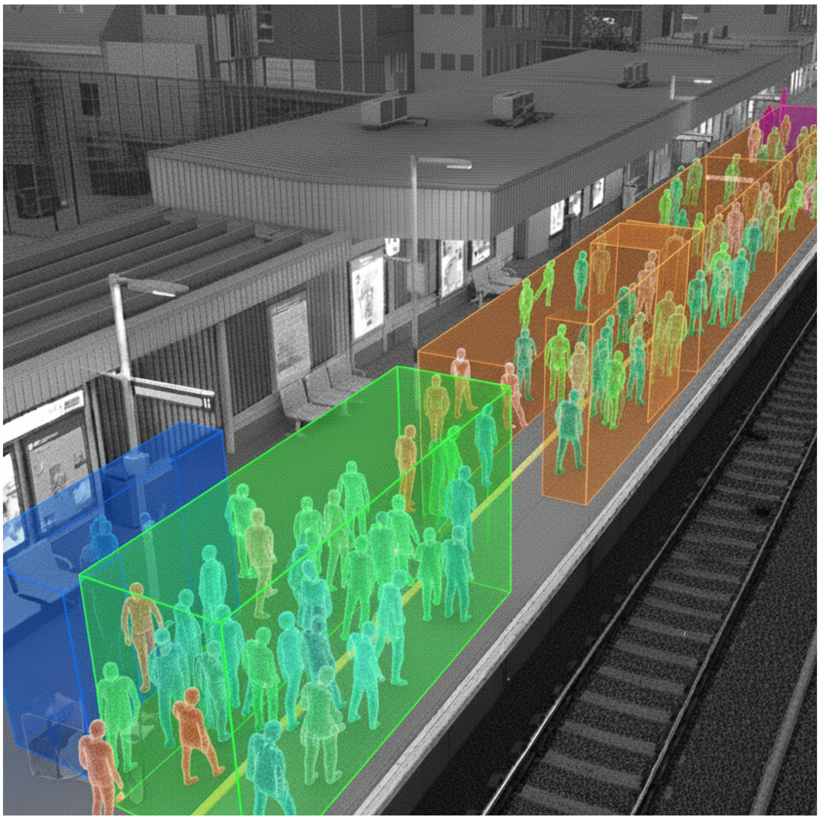
Msc Applied Ai

Write Up Draft

Oluwatoni Esan



Facilitating rail station dwell times improvement using machine Vision

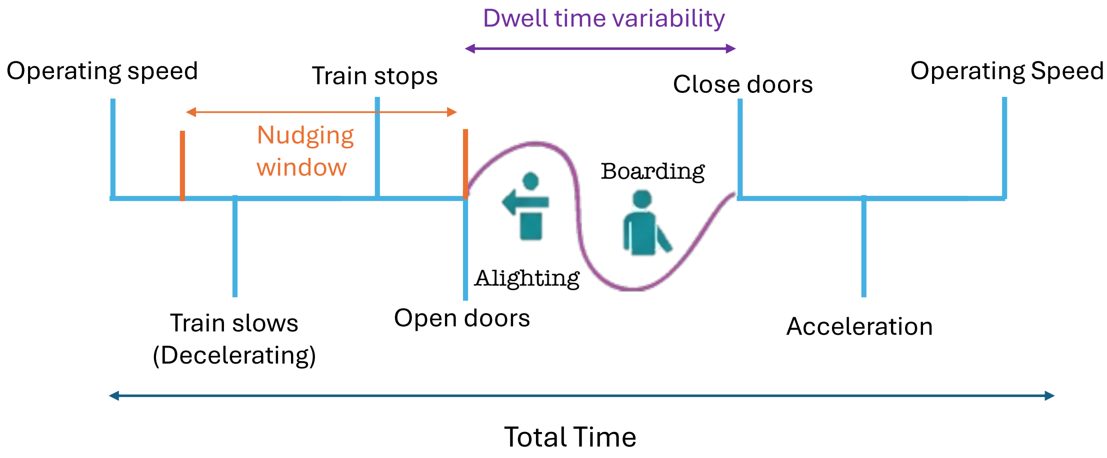
Prof Argyrios Zolotas

Prof Angelos Plastropoulos

1 Introduction

1. Introduction

The efficiency and punctuality of railway networks are fundamental to modern public transportation, directly impacting economic productivity and passenger satisfaction. A primary constraint on network capacity is train dwell time—the period a train remains stationary at a platform—which has been identified as a critical operational bottleneck. Minor delays at a single station can propagate through the network, leading to significant cascading disruptions and a reduction in overall service reliability (Vromans, 2005). While mechanical factors contribute to dwell time, research increasingly points to passenger behaviour as a key source of inefficiency, specifically the phenomenon of “concentrated boarding” (Gong, et al., 2018). Passengers instinctively cluster near primary platform access points such as stairs and lifts, a predictable heuristic that leads to a sub-optimal distribution across carriages (Fang, et al., 2019). This behaviour overwhelms a few train doors while leaving others underutilised, creating localised congestion that slows the boarding process and poses safety risks in crowded conditions.



To mitigate this, operators have shifted from costly infrastructure changes towards implementing behavioural interventions. Traditional methods include static signage, audio announcements, and painted platform markings designed to guide passenger distribution (Offiaeli and Yaman, 2021). While beneficial, the effectiveness of these static "nudges" is limited as they cannot adapt to real-time conditions. This has driven the development of dynamic, data-driven systems. Early iterations relied on historical train load data or camera-based computer vision systems to identify passenger hotspots (CIO, 2017). However, camera-based solutions raise significant privacy concerns. Consequently, LiDAR (Light Detection and Ranging) has emerged as a state-of-the-art, privacy-by-design sensor technology. By generating anonymised 3D point clouds of the platform, LiDAR enables the accurate, real-time tracking of crowd density and flow without capturing personally identifiable information. Its viability has been successfully demonstrated in trials at major UK stations, including Bristol Temple Meads (Createc, 2022; IMechE, 2023), establishing it as a feasible technology for live passenger guidance.

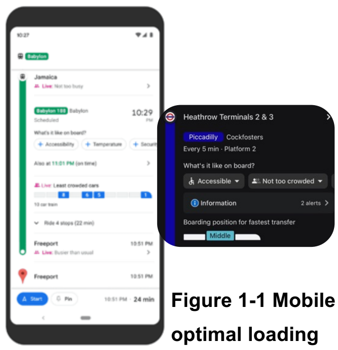
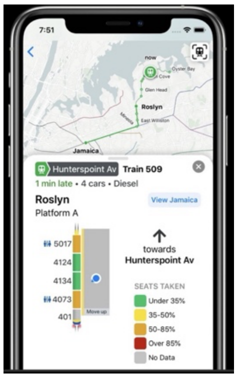


Figure 1-1 Mobile app optimal loading communication

Before such advanced nudging systems can be deployed, their complex dynamics and potential effectiveness must be rigorously evaluated. Agent-Based Modelling (ABM) is the standard methodology for this task, enabling the simulation of microscopic, heterogeneous passenger interactions that give rise to emergent crowd behaviour (Seriani and Fernandez, 2015). However, a critical gap persists in the existing literature. Previous ABM studies of passenger nudging have often suffered from three key limitations: (1) they are seldom grounded in real-world, real-time sensor data, relying instead on assumed or historical distributions; (2) they frequently model passengers as a homogenous crowd, failing to capture the diverse strategies and behaviours of different individuals; and (3) they have not yet quantified the complex, non-linear relationship between the level of passenger compliance with a nudge and its ultimate impact on boarding efficiency.



