makes possible for comprehensive pedestrian studies to be conducted (Bauer *et al.,* 2009). Empirical data amassed through the means of pedestrian tracking describe the essential parts of pedestrian dynamics, and are therefore crucial to the process of model formulation, calibration and validation.

The investigation of pedestrian movement behavior based on empirical observations has revealed a large variety of interesting collective effects and self-organization phenomena, such as lane formation (Daamen and Hoogendoorn, 2003, Hoogendoorn and Daamen, 2005, Navin and Wheeler, 1969), group formation (Shi *et al.,* 2007), density waves similar to the stop-and-go phenomenon from vehicular traffic (Helbing and Molnar, 1995), short-lived roundabouts at intersections and oscillations at bottlenecks (Schadschneider *et al.,* 2009). The understanding of these emergent patterns is a prerequisite to predicting how the flow will behave under different circumstances. On the other hand, several studies have been performed so as to obtain the quantitative analyses of pedestrian flow characteristics based on the empirical data. In that manner, commonly used observables are flow (q), density (k) and speed (v). Many researchers have proposed the relationship between these characteristics known as fundamental diagram. The fundamental diagram plays an important role in the field of pedestrian flows. It is used in the planning and design of pedestrian facilities. Also, it is a required input or calibration criterion for models of pedestrian dynamics (Hughes, 2002, Blue and Adler, 2001, Helbing and Molnar, 1995). Nevertheless, large differences are found in the literature regarding the fundamental relationship. In particular, the reported maximal value of flow goes from 1.2 ped/ms to 1.8 ped/ms, jam-density goes from 3.8 ped/m^2 to 10 ped/m^2 and the reported maximum flow density range is from 1.7ped/m^2 to 7 ped/m^2 (Seyfried *et al.,* 2010). The researchers have suggested several explanations for these deviations some of which can be attributed to the cultural differences and the complex nature of pedestrian interactions (Helbing *et al.,* 2007), the differences between pedestrian facilities and the effects of the environment (Rastogi *et al.,* 2013), psychological factors, and measurement methods (Schadschneider *et al.,* 2009).

Most of the studies have focused only on unidirectional flows and the relationships between the indicators are usually developed on the grounds of drawing the parallels between pedestrian and vehicular traffic. Although the field of vehicular flow modeling is quite well established (Hoogendoorn and Bovy, 2001) this comparison is considered useful to some extent given large differences that exist between the two types of traffic flow. In comparison to roadways where vehicular flow is separated by directions, pedestrian flow is such that pedestrian facilities permit them to move in a multidirectional fashion. Vehicular traffic is regulated while pedestrians can choose their direction and speed depending on different factors (personal characteristics, trip purpose, etc.). The lack of strict rules for pedestrians to follow allows them to occupy any part of the walkway so they are able to change their speed and direction faster than vehicles. Nonetheless, the empirical studies that focus on bidirectional and multidirectional flow are very scarce. Navin and Wheeler (1969) stated that the directional imbalance causes the reduction of flow, in contrast to Fruin (1971) who has neglected the differences between the unidirectional and multidirectional flow. Zhang and Seyfried (2013) have analyzed the effect of bidirectional flow on the fundamental diagram. It was observed that for the higher density levels ($> 1 ped/m^2$) speed values are larger in the case of the unidirectional flow, compared to the bidirectional one. Kretz *et al.* (2006) has reported that counter flows are more efficient then unidirectional. This has resulted from the observation that in a counter flow the sum of fluxes is larger than the flux in any of the counter flow streams. Lam *et al.* (2003) have investigated the effects of the bidirectional flow from free-flow to congested regimes. The reduction in the walkway capacity and speed of minor streams were observed. Wong *et al.* (2010) has developed a model of bidirectional pedestrian streams, modeling the effect of intersecting angle on fundamental diagram. It was reported that the speed of a stream is more negatively affected by the intersection angle under greater total densities with the head-on conflict being the worst case.

The presented studies have provided better understanding of the effect that different types of pedestrian flows have on the fundamental relationship. However, the highly heterogeneous nature of pedestrian flows has been neglected and the empirical data have usually been used so as to fit certain deterministic curves. Therefore, the improved methodologies for the pedestrian flow characterization need to be developed, and it should be done on the basis of the detailed pedestrian flow data. In this paper, we propose a framework designed to address the question of heterogeneity among pedestrians which is to be reflected through pedestrian flow indicators. The study provides the specific methodological concepts and results have been obtained through the utilization of the great potential of the pedestrian tracking input.

2 Data

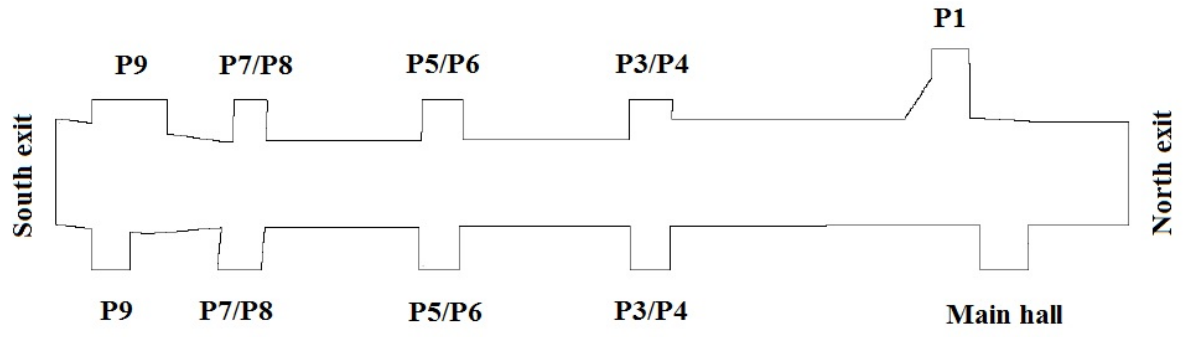
The data set employed in this study contains individual pedestrian trajectories collected in a pedestrian underpass West (PIW) of the Lausanne train station in Switzerland. Fig. 1 shows the layout of the studied area, which covers approximately 685 m^2. The underpass is frequently used especially during the morning and afternoon peak hours since it connects the exterior of the train station to the five main platforms. During the morning peak hour between 100 and 400 pedestrians can be observed every minute (Hänseler *et al.,* 2013). A large-scale network of thermal and range sensors has been employed to automatically locate pedestrians in PIW and to track them across time (Alahi *et al.,* 2011). The outcome of the complete vision processing is the set of the detailed trajectories of pedestrians:

