Available online at www.sciencedirect.com
ScienceDirect
Transportation Research Procedia 2 (2014) 715 – 723
Transportation Research Procedia
www.elsevier.com/locate/procedia

The Conference on Pedestrian and Evacuation Dynamics 2014 (PED2014)

Verification and validation of evacuation models - methodology expansion proposition

Robert Lubaś* [a], Marcin Mycek [a], Jakub Porzycki [a], Jarosław Wąs [a]

[a] *AGH University of Science and Technology, Department of Applied Computer Science, 30 Mickiewicza Av, 30-059 Kraków, Poland*

Abstract

In November 2013, Technical Note 1822: *The Process of Verification and Validation of Building Fire Evacuation Models* was released by NIST. The note was intended to open discussion about evacuation modeling, rather than provide definitive guidelines. The aim of our paper is to add creative contribution to V&V topic.

We propose adding some qualitative tests and distinguishing a base set of tests for all kinds of models from extended set of tests for specialized models. Moreover, the inclusion of additional tests is suggested, as well as the inclusion of some improvements and extensions to the tests proposed in the note.

Peer-review under responsibility of Department of Transport & Planning Faculty of Civil Engineering and Geosciences Delft University of Technology

Keywords: verification; validation; pedestrian dynamics models; cellular automata; evacuation

1. Introduction

Issue of reliability is crucial in the area of crowd dynamics models and simulations, and thus the process of validation and verification is a matter of great interest for researchers and practitioners. Currently, microscopic as well macroscopic approaches (Bellomo and Dogbe (2008); Hughes (2000); Lubas et al. (2013)) are applied in crowd modeling. However, it should be stressed that a vast majority of models are based on microscopic approaches - which take into account behavior of individuals. All individual-based models (continuous and discrete), can be classified as agent-based models eg. Bandini et al. (2009); Wąs and Lubaś (2014); Heliovaara et al. (2012). Among continuous models, the Social Force method (Helbing and Molnár (1995)) is especially popular and developed by numerous teams (Heliovaara et al. (2012); Wei-Guo et al. (2006); Yang et al. (2014); Dietrich et al. (2014)). On the other hand discrete modeling paradigms are also applied, for instance Cellular Automata (CA) (Nishinari et al. (2005); Leng et al. (2014); Wąs and Lubaś (2013)). The reliability of CA models in terms of verification and validation was recently discussed in our paper Porzycki et al. (2014).

* Corresponding author. Tel.: +48 617-52-18
E-mail address: rlubas@agh.edu.pl

Peer-review under responsibility of Department of Transport & Planning Faculty of Civil Engineering and Geosciences Delft University of Technology
doi:10.1016/j.trpro.2014.09.079

Regardless of the type of applied/developed model, each developer must take into account crucial aspects of the model's reliability, namely validation and verification (V&V). The process of V&V includes appearance of qualitative phenomena such as: jamming, lanes formation, density waves or bottleneck oscillations etc., as well as an analysis of quantitative criteria like proper fundamental diagrams (speed/density or flow/density relationships). Developed models should also satisfy basic verification requirements: pedestrians ability to maintain assigned walking speed or avoiding boundaries etc. In recent years several documents related to validation and verification have been released: a guide by the International Maritime Organization MSC/Circ.1238 (IMO (2007)), a guide by the RiMEA project (Rimea (2009)) (available as a whole only in German), works by Rogsch et al. (2009, 2010), as well as several ISO standards.

It should be stressed that the mentioned above documents do not constitute comprehensive methodological guidance. In order to address this issue the National Institute of Standards and Technology (NIST) released Technical Note 1822: *The Process of Verification and Validation of Building Fire Evacuation Models* (Ronchi et al. (2013)). The note summarizes current knowledge of building evacuation models V&V. In the note a set of seventeen verification tests, recommendations for validation data-sets and suggestions for a procedure of model uncertainty analysis are proposed. The aim of our paper is to analyze this document in the context of simulation practice and to discuss some of its crucial aspects.