This paper addresses this research gap by presenting a novel Agent-Based Model that directly integrates these missing elements. Grounded in real-world LiDAR data from Bristol Temple Meads, the model introduces four distinct passenger archetypes, informed by human factors research, to simulate a heterogeneous crowd. We develop and test a 'Combined Nudging Algorithm' which, in a novel approach, fuses static, pre-existing train load data with dynamic, simulated platform crowding data to issue optimal boarding guidance. The primary contribution of this work is to systematically quantify the non-linear relationship between nudge effectiveness rates and total boarding time, revealing a paradoxical outcome where perfect compliance is sub-optimal.

The remainder of this paper is structured as follows. Section 2 details the methodology, including the ABM architecture, the development of passenger archetypes, and the design of the nudging algorithm. Section 3 presents the simulation results, analysing the impact of varying nudge effectiveness on key performance indicators. Section 4 provides a discussion of these findings and their implications for real-world railway operations. Finally, Section 5 offers concluding remarks and outlines directions for future research.

2. Methodology

To investigate the impact of data-driven nudging on passenger boarding dynamics, a computational model was developed using an Agent-Based Modelling (ABM) paradigm. This approach was chosen over macroscopic equation-based models as it is uniquely suited for capturing the emergent, macro-level crowd behaviours (e.g., congestion, queuing) that arise from the non-linear, spatial interactions of individual, heterogeneous agents (Castle and Crooks, 2006). The model was implemented in the AnyLogic 8 simulation platform, selected for its powerful 2D/3D visualization engine for model verification and its Java-based extensibility, which was essential for implementing custom agent logic. The model represents a conceptual evolution of the passenger simulation framework developed by Perkins, Ryan, and Siebers (2015), adapting its core Social Force Model physics engine while introducing new behavioural archetypes, architectural components, and data-fusion algorithm.

2.1. Simulation Environment and Data Grounding

The simulation environment consists of a simplified but realistic model of a single UK train station platform, developed in accordance with general Network Rail design principles (Network Rail, 2011). As annotated in **Figure 3-3**, the platform is a rectangular area functionally divided into two main regions: a large farPlatform area for general waiting and congregation, and a narrow nearPlatform strip adjacent to the tracks, representing the final queuing space. A three-carriage train, with specifications based on the Class 387 "Electrostar" model, is simulated along the platform edge.

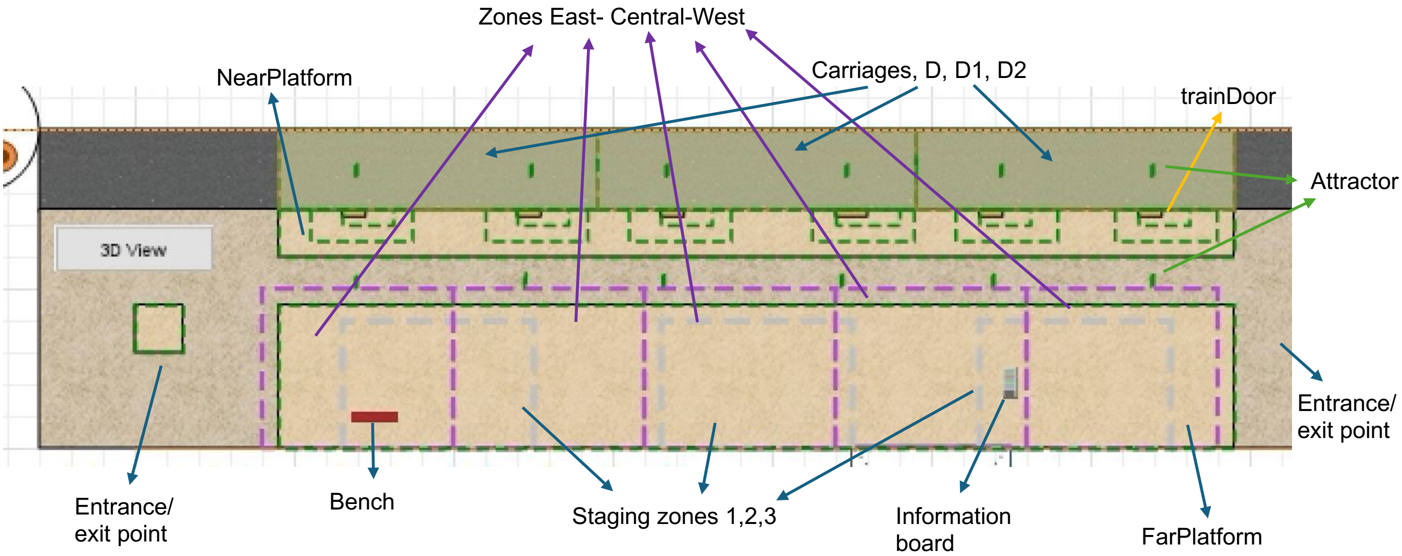


Figure 1-2 Annotated Simulation Environment in Anylogic

To manage agent behaviour and facilitate granular data collection, the environment incorporates two critical overlay systems. The first is a set of three logical Staging Zones, which correspond to the platform areas in front of each carriage. These zones do not represent physical markings but serve as intermediate targets for agents to position themselves strategically before the train's arrival. Zones, combined with timing algorithms, were found to be crucial to prevent unrealistic premature crowding at the platform edge. The second system is a set of five zones that partition the platform longitudinally for the sole purpose of data analysis, allowing for a quantitative assessment of passenger distribution at any given time.

A core principle of this research was to ground the model's baseline behaviour in real-world conditions. This was achieved through an Exploratory Data Analysis (EDA) of anonymised LiDAR sensor data, captured during a 2023 trial of the "Situate" system at Bristol Temple Meads station (Createc, 2022). Using Python libraries (pandas, matplotlib), the raw sensor data was analysed to quantify passenger dynamics. A temporal analysis confirmed distinct morning (07:00-09:00) and evening (17:00-18:30) activity peaks, justifying the passenger volumes selected for the experiments. More importantly, a spatial analysis, visualized in the density heatmap in **Figure 3-6**, revealed a highly non-uniform passenger distribution. As quantified in the analysis shown in **Figure 3-7**, there is significant and persistent clustering near platform entrances and information boards. This EDA empirically validated the "concentrated boarding" phenomenon and directly informed the inclusion of two Congregation Points (a bench and an information board) in the simulation to replicate this observed organic behaviour.

