

A hybrid optimization framework for dynamic rolling-stock assignment at congested railway terminals

Yash Kumar¹, Ramesh Chandra Sahoo¹, Prashant Dixit²

¹Department of Computer Science and Engineering, Manav Rachna International Institute of Research and Studies, Faridabad, India

²Department of Computer Science and Engineering, Galgotias University, Greater Noida, India

Article Info

Article history:

Received May 20, 2026

Revised Aug 26, 2026

Accepted Sep 4, 2026

Keywords:

Delay propagation

Dynamic scheduling

Indian railways

Rolling-stock assignment

Terminal optimization

ABSTRACT

Delayed arrival at busy stations results in cascading effects. Late arrivals mean that an incoming rake would not be available for reuse and the subsequent departure would be affected. The research investigates this issue for New Delhi's train station (NDLS), India's busiest rail terminus and a case study based on 58,698 train movements throughout 2023. It reveals that, as 79% of arriving express trains reached NDLS with delays (mean 53 minutes), these delays significantly transferred to departing trains ($r=0.491$). Of 6.86 million available buffer minutes, only 13% was actually utilized, rest were "frozen" by an inflexible static assignment method. This research formulates the rolling stock assignment problem as a constrained bipartite matching problem and develops a hybrid assignment algorithm (HAA) that integrates a cost matrix based optimization with auction-based assignment and a dynamic re-assignment process. HAA provides 98.9%, 99.9%, and 88.6% delay absorption capacities for three different levels of traffic congestion while reducing originating train delay by up to 39.8 percentage points compared to the present system. Thoughtful assignments not more capacity are the key to improving punctuality in a heavily utilized rail.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Ramesh Chandra Sahoo

Department of Computer Science and Engineering

Manav Rachna International Institute of Research and Studies

Faridabad, Haryana, India

Email: rcsahoo.set@mriu.edu.in

1. INTRODUCTION

The railway network faces severe operational challenges from delay propagation across interdependent infrastructure, rolling stock, and crews. Terminal stations are critical bottlenecks where incoming delays directly impair outgoing train departures. For instance, Indian Railways Mail/Express punctuality declined from 90.48% (FY 2021–2022) to 80.79% (FY 2022–2023) [1], [2], highlighting the urgent need for targeted operational interventions.

Static train assignment (STA) utilizes terminating rakes to specific originating services in the advance. When a train arrives late, its assigned outbound service is inevitably delayed, resulted leaving adjacent buffer times unutilized due to rigid linkages. To overcome this gap, we analyze a full year dataset of 58,698 movements in year 2023 from New Delhi Railway Station (NDLS) quantifying delay propagation and buffer underutilization. We formulate the rolling stock assignment problem (RSAP) as a constrained bipartite matching problem and introduce a dynamic train assignment (DTA) framework using a novel hybrid assignment algorithm (HAA).

To address these challenges and bridge the existing research gap, the primary contributions of this study are outlined as follows:

- Characterization and quantification of terminal level delay propagation and underutilized buffers using non-public IR data.
- A constrained bipartite-matching formulation for RSAP incorporating platform, maintenance, and departure constraints.
- Development of the HAA-driven DTA framework, reducing originating delays by up to 39.8 percentage points without requiring timetable alterations or capital expenditures.

The remainder of this paper is organized as: section 2 reviews the literature that is relevant to the findings; section 3 presents the empirical analysis; section 4 develops the mathematical formulation; section 5 discusses the HAA methodology; section 6 describes the experimental design and the metrics used in examining the analysis; section 7 presents and discusses the findings; and section 8 provides tactical insights from the research findings, as well as opportunities for further inquiry.

2. RELATED WORK

Research on rail operations has included work in timetable optimization, delay propagation, maintenance reliability, terminal operations, passenger metrics, and data-based prediction. Adaptive scheduling simulation [3], incremental timetable adjustments [4], and recovery planning [5] have been instrumental in mitigating delays.

Delay propagation is an area of intense investigation. Various studies have used stochastic modeling, Markov chains, and Bayesian networks in determining timetable fragility [6], [7], weather disruptions [8], incident propagation [9], [10], and suburban traffic flows [11]. Similarly, physical capacity limits have been investigated by studying the dynamics of track one congestion [12], delay propagation driven by dwell time delays [13], queuing at junctions [14], optimal train rescheduling in large passenger stations [15], and robust conceptual timetabling [16]. Research on service disruptions also examines passenger effects and connection recovery rescheduling [17]-[20].

In operational maintenance, research has moved towards predictive maintenance. Various predictive maintenance frameworks include data-driven maintenance analytics [21] and acoustic fault detection systems [22]. The cybersecurity threats in digital maintenance platforms have also been identified [23]. In terminal-level maintenance, developments include layover planning [24], shunting automation [25], and decomposition approaches for real-time rescheduling of the timetable [26].

In parallel to the above, delay prediction techniques have advanced from simple rule-based approaches to machine learning algorithms [27]. Neural network based methods have shown success in modeling nonlinear rail operations [28], prediction uncertainty [29], and in incorporating real-time data for better short-term predictions [30].

