ACCURACY OF PEDESTRIAN AND TRAFFIC FLOW MODELS: MEANINGFUL QUANTIFICATIONS

Femke van Wageningen-Kessels (corresponding author)
Delft University of Technology
Stevinweg 1, 2628CN Delft, The Netherlands
+31 6 2848 5258 / +968 9325 1057
f.l.m.vanwageningen-kessels@tudelft.nl

Winnie Daamen
Delft University of Technology
Stevinweg 1, 2628CN Delft, The Netherlands
+31 15 27 85927
w.daamen@tudelft.nl

Serge Hoogendoorn
Delft University of Technology
Stevinweg 1, 2628CN Delft, The Netherlands
+31 15 27 85475
s.p.hoogendoorn@tudelft.nl

July 25, 2014

ABSTRACT
Accuracy measures are used to quantify the accuracy of road traffic and pedestrian flow observations, models, simulations and predictions. We propose two new ways to measure the accuracy, by introducing and quantifying the phase error and the diffusion error. These errors take into account specific features that are of interest in traffic and pedestrian flow models. The phase error penalises a wrong estimate of the location of a high density region, the diffusion error penalises too smooth transitions between high density and low density regions. Combining the results of both accuracy measures, gives better insight than traditional accuracy measures into how the methodologies or models may be improved. Test cases show the application of the accuracy measures for the selection of a simulation method and for parameter estimation. Moreover, they show how the measures can be applied both for (one dimensional) road traffic flow and for (two dimensional) pedestrian flow.

INTRODUCTION

Practical applications of traffic and pedestrian flow theory include many steps in which accuracy is important: the flow is observed, a model and simulation tool are built and their parameters are estimated. The model can then be used for predictions, either or not based on real time data. In all these steps errors are made: there are discrepancies between reality and observation, between observations and models, between models and simulations based on them. One step further, Huber et al. (*1*) introduce methods to assess the accuracy of traffic information.

A lot of research effort has been put into assessing the accuracy of the steps. Qualitative assessments are used to show whether certain characteristics or phenomena are well represented (*2, 3, 4, 5*). Yet, more effort has been put into quantification of accuracy and their application in model choice, parameter calibration, validation and sensitivity analysis (*6, 7, 8, 9, 10, 11, 12, 13*). However, the way accuracy is quantified is often straightforward and does not provide information on the type of error that was made. For example, variables are compared on a one-to-one basis: the predicted value of a selected variable (e.g. speed, headway, density) at certain locations and times is compared to the reference value at the same locations and times. For each location and time the difference is computed and the difference is averaged over the whole space and time domain. The averaging can be done in many ways (*12*), including the mean error (ME), mean absolute error (MAE) and root mean square error (RMSE). Therefore, the accuracy measure does not give insight into what causes the error and how improvements in the applied methodologies can help to increase the accuracy.

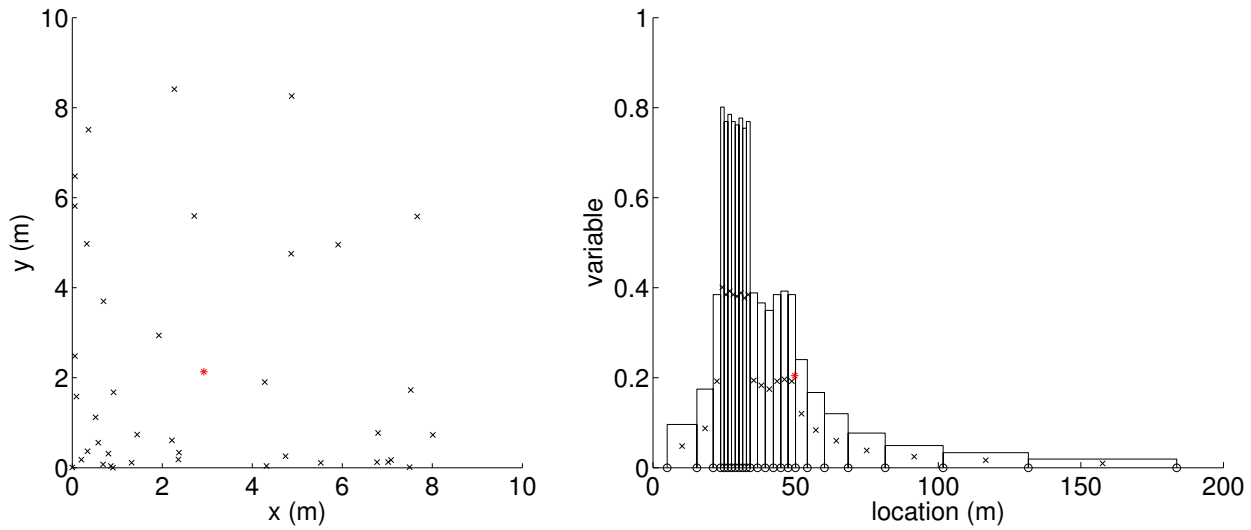
Our main contribution is the introduction of two new accuracy measures that do give insight into the type of error and take into account the specific features of traffic and pedestrian flow, such as chaotic behaviour and the importance of sharp transitions and the location and time when congestion occurs. The new accuracy measures apply the concepts of phase error and diffusion error and are introduced in the next section. They can be applied to both (one dimensional) road traffic flow and to (two dimensional) pedestrian flow, as we show in the Section 'Case Studies'. We conclude by discussing the main results and future research directions.

ACCURACY MEASURES

Traffic and pedestrian flows show large and sometimes sudden changes in the values of important variables. Congestion occurs locally and there may be congestion with low speeds just upstream of an on ramp, while 500 meter downstream traffic is free flowing and speeds are high. Additionally, stop and go waves are characterised by sudden and steep changes in density and speed. Pedestrian flows also show complex phenomena such as lane formation and clogging upstream of an exit. We introduce accuracy measures that can handle these key phenomena and include high penalties if they are represented badly. We focus on three important features:

1. key phenomena with steep state changes over time and/or space such as a sudden drops in speed (e.g. at the tail of congestion) or sudden speed increases (e.g. just upstream of a bottleneck) (*14, 15*),
2. key phenomena with states alternating over space and possibly time such as stop and go waves in road traffic or lanes in pedestrian flows (*16, 17, 18*),
3. the speed or velocity of high density areas, causing spill back and possible blockage of ramps, exits or other parts of networks (*2, 19, 20*).