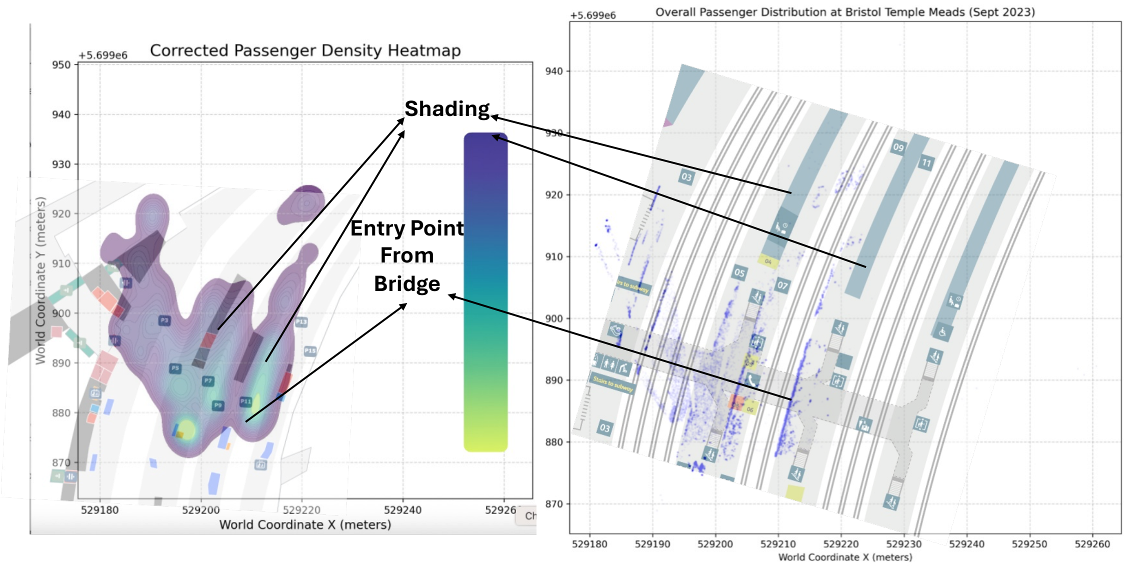


Figure 3-6

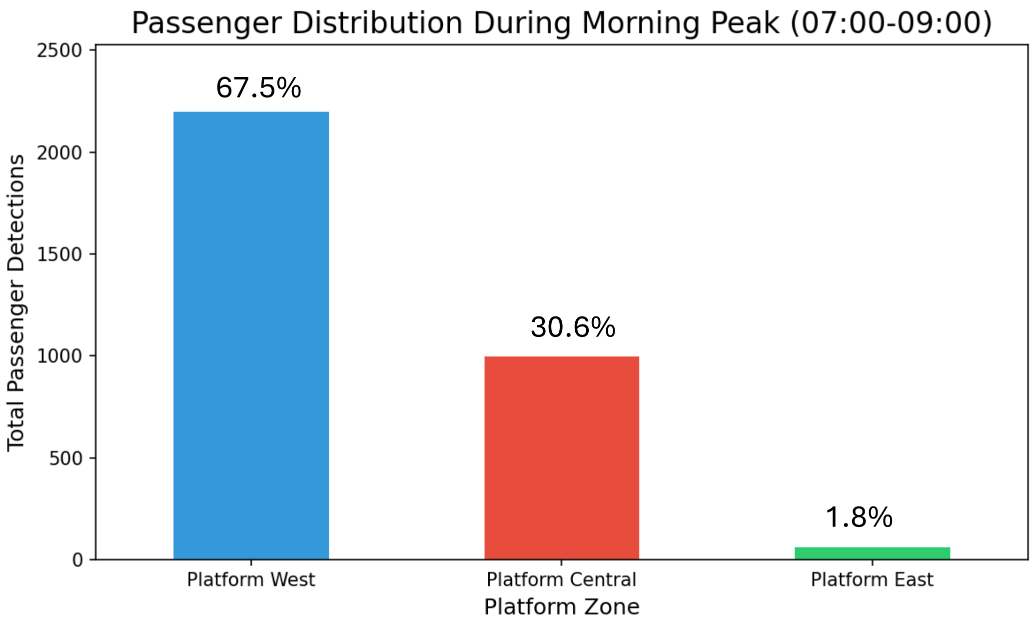


Figure 3-7

2.2. Agent Dynamics: Movement and Decision-Making

The behaviour of each passenger agent is governed by two distinct but interacting components: a movement model that dictates their physical motion, and a decision-making model that determines their goals. This separation allows for the simulation of a heterogeneous crowd where different types of passengers can share the same physical space but follow unique behavioural rules, as conceptualized in **Figure 3-1**.

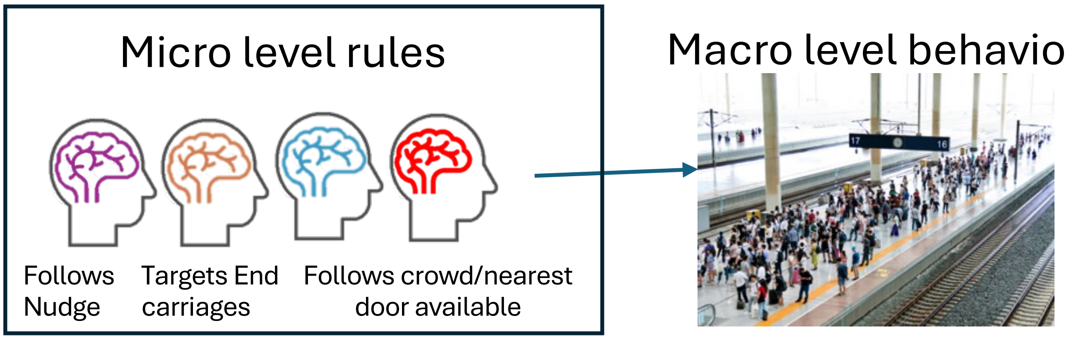


Figure 3-1

Movement: The Social Force Model

$$m_i \frac{dv_i}{dt} = m_i \frac{v_i^0(t) e_i^0(t) - v_i(t)}{\tau_i} + \sum_{j \neq i} f_{ij} + \sum_{w} f_{iW}$$

$$f_{ij} = f^{psy}ij + f^{phy}ij, \qquad f^{psy}ij = A_i \exp\left(\frac{rij - d_{ij}}{B_i}\right) n_{ij}$$

$$f^{phy}ij = kg(rij - d_{ij}) n_{ij} + \kappa g(r_{ij} - d_{ij}) \Delta v_{ji}^t \, t_{ij}$$

Equation 1

The physical movement of all agents within the 2D continuous space is governed by a custom implementation of the Social Force Model (SFM), based on the foundational work of Helbing and Molnar (1995). The SFM calculates each agent's acceleration at every time step as the vector sum of forces, formally defined in **Equation 1**. This includes an attractive force pulling the agent toward their current target, a repulsive force from other agents that increases exponentially at close distances to prevent overlapping, and a repulsive force from static obstacles like platform edges. This microscopic model produces realistic emergent crowd behaviours such as lane formation, collision avoidance, and congestion at bottlenecks, without being explicitly programmed.

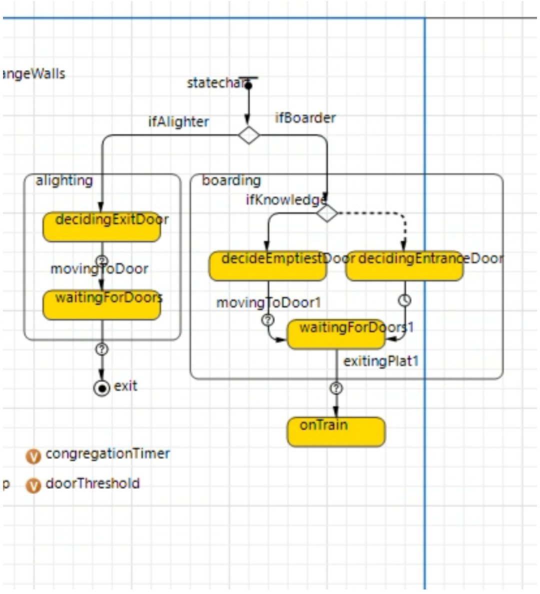


Figure 3-8 Passenger statechart from Anylogic

Decision-Making: Passenger Archetypes and State Logic
The core individual logic of each agent is managed by a statechart, depicted in **Figure 3-8**, which transitions agents through distinct behavioural states such as decidingEntranceDoor, movingToDoor, and waitingForDoors. A key part of this model is the implementation of a heterogeneous crowd through four distinct passenger archetypes, whose rules were informed by human factors research and consultation with the Rail Safety and Standards Board (RSSB).

```
For each carriage i:
    combinedLoad[i] = trueLoadData[i] +
liveStagingZoneCounts[i]

optimalCarriage = argmin(combinedLoad)
```

Combined Nudge Logic (Purple passenger)

Algorithm 1

```
if (randomTrue(0.5)):
    target = stageZone1  // End carriage strategy
else:
    target = stageZone3  // End carriage strategy
```

Combined Nudge Logic (Purple passenger)

Algorithm 2

1. **The Nudge Follower (Purple):** This archetype represents the segment of the population receptive to real-time information. As defined in **Algorithm 1**, these agents initially wait in the farPlatform area for instructions. Upon receiving a nudge, they update their target to the designated optimal staging zone.