2. Tests review

We start this section with carefully reviewing each verification test. A set of these tests covers five major components of evacuation models (Gwynne et al. (2012)):

- pedestrian pre-evacuation times,
- movement and navigation,
- exit choice,
- route availability,
- flow constraints.

2.1. Verif.1.1 Pre-evacuation time distribution

The first proposed test is devoted to verifying pedestrian pre-evacuation times. It is designed to test model ability, to adjust the given distribution of pre-evacuation time for the simulated population.

2.2. Verif.2.1. Speed in corridor and Verif.2.2. Speed on stairs

The most extended part of the verification test is connected with movement and navigation. Both tests: *2.1. Speed in corridor* and *2.2. Speed on stairs*, verify if occupants in the simulation can maintain a designated speed ($1\frac{m}{s}$) in a 40 meter corridor and on 100 meter long stairs. Due to discretization errors some models do not allow an exact speed ($1\frac{m}{s}$) or an exact length, therefore we propose to check if results are in a given range rather than exact results. Moreover this test is inapplicable to macroscopic models, where one deals with *crowd density* (not with individuals).

2.3. Verif.2.3. Movement around a corner

Test 2.3 checks the model's ability to simulate scenario boundaries. For given geometries, pedestrians should navigate around a corner without penetrating walls. This test seems to be designed especially for the force based models, where boundary penetration is possible. In other types of models like classical CA models or the hydrodynamic approach, walls penetration is prohibited by core model assumptions. It should be considered whether such a test should be executed for all types of pedestrian dynamic simulation.

Two sample movement trajectories for geometry given in test 2.3, generated by the Social Distance model, with different parameters, are shown in Fig. 1. Currently, there is no detailed knowledge about rules and patterns of corner movement, however, we think that the requirements of this test can be extended to include qualitative verification of pedestrian behavior eg. preference for the fastest rather than shortest paths.

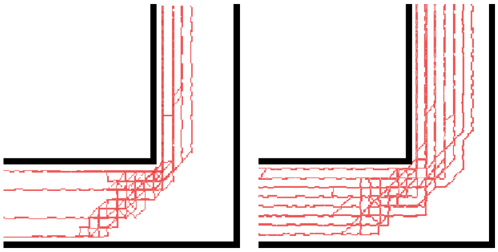

Fig. 1. Details of the geometry proposed in test 2.3. Space usage differs according to pedestrian preferences: shortest (a) or fastest path (b). Sample results captured by Social Distances model.

2.4. Verif.2.4. Assigned occupant demographics

This test, similarly to test 1.1, verifies the model's ability to assign given distribution of parameters to a whole population. In this case the tested parameter is free walking speed.

2.5. Verif.2.5 Reduced visibility vs. walking speed and Verif.2.6 Occupant incapacitation

Tests 2.5 and 2.6 are designed in order to test the influence of smoke on pedestrian movement. Such functionality is available in different types of models, however, many models focus only on pedestrian movement simulation. To ease verification of simple crowd dynamics model we suggest moving these tests to the *advance set* of tests.

2.6. Verif.2.7 Elevator usage

It is a simple verification if the model is able to transfer pedestrian from one floor to another using an elevator. This test is designed for specific evacuation scenarios, therefore similarly to tests 2.5 and 2.6 it also can be considered to be moved to the *advance set* of tests.

It is worth noting that the elevators working scheme changes significantly during evacuation caused by eg. fire. This test should be extended to verify whether a model is able to simulate the specific working schemes of elevators during some evacuation scenarios.

2.7. Verif.2.8. Horizontal counter-flows (rooms)

In terms of crowd dynamics simulation, one of the most important tests appears to be the verification of horizontal counter-flows. This test examines if a model is able to simulate counter-flow and its influence on pedestrian's speed. However, the only requirements in the note is that the passage time should decrease while the number of people in counter-flow increases.

