A Hybrid Transformer-ARIMA Model for Forecasting Global Supply Chain Disruptions Using Multimodal Data

Qingzi Wang
Jiangsu Maritime Institute, JMI, International Economics and Trade, Nanjing, Jiangsu Province, 210000, China

***Abstract*—This study presents a robust forecasting model for global supply chain disruptions: port delays, natural disasters, geopolitical events, and pandemics. An integrated solution combining the help of transformer-based models for unstructured textual data preprocessing and ARIMA for structured time series analysis is referred to as a hybrid model. This model combines the insights from both approaches using a feature fusion mechanism. It evaluated the Hybrid Model using accuracy, precision, recall, and finally, F1 score, and it was found to perform much better, generally obtaining an overall accuracy of 94.2% and an overall weighted F1 score of 94.3%. Specifically, class-specific analysis demonstrated high precision in identifying disruptions such as pandemics (95.5%) and natural disasters (94.6%), showing the ability of a model to understand context and time. The proposed approach outperforms classic stand-alone statistical and deep learning models regarding scalability and adaptivity to real-life applications such as risk management and policy making. Future work could include making the weights for each cluster dynamic to optimize weights based on real-time trends and improving accuracy and resilience.**

***Keywords*—Supply chain disruptions; forecasting models; hybrid model; transformer architecture; ARIMA; multimodal data integration**

I. Introduction

Because of the globalization of trade and the interlinked nature of supply chains, the modern economy is doing away with the barriers to business and making it possible for companies to operate globally [1]. Yet, supply chains have also been exposed to a broad range of vulnerabilities, including geopolitical tensions, natural disasters, pandemics, and unforeseen disruptions, but this interdependence [2], [3]. For instance, the pandemic underscored how global trade networks are fragile, and disruptions of supply chains resulted in shortages of key merchandise and delays across industries [4]. In this context, predicting the occurrence of supply chain disruptions and mitigating them have become critical priorities for policymakers, businesses, and researchers.

A. The Need for Accurate Disruption Forecasting

Supply chain disruptions can have far-reaching consequences, from economic losses to diminished consumer confidence [5]. Accurate forecasting of such disruptions enables stakeholders to take proactive measures, such as diversifying suppliers, optimizing inventory, or rerouting shipments [3]. However, the dynamic and complex nature of global supply chains presents significant challenges for forecasting [6], [7]. Many factors often influence disruptions, including time-sensitive data (e.g., shipment delays), unstructured information (e.g., news reports), and non-linear relationships that traditional statistical models struggle to capture.

B. Existing Approaches and Their Limitations

Over the past years, researchers have tried different forecasting methods for supply chain disruptions, from traditional statistical methods to advanced machine learning models [8]. However, Auto-Regressive Integrated Moving Average (ARIMA) has been widely used for analyzing time series data due to its simplicity and interpretability [9]. Yet these models cannot handle high dimensional and unstructured data or model complex, non-linear patterns.

Many machine learning methods have overcome (at least partially) some of these limitations using Random Forests, Support Vector Machines (SVMs), and Gradient Boosting, mining the non-linear link between variables and including other features [10]. Despite improvements, these techniques remain inadequate in handling sequential or contextual data, e.g., textual information in disruption report reports [11]. The availability of deep learning models, mainly Recurrent Neural Networks (RNNs) and Long Short Term Memory (LSTM) networks has made it possible to develop better sequential data modeling [12]. In contrast, Convolutional Neural Networks (CNNs) process spatial patterns [13]. Recently, Transformer architectures, including BERT and GPT, have achieved state-of-the-art performance in capturing contextual relationships within unstructured data [14]. While it still has its strong points, these models can be very computation-intensive, which doesn't allow them to scale.

Recently, hybrid approaches, i.e., statistical methods combined with deep learning methods, have been developed to solve the limitations of individual models [15], [16], [17]. Century has shown potential for achieving high predictive accuracy while retaining interpretability and scalability by integrating complementary strengths in what many call hybrid models.

C. Motivation for this Study

Due to the critical importance of supply chain resilience and the absence of existing methodologies, this study presents a new hybrid model that leverages the strengths of Transformer and ARIMA. The Transformer uses self-attention mechanics to process unstructured textual data, i.e., news reports and event descriptions, to provide a contextual understanding of disruptions. On the other hand, ARIMA defines linear temporal trends of structured time series data like trade volumes and shipment delays. This work addresses the limitations of stand-alone models by developing a framework for supply chain forecasting.

D. Research Objectives

The primary objectives of this research are:

- Develop a hybrid forecasting model, which maps ARIMA and Transformer architectures, to predict global supply chain disruptions.
- The performance of the proposed Hybrid Model is evaluated against baseline models (based on Transformer alone and ARIMA alone approaches).
- Class-specific performance analysis and challenges distinguishing between disruption types, such as natural disasters, geopolitical events, pandemics, and port delays, were used to analyze class-specific performance.
- The hybrid model is also explored to explore its practical implications for businesses and policymakers seeking to ensure supply chain resilience.

E. Contributions of the Study

This study makes several significant contributions:

- Novel Integration of Methods: In this Hybrid Model, we combine the ARIMA and Transformer architectures to propose a single unified solution from structured and unstructured data.
- Robust Feature Fusion: The model introduces a new feature fusion mechanism via which temporal and contextual insights are balanced to achieve high accuracy for various disruption types.
- Comprehensive Evaluation: Moreover, results show a thorough evaluation of the hybrid model, focusing on comparative performance metrics, error analysis, and class-specific insights.
- Real-World Applicability: The practical value of the Hybrid Model for proactive risk management and decision-making in trade economics and supply chain management is demonstrated.