2. **The Strategic Planner (Orange):** This archetype models experienced commuters with a pre-determined plan based on prior knowledge (e.g., to be near the exit at their destination). Their logic, shown in **Algorithm 2**, is to ignore all external information and immediately target one of the two end staging zones with a 50% probability.

3. **The Crowd Followers (Blue & Red):** These two archetypes represent the majority of passengers who rely on simple heuristics of proximity and observable crowd density. They are initially drawn to one of the Congregation Points. Blue passengers will re-evaluate their choice of staging zone at t=62s and may switch to a less crowded adjacent zone. Red passengers exhibit stronger "herding" behaviour, remaining at congregation points for a longer, more variable duration, simulating tourists or unfamiliar travellers.

2.3. The Combined Nudging Algorithm

The primary technical innovation of this research is the development and implementation of the Combined Nudging Algorithm. This algorithm serves as the decision-making core for the Nudge Follower agents and is designed to simulate an intelligent, data-driven passenger information system. As conceptually illustrated in **Figure 3-10**, the fundamental purpose of the algorithm is to dynamically identify the optimal boarding location on the platform by fusing two distinct streams of data: static, pre-existing information about the arriving train, and dynamic, real-time information about the waiting crowd.

The algorithm's logic is executed once at a specific, pre-determined moment in the simulation timeline (t=40s). This timing is a critical design choice. Intervening too early would result in a decision based on an unsettled and unrepresentative crowd distribution. Intervening too late would find passengers already committed to a boarding location and thus less likely to respond. The t=40s trigger allows the initial crowd to settle into a stable, natural distribution, providing a realistic baseline for the intervention, while still allowing sufficient time for Nudge Follower agents to react and reposition before the train's arrival at t=60s. This single-scan-then-nudge approach was chosen over continuous tracking for three key reasons: computational efficiency (avoiding noisy recalculations), behavioural realism (a single, clear instruction is more likely to be followed), and implementation feasibility (mirroring how a real-world snapshot-based system would operate, although LiDAR will track passenger movement along platform).

```
// Executed at t=40s

// 1. Perform LiDAR Scan
For each stagingZone i from 1 to 3:
   liveStagingZoneCounts[i] = count_passengers_in(stagingZone[i])

// 2. Calculate Combined Load
For each carriage i from 1 to 3:
   combinedLoad[i] = trueLoadData[i] + liveStagingZoneCounts[i]

// 3. Find Optimal Carriage
optimalCarriage_index = find_index_of_minimum(combinedLoad)

// 4. Issue Nudge
For each passenger p in population:
   If p.archetype is NUDGE_FOLLOWER:
      p.target = staging_zone_of(optimalCarriage_index)
```

The algorithm calculates a "combined load" for each of the three train carriages, as detailed in the pseudocode of **Algorithm 4**. The formula is: CombinedLoad_i = TrueLoadData_i + LiveStagingZoneCount_i

Where:

- **TrueLoadData** is a static integer array representing the number of passengers already on board each carriage of the arriving train. In a real-world implementation, this would be sourced from the train's Automatic Passenger Counting (APC) systems. In the simulation, it is a configurable parameter.
- **LiveStagingZoneCount** is a dynamic integer array that stores the number of passengers present in each of the three logical Staging Zones at the moment the LiDAR scan is performed. This simulates the output of a real-world LiDAR system.

After calculating the CombinedLoad for all carriages, the algorithm identifies the index of the carriage with the minimum value and directs all Nudge Follower agents to the corresponding staging zone. This data fusion approach allows for a more robust and contextually-aware decision than a system based on either data stream alone, aiming to balance the total passenger load both on the platform and within the train.

2.4. Experimental Design and Procedure

A structured experimental design was developed to test the primary research hypothesis: that an increase in passenger compliance with data-driven nudging directives leads to a measurable reduction in total boarding time.

The simulation follows a repeatable, event-driven timeline. At t=0s, the initial population of boarding passengers is created, and agents move from their spawn points to their initial positions based on their assigned archetype. Between t=0s and t=40s, the crowd settles into an organic, non-uniform distribution. At t=40s, the Combined Nudging Algorithmis executed, and Nudge Follower agents are given new targets. The train arrives at t=60s, at which point a fixed population of

17 "Alighter" agents is spawned inside the carriages. The train doors open at t=90s, and the "Alighters First" policy is enforced by preventing boarders from moving until the last alighter has cleared the platform (monitored by the monitorAlighters event, typically completing around t=99s). Once boarding is permitted (boardingAllowed = true), all boarding agents abandon their staging zone targets and activate their final door-seeking behaviour, governed by the SFM. The simulation concludes when the last boarding agent has been removed from the platform population.

- **Independent Variables:**
    - *Primary:* **Nudge Effectiveness** was the most critical variable, representing the proportion of the population assigned the Nudge Follower archetype. This double parameter was systematically tested at levels of 0%, 20%, 50%, 80%, and 100%.
    - *Secondary:* **Passenger Population Density** was tested at two levels: a standard-density scenario (49 boarders) and a high-density scenario (150 boarders) to assess scalability.
    - *Tertiary:* **Initial Train Load** was configured in an "End-Heavy" distribution to present a challenging and realistic problem for the algorithm to solve.
- **Dependent Variables (Key Performance Indicators):**
    - **Total Boarding Time:** The primary KPI, calculated as the time elapsed from the moment the boardingAllowed gate is opened until the final boarding passenger is considered "on the train".
    - **Passenger Distribution Efficiency:** To understand the mechanism of the nudge, this was quantified by the standard deviation of passenger counts across the five Coordinate Zones at t=89s (the moment just before doors open). A lower standard deviation indicates a more balanced platform and thus a more effective nudge.

To account for the stochastic elements inherent in the model (e.g., random initial positions, probabilistic archetype assignment), each unique combination of independent variables was simulated multiple times. The results presented in the following chapter represent the average outcomes across these replications to ensure statistical robustness.

