Alexandria Engineering Journal 128 (2025) 231–246

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

Enhancing supply chain resilience with data envelopment analysis and temporal convolutional networks for supplier efficiency and late delivery risk prediction

Mastoor M. Abushaega [a,*], Osamah Y. Moshebah [b], Ahmed Hamzi [a], Saleh Y. Alghamdi [b]

[a] Department of Industrial Engineering, College of Engineering and Computer Science, Jazan University, Jazan 45142, Saudi Arabia
[b] Department of Industrial Engineering, College of Engineering, King Khalid University, Abha 61421, Saudi Arabia

ARTICLE INFO

Keywords:
Supply chain
Supplier evaluation
Delivery risk
Data Envelopment Analysis
Temporal Convolutional Networks

ABSTRACT

The direct impact of late deliveries on supply-chain performance is critical, as it directly affects customer satisfaction and business income. Most conventional machine-learning-based risk-predictive models do not allow for transparency regarding determining supplier performance or identifying the reasons behind delivery delays. This research thus attempts to fill in the gaps of the models above by proposing an integrated analytical framework for real-time supplier evaluation and disruption forecasting. The hybrid model proposed utilizes Data Envelopment Analysis and Temporal Convolutional Networks for supplier-efficiency assessment with late-delivery risk prediction. The DEA model, particularly the Banker-Charnes-Cooper model, finds efficiency scores based on inputs such as lead time and costs per shipment and outputs such as on-time delivery rate and sales per customer. These efficiency scores are further fed, along with temporal features, to the TCN model, which captures historical patterns with dilated causal convolutions and residual connections to predict the probability of incidence of late deliveries. The integrated DEA-TCN model surpasses conventional approaches, with 99 % accuracy, 99 % precision, and 98 % recall. It serves to distinguish inefficient suppliers as well as high-risk transactions for making pre-emptive decision-making. Experimental evidence substantiates the viability of forecasting disruption and improving supply chain resilience. This transparent, data-driven framework gives managers actionable insights to optimize supplier selection, diminish operational risks, and improve delivery reliability. The matrix with a DEA-TCN foresees the strategic planning and strengthens the adaptability of supply chains in highly dynamic environments.

1. Introduction

Supply chain management is a backbone to all industries since it supplies products and services from producers to consumers. Today's supply chains are particularly challenging to manage due to globalization, the expansion of customers' needs, and the complexity of processes. The current supply chain models, particularly those that utilize manual processes and disconnected databases, can successfully tackle these challenges [1]. This is so because challenges associated with the system include a lack of real-time visibility of stock, poor management of inventories, delayed shipments, and high operating costs, among others. These challenges are compounded by the earnings pressure and flexibility in a market that changing customer expectations demand [2]. The rapidly emerging field of the IoT and the use of big data analytics present an opportunity to overcome these difficulties. For instance, IoT devices like sensors and RFID tags can gather massive quantities of real information at different points of the supply chain [3]. Such data can be used to get some insight into the operations and make some forecasts, as well as make sound decisions [4]. Risk is inherent in supply chain management, and international trade leads to developing international supply chain networks. Growth in international trade and globalization increases the risk and degrades the supply chain management risk [5].

The most valued expense in manufacturing companies is the cost of the supply chain [6]. However, in the banking and insurance sectors, risk in supply chain management is taken to mean that it does not directly affect the financial outcome in terms of risks most generally

* Corresponding author.
E-mail addresses: mabushaega@jazanu.edu.sa (M.M. Abushaega), omoshebah@kku.edu.sa (O.Y. Moshebah), amhamzi@jazanu.edu.sa (A. Hamzi), syalghamdi@kku.edu.sa (S.Y. Alghamdi).