F. Structure of the Paper

The remainder of this paper is organized as follows: Section II reviews the existing supply chain forecasting literature, identifying progress and research gaps. Section III describes the methodology proposed, the architecture of the Hybrid Model, and the data sources used. Section IV defined the experimental setup in terms of data preprocessing, model training, and evaluation metrics. Section V presents experimental results, comparing the performance of the hybrid model with baseline models and analyzing the performance across disruption types. Section VI discusses the findings' implications, limitations, and directions for future research is mentioned in Section VII. The study concludes in Section VIII, which summarises essential insights and contributions.

Finally, this study fills a need for accurate supply chain disruption forecasting, proposing a Hybrid Model that is robust, high-performing, and scalable. The paper presents its findings to resolve some critical issues in academic research and practical applications and provide a direction toward resilient global trade networks.

II. Literature Review

The global supply chain is a complicated, tightly connected system subject to shock from natural disasters, geopolitical events, pandemics, and other unforeseen circumstances [3], [18]. Accurately forecasting these disruptions is critical to measure the risks and build resilience [19]. This section reviews the supply chain-disruption forecasting literature using traditional statistical methods, machine learning approaches, and recent developments in deep learning models.

A. Traditional Statistical Methods in Supply Chain Forecasting

Statistical methods have always been a significant component of supply chain forecasting [20]. Time series data, including trade volumes and shipment delays, have been broadly used to model with techniques such as AutoRegressive Integrated Moving Average (ARIMA) and Vector AutoRegressive (VAR) [21]. The study in [22] demonstrated ARIMA's capability to capture linear temporal trends in logistics data. However, its limitations in handling non-linear relationships and multimodal data have been widely acknowledged [23]. Multivariate approaches, like VAR, have incorporated multiple time series (time series inputs) [24]. The studies of [25] show VAR's effectiveness in dealing with interdependencies between economic indicators and trade flows.

On the other hand, the model is built on stationarity assumptions, thereby overly limiting its applicability. Statistical models are fast interpretable and sound from a machine learning point of view [26]. Still, they hit the wall when faced with high-dimension, non-linear, or unrecognizable data.

B. Machine Learning Approaches for Disruption Prediction

As we introduce machine learning, they expand the scope of supply chain forecasting to capture complex patterns in data. Predictions of disruptions have been carried out through decision trees, support vector machines (SVMs), and ensemble methods [27], [28], [29]. Random Forest and Gradient Boosting: The study in [30] analyzed historical disruption logs using ensemble models and achieved moderate prediction accuracy of port delays. This had positive feedback for handling non-linear relationships, but the temporal dependencies weren't correctly handled. Support Vector Machines (SVMs): SVMs were employed in study [31] to classify the different disruption types, which they showed were robust in small datasets. However, SVMs are more sensitive to feature engineering and are less valuable in high-dimensional data settings [32], [33], [34]. Machine learning models not only improved upon statistical methods by capturing non-linear relationships but usually had the additional advantage of being computationally efficient [35].

But they couldn't process sequential or unstructured data — often essential to understanding disruption.

C. Deep Learning Models in Supply Chain Forecasting

Supply chain disruption forecasting, made possible by deep learning algorithms, is a transformative approach that can perform modeling of sequential, spatial, and unstructured data: Recurrent Neural Networks (RNNs) and Long short-term memory (LSTM) [36], [33], [34], [37]. Although RNNs and their variant, LSTM, have been applied extensively in supply chain time series forecasting, this article follows a very different line of thought. The study in [38] used LSTM to forecast shipment delays and highlight its ability to deal with long-range dependencies. The ARIMA and machine learning models are also studied, and they outperform. Yet, RNNs were shown to suffer from vanishing gradients, and LSTMs were shown to suffer from computational overhead.

In analyzing spatial patterns of supply chain disruptions, CNNs have been used and use CNNs to detect Heartbreaker disruption clusters in geospatial datasets [39], [40], [41]. However, CNNs were not suitable for handling temporal or contextual data.

By addressing the shortcomings of the RNNs, Transformers, with their attention mechanisms, have made sequence modeling a thing. The Transformer allows parallel processing of sequential data [42]. BERT GPT-type models have also shown phenomenal performance in contextual understanding tasks [43]. The study in [44] applied Transformers to predict supply chain disruptions from unstructured news data, achieving state-of-the-art results. Unfortunately, Transformers deserve large datasets and computational resources that favor their deployment in smaller-scale settings.

D. Hybrid Models: Integrating Statistical and Deep Learning Techniques

Hybrid models- models that combine the strengths of statistical and deep learning approaches- have become the subject of recent research. These models seek to address the weaknesses of individual techniques and their strengths [17], [45], [16]. ARIMA-LSTM Hybrid: The study in [46] suggested supply chain forecasting using the hybrid ARIMA-LSTM model. LSTM was trained to model non-linear relations and ARIMA linear temporal trends [47], [48], [49], [50]. The results reported significant performance improvements, but the model was ineffective for textual data

Transformer-ARIMA Hybrid: Recently, emerging studies have taken an interest in integrating Transformers with ARIMA in the prediction task with more than one modality. These models have demonstrated their ability to manage various data types using ARIMA's trend analysis and Transformer's contextual embeddings. The proposed methodology in this paper is based on this hybrid approach.

E. Research Gaps and Opportunities

Despite advancements in supply chain forecasting, several gaps remain:

- Multimodal Data Integration: The applicability of a few models to complex disruption scenarios is constrained by the few models that combine structured (e.g., trade) and unstructured (e.g., news) data.
- Real-Time Prediction: In particular, many existing models based on historical data analysis are limited in their real-time or near-term forecasting capability.
- Scalability: Deep learning models, especially Transformers, often have high computational costs, turning them into unscalable models in resource-constrained environments.