Authors believe that in this case, the test requirements can be extended. Without any changes in test scenario and geometry one can add the following expected results:

- observation of line formation phenomena,
- detailed expectation about passage time (based eg. on Kretz et al. (2006)),
- lack of long term blocking.

2.8. Verif.2.9. Group Behaviours

Test 2.9 verifies if the model implements group behavior, in terms of adjusting group movement speed to its slowest member – the whole group is supposed to leave the test area in given range of time between the first and last group member.

We propose to extend group behavior testing to check group ability to keep together in situations of path choice. A detailed description of the test proposed is available in Sec. 4.2.

2.9. Verif.2.10. People with movement disabilities

The last test in the movement and navigation section, verifies if the model is able to simulate the influence of disabled people on the evacuation process. In this scenario occupants should move from one room to another, with and without occurrence of one disabled person.

As is written in the note, this test is designed only for models that allow agents with different dimensions - therefore it can also be moved to the *extended set* of tests. The expected result is that in first case pedestrians will leave the room faster than in the second. However, we believe that the expected results can be specified more precisely, and one can execute simple experiments to acquire empirical data about pedestrian behavior in such a situation.

2.10. Verif.3.1 Exit choice/usage

Deterministic assignment of exits is verified by test 3.1. Occupants are expected to choose the exit they are assigned to. Owing to the fact that many models calculate exit assignment (according to distance, density, travel speed, etc), this test should be used carefully - even a model with realistic selection/usage of exits may fail this test.

2.11. Verif.3.2 Social influence and Verif.3.3 Affiliation

Test 3.2 verifies a model's ability to simulate the impact of social influence on exit choice - two exits at the opposite sides of a room are available. It is expected that if one pedestrian chooses given exit, it will increase the usage of this exit by another pedestrian. One should notice, that in many models exit choice is separated from movement algorithm, therefore using the same exit may be uncorrelated with social influence.

The next test is executed in the same geometry, but this time a pedestrian is expected to choose one exit more often than another due to affiliation. We believe that such a feature is valuable, however, it is debatable if it should be included in the basic set of tests.

2.12. Verif.4.1 Dynamic availability of exits

This test verifies pedestrian ability to change exit choice as a response to a change in condition, namely in situations when one exit becomes unavailable. In this test it is expected that closed exit is not used by an occupant. One should note that this is the crucial test in terms of the usage of given model in a real-time, data-driven simulation.

It is also worth considering a more sophisticated test of the availability of dynamic exits, that will verify how information about exit availability spreads among a population - should pedestrians know immediately about the fact of closed doors, even few rooms away?

2.13. Verif.5.1. Congestion

The last section of the test focuses on flow constraints. Test 5.1. inspects if model simulates congestion on staircase example, for details of geometry see Fig. 2. Congestion is expected at the exit from the room and at the base of the stairs. In the note authors propose visual evaluation of this test, however, one can consider if a more impartial method of evaluation is possible. We propose to compare average density in consecutive parts of a room and corridor during simulation.

2.14. Verif.5.2. Maximum flow rates

One of model's crucial characteristics is the pedestrian flow it generate. Test 5.2 is designed in order to verify this parameter. Occupants are expected to leave a room using a 1 m wide exit. Flow should not exceed a given threshold (recommended: 1.33 person per meter per second).

We wonder why only maximum flow is tested. Our suggestion here is to verify obtained results using more sophisticated characteristics like fundamental diagrams.

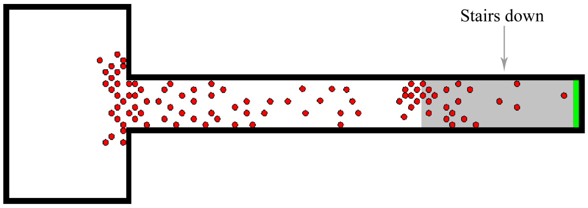

Fig. 2. A staircase congestion test. Sample results generated by the Social Distances model.

