Palacký University Olomouc

Transactions on Transport Sciences
Peer-Reviewed Open Access Journal
Vol. 3/2021 DOI: 10.5507/tots.2022.001
journal homepage: www.tots.upol.cz

Three-Step Performance Assessment of a Pedestrian Crossing Time Prediction Model

CHIARA GRUDEN[a], IRENA IŠTOKA OTKOVIĆ[b], MATJAŽ ŠRAML[a*]

a. Faculty of Civil Engineering, Transportation Engineering and Architecture, University of Maribor, Smetanova ulica 17, 2000 Maribor, Slovenia
b. Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek, Vladimira Preloga 3, 31000 Osijek, Croatia

KEYWORDS: Pedestrian; micro-simulation; neural network; crossing time; validation

ABSTRACT: Pedestrian behavior and safety are emerging issues in current transportation. One way to safely study pedestrian dynamics, especially at potential conflict points such as crosswalks, is through micro-simulation. This tool provides the opportunity to repeatedly study pedestrian behavior and safety under different scenarios of interest. However, to obtain reliable results, micro-simulation models need to be calibrated and their parameters fine-tuned. One way to methodically calibrate these models is to identify the outcomes of interest, develop a predictive model for those specific outcomes, and use it as a tool to fine-tune the input parameters of the micro-simulation model. To be reliable, the results of the predictive model should be comparable to those of the micro-simulation model, and these should be validated.

The aim of this research is to present a predictive model of pedestrian behavior and to evaluate this model and a conventional micro-simulation model developed using Vissim/Viswalk, given that the chosen common output is pedestrian crossing time.

To achieve this goal, a multi-step procedure is followed, which is part of a more general methodological framework for calibrating the Vissim/Viswalk micro-simulation model. This evaluation consisted in a three-step validation procedure, i.e. visual, conceptual and operational validation. Operational (statistical) validation was performed by comparing the variances of the results to understand whether the predicted sample is representative of the simulated sample.

A correlation of 97% have been found between the predicted and micro-simulated crossing time values, with mean values of 6.41s and 6.32s for the simulated and predicted crossing times, respectively. Furthermore, both the predicted and simulated crossing time values fall within the ranges found in the literature for field measurements of this variable, indicating good agreement with real observed pedestrian behavior.

1. INTRODUCTION

A sophisticated tool for transport infrastructure planning and design is micro-simulation: it allows to study different infrastructural options without their actual realization, in order to better meet the needs of the considered road users. However, to be useful, the accuracy and reliability of the developed models and the obtained results should be verified through the steps of calibration and validation. This paper is part of a broader research aimed at developing a calibration method for pedestrian micro-simulation models by applying a specifically structured neural network as a prediction tool. This research topic is specifically studied in relation to the action of pedestrians crossing the road at a roundabout entry leg. Although this type of location has been extensively studied from vehicular traffic perspective, focusing on the impact of pedestrians on vehicular capacity (Bak, & Kiec 2012; Brilon, & Miltner, 2005; Chen, Shao, & Hao 2008), research on pedestrian behavior at this type of urban location is still ongoing (Obsu, Meurer, Kassa, & Klar, 2016; Thakur, & Biswas, 2019). Also, the widespread use of this type of infrastructure in today's transport networks, highlighting the benefits in terms of safety and capacity for motorized traffic (Ambros, Novak, Borsos, Hoz, Kiec, Machcinik, & Ondrejka, 2016; Chen, 2013), makes it even more important to examine these benefits for the most vulnerable road users as well.

The paper, which presents the validation of the neural network prediction model results, is divided into 6 sections: after the introduction, an overview of the existing models, of the specifics of Social Force Model and of its implementation in Vissim/Viswalk is given; Section 2 explains the addressed problem, the study site and the developed micro-simulation model; Section 3 explains the process of parameter selection and defines the ranges of the same. Section 4 gives a brief overview of the formulated predictive model and its training. Section 5 explains the method used to evaluate the results obtained by the predictive model and validates them applying three different methods: visual, conceptual and operational validation. The last two sections deal with the discussion and conclusions of the research.

2. OVERVIEW OF EXISTING MICROSCOPIC MODELS

Micro-simulation models are very powerful tools for investigating pedestrian behavior and safety. Three main modelling approaches can be identified:

- Physical force based models;
- Cellular automata (CA) models;
- Queue models.

Physical models, which include Helbing's Social Force model (Helbing, 1998a), the utility maximization model, and Okazaki's magnetic force model (Teknomo, Takeyama, & Inamura, 2000), follow a common structure characterized by two terms: a factor that guides pedestrians to their destination and a term that represents the repulsive effects among pedestrians themselves and pedestrians and obstacles (Helbing, 1998b; Shiwakoti, Sarvi, & Rose, 2008).

Cellular automata, which include the cellular automata model of Blue and Adler (Blue, & Adler, 2001) and the benefit-cost model of Gipps and Marksjö (Gipps, & Marksjö, 1985), are based more on the discretization of space into cells. Each cell can be occupied by only one agent and its dimensions are determined by the researchers (Kouskoulis, Spyropoulou, & Antoniou, 2018).

Queueing models, mainly used to study evacuation dynamics, govern pedestrian movement by modeling the facility as a network of arcs (openings) and nodes (rooms) and pedestrians as individual flow objects. In these models, the recorded evacuation time and the initial reaction time play the main role (Teknomo et al., 2000).

On the other hand, artificial intelligence tools, such as neural networks, have been widely used recently, including in the field of transportation, especially to overcome issues related to the lack of analytical description of a problem or the difficulty of evaluating the factors/parameters that govern the problem itself.

Some efforts have been made by researchers to implement neural networks in the process of micro-simulation model calibration: (Li, & Yeh, 2001) applied this tool to calibrate a CA traffic model, while (Otković, Tollazzi, & Šraml, 2013) used them in the calibration process of the car-following model.

The aim of this research is to implement the neural network prediction function in the calibration process of a pedestrian micro-simulation model. To achieve this, the first important result is to validate the outputs predicted by the neural network and those obtained by the micro-simulation.

2.1 The Social Force Model and its parameters

The Social Force Model (SFM) was first introduced by (Helbing, 1998b) and it is nowadays implemented in the simulation software Vissim/Viswalk (PTV, 2018).

The basic concept of the SFM is that each pedestrian, defined by its desired speed and target time, moves towards its destination ruled by the so-called social forces (Eq.1).