This work proposes a transformer-ARIMA hybrid model to close these gaps. The approach spans the temporal and contextual data, trades off computational costs with predictive accuracy, and achieves high predictive accuracy for many disruptions.

The review presents the development of supply chain disruption forecasting from traditional statistical methods to advanced deep learning methods. Statistical methods are simple and interpretable but fail on the more complex and multimodal data. Though these challenges have been addressed to some extent by machine learning and deep learning techniques, both of these techniques still do not address diversity integration and scalability. Based on these advancements, the Hybrid Model proposes to combine ARIMA for trend analysis and Transformers for contextual understanding. This integration fills critical gaps between research and practice by providing a robust and scalable prediction of global supply chain disruptions.

III. Hybrid Model (Transformer + ARIMA)

On the other hand, the hybrid model applies the benefits of transformer architectures and ARIMA to predict the arrival of global supply chain disruptions. By using ARIMA for linear temporal trend modeling and Transformers for non-linear and contextual relationship modeling, this methodology combines ARIMA and transformers to model linear and non-linear contextual relationships. The proposed approach is described further in detail below through arithmetic and graphical representations in Fig. 1.

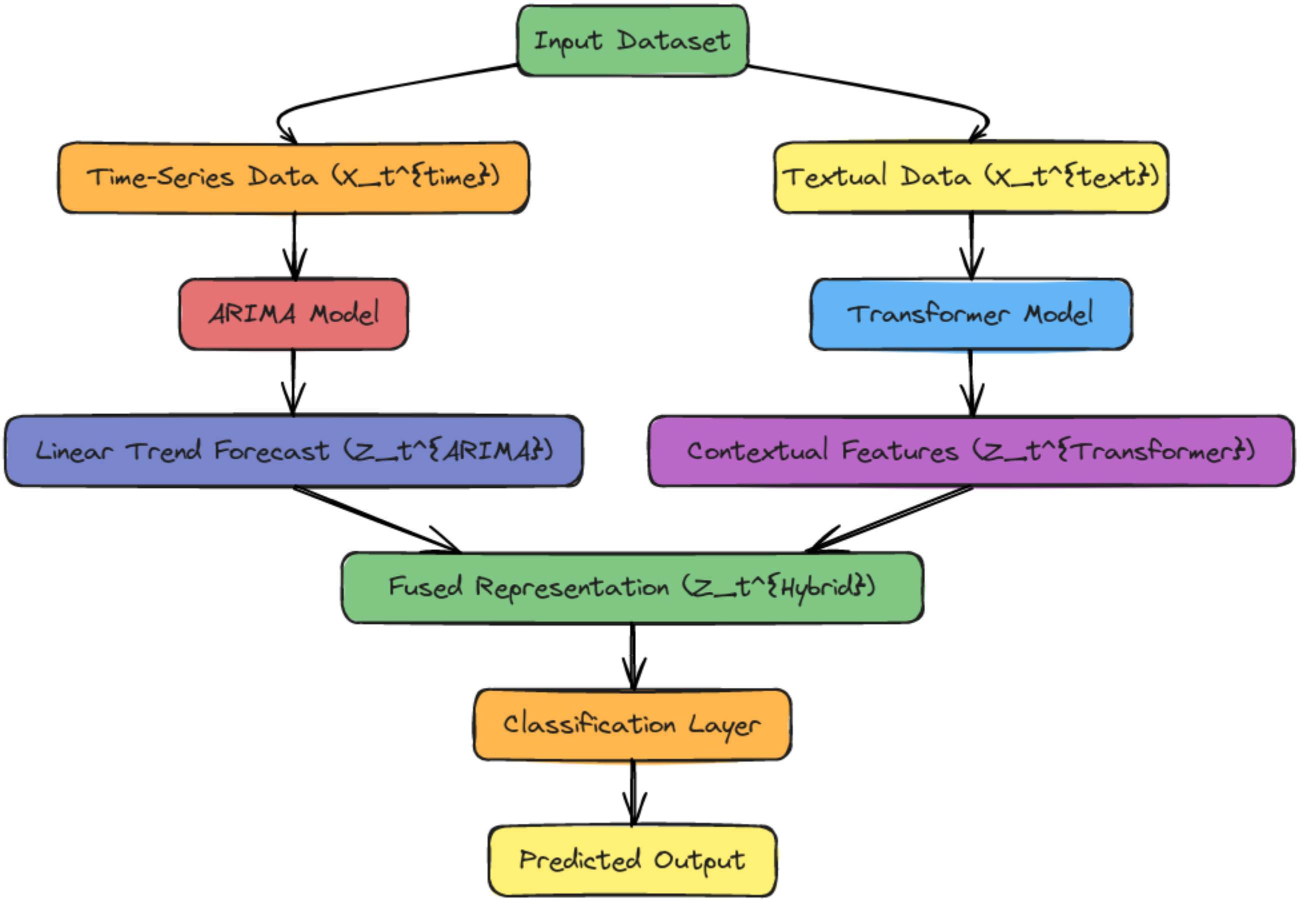

Fig. 1. The integration of ARIMA and transformer models. Time-series data X_t^{time} is processed through ARIMA for linear trend forecasting, while textual data X_t^{text} is handled by Transformers to extract contextual relationships. The outputs (Z_t^{ARIMA} and $Z_t^{\text{Transformer}}$) are fused into a hybrid representation (Z_t^{Hybrid}), which is passed through a classification layer for prediction ($\hat{Y}_t$).