However, despite the above advancements, the current literature on rolling stock operations focuses extensively on tactical planning and line-haul circulation using offline solvers such as integer linear programming (ILP) and Hungarian algorithm (HA). Such methodologies need complete prior visibility and cannot be used during terminal-level disruptions in real time. Hence, this research attempts to fill this gap by presenting a lightweight, sub-second dynamic reallocation methodology for congested terminals. The proposed methodology does not require any modifications to the timetable or expensive infrastructure expansions.

3. EMPIRICAL ANALYSIS OF DELAY BEHAVIOUR AT NDLS

NDLS, serves as the focal site for this study. It is located in Delhi, in the National Capital Region, and is one of the major railway stations in India. NDLS operates 16 physically separate but inter-connected active platforms with trains arriving and departing from and to them at about 250 train movements each day (with some peak periods seeing as many as 700,000 passengers in a day). Thus, all of this operational activity on platforms creates an incredibly high level of competition and berthing conflicts, which makes NDLS an optimal place to study delay dynamics at the terminal level.

3.1. Source and pre-processing of data

The main data set for analyzing NDLS is the "Station Train Performance" data set acquired from the Ministry of Railways, Government of India through a request in which data for the entire year 2023 (a total of 58,698 data points) was provided. The data comprises scheduled and actual times of train arrivals and departures along with the delay of each train in minutes, which is the absolute difference between scheduled and actual times. Data points with missing or faulty timestamp values, freight trains and those with unsatisfactory rake linkages have been filtered out. Although an entire year data set helps in the descriptive analysis, for the purpose of experiments and simulations, three one week time periods from January, July, and December 2023 were chosen.

3.2. Service-wise delay decomposition

From the analysis of the operational delay pattern at NDLS, there exist great differences based on the type of services. In terms of trains terminating at NDLS, the pre-arrival delay is high and is estimated to be about 53 minutes. On the other hand, the rail transit system is not much affected by terminal operations since it has little contact with the platform and the arrival delay is very low and estimated to be about 4 minutes. As for the originating trains from NDLS, there is a good operational recovery where the train recovers an average delay of 13 minutes post-departure, though delay at departure is still prevalent since 59% of the departing trains are still delayed at departure as shown in Table 1. Overall, delay characteristics varies substantially depending on whether movements are terminating, originating, or passing through transit.

Table 1. Operational delay profile by service type, NDLS 2023

Aspect	Parameter	Value	Service type	Remarks
Dataset	Total trains	58,698	All	Network-wide annual count
	Total cumulative delay	1,544,219 min	All	Unequally distributed across service types
Terminating	Total trains	22,678	Terminating	
	Mean delay	53 min		Severe average arrival delay
	On-time arrival	21% (4,809)		Low punctuality; 79% arrive late
	Delay 10–60 min	45%		Largest delay band
	Delay >3 h	9.4%		Significant long-duration disruptions
Originating	Total trains	22,678	Originating	
	Mean delay	13 min		Partial downstream recovery
	On-time departure	41% (9,346)		59% depart late; 83% of late < 10 min
Transit	Total trains	13,342	Transit	
	Mean delay	4 min		Minimal terminal interaction; largely unaffected

3.3. Seasonal variations of delays

Figure 1 displays monthly baselines of average delay patterns for terminating and originating rail services. There are two peaks that are evident in the month-to-month analysis, namely, December and January have the highest average delinquency due to heavy winter fog across northern India that created widespread delays due to visibility; while July has the second highest average delinquency as this is the monsoon season in northern India. Outside of the peak months of December, January, and July, terminating delays averaged between 30-60 minutes, while originating delays averaged less than 20 minutes; collectively these results suggest that there was a more stable operating environment during the months of February through November. Because of the peak and off-peak monthly averages, January, July, and December were selected to represent the experimental weeks for the analysis that takes place in section 6.

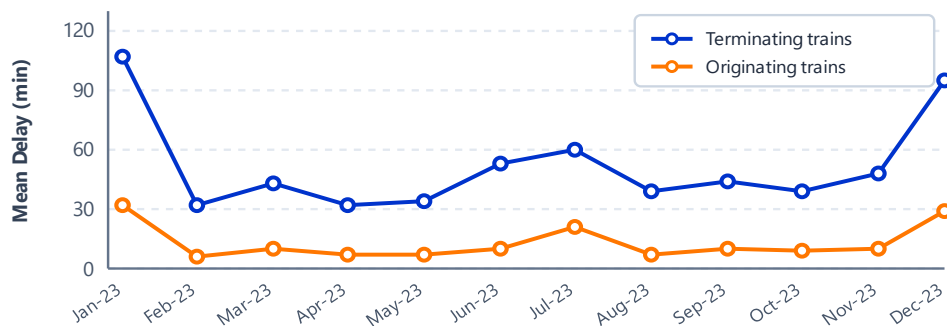


Figure 1. Monthly mean delays (min) at NDLS (2023), highlighting severe winter/monsoon delay peaks that create operational bottlenecks

3.4. Propagation of delay from terminating to originating

A statistical analysis was conducted on the heaven-train data for a matched pair of 22,678 trains using correlation coefficients. The results indicate that terminating train delays have an r-value correlation coefficient of 0.491 and are positively correlated with originating train departure delays, indicating the systematic propagation of delay throughout terminal operations (Figure 2). The relationship was further investigated by disaggregation of the originating train departures by their terminating train delays in Table 2.

