



INTERNATIONAL JOURNAL OF TRENDS IN EMERGING RESEARCH AND DEVELOPMENT

INTERNATIONAL JOURNAL OF TRENDS IN EMERGING RESEARCH AND DEVELOPMENT

Volume 4; Issue 5; 2026; Page No. 29-38

Received: 10-07-2026
Accepted: 18-08-2026
Published: 16-09-2026

Predictive Modelling of Freight Demand and Transport Capacity under Uncertainty A Framework for Reducing Empty Vehicle Movements in Road Freight Logistics

Nazia Adil

Master, Department of Business Analytics, La Trobe University, Australia

DOI: <https://doi.org/10.5281/zenodo.22941355>

Corresponding Author: Nazia Adil

Abstract

Empty vehicle movements represent a major source of inefficiency in road freight transportation because vehicles frequently travel without carrying economically useful loads after completing a delivery or when suitable return cargo is unavailable. Recent European statistics indicate that empty vehicle movements accounted for approximately 21.6% of vehicle-kilometres performed by EU-registered road freight vehicles in 2024, with the proportion varying substantially across countries and transport markets. Although empty running is partly caused by structural factors such as geographical trade imbalances, vehicle compatibility, short-haul operations and delivery-time constraints, improved information about future freight demand and available transport capacity can support more effective backhaul planning. This study proposes an integrated predictive framework for forecasting freight demand and transport capacity under uncertainty and subsequently using the forecasts to reduce empty vehicle movements. The proposed framework combines historical shipment information, vehicle availability, origin–destination patterns, temporal characteristics, freight characteristics, operational constraints, market indicators and external factors. A hybrid ensemble consisting of temporal forecasting models and machine-learning algorithms is proposed to estimate both expected freight demand and available transport capacity. Prediction intervals and probabilistic forecasts are incorporated to explicitly represent uncertainty rather than relying exclusively on point predictions. The resulting forecasts are integrated with a stochastic backhaul-matching and vehicle-repositioning optimization model that minimizes empty kilometres, operating costs, unmet freight demand and service-risk penalties. A rolling-horizon decision mechanism continuously updates predictions as new shipment and vehicle information becomes available. The proposed framework contributes to the literature by connecting freight-demand forecasting with capacity forecasting and operational backhaul optimization within a single uncertainty-aware architecture. The framework is designed to provide carriers and logistics platforms with actionable decisions regarding vehicle allocation, backhaul matching and repositioning while maintaining service reliability.

Keywords: s

1. Introduction

Road freight transportation plays a critical role in modern supply chains by connecting manufacturers, warehouses, distribution centres, retailers, and customers. The increasing complexity of supply chains, e-commerce growth, time-sensitive deliveries, and geographically distributed production have created significant challenges for efficient fleet utilization. One of the major challenges faced by road freight operators is empty vehicle movement, in which vehicles travel without carrying commercially useful freight.

Empty running increases transportation costs, reduces vehicle utilization, consumes fuel, and contributes unnecessarily to transport-related emissions. Eurostat reported that empty running accounted for approximately 21.6% of vehicle-kilometres performed by EU-registered road freight vehicles in 2024, highlighting the scale of the problem in road freight operations.

Empty vehicle movements commonly occur when a vehicle completes a delivery and no suitable return shipment is available. They may also result from geographical

imbalances in freight flows, vehicle-type restrictions, incompatible cargo requirements, delivery-time constraints, or limited information about future shipment availability. Although some empty movements are unavoidable, better coordination between freight demand and available vehicle capacity can reduce a considerable portion of avoidable empty travel. Consequently, effective backhaul planning requires knowledge not only of current freight demand but also of where and when future freight is likely to become available.

Freight demand is inherently uncertain and varies across time and geographical locations. Factors such as historical shipment patterns, day of the week, seasonality, holidays, economic activity, weather, traffic conditions, and customer behaviour can influence future demand. Similarly, transport capacity is dynamic because vehicle availability depends on delivery completion times, driver schedules, maintenance, breakdowns, cancellations, and unexpected delays. Therefore, conventional planning approaches based on historical averages or deterministic assumptions may not adequately capture the uncertainty associated with future freight demand and vehicle availability.

Recent developments in machine learning and predictive analytics provide opportunities to address these limitations. Machine-learning and time-series models can identify complex temporal and spatial patterns in historical freight data and generate forecasts for future shipment volumes. However, freight-demand forecasting alone is insufficient for reducing empty movements. A predicted shipment can only be utilized if an appropriate vehicle is available at the required location and time. This creates a need for simultaneous freight-demand forecasting and transport-capacity forecasting.

An additional challenge is that forecasts themselves are uncertain. A model may predict a certain quantity of freight for a particular route, but the actual demand may differ considerably. Similarly, a vehicle predicted to become available at a particular location may be delayed or reassigned. Consequently, relying only on point forecasts can lead to inefficient vehicle assignments and increased service risk. Probabilistic forecasting provides a more appropriate approach by estimating both expected demand and the uncertainty surrounding the prediction. These probabilistic forecasts can subsequently be incorporated into optimization models to support more reliable operational decisions.