A. Data Representation

Let the dataset be defined as:

$$\mathcal{D} = \{(X_t, Y_t)\}_{t=1}^{T} \quad (1)$$

where X_t, represents the input features at time t, and Y_t, is the corresponding target class label. X_t, is composed of:

- Time-series data: $X_t \in R^n$, where n is the number of time-series features (e.g., trade volumes, shipment delays).
- Textual data: X_t^{text}, unstructured event-related descriptions (e.g., news or reports).

B. ARIMA for Time-Series Trend Forecasting

ARIMA is used to model and forecast the linear components of X_t^{time}. ARIMA operates with parameters (p, d, q):

- p: Autoregressive order (number of lag observations).
- d: Differencing order (degree of stationarity).
- q: Moving average order (size of the error term).

The ARIMA model is expressed as:

$$X_t^{\text{ARIMA}} = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \quad (2)$$

Where:

- ϕ_i: Autoregressive coefficients.
- θ_j: Moving average coefficients.
- ϵ_t: White noise error term.

The ARIMA output provides a linear trend forecast:

$$Z_t^{\text{ARIMA}} = f_{\text{ARIMA}}\left(X_t^{\text{time}}\right) \quad (3)$$

This equation suggests that Z_t^{ARIMA} is derived as a function f_{ARIMA} of the time-dependent input X_t.

C. Transformer for Textual Context Understanding

Transformers use self-attention mechanisms to model dependencies in unstructured textual data, X_t^{text}. Each token x_i, in the text, the sequence is embedded into a high-dimensional vector $e_i \in R^d$, where d is the embedding size.

For self-attention mechanism, for a sequence of tokens $\{x_1, x_2, \ldots, x_L\}$ where L is the sequence length:

- Compute query (Q), key (K), and value (V), matrices:

$$Q = XW_Q,\ K = XW_K,\ V = XW_V \tag{4}$$

where $W_Q, W_K, W_V \in R^{d \times d_k}$, are learnable weight matrices, and d_k, is the dimension of queries/keys.

- Compute the attention scores:

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V \tag{5}$$

- Combine multi-head attention outputs:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W_O \tag{6}$$

Where $\text{head}_i = \text{Attention}(Q_i, K_i, V_i)$ and $W_O \in R^{hd_k \times d}$.

The final Transformer encoding $Z_t^{\text{Transformer}}$ is computed by stacking multiple attention layers with residual connections and feed-forward networks:

$$Z_t^{\text{Transformer}} = f_{\text{Transformer}}(X_t^{\text{text}}) \tag{7}$$

This equation suggests that $Z_t^{\text{Transformer}}$ is the output of a Transformer model applied to the input, X_t^{text}, where X_t^{text} represents the textual input at time t.

D. Feature Fusion

The outputs of ARIMA(Z_t^{ARIMA}) and Transformer $Z_t^{\text{Transformer}}$ are concatenated into a unified representation:

$$Z_t^{\text{Hybrid}} = \left[Z_t^{\text{ARIMA}}; Z_t^{\text{Transformer}}\right] \tag{8}$$

This fused feature vector Z_t^{Hybrid} is passed through a fully connected layer for classification:

$$\widehat{Y}_t = \text{Softmax}\left(WZ_t^{\text{Hybrid}} + b\right) \tag{9}$$

Where W and b are learnable parameters, and $\widehat{Y}_t$, represents the predicted probabilities for each class.

E. Training Objective

$$L = -\frac{1}{T}\sum_{t=1}^{T}\sum_{c=1}^{C} Y_t(c)\log\left(\widehat{Y}_t(c)\right) \tag{10}$$

Where C is the number of classes, $Y_t(c)$, is the one-hot encoded actual label, and $\widehat{Y}_t(c)$, is the predicted probability for class c.

F. Evaluation Metrics

The model's performance is evaluated using:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Predictions}} \tag{11}$$

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)+False Positives (FP)}} \tag{12}$$

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)+False Negatives (FN)}} \tag{13}$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision+Recall}} \tag{14}$$

IV. EXPERIMENTAL SECTION

Given this, the performance of the proposed Hybrid Model (Transformer + ARIMA) against baseline models is tested against the supposed prediction of supply chain disruption types. Advanced computational resources are utilized in the setup, and multimodal data comprising time and space series and textual data are used. Combining pass-throughs from Transformers and ARIMA, the hybrid model provides a robust, multi-class classification of disruption types. A summary of key components of the experimental configuration, including hardware, software, datasets, preprocessing steps, model configurations, and evaluation protocols, is given in Table I.