We select those features because they characterise traffic and pedestrian flows and the features distinguish them from many other types of flows such as most fluid flows.



(a) Centre of mass for phase error. Crosses indicate the positions of particles (e.g. pedestrians). The red star indicates the centre of mass of all pedestrians.

(b) Centre of mass for diffusion error. Circles on the x-axis indicate the position of vehicles. The height of the bars indicate the approximated density. The crosses are the centres of mass of the individual bars. The red star indicates the centre of mass of all bars.

FIGURE 1 Examples of centre of mass used to define the phase error and diffusion error.

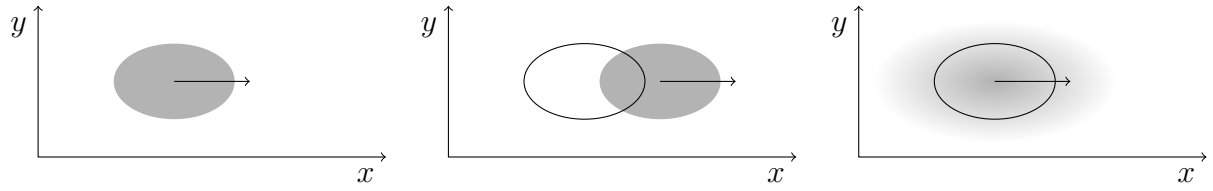
To measure the accuracy of a data set, a model, a simulation or a prediction, its variables (e.g. speed, velocities, densities or headways) are compared to those of a reference or 'ground truth'. We propose two new accuracy measures for traffic and pedestrian flows. They are based on the concepts of phase error and diffusion error, as introduced in the next section. Unlike the traditional measures, the new measures take into account the specific features of traffic and pedestrian flows as described above. They do this by measuring whether the location of high density areas is reproduced correctly (phase error) and by measuring how well sharp transitions are reproduced (diffusion error).

Location of Centre of Mass and Phase Error

Both the phase error and the diffusion error use the concept of centre of mass to quantify the accuracy. In physics, the centre of mass is defined as the point at which the entire weight of a body may be considered as concentrated so that if supported at this point the body would remain in equilibrium in any position.We translate the location of the centre of mass $\vec{X}$ as the point where all particles (vehicles or pedestrians) may be considered as concentrated. The concept is illustrated in Figure 1(a). We define the location of the centre of mass at time t. The location $\vec{x}_n$ of the n-th particle is supposed to be known for all particles $n \in N$ present at time t. The location of the centre of mass $\vec{X}_{\text{particle}}$ is defined as follows:

$$\vec{X}_{\text{particle}} = \frac{1}{N} \sum_{n=1}^{N} \vec{x}_n \tag{1}$$

with $\vec{X}_{\text{particle}}$ and $\vec{x}_n$ scalar in one dimensional road traffic flow and column vectors with two rows in two dimensional pedestrian flow.



(a) Reference solution (b) Solution with phase error (c) Solution with diffusion error

FIGURE 2 Three possible states of a crowd. The reference solution is considered as the ground truth, where all pedestrians are gathered within an ellipse shaped region (a). If a model predicts the location of the ellipse incorrectly, there is a phase error (b). If a model predicts that the pedestrians are more spread out over the region and the transition at the edge of the ellipse is too smooth, there is a diffusion error (c).

In a more generic case, location of the centre of mass within the domain Ω can be defined as:

$$\vec{X} = \frac{\int_{\vec{x}\in\Omega} \rho(\vec{x})\vec{x}\ \mathrm{d}\vec{x}}{\int_{\vec{x}\in\Omega} \rho(\vec{x})\ \mathrm{d}\vec{x}} \tag{2}$$

In a continuum model $\rho(\vec{x})$ would be the density at location $\vec{x}$, but it can also be interpreted as the probability that a particle is present at location $\vec{x}$. This definition is useful if the density at every point in the domain is known, or can be derived. If the domain can be subdivided into $i \in [1 : I]$ cells (road segments with a certain length or regions with a certain area), and the cells have constant density ρ_i, then (2) reduces to:

$$\vec{X}_{\text{continuum}} = \frac{\sum_{i=1}^{I} A_i \rho_i \vec{x}_i}{\sum_{i=1}^{I} A_i \rho_i} \tag{3}$$

with A_i the length (in one dimensional flow) or area (in two dimensional flow) of the i-th cell. Therefore, $A_i\rho_i$ is the (expected) number of particles within cell i. $\vec{x}_i$ denotes the location of the centre of mass of the cell, which can be computed easily if the cell has a simple shape and ρ_i is constant within the cell. The definition in (3) is useful if data from loop detector is available or if continuum flow models are applied.

Phase Error

A phase error occurs if the average speed of the particles is too low or too high or if they move in the wrong direction, see Figure 2(b). This results in an inaccuracy in the location of patterns such as a high density/low speed region in a congestion pattern or in lanes formed in bi-directional pedestrian flow. The difference in the location of the centre of mass between the reference solution and the test solution is therefore a measure for the phase error:

$$\vec{E}_{\text{phase}} = \vec{X}_{\text{test}} - \vec{X}_{\text{ref}} \tag{4}$$

We note that in some cases a phase error is something to avoid, for example when one wants to predict if or when the tail of a queue upstream of a traffic light will reach the next upstream intersection. In other cases, such as with lanes formed in bi-directional pedestrian flow, a phase error

may only indicate that the high and low density/speed regions have 'swapped': e.g. pedestrians walk from left to right where they were supposed to walk from left to right and vice versa. This type of phase error does not pose a problem in most applications. The difference between two causes for phase errors implies that they always need an interpretation to understand whether the phase error should be avoided or it can be ignored.