$$p(t) = (x(t), y(t), t, pedestrian_{id}) \tag{1}$$

so that a given pedestrian is at position (x, y) at time t. The spatial and temporal resolution of less than a meter and a few hundred milliseconds, respectively, enables the detailed study of the different types of pedestrian flows in public spaces.

In this study it has been exploited in order to achieve the following goals: (i) the derivation of the definitions of the pedestrian traffic fundamental indicators, (ii) the modeling of the relationships between them and (iii) model estimation and validation. Pedestrian tracking data collected in the time period between 07:00 and 08:00 on February 12, 13, 14, 15 and 18, 2013 is used for estimation purposes (Section 4.2). Additional data set that is collected on February 06, 2013 is used for validation (Section 4.3).

Figure 1: Pedestrian underpass West - layout



3 Fundamental flow indicators

The existing approaches to measuring and specifying pedestrian flow indicators consider pedestrian population as being completely homogenous usually constraining their flow patterns to the unidirectional case. Here we propose the methodological framework that is specifically designed so as to cope with the first drawback - the heterogeneity of pedestrian population which is to be reflected through the pedestrian flow indicators. The disaggregated and data-driven discretization framework is specifically designed to provide the comprehensive pedestrian-oriented definitions of density and speed indicators in the first place.

3.1 Data-driven space discretization

The space discretization is an important aspect that affects the pedestrian flow analysis. The physical space, within which the pedestrians perform their movement and interactions with each other and with the surrounding environment, could be represented in different ways (continuous, grid, network). When measuring pedestrian flow characteristics, one usual approach is based on the cell-based space representation. The cell-based method transforms the space into regions where each cell is seen as being entirely homogenous. In this case, the discretization that results in units that are too small might leave many intervals empty. On the other hand, discretization units which are too large might lead to the loss of heterogeneity. This phenomenon is known as Modifiable Areal Unit Problem (Openshaw, 1984), meaning that each size and boundary change affects the proportion of the number of pedestrians related to a specific space unit. These effects can be alleviated by the decomposition of the space at an individual pedestrian level through the use of Voronoi tessellations (Okabe *et al.,* 2009). The Voronoi-based structures have been investigated in the pedestrian flow theory (Nakamura *et al.,* 2011, Steffen and Seyfried, 2010), but in this paper they are employed for the sake of enabling flow characterization through the pedestrian-oriented approach. We consider space-time representation, where $p_i = (x_i, y_i, t_i)$ refers to physical location (x_i, y_i) of a given pedestrian i at a specific time (t_i). Voronoi space decomposition assigns the personal region $(V(p_i))$ to each pedestrian (p_i), in such a way that all space locations (p), are associated with the closest pedestrian in respect of the Euclidean distance:

$$V_i = V(p_i) = \left\{ p \middle| \|p - p_i\| \leq \|p - p_j\|, i \neq j \right\} \tag{2}$$

The ordinary Voronoi diagram is defined in an unbounded plane. In our approach however we deal with a bounded region where pedestrians are located. Therefore, we consider bounded

Voronoi diagram:

$$V = \{V(p_1) \cap RoI, V(p_2) \cap RoI, \ldots, V(p_n) \cap RoI\} \tag{3}$$

where RoI refers to Region of Interest.

3.1.1 Merging cells

When it comes to Voronoi space decomposition, certain issues can arise due to the finite arithmetics of computers and noisy observations in data. These issues lead to the degeneration of Voronoi cells, reflected through unrealistically small pedestrian-allocated polygons. Therefore, it is necessary to define a threshold distance between neighboring pedestrians when more polygons, which are critical in terms of the area, merge together. This is necessary in order to guarantee numerical stability of results. For this purpose, Delaunay triangulation(Okabe *et al.,* 2009) is used so as to pinpoint a threshold distance (ξ) between neighboring pedestrians (p_i, p_j) which will trigger the clustering event:

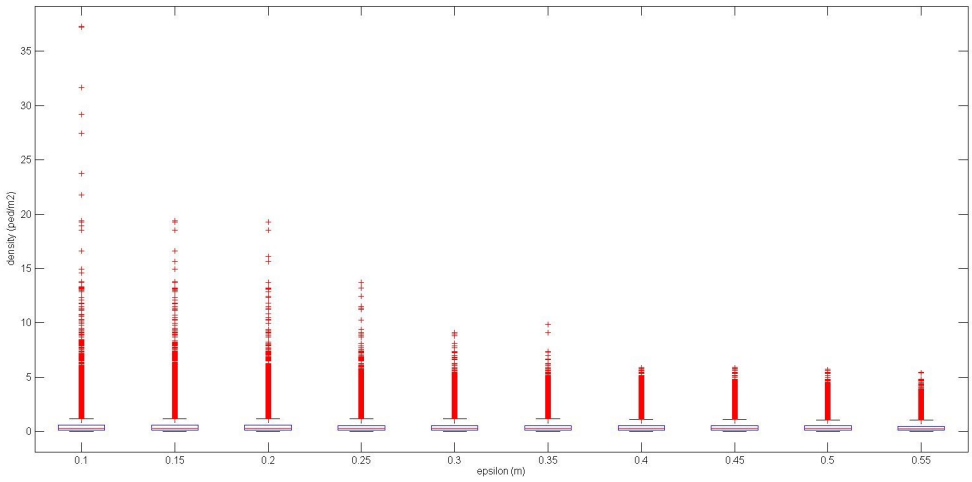
$$d\left(p_i, p_j\right) < \xi, \forall i, j. \tag{4}$$