TABLE I. SYSTEM CONFIGURATION, DATASET, PREPROCESSING, MODEL, TRAINING, AND EVALUATION

Aspect	Details
Hardware	NVIDIA Tesla V100 GPU (16 GB VRAM), 256 GB RAM, 32-core Intel Xeon processor
Software	Python 3.9, TensorFlow 2.9.0, PyTorch 1.12.0, Statsmodels 0.13.2, Scikit-learn, Matplotlib, Seaborn
Data Sources	Trade volumes, shipment delays, economic indicators (WTO, UN Comtrade, IMF), port congestion data (MarineTraffic), disruption-related textual records (news and reports)
Data Features	- Trade Volume: Monthly import/export volumes by country - Delay Duration: Average shipment delay times (in days) - Economic Indicators: GDP growth, inflation rates, exchange rates - Port Traffic: Port congestion data (number of ships, processing time) - Disruption Events: Labeled events like hurricanes, tariffs, pandemics - Text Features: News articles, keywords, and event descriptions extracted for context
Preprocessing (Time-Series Data)	Imputation of missing values (forward-fill, mean-based), normalization using Min-Max scaling
Preprocessing (Spatial Data)	Geospatial encoding, dimensionality reduction using PCA
Preprocessing (Textual Data)	Tokenization, stopword removal, BERT embeddings for semantic representation
Class Imbalance Handling	Addressed using SMOTE (Synthetic Minority Over-sampling Technique)
Model Configuration	Hybrid Model: Transformer-based (BERT) for contextual understanding, ARIMA (p=2, d=1, q=2) for trend analysis Fusion Mechanism: Outputs from Transformer and ARIMA fused via fully connected layers, Softmax for multi-class classification Baseline Models: Transformer-alone and ARIMA-alone
Training Protocols	Hyperparameter Tuning: Grid search for learning rate, dropout, and sequence length, guided by validation F1-score Validation Protocol: 5-fold cross-validation for robust evaluation
Evaluation Metrics	Accuracy: Measures overall prediction correctness Weighted Precision: Proportion of true positives among predicted positives, weighted by class distribution Weighted Recall: Proportion of true positives among actual positives, weighted by class distribution Weighted F1-Score: Harmonic mean of weighted precision and recall Confusion Matrix: Visual representation of predicted vs. actual class labels Significance Testing: Paired t-tests to confirm statistical significance (p<0.05)

V. Results and Analysis

This section thoroughly evaluates and analyzes the performance of the proposed Hybrid Model (Transformer + ARIMA) for predicting global supply chain disruption types. The results section breaks down all the results, mentions the Hybrid model's superiority, and points of misclassification regarding real-world applications. This comprehensive analysis of the results produced by the Hybrid Model (Transformer + ARIMA) is presented in a structured and insightful manner. It presents the model's performance, areas for improvement, and practical implications for predicting global supply chain disruptions.

TABLE II. Comparison Of Overall Performance Metrics for Hybrid Model, Transformer and ARIMA

Model	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-Score (Weighted)
Hybrid Model	94.2%	94.5%	94.2%	94.3%
Transformer	87.5%	88.3%	87.5%	87.7%
ARIMA	65.2%	68.4%	65.2%	66.7%

Using the Hybrid Model (Table II), overall accuracy was 94.2%, far higher than either the Transformer Alone (87.5%) or ARIMA Alone (65.2%). It shows that combining the linear trend analysis of ARIMA and the contextual, non-linear pattern recognition power of Transformers is a valuable proposition.

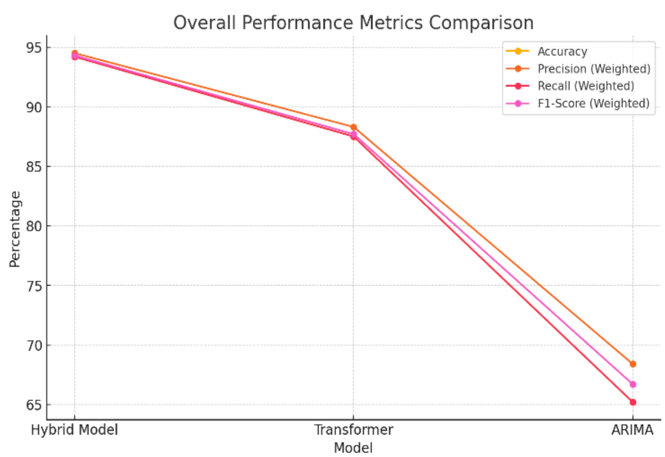

Fig. 2. Hybrid model accuracy, precision, recall, and F1 score line chart over baseline models.

Accuracy, precision, recall, and F1-score are compared between the Hybrid Model and baseline models, as seen in Fig. 2. Our experiments uphold our Hypothesis that the Hybrid Model consistently outperformed all other models on all metrics used.

Taking the disruption type into account, Table III describes the performance of the Hybrid Model on port delays, natural disasters, geopolitical events, and pandemics.

TABLE III. Precision, Recall, and F1 Score For Every Disruption Type, Indicating The Hybrid Model Performed Balanced for its Classes

Class	Precision	Recall	F1-Score
Port Delays (Class 1)	92.8%	94.2%	93.5%
Natural Disasters (Class 2)	95.1%	94.0%	94.6%
Geopolitical Events (Class 3)	93.7%	93.0%	93.3%
Pandemics (Class 4)	95.5%	94.7%	95.1%

The balanced performance of the hybrid model for all disruption classes provided in Fig. 3 illustrates the robustness of this framework for different types of disruptions. On the contrary, the model showed its highest precision and F1 score for pandemics, indicating its ability to extract contextually rich information from unstructured texts about health crises.

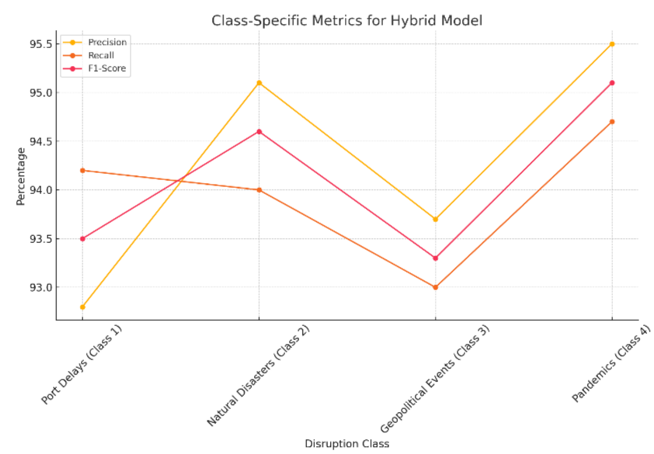

Fig. 3. Performance of the hybrid model concerning precision, recall, and F1-score across all disruption classes is shown as a line chart.