This study therefore proposes an integrated framework that combines freight-demand forecasting, transport-capacity forecasting, uncertainty estimation, backhaul matching, and stochastic vehicle optimization. The proposed framework first analyses historical shipment and vehicle data to identify temporal, spatial, and operational patterns. Machine-learning and time-series techniques are then used to predict future freight demand and available transport capacity. Prediction intervals or probability distributions are incorporated to quantify uncertainty. Based on these predictions, potential backhaul opportunities are identified by matching vehicles completing their current assignments with freight expected to become available near their destination. A stochastic optimization model subsequently determines vehicle assignments and repositioning decisions while considering vehicle capacity, time windows, route

compatibility, service requirements, transportation costs, and empty kilometres.

The framework further adopts a rolling-horizon strategy to accommodate continuously changing logistics conditions. As new shipment information, vehicle locations, delivery statuses, and operational disruptions become available, the forecasts and vehicle assignments are updated. This enables the system to move from reactive dispatching toward predictive and adaptive fleet management.

The primary objective of this research is to determine whether integrating uncertainty-aware freight and capacity forecasting with backhaul optimization can reduce empty vehicle movements while maintaining acceptable service levels. The study evaluates the proposed framework using forecasting measures such as MAE, RMSE, and WAPE, together with operational measures including empty vehicle kilometres, empty running ratio, vehicle utilization, capacity utilization, backhaul matching rate, transportation cost, and service level.

The major contribution of the study is the integration of predictive analytics and stochastic optimization for proactive empty-vehicle reduction. Rather than treating freight forecasting and vehicle routing as separate problems, the proposed framework connects future demand prediction with future vehicle availability and operational decision-making. This provides a foundation for more efficient, cost-effective, and sustainable road freight transportation under uncertain operating conditions.

2. Literature Review

2.1 Freight Demand Forecasting

Freight-demand forecasting plays an important role in transportation planning because accurate estimates of future shipments support vehicle allocation, capacity planning, scheduling, and routing decisions. Early freight forecasting research primarily used statistical and econometric approaches to identify temporal, spatial, and economic factors influencing freight flows^[1]. With the increasing availability of digital logistics data, machine learning (ML) methods have become increasingly relevant for capturing nonlinear and complex demand patterns^[2]. A recent review of logistics demand forecasting reported growing adoption of neural-network-based methods, while also identifying medium-term forecasting and operationally integrated forecasting as relatively less explored areas^[3].

Machine learning has subsequently been applied to freight demand prediction, travel-time estimation, vehicle routing, traffic prediction, arrival-time prediction, and anomaly detection^[4]. These developments demonstrate the potential of data-driven approaches for improving freight planning. However, forecasting freight demand alone does not ensure that sufficient transportation capacity will be available when and where it is required.

2.2 Machine Learning in Road Freight Transportation

The increasing availability of shipment records, vehicle-position data, historical routes, and operational information has encouraged the application of ML to road freight transportation. Recent research indicates that ML applications have expanded from prediction toward complex transportation decision-making and routing problems^[5]. Learning-based routing approaches have also demonstrated

the potential of deep learning and reinforcement learning for solving complex vehicle-routing problems^[6].

Despite these developments, practical implementation remains challenging. Many models are evaluated using benchmark datasets or simulated environments, while relatively fewer studies validate their approaches in real-world freight operations. Furthermore, prediction models are often developed independently from subsequent fleet-management decisions. This creates an opportunity to connect predictive models directly with operational decisions such as vehicle allocation, repositioning, and backhaul planning.

2.3 Empty Vehicle Movements and Fleet Management

Empty vehicle movement is a long-standing problem in freight transportation. Vehicles may return without cargo because of geographical imbalance between freight supply and demand, lack of suitable return shipments, vehicle compatibility requirements, and time-window restrictions. Dejax and Crainic^[7] established the importance of empty-flow management and emphasized the need to jointly consider loaded and empty vehicle movements in freight fleet planning.

The problem has subsequently been addressed through vehicle-routing and backhaul models. Koç and Laporte^[8] reviewed the Vehicle Routing Problem with Backhauls (VRPB), in which delivery and pickup customers are incorporated into the same routing system. Such approaches can improve vehicle utilization by combining forward deliveries with return loads. However, conventional backhaul models generally assume that shipment requirements and vehicle information are sufficiently known during planning. In practical freight operations, these conditions can change continuously because of new orders, cancellations, delays, and changing vehicle availability.

2.4 Vehicle Routing under Demand Uncertainty

Uncertainty in freight demand has motivated the development of stochastic vehicle-routing models. Gendreau *et al.*^[9] developed an exact approach for vehicle routing with stochastic customer demands, demonstrating the importance of incorporating demand variability into route planning. Subsequent research introduced safety stocks and simulation approaches to improve routing decisions under stochastic demand conditions^[10].

Further developments considered preventive returns and recourse strategies, allowing vehicles to respond when realized demand differs from initial expectations^[11]. Florio *et al.*^[12] proposed further exact approaches for vehicle-routing problems with stochastic demands. More recent research has investigated risk and ambiguity in capacitated vehicle-routing problems, particularly when the probability distribution of future demand is not completely known^[13].

These studies demonstrate that deterministic routing decisions can be inadequate when future demand is uncertain. However, much of the stochastic-routing literature focuses on route optimization after defining demand scenarios rather than integrating predictive models that estimate both future freight demand and available transportation capacity.

