Automated Quality Assessment of Space-Continuous Models for Pedestrian Dynamics

Valentina Kurtc, Mohcine Chraibi and Antoine Tordeux

Abstract printed version.

In this work we propose a methodology for assessment of pedestrian models continuous in space. With respect to the Kolmogorov-Smirnov distance between two data clouds, representing for instance simulated and the corresponding empirical data, we calculate an evaluation factor between zero and one. Based on the value of the herein developed factor, we make a statement about the goodness of the model under evaluation. Moreover this process can be repeated in an automatic way in order to maximize the above mentioned factor and hence determine the optimal set of model parameters.

1 Introduction

Mathematical models have been developed to describe and simulate the dynamics of complex systems. For pedestrian dynamics several kinds of models have been built with different intentions. For an overview of the existing models the reader is referred to [20]. In order to develop and use mathematical models for pedestrian dynamics with high fidelity level, the verification and validation process should be considered as key part in the development cycle.

According to conventional terminology [19] **verification** is the process of assuring if the computer programming and implementation of a mathematical model is correct. The verification process does not assess the quality of mathematical models

Valentina Kurtc
Peter the Great St. Petersburg Polytechnic University, Polytechnicheskaya, 29, St. Petersburg, Russia, e-mail: kurtsvv@gmail.com

Mohcine Chraibi
Forschungszentrum Jülich e-mail: m.chraibi@fz-juelich.de

Antoine Tordeux
University of Wuppertal, e-mail: tordeux@uni-wuppertal.de

nor does it allow to make any conclusions related to its realism. In the literature some verification-driven works regarding the goodness-of-fit of a pedestrian model have been published. In [16] several verification tests were recommended to assess the quality of simulations of ships. In the RiMEA project [17] verification tests were proposed to evaluate the quality of pedestrian simulations. Based on the abovementioned works Ronchi et al. [18] suggested some modified and new verification tests with the aim to additionally consider the movement of pedestrians in case of fire.

Having verified the numerical implementation of the given model, the question that should be of concern is "does the implemented model show sufficient accuracy in emulating the targeted system?" In order to answer this question reliably, an iterative **validation** process is required, that is, comparing the model/simulation results with empirical findings and evaluating their discrepancies to enhance the behavior of the model. Repeating this process leads to a better description of the targeted system (pedestrian dynamics). The process of validation in pedestrian dynamics can benefit from the tremendous development of empirical research. Several experiments were conducted to investigate crowd performances and to build a well-documented database. Empirical research does not only contain data of controlled experiments [9, 12, 13, 14, 7, 15, 2, 1, 8, 5], but also data issued from field studies [4, 3, 21]. These empirical works capture relevant properties of the crowd in *normal* situations and for basic geometries. Examples are unidirectional flows (fundamental diagram), bottleneck (jam formation), counter-flow (lane formation), intersection (e.g. T-junction).

The goal of this paper is to develop an automatic assessment of the predictive capability of pedestrian models. We introduce new quantities, in order to quantify the degree of success (or failure) of a model/simulation with respect to a given set of empirical data. The introduced "validity factor" is a measure for the verification *and* the validation process of a given implementation of a model. In the first stage of the model's assessment (verification) we use tests mostly based on the RiMEA guideline [17]. While in the second stage (validation) we compare the model with experimental data issued from six different experiments: Uni-directional flow (1D and 2D), bi-directional flow, corners, bottlenecks and T-junctions. Comparing the model results with empirical findings and automatically evaluating their discrepancies is important for the validation process as well as for making a precise estimation of its goodness. 1

2 Verification and Validation Tests

We implement two different kinds of tests: verification and validation tests. In the first stage of the pedestrian model assessment, the verification step, we use tests based on the RiMEA guideline [17]. During the validation process we propose to compare the simulation results with experimental data issued from six different experiments: Uni-directional flow (1D and 2D), bi-directional flow, corners, bottle-

necks and T-junctions (Fig. 1). The proposed geometries are different in order to