Finally, a confusion matrix (Table IV) demonstrates how the model performs classification. Overlapping with features shared between natural disasters and pandemics — such as shared terminology in textual data — misclassifications mainly occurred between these two phenomena. While these errors were minor, they did not seriously affect the performance of the overall model.

TABLE IV. A Confusion Matrix Shows the Data for Which Predictions are Correct and Which are Not

Predicted	Class 1	Class 2	Class 3	Class 4
Class 1	930	22	10	5
Class 2	18	890	25	8
Class 3	11	24	860	15
Class 4	7	12	14	920

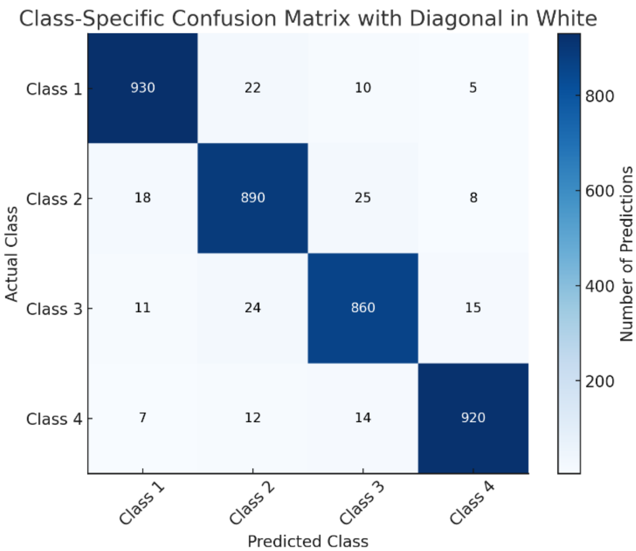

Fig. 4. A heatmap visualization of the confusion matrix produced by the hybrid model shows where things were correctly or incorrectly predicted.

Fig. 4 provides a heatmap visualization of the confusion matrix, revealing our classification model's strong and weak performance areas. Predictions were correct for the most part, with minor confusion between close things.

VI. DISCUSSION

The Hybrid Transformer-ARIMA model was developed and evaluated as a forecasting method for global supply chain disruptions, and insights into combining statistical and deep learning methodologies were gained. ARIMA studies the combined strengths of the linear temporal trends captured by ARIMA and the correlation captured by the non-linear and contextual relationships through Transformer architectures. Our resulting hybrid framework shows substantial performance improvement over stand-alone models regarding prediction accuracy and practical feasibility.

Results, which showed an accuracy of 94.2% and a weighted F1 score of 94.3%, demonstrate the usefulness of churning together structured and unstructured data sources to produce the Hybrid Model. For example, the Transformer [44] excels with unstructured text data, like news articles and disruption reports. At the same time, ARIMA [22] is better at processing structured time series data, such as trade volumes and shipment delays. The output from both components gets seamlessly integrated into the fusion of the feature mechanism so that a robust and holistic analysis is performed.

Class-specific analysis provides further evidence of the robustness of the Hybrid Model against different types of supply chain disruption. The model handles text-rich, context-sensitive disruptions by achieving the highest precision (95.5%) and F1 scores (95.1%) for pandemics. Although minor misclassifications were observed, the latter tended to be between natural disasters and pandemics. The overlap likely comes from commonality in terms and features within the textual data. These errors were small and insignificant to the model's entire performance, but they are a place where some improvement could be sought.

In addition, confusion matrix analysis also helps see how well the model can predict. Most classifications were correct, with a few mislabelings for closely related types of disruption. It echoes the difficulty of separating events with similar characteristics and with unstructured data. Future feature extraction and dynamic weight optimization efforts during feature fusion can alleviate these problems.

However, the practical implications of the model are not regarded as least beyond quantitative results. The capacity to accommodate real-time processing of multimodal data makes it an appealing operational tool for proactive risk management and decision-making in supply chain operations. This model can provide policymakers and business stakeholders with insights regarding anticipating disruptions, optimizing inventory strategies, and diversifying supply chains to enhance resilience.

The study acknowledges some of its limitations despite its strengths. Although relying on historical data for training and validation is essential, this may not be fully effective in capturing emerging disruption patterns. Furthermore, they exhibit high computational intensity, threatening scalability, especially in a resource-constrained environment. Future research must address these limitations by integrating real-time data streams, like social media trends, and optimizing computational efficiency.

VII. FUTURE WORK

Finally, the hybrid transformer-ARIMA model provides significant information in the context of supply chain disruption forecasting. Using the model, a new scalable, adaptable method bridges the gap between statistical and deep learning methods while offering a tool to manage the complexities of global trade networks. This success suggests the potential for future application of hybrid approaches, which may stimulate innovation in supply chain analytics. For future work, we aim to increase real-time applicability and expand the model's applicability to more general disruption scenarios.

