Bridging Empirical Research and Standards in Pedestrian Dynamics: Towards Enhanced Verification and Validation

Mohcine Chraibi*[1], Alessandro Corbetta†[2], Claudio Feliciani[3], and Bryan Klein[4]

[1]Forschungszentrum Jülich, Germany
[2]Eindhoven University of Technology, Netherlands
[3]The University of Tokyo, Japan
[4]Thunderhead Engineering, USA

Abstract As our quantitative knowledge of pedestrian dynamics keeps growing thanks to more and more accurate measurements, so increase our capability of validating models. We introduce here a benchmarking framework. It hinges on one side on the latest experimental evidence, and on the other on a Continuous Integration (CI) pattern for automatic and independent simulation software testing during development and maintenance phases. Ultimately, the framework allows researchers, practitioners, and standardization organizations to work together to improve the reliability and applicability of pedestrian simulations in real-life scenarios.

Keywords Verification, Validation, Pedestrian dynamics, Evacuation Software, Standards

Introduction

Mathematical models, a cornerstone of the scientific method, are developed to investigate, explain, and possibly predict the properties of real-world systems. Pedestrian dynamics is no exception: models are developed to analyze crowd behavior at all scales. This ranges from microscopic-level individual locomotion and interactions to coarse-grained properties, e.g. the capacity of a train station to operate within safety limits. Reproducing with models emergent features measured in real world systems that have not been themselves explicitly included in the model underlies validation, and to probe for limitations. Ultimately, models are translated into software to provide crowd management managers with convenient tools to assess design scenarios and evaluate potential interventions.

Recent advances in empirical research on pedestrian and crowd dynamics unlock a wealth of new data that are describing the behavior of crowds with higher and higher levels of detail. These include characterization of stochastic behavior, universal features, dependency on type of crowd, surrounding geometry, and more [2, 5]. Naturally, these provide invaluable sources for significantly improving the way current models are designed and, more importantly, validated and verified. Existing standards, such as those described in [3, 1], continue to play an important role in providing foundational benchmarks (for example, ensuring that agents walking 10 meters at 1 m/s cover that distance in 10 seconds). However, they often predate recent experimental findings and may miss critical nuances of pedestrian flow. To address this gap, new test frameworks, supported by up-to-date empirical studies, are necessary to capture complex behaviors more faithfully.

This contribution critically discusses recent pedestrian dynamics experimental data and how these can be tailored into the next generation of references and benchmarks to validate the effectiveness of current and forthcoming crowd models. In particular, we will address the following:

- Model validation benchmarks; i.e. how well the model—and its software implementation—reflect (statistic-level) empirical observations of actual systems (cf. e.g., Fig. 1 (bcd)).
- Model verification tests; i.e., to what extent the implemented software correctly simulates the behaviors prescribed by the underlying mathematical model and whether those simulation outcomes reliably mirror the real-world system.

Although validation and verification are closely related, they have distinct goals. The former is about capturing physical behavior whereas the latter is about building trust around the consistency of the model

*Email of the corresponding author: m.chraibi@fz-juelich.de
†Email of the corresponding author: a.corbetta@tue.nl

implementation. In practical terms, this contribution will introduce an automated online testing suite, using the well-established Continuous Integration development paradigm to enable automated, online, and unbiased testing, standard for all.