(Eq. 1) $F_\alpha\ (t) = F_\alpha^0 + F_{\alpha\beta} + F_{\alpha B} + F_{\alpha i}$

These effects are:

The attractive forces leading each pedestrian to its goal/destination F_α^0 (Eq.2);

The repulsive forces among pedestrians $F_{\alpha\beta}$ and among the individual and obstacles $F_{\alpha\beta}$ (Eq. 3-4);

The attractive forces due to other pedestrians/objects $F_{\alpha i}$ (Eq. 5).

Mathematically, they are expressed as follows:

(Eq. 2) $F_\alpha^0(v_\alpha, v_\alpha^0 e_\alpha) = \frac{1}{\tau_\alpha}(v_\alpha^0 e_\alpha - v_\alpha)$

(Eq. 3) $F_{\alpha\beta}(r_{\alpha\beta}) = -\nabla_{r_{\alpha\beta}} V_{\alpha\beta}[b(r_{\alpha\beta})]$

(Eq. 4) $F_{\alpha B}(r_{\alpha B}) = -\nabla_{r_{\alpha B}} U_{\alpha B}(\|r_{\alpha B}\|)$

(Eq. 5) $F_{\alpha i}(\|r_{\alpha i}\|, t) := -\nabla_{r_{\alpha i}} W_{\alpha i}(\|r_{\alpha i}\|, t)]$

where:

v_α pedestrian actual velocity

v_α^0 pedestrian desired velocity

e_α pedestrian desired direction

τ_α relaxation time

$r_{\alpha\beta}$ distance

$V_{\alpha\beta}(b), U_{\alpha B}(\|r_{\alpha B}\|)$ repulsive potentials

$W_{\alpha i}(\|r_{\alpha i}\|)$ attractive potential

The original Social Force Model expression was adapted to the simulation needs in Vissim/Viswalk, adding some parameters, which govern pedestrian walking behavior and influence the equation (Eq.6):

(Eq.6) $F = A_{soc_{isotropic}}\omega(\lambda)\exp\left(-\frac{d}{B_{soc_{isotropic}}}\right)n + A_{soc_{mean}}\exp\left(-\frac{d}{B_{soc_{mean}}}\right)n$

where: $\omega(\lambda) = (\lambda + (1+\lambda)(1+cos(\varphi))/2$

The highlighted parameters can be defined as follows (PTV, 2018):

TAU (τ) relaxation time as expressed in Helbing's original model.

LAMBDA_MEAN (λ) amount of anisotropy. It regulates the effect of phenomena that take place in the back of the considered pedestrian.

$A_{\_soc\_isotropic}$ and $B_{\_soc\_isotropic}$ non-measurable parameters, that control the two forces among pedestrians.

$A_{\_soc\_mean}$ and $B_{\_soc\_mean}$ define respectively the strength and typical range of the social force between pedestrians.

NOISE introduces the random forces, which are systematically added to the calculated forces.

REACT_TO_N number of pedestrians considered for the calculation of the forces.

SIDE_PREFERENCE defines whether opposing pedestrians prefer using the right or left side when passing each other.

QUEUE_ORDER and QUEUE_STRAIGHTNESS specify the properties of the queue.

It should be pointed out that these four last parameters do not directly appear in Eq. 6, nevertheless they are at the basis of pedestrian walking behavior and are nested in other terms (e.g. noise is at the basis of the calculation of random forces added to the systematically calculated total forces; react_to_n is used in the calculation of the total force to consider the effect of surrounding pedestrians, etc.).

As (Kretz, Lohmiller, & Sukennik, 2018) states, although the model is conceptually easy to understand, the parameters governing it are mainly abstract and difficult to be measured and evaluated. In the next sections, after describing the study site, the parameter selection is explained, which provides an initial selection of parameters and their ranges relevant to pedestrian studies.

3. METHODOLOGY

3.1 Addressed problem and case study location

The problem addressed in this research is the validation of the results provided by a predictive model that reproduces the crossing action - in terms of time - of pedestrians, as simulated by Vissim/Viswalk. To achieve this goal, a multi-step procedure (Figure 1) was applied, which is part of the general methodological framework used to calibrate Vissim/Viswalk. The method followed consists of three pillars: real world behavior, micro-simulation and prediction. Each of these steps is independent (in terms of workflow), but defined features of each pillar influence the others.

In fact, real-world recordings were used to obtain geometry and flow data to define the micro-simulation model and the real-world crossing time, which are used in further passages to obtain the calibrated model. Vissim/Viswalk simulations were run to obtain a set of simulated crossing times that could be compared to the predicted ones, and to generate a dataset consisting of combinations of input parameters and associated crossing times, which was used as a training set for the formulated neural network.

A prerequisite for the creation of the two training and prediction databases is, of course, the selection of the input parameters, which is described in sub-section 3.3. These two databases were used to let the formulated neural network learn the relationship between the selected input parameters and the selected output (training database) and to validate it

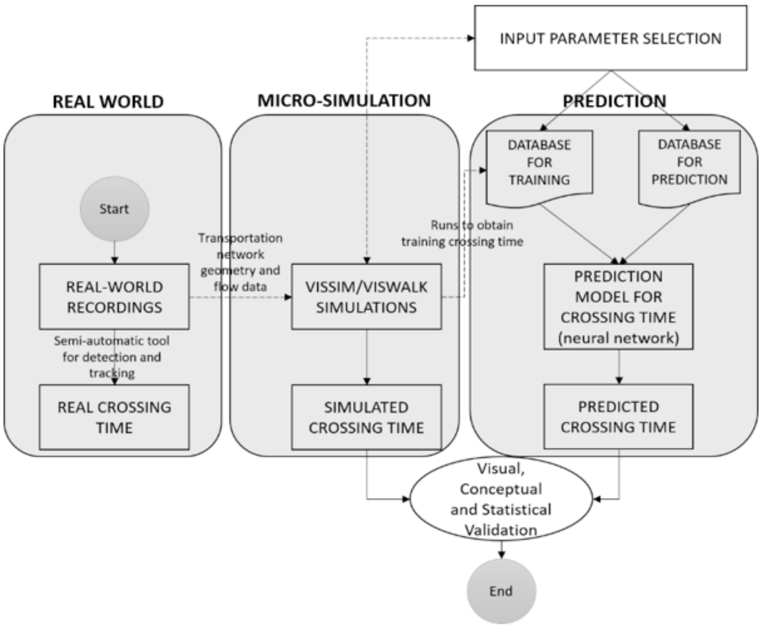

Figure 1 First steps for the development of the calibration methodology

Figure 2 Study location: (a) urban roundabout set in Monfalcone (GO) Italy; (b) the crosswalk under study modelled in Vissim/Viswalk

on the basis of new input combinations (prediction database) for which the network should predict the crossing time. Finally, the predicted result and the simulated crossing time are compared and statistically validated.

