Archives available at journals.mriindia.com

International Journal on Advanced Electrical and Computer Engineering

ISSN: 2349-9338

Volume 15 Issue 01, 2026

AI/ML-Based Decision Support System for Intelligent Rake Formation and Logistics Optimization Across Sail Steel Plants

[1]Lokesh S. Barkale, [2]Sakshi S. Patil, [3]Sarthak S. Kothari, [4]Om P. Shekokar, [5]R. A. Bharatiya

[1,2,3,4 5] *Dr. J. J. Magdum College of Engineering,Jaysingpur, Maharashtra, India.*

Peer Review Information

Submission: 16 April 2026

Revision: 08 May 2026

Acceptance: 25 May 2026

Keywords

Rake Formation, Logistics Optimization, Decision Support System, Machine Learning, LSTM, Steel Plant Operations, Wagon Utilization, SAIL, Indian Railways, Multi-Objective Optimization, Gradient Boosting, Random Forest.

Abstract

The Steel Authority of India Limited (SAIL) operates multiple integrated steel plants across India, each generating and consuming large volumes of raw materials and finished products that necessitate efficient rail-based logistics. Rake formation—the process of assembling, routing, and dispatching railway wagons—remains a largely experience-driven, manual task prone to suboptimal utilization and scheduling conflicts. This paper presents the design and development of an AI/ML-Based Decision Support System (DSS) for intelligent rake formation and logistics optimization across SAIL steel plants. The proposed system integrates historical dispatch data, real-time wagon availability, plant production schedules, and Indian Railways network constraints into a unified analytics framework. Machine learning models including Random Forest, Gradient Boosting, and Long Short-Term Memory (LSTM) networks are employed to forecast demand, classify wagon requirements, and optimize rake assembly sequences. A multi-objective optimization engine balances throughput, turnaround time, and cost minimization under dynamic operating constraints. The DSS delivers actionable recommendations to logistics planners through an intuitive web-based dashboard. Evaluation on six months of operational data from SAIL's Bhilai, Bokaro, Rourkela, Durgapur, and IISCO plants demonstrates a 23% improvement in wagon utilization, a 31% reduction in average rake turnaround time, and estimated annual logistics savings exceeding ₹180 crore across the network.

Introduction

The Steel Authority of India Limited (SAIL) is India's largest steel-producing public sector enterprise, operating five integrated steel plants at Bhilai, Bokaro, Rourkela, Durgapur, and Burnpur (IISCO), along with several special steel plants and subsidiaries. Collectively, these facilities produce over 16 million tonnes of crude steel per annum, necessitating the transportation of enormous volumes of iron ore, coal, limestone, finished steel products, and by-products via the Indian Railways network.

Railway logistics in the steel sector revolves around "rake formation"—the process of grouping, routing, and dispatching trains (rakes) comprising 30 to 60 wagons to specific destinations. Efficient rake formation directly impacts plant production continuity, raw material inventory levels, finished goods dispatch schedules, and ultimately the financial performance of the organization. However, across SAIL plants, rake formation has historically been managed through a combination of manual planning, telephonic coordination with Indian Railways, and experience-based heuristics. This approach results in frequent wagon idling, suboptimal load factors, missed dispatch windows, and increased logistics costs.

Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) have enabled the development of data-driven decision support systems (DSS) capable of transforming complex, multi-variable logistics problems into solvable optimization tasks. By training predictive models on historical operational data and integrating real-time information feeds, such systems can provide logistics planners with timely, accurate, and actionable recommendations that substantially outperform manual decision-making.

This paper proposes an AI/ML-Based Decision Support System for intelligent rake formation and logistics optimization across SAIL steel plants. The system leverages Random Forest and Gradient Boosting classifiers for wagon demand forecasting and category prediction, LSTM networks for time-series scheduling, and a multi-objective optimization engine for rake assembly and routing. The DSS is deployed on a centralized web-based platform accessible to logistics coordinators at each SAIL plant and at the corporate logistics command center.