https://doi.org/10.1016/j.aej.2025.05.062
Received 10 March 2025; Received in revised form 4 May 2025; Accepted 18 May 2025
Available online 28 May 2025

understood. Recognizing the risks involved in managing a global supply chain has raised moderate interest among scholars and practitioners [7]. Various risks associated with business enterprises in supply chain management are increasing, posing threats to the economy and the pattern of the network [8]. The supply chain has always been disrupted, even before the term "supply chain management" was invented. Over the decades, disruptions have changed, ranging from supply chain disruptions brought on by shifting ecosystems and business models to cyberattacks by new technology. Global supply chains are becoming more fragile due to disruption risks that threaten societal stability and security either directly or indirectly. For instance, the global movement of persons and goods has been impacted by the proliferation of COVID-19 [9]. The ensuing supply chain interruptions have caused product shortages and delivery delays. At the same time, demand for some products used to combat epidemics has exploded, and it is now more difficult to predict demand trends for a wide range of consumer goods. Following the difficulties brought on by the need to stop the virus's spread, companies' largest difficulty is resuming their supply networks. The risk associated with supply chain disruptions can take various forms, but in recent years, the academic community has focused on those associated with cyberattacks [10].

Supply chain management is central to increasing business relationships in improving operational effectiveness and satisfying customers within diverse economic sectors [11]. Accurate management of the supply chain results in cost savings, service quality enhancements, and long-term competitive advantage for the firm. As organizations are under pressure to provide transparency, flexibility, and resilience, more industries are opting for new technologies, including Artificial Intelligence, Machine Learning, Blockchain, and Advanced Analytics, transforming traditional supply chain models [12]. They improve the extent of the predictor's credibility, the credibility of stock control, and the solidity of the risk management plans [13]. Due to these risks and a constantly changing business environment, there is a need for new ideas that will help organizations increase their supply chain operational performance and reduce risks [14]. Data Envelopment Analysis (DEA) represents a flexible evaluation technique that institutions, companies, healthcare organizations, airlines, and government entities utilize to determine the efficiency of decision-making units (DMUs) [15]. The assessment method extends its application to determine organizational risk levels [16]. DEA gives organizations an analytical system to effectively measure their operational risk levels [17].

1.1. Problem statement

Modern supply chains face many operational efficiencies and risks related to costs, such as supplier delivery delays and order fulfillment failures. Many existing predictive models, including classical machine learning approaches, fail to meet the requirements primarily because of their limited transparency, poor adaptation to time-series data, and being out of scope for real-time decision priorities. Moreover, common models tend to overlook the dual requirement of assessing supplier performance and predicting delivery risks [18]. Although DEA provides a very effective, interpretable method for evaluating supplier efficiency, its use is short-sighted from a predictive standpoint. Conversely, TCN is typically a good fit for forecasting time-series data and has yet to be joined with efficiency analytics. A complete literature survey finds a dearth of hybrid models, denoting the combination of DEA and deep learning for improving resourcing supply chains' resilience. This study will fill in this gap by proposing a novel DEA-TCN model, which integrates the predictability of TCN with the interpretative ability of DEA, supplying real-time risk measures and actionable insights about suppliers, hence having more informed, resilient supply chain decision-making [19].

1.2. Research motivation

The proposed model integrated DEA and temporal convolution networks (TCN) to maximize the complementary advantage of such techniques to predict risks in supply chains. DEA is well-known as a recognized application, and this quantifies relative efficiencies among the most effective decision-making units in terms of multiple input-output relationships. They are very useful in identifying the underperforming suppliers and optimizing resource allocation by using transparent and interpretable efficiency evaluations. Research has modelled DEA only temporally, lacking temporal contextuality, which would increase the predictive power. Thus, TCN will be incorporated here, and it is known to capture long dependencies within time series data using its dilated causal convolutions and residual connections. It can be trained as fast as possible, parallelized well, and has increased stability for forecasting over long sequences, compared to recurrent models like LSTM. Hence, DEA-TCN is a promising real-time efficiency evaluation and risk assessment model.

1.3. Research significance

Through its combination of DEA and TCN, the proposed model detects inefficient suppliers and accurately forecasts delivery delays, boosting supply chain reliability. The research combines DEA for transparent assessments with TCN for time-series forecasting to enhance decision quality, optimize logistics choices and supplier selection, and manage supply chain risk.