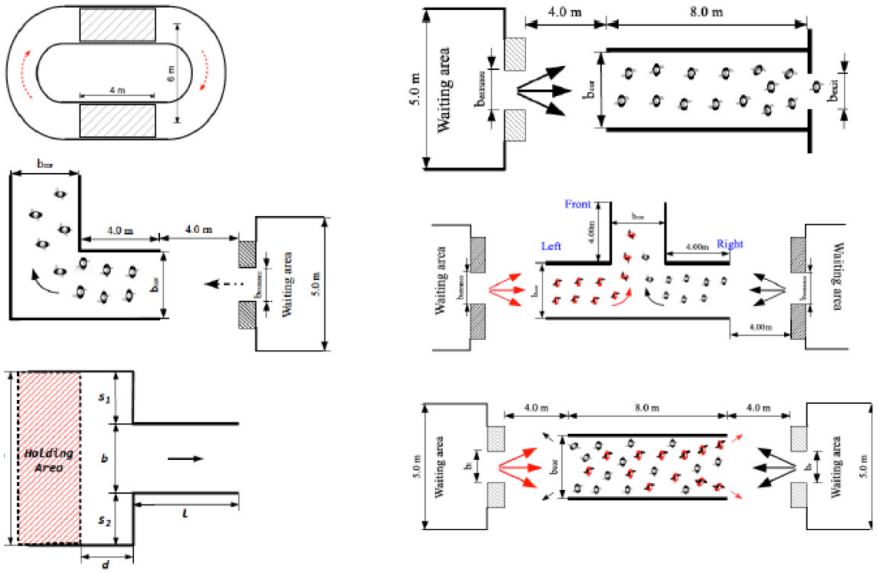

Fig. 1 Validation tests: six different experimental set-ups suggested to validate a model.

cover various dynamics of pedestrians and hence, the validation results may show more significance with respect to empirical findings.

Considering the pedestrian model and specifying area geometry and initial conditions, which are speed and position of pedestrians, one can compute trajectories of all pedestrians. The simulations are performed with JuPedSim [11, 10], using models [23, 22].

Afterwards, both for simulated and empirical trajectories the fundamental diagram is calculated using one of the measurement methods implemented in the JPSreport module of JuPedSim. For an accurate comparison it is important to use the same measurement method for both the experimental as well as the simulation data.

3 Methodology

This section presents the methodology for quantitative comparison of results obtained from experiments and simulations on basis of the verification and validation tests introduced earlier. Firstly, we suggest the approach for similarity assessment of two fundamental diagrams (speed-density relations), which are the main outcome of the validation tests. This approach exploits the data-binning and cumulative distribution functions calculation. Secondly, the quantitative quality assessment metric

is formulated, which aggregates the comparison results of experiments and simulations both for verification and validation tests.

3.1 Comparison of two data clouds

Assessment of validation results is made by means of comparison of two fundamental diagrams (FD), i.e. the speed-density relation. Given observation points of speed and density from experiments $\{[\rho_i^{\text{data}}, v_i^{\text{data}}]\}$ and simulations $\{[\rho_i^{\text{model}}, v_i^{\text{model}}]\}$, our goal is to quantify the degree of similarity among these two point clouds. Therefore, we create a partitioning of the data points by filtering them according to N equally-spaced density intervals (Fig. 2),

$$V_j^{\text{src}} = \{v_i^{\text{src}} : \rho_i^{\text{src}} \in [\rho_j, \rho_{j+1}], i = 1, ..., N^{\text{src}}\}, \tag{1}$$

where $j = 1, ..., N, \text{src} = \{\text{data, model}\}$ and N^{src} is the number of observations. The key idea here is to interpret V_j^{src} as a one-dimensional random variable. Now we calculate the cumulative distribution functions (CDFs) both for the experiment, $F_{V_j^{\text{data}}}(x)$ as well as for the simulations, $F_{V_j^{\text{model}}}(x)$, and calculate the Kolmogorov-Smirnov distance

$$D_j = \sup_x |F_{V_j^{\text{data}}}(x) - F_{V_j^{\text{model}}}(x)|, j = 1, ..., N. \tag{2}$$