3. Basic and advanced sets of tests - proposal of a division

As has been outlined in Section 2, while we believe that all the proposed tests are reasonable, some of them could be omitted in the validation of models designed for specific use cases. We therefore propose a division of the tests into two groups:

- The basic set - applicable for all types of models, do not include tests for specific conditions or features
- Advanced set - tests that are dependent on:
  - model type (continuous/discrete, macroscopic/microscopic, etc.),
  - specific features or conditions (smoke spreading, reduced visibility, elevator usage etc.),
  - model applications (evacuation, comfort optimization etc.).

A model should always be able to pass *a basic set* of tests and to declare a set implemented features to be verified and validated along with type specific tests. In this way a model can be verified for most basic aspects with *a basic set* of tests and depending on the use case or model type with a subset of *the advanced set* of tests. In this way a model is not forced to have aspects not needed for a specific use case. A model aiming to constitute a complete building fire evacuation model, should be able to pass both *the basic set* and *the advanced set* of tests, while the tests themselves could be applied in an organized way to verify a variety of models not only fire evacuation ones.

It should be noted that such a division can be one step toward the unification of V&V methodology of pedestrian evacuation/movement for all fields of applications.

4. Test extensions and new tests proposal

The process of verification and validation (V&V) is a very important stage in the development and assessment of the reliability of pedestrian dynamics models. Verification can be easily achieved by employing numerous (sometimes "artificial") test cases, however, the validation process mainly relies on experimental/empirical data. Authors claims that the verification test should be adjusted to the experiments setup as much as possible.

4.1. Density/velocity relation

A test is suggested to measure velocity-density relation for a single-file movement. The idea was taken from Seyfried et al. (2007). Furthermore, the empirical data are available for validation. This verification and validation test is appropriate to verify, whether microscopic models are able to reproduce the empirical relation between velocity and density in simple geometry. The test method is a quantitative evaluation of model results, i.e. the comparison between the results produced by the model and the empirical data.

Geometry: The geometry of this test is presented in Fig. 3

Scenario: To achieve diverse densities, the scenarios with 15, 20, 25, 34 pedestrians can be executed. Pedestrians are uniformly distributed through the corridor. They should move around passageway three times and avoid passing other pedestrians.

Expected result: The expected result is that the measured individual velocity should decrease (linearly) with increasing density, see Fig. 4.

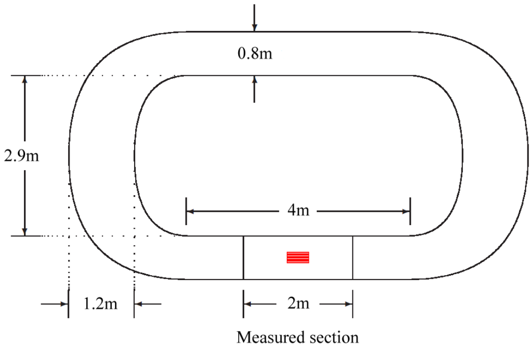

Fig. 3. The geometry of the proposed velocity-density relation test.

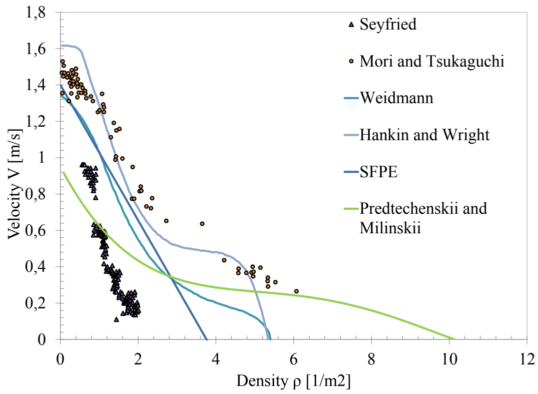

Fig. 4. Fundamental diagram of pedestrians movement on flat surfaces with empirical data by SFPE Handbook (Association and of Fire Protection Engineers, 2002), (Predtechenskii and Milinskii, 1978), (Weidmann, 1993), (Seyfried et al., 2007), (Mori and Tsukaguchi, 1987)

4.2. Group coherence

The following test is designed in order to verify the model's ability to maintain group coherence. In situation of more than one available solution, all groups of pedestrians usually chose the same path.