The key contributions of the proposed study are as follows:

- Established an integrated DEA and CNN for efficient supply chain risk prediction, combining efficiency analysis with deep learning.
- Applied DEA (BCC model) to evaluate supplier performance, identifying inefficient suppliers contributing to late deliveries.
- Implemented TCN to analyze historical supply chain trends, providing accurate predictions of late delivery risk for proactive decision-making.
- Provided data-driven risk mitigation strategies, enabling real-time supplier monitoring, optimizing logistics, and improving on-time delivery rates.

The remainder of the paper is arranged as follows. In Section 2, related studies are presented. The proposed methodology is presented in Section 3. Findings and summary are provided in Section 4. In Section 5, the conclusion and further research are provided.

2. Literature review

Mittal et al.,[20] study focuses on the complexities and weaknesses in supply chains, which are frequently impacted by outside disturbances like inflation, pandemics, and conflict situations. The goal is to create an AI-powered system that can precisely assess these nuances in the field and successfully address their vulnerabilities. Using datasets from many investigations, the work uses an empirical approach to build Deep Learning (DL) and Machine Learning (ML) models. With an accuracy rate of almost 90 %, the Deep Convolution Neural Network (DCNN) regression model performs better than the other models in forecasting supply chain hazards. Complex and nonlinear interactions between the variables will be easier for the constructed model to handle. The findings demonstrate how ML and DL may improve supply chain networks' predictability and resilience in the face of growing hazards.

Kosasih et al. [21], provide a neuromyotonic machine learning method to proactively find new data and hidden supply chain vulnerabilities. Knowledge graph reasoning and graph neural networks are combined within the approach. The ability of the model to infer various hidden connection risks sets it apart from previous studies and represents a significant advancement in computerized supply chain surveillance. Two empirical datasets from the energy and automotive industries were used to test the approach, showing that it can infer information from various links, including companies, goods, production capabilities, and certifications. This allows for more complex queries than just who supplies whom. As a result, graph structure can reveal more risk insights, giving practitioners new insights.

J. Thomas et al., [22] state that firms are rethinking data analysis and supply chain modifications with the advent of AI. Supply chain management researchers have attempted to explore and mitigate the risks of ML and DL in the prediction and classification of data. This research investigates several approaches of DL, specifically Bi-LSTM, to ascertain whether late deliveries are a consequence of unidentified drivers in intricate supply chain networks. The Kaggle dataset, "Datacom Smart Supply Chain for Big Data Analytics," is preprocessed, normalized, and divided into training and test sets. A BiLSTM model predicts major drivers of delayed deliveries. Performance measures are an F1-score of 90.11 %, AUC-ROC of 14.01, and accuracy of 97.59 %. Comparisons with KNN and Decision Trees emphasize BiLSTM's effectiveness in reducing supply chain risks.

Disrupting events present big problems for the SCNs because they rely on transport networks. To achieve an optimum SCN performance, these strategies have to be synchronized and planned according to efficient restoration plans for the transportation network (TN) and distribution plans for downstream SCN in the Abushaega et al. study [18]. To address disruption in transport networks in addition to addressing the issue of fairness, a novel nonlinear mixed integer programming (NMIP) model is developed to solve the problem. In the model, the TN is given optimal restoration strategies following unpredictable events, considering cost-benefit analysis per demand node in the complex network. Commodity transportation in Colombia shows that if a distribution is based on fairness, then the satisfaction rates attained are faster than when based on cost-effective restoration plans.

C. Ju et al.,[23] study on healthcare supply chains faces significant disruption risks due to the need to handle varying demand levels and delivery issues in addition to unforeseen emergencies. A study analysis assesses the performance of Long-Short Term Memory (LSTM) networks with attention mechanisms for forecasting supply chain disruption risks from various data streams. Operation on three data sets spanning 2021–2024 produces results as high as 91.8 % when contrasted with distinct-source methods, which return accuracy at 15.4 % higher. The test proves how the system repeatedly executes with accurate outcomes during all observed durations. The flexible solution yields more reliable forecasts and functionality value for managing healthcare supply chain risk.