The specific crossing action addressed in this study involves unsignalized pedestrian crossings within the roundabout's influential area. A roundabout located in the urban area of the Italian city of Monfalcone was selected to study pedestrian behavior at this specific intersection typology. This location (Figure 2) is particularly interesting because there are large flows of both motorized - vehicles - and non-motorized - pedestrians - users. The crosswalk passes over two traffic lanes and is characterized by a length of 10.25 m and a width of 4 m.

This site was reproduced in Vissim/Viswalk by inputting its geometric properties and observed inflows. The vehicular movement was simulated by applying specific speed reduction areas, which allowed to simulate the particular yielding behavior of Italian drivers, who do not normally stop suddenly at crossings. On the other hand, pedestrian behavior was modelled by implementing the option of links used as pedestrian areas, which visually better reproduced the observed movement (for a more detailed description of the model, refer to (Gruden, Otković, & Šraml, 2020)).

3.2 Parameter selection and database creation

The issue of parameter selection for fine-tuning them, has been addressed by various authors (Kretz et al., 2018; Liao, Chraibi, Seyfried, Zhang, Zheng, & Zhao, 2017; Rudloff, Matyus, Seer, & Bauer, 2011; Zhong, Hu, Cai, Lees, & Luo, 2015). Limiting the review to the Social Force Model, Table 1 reports on some investigations and the calibrated parameters.

A considerable work was done by (Kretz et al., 2018) who, starting from an initial set of 13 parameters, through 3 steps reduced them to 4, changing also their values.

In this study, since the examined condition is strictly linked to the interaction between pedestrians and vehicles, 5 pedestrian parameters and 3 car-following model parameters were selected and are listed in Table 2.

With these parameters, the databases for the training of the neural network and its application were created. Two databases were developed: the training database consisted of 100 combinations of input parameters with a changing step of 0.1. Each combination was simulated via Vissim/Viswalk and the obtained crossing time was entered into the database. A second database, corresponding to the 20% of the previous one, was created for the prediction of the neural network and for the evaluation of its ability to generalize.

Authors	Calibrated parameters
(Liao et al., 2014)	free speed, anisotropy of social force; interaction strength and its range
(Rudloff et al., 2011)	parameters related to attractive and repulsive forces
(Zhong et al., 2015)	interaction strength and range, obstruction effects of physical interactions, 4 social force parameters + 8 scenario specific parameters
(Kretz et al., n.d.)	radius of pedestrians, A social, B social, B physical, border, A social Isotropic, B social Isotropic, τ, friction force, side preference right, velocity dependence, λ, longitudinal scale consider at maximum n pedestrians
(Hongfei et al., 2009)	pedestrian size, desired speed, time pressure
(Johansson, Helbing, Shukla, 2008)	interaction strength and range, anisotropy
(Zeng et al., 2017)	Interaction strengths and ranges for repulsive and attractive forces, relative distance, relative conflicting time, "footprint" effect.

Table 1 Researches on pedestrian simulation parameter fine-tuning

Input	Name	Description	Min	Max
I1	Tau	Relaxation time	0.05	2
I2	Lambda	Amount of anisotropy	0	0.4
I3	Asoc_iso	Parameter governing pedestrian forces	3	7
I4	Bsoc_iso	Parameter governing pedestrian forces	0.1	10
I5	Side_pref	Side preference	-1	1
I6	Avg standstill distance [m]	Average standstill distance	1	3
I7	Additive part of safety distance [m]	Additive part of safety distance	1	5
I8	Multiplicative part of safety distance [m]	Multiplicative part of safety distance	1	6

Table 2 Pedestrian and vehicular micro-simulation parameters selected in the current study