Height of Centre of Mass and Diffusion Error

To quantify the diffusion error we use the concept of centre of mass in a slightly different way and use it to define the centre of mass in the direction of a fundamental variable such as density, speed or x- or y-velocity. The concept is illustrated in Figure 1(b). We define the centre of mass in the direction of the fundamental variable, as the centre of mass of the area under its profile. In the figure, the position of $N+1$ vehicles is depicted, with x_n the location of the n-th particle. In this case, each bar indicates one vehicle. The centre of mass in the direction of variable f is the height of the combined centre of mass of all N bars:

$$F = \frac{1}{N} \sum_{n=1}^{N} \frac{f_n}{2} \tag{5}$$

We note that (5) also holds in two dimensional cases. Generalization of (5) without assuming each cell contains exactly one particle, but instead taking a continuum approach as in (3), leads to:

$$F_{\text{continuum}} = \frac{\sum_{i=1}^{I} A_i \rho_i f_i}{2 \sum_{i=1}^{I} A_i \rho_i} \tag{6}$$

with A_i the area of the i-th cell and ρ_i its density, which is assumed to be constant. Consequently, $A_i \rho_i$ is the number of particles in the i-th cell. Finally, further generalisation gives, similar to (2):

$$F = \frac{\int_{\vec{x} \in \Omega} \rho(\vec{x}) f(\vec{x}) \, \mathrm{d}\vec{x}}{2 \int_{\vec{x} \in \Omega} \rho(\vec{x}) \, \mathrm{d}\vec{x}} \tag{7}$$

Diffusion error

A diffusion error occurs if transitions between areas with a high value of the fundamental variable and with a low value become too smooth, see Figure 2(c). If the fundamental variable is density, high density regions get a lower density and vice versa. Therefore, if transitions are sharp, the centre of mass in density direction is higher than when transitions are smoother. The difference in the centre of mass in the direction of the fundamental variable between the reference solution and the test solution is a measure for the diffusion error:

$$E_{\text{diff}} = F_{\text{test}} - F_{\text{ref}} \tag{8}$$

CASE STUDIES

We discuss the concepts of phase and diffusion error further using some case studies. The case studies compare simulation results with analytical model solution (for road traffic, case 1), and with experimental data (for pedestrian flows, case 2). Errors are quantified using our newly developed accuracy measures and traditional ones. The results show that the new accuracy measures give insight into how accurate simulation results are and help to identify the cause of inaccuracies.

Case 1: Queue Upstream of Traffic Light

This case illustrates the application of the new accuracy measures to compare simulation results with the analytical model solution. We consider a hypothetical one lane road with a queue (density is jam density) of vehicles in front of a red light. Downstream of the queue the density is initially zero, the length of the queue is 2 kilometre and upstream of the queue the initial density is half of critical density. At time $t = 0$, the red light turns green and vehicles start driving. To determine the traffic state until the queue has solved, we use the LWR model (*19, 21*) with a parabolic-linear fundamental diagram (*22*) with maximum speed 33.3 m/s, critical speed 20.8 m/s, critical density 0.33 vehicles/m and jam density 0.2 vehicles/m.

Analytical and numerical solution

We solve this problem both analytically and by simulation. We use three different simulation methods: a minimum supply demand method with explicit time stepping (*23, 24*), an upwind method with explicit time stepping (*25*) and an upwind method with implicit time stepping (*26*). Furthermore, we vary the time step size to study the influence of the Courant-Friedrichs-Lewy (CFL)-number (*27*). With a small time step size, the CFL-number is small and a simulation takes long. However, with a too large time step size, the CFL-number is larger than 1 and the minimum supply demand method and the upwind method get unstable, resulting in unrealistic simulation results. The simulation results and the exact solution are shown in Figure 3 and 4.

Accuracy

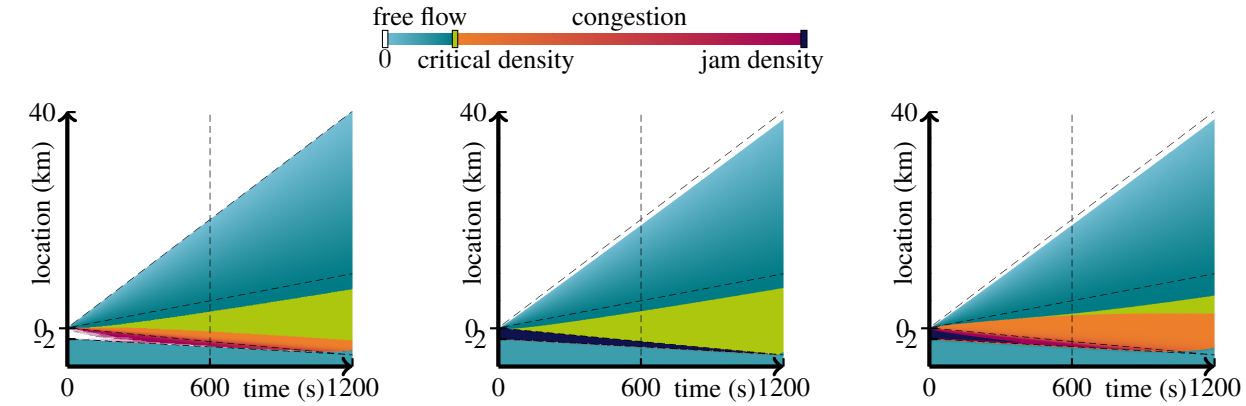
Analysis of the numerical methods show that all methods are first order accurate, implying that the global error is proportional to the numerical resolution (*28*). However this does not mean that, in practice, the accuracy of the methods is similar. To measure the accuracy of the simulation results, they are compared to the exact solution. We do this only at time $t = 600$ s, which is representative for the whole time period. For all solutions we compute the centre of mass of the cross sections shown inFigure 4. The phase error and diffusion error are computed as in (4) and (8), respectively. For the diffusion error we use the density as variable: $f = \rho$. The errors are shown in Figure 5.

Results

The results in Figure 5 show that the minimum supply demand method always performs worse than the upwind explicit method. The performance of the upwind implicit method depends on the time step size: it gives relatively high accuracy if the CFL-number is small, while accuracy is low (errors are large) if it is large. For CFL-numbers larger than 1, the minimum supply demand method and the upwind explicit method become unstable and give very unrealistic results and they are therefore omitted from the figure. The accuracy of the upwind implicit method decreases with increasing CFL-number, but the accuracy is of the same order as that of the minimum supply demand method. Furthermore, the results show that, for the minimum supply demand and the upwind implicit method, the phase error depends more strongly on the CFL-number than the diffusion error. For the upwind explicit method the difference is not clearly visible from the graphs.