In order to identify the appropriate threshold distance, the merging of cells is performed by utilizing the different values of distances. The corresponding density values are analyzed through the means of sensitivity analyses. The value ($\xi = 0.4m$) that guarantees the numerical stability of the results is selected according to Fig. 2 which shows that the range of density values stabilizes for $\xi >= 0.4m$. The merged cells are considered as one polygon associated with a weight (ω_i), representing the group of pedestrians occupying the corresponding space.

3.1.2 Dealing with obstacles

The underlying assumption for the proposed space tessellation is that it is possible to connect two generator points by a straight line. This is not the case where the plane is not obstacle-free. As for this particular case we combine the ideas of Voronoi diagram for the set of points and Voronoi diagram for the set of areas (Okabe *et al.,* 2009), in order to create personal regions which do not overlap with obstacles. In this case points refer to pedestrian positions in the space, whereas the areas determine the space occupied by obstacles. To obtain the bisector between a pedestrian (p_i) and an obstacle (O), the distance is defined as the shortest Euclidean distance

Figure 2: Treshold distance determination

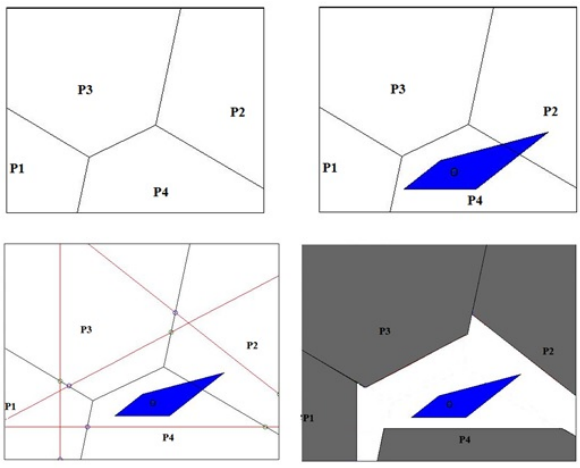


from p_i to O :

$$d\left(p_i, O\right) = min_{o_j}\left\{\|p_i - o_j\| \middle| o_j \in O\right\} \tag{5}$$

The outcome is such a spatial tessellation where each of the aforementioned has its own 'personal' polygon while pedestrians' polygons remain convex (Fig. 3). The advantage of having convex polygons lies in the simple calculation of their areas and natural looking personal regions.

Figure 3: Obstacles: (up-left) Obstacle-free space; (up-right) Pedestrian polygons that overlap with an obstacle; (down-left) Bisectors between obstacle and pedestrians; (down-right) Updated Voronoi diagram



3.1.3 Density

Given the space discretization specified above, the density of the cell V_i around pedestrian i positioned at $p_i = (x_i, y_i, t_i)$ is defined as the inverse of the area of the personal space:

$$k_i = \frac{\omega_i}{|V_i|}. \tag{6}$$

$|V_i|$ refers to the area of the space assigned to the group of w_i pedestrians, in the case of the merging of cells. The unit is the number of pedestrians per square meter. This way a good resolution in space is obtained (Fig. 4) by shedding light to a phenomenon not observable with the other methods (Fig. 5 shows the density maps obtained by utilizing the cell-based method for the same data set as considered in Fig. 4).

Figure 4: Voronoi based density map

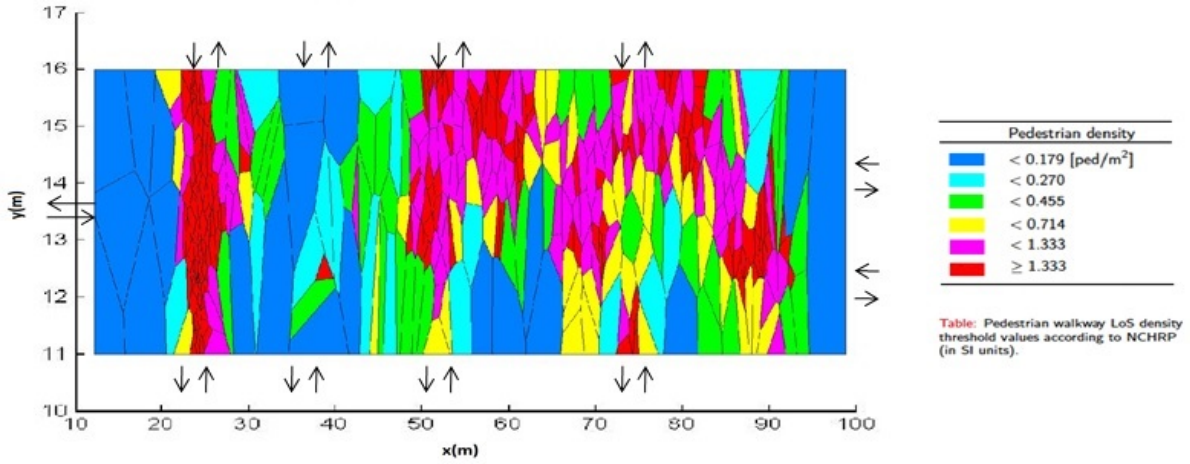
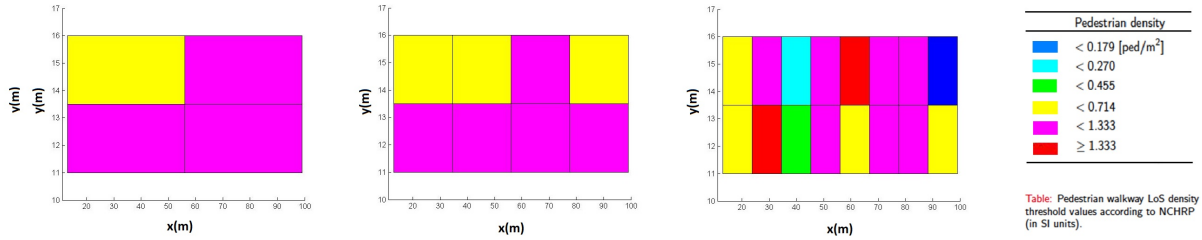