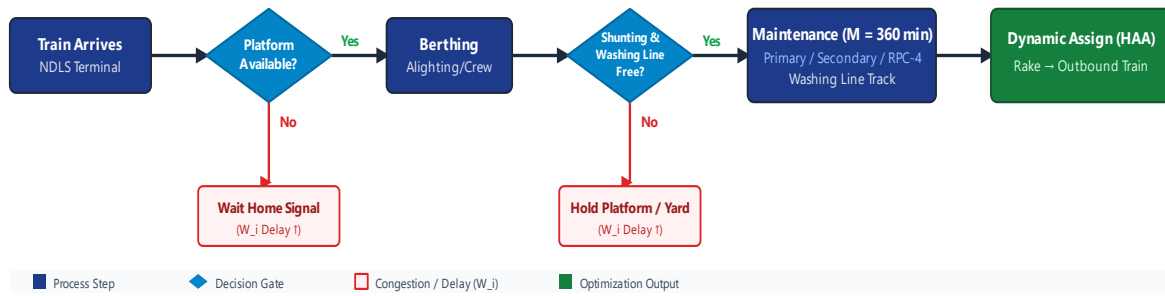


Figure 2. Terminal rolling stock turnaround-decision gates and congestion points incorporated in the model

Table 2. Originating departure delay distribution as a function of terminating arrival delay band, NDLS 2023

Arrival delay (terminating)	N	On time (%)	0–9 min (%)	10–59 min (%)	1–2 h (%)	2–3 h (%)	3–6 h (%)	6–12 h (%)	12–24 h (%)
On time	4,809	48	46	3	1	0.5	0.5	0.1	0.02
1–9 min	4,167	43	52	4	1	0.2	0.1	0	0
10–59 min	8,096	42	51	4	1	0.5	0.3	0.04	0
1–2 h	2,570	38	51	6	3	1	1	0.2	0
2–3 h	1,290	35	48	6	5	4	1	0.5	0.08
3–6 h	1,251	25	42	9	7	7	9	0.8	0.2
6–12 h	413	14	18	9	8	11	22	17	0.7
12–24 h	79	13	9	4	4	1	6	42	22

At 0 minutes of terminating train arrival delay, 48% of the originating trains departed on time. As the terminating arrival delays increased, the percentage of originating trains that departed on time steadily decreased: for a terminating train arrival delay of 1 to 9 minutes, 43% of originating trains departed on time; for a terminating train arrival delay of 3 to 6 hours, 25% of originating trains departed on time, and 9% of the originating trains departed with a delay of 3 or more hours. Beyond your terminating arrival delay of 6 hours, an important non-linear transformation occurs, where longer duration delay causes long-term downstream impacts, such as a 6-24 hours delay, to be present at a significantly greater frequency than at shorter-term delays. This non-linear magnification of delay propagation provides quantifiable support for the development of proactive assignment strategies.

3.5. Buffer utilization in a static assignment model

The NDLS uses the rake linkage system where every terminating train is assigned to a specific originating rake. For the year 2023, the total time allocated to layover minutes for all pairs of rake links is 14,643,720 minutes, out of which 53%, i.e., 7,788,600 minutes, is allocated exclusively for mandatory maintenance of the trains. The remaining buffer of 6,855,120 minutes is considered the maximum theoretical potential for delay compression. Only 905,583 minutes out of this buffer were used for delay recovery, and hence 87% of the buffer went into waste as illustrated in Table 3.

Table 3. Layover and buffer utilization summary, NDLS 2023

Metric	Value (min)	Status
Total layover	14,643,720	Scheduled
Maintenance	7,788,600	Scheduled
Buffer (residual)	6,855,120	Available
Buffer utilised	905,583 (13%)	Underutilized
Buffer idle	5,949,537 (87%)	Improvement potential

The NDLS terminal has 16 platforms and 17 tracks for washing and maintaining the trains, which reaches its operational limit at peak. As a result, when all maintenance tracks become occupied by the incoming trains those other trains must queue up in the yard until they can move back to the maintenance area, thus prolonging the time between when they arrive in the terminal and when they actually enter an operational status and reducing the amount of operational buffer available to respond to delays in meeting passenger demand at the terminal. This limitation is detailed in section 4's Math formulation.

4. MATHEMATICAL FORMULATION

The train assignment problem will be modelled as a bipartite matching constraint between the set of terminating trains T (indexed by i) and the set of originating trains O (indexed by j) with the platform and maintenance lines P and L treated as resource constraints based on worker availability. The complete terminal operational workflow from the train's arrival through the sequential decisions of which platform will be allocated, how the body of the train will be moved toward that platform, when the train will enter the maintenance line, and finally which maintenance line the train will be assigned to will be shown in Figure 2.