Discussion

These results give insight into when to use which numerical method. A large error, has to be weighted against a large computation time (i.e. small time steps). Furthermore, if the phase error is important in the application (for example if one wants to know whether congestion will spill

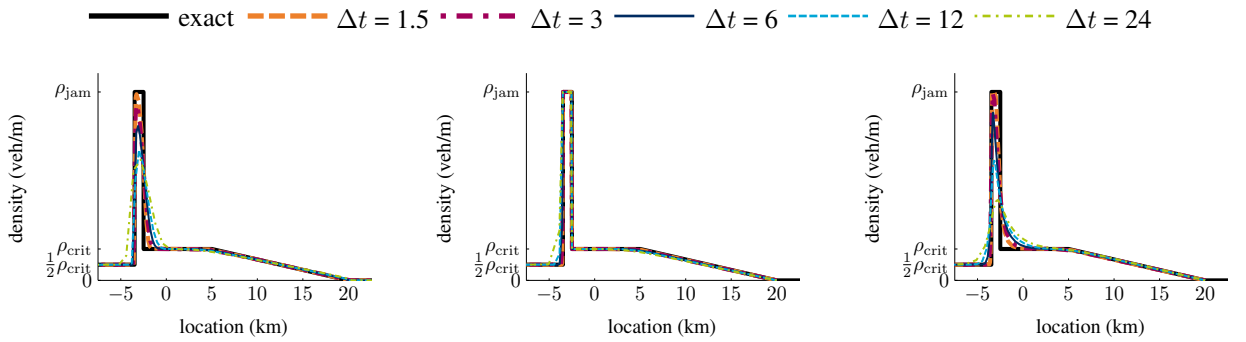


(a) Minimum supply demand method.

(b) Upwind method with explicit time stepping.

(c) Upwind method with implicit time stepping.

FIGURE 3 Densities computed with different numerical methods and fixed time step size. The vertical dashed line indicate the time for which cross sections are shown in Figure 4. The other dashed lines indicate sharp boundaries in the analytical solution.



(a) Minimum supply demand method.

(b) Upwind method with explicit time stepping.

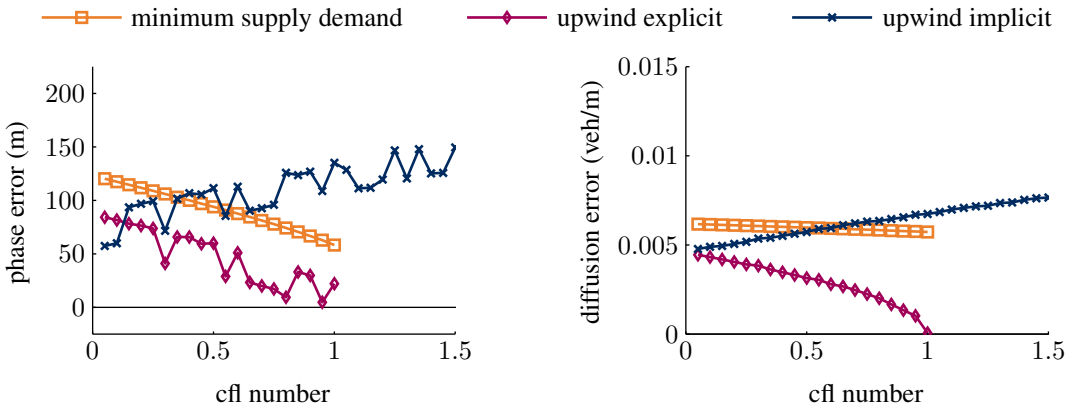
(c) Upwind method with implicit time stepping.

FIGURE 4 Density profiles at time $t = 600$ s with different numerical methods and time step sizes.

back until a certain off ramp), than also the CFL-number should be taken into account. If a small diffusion error is more important, than it is less important to chose a good CFL-number, as long as it is not larger than 1 with the minimum supply demand method or the upwind explicit method. We note that no final conclusion on an appropriate numerical method can be drawn from this very limited test case. Instead a wider range of tests such as in (*26*) should be performed. However, the test case shows how the new accuracy measures can be applied to support the choice of simulation method.

Case 2: Bidirectional Pedestrian Flow

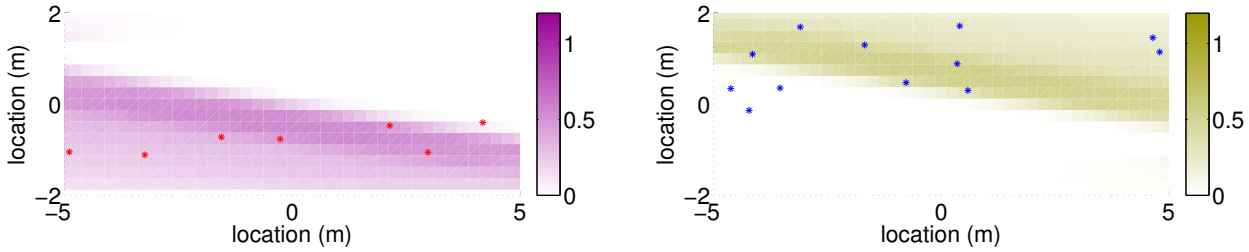
In this case we illustrate how to compare data from an experiment with bidirectional pedestrian flows (*29*) with simulation results and how to apply the results in parameter estimation. We consider a 10 meter long and 4 meter wide corridor. Two groups of pedestrians cross the corridor in opposite directions, thus creating a bi-directional flow.



(a) Phase error in queue test case.

(b) Diffusion error in queue test case.

FIGURE 5 Accuracy of solutions at time $t = 600$ s with different numerical methods and time step sizes/CFL numbers.



(a) Group 1 (left to right)

(b) Group 2 (right to left)

FIGURE 6 Typical class specific densities in simulation (coloured areas, purple and yellow) and positions of pedestrians in preprocessed experimental data (red and purple stars), both at time $t = 350$ s.

Experimental data

The data that was gathered from this experiment and is used here, includes positions in both x- and y-direction every 0.1 seconds, for 7 minutes. The data also includes identification numbers to trace individual pedestrians. We preprocessed the data to extract velocities of individual pedestrians every 0.1 seconds and to assign each pedestrian to a group: group 1 when entering on the left and leaving on the right, group 2 for pedestrians walking in the opposite direction. The preprocessed data shows that the setup resulted in lane formation. Most frequently, there are 2 lanes with pedestrians walking on their right hand side of the corridor. In some situations, there are 1, 3 or 4 lanes (*30*). A typical situation is shown in Figure 6.

Simulation results

The experimental data is compared to simulation results. These were obtained using a continuum pedestrian flow model (*31*) and simulation method (*32*). The simulation was performed for a

TABLE 1 Settings and some results of pedestrian simulation.

parameter set	β_u	β_o	N	resulting lanes
1	0.80	2.30	705	right
2	0.70	1.36	708	left
3	0.63	0.63	707	none
experimental data	na	na	709	right

corridor of the same width (4 meters) but longer (50 meters). Only the results for the centre 10 meter were used for comparison with the experimental data, see Figure 6.