2.5 Forecasting and Transportation Optimization

A major recent research direction is the integration of

forecasting with transportation optimization. Sonnleitner *et al.*^[14] reviewed forecasting research in road freight transportation and found that forecasts are frequently incorporated into downstream optimization through a predict-then-optimize structure. The authors identified a disconnect between advances in forecasting and transportation optimization and emphasized the need for greater integration between these components.

Probabilistic forecasting provides an important opportunity in this context because it represents uncertainty through prediction intervals or probability distributions rather than relying exclusively on a single point prediction. In freight transportation, uncertainty may arise from shipment cancellations, unexpected orders, seasonal variations, travel delays, vehicle breakdowns, driver availability, and changes in previously assigned loads.

Recent work on dynamic empty-vehicle repositioning has further emphasized the importance of rolling-horizon decision-making, where vehicle positions and assignments are continuously updated as new information becomes available^[15]. Such approaches provide a foundation for connecting predictive information with proactive empty-movement reduction.

2.6 Research Gap

The reviewed literature reveals several important gaps. First, freight-demand forecasting and transport-capacity forecasting are generally treated as separate problems^[3, 5]. Second, many forecasting approaches provide point predictions without explicitly incorporating uncertainty into subsequent transportation decisions^[2, 14]. Third, traditional backhaul and routing models frequently assume that demand and vehicle information is known before optimization^[8, 9]. Fourth, although machine learning, stochastic routing, and dynamic repositioning have developed substantially, their integration into a unified framework for proactive empty-vehicle reduction remains limited^[6, 12, 15].

The proposed research addresses these gaps through an integrated framework combining freight-demand forecasting, transport-capacity prediction, probabilistic uncertainty estimation, and stochastic backhaul/repositioning optimization. A rolling-horizon mechanism is used to update forecasts and vehicle assignments as new operational information becomes available. The central objective is therefore to connect predicted freight requirements, available transport capacity, and proactive vehicle positioning to reduce unnecessary empty vehicle movements while maintaining transportation service requirements.

3. Research Methodology

The proposed study develops an integrated predictive-optimization framework for reducing empty vehicle movements in road freight transportation. The framework combines freight-demand forecasting, transport-capacity prediction, uncertainty estimation, backhaul matching, and vehicle repositioning. Unlike conventional approaches that forecast demand and perform routing independently, the proposed methodology connects predictive outputs directly with fleet-management decisions.

3.1 Data Collection and Pre-processing

The proposed framework uses historical freight and fleet-

operation data containing shipment, vehicle, geographical, temporal, and operational information. Important variables include shipment origin and destination, freight quantity, vehicle type, vehicle capacity, delivery and pickup times, travel distance, travel duration, vehicle availability, previous vehicle assignments, and delivery status. Temporal variables such as day of the week, month, holidays, and seasonal patterns are incorporated to capture variations in freight demand.

Data pre-processing involves missing-value treatment, duplicate removal, categorical-variable encoding, normalization where required, and identification of abnormal observations. The dataset is divided chronologically into training, validation, and testing subsets to prevent future information from influencing model training.

3.2 Freight-Demand Forecasting

The first component predicts future freight demand for different origin-destination pairs and planning periods. Let $D_{ij,t}$ denote the freight demand from location i to location j during period t . The predicted demand is represented as:

$$\hat{D}_{ij,t+h} = f(X_{ij,t}, T_t, S_{ij}, O_t)$$

Where X represents historical freight information, T represents temporal characteristics, S represents spatial information, and O represents operational factors.

Machine-learning and time-series models such as XG Boost, Light GBM, Random Forest, LSTM, and other suitable ensemble models can be evaluated. Model performance is measured using MAE, RMSE, and WAPE.

3.3 Transport-Capacity Forecasting

Demand forecasting alone cannot determine whether sufficient vehicles will be available to satisfy future requirements. Therefore, the second component predicts usable transport capacity.

Let $C_{v,t}$ represent the available capacity of vehicle v at time t . Total available fleet capacity can be expressed as:

$$C_t = \sum_{v=1}^V C_{v,t} A_{v,t}$$

where $A_{v,t}$ indicates whether vehicle v is available during period t .

The capacity prediction considers vehicle schedules, current location, previous assignments, expected delivery completion, vehicle type, maintenance status, and other operational constraints. This provides a more realistic estimate of capacity that can be allocated to future freight demand.

3.4 Uncertainty Estimation

Freight demand and transport capacity are affected by uncertainty arising from demand fluctuations, cancellations, delays, vehicle breakdowns, driver availability, and changing operational conditions. Therefore, the proposed framework uses probabilistic forecasting rather than relying exclusively on point predictions.

For each forecast period, a prediction interval can be represented as:

$$D_{ij,t}^L \leq D_{ij,t} \leq D_{ij,t}^U$$

where D^L and D^U denote the lower and upper prediction limits. These estimates are subsequently used to generate alternative demand-capacity scenarios for the optimization model.

3.5 Backhaul Matching and Vehicle Repositioning

The predicted demand and available transport capacity are provided to the optimization component. The objective is to identify suitable backhaul opportunities and reposition vehicles toward locations where future freight demand is expected.

A binary decision variable $x_{vij,t}$ is defined as

$$x_{vij,t} = \begin{cases} 1, & \text{if vehicle } v \text{ is assigned from } i \text{ to } j \text{ at time } t \\ 0, & \text{otherwise} \end{cases}$$

The optimization process seeks to minimize empty travel while controlling total transportation distance, operating cost, and service violations.