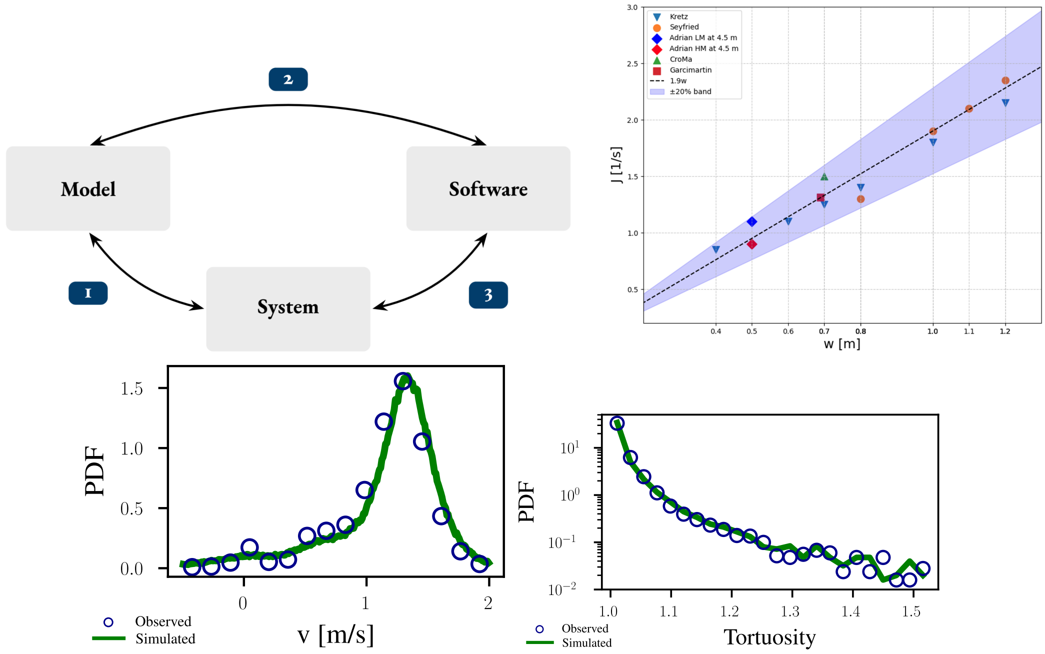

Figure 1: (a) Validation process involving a model, software, and real system. [1] and [3] validate the results of the model and its implementation against real-world observations, while [2] verifies that the model is implemented faithfully in the software. (b) Flow versus width of bottleneck for different experiments from the literature. (c,d) Examples of complex statistical features of pedestrian dynamics, here velocity and trajectory tortuosity. Figures reproduced from [4].

Figure 1(a) illustrates how verification and validation interact among the three main elements: the real-world system, the model, and the software. Steps [1] and [3] deal with validating outcomes against real-world data, whereas step [2] ensures that the software correctly implements the model's logic. Figure 1(b) illustrates an experiment that highlights the linear relationship between flow and the width of the bottleneck. Simulation results falling within the blue band are considered realistic. Figure 1(c,d) show complex statistics of pedestrians trajectories, such as velocity or tortuosity, that realistic simulations should be able to capture.

By backing each of these steps with state-of-the-art empirical insights and by carefully choosing which data to include, researchers, software developers, and standards bodies can collaboratively strengthen the reliability and realism of pedestrian models. Ultimately, this improves confidence in simulated outcomes, enabling better-informed decisions for crowd management and urban planning.

References

[1] ISO/TC 92/SC 4 20414:2020. *Fire Safety Engineering — Verification and Validation Protocol for Building Fire Evacuation Models.* Standard 20414. ISO, Nov. 2020, p. 69.

[2] A. Corbetta and F. Toschi. "Physics of human crowds". In: *Ann. Rev. Conden. Matter Phys.* 14.1 (2023), pp. 311–333.

[3] IMO. *Revised guidelines on evacuation analysis for new and existing passenger ships.* Tech. rep. 2016.

[4] Koen Minartz et al. "Discovering interaction mechanisms in crowds via deep generative surrogate experiments". In: *Scientific Reports, to appear. arXiv:2412.01491* (2025).

[5] A. Schadschneider et al. "Pedestrian Dynamics: From Empirical Results to Modeling". In: *Crowd Dynamics, Vol. 1.* Ed. by L. Gibelli. Modeling and Simulation in Science, Engineering and Technology. Springer, 2018, pp. 63–102.