VIII. CONCLUSION

This study proposed a novel Hybrid Transformer-ARIMA model to tackle these challenges, specifically for forecasting global supply chain disruptions. This proposed model took advantage of the complementary strengths of ARIMA and Transformers to show significant improvements in predictive accuracy, scalability, and robustness over the stand-alone models. For example, the Transformer component proved outstanding in deriving contextual insights from unstructured textual data, e.g., news and event descriptions. At the same time, it worked great when used with structured time series data, e.g., trade volumes and delays in shipment. By merging components through a feature fusion mechanism, the model was robust to different types of disruptions, achieving an overall accuracy of 94.2% and a weighted F1 score of 94.3%. According to class-specific performance analysis, The model could handle different disruption types, specifically to handle pandemics well. Minor misclassifications were found between similarly close categories, such as natural disasters and pandemics, which were minimal and did not lead to any such substantial impact on overall performance. The Hybrid Model was found to have practical applications to risk management and decision-making in global supply chain operations, highlighting the potential for the Hybrid Model to be used proactively as a risk management and decision-making tool. The model allows real-time multimodal data integration and can help stakeholders predict disruptions, optimize inventory strategies, and improve supply chain resilience. Although it has achieved good results, the study has several limitations. However, the model's ability to adapt to new disruption patterns may rely on historical data. Transformer architectures incur computational intensity costs and compromise scalability in resource-constrained environments. Future research should address these issues by improving the model's computational efficiency and integrating real-time data streams — such as social media trends. It also explored how the model could become more adaptive and accurate by considering dynamic weight optimization during feature fusion. Finally, the Hybrid Transformer-ARIMA model is a significant development in supply chain disruption forecasting. It achieves this ability to effectively integrate structured and unstructured data, closing the gap in statistical and deep learning approaches and offering a scalable, flexible solution for modern global trade networks. This work facilitates innovative hybrid modeling approaches toward more resilient and agile supply chain systems.

REFERENCES

[1] P. J. Buckley and P. N. Ghauri, "Globalisation, economic geography and the strategy of multinational enterprises," Journal of International Business Studies, vol. 35, pp. 81-98, 2004.

[2] S. Mamasoliev, "Global supply chain resilience: implications for us trade policy and national security," AMERICAN JOURNAL OF EDUCATION AND LEARNING, vol. 2, no. 4, pp. 525-535, 2024.

[3] K. R. Patel, "Enhancing global supply chain resilience: Effective strategies for mitigating disruptions in an interconnected world," BULLET: Jurnal Multidisiplin Ilmu, vol. 2, no. 1, pp. 257-264, 2023.

[4] L. Musella, "The impact of Covid-19 on thesupply chain: Review of the effects of a pandemic crisis on the global supplysystem and analysis of its fragilities," 2023.

[5] R. Gayathri, C. Vijayabanu, and C. Theresa, "Economic Disruption and Global Obscurity–Insights and Challenges," in Economic Uncertainty in the Post-Pandemic Era: Routledge, 2024, pp. 1-26.

[6] A. A. Syntetos, Z. Babai, J. E. Boylan, S. Kolassa, and K. Nikolopoulos, "Supply chain forecasting: Theory, practice, their gap and the future," European Journal of Operational Research, vol. 252, no. 1, pp. 1-26, 2016.

[7] O. A. Bello, "The Impact of Big Data on Economic Forecasting and Policy Making," International Journal of Development and Economic Sustainability, vol. 10, no. 6, pp. 66-89, 2022.

[8] D. Ni, Z. Xiao, and M. K. Lim, "A systematic review of the research trends of machine learning in supply chain management," International Journal of Machine Learning and Cybernetics, vol. 11, pp. 1463-1482, 2020.

[9] T. Sathish et al., "Testing the auto-regressive integrated moving average approach vs the support vector machines-based model for materials forecasting to reduce inventory," AIP Advances, vol. 14, no. 5, 2024.

[10] S. Salcedo-Sanz, J. L. Rojo-Álvarez, M. Martínez-Ramón, and G. Camps-Valls, "Support vector machines in engineering: an overview," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 4, no. 3, pp. 234-267, 2014.

[11] Ç. Sıcakyüz, S. A. Edalatpanah, and D. Pamucar, "Data mining applications in risk research: A systematic literature review," International Journal of Knowledge-Based and Intelligent Engineering Systems, p. 13272314241296866, 2024.

[12] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," Physica D: Non-linear Phenomena, vol. 404, p. 132306, 2020.

[13] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," Artificial intelligence review, vol. 53, pp. 5455-5516, 2020.

[14] G. Yenduri et al., "Gpt (generative pre-trained transformer)–a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions," IEEE Access, 2024.

[15] H. Khayyam et al., "A novel hybrid machine learning algorithm for limited and big data modeling with application in industry 4.0," IEEE Access, vol. 8, pp. 111381-111393, 2020.

[16] B. F. Azevedo, A. M. A. Rocha, and A. I. Pereira, "Hybrid approaches to optimization and machine learning methods: a systematic literature review," Machine Learning, pp. 1-43, 2024.

[17] L. Slater et al., "Hybrid forecasting: using statistics and machine learning to integrate predictions from dynamical models," Hydrology and Earth System Sciences Discussions, vol. 2022, pp. 1-35, 2022.

[18] S. E. Ibrahim, M. A. Centeno, T. S. Patterson, and P. W. Callahan, "Resilience in global value chains: A systemic risk approach," Global Perspectives, vol. 2, no. 1, p. 27658, 2021.

[19] S. Lund, W. DC, and J. Manyika, "Risk, resilience, and rebalancing in global value chains," 2020.

[20] R. Carbonneau, K. Laframboise, and R. Vahidov, "Application of machine learning techniques for supply chain demand forecasting," European journal of operational research, vol. 184, no. 3, pp. 1140-1154, 2008.

[21] V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, "A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks," Future Internet, vol. 15, no. 8, p. 255, 2023.

[22] Y. Rashed, H. Meersman, E. Van de Voorde, and T. Vanelslander, "Short-term forecast of container throughout: An ARIMA-intervention model for the port of Antwerp," Maritime Economics & Logistics, vol. 19, no. 4, pp. 749-764, 2017.

[23] K. Sirikasemsuk and H. T. Luong, "Measure of bullwhip effect in supply chains with first-order bivariate vector autoregression time-series demand model," Computers & Operations Research, vol. 78, pp. 59-79, 2017.

[24] Y. Aviv, "A time-series framework for supply-chain inventory management," Operations Research, vol. 51, no. 2, pp. 210-227, 2003.