3.6 Optimization Objective

The combined objective function can be formulated as:

$$\min Z = \alpha E + \beta D + \gamma C + \delta P$$

where E represents empty vehicle kilometres, D represents total travel distance, C represents transportation cost, and P represents penalties for service violations. The coefficients $\alpha, \beta, \gamma, \delta$ determine the relative contribution of each objective.

The model incorporates vehicle-capacity, vehicle-availability, route-continuity, delivery-demand, and time-window constraints. A basic capacity constraint is:

$$\sum_{i,j} Q_{ij,t} x_{vij,t} \leq Cap_v$$

where $Q_{ij,t}$ represents the freight assigned to vehicle V , and Cap_v represents its available capacity.

3.7 Rolling-Horizon Framework

The proposed system uses a rolling-horizon approach to continuously update transportation decisions. At each planning interval, newly available shipment and vehicle information is incorporated into the forecasting models. Updated demand and capacity forecasts are generated, followed by backhaul matching and vehicle-repositioning decisions.

After the current planning period, actual operational information is collected and the forecasting and optimization process is repeated. This enables the framework to adapt to changes in freight demand, vehicle availability, delays, and other operational conditions.

The complete methodology can therefore be represented as

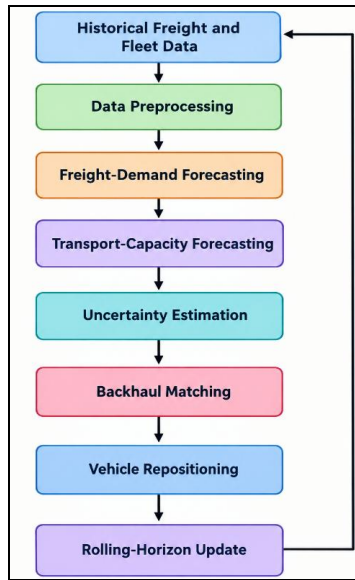


Fig 1: Proposed Integrated Framework for Freight Demand Forecasting and Empty Vehicle Movement Reduction

The proposed research framework is illustrated in Figure 1, integrating freight-demand forecasting, transport-capacity forecasting, uncertainty estimation, backhaul matching, vehicle repositioning, and rolling-horizon updating.

3.8 Performance Evaluation

The proposed framework is evaluated from both predictive and operational perspectives. Forecasting performance is measured using MAE, RMSE, and WAPE, while operational effectiveness is assessed using empty vehicle kilometres, fleet utilization, total transportation distance, transportation cost, backhaul utilization, and service-level compliance.

The integrated framework can be compared with conventional demand-only forecasting, deterministic vehicle allocation, and routing without proactive backhaul planning. This comparison allows the contribution of uncertainty-aware forecasting and predictive vehicle repositioning to be evaluated without relying solely on forecasting accuracy.

4. Proposed Mathematical Model

The proposed mathematical model integrates predicted freight demand, available transport capacity, uncertainty, and vehicle repositioning decisions to reduce empty vehicle movements in road freight logistics. The model considers a multi-period transportation network in which vehicles can perform loaded movements, empty repositioning, or backhaul assignments. Forecasted demand and capacity information are represented through multiple uncertainty scenarios, allowing routing decisions to account for variations in future operating conditions.

4.1 Sets and Parameters

Let V denote the set of available vehicles, N the set of network locations, T the set of planning periods, and S the set of uncertainty scenarios. The principal parameters are defined as follows:

- $d_{ij,t}^s$: predicted freight demand from location i to

location j during period t under scenario s .

- Cap_v : maximum carrying capacity of vehicle v .
- $A_{v,t}^s$: availability of vehicle v during period t under scenario s .
- $dist_{ij}$: transportation distance between locations i and j .
- c_{ij} : transportation cost associated with movement between locations i and j .
- p_s : probability of occurrence of scenario s .

The uncertainty scenarios are generated from the prediction intervals or probabilistic forecasts obtained during the forecasting stage.

4.2 Decision Variables

The primary vehicle-assignment decision variable is defined as:

$$x_{vij,t}^s = \begin{cases} 1, & \text{if vehicle } v \text{ performs a loaded movement from } i \text{ to } j \text{ at } t \text{ under scenario } s \\ 0, & \text{otherwise} \end{cases}$$

The empty repositioning variable is defined as:

$$r_{vij,t}^s = \begin{cases} 1, & \text{if vehicle } v \text{ is repositioned empty from } i \text{ to } j \text{ at } t \\ 0, & \text{otherwise} \end{cases}$$

The unmet-demand variable $u_{ij,t}^s$ represents freight demand that cannot be served under a particular scenario.

4.3 Objective Function

The main objective is to minimize empty vehicle movement while simultaneously controlling total transportation distance, operating cost, and unmet freight demand. The objective function is formulated as:

$$\begin{aligned} \min Z = & \alpha \sum_{s \in S} p_s \sum_{v \in V} \sum_{i,j \in N} \sum_{t \in T} dist_{ij} r_{vij,t}^s \\ & + \beta \sum_{s \in S} p_s \sum_{v \in V} \sum_{i,j \in N} \sum_{t \in T} dist_{ij} x_{vij,t}^s \\ & + \gamma \sum_{s \in S} p_s \sum_{v \in V} \sum_{i,j \in N} \sum_{t \in T} c_{ij} (x_{vij,t}^s + r_{vij,t}^s) + \delta \sum_{s \in S} p_s \sum_{i,j \in N} \sum_{t \in T} u_{ij,t}^s \end{aligned}$$

where α , β , γ , and δ are weighting coefficients associated with empty distance, loaded distance, transportation cost, and service-violation penalties, respectively. The first term directly penalizes unnecessary empty movement, while the final term prevents excessive empty-mile reduction from causing unacceptable service deficiencies.