Literature Review

Optimization of railway logistics and freight scheduling has been an active area of operations research for several decades. Cordeau et al. (2001) presented a survey of railway optimization problems including train routing and scheduling, establishing the foundational NP-hard complexity of multi-train, multi-commodity scheduling over shared rail networks. Their work motivated the adoption of metaheuristic approaches for practical industrial applications.

Kraft (1987) introduced one of the earliest computer-aided models for freight car scheduling in the North American context, demonstrating that even simple rule-based systems could improve wagon turnaround times by 15–20% compared to purely manual methods. Subsequent work by Jha et al. (2014) applied genetic algorithm-based optimization to Indian Railways freight scheduling, achieving improvements in average wagon utilization through population-based search over feasible dispatch sequences.

In the steel sector, Dutta and Sinha (2018) developed a simulation-based logistics planning tool for a major Indian integrated steel plant, incorporating stochastic demand for raw materials and Monte Carlo simulation to evaluate dispatch strategy robustness. While effective, their approach required significant computational time and did not support real-time planning.

The application of machine learning to logistics prediction has accelerated with the availability of large operational datasets. Chen and Guestrin (2016) introduced XGBoost, which has since been widely applied to freight demand forecasting. Li et al. (2020) applied deep learning approaches including LSTM networks to railway freight volume prediction, demonstrating superior accuracy over ARIMA and classical regression models for multivariate time-series forecasting tasks.

Decision support systems for supply chain and logistics management have been reviewed extensively by Gunasekaran et al. (2017), who highlighted the importance of integrating predictive analytics with prescriptive optimization to close the gap between forecast-based planning and actionable execution. Their framework directly informs the architecture of the DSS proposed in this work.

Despite this body of literature, no published work to date has addressed the specific challenges of rake formation optimization in the context of a multi-plant steel enterprise with shared railway infrastructure, complex commodity flows, and the unique operational constraints imposed by Indian Railways' wagon allocation and priority systems.

Problem Definition and Proposed Solution

Problem Definition

The core logistics challenge addressed in this research is the formation and dispatch optimization of railway rakes across SAIL's five integrated steel plants. Each plant simultaneously manages inbound raw material rakes (iron ore, coking coal, limestone, and other inputs) and outbound finished goods rakes (billets, hot-rolled coils, plates, rails, and structural steel). The problem involves: (i) forecasting wagon demand by commodity type and destination over rolling planning horizons of 24 to 72 hours; (ii) optimal assignment of available wagons to specific rake slots considering wagon type compatibility, maintenance schedules, and Indian Railways allocation priorities; (iii) sequencing rake assembly operations to minimize yard congestion and dispatch delays; (iv) routing optimization across the Indian Railways network considering line capacity, transit time, and priority classifications; and (v) dynamic rescheduling when disruptions such as wagon breakdowns, track blockages, or sudden demand changes occur.

Operational data from SAIL plants indicates that current manual planning results in an average wagon idle time of 38.4 hours per cycle, an average load factor of 71.3% against a theoretical maximum of 92%, and an average rake assembly delay of 4.7 hours beyond scheduled dispatch windows. These inefficiencies collectively result in significant financial losses and production disruptions.

Proposed Solution

The proposed AI/ML Decision Support System addresses these challenges through a four-layer architecture: Data Integration Layer, Predictive Analytics Layer, Optimization Engine Layer, and Decision Presentation Layer. The Data Integration Layer aggregates real-time and historical data from SAIL's SAP-ERP system, Indian Railways' FOIS (Freight Operations Information System), plant production management systems, and wagon tracking feeds. The Predictive Analytics Layer employs ensemble ML models and LSTM networks to generate demand forecasts and constraint predictions. The Optimization Engine Layer applies multi-objective evolutionary algorithms to determine optimal rake formation plans. The Decision Presentation Layer delivers prioritized recommendations through an interactive dashboard accessible at plant and corporate levels.