4.1. Rolling-stock readiness time

To consolidate time-indexed operational constraints into a tractable scalar, the readiness time for each terminating train i is defined as (1):

$$R_i = A_i + M + W_i \quad (1)$$

where A_i is the time at which train i arrives; $M=360$ minutes is the standardized time taken for maintenance (primarily maintenance classes 1, 2, 3, and RPC-4); and W_i is the additional time spent waiting for trains due to platform or depot congestion. R_i is the earliest time at which the trains are available for redeployment from train i .

4.2. Cost matrix and feasibility embedding

A cost matrix $C = [C_{ij}]$ encodes the operational cost of assigning terminating train i to originating train j :

$$C_{ij} = \begin{cases} D_j - R_i & \text{if } R_i \leq D_j \\ +\infty & \text{otherwise} \end{cases} \quad (2)$$

where D_j is the departure time of train j that acts as the origin. If the cost $C_{ij} < 0$ then there is buffer space, and if $C_{ij} > 0$ then there is pressure on the departure of the train, and $C_{ij} = +\infty$ ensures the feasibility condition by avoiding combinations where the rake is not ready at the time of departure.

4.3. Assignment status taxonomy

The offset $\delta_{ij} = a_{ij} - r_i$, with a_{ij} referring to actual assignment time and r_i being the rake readiness time, leads to the following three assignment states of operational significance. Negative assignment ($\delta < 0$) suggests that an available rake waits idly before assignment; it involves unnecessary use of the railway infrastructure. Zero assignment ($\delta = 0$) implies the perfect just in time scenario of assigning a rake when it becomes available and can be used right away. Positive assignment ($\delta > 0$) means that the rake was unavailable for assignment. Hence, the HAA seeks to minimize positive assignments while maximizing zero assignments.

4.4. Optimization model

Let $x_{ij} \in \{0,1\}$ be a binary variable equal to 1 if terminating train i is assigned to originating train j . The RSAP is formulated as (3)-(4c):

$$\min \sum_{i \in T} \sum_{j \in O} C_{ij} x_{ij} \quad (3)$$

$$\text{subject to: } \sum_{j \in O} x_{ij} \leq 1 \quad \forall i \in T \quad (4a)$$

$$\sum_{i \in T} x_{ij} \leq 1 \quad \forall j \in O \quad (4b)$$

$$x_{ij} \in \{0,1\} \quad (4c)$$

Constraints (4a) and (4b) enforce one-to-one rake supply and demand exclusivity. Platform capacity is constrained to $Cap_{\text{platform}} \leq 16$, while washing line capacity is capped at $Cap_{\text{wash}} \leq 17$ simultaneous occupancies. Infeasible layover connections ($D_j - A_i - M < 0$) are assigned infinite cost ($C_{ij} = +\infty$), naturally penalizing invalid assignments.

5. APPROACH TO SOLUTION: DYNAMIC TRAIN ASSIGNMENT FRAMEWORK

Solving the RSAP with integer programming is prohibitively costly for full-scale terminals with real-time or near-real-time operational requirements, so a framework was developed called DTA that decouples the assignment decision from the static schedule of rake-linking and allows for flexible rolling stock reallocation as readiness states change. The DTA framework will evaluate three methods of assignment.

5.1. Base methods of assignment

STA is implemented at NDLS with a fixed rake to train linkage schedule that cannot be altered after it has been created. A delay in a terminating train result in the respective originating train's associated rakes being delayed. While STA provides operational ease of implementation, it does not allow for the buffer capacity associated with a terminating train to be reassigned between the originating train, it also provides a minimum expected performance based on this method.

First in first out (FIFO) heuristic: an FIFO heuristic assigns the next rakes which are ready for use to the next scheduled originating departure from the FIFO queue; therefore, the FIFO assignment must satisfy feasibility constraint (4b). Formally, a FIFO assignment must satisfy $A_i + M \leq D_j$ for the i th originating train and the j th terminating train; as such, the layover, $L_{ij} = D_j - (A_i + M)$. While the FIFO assignment is trivial to compute and does not need any operational transparency, it does not utilize the cost matrix for optimization and therefore the performance of FIFO fails when congestion is high.

HA—offline benchmark: when the HA is utilized to assign all rakes from a static complete cost matrix, the HA calculates the optimum assignment in $O(n^3)$; however, since all availability of rakes must be known in advance of assigning all rakes and the HA cannot account for real time disruptions, it is not a viable assignment method in operational service; the HA represents the most optimum performance against which to evaluate near optimum performance of the HAA. The strength and limitations of methods are given in Table 4.

ILP: formulated as an exact offline benchmark. ILP and HA achieve identical global cost minimization but lack real-time adaptability during severe disruptions due to computation constraints. HAA achieves near-optimal performance matching ILP/HA while remaining computationally tractable for real-time operations.