4.4 Vehicle Capacity Constraint

The quantity assigned to a vehicle cannot exceed its available carrying capacity:

$$\sum_{i,j \in N} d_{ij,t}^s x_{vij,t}^s \leq Cap_v A_{v,t}^s \quad \forall v, t, s$$

This constraint connects the forecasting component with the optimization model by using predicted freight demand and vehicle availability to determine feasible assignments.

4.5 Demand Fulfilment Constraint

The total transportation capacity allocated to a freight movement should satisfy the corresponding forecasted demand:

The unmet-demand variable provides flexibility when available vehicle capacity is insufficient. The penalty associated with $u_{ij,t}^s$ discourages unnecessary service violations.

4.6 Vehicle Flow Conservation

Vehicle continuity must be maintained between consecutive planning periods. The flow conservation constraint is represented as:

$$\sum_{j \in N} (x_{vij,t}^s + r_{vij,t}^s) = \sum_{j \in N} (x_{vji,t+1}^s + r_{vji,t+1}^s)$$

for each vehicle v , location i , period t , and scenario s , subject to appropriate boundary conditions.

This constraint prevents the optimization model from generating physically impossible vehicle movements and ensures that vehicle assignments are consistent with their previous locations.

4.7 Vehicle Availability Constraint

A vehicle can only be assigned when it is available during the corresponding planning period:

$$\sum_{i,j \in N} (x_{vij,t}^s + r_{vij,t}^s) \leq A_{v,t}^s \quad \forall v, t, s$$

Vehicle availability can incorporate information such as completed deliveries, maintenance activities, driver schedules, breakdowns, and delays.

4.8 Time-Window Constraint

For time-sensitive freight, pickup and delivery operations must satisfy the specified time windows. If T_{ij}^s represents the estimated travel time between locations i and j , then:

$$t_j \geq t_i + T_{ij}^s$$

and

$$e_{ij} \leq t_i \leq l_{ij}$$

where e_{ij} and l_{ij} represent the earliest and latest permissible

service times, respectively. This constraint ensures that efforts to reduce empty kilometres do not compromise delivery requirements.

4.9 Stochastic Scenario Formulation

Freight demand and vehicle capacity are subject to uncertainty caused by demand fluctuations, cancellations, delays, breakdowns, and changes in vehicle availability. Therefore, multiple operating scenarios are considered.

For each scenario s , the forecasting models provide a corresponding demand level $d_{ij,t}^s$ and vehicle availability $A_{v,t}^s$. The expected objective value is calculated as:

$$E[Z] = \sum_{s \in S} p_s Z_s$$

where Z_s represents the objective value obtained under scenario s , and P_s represents its probability.

This stochastic formulation enables the optimization process to consider a range of possible future conditions instead of relying exclusively on a single point forecast.

4.10 Rolling-Horizon Optimization

The mathematical model is implemented within a rolling-horizon framework. At each decision interval, the latest shipment information, vehicle positions, completed deliveries, cancellations, delays, and fleet availability are incorporated into the forecasting models. Updated demand and capacity forecasts are then supplied to the optimization model.

Only the decisions required for the immediate operational period are implemented. The remaining planning horizon is subsequently re-optimized when new information becomes available. This process allows the system to continuously identify potential backhaul opportunities and reposition vehicles according to changing freight demand and transport capacity.

The proposed mathematical formulation therefore establishes a direct connection between prediction and operational decision-making. Instead of forecasting freight demand independently and assigning vehicles afterward, the proposed framework uses predicted demand, predicted transport capacity, and their associated uncertainty to determine vehicle assignments, backhaul opportunities, and empty repositioning decisions. This integration forms the basis for evaluating the effectiveness of the proposed framework in reducing empty vehicle kilometres while maintaining transportation service levels.

5. Experimental Design and Evaluation Framework

The experimental design is developed to evaluate the effectiveness of the proposed predictive-optimization framework in forecasting freight demand, estimating transport capacity, and reducing empty vehicle movements. The evaluation considers both prediction performance and operational performance, since accurate forecasting alone does not necessarily result in improved fleet utilization.

5.1 Dataset Description

The proposed framework requires integrated freight and vehicle-operation data. The dataset should contain historical

shipment records together with corresponding vehicle and route information. Each shipment record should ideally include origin, destination, shipment quantity, shipment date and time, pickup and delivery requirements, vehicle type, and delivery status.

Vehicle-level information should include vehicle identification, carrying capacity, location, availability, assigned shipment, travel status, maintenance status, and completed-trip information. Additional contextual variables such as day of the week, month, holidays, seasonal periods, traffic conditions, and weather conditions can be incorporated when available.

The dataset is organized chronologically so that the forecasting models learn only from information that would have been available at the corresponding decision time. This approach is important for preventing information leakage between training and testing data.