Finally, the weighted arithmetic mean of distances D_j is used as a quantitative

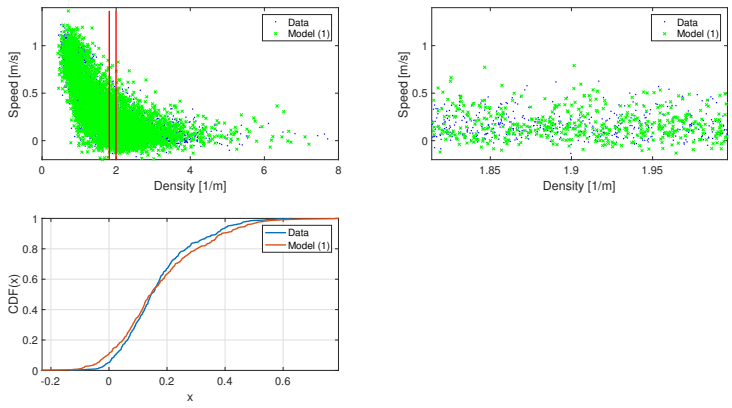

Fig. 2 Comparison of simulated (green) and empirical (blue) speed-density relations with data-binning method (upper plots). Cumulative distribution functions for specific bin (bottom plot).

metric estimating the similarity of two data-clouds

$$D^* = \frac{\sum_{j=1}^{N}(N_j^{\text{data}} + N_j^{\text{model}})D_j}{N^{\text{data}} + N^{\text{model}}}, \tag{3}$$

where N_j^{data} and N_j^{model} are the number of observations in the jth bin for data and model respectively. In other words, metric D^* quantifies the degree of success (or failure) of the validation process.

3.2 Validity factor

Verification tests are considered to be relatively simple and are expected to be fulfilled. For each verification test there are only two possible outputs which are 0 (failure) and 1 (success). The quantitative metric for N_{ver} verification tests is

$$\delta_{\text{ver}} = \prod_{k=1}^{N_{\text{ver}}} v_k^{\text{ver}}, \tag{4}$$

where v_k^{ver} is the outcome of the kth verification test.

We quantify the degree of success of the considered pedestrian model for the k-th validation test with $v_k^{\text{val}} \in [0,1]$. 0 and 1 corresponds to complete failure and absolute success respectively. The value v_k^{val} is computed according to the methodology described in Section 3.1, that is $v_k^{\text{val}} = 1 - D^*$. To quantify the degree of success for N_{val} validation tests we write

$$\delta_{\text{val}} = \frac{\sum_{k=1}^{N_{\text{val}}} w_k \cdot v_k^{\text{val}}}{N_{\text{val}}}, \tag{5}$$

with a weight w_k equal to 1.

However, in general the parameter values of the model are obtained after a calibration process involving several scenarios simultaneously [6]. In this case we calculate w_k as follows

$$w_k = \frac{N_{\text{cal}}}{N_{\text{val}}}, \tag{6}$$

where $N_{\text{cal}} \in [0, N_{\text{val}}]$ is a number of scenarios used simultaneously in the calibration procedure. Eq. (6) expresses the following assumption: it is less beneficial to get the value v_k^{val} by calibration obtained on the basis of one scenario, that is $w_k = \frac{1}{N_{\text{val}}}$, than if we get the same value v_k^{val} with the parameter set from calibration with all scenarios simultaneously ($w_k = 1$).

Finally the success estimator of the considered model on the basis of verification and validation tests is as follows

$$\delta = \delta_{\text{ver}} \times \delta_{\text{val}}. \tag{7}$$

4 Results

We tested this methodology on two continuous pedestrian speed models based on optimal velocity function and specific stochastic additive noises. The noise is white for the first model (Model (1)) [23], while it is determined by the inertial Ornstein-Uhlenbeck process for the second model (Model (2)) [22]. In this section we show only results of one of the validation tests – uni-directional 1D flow. In other validation test (except bottleneck scenario) the fundamental diagram is produced by the measurement module (Jpsreport) using Voronoi diagrams (method D). Figure 3 shows the results. The first row represents speed-density relations calculated on the basis of real trajectories and obtained from simulations of Model (1) and Model (2).

We consider 20 bins for the density partitioning of the validation test and limit the observations for densities not higher than $4\,m^{-2}$. For both models and each density bin the Kolmogorov-Smirnov distance is calculated (lower-left subplot of Fig. 3), the mean value of which (Eq. 3) is equal to 0.13 and 0.11 for Model (1) and Model (2) respectively.

Moreover, we investigated how the number of density bins N influences the final value D^*. Considering values of N from 20 to 470 we calculated corresponding values of D^* (bottom-right subplot of Fig. 3). Firstly, one can conclude that Model (2) is always better than Model (1), which is conform with the results in [22]. Secondly, for both models the value of D^* is increasing when the number of bins N is increasing. This can be explained by the fact that, with increasing number of bins, the number of points each bin contains, becomes more scarce. For example, if a bin contains only two points from each compared data set it is obvious that the Kolmogorov-Smirnov distance (Eq. 2) will be large.

