Predictive Disruption Modeling in Cross-Border Logistics: A Machine Learning Approach

Sarah L. Chen | Raj Patel | Ines Ferreira

PLIANT Institute, Palo Alto, CA | Stanford Center for Global Logistics

International Journal of Logistics Research and Applications, Vol. 28, No. 1 (April 2025), pp. 112–139

Abstract

Disruptions in cross-border logistics networks impose significant costs on global trade — estimated at $1.5 trillion annually in direct losses — yet the analytical tools available to anticipate them remain limited relative to the strategic significance of the problem. This paper presents a machine learning framework for predicting supply chain disruptions across a panel of 340 international logistics corridors over a six-year period. Using a gradient-boosted ensemble model trained on geopolitical risk indicators, port throughput data, weather patterns, and trade flow anomalies, we achieve a disruption prediction accuracy of 78.4% at a 30-day forecast horizon. We discuss model architecture, feature engineering, interpretability, data sourcing constraints, regional performance variation, and the operational implications of deploying predictive tools in time-sensitive logistics environments.

Keywords: predictive modeling, supply chain disruption, machine learning, cross-border logistics, geopolitical risk

---

1. Introduction

The past decade has seen an acceleration in both the frequency and severity of supply chain disruptions. Events ranging from port congestion and labor actions to geopolitical conflicts, pandemic-related shutdowns, and increasingly sophisticated cyber-physical attacks have exposed the brittleness of globally integrated logistics networks. Despite substantial industry investment in supply chain visibility tools, the capacity to predict disruptions before they materialize remains limited.

Existing predictive approaches tend to be sector-specific, reliant on proprietary data, or focused on single-node risk assessment rather than corridor-level dynamics. Commercial supply chain platforms increasingly offer disruption alerting, but these systems are predominantly reactive — they detect disruptions as they begin to materialize rather than anticipating them. The gap between detection and prediction is operationally consequential, particularly for high-cost pre-positioning decisions that require advance notice measured in weeks rather than hours.

This paper contributes a corridor-level predictive model that integrates heterogeneous data streams to generate actionable disruption forecasts at operationally relevant time horizons. Our framework draws explicitly on the disruption modeling tradition in operations research while extending it to incorporate the geopolitical, meteorological, and operational data streams that have become available through recent advances in commercial data infrastructure.

2. Related Work

Predictive modeling of supply chain disruption draws on several distinct literatures. The operations research literature on supply chain risk management (Tang 2006; Sodhi & Tang 2012) provides theoretical foundations but has historically been data-poor. The recent emergence of large-scale logistics datasets — AIS vessel tracking, GDELT geopolitical event data, and commercial throughput aggregations — has enabled empirical work at a scale previously impossible.

Machine learning applications to supply chain prediction have grown substantially in the past five years. Cavalcante et al. (2019) applied neural networks to demand forecasting; Baryannis et al. (2019) reviewed the state of the art in supply chain risk modeling; and several recent papers have applied gradient boosting and ensemble methods to disruption-adjacent prediction problems. Our work extends this literature by combining operational logistics data with geopolitical risk indicators in a unified corridor-level framework.

3. Data Sources and Feature Engineering

Our model draws on five primary data sources. AIS vessel tracking data, aggregated at the port-pair level over the 2017–2023 observation period, provides high-frequency information about actual logistics network performance. GDELT geopolitical event data, filtered for logistics-relevant event categories using a supervised classifier trained on hand-labeled incidents, provides upstream signal on geopolitical disruption drivers. NOAA weather and sea state data for maritime corridors captures meteorological risk factors with established disruption correlations.

UN Comtrade trade flow statistics provide longer-horizon context for corridor activity. Proprietary throughput data obtained under data-sharing agreements with six major terminal operators completes the input set with operationally granular data that is not available through public channels. The combination of public and proprietary sources is a deliberate methodological choice: each addresses limitations of the others, and the joint dataset is substantially more informative than any single source.

3.1 Feature Engineering Strategy

Feature engineering emphasized three categories of derived signals. Lag structures capture the temporal dependencies in disruption dynamics — a major weather event today influences disruption probability in subsequent days through known propagation channels. Rolling volatility measures capture regime shifts that may precede disruption onset. Cross-corridor spillover indicators capture the empirical reality that disruptions at major nodes propagate through adjacent corridors with predictable timing and magnitude.

Geographic coverage spans 340 corridors across all primary trade routes, with somewhat denser coverage in the trans-Pacific and North Atlantic regions reflecting both data availability and PLIANT's research priorities. Coverage of corridors involving certain regional jurisdictions remains limited and represents a priority gap for future work — we return to this point in the limitations discussion.

4. Model Architecture and Validation

We employ a gradient-boosted ensemble (XGBoost) as our primary architecture. The choice reflects three considerations: strong performance on tabular data with heterogeneous feature types; relative interpretability compared to deep learning alternatives; and computational efficiency that permits frequent retraining as new data becomes available. Hyperparameter selection followed a standard cross-validated grid search procedure with explicit attention to overfitting risks given the tail-heavy nature of the prediction target.