In the simulation, pedestrians enter from both sides, with inflow demand equal to capacity. Group 1 moves from the left to the right, group 2 moves from right to left. Parameters were chosen to have realistic values: free flow speed is 1.34 m/s and jam density is 5.4 pedestrians/m^2. However, to compare the accuracy for different simulation results we used three different sets of parameter values for the avoidance of the own group (β_u) and the avoidance of the other group (β_o). For more details about the model and its parameters we refer to (*31*). We chose the parameter values such that the total number of pedestrians crossing the area would be very similar to those in the experimental data with $N = 709$. Furthermore, the parameters were chosen to give clearly distinguishable results, with either two lanes with pedestrians walking on their right hand side of the corridor (set 1) or on their left hand side (set 2), or with no lanes at all (set 3). We did not consider parameter settings resulting in more than two lanes. The settings are summarised in Table 1.

Accuracy

For both the experimental data and the simulation results, the centres of mass are computed and compared. We compute the phase error as in (4), but for both groups separately. For the diffusion error, we apply (8), again for both groups separately. We use the centre of mass of the reference solution F_{ref} as in (5) with $f_n = v_{x,n,u}$ the velocity in x-direction of the n-th pedestrian of group u and with N the total number of pedestrians of group u. Furthermore, the centre of mass of the test solution F_{test} is as in (6) with $f_i = v_{x,i,u}$ the velocity in x-direction of group u in the i-th cell. Also for the density ρ_i only pedestrians of group u are taken into account, i.e. ρ_i is the number of pedestrians of group u in cell i. We use the centre of mass in x-velocity direction instead of in density (ρ) direction because the density is not straightforward to determine from the experimental data, while the x-velocity is relatively simple to extract from the position data. Furthermore, it gives better insight into the differences in the results than the speed or y-velocity, because speed is always positive and in the experimental data close to free flow speed of the experiment (1.34 m/s) and because y-velocity is mostly close to zero.

We also compared the parameter sets with respect to traditional error measures. To compute the errors, at each time step, the experimental data for each pedestrian was compared with the simulations results at that location. The RMSE of the speed of each group at time t is defined as:

$$E_{\text{RMSE},|\text{v}_\text{u}|}(t) = \frac{\sqrt{\sum_{i \in I} \left[|v_{u,\text{ref}}(x_i, y_i, t)| - |v_{u,\text{test}}(x_i, y_i, t)| \right]^2}}{N_{\text{test}}(t)} \tag{9}$$

with (x_i, y_i, t) the position of the i-th pedestrian in the experimental data at time t, $u \in \{1, 2\}$ the group, $|v_{u,\text{ref}}(x_i, y_i, t)|$ its speed, $|v_{u,\text{test}}(x_i, y_i, t)|$ the speed at this location and time according

to the simulation and $N_{\text{test}}(t)$ the number of pedestrians in the area at time t according to the experimental data. The ME of the speed of each group at time t is defined as:

$$E_{\text{ME},|\text{v}_\text{u}|}(t) = \frac{\sum_{i \in I} \left[|v_{u,\text{ref}}(x_i, y_i, t)| - |v_{u,\text{test}}(x_i, y_i, t)| \right]}{N_{\text{test}}(t)} \tag{10}$$

The RMSE of the velocity in x-direction is defined as:

$$E_{\text{RMSE},\text{v}_\text{x}}(t) = \frac{\sqrt{\sum_{i \in I} \left[v_{x,\text{ref}}(x_i, y_i, t) - v_{x,\text{test}}(x_i, y_i, t) \right]^2}}{N_{\text{test}}(t)} \tag{11}$$

with the velocity in x-direction the weighted average velocity of both groups in the simulation:

$$v_{x,\text{test}}(x_i, y_i, t) = \frac{\rho_1 v_{x,1} + \rho_2 v_{x,2}}{\rho_1 + \rho_2} \tag{12}$$

The MAE of the velocity in x-direction is defined as:

$$E_{\text{MAE},\text{v}_\text{x}}(t) = \frac{\sum_{i \in I} |v_{x,\text{ref}}(x_i, y_i, t) - v_{x,\text{test}}(x_i, y_i, t)|}{N_{\text{test}}(t)} \tag{13}$$

Results

The errors are shown in Figures 7–9 and Tables 2 and 3. The figures show the development of the errors over time, the table shows the average error over the last 200 seconds. We note that the values of the errors vary largely over time, unlike in the road traffic test case. This shows that we would not be able to draw reasonable conclusions if we would only study the error at one moment in time as we did in the previous case.

TABLE 2 Errors in the pedestrian simulation, traditional measures

parameter set	$E_{\text{RMSE},\|\text{v}_1\|}$	$E_{\text{RMSE},\|\text{v}_2\|}$	$E_{\text{RMSE},\text{v}_\text{x}}$	$E_{\text{ME},\|\text{v}_1\|}$	$E_{\text{ME},\|\text{v}_2\|}$	$E_{\text{MAE},\text{v}_\text{x}}$
1	0.066	0.070	0.31	-0.028	-0.046	0.73
2	0.065	0.069	0.54	-0.030	-0.049	1.86
3	0.065	0.068	0.34	-0.027	-0.046	1.27

TABLE 3 Errors in the pedestrian simulation, new measures

parameter set	E_{phase}, x_1	E_{phase}, x_2	E_{phase}, y_1	E_{phase}, y_2	E_{diff}, v_1	E_{diff}, v_2
1	-0.57	0.64	-0.42	0.19	0.033	0.049
2	-0.68	0.51	1.27	-1.25	0.032	0.044
3	-0.45	0.52	0.41	-0.64	0.026	0.041

Most traditional error measures do not discriminate between the parameter sets: the RMSE of the group specific speeds ($|v_1|$ and $|v_2|$) are almost the same for each parameter set. The same holds for the ME of the group specific speeds. The only difference can be found in the velocities in x-direction, both the RMSE and the MAE show different values for different parameter sets. However, the error is largest if the lanes are swapped (i.e. pedestrians walk on their left hand side