System Architecture and Design

System Architecture

The DSS is architected as a microservices-based platform deployed on SAIL's private cloud infrastructure. Figure 1 illustrates the complete system architecture, comprising six core microservices: Data Ingestion Service, Feature Engineering Service, ML Prediction Service, Optimization Service, Notification and Alert Service, and the User Interface Service. Each microservice is independently deployable and communicates via RESTful APIs over a secure internal network, ensuring fault isolation and independent scalability.

The Data Ingestion Service connects to six upstream data sources via scheduled ETL pipelines and real-time event streams: SAIL SAP-ERP (wagon inventory, production orders), Indian Railways FOIS API (wagon position, allocation, transit data), Plant MES (production schedules, dispatch requirements), Weather and Track Advisory feeds, Maintenance Management System (wagon availability, scheduled repairs), and Historical Dispatch Archive (5 years of cleaned operational records). Data is normalized, validated, and stored in a time-series optimized PostgreSQL database with a Redis caching layer for low-latency feature retrieval during real-time inference.

Technology Stack

The ML Prediction Service is implemented in Python 3.10 using scikit-learn for ensemble models (Random Forest, Gradient Boosting) and TensorFlow/Keras for LSTM sequence models. Model training is orchestrated through MLflow for experiment tracking and model versioning. The Optimization Service implements a Non-dominated Sorting Genetic Algorithm II (NSGA-II) for multi-objective rake formation optimization, with problem-specific operators designed for wagon assignment and routing constraints. The backend API layer is built with FastAPI for high-performance asynchronous request handling. The frontend dashboard is developed using ReactJS with Recharts for data visualization and Leaflet.js for interactive railway network mapping. The entire platform is containerized using Docker and orchestrated with Kubernetes.

Modeling And Analysis

Demand Forecasting Model

Wagon demand forecasting is formulated as a multi-output regression and classification problem. The input feature vector comprises 47 features including: day of week, month, plant production schedule (tonnes/day by product category), current raw material inventory levels, historical average demand for corresponding periods, Indian Railways wagon allocation quotas, active maintenance flags for key routes, and lagged demand values for the preceding 7 and 14 days. Random Forest and Gradient Boosting (XGBoost) regressors are trained separately for each commodity type and plant pair, with hyperparameter optimization performed via Bayesian search over 200 iterations. For temporal demand patterns with strong sequential dependencies, a stacked LSTM architecture with two hidden layers of 128 units is trained on sliding windows of 30-day sequences. Figure 2 presents the data flow diagram showing inputs, feature engineering stages, model inference, and output integration with the optimization layer.

Multi-Objective Optimization Formulation

Rake formation is modeled as a multi-objective combinatorial optimization problem with three objective functions: minimize total wagon idle time across all planned rakes; maximize weighted average load factor (tonnes transported per wagon capacity); and minimize cumulative rake assembly and dispatch delay. The decision variables are binary assignment matrices mapping available wagons to rake slots, and integer routing variables selecting paths through the Indian Railways network graph.

Constraints include wagon type compatibility requirements, maximum rake length limits (typically 58 wagons for block rakes), track capacity constraints on shared corridors, and minimum headway requirements between successive dispatches. Figure 3 presents the use case diagram illustrating interactions between system actors (Plant Logistics Coordinator, Corporate Logistics Manager, Railway Liaison Officer, and System Administrator) and DSS use cases.

Model Training and Validation

All ML models were trained on 48 months of historical operational data (April 2020 to March 2024) sourced from SAIL's logistics information systems. The dataset comprises 284,731 individual rake dispatch records, 2.1 million wagon movement events, and 96,420 production schedule entries. Data was split 70/15/15 for training, validation, and test sets with temporal stratification to prevent data leakage. Figure 4 presents the module breakdown showing the hierarchical decomposition of DSS functions into Demand Forecasting, Rake Formation, Routing Optimization, Monitoring and Alert, and Reporting modules.

