Research on Timeliness Prediction and Reliability Analysis of Cross-border Logistics

Xiaodong Wan* **Quanliang Zhao**

Ziyang College of Environmental Science and Technology, Ziyang, 641300, China
**Corresponding author: 13540871464@163.com*

Abstract: *In the context of the deep integration of global supply chains, the accurate prediction of cross-border logistics timeliness and the systematic analysis of reliability have become core issues for optimizing supply chain resilience and efficiency. This study first abstracts the cross-border logistics system as a complex network with scale-free characteristics, analyzing its topological structure, multimodal coordination mechanisms, and the transmission pathways of external uncertainties, thereby clarifying the systemic causes of timeliness fluctuations. Subsequently, it constructs an engineering framework that integrates machine learning time-series prediction algorithms with multi-source heterogeneous data features, focusing on model adaptability, interpretability, and the propagation mechanisms of prediction errors. On this basis, a multi-dimensional dynamic reliability assessment system is proposed, which identifies vulnerable nodes in the network from the perspective of supply chain resilience, ultimately forming an integrated decision-support framework that links timeliness prediction with reliability assessment. This research aims to provide a systematic theoretical methodology for quantifying the performance of cross-border logistics systems, enabling proactive risk management and optimizing operational decision-making.*

Keywords: *Cross-border logistics; Timeliness prediction; Reliability analysis; Complex networks; Machine learning; Supply chain resilience*

Introduction

The operational efficiency of cross-border logistics systems directly impacts the smoothness and cost structure of global trade, with their timeliness performance constrained by complex network structures, multi-link coupling, and highly uncertain environments. Traditional timeliness assessment methods based on experience or simple statistics struggle to address systemic dynamics and sudden disruptions, exhibiting significant limitations in prediction accuracy and reliability quantification. Therefore, constructing a comprehensive research framework capable of analyzing the mechanisms behind timeliness generation, achieving high-precision prediction, and scientifically evaluating system reliability holds dual theoretical and practical necessity. The significance of this study lies in its systematic integration of complex network theory, advanced machine learning methods, and supply chain resilience analysis to link the research chain from "feature analysis" and "model construction" to "assessment and optimization." This approach not only deepens the understanding of the inherent patterns within cross-border logistics systems but also provides methodological support for transitioning decision-making models from passive reaction to proactive management.

1. Analysis of Cross-border Logistics System Characteristics and Timeliness Influencing Factors

1.1 Topological Structure and Dynamic Characteristics of Cross-border Logistics Networks

The cross-border logistics system can be physically abstracted as a complex network composed of nodes and edges. Nodes represent critical infrastructure such as ports, airports, logistics hubs, customs clearance points, and last-mile distribution centers, while edges represent the connections between these nodes, including shipping routes, air corridors, land transport links, and information flow paths. This network typically exhibits scale-free characteristics, meaning a small number of core hub nodes handle the vast majority of connections and flows, while a large number of peripheral nodes show sparse connectivity. This topological structure results in a coexistence of overall network efficiency and vulnerability; congestion or disruption at core nodes can generate significant cascading effects through the connecting edges, thereby exerting a nonlinear impact on global timeliness. The dynamic characteristics of the network are reflected in the time-varying nature of node processing capacities and edge transit capacities, such as periodic fluctuations in port operational efficiency, seasonal adjustments to flight frequencies, and real-time changes in transport vehicle availability. Together, these factors constitute a dynamically evolving network flow system.

The dynamic evolution of the network topology is further driven by commercial decisions and market conditions. Carriers' optimization and restructuring of route networks based on revenue management, as well as logistics service providers' network expansion or contraction to improve coverage, lead to changes in network connection relationships and weight parameters. This structural dynamic adjustment causes the availability and cost-effectiveness of logistics paths to be in a state of continuous change, thereby affecting the baseline value and stability of timeliness. Understanding the static characteristics and dynamic evolution patterns of the network structure provides the topological foundation for analyzing the generation mechanism of logistics timeliness and constructing prediction models. Network resilience analysis offers a structural perspective for identifying timeliness vulnerabilities[1].

1.2 Timeliness Fluctuation Mechanisms in Multi-modal Transportation Coordination

Cross-border logistics chains typically involve the sequential or parallel combination of multiple transportation modes, such as maritime shipping, air freight, rail, and road transport. Different modes exhibit inherent differences in transport speed, batch capacity, scheduling flexibility, and cost structure, making the interface and transfer processes between them a primary source of timeliness fluctuations and uncertainty. The physical manifestation of modal coordination lies in the operations at transshipment nodes, including cargo handling, vehicle transfer, documentation exchange, and customs inspection. The efficiency of these interface operations is constrained by multiple factors, such as the compatibility of facilities and equipment, the degree of standardization in operational procedures, and the level of information synchronization. Delays in any single link can accumulate and amplify throughout the subsequent chain.