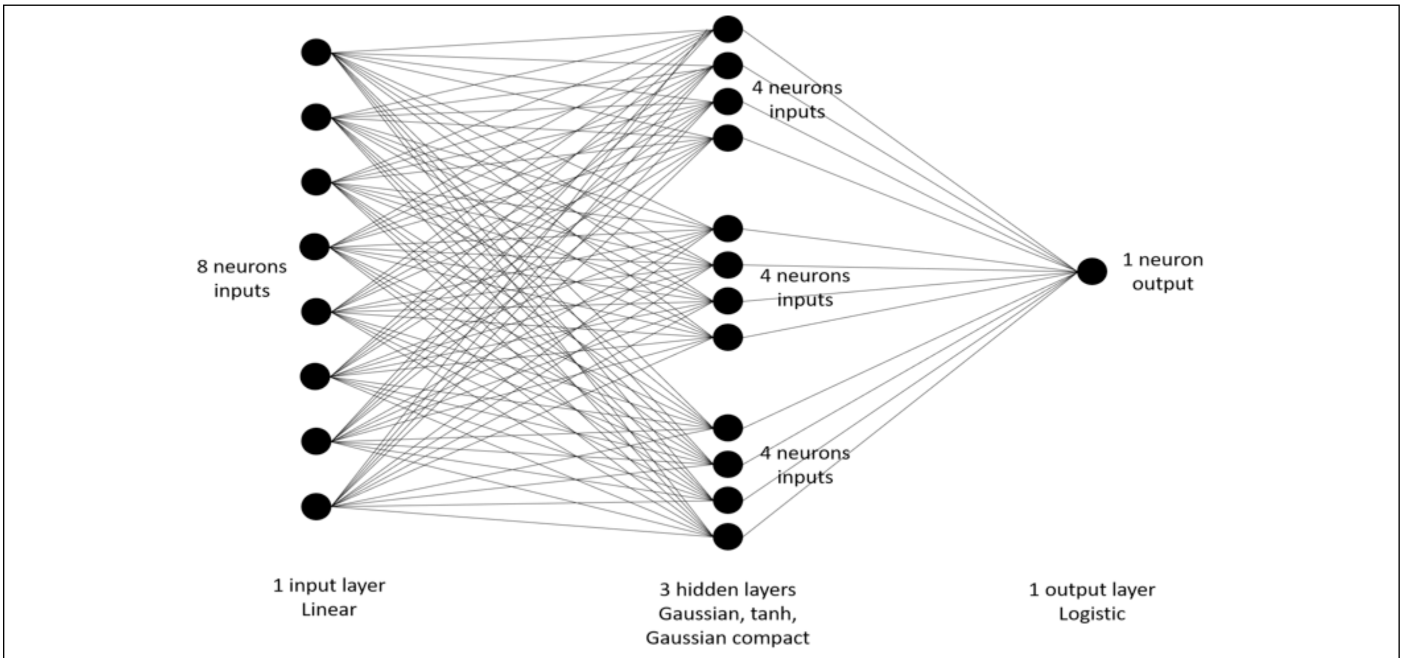

Figure 3 Structure of the chosen ward network

3.3 The prediction model

Neural networks were investigated as tool to develop the prediction model for pedestrian crossing time. The formulation of the neural network was performed using NeuroShell2 (Advanced Neural Network and Genetic Algorithm Software, n.d.). Based on previous author's experiences (Otković et al., 2013; Otković, Varevac, & Šraml, 2015), a first screen of all available network structures was made and a ward network was selected (Figure 1). The structure of the chosen neural network consists of 5 layers. The 8 input parameters chosen for the network are represented by the neurons of the input layer, the chosen output - crossing time - is concretized by 1 neuron of the output layer; 3 hidden layers consist of 4 neurons each, but they differ for the prediction function (Figure 3). The fundamental feature of the selected ward network is the composition of the activation functions: although the input slab is connected to the hidden layers by a linear function, the hidden layers are connected to the output slab by three different functions, namely the Gaussian, the hyperbolic tangent and the Gaussian compact functions (Figure 3). The neural network was trained on the training database obtained from the results of Vissim/Viswalk simulations. The database, briefly introduced in Sections 2 and 3, was used for the learning and validation steps of the network, 80% for training and the remaining 20% as a testbed for the initial evaluations. An additional database, developed according to the same criteria as the first one and corresponding to the 20% of the training dataset, was generated after training and testing the network to validate it for generalization.

4. VALIDATION OF THE RESULTS

4.1 Visual validation

Visual validation was performed by comparing the results of the neural network and the simulation model first via descriptive statistics and then via graphs. In Figure 4, the distributions of the two samples are reported. Descriptive statistics of the samples indicate mean crossing times of 6.41 s and 6.32 s for simulated and predicted outputs, respectively, and standard deviations of 3.50 and 3.43. Also, predicted crossing time values stand within the range [3.4; 12.51] s, similarly to the simulated ones, which belong to the range [2.44; 14.5] s. For the training database, a correlation of 97% was obtained between the simulated and predicted data, and a mean absolute error of 0.559 s was calculated.

The statistics worked out on the results of the generalization step, which used a much more restricted dataset than the training dataset, also confirm the previous performance, returning a 94% correlation and a mean absolute error of 1.76 s. Since the asymmetry calculated in the descriptive statistics indicates a probable non-normal distribution of the samples, their normality was checked by applying the Anderson-Darling test. This test consists in evaluating whether the p-value calculated for the selected sample is lower than the selected significance level (5%). If it is, the distribution is not normal, otherwise the non-normality of the sample cannot be established and additional tests should be devised. The results reported in Table 3 strengthen the hypothesis of non-normality, providing p-values lower than the significance level for both simulated and predicted crossing times.

	Simulated (Vissim)	Predicted (Neural Network)
AD	6.355	10.424
P-Value	<0.005	<0.005

Table 3 Results of Anderson-Darling test

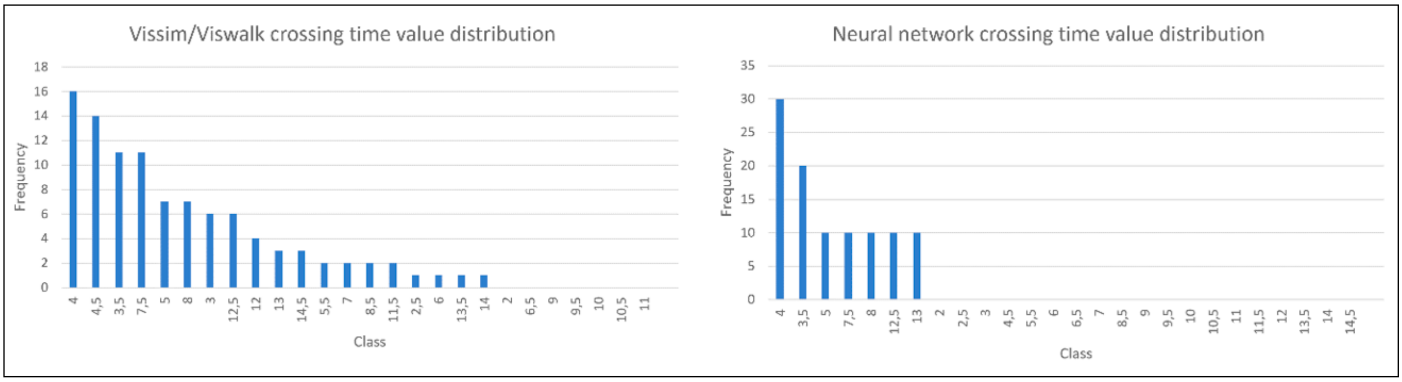

Figure 4 Distribution of crossing time values simulated with Vissim/Viswalk and calculated by the neural network

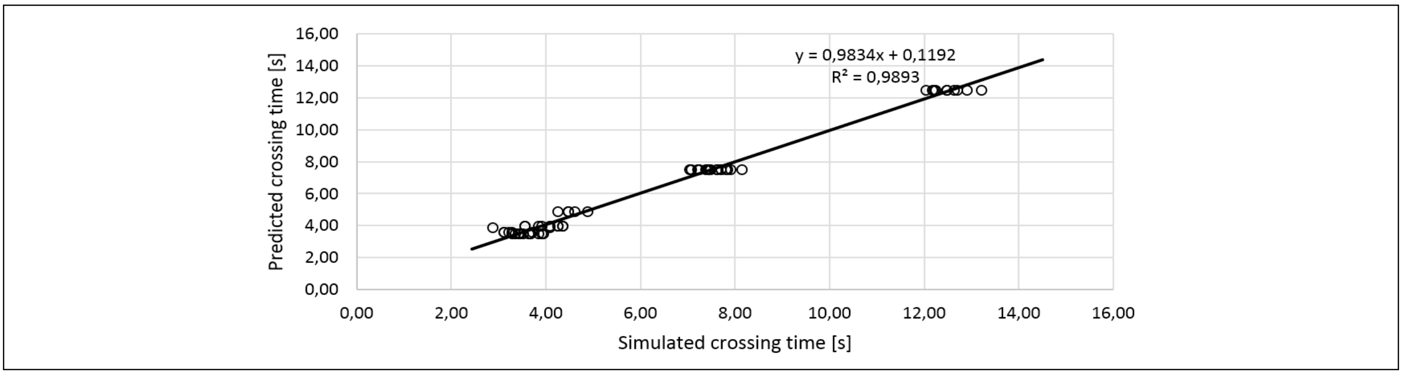

Figure 5 Comparison between the simulated outputs and the accepted values, predicted by the ward network