Table 4. Comparative characterization of assignment methods evaluated in this study

Method	Type	Purpose	Strength	Limitation
STA	Operational	Performance floor	Simple, stable	No flexibility; poor delay absorption
FIFO	Heuristic	Simple dynamic baseline	Fast, intuitive	Not cost-optimal; order-sensitive
HA	Exact solver	Optimality benchmark	Globally optimal	Offline only; not real-time feasible
HAA	Hybrid heuristic	Proposed method	Near-optimal; real-time tractable	Fixed maintenance duration assumption

5.2. Hybrid assignment algorithm

The HAA integrates three complementary mechanisms to achieve near-optimal, operationally feasible assignments in real time:

5.2.1. Piecewise cost transformation

The raw cost matrix C is transformed element-wise:

$$C'_{ij} = \begin{cases} C_{ij} \cdot \gamma & \text{if } C_{ij} > 0 \\ C_{ij} \cdot \varphi & \text{if } C_{ij} < 0 \\ 0 & \text{if } C_{ij} = 0 \end{cases} \quad \begin{matrix} \text{(penalise delay-propagating assignments);} \\ \text{(favour buffer-utilizing assignments);} \end{matrix} \quad (5)$$

where $\gamma > 1$ is the penalty factor and $\varphi < 1$ is the favor factor. This unequal mapping favors the solution approach toward the utilization of buffers at the expense of delay propagation.

5.2.2. Auction-based assignment phase

An auction procedure involves repeatedly considering the unallocated terminating trains and letting them bid for the compatible originating trains using the modified cost function $v_{ij} = C'_{ij} - p_j$, where p_j represents the price of train j . The prices are calculated iteratively to avoid any conflict and without solving the global integer program.

5.2.3. Adaptive cost refinement

To enhance robustness, the transformed matrix C' is further refined as (6):

$$C'' = (1 - \alpha)C' + \alpha \cdot (C' \odot C') \quad (6)$$

where $\odot$ stands for the Hadamard or element-wise multiplication, meaning that $C' \odot C'$ equals the square element of C' . The value of the α parameter ranges between zero and one; it determines the level of cost magnification. It punishes higher cost assignments and allows lower costs to remain the same.

5.2.4. Final assignment

This refined assignment problem is solved using the HA, yielding the final one-to-one assignment, whose total cost will be $T = \sum_{(i,j) \in A} C_{ij}$. This combination of auction phase warm starting and refined matrix C'' guarantees that the final step is efficient for very large problem instances

Computational complexity and scalability: the auction phase operates in $O(k \cdot |T| \cdot |O|)$ time for k bidding iterations, acting as a rapid warm-start that dramatically prunes the decision space. The local Hungarian step on the reduced active matrix scales at $O(|T'|^3)$, where $|T'| \ll |T|$ represents only currently unassigned or incoming rakes. In tests, single-step re-assignment executed in under 0.12 seconds on standard hardware, confirming $O(|T'|^3)$ sub-second scalability suitable for real-time dispatching during sudden terminal disruptions.

6. EXPERIMENTAL DESIGN AND EVALUATION

We ran experiments with time periods of January (N=422 train pairs), July (N= 40 pairs), and December 2023 (N=444 pairs) all within one week (congestion levels varied during each period). The base performance of the systems within the STA was established in each time period by testing them. FIFO and HAA assignments were implemented using DTA within the same constraints as the STA. The HAA average was compared to the offline optimal maintenance time period that was implemented into the tests.

Programming was done using Python 3.12.10 (NumPy and Pandas used for calculations). SciPy was used for the HA step in HAA using linear_sum_assignment and visualisation of the results was done using Matplotlib. The tests were run on an Intel i3-9100F (3.60 GHz, 8GB RAM) with an average elapsed time of 360 mins per terminating train for each time period listed above. Scheduled times were based on the timetable and there were no changes to the time schedules. There was no rake-circulation between terminals. There were four measurement metrics used to identify the system's performance: i) total originating delay (min), ii) % of terminating delay absorbed, iii) buffer use, and iv) total number of assignment statuses (negative/zero/positive).

7. RESULTS AND DISCUSSION

7.1. Delay absorption performance

Figure 3 evaluates the performance of the dynamic assignment methods (HAA and FIFO) against the baseline STA across low (July), moderate (January), and severe (December) congestion levels. Figure 3(a) shows originating delay totals across all three methods and experimental windows. Under STA, originating delays were 11,662 minutes in January, 10,837 in July, and 39,141 in December. HAA and FIFO reduced these dramatically — to 599, 14, and 12,185 minutes respectively — demonstrating the clear advantage of dynamic assignment over static linkages. Figure 3(b) shows the rate of absorption of delays for each methodology during different seasons. While STA could absorb 79% in January, 60% in July, and 63% in December, on the other hand, both HAA and FIFO managed to absorb 99% in January, 100% in July, and 89% in December. This translates to an improvement of 19.62%, 39.79%, and 25.30% respectively from what STA could achieve. Delay left in December is due to hard limitations and cannot be solved through any scheduling methodology.