The timeliness fluctuation mechanism is deeply rooted in the asynchronicity and dependencies within the multimodal system. For instance, the relative rigidity of shipping schedules, when mismatched with the flexible demands of upstream cargo consolidation or downstream distribution, can generate waiting times or expedited handling costs. The coordination between air and land transport places extremely high demands on the precision of time windows, where even minor deviations between flight arrival times and ground fleet scheduling can lead to subsequent delays. Furthermore, the division of responsibilities based on contracts and the existence of information barriers among different carriers may lead to decentralized decision-making and coordination difficulties, thereby generating systemic fluctuations internally. Consequently, logistics timeliness under multimodal coordination is not a simple summation of individual transport segment durations; rather, it is an emergent property influenced by the coupling effects of a complex system. Quantifying its volatility must be based on characterizing the interactions and feedback mechanisms among different modes[2].

1.3 Transmission Pathways of External Environmental Uncertainties on Logistics Timeliness

The cross-border logistics system is an open system, continuously subjected to disturbances from a series of external environmental uncertainties. These sources of uncertainty can be categorized into multiple dimensions, including natural factors, technical operational conditions, and the broader market environment. Natural factors, such as extreme weather and geological disasters, primarily disrupt or delay logistics processes by directly damaging transportation infrastructure, forcing route changes, or reducing operational efficiency. Uncertainties at the level of technical operational conditions encompass sudden failures of critical equipment, disruptions in communication systems, and fluctuations in energy supply, which directly impact the normal operational capacity of logistics nodes and routes.

The impact of external shocks on logistics timeliness is not a simple linear disruption, but rather a complex process of transmission and evolution through the system's internal structure. A shock first causes an initial delay at the locally affected node or path, and then spreads upstream and downstream through network connections. The speed of transmission and the scope of impact depend on the network's connection density, the buffer capacity of nodes, and emergency response strategies. For example, the sudden closure of a major port forces cargo flows to reroute, leading to congestion on alternative paths and triggering a surge in demand for transport capacity along with freight rate fluctuations. These secondary effects further distort the timeliness distribution across the entire network. Uncertainties in the market environment, such as sudden demand shifts, dramatic changes in trade volume, or structural shortages in transport capacity, affect timeliness by altering the state of resource competition. Their transmission paths are often intertwined with price signals and resource allocation efficiency. Understanding the topological structure and dynamic characteristics of these transmission pathways is key to assessing logistics system reliability and constructing predictive models with disturbance-resistant capabilities.

2. Construction of Cross-border Logistics Timeliness Prediction Models and Methodology

2.1 Analysis on the Applicability of Time-Series Prediction Algorithms Based on Machine Learning

Machine learning algorithms provide a methodological foundation for modeling the nonlinear and high-dimensional characteristics of cross-border logistics timeliness. Classical time-series models, such as the Autoregressive Integrated Moving Average model, offer theoretical advantages when handling stable sequences with clear linear trends and seasonality. However, they often prove inadequate in fitting cross-border logistics time-series data, which is subject to multifactorial random disturbances and potential structural breaks. Ensemble learning algorithms, such as Gradient Boosting Decision Trees and Random Forests, can effectively capture the complex interactive relationships and non-additive effects between feature variables and timeliness by constructing and aggregating multiple decision trees. These algorithms demonstrate strong robustness in handling data with multicollinearity among features. They require fewer assumptions about data distribution and also provide intuitive feature importance ranking[3].

However, cross-border logistics timeliness data is inherently a sequence with spatio-temporal dependencies. Deep learning methods, particularly recurrent neural networks and their variants such as long short-term memory networks, demonstrate theoretical potential in modeling long-range temporal dependencies due to their inherent gated memory mechanisms. They can take historical timeliness sequences and operational event sequences as inputs to learn their dynamic evolution patterns. Graph neural networks offer a pathway to explicitly encode the logistics network topology into the model, predicting the flow timeliness of the entire network by learning the representations of nodes and edges. The core of algorithm adaptability lies in identifying the essential characteristics of the data generation process — whether it relies more on the interaction of high-dimensional static features, is more constrained by dynamic temporal evolution patterns, or involves both — and accordingly selecting or integrating model architectures with the corresponding inductive bias.

2.2 Feature Engineering Framework Based on Multi-source Heterogeneous Data Fusion

The prerequisite for constructing high-performance prediction models is to establish a feature engineering framework that comprehensively characterizes the system state. The influencing factors of cross-border logistics timeliness are scattered across multi-source heterogeneous data. This includes data describing cargo flow, such as order attributes, physical cargo properties, and planned versus actual operational timestamps at nodes; data describing carrier flow, such as vessel/flight identification codes, positioning data, and status information; data describing network state, such as infrastructure utilization rates and route booking rates; as well as external data like regional meteorological conditions and geospatial information. Significant differences exist in the granularity, timeliness, accuracy, and schema of these data, making their fusion subject to core challenges such as entity alignment, temporal synchronization, and scale unification.