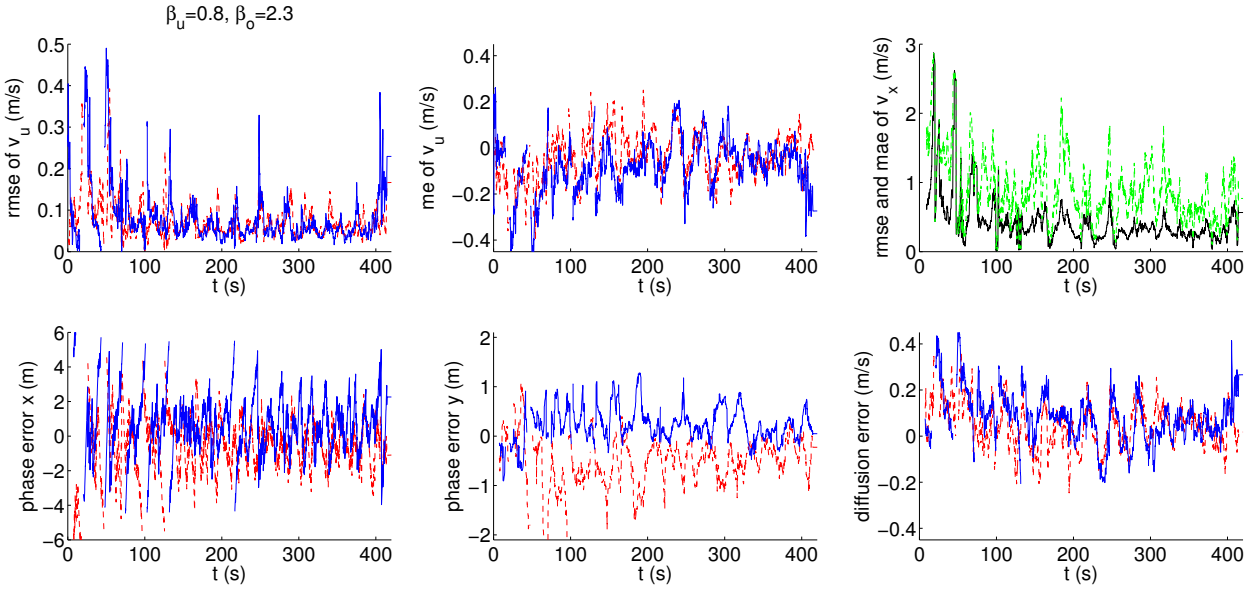


FIGURE 7 Errors using parameter set 1. Red dashed: group 1 (walking from left to right), blue solid: group 2 (walking from right to left), black solid: RMSE of v_x, green dashed: MAE of v_x.

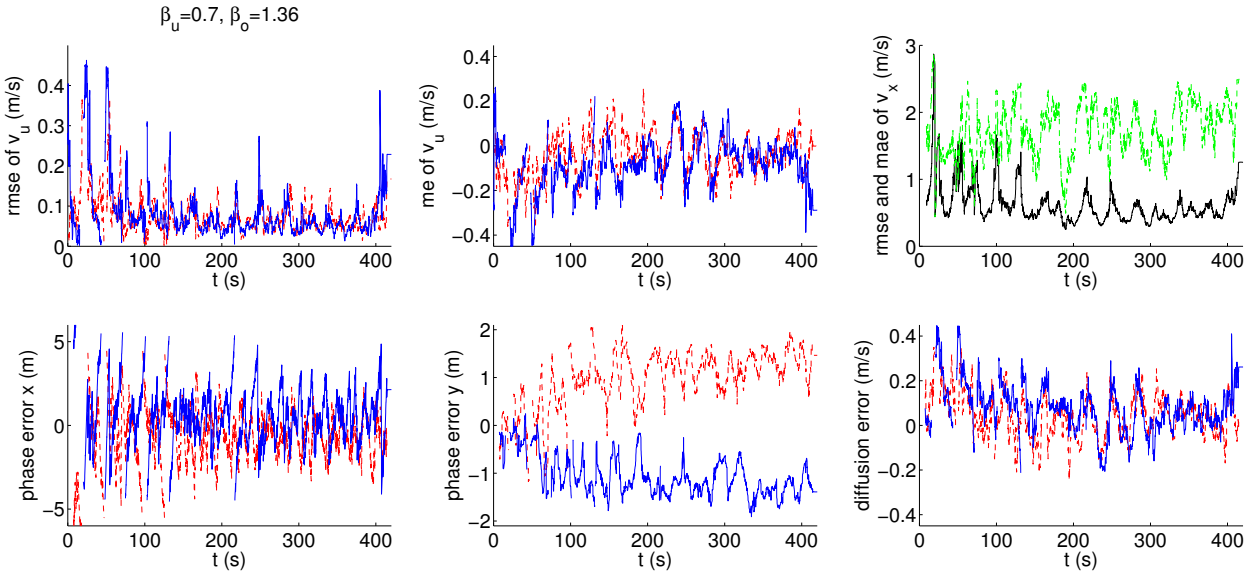


FIGURE 8 Errors using parameter set 2. Red dashed: group 1 (walking from left to right), blue solid: group 2 (walking from right to left), black solid: RMSE of v_x, green dashed: MAE of v_x.