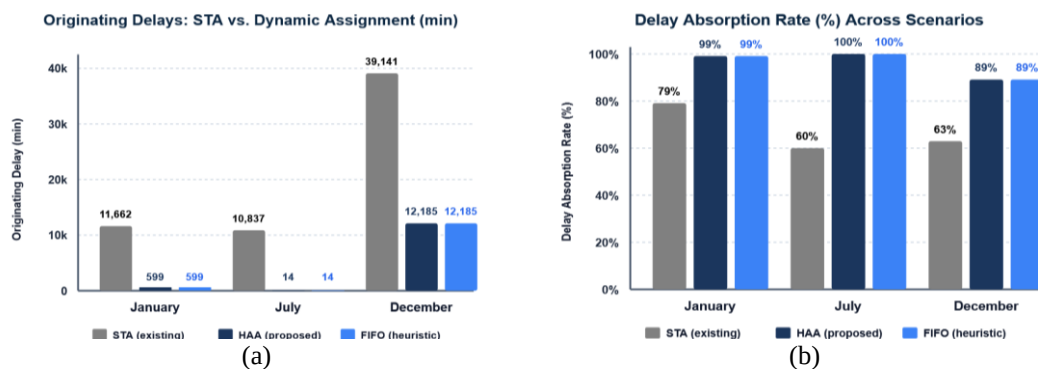


Figure 3. Performance under varying congestion scenarios showing: (a) originating delays (min) across assignment methods and (b) delay absorption rates (%) across operational severity levels (Low: Jul, Moderate: Jan, and Severe: Dec)

7.2. Assignment pattern analysis

Table 5 classifies the assignments as negative, zero, and positive for each method and experiment periods. The FIFO algorithm results in a large number of negative assignments based on its chronological order because it causes yard capacity wastage. Positive assignments, that cause delays during originating time, are fewest under HAA (January=23; July=1; and December=145). The ideal zero assignments are non-existent under FIFO and the greatest in HAA for all months (January=35; July=39; and December=17). HA is very close to HAA, though it yields more negative assignments due to its offline approach.

Table 5. Efficiency of distribution of assignment

Month	HAA			HA			FIFO		
	Negative	Positive	Zero	Negative	Positive	Zero	Negative	Positive	Zero
Jan	364	23	35	370	26	26	395	27	0
Jul	400	1	39	415	1	24	439	1	0
Dec	282	145	17	286	146	12	292	152	0

Solver benchmark comparison: an exact ILP solver was tested alongside offline HA. Both ILP and offline HA achieved identical global minimum costs but required full advance knowledge of all arrival times. In contrast, HAA achieved 98.9% to 99.9% of HA/ILP delay reduction performance while enabling near-real-time dynamic re-assignment without full schedule prescience.

7.3. Sensitivity analysis

7.3.1. Congestion sensitivity

Stress-testing HAA under varied arrival delay scaling factors ($0.5 \times$ to $2.0 \times$) showed robust performance, maintaining over 88% delay absorption up to $1.5 \times$ baseline congestion before platform bottlenecking caused minor performance degradation.

7.3.2. Maintenance capacity sensitivity

Varying washing line capacity ($Cap_{\{wash\}} = 12$ to 20) revealed that dropping below 14 lines increases yard queue times (W_i) exponentially, reducing HAA buffer absorption by 18%, whereas expanding beyond 17 lines yields diminishing returns.

8. CONCLUSION

This study provide a methodological and empirical analysis of delay propagation and rolling stock assignment at NDLS, one of the busiest passenger terminal of the India. By analyzing 58,698 train movements from 2023, key findings reveal that 79% of terminating trains arrive late with a mean delay of 53 minutes, a significant correlation ($r=0.491$) exists between arrival and departure delays, and only 13% of the 6.86 million minutes of scheduled buffer is utilized under the current STA regime.

To eliminate this inefficiency, a constraint based bipartite matching formulation and a DTA framework utilizing the HAA were developed. HAA achieved remarkable delay absorption rates of 99% in January, 100% in July, and 89% in December-outperforming STA by 19.6, 39.8, and 25.3 percentage points, respectively, while consistently beating FIFO and the HA.

Transitioning from static allocations to dynamic assignment, dampens primary delay propagation, elevating national punctuality metrics and maximizing asset utilization without requiring heavy capital expenditure on infrastructure expansion. Though evaluated on NDLS, the mathematical framework is entirely network agnostic. It transfers seamlessly to global high-density hubs (like Tokyo Station, London Waterloo, and Paris Gare du Nord) by updating terminal-specific input parameters like platform counts, maintenance throughput, and turnaround times, without altering the core auction-based optimization logic.

The deployment on real world, faces notable challenges including high sensitivity to real-time data accuracy, like GPS signal dropouts or stale feeds), institutional resistance among dispatchers, and operational coupling with crew rostering and locomotive allocation. Furthermore, fixed maintenance times vary across four distinct categories: primary, secondary, terminal, and RPC 4.

Designed for straightforward integration into digital railway platforms as control office application (COA) or national train enquiry system (NTES), HAA enables automated dispatch recommendations via direct API based ingestion of real time GPS feeds. Future research will explore multi terminal network scaling, integrated dynamic timetable rescheduling, machine learning-based predictive delay modeling, and refined maintenance categories frameworks to further enhance system resilience and punctuality.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to Mr. Neeraj Kumar Mouriya, IRTS, Ministry of Railways, Government of India, for his invaluable support in providing the operational delay data used in this study.