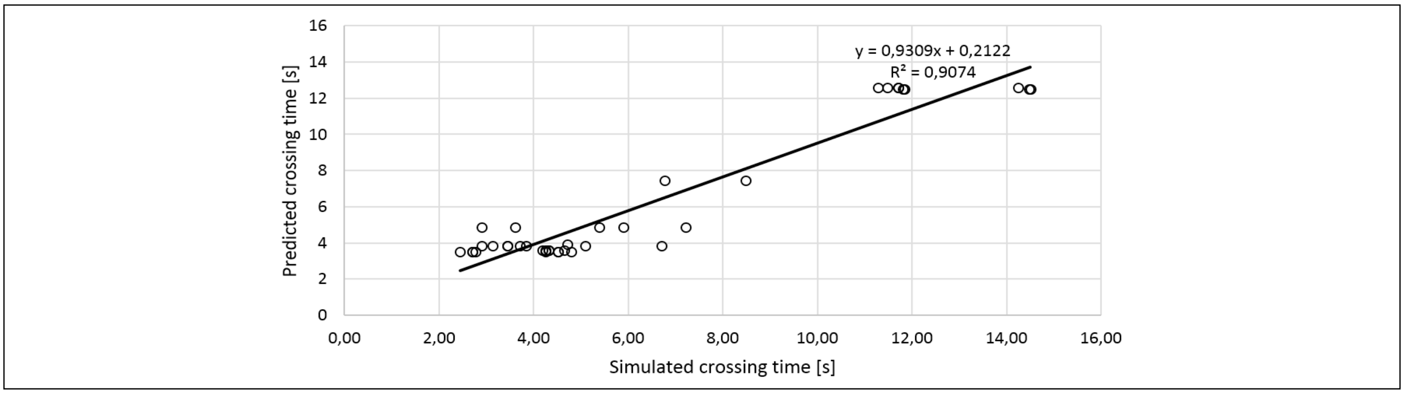

Figure 6 Comparison between the simulated outputs and the not accepted values, predicted by the ward network

Figure 5 and Figure 6 show the graphical comparison of the 2 model outputs: in both graphs, the x-axis represents the simulated outputs, while the y-axis shows the predicted ones. Figure 5 shows the accepted predicted values compared to the simulated ones, while Figure 6 represents the rejected pairs.

Figure 5 shows a high level of correspondence between the simulated and predicted results, confirming the findings obtained through descriptive statistics; the regression equation also provides promising parameters, in particular a coefficient of determination of 98.8%. This value is higher than the 97% reported by previous statistics because, unlike these elaborations (which considered the entire datasets), only the accepted values are considered here. Moreover, three clusters of outputs can be identified: they correspond to [2;4] s, [6;8] s and [12;14] s. In particular, the cluster [2;4] s corresponds to critical parameter combinations that were expected to give unrealistic results, nevertheless it was decided to use them in order to train the network on them as well and avoid future errors. As for the results in Figure 5, they were not accepted because they have a percentage error higher than 5%, yet it can be seen from the regression equation that they still have a good coefficient of determination index R^2=90%. Moreover, the graph in Figure 5 shows that the highest errors occur in the most critical cluster, while they are less pronounced in the other two clusters.

Additionally, a Kruskal-Wallis test was developed to understand, which of the input parameter most significantly affect crossing time. This test, comparing p-value to the set significance level – in this case 0.05, states if the null hypothesis „the population medians are all equal" is valid or must be rejected. If the p-value is less than the chosen significance level, the null hypothesis must be rejected, otherwise it can be confirmed. In the considered study, Kruskall-Wallis has been applied to all the selected input parameters to understand their influence on the output „crossing time". It turned out that all p-values are lower than the significance level (Table 6), and therefore all parameters significantly affect crossing time. For seek of completeness it has to be clarified that P-values reported in Table 4 are all null: actually, they differentiate one from the other for such low values, that do not influence the results, when compared to the significance level and thus they have been omitted. To understand which parameter influences the most the crossing time, the same test provides H-values. The greater this value is, the most influential the relative parameter. Table 6 summarizes also these results.

Input parameters	Description	H-values	P-values	Significance level
I1	Tau	80.63	0	0.05
I2	Lambda	47.67	0	0.05
I3	Asoc_iso	80.65	0	0.05
I4	Bsoc_iso	69.69	0	0.05
I5	Side_pref	27.15	0	0.05
I6	Avg standstill distance	48.04	0	0.05
I7	Additive part of safety distance	53.55	0	0.05
I8	Multiplicative part of safety distance	80.68	0	0.05

Table 4 Results of Kruskal-Wallis test on the selected parameters

From Table 4 it can be inferred that the 3 most influential parameters are I1, I3 and I8, i.e. τ, Asoc_iso and multiplicative part of safety distance.

4.2 Conceptual validation

As mentioned in the previous sections, the addressed measure of performance is crossing time. Conceptual validation must confirm that the results of the models are consistent with the accepted concepts behind them. In this case, conceptual validation of the results requires verifying that they stand in appropriate ranges for the chosen infrastructure. Several papers found in the literature address pedestrian crossing time at both signalized and unsignalized crosswalks (Virkler, & Guell, 1984; Jain, Gupta, & Rastogi, 2014; Thompson, Rivara, Ayyagari, & Ebel, 2013). Manuals such as (HCM, 2016) and (National Joint Committee on Uniform Traffic Control Devices (U.S.). (1971)) provide guidance and equations for calculating crossing time at signalized intersections.