3. Results

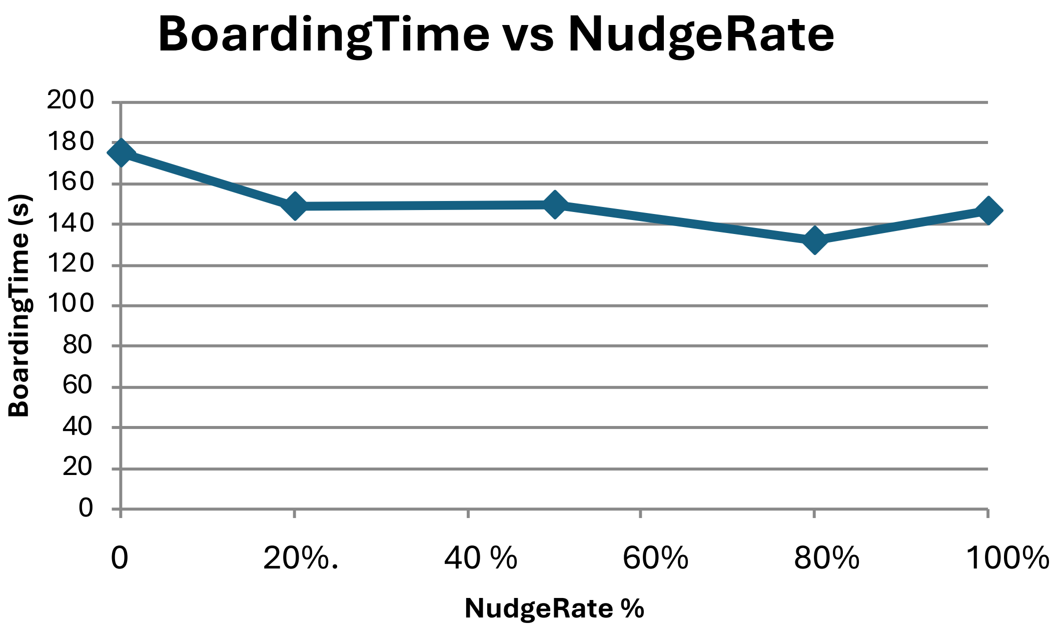
The simulation experiments were conducted to quantify the relationship between the effectiveness of the Combined Nudging Algorithm and the efficiency of the passenger boarding process. This section presents the results from the standard-density scenario (49 boarders), beginning with the primary key performance indicators (KPIs) and followed by an analysis of the underlying passenger distribution dynamics.

3.1. The Non-Linear Impact of Nudging on Boarding Time

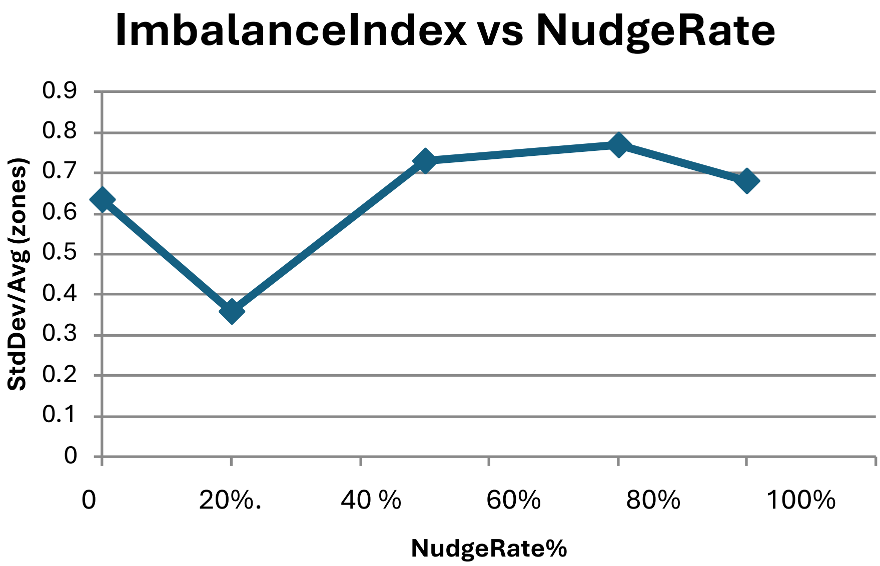
The primary KPI, total boarding time, reveals a complex, non-linear relationship with nudge effectiveness. As illustrated in **[Graph 1]**, the nudging strategy successfully reduces total boarding time, but its effectiveness does not scale linearly with passenger compliance. In the baseline scenario (0% nudge effectiveness), with no intervention, the mean boarding time was approximately 170 seconds. A modest nudge rate of 20% yielded a notable improvement, reducing the mean time by over 10% to around 150 seconds. The optimal performance was observed at an 80% nudge effectiveness rate, which achieved the lowest mean boarding time of approximately 145 seconds, a reduction of nearly 15% from the baseline.

Critically, the results reveal a paradoxical outcome at maximum compliance. A theoretically "perfect" 100% effective nudge was detrimental to performance, causing the mean boarding time to increase again to a level comparable with the

20% scenario. This demonstrates that simply maximizing compliance does not guarantee optimal efficiency.



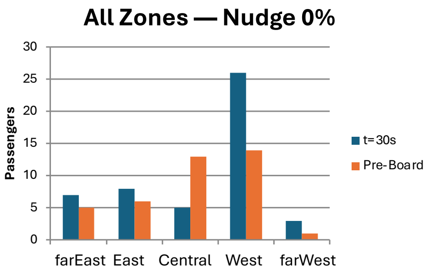
The underlying cause of this U-shaped performance curve is explained by the platform's Imbalance Index, a metric that quantifies the evenness of passenger distribution, with lower values indicating a more balanced platform. As shown in **[Graph 2]**, the relationship between nudging and platform balance is not straightforward. While a 20% nudge rate improves the platform balance compared to the baseline, nudge rates of 40% and higher cause the platform to become progressively more imbalanced. This is an expected emergent property of the algorithm's design; with an initial train load configuration of (25, 5, 20), the central carriage is consistently identified as the optimal target. Consequently, higher nudge rates do not disperse passengers but rather concentrate them in the central zone, shifting the hotspot instead of eliminating it.

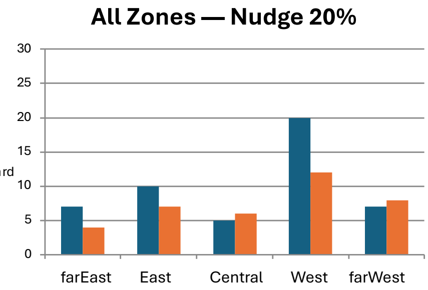


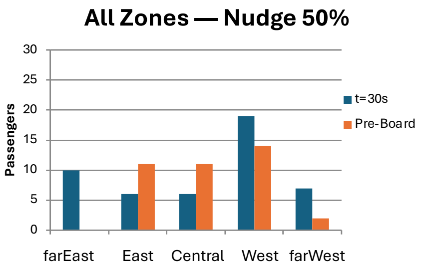
3.2. Visualising Passenger Redistribution on the Platform

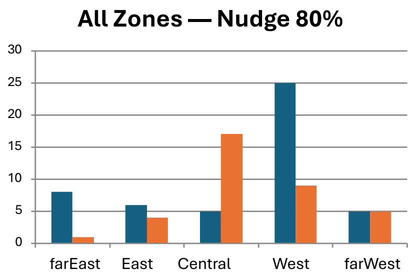
The direct impact of the nudging algorithm on passenger movement is visualized by comparing the crowd's spatial distribution before the nudge (t=30s) and just before boarding (t=89s). **Figure [Graph 4]** illustrates this macro-level redistribution across the five coordinate zones for different nudge effectiveness rates.

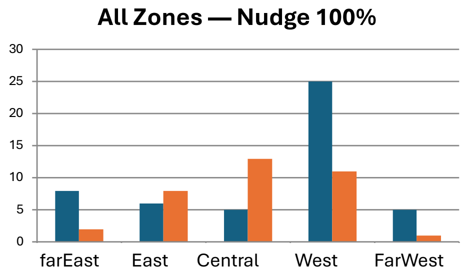
The 0% nudge scenario serves as a crucial baseline, revealing the organic, imbalanced distribution of the crowd. A high concentration of passengers is observed in the "West" zone, a direct result of the Crowd Follower agents being drawn to the information screen located in that area. As the nudge rate increases, the effectiveness of the intervention becomes visually apparent. In the 80% nudge scenario, the number of passengers in the "West" zone is substantially lower in the pre-boarding phase (orange bar) compared to the baseline (blue bar). Conversely, the "Central" zone, which was initially less populated, sees a significant influx of passengers. This provides clear, visual evidence of the causal link: the nudge successfully mitigates the natural hotspot by redirecting passengers to the algorithmically determined optimal central zone.