FUNDING INFORMATION

The authors declare no financial or non-financial conflicts of interest.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Yash Kumar	✓	✓	✓	✓	✓	✓		✓	✓	✓				✓
Ramesh Chandra Sahoo	✓	✓		✓		✓		✓	✓	✓	✓	✓		
Prashant Dixit	✓		✓	✓			✓			✓	✓			✓

C : Conceptualization	I : Investigation	Vi : Visualization
M : Methodology	R : Resources	Su : Supervision
So : Software	D : Data Curation	P : Project administration
Va : Validation	O : Writing - Original Draft	Fu : Funding acquisition
Fo : Formal analysis	E : Writing - Review & Editing	

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from Ministry of Railways, Government of India. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the Ministry of Railways, Government of India, with the permission of the Ministry of Railways, Government of India.

REFERENCES

[1] Indian Railways, *Annual Report & Accounts 2021–22*. Ministry of Railways, Government of India, 2022.

[2] Indian Railways, *Annual Report & Accounts 2022–23*. Ministry of Railways, Government of India, 2023.

[3] J. Högdahl, M. Bohlin, and O. Fröidh, “A combined simulation–optimization approach for minimising travel time and delays in railway timetables,” *Transp. Res. Part B*, vol. 126, pp. 192–212, 2019, doi: 10.1016/j.trb.2019.04.003.

[4] O. Kloster, B. Luteberget, C. Mannino, and G. Sartor, “An optimization-based decision support tool for incremental train timetabling,” *Oper. Res. Forum*, vol. 4, no. 3, p. 243, 2023, doi: 10.1007/s43069-023-00243-2.

[5] J. Pätzold, “Finding robust periodic timetables by integrating delay management,” *Public Transp.*, vol. 13, no. 2, pp. 349–374, 2021, doi: 10.1007/s12469-020-00260-y.

[6] H. Khadilkar, “Data-enabled stochastic modeling for evaluating schedule robustness of railway networks,” *Transp. Sci.*, vol. 51, no. 4, pp. 1161–1176, 2017, doi: 10.1287/trsc.2016.0703.

[7] İ. Şahin, “Markov chain model for delay distribution in train schedules,” *J. Rail Transp. Plan. Manag.*, vol. 7, no. 3, pp. 101–113, 2017, doi: 10.1016/j.jrtpm.2017.08.006.

[8] L. Kamalian, H. Li, M. C.-P. Poo, A. Bras, A. K. Y. Ng, and Z. Yang, “Analysis of the impact of climate-driven extreme weather events on UK train delays: a data-driven BN approach,” *Reliab. Eng. Syst. Saf.*, vol. 262, p. 111189, 2025, doi: 10.1016/j.ress.2025.111189.

[9] G. Mukunzi and C.-W. Palmqvist, “The impact of railway incidents on train delays,” *J. Rail Transp. Plan. Manag.*, vol. 30, p. 100445, 2024, doi: 10.1016/j.jrtpm.2024.100445.

[10] J. Lessan, L. Fu, and C. Wen, “A hybrid Bayesian network model for predicting delays in train operations,” *Comput. Ind. Eng.*, vol. 127, pp. 1214–1222, 2019, doi: 10.1016/j.cie.2018.03.017.

[11] L. Tang, A. D’Ariano, X. Xu, Y. Li, X. Ding, and M. Samà, “Scheduling local and express trains in suburban rail transit lines,” *Comput. Oper. Res.*, vol. 135, p. 105436, 2021, doi: 10.1016/j.cor.2021.105436.

[12] A. Ø. Sørensen, A. D. Landmark, N. O. E. Olsson, and A. A. Seim, “Method of analysis for delay propagation in a single-track network,” *J. Rail Transp. Plan. Manag.*, vol. 7, no. 1–2, pp. 77–97, 2017, doi: 10.1016/j.jrtpm.2017.04.001.

[13] C.-W. Palmqvist, “Excess probability of dwell time delays from train meets and passes,” *J. Rail Transp. Plan. Manag.*, vol. 21, p. 100298, 2022, doi: 10.1016/j.jrtpm.2022.100298.