(HCM, 2016) gives an equation (Eq. 7) for calculating pedestrian crossing time at signalized crossings that takes into account the presence of a platoon with more than 15 people:

$$\text{(Eq. 7)} \quad t_c = t_s + \frac{L}{v_p} + \left(a * \frac{N_{ped}}{W}\right)$$

Where:
t_s is pedestrian start-up time, set to 3.2 s;
L is the length of the crosswalk;
v_p is pedestrian speed;
a is a parameter set to 2.7 and 0.27 respectively if the crosswalk width is larger of 3m (10ft) or not;
W is the width of the crosswalk. It is set equal to the real width, if it is more than 3 m, while it set equal to 1 if it is less than 3m.
N_{ped} is the number of pedestrians crossing in an interval, and it depends from the cycle and green durations.

Also, when dealing with unsignalized intersections, (HCM, 2016) provides a formula (Eq. 8) to calculated pedestrian critical gap, i.e. the time below that a pedestrian will not attempt to begin the action of crossing. The critical gap is defined as following:

$$\text{(Eq. 8)} \quad t_c = \frac{L}{v_p} + t_s$$

Where t_c stands for critical gap time, while the other symbols are the same previously mentioned.

(National Joint Committee on Uniform Traffic Control Devices (U.S.). (1971)) points out that the crossing time - considered as the sum of the so-called walk interval and the pedestrian clearance time - should be calculated as the time it takes a pedestrian to travel the length crossed at a speed of 0.91 m/s (3 ft/s).

Various authors (Virkler et al., 1984; Jain et al., 2014; Malinovskiy Wu, & Wang, 2008) also dealt with the measurement of crossing times at different locations.

(Virkler et al., 1984) studied 6 locations, 4 of which were signalized ones in the downtown, while two were unsignalized intersections near a stadium and traffic was controlled by a police officer. They found crossing time values ranging from 5.9 to 9.8 s at the signalized crossings, and values around 15 s for the unsignalized ones. (Jain et al., 2014) analyzed pedestrian crossing behavior, highlighting the influence of various factors on it. They considered uncontrolled intersections where pedestrians can cross in one or two steps modifying in this way their crossing time. They found a range of crossing time from 5 to 9 s for one-step crossings and from 3 to 12 s for two-step crossings. (Thompson et al., 2013) focused their work on the distraction rate caused by the use of technological devices and also reported the crossing times. They found a mean crossing time of 10.4 s. (Malinovskiy et al., 2008) also developed an automatic approach to capture pedestrian parameters such as waiting time, crossing time,

Authors	Crossing typology	Crossing time range [s]
Present findings	Unsignalized, on roundabout entry leg	6.318 (predicted - NN); 6.407 (simulated - Vissim) 4.0-14.0 (mean 8.27) (measured on field)
(HCM, 2016)	Unsignalized crossings (time gap – calculated for the given geometry)	11.6
(National Joint Committee on Uniform Traffic Control Devices (U.S.). (1971))	Signalized crossings	11.26
(Jain et al., 2014)	Uncontrolled crossing – one step	5.0-9.0
(Jain et al., 2014)	Uncontrolled crossing – two steps	3.0-12.0
(Virkler et al., 1984)	Signalized crossing	9.0-15.6
(Thompson et al., 2013)	Various intersections	10.4 (average value)
(Malinovskiy et al., 2008)	Signalized intersections	4.5-9.2

Table 5 Ranges of crossing time found in literature

and arrival rate, and they provide some ranges for these quantities. As for the crossing time, values ranging from 4.5 to 9.2 s were given.

All listed values are shown in Table 5, where the comparison between the mean crossing time values provided by the simulation and prediction models and the crossing time ranges on intersections found in the literature is provided.

It should be noted that the ranges are quite variable, depending on both the nature of the intersection, its geometric properties, and the properties of the action itself (1- or 2-step). Anyway, the simulated and predicted values belong to the ranges of crossing time reported in the literature and based on real measurements. Considering the measurements done at the chosen location, it can be seen that though the mean value of the measured sample is higher than the simulated and predicted crossing times, both of them stands in the range of the measurements. At this point there is a difference of 1.86 s and 1.95 s between simulated and measured values and predicted and measured crossing time respectively. Nevertheless, it should be considered that these values are not related to the combination of input parameters best matching the on-field measurements.For seek of completeness, Table 6 summarizes the descriptive statistics of the crossing time sample measured on field.

	Measured crossing time
Mean	8.27
Standard error	0.13
Median	8.0
Standard deviation	1.54
Asimmetry	0.73
Min	4
Max	14
Confidence level	0.26

Table 6 Statistics of the on-field measured crossing times [s]

These findings give good expectations for further steps of the research. In fact, the final goal will be that the simulation and prediction results obtained with the best combination of input parameters agree with the real conditions.

4.3 Operational validation

The operational validation consists of statistical testing of the model results. For this purpose, tests for two variances were applied. This type of test generally compares independent samples or paired samples by calculating the difference between the absolute deviations and the group (Derrick, Ruck, Toher, & White, 2018). Examples of this test typology include the F-test, which is used to evaluate two independent samples and assumes normality of the data as a precondition; the Pitman-Morgan test is more commonly used to evaluate paired samples under normality conditions. When the normality precondition is violated, these tests lack in robustness. Moreover, they cannot be applied when both paired and independent samples are present (Derrick et al., 2018). Another set of two-variance tests is based on the deviations from the median. This category includes the Brown and Forsythe test-better known as Levene's test since it is a modification of Levene's original calculation- (Derrick et al., 2018) and the Bonnett test (Bonnett, 2006). These tests are robust and powerful even when the data are not normally distributed and can be used when the samples are paired and independent, while they lose their power when the samples are either paired only or independent only (Derrick et al., 2018).

Due to the non-normal distribution of the data, the F-test and Pitman-Morgan -test were excluded and both Levene's test and Bonett's test were performed. This decision was based on the generally higher reliability of the Bonett test, but also on the trustworthy results of the Levene test when applied to highly skewed and tailed distributions (Kitchen, 2009; Marusteri, & Bacarea 2010; McDonald, 2014). The first step in conducting the tests was to calculate the variance, standard deviation, and confidence interval. The significance level for both tests was set at 5%. The ratio of the standard deviations was found to be 1.021 and that of the variances was found to be 1.043. By applying the two recalled tests, the results summarised in Table 7 were obtained.