Geometry: Details of test geometry are presented in Fig. 5

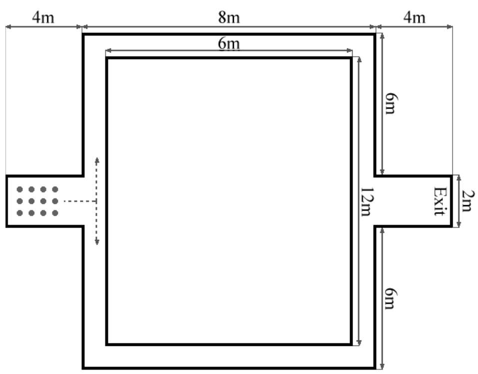

Fig. 5. Geometry layout of the proposed group coherence test.

Scenario: A group of 12 occupants are located in the beginning of the corridor, they should move to the exit. Pedestrian speed and pre-movement time should be assigned using the chosen distribution. Occupants do not have any affiliation to a specific path. There is no leader in the group.

Expected results: All members of group are expected to choose one path to exit. Moreover, similarly to test 2.9 they should reach the exit together in a given range of time between the first and last pedestrian.

4.3. Discretization errors

There is huge group of crowd dynamics models that uses space discretization. Such an approach has many advantages. However, among its drawback are discretization errors that can affect simulated crowd parameters. This test is designed for discrete method to verify how discretization affect pedestrian flow.

Geometry: A 6 m by 8 m room connected with a 10 m long corridor. The door to the corridor is located in the middle of one wall. In consecutive runs of the experiment the width of corridor should increase by 10 cm from 0.5 m to 1.5 m. Details of test geometry are presented in Fig.: 6

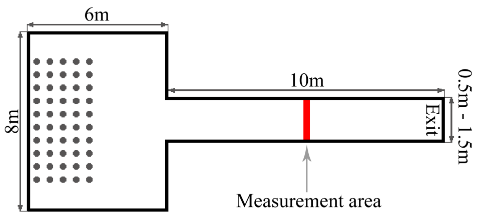

Fig. 6. Geometry layout of proposed discretization errors test.

Scenario: Populate the room with 50 occupants. Pedestrian speed should be assigned using the chosen distribution, and no pre-movement time. All occupants should leave the room and corridor through the exit at the end of corridor.

Expected results: Pedestrian flow in the middle of the corridor should increase linearly with the growth of corridor width, as has been shown in Seyfried et al. (2009). A discretization error can disturb this linear growth. We therefore propose a linear correlation coefficient between the flows in consecutive runs of experiments as a simple method to measure the influence of a discretization error on pedestrian flow.

4.4. Pedestrian dynamics phenomena

To assess qualitative verification several important phenomena need to be captured by the model, for example:

- density waves (stop-and-go waves) - these typically occur in a densely crowded corridor (close to the density that causes the total absence of movement). We can use a *Verif.2.1. Speed in a corridor* test with slight modifications. Uniformly distributed pedestrians with high density up to $7\frac{P}{m^2}$ moving in the same direction.
- oscillations - in counter-flow at bottlenecks, e.g. doors, one can sometimes observe oscillatory changes in the direction of motion. We should observe this on the occasion of *Verif.2.8. Horizontal counter-flows (rooms)* test.
- freezing-by-heating effect - at sufficiently high crowd density and under extreme conditions (panics) pedestrians lanes are disturbed by increasing fluctuation. This leads to the formation of blockage which may a have regular structure Helbing et al. (2000). The *Verif.2.8. Horizontal counter-flows (rooms)* test with modification, may be used to observe this effect, e.g. in a situation where there are 100 pedestrians in each room moving to the opposite room.
- jamming - usually occurs at high crowd density and in places where the capacity is reduced (bottleneck). This phenomenon can be observed in the *Verif.5.1. Congestion* test.
- lane formation - in counter-flow when two groups of pedestrians are moving in opposite direction, lanes are formed for those people who move in same direction. The number of lanes is not constant and changes over time. Some quantitative empirical data and experimental setup can be find in Kretz et al. (2006).