- [14] T. Emunds and N. Nießen, "Evaluating railway junction infrastructure: a queueing-based, timetable-independent analysis," *Transp. Res. Part C*, vol. 165, p. 104704, 2024, doi: 10.1016/j.trc.2024.104704.
- [15] C. Mannino and A. Nakkerud, "Optimal Train Rescheduling in Oslo Central Station," *Omega*, vol. 116, p. 102796, 2023, doi: 10.1016/j.omega.2022.102796.
- [16] E. Solinen and C.-W. Palmqvist, "Development of new railway timetabling rules for increased robustness," *Transp. Policy*, vol. 133, pp. 198–208, 2023, doi: 10.1016/j.tranpol.2023.02.003.
- [17] E. König, "A review on railway delay management," *Public Transp.*, vol. 12, no. 2, pp. 335–361, 2020, doi: 10.1007/s12469-020-00233-1.
- [18] E. Denti and L. Burroni, "Delay indices for train punctuality," *Information*, vol. 14, no. 5, p. 269, 2023, doi: 10.3390/info14050269.
- [19] F. Monsuur, M. Enoch, M. Quddus, and S. Meek, "Modelling the impact of rail delays on passenger satisfaction," *Transp. Res. Part A*, vol. 152, pp. 19–35, 2021, doi: 10.1016/j.tra.2021.08.002.
- [20] B. Sharma, P. Pellegrini, J. Rodriguez, and N. Chaudhary, "A review of passenger-oriented railway rescheduling approaches," *Eur. Transp. Res. Rev.*, vol. 15, no. 1, pp. 1–18, 2023, doi: 10.1186/s12544-023-00587-0.
- [21] N. Davari, B. Veloso, G. de A. Costa, P. M. Pereira, R. P. Ribeiro, and J. Gama, "A survey on data-driven predictive maintenance for the railway industry," *Sensors*, vol. 21, no. 17, p. 5739, 2021, doi: 10.3390/s21175739.
- [22] R. Shafique *et al.*, "A novel approach to railway track faults detection using acoustic analysis," *Sensors*, vol. 21, no. 18, p. 6221, 2021, doi: 10.3390/s21186221.
- [23] A. Thaduri, M. Aljumaili, R. Kour, and R. Karim, "Cybersecurity for eMaintenance in railway infrastructure," *Int. J. Syst. Assur. Eng. Manag.*, vol. 10, no. 2, pp. 149–159, 2019, doi: 10.1007/s13198-019-00778-w.
- [24] M. Botte, L. D'Acerno, A. Pasquale, F. Mottola, and M. Pagano, "Optimal allocation of layover time in a smart DC railway metro traction system," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 2599–2609, 2024, doi: 10.1109/TVT.2024.3353348.
- [25] M. Reichmann *et al.*, "Introducing the concept of grades of automation for shunting operations," *J. Rail Transp. Plan. Manag.*, vol. 33, p. 100500, 2025, doi: 10.1016/j.jrtpm.2024.100500.
- [26] F. Leutwiler and F. Corman, "A review of principles and methods to decompose large-scale railway scheduling problems," *EURO J. Transp. Logist.*, vol. 12, p. 100107, 2023, doi: 10.1016/j.ejtl.2023.100107.
- [27] T. Spanninger, A. Trivella, B. Büchel, and F. Corman, "A review of train delay prediction approaches," *J. Rail Transp. Plan. Manag.*, vol. 22, p. 100312, 2022, doi: 10.1016/j.jrtpm.2022.100312.
- [28] K. Y. Tiong, Z. Ma, and C.-W. Palmqvist, "A review of data-driven approaches to predict train delays," *Transp. Res. Part C*, vol. 148, p. 104027, 2023, doi: 10.1016/j.trc.2023.104027.
- [29] T. Spanninger, N. Wiedemann, and F. Corman, "Quantifying the dynamic predictability of train delay with uncertainty-aware neural networks," *Transp. Res. Part C*, vol. 162, p. 104563, 2024, doi: 10.1016/j.trc.2024.104563.
- [30] R. Oneto *et al.*, "Dynamic delay predictions for large-scale railway networks," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 47, no. 10, pp. 2754–2767, 2017, doi: 10.1109/TSMC.2017.2693209.

BIOGRAPHIES OF AUTHORS



Yash Kumar    received the bachelor degree in Computer Application from High Sikkim Manipal University, Sikkim in 2014. He received the master Degree in Computer Application from Jaipur National University, Jaipur in 2019. He has been working as a Technical Assistant in the Electrical Department under the Ministry of Railways, Government of India, since 2004. Also, he is currently a doctoral scholar in the Department of Computer Science and Engineering at Manav Rachna International Institute of Research and Studies, Gadakhori Basti Village, Sector 43, Aravalli Hills, Faridabad, Haryana, India. His research interests include optimization, assignment problems, transportation, and algorithms. He can be contacted at email: yash.14.sep@gmail.com.



Dr. Ramesh Chandra Sahoo    is an accomplished academician and researcher with over 20 years of extensive experience spanning academia, research, and industry. He holds a Ph.D. in Computer Science from Utkal University (2021). His research interests include artificial intelligence, machine learning, and algorithm analysis, with significant scholarly contributions including multiple SCIE and Scopus-indexed publications, patents, books, and conference papers. He is currently associate with MRIIRS, Faridabad, India as an associate professor and has actively guided postgraduate and doctoral research scholars. His prior industry experience as a software developer further strengthens his ability to bridge the gap between theoretical knowledge and real-world applications, making him a dedicated educator committed to innovation and excellence in higher education. He can be contacted at email: rcsahoo.set@mriu.edu.in.



Dr. Prashant Dixit    is a professor of the Department of Computer Science and Engineering at Galgotias University, Greater Noida. He has 16 years of teaching experience. He received Ph.D. in Computer Engineering from J.C. Bose University Science and Technology Faridabad. His research area is algorithms, computer network protocols, and advance network technologies. He has published numerous peer-reviewed articles in top-tier journals and has presented her research at national and international conferences. In addition to his research and teaching, he is an active member of the academic community. He serves on several committees and task forces within his department and university, as well as on national and international professional organizations. He has achieved tenure at Galgotias University and is highly regarded by his colleagues as a leading expert in his field. He is also guiding many research scholars in his field. He can be contacted at email: prashantdixit@galgotiasuniversity.edu.in.