N. Rezki et al.,[24] research on the rapid business environment of today makes supplier tardiness in fulfilling orders one of the main drivers that adversely impact supply chain efficiency. Risk mitigation requires identifying risky suppliers and accurate predictions of possible disruptions. Machine learning (ML) applications enable companies to implement predictive risk monitoring because such systems generate future delay forecasts alongside suitable intervention methods. ML classification methods for predictive analysis of supplier performance rates. The predictive models with the best performance were RF Regression and GB Regression according to their metrics (RMSE: 1.81, MAE: 1.47, RMSE: 1.66, and MAE: 1.37). Logistic Regression and GB Classifier achieve the best results in precision, recall, and F1-score evaluations, thus creating higher supply chain resilience through operational disruption reduction.

A. Hatami-Marbini et al.,[25] created a risk management framework to enhance optimized Nigerian oil supply chain mitigation strategies. The framework provides complex support to researchers and decision-makers who must investigate new risk management strategies and their related challenges. The Data Envelopment Analysis methodology evaluates risk factors to help identify the best possible response solutions. Analytical results demonstrate that criminal activities, together with terrorist assaults, represent the major security dangers. Effective strategies for risk reduction involve transferring risks through proper planning, implementing alternative energy carriers, improving energy efficiency, creating rescue plans, addressing expected shortages, and maintaining diplomatic relationships. The analyzed research provides clear insights about the Nigerian oil sector's risk complications, which enables better strategic decision-making and operational planning in supply chains.

A. Alzahrani et al.,[26] demonstrated that IoT technology has substantially enhanced logistics aspects, including storage management and communication systems, as well as service quality and supply chain management systems. The research creates an intelligent supply chain management system that enables proper decision making for effective IoT-based logistics operations. Studies that evaluate shipping risks in natural disasters using traditional feature-encoding and machine learning techniques have resulted in unexpected findings, thus compromising the research validity. The former research used deep neural models to extract features but did not monitor sequence data. The present research implements a CNN-BiGRU deep learning combination to evaluate shipping possibilities between geographic locations. The created framework delivers results with an accuracy exceeding 94 % above standard baseline performance metrics.

Table 1 compares different supply chain optimization techniques, outcomes, and drawbacks. While extensive studies have been conducted using various deep learning techniques for modeling supply chain risk prediction, such as LSTM, CNN, and hybrid architectures, Temporal Convolutional Networks (TCN) have been considered only faintly, even though they are most advantageous in the modeling of long-range

Table 1
Literature review summary.

Reference	Method	Findings	Limitations
Mittal et al., [20]	Deep CNN Regression, Fuzzy C-means	Achieved ~90 % accuracy in predicting supply chain hazards, improving stability using AI models.	Does not address real-time adaptability to sudden supply chain disruptions.
Kosasih et al., [21]	Neurosymbolic Machine Learning, GNN, Knowledge Graph Reasoning	Enhanced detection of hidden supply chain vulnerabilities using knowledge graph insights.	Limited to structured datasets; real-world unstructured data integration is unclear.
J. Thomas et al., [22]	Bi-LSTM	Bi-LSTM achieved 97.59 % accuracy, 90.11 % F1-score, and 14.01 AUC-ROC.	Requires high computational resources for training.
Abushaega et al., [18]	Nonlinear Mixed Integer Programming	Fairness-based distribution strategy improves supply chain recovery rates.	Focuses on Columbia's transport network; applicability to other regions is unclear.
C. Ju et al., [23]	LSTM with Attention Mechanism	Achieved 91.8 % accuracy in predicting healthcare supply chain risks.	Model performance may vary with different datasets.
N. Rezki et al.,[24]	KNN, RF, GB, DT, SVM, Logistic Regression	RF and GB provide superior predictions of supplier delays.	No evaluation of real-time decision-making capabilities.
A. Hatami-Marbini et al.,[25]	Data Envelopment Analysis	Identifies criminality and terrorism as primary risks in the Nigerian oil supply chain.	The DEA model assumes static risk factors and lacks real-time adaptability.
A. Alzahrani et al.,[26]	CNN + BiGRU Hybrid Model	94 % accuracy in IoT-based supply chain risk prediction.	Feature extraction methods need further validation for diverse scenarios.