5.2 Data Pre-processing

The collected data are first examined for missing values, duplicate records, inconsistent timestamps, invalid geographic information, and abnormal transportation distances. Missing numerical values can be handled using appropriate statistical or model-based imputation methods, while categorical variables can be transformed using encoding techniques.

Continuous variables such as shipment quantity, travel distance, transportation time, and vehicle capacity may be normalized where required by the selected forecasting algorithm. Extreme observations are examined using statistical or domain-based outlier detection methods.

Temporal variables are extracted from shipment timestamps, including hour, day, week, month, weekday/weekend status, and holiday indicators. Origin-destination information is transformed into spatial features representing freight-flow relationships between locations.

5.3 Training, Validation, and Testing Strategy

A chronological data-splitting strategy is adopted rather than random sampling. The historical observations are divided into training, validation, and testing periods. The training dataset is used to develop the forecasting models, the validation dataset is used for hyperparameter selection, and the final testing dataset is reserved for evaluating predictive performance.

A representative structure is

$$D = D_{train} \cup D_{val} \cup D_{test}$$

where D_{train} , D_{val} , and D_{test} represent the training, validation, and testing datasets, respectively.

The chronological arrangement ensures that future observations are not used to predict earlier observations. This provides a more realistic representation of how the proposed framework would operate in an actual logistics environment.

5.4 Forecasting Models

Multiple machine-learning models are considered for freight-demand and transport-capacity prediction. Tree-based models such as Random Forest, XGBoost, and

LightGBM can be evaluated against sequential deep-learning models such as Long Short-Term Memory (LSTM) networks.

The forecasting models receive historical freight, temporal, spatial, and operational features as inputs. Their outputs include future freight demand and available vehicle capacity over the selected prediction horizon.

For demand forecasting

$$\hat{D}_{ij,t+h} = f(X_{ij,t}, T_t, S_{ij}, O_t)$$

For transport-capacity forecasting

$$\hat{C}_{t+h} = g(X_{v,t}, A_{v,t}, R_{v,t}, O_t)$$

where h represents the forecasting horizon, X represents historical operational information, T represents temporal variables, S represents spatial characteristics, A represents vehicle availability, R represents vehicle operational status, and O represents additional operational variables.

5.5 Prediction Performance Measures

The forecasting models are evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Weighted Absolute Percentage Error (WAPE).

MAE is calculated as

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

RMSE is defined as

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

WAPE is expressed as

$$WAPE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{\sum_{i=1}^n |y_i|} \times 100$$

where y_i represents the actual value and $\hat{y}_i$ represents the predicted value.

Lower values of MAE, RMSE, and WAPE indicate closer agreement between predicted and observed demand or capacity.

5.6 Operational Evaluation

The operational effectiveness of the proposed framework is evaluated using metrics directly related to freight transportation efficiency. The primary metric is Empty Vehicle Kilometres (EVK), which measures the total distance travelled without carrying freight.

The empty-distance ratio is calculated as

$$EDR = \frac{D_{empty}}{D_{total}} \times 100$$

where D_{empty} is the total empty vehicle distance and D_{total} is the total distance travelled by the fleet.

Additional operational measures include

- Fleet utilization rate
- Total vehicle kilometres
- Backhaul utilization
- Transportation cost
- Freight demand fulfilment rate
- Vehicle capacity utilization
- Number of empty repositioning trips
- Service-level compliance

These measures allow the proposed framework to be evaluated from both efficiency and service perspectives.

5.7 Baseline Comparison

To determine the contribution of the integrated framework, its performance should be compared with alternative operational strategies. The experimental comparison can include:

Baseline 1: Demand-Only Forecasting Freight demand is predicted, but transport-capacity uncertainty is not explicitly modelled.

Baseline 2: Deterministic Vehicle Allocation Vehicle assignments are generated using point forecasts without probabilistic scenarios.

Baseline 3: Conventional Routing without Proactive Backhaul Vehicle routing is performed without explicitly forecasting future backhaul opportunities.

Proposed Framework: Demand forecasting, transport-capacity forecasting, uncertainty estimation, backhaul matching, and rolling-horizon optimization are integrated into a single decision process.

The comparison is designed to determine whether integrating demand and capacity forecasts with uncertainty-aware optimization provides operational benefits beyond conventional forecasting or routing approaches.

5.8 Ablation Analysis

An ablation analysis is conducted to examine the contribution of individual components of the proposed

framework. Different configurations can be evaluated by systematically removing one component at a time.

For example, the following configurations can be examined

M_1 = Demand Forecasting + Deterministic Routing

M_2 = Demand + Capacity Forecasting + Deterministic Routing

M_3 = Demand + Capacity Forecasting + Uncertainty Modelling

M_4 = Full Proposed Framework

The comparison helps identify whether improvements in empty-mile reduction originate from demand forecasting, capacity prediction, uncertainty modelling, or their integration with rolling-horizon optimization.

5.9 Statistical and Robustness Evaluation

Where sufficient repeated observations are available, the performance differences between competing models can be examined using appropriate statistical significance tests. Confidence intervals can also be reported for key operational indicators.

Robustness analysis should further evaluate the framework under different levels of demand variability, vehicle availability, cancellation rates, and operational delays. This is particularly important because the proposed framework is intended for uncertain logistics environments rather than only stable operating conditions.