5. Summary

We believe, that Technical Note 1822, released by NIST, is a valuable initiative concerning the verification and validation of evacuation models. It is not an easy task to propose a unified guide, due to the fact of the existence of different evacuation models and methodologies. We analyze all proposed stages of validation and verification in the context of the application of different models of crowd dynamics and finally we propose some amendments.

Due to the fact that pedestrian dynamics models have diverse field of application and duration of validity, the authors claim that the set of verification tests should be divided into two independent subsets. One of them - *the basic set* of tests - verifies basic crowd dynamics properties. The second *advanced set* should cover the specific tests which depend on model type, application, etc.

We propose three additional tests that examine: the density/velocity relation, group coherence and discretization errors. It should be stressed that using empirical data for geometry and test scenarios is a key factor for the validation process. Moreover we propose for further discussion the fact that qualitative validation of emerging phenomena (density waves, oscillations, lane formation, jamming, freezing by heating) is possible on the base of the existing test.

The authors believe that the Verification and Validation process has a crucial role in crowd dynamics modeling. NIST's Technical Note 1822 is a worthwhile initiative to unify the question of crowd dynamics models testing. As the note was intended to open discussion of this topic, this paper is intended as a voice in that debate.

References

Association, N. F. P., of Fire Protection Engineers, S., 2002. SFPE handbook of fire protection engineering. National Fire Protection Association.

Bandini, S., Manzoni, S., Vizzari, G., 2009. Agent based modeling and simulation: an informatics perspective. Journal of Artificial Societies and Social Simulation 12 (4), 4.

Bellomo, N., Dogbe, C., 2008. On the modelling crowd dynamics from scaling to hyperbolic macroscopic models. Mathematical Models and Methods in Applied Sciences 18, 1317–1345.

Dietrich, F., Köster, G., Seitz, M., von Sivers, I., 2014. Bridging the gap: From cellular automata to differential equation models for pedestrian dynamics. Journal of Computational Science.

Gwynne, S., Kuligowski, E., Spearpoint, M., 2012. More thoughts on defaults. Presented at the Fifth International Symposium on Human Behaviour in Fire.

Helbing, D., Farkas, I. J., Vicsek, T., Feb 2000. Freezing by heating in a driven mesoscopic system. Phys. Rev. Lett. 84, 1240–1243.

Helbing, D., Molnár, P., May 1995. Social force model for pedestrian dynamics. Phys. Rev. E 51, 4282–4286.

Heliovaara, S., Korhonen, T., Hostikka, S., Ehtamo, H., 2012. Counterflow model for agent-based simulation of crowd dynamics. Building and Environment 48 (0), 89 – 100.

Hughes, R., 2000. The flow of large crowds of pedestrians. Mathematics and Computers in Simulation 53 (4-6), 367 – 370.

IMO, 2007. Guidelines for evacuation analysis for new and existing passenger ships. msc.1/circ.1238. Tech. rep., International Maritime Organization.

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

Leng, B., Wang, J., Zhao, W., Xiong, Z., 2014. An extended floor field model based on regular hexagonal cells for pedestrian simulation. Physica A: Statistical Mechanics and its Applications 402 (0), 119 – 133.

Lubas, R., Miller, J., Mycek, M., Porzycki, J., Was, J., 2013. Three different approaches in pedestrian dynamics modeling. case study. In: Zamojski, W., Mazurkiewicz, J., Sugier, J., Walkowiak, T., Kacprzyk, J. (Eds.), New Results in Dependability and Computer Systems. Vol. 224 of Advances in Intelligent Systems and Computing. Springer International Publishing, pp. 285–294.