Figure 5: Cell based density maps



3.2 Data-driven time discretization

The choice of the time interval (Δt) for speed computation affects the statistical properties of the corresponding speed distributuion. In the literature, however, it is usually based on an arbitrarily chosen value. The data set employed in order to analyze this effect contains pedestrian trajectories collected on February 13, 2013, in the time period between 07:00 and 08:00 in PIW. The statistics obtained by utilizing a different time interval shows that the values of the mean, mode, median and quantiles are rather stable (Table 1). The difference in the maximum values of the speed obtained by using a different discretization, however, shows the existence of values that are highly unrealistic. They can be classified as outliers due to the fact that they occur infrequently. The source of these values lies in the noisy observations induced at the different levels of the pedestrian tracking pipe.

Table 1: The effect of Δt on speed values

Statistic	$\Delta t = 0.1s$	$\Delta t = 0.2$	$\Delta t = 0.3s$	$\Delta t = 0.4s$	$\Delta t = 0.5s$	$\Delta t = 0.6s$	$\Delta t = 0.7s$	$\Delta t = 0.8s$	$\Delta t = 0.9s$	$\Delta t = 1s$
min	0.0050	0.0025	0.0017	0.0013	0.0010	0.0008	0.0017	0.0036	0.0039	0.0052
mean	1.1155	1.1154	1.1153	1.1153	1.1152	1.1151	1.1150	1.1149	1.1149	1.1148
max	34.1162	17.7847	12.3805	9.7226	8.1288	7.1808	6.5927	6.2454	5.8898	5.6450
median	1.1173	1.1269	1.1296	1.1298	1.1293	1.1285	1.1273	1.1262	1.1253	1.1245
mode	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000	1.1000
$Q_{0.9}$	1.7022	1.6626	1.6478	1.6385	1.6321	1.6276	1.6245	1.6206	1.6175	1.6135
$Q_{0.95}$	2.0096	1.9470	1.9231	1.9087	1.8992	1.8910	1.8848	1.8785	1.8700	1.8615
$Q_{0.99}$	3.0419	2.9679	2.9361	2.8915	2.8301	2.7515	2.6874	2.6295	2.5827	2.5395

This poses the question of appropriate time discretization. The criterion to be used is the one that provides the aggregation which is: (i) not too large because it can lead to the loss of information; (ii) not too small since it is not resistant to the presence of noise. The choice of the time discretization in this paper is justified on the grounds of the statistics based approach. The speed distributions obtained through different time discretization are analyzed on the basis of the statistical test of moments. The outcome is the range of appropriate time discretization intervals that do not cause 'critical loss' of information (Table 2). The Kruskal-Wallis test (95% confidence level) on these set of moments revealed that they represent the same population (H=4.61, df=9, p=0.87).

Table 2: Raw moments of the speed distributions

Moment	$v_{\Delta t=0.1s}$	$v_{\Delta t=0.2}$	$v_{\Delta t=0.3s}$	$v_{\Delta t=0.4s}$	$v_{\Delta t=0.5s}$	$v_{\Delta t=0.6s}$	$v_{\Delta t=0.7s}$	$v_{\Delta t=0.8s}$	$v_{\Delta t=0.9s}$	$v_{\Delta t=1s}$
1	1.1161	1.1158	1.1156	1.1155	1.1153	1.1152	1.1150	1.1149	1.1148	1.1147
2	0.4175	0.3296	0.2956	0.2747	0.2591	0.2465	0.2358	0.2263	0.2179	0.2104
3	5.7853	2.5957	1.7703	1.4310	1.2544	1.1476	1.0740	1.0188	0.9744	0.9363
4	134.4926	31.2621	15.5319	10.9042	9.0167	8.0657	7.4917	7.0994	6.8045	6.5660

3.2.1 Speed

The speed at $p = (x, y, t)$ is defined as the speed of a pedestrian i (or the group of pedestrians in the case of the merging of cells) characterizing the Voronoi cell containing p:

$$v(p) = v_p \cdot d \tag{7}$$

where d refers to direction and v_p is given by

$$v_p = \frac{\|p_i\,(t + \Delta t) - p_i\,(t - \Delta t)\|}{2\Delta t}. \tag{8}$$