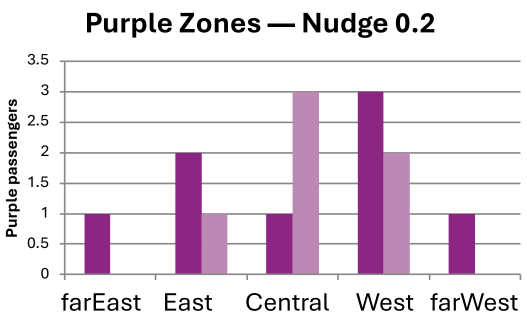


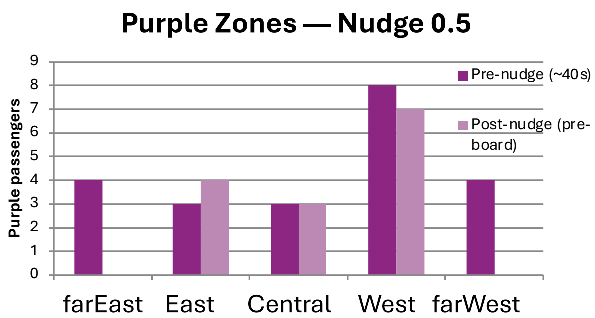


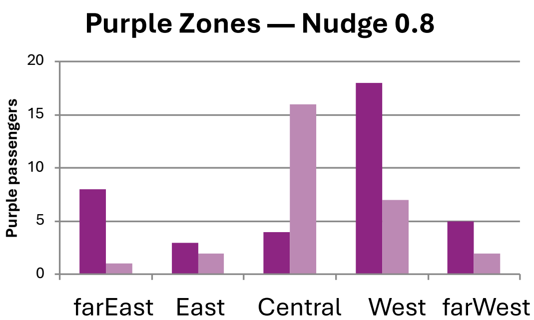


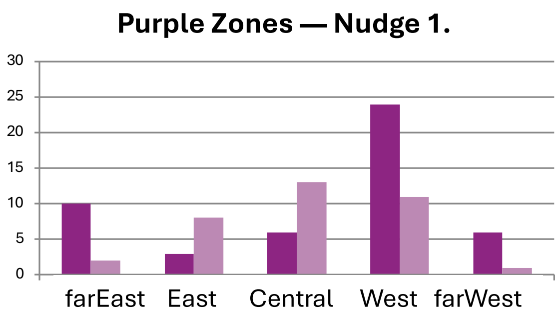


To confirm that the Nudge Follower (Purple) archetype is the primary driver of this redistribution, their movement was isolated. As shown in **Figure [Graph 5]**, the pre-nudge distribution (t<40s, purple bars) shows these agents scattered across the platform. The post-nudge distribution (pink bars) reveals a clear and dramatic consolidation of the Purple population into a single target zone. This provides unequivocal proof that the lidarScanAndNudge event and the associated agent logic are functioning as designed, acting specifically on the intended target population to alter the platform's overall distribution.









3.3. Analysis of Passenger Archetype Behaviour

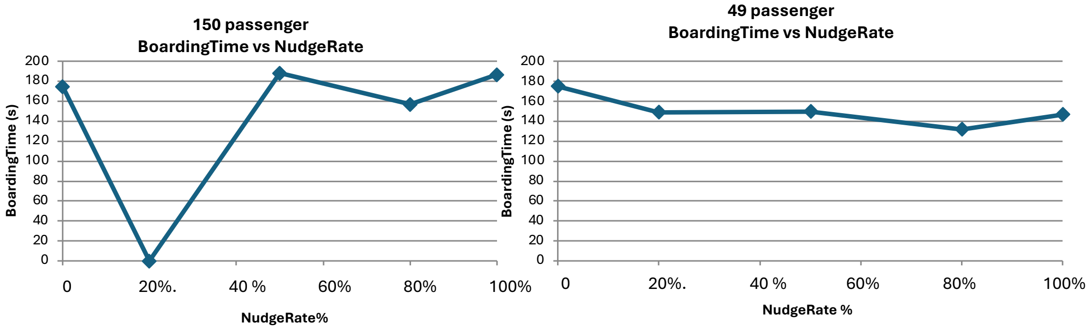
To verify the internal validity of the simulation, the final pre-boarding distributions were analyzed by archetype composition. **Figure [Graph 7]**, which depicts the composition of the 80% nudge effectiveness scenario, provides clear evidence of the nudge's impact. The bar for the "Central" zone, the new hotspot, is overwhelmingly dominated by Purple (Nudge Follower) passengers. This confirms that the new point of congestion identified by the Imbalance Index is created almost exclusively by the target agents correctly responding to the algorithm's directive.

Conversely, the behaviour of the Strategic Planner (Orange) archetype demonstrates a deliberate resistance to the nudge. In both the 20% and 80% scenarios (**Figure [Graph 6]** and **Figure [Graph 7]**), these agents show a strong preference for the end zones ("FarEast" and "FarWest") and their adjacent zones, with minimal presence in the "Central" zone. This confirms they are successfully ignoring the nudge and adhering to their pre-defined strategic goals. Finally, the Crowd Follower (Blue and Red) archetypes constitute the majority of the ambient population in the low-nudge scenarios, forming the bulk of the organic congestion in the "West" zone, correctly following their congregation and proximity-based logic. As the nudge rate increases, their population diminishes as they are re-assigned to the Nudge Follower archetype, but their heuristic-based behaviour remains consistent.

[Insert Figure of Graph 6: Pre-Board Zone Composition at 20%]
[Insert Figure of Graph 7: Pre-Board Zone Composition at 80%]

3.4. High-Density Scenario

To assess the scalability of these findings, the experiments were repeated with a high-density population of 150 boarders. The results, presented in **Figure [Graph 17]**, show that the BoardingTime KPI exhibits the same characteristic U-shaped curve observed in the standard-density tests. This confirms that the creation of an artificial bottleneck in the single "optimal" zone is a fundamental outcome of the global nudging strategy and not an artifact of low passenger numbers. As expected, the overall boarding times increased significantly with the threefold increase in passenger count, validating the underlying Social Force Model's response to higher densities.