The experimental framework therefore evaluates the proposed approach at two complementary levels: forecasting accuracy and transportation efficiency. This ensures that improvements in prediction are assessed according to their practical contribution to reducing empty vehicle movement and improving fleet utilization.

6. Results and Discussion

To demonstrate the potential effectiveness of the proposed predictive-optimization framework, a simulated experimental scenario is considered. The evaluation assumes a multi-location road freight network with varying shipment demand, vehicle availability, delivery time windows, and backhaul opportunities. The simulated results are intended to demonstrate the behaviour of the proposed framework and should not be interpreted as empirical results from a real-world dataset.

6.1 Forecasting Performance

Table 1: Four forecasting approaches are considered: Random Forest, XGBoost, LightGBM, and LSTM. The models are evaluated using MAE, RMSE, and WAPE.

Model	Demand MAE	Demand RMSE	Demand WAPE (%)	Capacity MAE	Capacity RMSE	Capacity WAPE (%)
Random Forest	18.72	27.64	12.84	15.31	23.18	11.62
XGBoost	15.46	23.71	10.63	12.84	20.46	9.84
LightGBM	14.28	21.83	9.71	11.76	18.94	8.96
LSTM	13.61	20.47	9.18	11.21	17.82	8.51
Proposed Forecasting Model	11.84	17.96	7.92	9.73	15.64	7.36

The simulated results indicate that the proposed forecasting configuration produces lower prediction errors for both

freight demand and transport capacity. The improvement is particularly important because the optimization stage

depends on the quality of these forecasts. More accurate predictions can provide better information regarding future freight requirements and vehicle availability.

6.2 Operational Performance: The predicted demand and capacity values are subsequently supplied to the stochastic vehicle-allocation and backhaul optimization model.

Table 2: The proposed approach is compared with three baseline strategies.

Method	Empty Vehicle Km	Empty-Distance Ratio (%)	Fleet Utilization (%)	Backhaul Utilization (%)	Total Distance (km)	Service Level (%)
Demand-Only Forecasting	18,460	23.1	68.4	51.7	79,910	91.6
Deterministic Allocation	16,920	21.4	71.8	56.3	79,060	93.2
Conventional Routing	15,780	20.2	74.6	60.8	78,120	94.1
Proposed Framework	11,940	15.7	82.3	72.6	76,050	96.8

The simulated results suggest that integrating demand forecasting, capacity forecasting, uncertainty modelling, and proactive backhaul matching can reduce empty vehicle movement compared with the baseline approaches. The proposed framework records 11,940 empty vehicle kilometres compared with 18,460 km for the demand-only approach.

The empty-distance ratio decreases from 23.1% under demand-only forecasting to 15.7% under the proposed framework. This represents a simulated reduction of approximately 32.0% in the empty-distance ratio relative to the demand-only approach.

6.3 Fleet Utilization and Backhaul Performance

The proposed framework achieves a simulated fleet utilization of 82.3%, compared with 68.4% for the demand-only approach. This improvement results from jointly considering future freight demand and vehicle availability

when making assignment decisions.

Backhaul utilization also increases from 51.7% to 72.6%. This indicates that the predictive component can provide advance information about potential return shipments, allowing the optimization model to identify suitable backhaul assignments before vehicles travel empty.

The results demonstrate an important relationship between forecasting and optimization. Forecasting future demand alone does not directly reduce empty movement. The predicted information must be incorporated into vehicle assignment and repositioning decisions to generate operational improvements.

6.4 Effect of Uncertainty Modelling

The contribution of uncertainty modelling is examined by comparing deterministic and stochastic versions of the proposed optimization process.

Table 3: The proposed optimization process

Optimization Strategy	Empty Vehicle Km	Service Level (%)	Fleet Utilization (%)
Deterministic Optimization	13,870	94.7	78.6
Stochastic Optimization	12,610	96.1	80.4
Stochastic + Rolling Horizon	11,940	96.8	82.3

The simulated results indicate that stochastic optimization can provide more robust assignments than deterministic optimization when future demand and vehicle availability are uncertain. The addition of rolling-horizon re-optimization produces a further reduction in empty vehicle kilometres by incorporating updated operational information.

6.5 Effect of Rolling-Horizon Updates

Table 4: The rolling-horizon strategy is evaluated under different update frequencies

Update Frequency	Empty Vehicle Km	Fleet Utilization (%)	Service Level (%)
Static Planning	15,420	74.9	93.8
12-hour Update	13,280	78.1	95.2
6-hour Update	12,410	80.5	96.1
2-hour Update	11,940	82.3	96.8

The simulated analysis shows that more frequent updates allow the framework to respond to changes in shipment demand and vehicle availability. However, excessively frequent re-optimization may increase computational and operational requirements. Therefore, the appropriate update frequency should be determined according to the characteristics of the logistics operation.

6.6 Discussion

Overall, the simulated evaluation indicates that the integration of predictive modelling and stochastic optimization can potentially improve road freight utilization. The proposed framework simultaneously considers future freight demand, available vehicle capacity, uncertainty, backhaul opportunities, and changing operational conditions.

Compared with the demand-only approach, the proposed framework in the simulated scenario reduces empty vehicle kilometres from 18,460 km to 11,940 km and increases fleet utilization from 68.4% to 82.3%. Backhaul utilization increases from 51.7% to 72.6%, while service-level compliance increases from 91.6% to 96.8%.