Test	Confidence interval for StDev ratio	Confidence interval for variance ratio	P-Value
Bonett	(0.845; 1.228)	(0.715; 1.509)	0.822
Levene	(0.760; 1.402)	(0.578; 1.966)	0.846

Table 7 Results of Bonett's and Levene's tests

Both the Bonett and Levene tests are based on calculating the p-value and accepting or rejecting the null hypothesis, which states that the ratio of standard deviations or variances is not statistically significant. If the p-value is lower than the significance level, the null hypothesis must be rejected; otherwise, it can be accepted. As can be seen from Table 7, in both cases the p-value exceeds the set significance level (5%), therefore it can be judged that the two ratios are not statistically significant, which confirms the good fit of the result predicted by the neural network to the crossing time simulated by Vissim.

5. DISCUSSION

All three steps of the validation process yielded very valuable results. Visual validation made it possible to determine acceptable and unacceptable results and, based on descriptive statistics, to hypothesize further necessary analyzes to be developed on the data samples. Moreover, optimal correlation indices (97% for the training data sample and 94% for the generalization sample) were obtained. The conceptual validation showed the agreement between the results of the models and the ranges measured in the literature and confirmed that the simulated and predicted crossing time can effectively reproduce the real observed pedestrian behavior. The results of the statistical analysis not only confirm the high degree of correlation between the two data sets, but also show a very high degree of overlap between the two modeling results. Assuming 95% confidence interval, it can be outlined that the ratio between the two variances is very similar to 1, namely 1.021, which confirms the good fit of the compared databases. This preliminary evaluation is very important as it confirms the appropriate selection of the neural network, its input parameters and the correctness and accuracy of the training and validation steps.

6. CONCLUSIONS

In this paper, a three-step validation of a pedestrian crossing time prediction model is performed, which consists in a visual, conceptual and operational validation. The presented prediction model is based on the application of neural networks to reproduce a selected output of Vissim/Viswalk micro-simulation model, i.e. crossing time. The selection of input parameters for the prediction model is also briefly described. A ward neural network was formulated, trained and validated on a large data set consisting of random combinations of these parameters. Visual validation, consisting of comparison among the crossing times simulated by Vissim/Viswalk and the ones predicted by the ward net, showed a very high level of correlation - 97% on the training dataset and 94% on the validation dataset. The conceptual validation demonstrated the agreement of the simulated and predicted data with the real-world measurements and the ranges of crossing time found in the literature. Finally, the operational validation was elaborated, consisting in the development of statistical tests based on the comparison of the variances and standard deviations of the two populations. The evaluations carried out confirmed the well-fitting of the predicted results with the simulated ones, by showing a non-significant statistical difference between the two data sets. Thus, it can be concluded that a reliable prediction model has been formulated that well reproduces pedestrian crossing time as modeled by Vissim/Viswalk. Moreover, the obtained values are comparable to measurements in the field and therefore suitable for real-world applications.

All the summarized results are very influential for the development of the predictive model and for its further applications. Indeed, the formulated neural network will be applied in the next steps in the process of calibration of the micro-simulation model implemented in Vissim/Viswalk, and a good fit of these preliminary results is essential for the evaluation of the whole methodology to be developed.

REFERENCES

Advanced Neural Network and Genetic Algorithm Software. (n.d.). Retrieved February 13, 2020 from http://www.wardsystems.com/neuroshell2.asp.

Ambros, J., Novak, J., Borsos, A., Hoz, E., Kiec, M., Machcinik, Š., & Ondrejka, R. (2016). Central European Comparative Study on traffic Safety on Roundabouts. *Transportation Research Procedia*, 14, 4200-4208.

Bak, R., & Kiec, M. (2012). Influence of Midblock Pedestrian Crossings on Urban Street Capacity. *Transportation Research Record: Journal of the Transportation Research Board*, 2316, 76-83.

Blue, V.J., & Adler, J.L. (2001). Cellular automata micro-simulation for modeling bi-directional pedestrian walkways. *Transportation Research Bart B: Methodological*, 35(3), 293-312. doi:10.1016/S0191-2615(99)00052-1.

Bonnett, D.G. (2006). Robust confidence interval for a ration of standard deviations. *Applied Psychological Measurement*, 30(5), 432-439. doi:10.1177/0146621605279551.

Brilon, W., & Miltner, T. (2005). Capacity at Intersections Without Traffic Signals. *Transportation Research Record: Journal of the Transportation Research Board*, 1920, 32-40.

Chen, X., Shao, C., & Hao, Y., Influence of Pedestrian Traffic on Capacity of Right Turning Movements at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, 2073, 114-124.

Chen, Y., Persaud, B., Sacchi, E., & Bassani, M. (2013). Investigation of models for relating roundabout safety to predicted speed. *Accident Analysis and Prevention*, 50, 196-203.

Derrick, B., Ruck, A., Toher, D.; & White, P. (2018). Tests for equality of variances between two samples which contain both paired observations and independent observations. *Journal of Applied Quantitative Methods*, 13(2), 36-47. Retrieved from http://jaqm.ro/issues/volume-13,issue-2/pdfs/3_BE_AN_DE_PA_.pdf.

Gipps, P.G., & Marksjö, B. (1985). A micro-simulation model for pedestrian flows. Mathematics and Computers in *Simulation*, 27(2–3), 95-105. doi:10.1016/0378-4754(85)90027-8.

Gorrini, A., Vizzari, G., & Bandini, S. (2016). Age and Group-driven Pedestrian Behavior: from Observations to Simulations. Collective Dynamics, 1 (A3), 1-16.

Gruden, C., Otković, I.I., & Šraml, M. (2020). Neural Networks Applied to Micro-simulation: A Prediction Model for Pedestrian Crossing Time. Sustainability, 12(13), 5355. doi: 10.3390/su12135355.

HCM 2016. *Highway Capacity Manual*. (2016). Washington, D.C.: Transportation Research Board.

Helbing, D. (1998a). A mathematical Model for the behavior of pedestrians. Retrieved from http://www.theo2.physik.uni-stuttgart.de/helbing.html.

Helbing, D. (1998b). Models for Pedestrian Behavior. Retrieved from http://www.theo2.physik.uni-stuttgart.de/helbing.html.

Hongfei,J., Lili, Y., & Tang, M. (2009). Pedestrian Flow Characteristics Analysis and Model Parameter Calibration in Comprehensive Transport Terminal. *Journal of Transportation Systems Engineering and Information Technology*, 9(5), 117-123. doi:10.1016/ S1570-6672(08)60082-3.