The model was trained on data from 2017–2022 and validated on a held-out 2023 test set. Cross-validation within the training period was conducted using temporally ordered folds to guard against temporal leakage — a particular risk in time-series prediction problems where standard random cross-validation can produce misleadingly optimistic performance estimates. SHAP values were computed for feature attribution across the full test set, supporting interpretability analysis at both individual prediction and global feature importance levels.

5. Results

The ensemble model achieves a 30-day disruption prediction accuracy of 78.4% (AUC: 0.83) on the held-out test set, representing a 19-point improvement over the persistence baseline and a 12-point improvement over a linear regression benchmark using the same feature set. Performance degrades modestly at the 60-day horizon (AUC: 0.74) and more substantially at 90 days (AUC: 0.61), consistent with the bounded predictability of geopolitical dynamics.

Feature attribution analysis identifies geopolitical event indicators and port throughput anomalies as the strongest predictors, accounting for 41% and 28% of model explanatory power respectively. Weather and sea state indicators contribute 14%, trade flow statistics 11%, and proprietary throughput data the remaining 6% — though the latter contribution is concentrated in a small number of high-performance predictions where the proprietary data was decisive.

5.1 Regional Performance Variation

Model performance varies meaningfully across regional corridors. Trans-Pacific and North Atlantic corridors achieve the strongest performance (AUC 0.86 and 0.84 respectively), reflecting both data quality and the relative tractability of disruption dynamics in these regions. Middle East and North Africa corridors present substantially weaker performance (AUC 0.69), reflecting both data limitations and the high volatility of geopolitical dynamics in the region.

6. Discussion and Limitations

The dependence of model performance on proprietary terminal data represents a significant operationalization challenge. In deployment contexts where such data sharing agreements are not feasible, model accuracy is expected to decline by approximately 12–15 percentage points. Future work should investigate open data substitutes and federated learning approaches that preserve data confidentiality while enabling collaborative model training.

An additional limitation worth flagging: coverage of corridors involving the People's Republic of China, parts of Southeast Asia, and the broader Belt and Road infrastructure network is comparatively thin. Improving this coverage will require either expanded regional fieldwork or new data-sharing arrangements that have proven difficult to establish through formal institutional channels. The author intends to address this gap in subsequent work.

The ethical dimensions of deploying predictive risk tools — particularly with respect to the potential for self-fulfilling disruption dynamics — also warrant further examination. A predictive system that influences the behavior of the actors whose behavior it predicts is methodologically and ethically distinct from one that operates passively. The institutional design choices that govern model deployment have substantive implications for the risk landscape the model is intended to characterize.

7. Conclusion

Predictive disruption modeling has matured to the point where 30-day forecasts at the corridor level can achieve operationally meaningful accuracy. The framework presented here demonstrates that the combination of geopolitical, meteorological, and operational data streams produces predictions that materially exceed what any single data source would support. Realizing the operational value of these capabilities will require sustained investment in data infrastructure, institutional partnerships, and the analytical workforce required to operate and interpret predictive systems at scale.

References

Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2019). Supply Chain Risk Management and AI: A Systematic Literature Review. *International Journal of Production Research*, 57(7), 2179–2202.

Cavalcante, I. M., Frazzon, E. M., Forcellini, F. A., & Ivanov, D. (2019). A Supervised Machine Learning Approach to Data-Driven Simulation of Resilient Supplier Selection. *International Journal of Information Management*, 49, 86–97.

Chen, S. L., Zhang, W., & Holt, M. T. (2024). Information Asymmetry and Disclosure Gaps in Global Procurement Networks. *Journal of Supply Chain Management Research*, 18(3), 44–67.

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, 29(5), 1189–1232.

Ivanov, D., & Dolgui, A. (2020). A Digital Supply Chain Twin for Managing the Disruption Risks and Resilience in the Era of Industry 4.0. *Production Planning & Control*, 32(9), 775–788.

Leetaru, K., & Schrodt, P. A. (2013). GDELT: Global Data on Events, Location, and Tone, 1979–2012. *ISA Annual Convention*.

Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30.

Sodhi, M. S., & Tang, C. S. (2012). *Managing Supply Chain Risk*. Springer.

Tang, C. S. (2006). Perspectives in Supply Chain Risk Management. *International Journal of Production Economics*, 103(2), 451–488.

Wieland, A. (2021). Dancing the Supply Chain: Toward Transformative Supply Chain Management. *Journal of Supply Chain Management*, 57(1), 58–73.

Acknowledgments: This research was supported in part by PLIANT Institute internal research funds. Terminal throughput data was provided under confidentiality agreements; operators are not identified. The authors thank participants at the 2024 INFORMS Annual Meeting for feedback on an earlier version of this work. Correspondence: s.l.chen@projectpliant.com