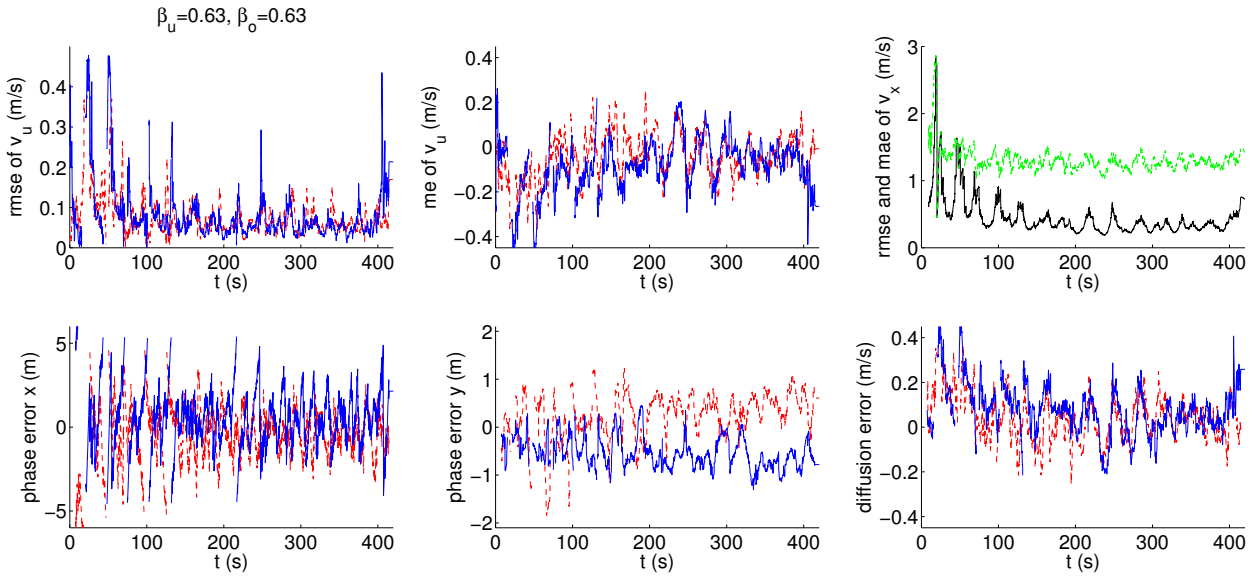


FIGURE 9 Errors using parameter set 3. Red dashed: group 1 (walking from left to right), blue solid: group 2 (walking from right to left), black solid: RMSE of v_x, green dashed: MAE of v_x.

of the corridor, instead of at their right hand side). The difference in RMSE of v_x between set 1 (correct lane formation) and set 3 (no lanes) is small.

The phase errors for both groups in x-direction are similar for each parameter set and we do not see any pattern. However, the other new measures do discriminate between the parameter sets. The absolute phase error (pe in y_1 and in y_2) is largest for parameter set 2, as expected. However, this is combined with a diffusion error of similar size than diffusion errors with the other parameter sets. The diffusion error is small, but clearly largest for set 1 and smallest for set 3.

Discussion

We note that in all simulations, the total number of pedestrians crossing the corridor is (almost) the same as in the experimental data. Therefore, an error measure based on the total flux (such as used by Kretz et al. (*9*)) would not be able to discriminate between the parameter sets. An other traditional error measure is based on the likelihood of the simulation result (*11*). This may be an appropriate approach to estimate parameters. However, it does not give any insight into the type of inaccuracies and how they may be reduced.

The results of the traditional accuracy measures give little insight in the cause of the error and would prefer parameter set 3 over set 2, while in most applications, the lane formation is essential and parameter set 2 would yield better results than set 3. The results for the phase error in y-direction are similar to those for the RMSE and MAE of v_x. However, the phase error gives insight into the type of error: the location of the lanes is predicted wrongly in set 2. The diffusion error does not add much insight because the speeds both in the experimental data and in the simulation are always close to the free flow speed of 1.34 m/s. Furthermore, in the experiments, three or four lanes were formed in some instances, while in the simulations there are only two. For a test case with more variation in the speeds or with a simulation setting with more lanes, the diffusion error may be able to discriminate between different parameter sets.

DISCUSSION, CONCLUSION AND OUTLOOK

We have introduced two new accuracy measures, phase error and diffusion error. They give better insight in the type of error and take into account specific features of traffic and pedestrian flows. The accuracy measures quantify the error using a method based on the centre of mass of a region with high density or speed. The new accuracy measures are more meaningful in traffic and pedestrian flow applications than traditional methods. Because of their generic formulation, they can be used in a wide range of applications: from the assessment of a data collection method to the quantification of the accuracy of traffic information.

We applied the new accuracy measures to two test cases. In the first case we compared different simulation methods for road traffic flow (one-dimensional), in the second case we compared different parameter settings of a pedestrian flow model. Both cases show that the way the accuracy is defined is important. A good result according to one accuracy measure, may be a bad result according to an other accuracy measure. Furthermore, some accuracy measures fail to discriminate between the different test settings and give similar accuracies independent of the simulation method or parameter set. The new accuracy measures give insight into the type of error by showing, for example, that the phase error is small, while the diffusion error is large. This indicates that, if one wants to increase the accuracy, the model or simulation method should be improved to reduce diffusion.

It is important to note that, especially for the phase error, the interpretation of the error remains important. In some cases (such as while predicting the spill back of a queue until the upstream intersection), a large phase error should be avoided, while in other cases (such as lane formation in bi-directional pedestrian flow) a phase error may only indicate that the lanes are swapped.

The current applications focus on simple problems and networks with different phenomena at different location have not yet been considered. The new accuracy measures could be used as building blocks for an assessment method for accuracy on networks. This can for example be done by splitting the networks into parts and isolating areas with different layout or flow characteristics. Furthermore, we have now explored the applications for one-dimensional road traffic flow and for two-dimensional pedestrian flow. However, the methods may be extended to other applications such as network flows using a Network Fundamental Diagram (*33*).

An other future research direction is to extend the measures to changes over time. In our current applications, either the error at just one moment was considered, or the average error over a longer period of time. Time could be considered explicitly in the measures by also comparing the centre of mass in time direction. This may especially improve the quantification of the phase error.

Finally, the measures may be extended for heterogeneous flow including for example trucks and passenger cars or different types of pedestrians. In (*26*) we have developed a method to give trucks a higher weight when computing the centre of mass than passenger cars. This method can be extended to pedestrian flows.

ACKNOWLEDGEMENT

This research is performed as part of the NWO Aspasia grant of Daamen (first and second author) and as part of the NWO-VICI project of Hoogendoorn (third author).

REFERENCES

[1] Huber, G., K. Bogenberger, and R. L. Bertini, New Methods for Quality Assessment of Real Time Traffic Information. In *Transportation Research Board 93th Annual Meeting Compendium of Papers*, Washington D.C., 2014.

[2] Ward, J. A. and R. E. Wilson, Criteria for convective versus absolute string instability in car-following models. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 2011.

[3] del Castillo, J., Three new models for the flow-density relationship: Derivation and testing for freeway and urban data. *Transportmetrica*, Vol. 8, No. 6, 2012, pp. 443–465.

[4] van Wageningen-Kessels, F. L. M., B. van 't Hof, S. P. Hoogendoorn, J. W. C. van Lint, and C. Vuik, Anisotropy in generic multi-class traffic flow models. *Transportmetrica A: Transport Science*, Vol. 9, No. 5, 2013, pp. 451–472.

[5] Duives, D. C., W. Daamen, and S. P. Hoogendoorn, State-of-the-art crowd motion simulation models. *Transportation Research Part C: Emerging Technologies*, Vol. 37, No. 0, 2013, pp. 193–209.