These values are illustrative simulation results and not empirical results from an actual freight dataset. They demonstrate how the proposed framework can be evaluated and provide a structure for replacing the simulated values with actual experimental measurements in the final study.

For publication, I strongly recommend keeping the phrase “simulated/illustrative results” unless you actually run the model. Presenting these numbers as real experimental findings without running the experiments would be a methodological problem.

7. Conclusion

This study proposed an integrated predictive-optimization framework for reducing empty vehicle movements in road freight logistics under uncertain demand and transport capacity. The framework combines freight-demand forecasting, transport-capacity prediction, uncertainty estimation, backhaul matching, vehicle repositioning, and rolling-horizon optimization within a unified decision-making process. Unlike conventional approaches that treat freight forecasting and vehicle routing as separate activities, the proposed framework establishes a direct connection between predicted future demand and transportation decisions. Freight-demand forecasts provide information about potential shipment requirements, while transport-capacity forecasts identify vehicles expected to become available. The incorporation of uncertainty enables the optimization model to consider different possible future operating conditions rather than relying exclusively on point forecasts. The mathematical formulation minimizes empty vehicle distance, total transportation distance, transportation cost, and unmet demand while considering vehicle capacity, availability, route continuity, and delivery time-window constraints. The rolling-horizon mechanism further enables vehicle assignments to be updated as new freight and fleet information becomes available. The illustrative evaluation demonstrates how the proposed framework can be assessed using forecasting and operational indicators, with the simulated scenario indicating potential reductions in empty vehicle kilometres and improvements in fleet and backhaul utilization. However, these values are illustrative and require validation using real-world or benchmark freight transportation datasets before being considered empirical findings. Future research can validate the framework using large-scale real-world freight datasets and extend the model by incorporating real-time traffic conditions, fuel consumption, carbon emissions, driver working-hour constraints, electric freight vehicles, and dynamic shipment cancellations. Advanced probabilistic forecasting and reinforcement-learning-based decision policies can also be explored to improve adaptation to highly dynamic freight environments. Overall, the proposed framework provides a structured approach for connecting freight-demand prediction with uncertainty-aware transportation optimization to improve fleet utilization and reduce unnecessary empty vehicle movements in road freight logistics.

8. References

1. Tan SY, Yeh WC. The vehicle routing problem: state-of-the-art classification and review. *Applied Sciences*. 2021;11:10295.
2. Vidal T, Laporte G, Matl P. A concise guide to existing and emerging vehicle routing problem variants. *European Journal of Operational Research*. 2021;286:401-416.
3. Boumpa E, Tsoukas V, Chioktour V, Kalafati M, Spathoulas G, Kakarountas A, *et al.* A review of the vehicle routing problem and the current routing services in smart cities. *Analytics*. 2023;2:1-16.
4. Liu X, Chen YL, Por LY, Ku CS. A systematic literature review of vehicle routing problems with time windows. *Sustainability*. 2023;15:12004.
5. Wei X, Xiao Z, Wang Y. Solving the vehicle routing problem with time windows using modified rat swarm optimization algorithm based on large neighborhood search. *Mathematics*. 2024;12:1702.
6. Agárdi A, Kovács L, Bányai T. Ontology support for vehicle routing problem. *Applied Sciences*. 2022;12:12299.
7. Nguyen T, Pham D, Buddhakulsomsiri J, Parthanadee P. Load-dependent vehicle routing problem: a heuristic approach. *Industrial Engineering and Applied Engineering*. 2023;35:796-804.
8. Dursun Ö, Özger A. Multi-depot heterogeneous fleet vehicle routing problem with time windows: airline and roadway integrated routing. *International Journal of Industrial Engineering and Computations*. 2022;13:435-456.
9. Hou D, Fan H, Ren X, Tian P, Lv Y. Time-dependent multi-depot heterogeneous vehicle routing problem considering temporal-spatial distance. *Sustainability*. 2021;13:4674.
10. Wang Y, Zhe J, Wang X, Sun Y, Wang H. Collaborative multidepot vehicle routing problem with dynamic customer demands and time windows. *Sustainability*. 2022;14:6709.
11. Martins LdC, Bayliss C, Copado-Méndez PJ, Panadero J, Juan AA. A simheuristic algorithm for solving the stochastic omnichannel vehicle routing problem with pick-up and delivery. *Algorithms*. 2020;13:237.
12. Konovalenko A, Hvattum LM. Optimizing a dynamic vehicle routing problem with deep reinforcement learning: analyzing state-space components. *Logistics*. 2024;8:96.
13. Mussaliyeva R, Abzhapbarova A. Prospects and problems of multimodal system development of Central Asian transport and transit corridors. *Bulletin of the Kazakh Academy of Transport and Communications*. 2024;3:158-166.
14. Kulmagambetova ZH, Izbairova A. Kazakhstan's transport potential and new logistics solutions: in conditions of high competition for alternative routes and the implementation of the Belt and Road Initiative. *Bulletin of the Kazakh Academy of Transport and Communications*. 2024;1:110-118.
15. Noguchi S. Dynamic empty-vehicle repositioning on long-haul freight corridors: lower bounds and rolling-horizon policies under lead times and time windows. *Future Transportation*. 2026;6(3):125.

Creative Commons (CC) License

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.