Mori, M., Tsukaguchi, H., 1987. A new method for evaluation of level of service in pedestrian facilities. Transportation Research Part A: General 21 (3), 223 – 234.

Nishinari, K., Kirchner, A., Namazi, A., Schadschneider, A., 2005. Simulations of evacuation by an extended floor field ca model. In: Hoogendoorn, S. P., Luding, S., Bovy, P. H., Schreckenberg, M., Wolf, D. E. (Eds.), Traffic and Granular Flow 2003. Springer Berlin Heidelberg, pp. 405–410.

Porzycki, J., Lubaś, R., Mycek, M., Wąs, J., 2014. Application of nist technical note 1822 to ca crowd dynamics models verification and validation. Proceedings of ACRI 2014.

Predtechenskii, V. M., Milinskii, A. I., 1978. Planning for foot traffic flow in buildings. Amerind.

Rimea, 2009. Rimea, richtlinie für mikroskopische entfluchtungsanalysen, ver: 2.2.1, (2009). Tech. rep., Rimea.

Rogsch, C., Klingsch, W., Seyfried, A., Weigel, H., 2009. Prediction accuracy of evacuation times for high-rise buildings and simple geometries by using different software-tools. In: Appert-Rolland, C., Chevoir, F., Gondret, P., Lassarre, S., Lebacque, J.-P., Schreckenberg, M. (Eds.), Traffic and Granular Flow 2007. Springer Berlin Heidelberg, pp. 395–400.

Rogsch, C., Weigel, H., Klingsch, W., 2010. Hand-calculation methods for evacuation calculation last chance for an old-fashioned approach or a real alternative to microscopic simulation tools? In: Klingsch, W. W. F., Rogsch, C., Schadschneider, A., Schreckenberg, M. (Eds.), Pedestrian and Evacuation Dynamics 2008. Springer Berlin Heidelberg, pp. 523–528.

Ronchi, E., Kuligowski, E. D., Reneke, P. A., Peacock, R. D., Nilsson, D., 2013. Nist technical note 1822, the process of verification and validation of building fire evacuation models. Tech. rep., NIST.

Seyfried, A., Passon, O., Steffen, B., Boltes, M., Rupprecht, T., Klingsch, W., Aug. 2009. New insights into pedestrian flow through bottlenecks. Transportation Science 43 (3), 395–406.

Seyfried, A., Steffen, B., Klingsch, W., Lippert, T., Boltes, M., 2007. The fundamental diagram of pedestrian movement revisited - empirical results and modelling, 305–314.

Wąs, J., Lubaś, R., 2013. Adapting social distances model for mass evacuation simulation. Journal of Cellular Automata, 395–405.

Wąs, J., Lubaś, R., 2014. Towards realistic and effective agent-based models of crowd dynamics. Neurocomputing. URL `http://dx.doi.org/10.1016/j.neucom.2014.04.057`

Wei-Guo, S., Yan-Fei, Y., Bing-Hong, W., Wei-Cheng, F., 2006. Evacuation behaviors at exit in {CA} model with force essentials: A comparison with social force model. Physica A: Statistical Mechanics and its Applications 371 (2), 658 – 666.

Weidmann, U., March 1993. Transporttechnik der fussgänger - transporttechnische eigenschaften des fussgängerverkehrs (literaturstudie). Literature Research 90, Institut füer Verkehrsplanung, Transporttechnik, Strassen- und Eisenbahnbau IVT an der ETH Zürich, ETH-Hönggerberg, CH-8093 Zürich, in German.

Yang, X., Dong, H., Wang, Q., Chen, Y., Hu, X., 2014. Guided crowd dynamics via modified social force model. Physica A: Statistical Mechanics and its Applications 411 (0), 63 – 73.