Jain, A., Gupta, A., & Rastogi, R. (2014). Pedestrian crossing behavior analysis at intersections. *International Journal for Traffic and Transport Engineering*, 4(1), 103-116. doi:10.7708/ijtte.2014.4(1).08.

Johansson, A., Helbing, D., & Shukla, P.K. (2008). Specification of a Microscopic Pedestrian Model by Evolutionary Adjustment to Video Tracking Data. *Advances in Complex Systems*. Retrieved from https://arxiv.org/pdf/0810.4587.pdf.

Kitchen, C.M.R. (200). Nonparametric vs Parametric Tests of Location in Biomedical Research. *American Journal of Ophtalmology*, 147, 571-572. doi:10.1016/j.ajo.2008.06.031.

Kouskoulis, G:, Spyropoulou, I., & Antoniou, C. (2018). Pedestrian simulation: Theoretical models vs. Data driven techniques. *International Journal of Transportation Science and Technology*, 7, 241-253. doi: 10.1016/j.ijtst.2018.09.001.

Kretz, T., Hnegst, S., &Vortisch, P. (n.d.). Pedestrian flow at bottlenecks validation and calibration of Vissim Social Force Model of Pedestrian Traffic and its empirical foundations. Retrieved from https://arxiv.org/ftp/arxiv/papers/0805/0805.1788.pdf.

Kretz, T., Lohmiller, J. & Sukennik, P. (2018). Some Indications on How to Calibrate the Social Force Model of Pedestrian Dynamics. *Transportation Research Record: Journal of the Transportation Research Board*, 2672(20), 228-238. doi:10.1177/0361198118786641

Li, X., & Yeh, A.G.-O. (2001). Calibration of Cellular Automata by Using Neural Networks for the Simulation of Complex Urban Systems. *Environment and Planning A: Economy and Space*, 33(8), 1445-1462. Doi:10.1068/a33210.

Liao, W., Chraibi, M., Seyfried, A., Zhang, J., Zheng, X., & Zhao, Y. (2014). Validation of FDS+Evac for pedestrian simulations in wide bottlenecks. *17th IEEE International Conference on Intelligent Transportation Systems*, ITSC 2014, 554-559. doi:10.1109/ITSC.2014.6957748.

Liao, W., Zhang, J., Zheng, X., & Zhao, Y. (2017). A generalized validation procedure for pedestrian models. *Simulation Modelling Practice and Theory*, 77, 20-31. doi:10.1016/j.simpat.2017.05.002.

Malinovskiy, Y., Wu, Y.J., & Wang, Y. (2008). Video-Based Monitoring of Pedestrian Movements at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, 2073(1), 11-17. doi:10.3141/2073-02.

Marusteri, M., &Bacarea, V. (2010). Comparing groups for statistical differences: how to choose the right statistical test? *Biochemia Medica*, 20(1), 15-32. Retrieved from https://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=73802.

McDonald, J.H. (2014). *Handbook of biological statistics third edition*. Retrieved from http://udel.edu/~mcdonald/statpermissions.html.

National Joint Committee on Uniform Traffic Control Devices (U.S.). (1971). *Manual on uniform traffic control devices for streets and highways*. Federal Highway Administration.

Obsu, L.L., Meurer, A., Kassa, S.M., & Klar, A. (2016). Modelling pedestrians' impact on the performance of a roundabout. *Traffic flow modelling View project. Neural, Parallel and Scientific Computations*, 24, 317-334.

Otković, I. I., Tollazzi, T., & Šraml, M. (2013). Calibration of micro-simulation traffic model using neural network approach. *Expert Systems with Applications*, 40(15), 5965-5974. doi: 10.1016/j.eswa.2013.05.003.

Otković, I.I., Varevac, D., & Šraml, M. (2015). Analysis of neural network responses in calibration of micro-simulation traffic model. *The Electronic Journal of the Faculty of Civil Engineering Osijek E-GFOS*, 10, 67-76. Doi:10.13167/2015.10.8.

PTV. (2018). *PTV VISSIM 11 User manual*. Karlsruhe, Germany.

Rudloff, C., Matyus, T., Seer,S., &Bauer, D. (2011). Can Walking Behavior be Predicted? Analysis of Calibration and Fot of Pedestrian Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2264(1), 101-109. doi:10.3141/2264-12.

Shiwakoti, N., Sarvi, M., & Rose, G. (2008). Modelling pedestrian behavior under emergency conditions – State-of-the-art and future directions. *31st Australasian Transport Research Forum*, 457-473. Retrieved from https://www.australasiantransportresearchforum.org.au/sites/default/files/2008_Shiwakoti_Sarvi_Rose.pdf.

Thakur, S., & Biswas, S. (2019). Assessment of pedestrian-vehicle interaction on urban roads: a critical review. *Archives of Transport*, 51 (3), 50-63.

Teknomo, K., Takeyama, Y., & Inamura, H. (2000). Review on Microscopic Pedestrian Simulation Model. *Proceedings Japan Society of Civil Engineering Conference*. Retrieved from https://arxiv.org/abs/1609.01808.

Thompson, L.L., Rivara, F.P., Ayyagari, R.C., & Ebel, B.E. (2013). Impact of social and technological distraction on pedestrian crossing behavior: an observational study. *Injury Prevention*, 19(4), 232-237. doi: 10.1136/injuryprev-2012-040601.

Toledo, T., Koutsopoulos, H. N., Davol, A., Ben-Akiva, M. E., Burghout, W., Andréasson, I., & Lundin, C. (2003). Calibration and Validation of Microscopic Traffic Simulation Tools: Stockholm Case Study. *Transportation Research Record: Journal of the Transportation Research Board*, 1831(1), 65-75. doi: 10.3141/1831-08.

Virkler, M. R., & Guell, D. L. (1984). Pedestrian Crossing-Time Requirements at Intersections. *Pedestrian and Bicycle facilities*, 47-51. Retrieved from http://onlinepubs.trb.org/Onlinepubs/trr/1984/959/959-007.pdf.

Zeng, W., Chen, P., Yu, G., & Wang, Y. (2017). Specification and calibration of a microscopic model for pedestrian dynamic simulation at signalized intersections: a hybrid approach. *Transport Research Part C: Emerging Technologies*, 80, 37-70. doi: 10.1016/j.trc.2017.04.009.

Zhong, J., Hu, N., Cai, W., Lees, M., & Luo, L. (2015). Density-based evolutionary framework for crowd model calibration. *Journal of Computational Science*, 6, 11-22. doi: 10.1016/j.jocs.2014.09.002.