Graph 1: 150 vs 50 Passenger boarding time

4. Discussion

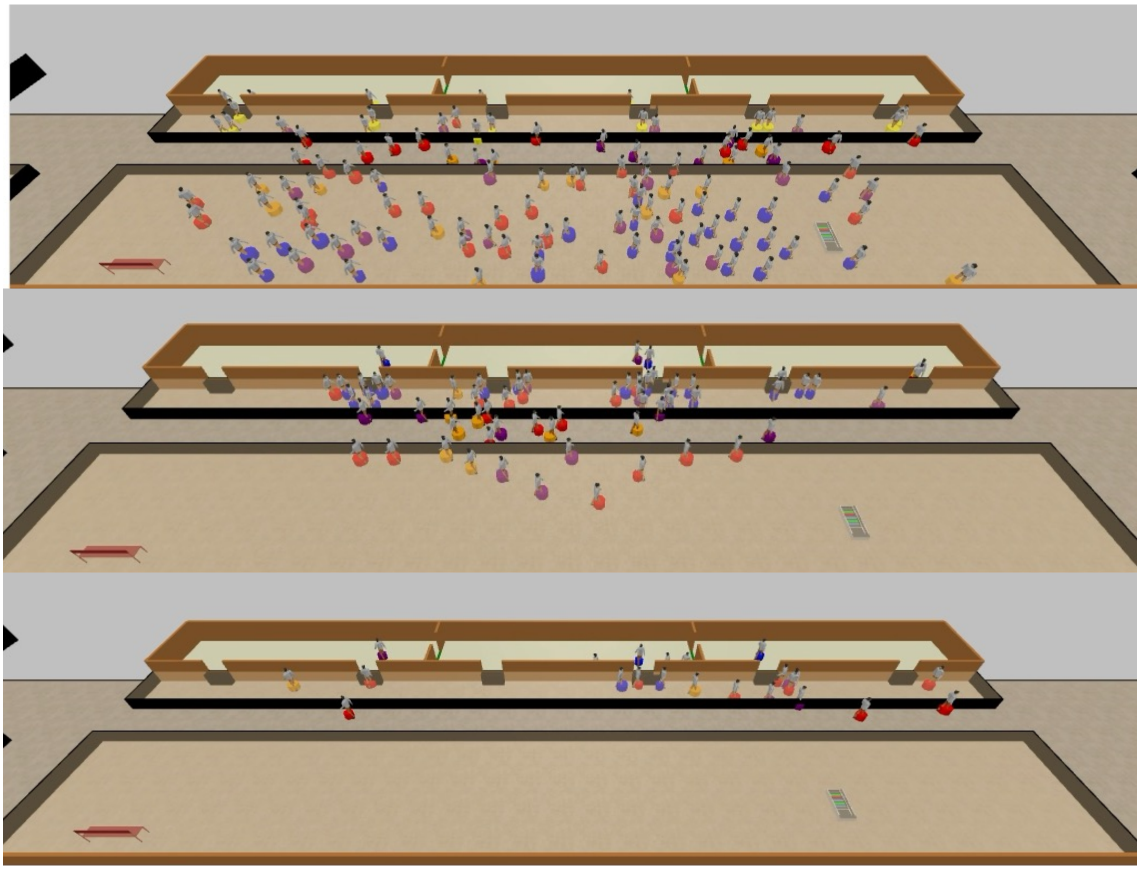
The simulation results demonstrate that a data-driven nudging strategy, which fuses static train load with dynamic platform data, can significantly reduce train boarding times. However, the primary finding is that the relationship between passenger compliance and boarding efficiency is non-linear and paradoxical. This section interprets these findings, discusses their practical implications for railway operations, and acknowledges the limitations of the current study.

4.1. The "Shifting the Hotspot" Phenomenon: Why Perfect is Not Optimal

The most significant finding of this study is the U-shaped performance curve for boarding time, where optimal efficiency was achieved at a high but imperfect 80% compliance rate, while a 100% compliance rate was detrimental. The underlying cause, supported by the Imbalance Index data and passenger distribution visualizations, can be described as the "shifting the hotspot" phenomenon.

In the absence of intervention, a natural hotspot forms due to passenger heuristics. A low-to-moderate nudge (e.g., 20%) successfully mitigates this by redirecting a portion of the crowd, leading to a more balanced platform and a notable reduction in boarding time. However, as compliance increases towards

100%, the global nature of the algorithm directs every compliant passenger to the single "optimal" zone. Rather than dispersing the crowd, the intervention simply re-concentrates it at a new location, creating an artificial bottleneck that throttles the rate of boarding. This finding has a critical implication: the naive application of a global "best-zone" instruction, even if perfectly followed, can be detrimental. The goal of an intelligent passenger guidance system should not be to concentrate all passengers, but to achieve a more uniform distribution across all available doors.



Visualisation of the passenger surge from staging zones to the nearPlatform area at t=60/120/150s, illustrating the formation of the pre-boarding queue

4.2. Temporal Dynamics and the Queuing Effect

The temporal analysis provides further insight into this bottlenecking effect. The data reveals that the peak number of passengers congregating at the platform edge actually increases with the nudge rate. While nudging creates a larger initial queue, it appears to be a more *organized* one, ready for boarding. However, at 100% compliance, the density of this initial surge becomes too great, and the physical constraints of the doorways become the limiting factor, slowing the queue's dissipation and increasing the overall time. This highlights a crucial tension between strategic pre-positioning and the physical realities of passenger flow through a constrained space.

4.3. Implications for Real-World Implementation

These findings offer several key recommendations for the practical deployment of passenger nudging systems. The different compliance rates tested in the simulation can be mapped to the likely effectiveness of different intervention technologies.

- **Low-to-Medium Compliance (0-30%):** This level may be achievable with passive systems like dynamic information screens or audio announcements. Existing carriage information screens, which often rely only on APC data, could be significantly augmented by fusing live LiDAR crowding data with historical and current ticket sales data to provide more accurate guidance.
- **Optimal High Compliance (40-80%):** Achieving this higher, optimal rate would likely require more direct and engaging technologies, such as dynamic LED lighting on the platform edge or personalized mobile notifications.
- **Overcoming the Bottleneck:** To overcome the "shifting the hotspot" problem, a deployed system should evolve beyond global commands towards **zonal-specific nudging**. A LiDAR-enabled system is uniquely poised for this. Instead of issuing a single instruction to everyone, it could target instructions *only* at passengers detected within the most congested areas, guiding them to the nearest, less-congested alternative, a capability already being explored in airport and stadium management.

Furthermore, two principles are key for future deployment. The first is **context-awareness**: the nudging system must be intelligent enough to deactivate itself during periods of low passenger density when no intervention is needed. The second is to consider alternate station design; dispersing congregation points like benches and screens could passively encourage a more even initial crowd layout.

4.4. Limitations and Future Research

This study provides valuable insights, but its limitations point towards important avenues for future research. Firstly, the simulation environment is a simplification, modelling a single platform section and making it most applicable to origin or terminus stations.

Secondly, the passenger archetypes, while an advance over homogenous models, do not capture the full spectrum of human factors. The model did not consider vulnerable groups (e.g., women at night who may prefer crowded areas for safety), passengers with mobility impairments or heavy luggage, or the impact of environmental conditions. Future models should incorporate these crucial human factors for greater realism. Comparing historical LiDAR crowding data against arriving train load data for specific stations would also greatly aid in providing an empirical baseline for these more nuanced models.

Finally, while this simulation provides a robust testbed, the ultimate validation lies in real-world application. Enabling operators to conduct live **A/B tests**, comparing different nudging technologies and strategies against a control, would provide invaluable data on their true effectiveness. Pairing such empirical results with simulation data would create a powerful feedback loop for continuously refining and optimizing intelligent passenger guidance systems.

5.Conclusion

This research successfully demonstrated through agent-based simulation that a data-driven nudging strategy, which fuses static train load data with real-time platform crowding information, can significantly influence passenger distribution and reduce train boarding times.

The central contribution of this work is the identification of a non-linear and paradoxical relationship between passenger compliance and boarding efficiency. Optimal performance, achieving a boarding time reduction of nearly 15%, was found at a high but imperfect 80% compliance rate. A theoretically perfect 100% compliance rate proved to be counter-productive, as the global nature of the nudge simply shifted the congestion hotspot, creating a new artificial bottleneck at the designated "optimal" zone.

These findings have significant implications for practice, showing that while passenger nudging is a powerful tool for reducing dwell times, its naive application can be detrimental. The future of intelligent passenger guidance lies not in concentrating all passengers, but in dispersing them. By moving towards targeted, context-aware algorithms and leveraging real-world data collection, train operators can evolve from simple crowd control to a proactive system of intelligent passenger management, unlocking significant gains in both operational efficiency and the passenger experience. Future work should focus on simulating these zonal-specific nudging strategies and validating the models through live, in-station A/B testing.