5 Discussion and Conclusions

In this paper we present a methodology for quantitative assessment of pedestrian models exploiting results of verification and validation tests. However, this procedure can be applied to any space-continuous model describing pedestrian dynamics. Two stochastic pedestrian models were compared using the results of 1D unidirectional flow test. We investigated stability properties of the proposed metric (3) by considering different number of bins for density partitioning. According to the results, the metric for quantitative assessment shows a monotonic increasing behavior.

More generally, the method of quantitative comparison of two data-clouds (here, two fundamental diagrams) using an averaged Kolmogorov-Smirnov distance D^*

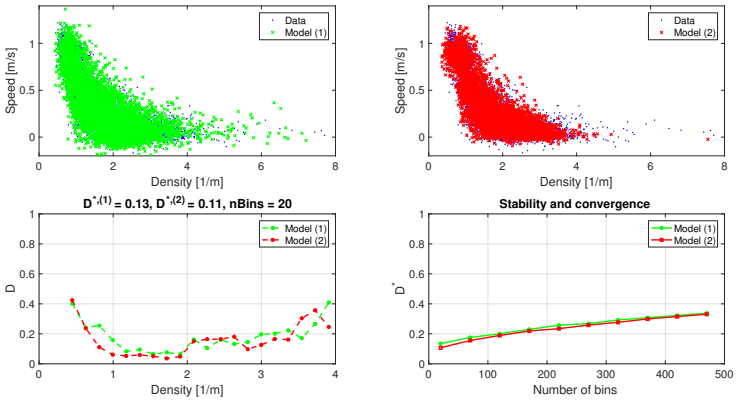

Fig. 3 Uni-directional 1D flow. Fundamental diagrams calculated on the basis of real trajectories and trajectories produced with Model (1) (upper-left) and Model (2) (upper-right). The Kolmogorov-Smirnov distance (2) for each density bin (lower-left). The dependence of the value D^* (3) on the number of density bins N.

(Section 4) can be interpreted as a separate result. In other words, this method can be used for comparison of two data-clouds of any nature.

The approach for comparison of two data-clouds allows a detailed analysis for different density ranges (lower densities, higher densities) by means of Kolmogorov-Smirnov distances per density interval D_j, $j = 1,...,N$. Analyzing values of D_j, it is possible to determine for which density ranges the underlying model performs better wrt. to experimental data.

Finally, the proposed metric (3) can be used in calibration procedure as the minimized objective function. We suggest to use the introduced quantity D^* (3) as a goodness of fit function in calibration procedure.

References

1. Asja, J., Appert-Rolland, C., Lemercier, S., Pettré, J.: Properties of pedestrians walking in line: Fundamental diagrams. Physical Review E **E 85**(85), 9 (2012). DOI 10.1103/PhysRevE.85.036111. URL http://pre.aps.org/abstract/PRE/v85/i3/e036111
2. Boltes, M., Zhang, J., Seyfried, A., Steffen, B.: T-junction: Experiments, trajectory collection, and analysis. In: IEEE Workshop on Modeling, Simulation and Visual Analysis of Large Crowds, 13. International Conference on Computer Vision (ICCV) (2011)
3. Burghardt, S.: Dynamik von personenströmen in sportstadien. Dissertation, Bergische Universität Wuppertal (2013)
4. Burghardt, S., Seyfried, A., Klingsch, W.: Improving egress design through measurement and correct interpretation of the fundamental diagram for stairs. In: M. Panda, U. Chattaraj (eds.) Developments in Road Transportation, pp. 181–187. Macmillan Publishers India Ltd (2010)
5. Burghardt, S., Seyfried, A., Klingsch, W.: Fundamental diagram of stairs: Critical review and topographical measurements. In: U. Weidmann, U. Kirsch, M. Schreckenberg (eds.) Pedes-