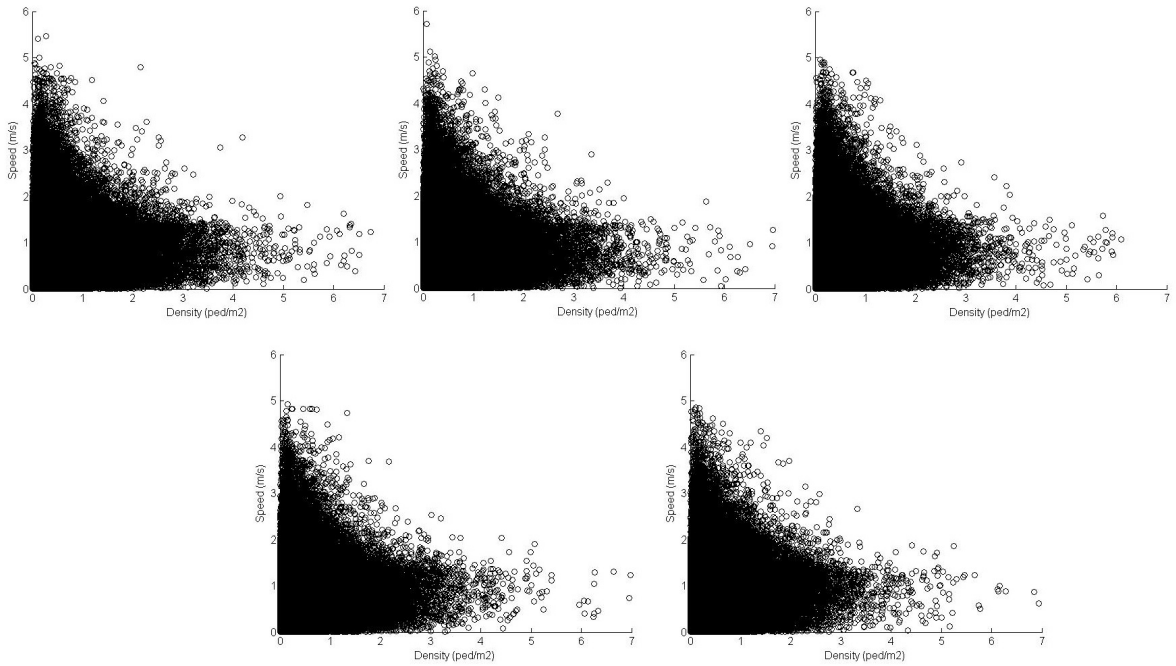
In this study the magnitude of the speed vector (v_p) is considered, where $\Delta t = 1s$ is chosen from the range of time interval values (Table 2) so that the information is preserved and the noise is removed to the greatest possible extent. The unit of v is meters per second.

4 Empirical speed-density relationship

The framework presented in the previous section has been designed in order to serve for defining and measuring speed and density indicators in a pedestrian-oriented manner. The empirical speed-density relationship is obtained through the use of pedestrian oriented definitions of the indicators given by Eq. (6) and Eq. (8). The employed data set consists of individual trajectories collected in PIW in the time period between 07:00 and 08:00 on February 12, 13, 14, 15 and 18, 2013.

Fig. 6 shows that empirical speed-density profiles exhibit a high scattering phenomenon. It can be observed that in the case of lower densities pedestrian movement behavior exhibit a higher level of randomness (based on speed values) compared to the situation where higher densities are observed. This can be explained by the fact that pedestrians tend to relish more freedom of movement when they are few in number whereas a crowded flow forces them to

Figure 6: Speed-density profiles



act within given circumstance. This can also explain the interesting phenomena related to the observation that speed values never reach zero for high densities.

According to our knowledge, in the pedestrian flow literature the speed-density relationship is always specified in a deterministic way, although the probabilistic one would be more appropriate so as to describe the effects of a large number of factors that involves randomness (personal characteristics, geometric settings, environmental conditions, etc.). Moreover, one week morning peak hour analyses established that the relationship on any weekday is essentially the same (the pedestrian behaviors are stable during the weekdays) suggesting that it can be modeled by using one single speed-density relationship. In order to explain the observed heterogeneity, theoretical distributions to be considered are those which are (i) bounded and (ii) have flexible shape. The first condition is motivated by the fact that the speed range is bounded by definition; the second one is needed to capture the observed randomness. The outcome of the analyses of a number of theoretical distributions is the Kumaraswamy specification (Kumaraswamy (1980)) due to the following reasons:

1. It is defined in the bounded region $[\ell,u]$, making it suitable for pedestrian speed modeling;
2. The shape of the distribution is controlled by two non-negative shape parameters α and β;
3. The simple closed form of both its probability density function ($f(v)$) and cumulative density function ($F(v)$), making it suitable for the use in simulation studies.

$$f(v) = \frac{\alpha \cdot \beta \cdot (v-\ell)^{\alpha-1} \cdot ((u-\ell)^{\alpha} - (v-\ell)^{\alpha})^{(\beta-1)}}{(u-\ell)^{\alpha \cdot \beta}} \tag{9}$$

$$F(v) = 1 - (1 - (\frac{v-\ell}{u-\ell})^{\alpha})^{\beta} \tag{10}$$

4.1 Model specification

The speed-density relationship is specified in probabilistic terms, such that the speed (V) represents a random variable which is Kumaraswamy-distributed:

$$V \sim f(\alpha(k), \beta(k), l(k), u(k)). \tag{11}$$

The parameters α and β affect the shape of the distribution, ℓ and u refer to lower boundary and upper boundary parameters respectively, while f is given by Eq. (9). Since the main goal is a unified model that would represent an empirical speed-density relationship, the distribution parameters (α, β, l and u) are to be specified as functions of density (they change with the density level).

Fig. 7 indicates relationships between the considered parameters and the explanatory variable density.According to the pattern they exhibit, the functional forms given in Table 3 have been considered. In both specifications parameter ℓ is fixed to zero since it represents the lower bound on speed values. It is based on the observation that for the whole density range lower speed values are either zero or very close to it. This way, the number of parameters to be estimated is also reduced while the rest of the distribution parameters is specified in such a way to capture the feature of the data using simple models. In the first specification, the exponential function is used in order to model the relationship between parameters β and u and the density level. The choice is motivated by the observed decrease in their values with an increase in density. The relationship between the parameter α and the density level is captured employing the polynomial function of the degree 3 because of its moderate flexibility of the shape and a simple form which is computationally easy to use. The second specification differs from the first one only in the specification of the parameter u in which case the polynomial function of the degree 3 is used.