References

Bosina, E., et al. (2015) ‘Distribution of passengers on railway platforms’, in Proceedings of the 15th Swiss Transport Research Conference. Monte Verità, Ascona, 2015.

Christoforou, Z., et al. (2020) ‘Factors affecting knock-on delays in metro lines’, Journal of Rail Transport Planning & Management, 16, 100196. doi: https://doi.org/10.1016/j.jrtpm.2020.100196.

CIO (2017) Sydney Trains commences Dwell Track trial to ease crowding, reduce delays. Available at: https://www.cio.com/article/201767/sydney-trains-commences-dwell-track-trial-to-ease-crowding-reduce-delays.html (Accessed: 23 August 2025).

Davidich, M., et al. (2013) ‘A validated model for pedestrian flow in train stations with waiting areas’, Transportation Research Part C: Emerging Technologies, 31, pp. 136–152. doi: https://doi.org/10.1016/j.trc.2013.02.016.

Fang, J., Fujiyama, T. and Wong, H. (2019) ‘Modelling passenger distribution on metro platforms based on passengers’ choices for boarding cars’, Transport Planning and Technology, 42(5), pp. 442–458. doi: https://doi.org/10.1080/03081060.2019.1609218.

Gong, L., et al. (2018) ‘Understanding users’ behaviours in relation to concentrated boarding in metro systems’, Safety Science, 108, pp. 232–242. doi: https://doi.org/10.1016/j.ssci.2018.04.017.

Harris, N.G. and Anderson, R.J. (2007) ‘An international comparison of urban rail boarding and alighting rates’, Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 221(4), pp. 521–526. doi: https://doi.org/10.1243/09544097JRRT79.

Holloway, C., et al. (2015) All Aboard! The Impact of Luggage on Boarding and Alighting Times. Available at: https://www.researchgate.net/publication/282855135_All_Aboard_The_Impact_of_Luggage_on_Bording_and_Alighting_Times (Accessed: 23 August 2025).

Offiaeli, N. and Yaman, F. (2021) The Green Nudge. Available at: https://www.green-nudges.com/green-lanes/ (Accessed: 23 August 2025).

Sutton, R. and Moncrieff, S. (2015) Train interior layout and its effect on passenger boarding and alighting times. Available at: https://www.researchgate.net/publication/303862283_Train_interior_layout_and_its_effect_on_passenger_boarding_and_alighting_times (Accessed: 23 August 2025).

Vromans, M. (2005) Modeling and scheduling of railway networks. PhD thesis, Technische Universiteit Eindhoven. Available at: https://research.tue.nl/en/publications/modeling-and-scheduling-of-railway-networks.

Weston, J. (1989) 'The role of operational research in London Underground', as cited in Harris, N.G. and Anderson, R.J. (2007).
Anagnostis, A., et al. (2022) ‘An Intelligent Vision System for Increasing Safety and Capacity on Railway Platforms’, Sensors, 22(3), 1184. doi: https://doi.org/10.3390/s22031184.

WCBS 880 Newsroom (2020). LIRR App Shows Crowding Data For Trains. [online] Audacy.com. Available at: https://www.audacy.com/wcbs880/articles/news/lirr-app-shows-crowding-data-for-trains# [Accessed 23 Aug. 2025].

CIO (2017) Sydney Trains commences Dwell Track trial to ease crowding, reduce delays. Available at: https://www.cio.com/article/201767/sydney-trains-commences-dwell-track-trial-to-ease-crowding-reduce-delays.html (Accessed: 23 August 2025).

Createc (2022) Crowd watch: DASA funded AI and LiDAR innovation helps enhance crowd safety. Available at: https://www.gov.uk/government/news/crowd-watch-dasa-funded-ai-and-lidar-innovation-helps-enhance-crowd-safety (Accessed: 23 August 2025).

Global Railway Review (2022) Q&A with DB: New real-time occupancy display systems. Available at: https://www.globalrailwayreview.com/article/141464/deutsche-bahn-qa-new-real-time-occupancy-display-systems/ (Accessed: 23 August 2025).

IMechE (2023) Lidar and AI to monitor crowds at Bristol train station. Available at: https://www.imeche.org/news/news-article/lidar-and-ai-to-monitor-crowds-at-bristol-train-station (Accessed: 23 August 2025).
Li, M., et al. (2024) ‘Train dwell time modelling based on automatic passenger counting data: A quantile regression approach’, Journal of Rail Transport Planning & Management, 29, 100431. doi: https://doi.org/10.1016/j.jrtpm.2024.100431.

Outsight (2022) LiDAR Solutions for Train Stations. Available at: https://www.outsight.ai/use-cases/railway-stations (Accessed: 23 August 2025).

Seriani, S. and Fernandez, R. (2015) ‘Modeling people behavior at boarding and alighting in metro stations’, Journal of Rail and Rapid Transit, 229(8), pp. 865–879. doi: https://doi.org/10.1007/s40864-015-0010-0.

Ahmed, A.A., et al. (2017) ‘An Agent-Based Modeling and Simulation of Bus Movement in a Bus Rapid Transit Station’, in 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 326–330. doi: https://doi.org/10.1109/IEEM.2017.8289912.

Perkins, P., et al. (2015) Reducing Dwell Times via Simulation: A Feasibility Study using Agent-Based Modelling. Available at: https://people.cs.nott.ac.uk/pszps/docs/MAS2015-Perkins.pdf (Accessed: 23 August 2025).

Ramanan, A., et al. (2024) Integrated Railway Station Simulation: Modeling Passenger and Train Dynamics with AnyLogic. Available at: https://www.researchgate.net/publication/379642133_Integrated_Railway_Station_Simulation_Modeling_Passenger_and_Train_Dynamics_with_AnyLogic (Accessed: 23 August 2025).

Sagerschnig, C., et al. (2024) ‘Spatiotemporal Trajectories of Pedestrian Mobility at the Train Station: evidence of 24 million trajectories’, Scientific Data, 11(1), 389. doi: https://doi.org/10.1038/s41597-024-04071-9.

Shamsheer-hub (2021) SimPy-Simulation-Model-for-Railway-Platform-Ticket. Available at: https://github.com/Shamsheer-hub/SimPy-Simulation-Model-for-Railway-Platform-Ticket (Accessed: 23 August 2025).

Simulating passenger movement at the Platform Train interface to understand the impact of station design on dwell time (2017) Rail Technical Strategy. Available at: https://railtechnicalstrategy.co.uk/case-study/simulating-passenger-movement-at-the-platform-train-interface-to-understand-the-impact-of-station-design-on-dwell-time/ (Accessed: 23 August 2025).

Zhang, Y., et al. (2022) M2DGR: A Multi-sensor and Multi-scenario SLAM Dataset for Ground Robots. Available at: https://doi.org/10.48550/arXiv.2204.01026.

Castle, C. and Crooks, A. (2006). Principles and Concepts of Agent-Based Modelling for Developing Geospatial Simulations. CASA Working Paper 110. London: University College London.

Helbing, D. and Molnar, P. (1995). Social force model for pedestrian dynamics. Physical Review E, 51(5), pp.4282-4286.

Perkins, A., Ryan, B. and Siebers, P.-O. (2015). Modelling and Simulation of Rail Passengers to Evaluate Methods to Reduce Dwell Times. Nottingham: University of Nottingham.

Class 365 Class 387. (n.d.). Available at: https://grumpytrainuser.wordpress.com/wp-content/uploads/2017/10/great_northern_standing.pdf [Accessed 20 Aug. 2025].

Network Rail (2011). Station Capacity Assessment Guidelines. Station Design Principles. London: Network Rail