A systematic feature engineering framework must encompass three stages: data fusion, feature construction, and feature selection. In the fusion stage, data association and alignment are performed based on unified spatio-temporal benchmarks and logistics entity identifiers. The feature construction phase is dedicated to extracting high-level features with predictive informativeness from raw data. Examples include calculating sliding window statistics for node processing efficiency from time-series operational data, deriving node centrality metrics from network topology data, and inferring path consistency indicators from trajectory data. The feature selection phase requires the application of domain knowledge and statistical methods to assess the correlation between features and the target variable, as well as their contribution to model stability. This process aims to eliminate redundant features and noise, thereby controlling model complexity and enhancing generalization capability. The output of this framework is a set of multi-dimensional feature vectors that comprehensively reflect the dynamics, state, and external environment of the logistics system.

2.3 Interpretability of Prediction Results and Error Propagation Analysis

Machine learning models, particularly complex ensemble models and deep learning models, are often regarded as "black boxes," which hinders the credible application of their prediction results in logistics decision-making that emphasizes causality and controllability. Therefore, enhancing the interpretability of prediction models is necessary. Post-hoc interpretation techniques, such as SHAP values, assign a contribution value to each feature prediction based on cooperative game theory, thereby quantifying the direction and magnitude of the influence of individual features or feature combinations on a single prediction outcome or the overall model output. This aids in identifying the key drivers affecting timeliness and their decision boundaries. Inherently interpretable model designs, such as deep learning models employing attention mechanisms, can reveal which parts of the input sequence the model focuses on when making predictions, providing insights for understanding the temporal patterns the model relies on[4].

The existence of prediction errors is inevitable, and analyzing their sources and propagation pathways is a crucial step in assessing the reliability of predictions. Errors may originate from data quality issues such as noise, missing values, and bias; from model-level issues like incorrect specifications or overfitting; or from the inherent randomness of the system itself. In the sequential processes of cross-border logistics, prediction errors from preceding stages can be transmitted as input errors to the prediction models of subsequent stages, potentially leading to the accumulation or even amplification of errors. Error propagation analysis aims to quantify this error transmission effect through methods such as sensitivity analysis or Monte Carlo simulation, thereby assessing the uncertainty interval of the final end-to-end timeliness prediction result. This not only provides a confidence measure for the predictions but also highlights directions for optimizing data collection priorities and improving the accuracy of sub-models, thereby enhancing the overall robustness and practicality of the prediction system.

3. Cross-border Logistics Reliability Assessment System and Optimization Dimensions

3.1 Construction of Reliability Measurement Indicator System and Weight Allocation

The reliability of a logistics system needs to be comprehensively characterized through a multi-dimensional, quantifiable indicator system. This system should go beyond traditional single metrics such as on-time rate, encompassing the statistical distribution characteristics of timeliness performance, system behavior under extreme fluctuation scenarios, and the spatial performance of service consistency. Core indicators include the mean and variance of timeliness, used to describe its central tendency and dispersion; specific percentile levels, used to define and measure the fulfillment level of service commitments; and order-level timeliness deviation, used to quantify the discrepancy between actual fulfillment and planned routes. Furthermore, introducing metrics such as time window compliance rate and cross-regional timeliness stability coefficient can respectively assess the system's ability to meet precise time requirements and its service balance within a complex network.

The allocation of indicator weights is not a subjective assignment process but should be based on the redundancy of information reflected by the indicators and their contribution to overall system failure. Objective weighting methods, such as the entropy weight method, determine weights based on the degree of variation in the observed value sequences of each indicator; greater variation is considered to indicate stronger discriminative power and thus a higher weight. Network analysis-based weighting methods treat the logistics network as a whole, quantifying the importance of nodes or edges by simulating the degree of disturbance to indicator values after their failure, thereby assigning higher weights to the corresponding performance indicators. This data-driven weight allocation mechanism ensures that the assessment system can adaptively capture the most critical dimensions of reliability under the current network structure and operational state, making the evaluation results more dynamic and targeted[5].

3.2 Identification of Reliability Vulnerable Nodes from a Supply Chain Resilience Perspective

Supply chain resilience focuses on a system's ability to maintain basic functions, absorb shocks, and recover to its original or a better state after being subjected to disturbances. From this perspective, reliability vulnerable nodes refer not only to those with a high probability of failure but, more importantly, to those whose failure would disproportionately and significantly impact the overall timeliness performance of the entire network and are difficult or slow to recover. Identifying such nodes requires a combined approach of stress testing and dynamic simulation. By constructing a digital simulation model of the logistics network, it is possible to systematically simulate the temporary functional degradation or complete failure of different nodes caused by operational disruptions, capacity saturation, or external shocks.