Figure 7: Estimated values of parameters α, β, ℓ and u per density level

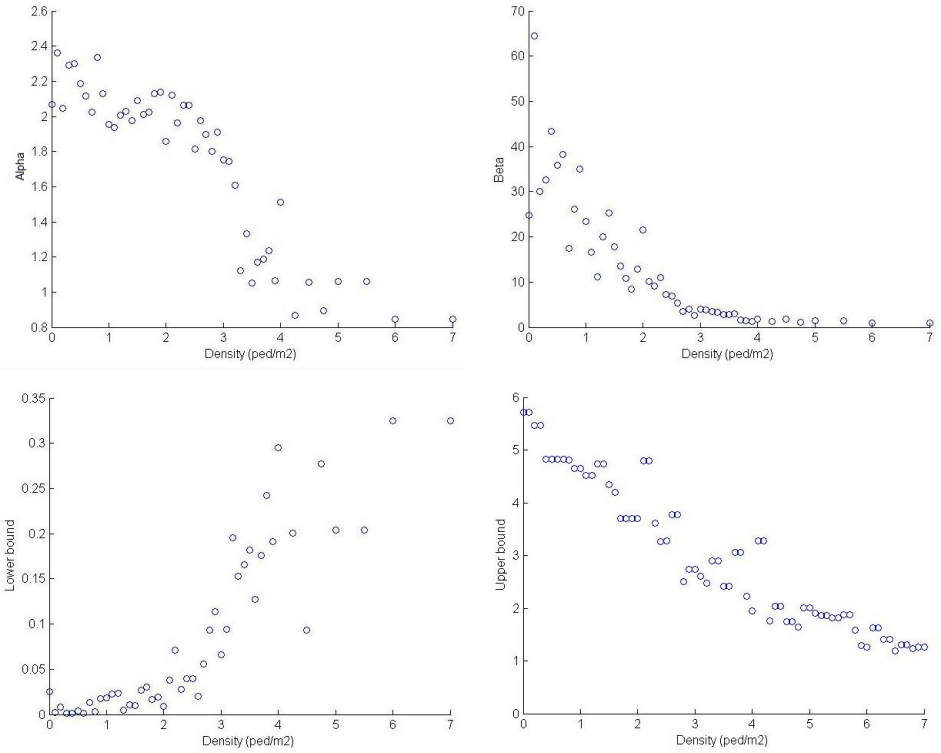


Table 3: Parameters specification

Parameter	Specification#1	Specification#2
$\alpha(k)$	$a_\alpha k^3 + b_\alpha k^2 + c_\alpha k + d_\alpha$	$a_\alpha k^3 + b_\alpha k^2 + c_\alpha k + d_\alpha$
$\beta(k)$	$a_\beta exp(b_\beta k)$	$a_\beta exp(b_\beta k)$
u(k)	$a_u exp(b_u k)$	$a_u k^3 + b_u k^2 + c_u k + d_u$
l(k)	0	0

4.2 Model estimation

For the estimation purposes, a maximum likelihood methodology (MLE) has been employed with the following log-likelihood function:

$$\begin{aligned} log\mathcal{L} = &\sum_{i=1}^{n} log(\alpha(k_i)) + \sum_{i=1}^{n} log(\beta(k_i)) + \sum_{i=1}^{n} (\alpha(k_i) - 1) log(v_i - l(k_i)) \\ &+ \sum_{i=1}^{n} (\beta(k_i) - 1) log((u(k_i) - l(k_i))^{\alpha(k_i)} - (v_i - l(k_i))^{\alpha(k_i)}) \\ &- \sum_{i=1}^{n} \alpha(k_i)\beta(k_i) log(u(k_i) - l(k_i)) \end{aligned} \tag{12}$$

In the Eq. (12) n is the sample size, k_i refers to a density and v_i is the corresponding speed value. The MLEs of the parameters are in this way obtained from the maximization of Eq. (12)) , on the basis of the iterative optimization method. The estimation results are given in Table 4. The resulting parameters′ specification is selected based on the log-likelihood value which gives a higher probability that the observations come from the Kumaraswamy distribution when the parameters are given by the first specification.

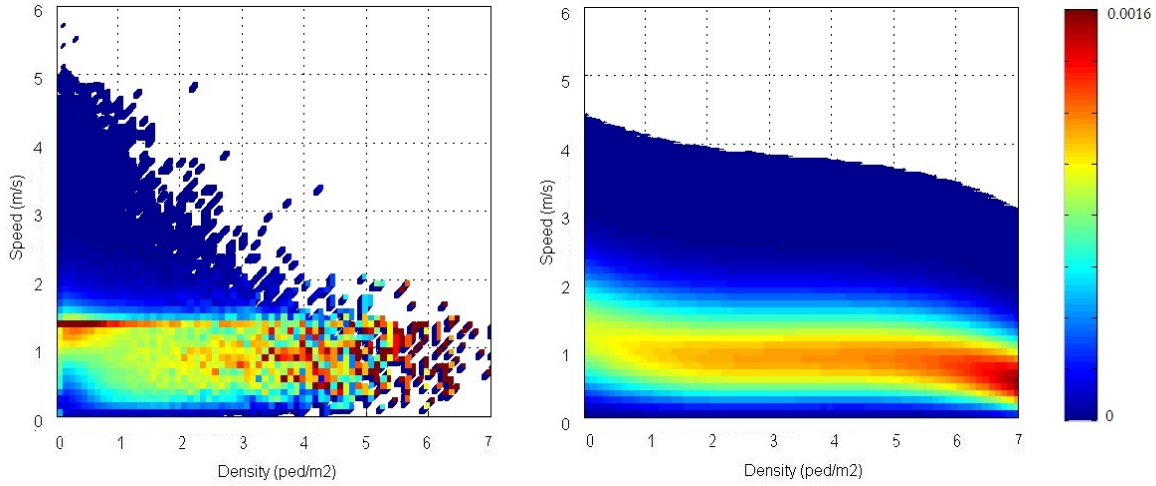
Table 4: Parameters estimates

Parameter	Specification#1	Specification#2
a_α	-0.007	0.049
b_α	0.096	-0.282
c_α	-0.378	-0.020
d_α	2.218	2.008
a_β	44.819	45.362
b_β	-0.105	-0.594
a_u	7	0.000
b_u	0	-0.000
c_u		-0.001
d_u		8.001
$log\mathcal{L}$	-891880	-932990