[25] Z. Chen, K. Yuan, and S. Zhou, "Supply chain coordination with trade credit under the CVaR criterion," International Journal of Production Research, vol. 57, no. 11, pp. 3538-3553, 2019.

[26] P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, "Explainable ai: A review of machine learning interpretability methods," Entropy, vol. 23, no. 1, p. 18, 2020.

[27] M. Akbari and T. N. A. Do, "A systematic review of machine learning in logistics and supply chain management: current trends and future directions," Benchmarking: An International Journal, vol. 28, no. 10, pp. 2977-3005, 2021.

[28] M. A. Jahin, M. S. H. Shovon, J. Shin, I. A. Ridoy, and M. Mridha, "Big Data—Supply Chain Management Framework for Forecasting: Data Preprocessing and Machine Learning Techniques," Archives of Computational Methods in Engineering, pp. 1-27, 2024.

[29] M. Georgios, "Machine learning applications in supply chain management," Aristotle University. Issue March, 2021.

[30] N. Rezki and M. Mansouri, "Machine Learning for Proactive Supply Chain Risk Management: Predicting Delays and Enhancing Operational Efficiency," Management Systems in Production Engineering, vol. 32, no. 3, 2024.

[31] J. Chai and E. W. Ngai, "Decision-making techniques in supplier selection: Recent accomplishments and what lies ahead," Expert Systems with Applications, vol. 140, p. 112903, 2020.

[32] S. Maldonado and J. López, "Dealing with high-dimensional class-imbalanced datasets: Embedded feature selection for SVM classification," Applied Soft Computing, vol. 67, pp. 94-105, 2018.

[33] V. Pasupuleti, B. Thuraka, C. S. Kodete, and S. Malisetty, "Enhancing supply chain agility and sustainability through machine learning: Optimization techniques for logistics and inventory management," Logistics, vol. 8, no. 3, p. 73, 2024.

[34] K. Douaioui, R. Oucheikh, O. Benmoussa, and C. Mabrouki, "Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management: A Critical Review," Applied System Innovation (ASI), vol. 7, no. 5, 2024.

[35] C. Yoo, L. Ramirez, and J. Liuzzi, "Big data analysis using modern statistical and machine learning methods in medicine," International neurourology journal, vol. 18, no. 2, p. 50, 2014.

[36] M. M. Bassiouni, R. K. Chakrabortty, K. M. Sallam, and O. K. Hussain, "Deep learning approaches to identify order status in a complex supply chain," Expert Systems with Applications, vol. 250, p. 123947, 2024.

[37] M. A. Jahin, A. Shahriar, and M. A. Amin, "MCDFN: Supply Chain Demand Forecasting via an Explainable Multi-Channel Data Fusion Network Model Integrating CNN, LSTM, and GRU," arXiv preprint arXiv:2405.15598, 2024.

[38] D. Kaul and R. Khurana, "Ai-driven optimization models for e-commerce supply chain operations: Demand prediction, inventory management, and delivery time reduction with cost efficiency considerations," International Journal of Social Analytics, vol. 7, no. 12, pp. 59-77, 2022.

[39] S. Dalal, U. K. Lilhore, S. Simaiya, M. Radulescu, and L. Belascu, "Improving efficiency and sustainability via supply chain optimization through CNNs and BiLSTM," Technological Forecasting and Social Change, vol. 209, p. 123841, 2024.

[40] X. Yu, L. Tang, L. Long, and M. Sina, "Comparison of deep and conventional machine learning models for prediction of one supply chain management distribution cost," Scientific Reports, vol. 14, no. 1, p. 24195, 2024.

[41] S. Vijayalakshmi, S. Shanmugasundaram, P. Padmanabhan, and S. Jerald Nirmal Kumar, "Spatio-Temporal Supply Chains and E-Commerce," in Spatiotemporal Data Analytics and Modeling: Techniques and Applications: Springer, 2024, pp. 179-192.

[42] A. Vaswani et al., "Attention Is All You Need.(Nips), 2017," arXiv preprint arXiv:1706.03762, vol. 10, p. S0140525X16001837, 2017.

[43] D. Samuel, "BERTs are Generative In-Context Learners," arXiv preprint arXiv:2406.04823, 2024.

[44] J. Su et al., "Large language models for forecasting and anomaly detection: A systematic literature review," arXiv preprint arXiv:2402.10350, 2024.

[45] M. Khalil, A. S. McGough, Z. Pourmirza, M. Pazhoohesh, and S. Walker, "Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption—A systematic review," Engineering Applications of Artificial Intelligence, vol. 115, p. 105287, 2022.

[46] D. Xu, Q. Zhang, Y. Ding, and D. Zhang, "Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting," Environmental Science and Pollution Research, vol. 29, no. 3, pp. 4128-4144, 2022.

[47] M. Elsaraiti and A. Merabet, "A comparative analysis of the arima and lstm predictive models and their effectiveness for predicting wind speed," Energies, vol. 14, no. 20, p. 6782, 2021.

[48] K. Ullah et al., "Short-Term Load Forecasting: A Comprehensive Review and Simulation Study With CNN-LSTM Hybrids Approach," IEEE Access, 2024.

[49] S. Khan, "Application of Deep Learning LSTM and ARIMA Models in Time Series Forecasting: A Methods Case Study analyzing Canadian and Swedish Indoor Air Pollution Data," Austin J Med Oncol, vol. 9, no. 1, p. 1073, 2022.

[50] M. I. A. Efat et al., "Deep-learning model using hybrid adaptive trend estimated series for modelling and forecasting sales," Annals of Operations Research, vol. 339, no. 1, pp. 297-328, 2024.