trian and Evacuation Dynamics 2012, pp. 329–344. Springer International Publishing (2014). DOI 10.1007/978-3-319-02447-9_27. URL http://dx.doi.org/10.1007/978-3-319-02447-9_27
6. Campanella, M., Hoogendoorn, S., Daamen, W.: Quantitative and Qualitative Validation Procedure for General Use of Pedestrian Models, pp. 891–905. Springer International Publishing, Cham (2014). DOI 10.1007/978-3-319-02447-9_74. URL https://doi.org/10.1007/978-3-319-02447-9_74
7. Chattaraj, U., Seyfried, A., Chakroborty, P.: Comparison of pedestrian fundamental diagram across cultures. Advances in Complex Systems **12**(3), 393–405 (2009). DOI 10.1142/S0219525909002209
8. Holl, S., Seyfried, A.: Laboratory experiments on crowd dynamics. inSiDE (Innovatives Supercomputing in Deutschland) **11**(2) (2013)
9. Hoogendoorn, S.P., Daamen, W.: Pedestrian behavior at bottlenecks. Transport. Sci. **39**(2), 147–159 (2005). DOI 10.1287/trsc.1040.0102
10. Kemloh, U., Chraibi, M., Zhang, J.: Jupedsim: Jülich pedestrian simulator. http://jupedsim.org. DOI https://doi.org/10.5281/zenodo.592209
11. Kemloh Wagoum, A.U., Chraibi, M., Zhang, J.: Jupedsim: An open framework for simulating and analyzing the dynamics of pedestrians. In: 3rd Conference of Transportation Research Group of India (2015)
12. Kretz, T., Grünebohm, A., Kaufman, M., Mazur, F., Schreckenberg, M.: Experimental study of pedestrian counterflow in a corridor. J. Stat. Mech. **10**, P10,001 (2006). DOI 10.1088/1742-5468/2006/10/P10001
13. Kretz, T., Grünebohm, A., Schreckenberg, M.: Experimental study of pedestrian flow through a bottleneck. Journal of Statistical Mechanics: Theory and Experiment **10**, 10,014 (2006). DOI 10.1088/1742-5468/2006/10/P10014
14. Liu, X., Song, W., Zhang, J.: Extraction and quantitative analysis of microscopic evacuation characteristics based on digital image processing. Physica A: Statistical Mechanics and its Applications **388**(13), 2717–2726 (2009). DOI 10.1016/j.physa.2009.03.017
15. Moussaid, M., Garnier, S., Theraulaz, G., Helbing, D.: Collective information processing and pattern formation in swarms, flocks and crowds. Topics in Cognitive Science **1**(3), 469–497 (2009)
16. Organization, I.M.: Guidelines for evacuation analysis for new and existing passenger ships (2007)
17. Rimea-richtlinie für mikroskopische entfluchtungs-analysen (2007). URL www.rimea.de. Www.rimea.de
18. Ronchi, E., KuKuligowski, E.D., Reneck, P.A., Peacock, R.D., Nilsson, D.: The process of verification and validation of building fire evacuation models. Tech. rep., National Institute of Standards and Technology (2013)
19. Sargent, R.G.: Verification and validation of simulation models. Journal of Simulation **7**(1), 12–24 (2013). DOI 10.1057/jos.2012.20. URL https://doi.org/10.1057/jos.2012.20
20. Schadschneider, A., Klingsch, W., Kluepfel, H., Kretz, T., Rogsch, C., Seyfried, A.: Evacuation dynamics: Empirical results, modeling and applications. In: R.A. Meyers (ed.) Encyclopedia of complexity and system science, vol. 5, pp. 3142–3176. Springer, Berlin, Heidelberg (2009)
21. Seer, S., Rudloff, C., Matyus, T., Brändle, N.: Validating social force based models with comprehensive real world motion data. Transportation Research Procedia **2**(0), 724 – 732 (2014). DOI http://dx.doi.org/10.1016/j.trpro.2014.09.080. URL http://www.sciencedirect.com/science/article/pii/S2352146514001161. The Conference on Pedestrian and Evacuation Dynamics 2014 (PED 2014), 22-24 October 2014, Delft, The Netherlands
22. Tordeux, A., A., S.: White and relaxed noises in optimal velocity models for pedestrian flow with stop-and-go waves. Journal of Physics A Mathematical and Theoretical (2016). DOI 10.1088/1751-8113/49/18/185101
23. Tordeux, A., Chraibi, M., Seyfried, A.: Collision-Free Speed Model for Pedestrian Dynamics, pp. 225–232. Springer International Publishing, Cham (2016). DOI 10.1007/978-3-319-33482-0_29. URL http://dx.doi.org/10.1007/978-3-319-33482-0_29