4.3 Model validation

The validation of the model is performed by the application on an additional data set (Fig. 8), that is not used in the model estimation process. It is performed with respect to the discrete joint

Figure 8: Validation: left - joint distribution of empirical observations; right - joint distribution of model predictions



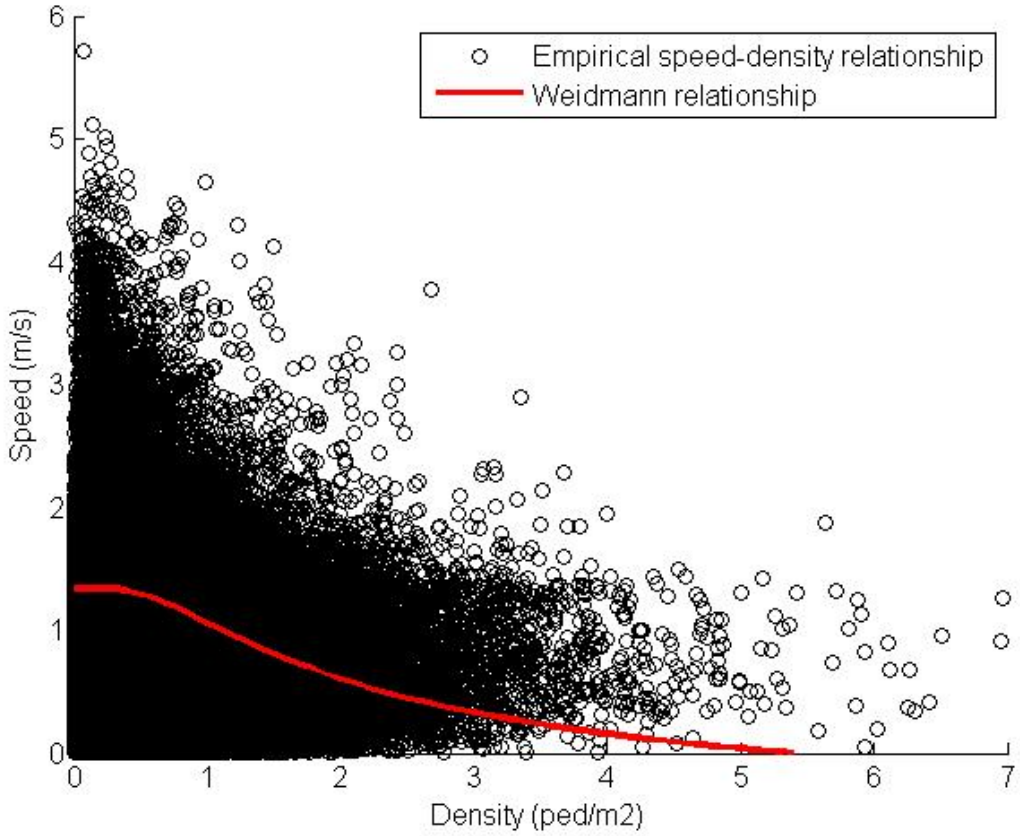
distribution of empirical observations and the moments derived. The moments are compared with the corresponding ones of model predictions using formal statistical test (Kruskal and Wallis (1952)). The raw moments are shown in Table 5. The Kruskal-Wallis test (H=0.33, df=1, p=0.5637) on these set of moments revealed that they represent the same population at 95% confidence level. The analysis results indicate a good performance of the model. Although the main pattern observed from data is reproduced, few observations are poorly predicted. In particular, the model does not predict speed values greater than approximately 4.5m/s. Given that these values rarely occur in the considered data set, it can be concluded that this observation is not critical with respect to the model performance.

Table 5: Raw moments - estimation data set

Moments	Data	Model prediction
1	0.9333	0.9856
2	0.1845	0.2376
3	0.0426	0.0648
4	0.1521	0.1769

Additionally, the performance of the proposed model is compared with the well-accepted model (Weidmann 1993, as quoted in (Zhang, 2012)) against the empirical data (Fig. 9).

Figure 9: Comparison to Weidmann model



Weidmann described the fundamental relationship between the speed and density using the Kladek formula:

$$v(k) = v_f \left[1 - exp\left(-\gamma\left(\frac{1}{k} - \frac{1}{k_j}\right)\right)\right] \tag{13}$$

where v_f refers to the free-flow speed ($1.34m/s$), k_j is the jam density ($5.4m^{-2}$) and γ is the shape parameter ($1.913m^{-2}$). The model is estimated by collecting 25 data sets. The maximum value of speed observed in our data set reaches $5.719m/s$ while the Weidmann model predicts the speed $1.34m/s$ in the free-flow regime that decreases exponentially with the density level. In the Weidmann model speed reaches zero at the jam density of $5.4m^{-2}$. However, the empirical speed-density relationship shows that people tend to walk at a speed that is as close as possible to their desired one even if the traffic conditions deteriorate. In this case, the distribution of empirical speed values is less spread, which is expected, given the fact that pedestrian movement is more constrained in times of increased congestion. Finally, let us note that the proposed

probabilistic model is able to explain the observed heterogeneity therefore addressing one of the main drawbacks of the existing deterministic approaches.