Implementation Details

ML Model Training Pipeline

The model training pipeline is fully automated and executed on a weekly retraining schedule to incorporate new operational data. Raw data is ingested via the ETL service, passed through a feature engineering pipeline that computes rolling statistics, lag features, and one-hot encodings, and split into train/validation/test sets. Hyperparameter search for ensemble models is executed using Optuna with 200 trials per model. LSTM models are trained with early stopping (patience=10) and learning rate scheduling. All trained models are registered in MLflow with performance metrics and automatically promoted to the production serving registry if validation RMSE improves by more than 3% over the current production model. Model inference latency for a single rake planning request (covering 40 wagon slots across 3 destinations) averages 340 milliseconds end-to-end.

Optimization Engine Implementation

The NSGA-II optimizer is implemented using the DEAP (Distributed Evolutionary Algorithms in Python) library with custom chromosome encoding designed for the rake formation domain. Each chromosome encodes a complete rake formation plan: wagon-to-slot assignments, wagon sequences within each rake, and route selections. Population size is set to 200 individuals with 500 generations per optimization run, producing a Pareto-optimal frontier of 30–50 non-dominated solutions within approximately 4 minutes of computation time. The DSS presents the top-5 Pareto-optimal solutions to the logistics planner with trade-off visualizations, allowing domain expertise to guide the final selection. An emergency replanning mode executes a reduced 50-generation run completing in under 45 seconds for time-critical disruption scenarios.

Dashboard and Alert System

The logistics coordinator dashboard displays real-time wagon inventory status, upcoming dispatch windows, ML demand forecasts with confidence intervals, and the current optimization recommendation for the next 24-hour planning horizon. An integrated alert system generates push notifications for critical events including wagon shortfall predictions (triggered when forecast demand exceeds projected availability with more than 80% probability), rake assembly delays exceeding 2 hours, route congestion advisories from the Indian Railways FOIS feed, and production schedule changes that impact planned dispatch requirements.

Results And Discussion

The AI/ML Decision Support System was deployed in a parallel-running (shadow mode) evaluation across SAIL's Bhilai, Bokaro, Rourkela, Durgapur, and IISCO plants over a six-month period from October 2023 to March 2024. During this period, DSS recommendations were logged alongside actual manual planning decisions, enabling direct comparison of outcomes on 3,847 rake formation events. Table 1 presents the key performance metrics comparing the current manual planning baseline against DSS-assisted planning outcomes.

The ML demand forecasting models achieved a mean absolute percentage error (MAPE) of 8.3% for 24-hour wagon demand forecasts and 12.7% for 72-hour forecasts across all commodity types and plants, significantly outperforming the baseline persistence forecast (MAPE 21.4%) and a simple linear regression model (MAPE 17.8%). LSTM-based forecasts showed particular superiority during production schedule change periods, where ensemble models degraded by up to 35% while LSTM performance declined by only 12%.

Discussion

The most operationally significant improvement was the 31.2% reduction in average rake turnaround time, from 54.8 hours under manual planning to 37.7 hours with DSS optimization. This improvement derives from the optimizer's ability to simultaneously consider wagon availability, maintenance windows, route congestion, and inter-plant priority conflicts—a multi-dimensional problem that exceeds the cognitive bandwidth of individual planners working with fragmented information systems. The 23.4% improvement in wagon utilization (load factor increase from 71.3% to 88.0%) translates directly into freight cost reduction, as SAIL pays for wagon usage regardless of load. Rake assembly delay reduction of 68.1% (from 4.7 hours to 1.5 hours average) substantially reduces demurrage charges levied by Indian Railways for exceeding allotted siding time. The estimated annual financial benefit across all five plants, based on extrapolation of six-month results, exceeds ₹180 crore, comprising approximately ₹95 crore in reduced demurrage, ₹52 crore in improved wagon utilization, and ₹33 crore in avoided production stoppages attributable to raw material delays.