[6] Punzo, V. and F. Simonelli, Analysis and Comparison of Microscopic Traffic Flow Models with Real Traffic Microscopic Data. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1934, 2005, pp. 53–63.

[7] Hoogendoorn, S. P. and W. Daamen, Microscopic Calibration and Validation of Pedestrian Models: Cross-Comparison of Models Using Experimental Data. In *Traffic and Granular Flow '05* (A. Schadschneider, T. Pöschel, R. Kühne, M. Schreckenberg, and D. Wolf, eds.), Springer Berlin Heidelberg, 2007, pp. 329–340.

[8] Ossen, S., S. P. Hoogendoorn, and B. G. H. Gorte, Interdriver Differences in Car-Following: A Vehicle Trajectory-Based Study. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1965, 2006, pp. 121–129.

[9] Kretz, T., S. Hengst, and P. Vortisch, Pedestrian Flow at Bottlenecks - Validation and Calibration of Vissim's Social Force Model of Pedestrian Traffic and its Empirical Foundations. In *3rd International Symposium on Transport Simulation (IST08): symposium proceedings*, 2008.

[10] Kesting, A. and M. Treiber, Calibrating Car-Following Models by Using Trajectory Data: Methodological Study. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2088, 2008, pp. 148–156.

[11] Robin, T., G. Antonini, M. Bierlaire, and J. Cruz, Specification, estimation and validation of a pedestrian walking behavior model. *Transportation Research Part B: Methodological*, Vol. 43, No. 1, 2009, pp. 36–56.

[12] Ciuffo, B., V. Punzo, and M. Montanino, *The Calibration of Traffic Simulation Models: Report on the assessment of different Goodness of Fit measures and Optimization Algorithms*. Joint Research Centre - European Commission, 2012.

[13] Ciuffo, B., J. Casas, M. Montanino, J. Perarnau, and V. Punzo, Gaussian Process Metamodels for Sensitivity Analysis of Traffic Simulation Models. *Transportation Research Records: Journal of the Transportation Research Board*, Vol. 2390, 2013, pp. 87–98.

[14] Daganzo, C. F., Requiem for second-order fluid approximations of traffic flow. *Transportation Research Part B: Methodological*, Vol. 29, No. 4, 1995, pp. 277–286.

[15] Chalons, C., P. Goatin, and N. Seguin, General constrained conservation laws. Application to pedestrian flow modeling. *Networks and Hetereogeous Media*, Vol. 8, No. 2, 2013, pp. 433–463.

[16] Zielke, B. A., R. L. Bertini, and M. Treiber, Empirical Measurement of Freeway Oscillation Characteristics: An International Comparison. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2088, 2008, pp. 57–67.

[17] Johansson, A., D. Helbing, H. Z. Al-Abideen, and S. Al-Bosta, From Crowd Dynamics to Crowd Safety: A Video-Based Analysis. *Advances in Complex Systems*, Vol. 25, 2008, pp. 9–41.

[18] Jiang, Y., T. Xiong, S. Wong, C.-W. Shu, M. Zhang, P. Zhang, and W. H. Lam, A reactive dynamic continuum user equilibrium model for bi-directional pedestrian flows. *Acta Mathematica Scientia*, Vol. 29, No. 6, 2009, pp. 1541 – 1555.

[19] Lighthill, M. J. and G. B. Whitham, On Kinematic Waves II: A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, Vol. 229, No. 1178, 1955, pp. 317–345.

[20] Colombo, R. M., P. Goatin, and M. D. Rosini, A macroscopic model for pedestrian flows in panic situations. In *4th Polish-Japan Days*, Madralin, Poland, 2010, Vol. 32, pp. 255–272.

[21] Richards, P. I., Shock Waves on the Highway. *Operations Research*, Vol. 4, No. 1, 1956, pp. 42–51.

[22] Smulders, S., Control of freeway traffic flow by variable speed signs. *Transportation Research Part B: Methodological*, Vol. 24, No. 2, 1990, pp. 111–132.

[23] Daganzo, C. F., The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, Vol. 28, No. 4, 1994, pp. 269–287.

[24] Lebacque, J.-P., The Godunov scheme and what it means for first order traffic flow models. In *Transportation and Traffic Theory: Proceedings of the 13th International Symposium on Transportation and Traffic Theory, 1996* (J.-B. Lesort, ed.), Pergamon, 1996, pp. 647–677.

[25] Leclercq, L., J. Laval, and E. Chevallier, The Lagrangian coordinates and what it means for first order traffic flow models. In *Transportation and Traffic Theory 2007* (R. E. Allsop, M. G. H. Bell, and B. G. Heydecker, eds.), Elsevier, Oxford, 2007, pp. 735–753.

[26] van Wageningen-Kessels, F. L. M., *Multi class continuum traffic flow models: Analysis and simulation methods*. Ph.D. thesis, Delft University of Technology/TRAIL Research school, Delft, 2013.

[27] Courant, R., K. Friedrichs, and H. Lewy, On the partial difference equations of mathematical physics. *IBM Journal*, 1967, pp. 215–234.

[28] LeVeque, R. J., *Finite volume methods for hyperbolic problems*. Cambridge texts in applied mathematics, Cambridge University Press, Cambridge, 2002.

[29] Daamen, W. and S. P. Hoogendoorn, Controlled experiments to derive walking behaviour. *Journal of Transport and Infrastructure Research*, Vol. 3, No. 1, 2003, pp. 39–59.

[30] Campanella, M. and W. D. S. P. Hoogendoorn, Calibration of pedestrian models with respect to lane formation self-organisation: Exploring chaos to obtain coherent behaviour. In *TRAIL Congress (DVD)*, Rotterdam, 2008.

[31] Hoogendoorn, S. P., F. L. M. van Wageningen-Kessels, W. Daamen, and D. C. Duives, Continuum Modelling of Pedestrian Flows: From Microscopic Principles to Self-Organised Macroscopic Phenomena. *Physica A*, in press.

[32] van Wageningen-Kessels, F. L. M., W. Daamen, and S. P. Hoogendoorn, The two-dimensional Godunov scheme and what it means for macroscopic pedestrian flow models, 2015, extended abstract accepted, full paper in preparation for submission to ISTTT and subsequent publication in Transportation Research series.

[33] Daganzo, C. F. and N. Geroliminis, An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological*, Vol. 42, No. 9, 2008, pp. 771–781.