5 Conclusion and future research

The methodology for the pedestrian-oriented flow characterization has been proposed. The speed and density indicators are formulated on the grounds of a data-driven disaggregated space and time representation. The probabilistic methodology has been employed so as to describe the observed heterogeneity among pedestrians with respect to the speed-density relationship. The real case study has been engaged for the purposes of model estimation and validation. The validation results prove the strength of the proposed methodology when it is performed against the data. The comparison with the Weidmann model against the empirical data shows the ability of the probabilistic approach to explain more than the traditional deterministic ones currently do. The presented work also has the implications on dynamic continuum and discrete models for pedestrians that combine a conservation principle with a fundamental diagram. Moreover, the evaluation and optimization of the level of service of pedestrian facilities can largely benefit from the findings presented in this paper.

The described framework as such is however insufficient to explain the multidirectional nature of pedestrian flows. As further steps we will explore the possibility of addressing this issue through a stream-based concept. The final objectives will be the integration of the stream-based concept with the developed probabilistic framework in order to comply with the observed pedestrian heterogeneity.

6 References

Alahi, A., L. Jacques, Y. Boursier and P. Vandergheynst (2011) Sparsity driven people localization with a heterogeneous network of cameras, *Journal of Mathematical Imaging and Vision*, **41** (1-2) 39–58.

Bauer, D., N. Brandle, S. Seer, M. Ray and K. Kitazawa (2009) Measurement of pedestrian movements: A comparative study on various existing systems, *Pedestrian Behavior: Models, Data Collection and Applications: Models, Data Collection and Applications*, 325.

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

Daamen, W. and S. P. Hoogendoorn (2003) Experimental research of pedestrian walking behavior, *Transportation Research Record: Journal of the Transportation Research Board*, **1828** (1) 20–30.

Fruin, J. J. (1971) Pedestrian planning and design, *Technical Report*.

Hänseler, F., M. Bierlaire, B. Farooq and T. Mühlematter (2013) An aggregate model for transient and multi-directional pedestrian flows in public walking areas, *Technical Report*, **TRANSP-OR 131219**, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

Helbing, D., A. Johansson and H. Z. Al-Abideen (2007) Crowd turbulence: the physics of crowd disasters, *arXiv preprint arXiv:0708.3339*.

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

Hoogendoorn, S. and W. Daamen (2005) Self-organization in pedestrian flow, in *Traffic and Granular Flow'03*, 373–382, Springer.

Hoogendoorn, S. P. and P. H. Bovy (2001) State-of-the-art of vehicular traffic flow modelling, *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, **215** (4) 283–303.

Hughes, R. L. (2002) A continuum theory for the flow of pedestrians, *Transportation Research Part B: Methodological*, **36** (6) 507–535.

Kretz, T., A. Grünebohm, M. Kaufman, F. Mazur and M. Schreckenberg (2006) Experimental study of pedestrian counterflow in a corridor, *Journal of Statistical Mechanics: Theory and Experiment*, **2006** (10) P10001.

Kruskal, W. H. and W. A. Wallis (1952) Use of ranks in one-criterion variance analysis, *Journal of the American statistical Association*, **47** (260) 583–621.

Kumaraswamy, P. (1980) A generalized probability density function for double-bounded random processes, *Journal of Hydrology*, **46** (1) 79–88.

Lam, W. H., J. Lee, K. Chan and P. Goh (2003) A generalised function for modeling bi-directional flow effects on indoor walkways in hong kong, *Transportation Research Part A: Policy and Practice*, **37** (9) 789–810.

Nakamura, A., M. Ishii and H. Hiyoshi (2011) Uni-directional pedestrian movement model based on voronoi diagrams, paper presented at the *Voronoi Diagrams in Science and Engineering (ISVD), 2011 Eighth International Symposium on*, 123–126.

Navin, F. and R. Wheeler (1969) Pedestrian flow characteristics, *Traffic Engineering, Inst Traffic Engr*, **39**.

Okabe, A., B. Boots, K. Sugihara and S. N. Chiu (2009) *Spatial tessellations: concepts and applications of Voronoi diagrams*, vol. 501, John Wiley & Sons.

Openshaw, S. (1984) Concepts and techniques in modern geography number 38: the modifiable areal unit problem, *Norwick: Geo Books*.

Rastogi, R., S. Chandra *et al.* (2013) Pedestrian flow characteristics for different pedestrian facilities and situations.

Schadschneider, A., W. Klingsch, H. Klüpfel, T. Kretz, C. Rogsch and A. Seyfried (2009) Evacuation dynamics: Empirical results, modeling and applications, in *Encyclopedia of complexity and systems science*, 3142–3176, Springer.

Seyfried, A., M. Boltes, J. Kähler, W. Klingsch, A. Portz, T. Rupprecht, A. Schadschneider, B. Steffen and A. Winkens (2010) Enhanced empirical data for the fundamental diagram and the flow through bottlenecks, in *Pedestrian and Evacuation Dynamics 2008*, 145–156, Springer.

Shi, J., Y. Chen, F. Ren *et al.* (2007) Research on pedestrian crowd characteristics and behaviours in peak-time on chinese campus, in *Pedestrian and Evacuation Dynamics 2005*, 79–90, Springer.

Steffen, B. and A. Seyfried (2010) Methods for measuring pedestrian density, flow, speed and direction with minimal scatter, *Physica A: Statistical mechanics and its applications*, **389** (9) 1902–1910.

Wong, S., W. Leung, S. Chan, W. H. Lam, N. H. Yung, C. Liu and P. Zhang (2010) Bidirectional pedestrian stream model with oblique intersecting angle, *Journal of transportation Engineering*, **136** (3) 234–242.

Zhang, J. and A. Seyfried (2013) Empirical characteristics of different types of pedestrian streams, *Procedia Engineering*, **62**, 655–662.