The identification of vulnerability is achieved by analyzing the network's performance response curve under disturbance. Key indicators include the depth of performance degradation, the rate at which performance reaches its trough, the trajectory of the recovery path, and the time required for recovery. The failure of a highly vulnerable node triggers a sharp, steep decline in network performance and a prolonged recovery cycle, potentially inducing secondary failures. Further analysis can introduce metrics such as the area under the resilience curve or the recovery time based on a specific resilience threshold as quantitative criteria. The identified vulnerable nodes are typically hubs that occupy high-centrality positions in the network topology, handle enormous flow volumes, and lack effective redundant paths or backup capacity. Strengthening these nodes represents a leverage point for enhancing the reliability of the entire cross-border logistics network[6].

3.3 Decision-Support Framework Integrating Prediction and Reliability

The integration of the timeliness prediction model and the reliability assessment system forms a dynamic decision-support framework oriented towards risk pre-control. The prediction model provides the probability distribution or point estimate of timeliness for a specific future period, along with its confidence interval, which is essentially an uncertain description of future states. The reliability assessment system offers a static and dynamic profile of the system's robustness under historical and current conditions. The linkage between the two lies in feeding the uncertainty information output by the prediction model as input into the network simulation model of the reliability assessment, thereby enabling a prospective evaluation of "future reliability" or "reliability risk."

The operational workflow of this integrated framework is iterative and closed-loop. Based on current data and the prediction model, timeliness estimates and their distributions for multiple future time windows are generated. Subsequently, within the simulation environment, parameters such as network load and node processing time are probabilistically configured according to these predictive distributions, and potential disturbance scenarios are simulated. The simulation output includes not only the expected average performance but, more importantly, provides the probable range of system reliability indicators (such as on-time delivery probability) and their associated risk probabilities for each future period. Decision-makers can use this information to assess the potential reliability risks of different routes, modes, or supplier choices, enabling a quantified trade-off between cost and reliability risk. Ultimately, this framework supports the formulation of proactive strategies, such as dynamic route selection, resilient inventory deployment, or prioritized resource allocation, thereby facilitating a paradigm shift from passive reaction to proactive management.

Conclusion

This study systematically constructs a research framework for cross-border logistics timeliness prediction and reliability analysis by integrating complex systems theory, data science, and operations research methods. The research elucidates the dynamic characteristics of logistics network topology, the fluctuation mechanisms in multimodal coordination, and the transmission pathways of external shocks, thereby establishing a theoretical foundation for understanding timeliness uncertainty. On the methodological level, it demonstrates the applicability of various machine learning algorithms in time-series prediction scenarios, proposes a feature engineering framework under multi-source data fusion, and emphasizes the value of model interpretability and error analysis for prediction credibility. Furthermore, the constructed reliability assessment system, which integrates multi-dimensional statistical indicators and network resilience, can effectively identify key vulnerable nodes. Ultimately, the integrated framework linking timeliness prediction and reliability assessment provides a viable pathway for closed-loop decision-making that enables risk anticipation and strategy optimization. Future research may focus on constructing higher-fidelity digital twin systems, developing multi-objective collaborative optimization algorithms under uncertainty, and empirically examining the effects of cross-organizational data collaboration mechanisms on enhancing prediction and reliability.

References

[1] Yi Shuxin, and Zhang Zilin. "Research on the Evaluation, Spatio-Temporal Differences, and Promotion Mechanisms of Cross-border Logistics Collaborative Development in the Beijing-Tianjin-Hebei Region." Price Monthly, no. 05 (2025): 52-61.
[2] He Qian, and Qiao Xinyi. "Controlling Costs and Competing on Timeliness: The Race in Cross-border Logistics." Beijing Business Today, May 27, 2024, sec. 004, Business.
[3] Huang Lirong, and Tian Jun. "Discussion on Cross-border Logistics Risk Management for E-commerce Platforms." Journal of Commercial Economics, no. 21 (2023): 105-107.
[4] Li Jingjing, Zhou Yuliang, and Zhou Minwei. "Improving Customs Clearance Efficiency and Reducing Logistics Costs." China Inspection and Quarantine Times, October 13, 2022, sec. 002, General News.
[5] Shu Chang. "Research on the Coordinated Development of Cross-border E-commerce and Cross-border Logistics in China under the New Dual-Circulation Development Pattern." Studies on Party and Government, no. 02 (2021): 121-128.
[6] Li Junbo, Chen Mingda, and Qin Chunlian. "Research on the Evaluation of China's Cross-border E-commerce Development Level—Based on Fuzzy Network Analysis." Price: Theory & Practice, no. 11 (2020): 157-160.