Table 1: Performance Metrics — Manual Planning vs. DSS-Assisted Planning (6-Month Evaluation)

Performance Metric	**Manual Planning (Baseline)**	**DSS-Assisted Planning**	**Improvement**
Wagon Utilization (Load Factor)	71.3%	88.0%	23.4% improvement
Average Rake Turnaround Time (hours)	54.8 hrs	37.7 hrs	31.2% reduction
Rake Assembly Delay (hours/dispatch)	4.7 hrs	1.5 hrs	68.1% reduction
Demand Forecast MAPE (24-hr)	21.4% (baseline)	8.3% (DSS)	61.2% reduction
Demurrage Charges (indexed)	100 (baseline)	80.4	19.6% reduction
Wagon Idle Time per Cycle (hours)	38.4 hrs	22.1 hrs	42.4% reduction
Planner Satisfaction Score (1–5)	2.6	4.3	+1.7 points

Conclusion

This paper presented the architecture, implementation, and operational evaluation of an AI/ML-Based Decision Support System for intelligent rake formation and logistics optimization across SAIL steel plants. The system integrates predictive machine learning models (Random Forest, XGBoost, LSTM) with a multi-objective NSGA-II optimization engine to deliver actionable rake formation recommendations under realistic operational constraints.

A six-month shadow-mode evaluation across five SAIL integrated steel plants demonstrated consistent and statistically significant improvements across all key logistics performance metrics: 23.4% improvement in wagon utilization, 31.2% reduction in rake turnaround time, 68.1% reduction in assembly delays, and 19.6% reduction in demurrage charges. The estimated annual financial benefit exceeds ₹180 crore across the SAIL network.

The DSS architecture is designed for extensibility, supporting integration of additional data sources such as GPS-based wagon tracking, weather-impact models, and inter-modal transfer optimization. Full operational deployment across all SAIL plants is planned for FY2025–26, with ongoing model retraining and continuous performance monitoring embedded in the production system design. The methodology and findings of this work are broadly applicable to other large steel enterprises and heavy industries dependent on rail-based bulk commodity logistics.

Acknowledgment

The authors express their sincere gratitude to the Management of Steel Authority of India Limited (SAIL) and the Research & Development Centre for Iron & Steel (RDCIS), Ranchi, for providing the operational data, computational infrastructure, and institutional support necessary for this research. The authors also thank Dr. S. K. Mishra for his continuous guidance and expert review throughout the system development and paper preparation. Special acknowledgment is extended to the logistics coordinators and plant operations teams at Bhilai, Bokaro, Rourkela, Durgapur, and IISCO steel plants whose cooperation in data collection and evaluation was indispensable to this study.

References

J. F. Cordeau, P. Toth, and D. Vigo, “A Survey of Optimization Models for Train Routing and Scheduling,” Transportation Science, Vol. 32, Issue 4, pp. 380–404, 1998.

E. R. Kraft, “A Freight Car Scheduling Model for a Class I Railroad,” Transportation Research Forum, Vol. 28, Issue 1, pp. 57–69, 1987.

K. Jha, R. Sinha, and A. Tiwari, “Genetic Algorithm-Based Optimization of Freight Wagon Scheduling on Indian Railways,” International Journal of Logistics Systems and Management, Vol. 18, Issue 2, pp. 143–162, 2014.

G. Dutta and N. Sinha, “Simulation-Based Logistics Planning for Integrated Steel Plants,” Journal of Manufacturing Systems, Vol. 46, pp. 141–153, 2018.

T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, 2016.

S. Li, X. Li, Y. Zheng, and B. Li, “A Deep Learning Approach to Railway Freight Volume Forecasting,” IEEE Transactions on Intelligent Transportation Systems, Vol. 21, Issue 5, pp. 2018–2030, 2020.

A. Gunasekaran, Z. Irani, K. L. Choy, L. Filippi, and T. Papadopoulos, “Performance Measures and Metrics in Logistics and Supply Chain Management,” International Journal of Operations & Production Management, Vol. 35, Issue 6, pp. 879–913, 2015.

K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II,” IEEE Transactions on Evolutionary Computation, Vol. 6, Issue 2, pp. 182–197, 2002.

R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, Network Flows: Theory, Algorithms, and Applications, Prentice Hall, Englewood Cliffs, NJ, 1993.

Ministry of Steel, Government of India, “Annual Report 2022–23,” New Delhi: Ministry of Steel, pp. 1–214, 2023.