temporal dependencies in time-series data. Data Envelopment Analysis (DEA), on the other hand, is mostly applied in different contexts of supply chains for performance evaluation, such as in supplier selection and operational benchmarking, leaving a few uses for real-time risk prediction coupled with it. An evident point in the literature is the gap about hybrid models that would empower DEA in measuring efficiency in tandem with deep learning methods' predictive capabilities, specifically TCNs. Most literature either undertakes predictive modeling, ignores interpretability, or focuses on efficiency analysis, with no concern for historical contexts in making predictions. This, therefore, presents an exciting avenue for research in developing a unified framework featuring both the explanation of supplier performance and the prediction of delivery risks on the other side in such an ever-dynamic environment as the supply chain. In bridging the gap, DEA-TCN analytics integrates DEA efficiency scoring with TCN for sequence learning, thus providing an all-encompassing supply chain monitoring, supplier evaluation, and disruption forecasting scheme.

3. Proposed methodology

The proposed DEA-TCN hybrid model combines data envelope analysis to evaluate efficiency with temporal convolutional networks for a time-series forecast of risks to maximize supply chain performance. A supply chain dataset with shipment information, prices, and delivery times is first preprocessed by treating missing values, feature normalization, and reorganizing data into a form suitable for time series. Critical supply chain measures like shipping days (actual and planned), benefit per order, customer sales per customer, and risk of late delivery are derived. Second, DEA under the BCC model assesses the operational performance of supply chain units, including suppliers and logistics centers. Lead time, cost per shipment, and inventory turnover are inputs; on-time delivery rate, late delivery risk, and sales per customer are outputs. DEA calculates efficiency scores in which the value is closer to 1 to represent best performance, and lower than 1 to represent inefficiencies. Historical DEA efficiency scores are combined with time-series features in the TCN model to forecast future supply chain risks. Dilated causal convolutions preserve long dependencies, and residual connections allow free gradient flow. The fully connected layer gives output as the likelihood of late shipments and operational inefficiencies. Lastly, ineffective suppliers are highlighted for intervention, allowing for proactive supply chain optimization and risk mitigation measures to ensure better logistics performance.

Fig. 1 defines a supply chain risk forecasting process. Beginning with data gathering from different sources, it goes to data preprocessing, such as missing value handling, normalization, and feature engineering. Efficiency scores are computed by Data Envelopment Analysis, which a TCN model further utilizes to forecast delivery risk. The prediction identifies deliveries as "late" or "no risk of late delivery."

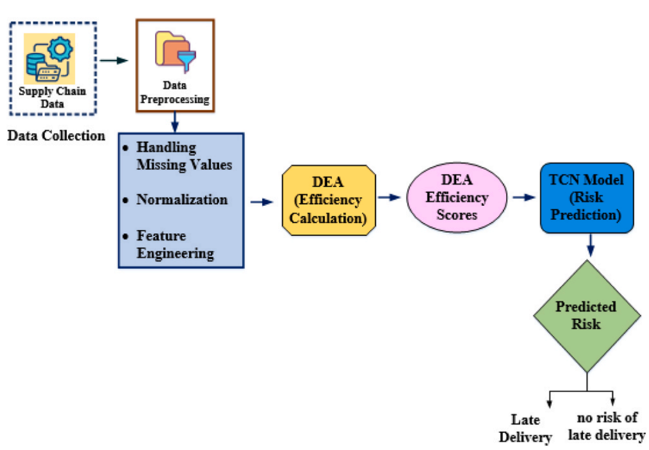

Fig. 1. Supply Chain Risk Assessment Framework.

3.1. Data collection

The proposed study utilizes supply chain data from the public repositories belonging to Kaggle during the data collection stage [27]. The study uses supply chain transaction data from the 'Dataco Smart Supply Chain' dataset on Kaggle, which speaks to supplier-to-customer interactions. This model was trained and tested on a single public dataset, 'Dataco Smart Supply Chain' from Kaggle, which contained more than 200,000 transactions. Thorough time-series analysis depends on the "Dataco Smart Supply Chain" dataset, which includes multiple vital elements, including order dates, ship dates, and delivery statuses. The supply chain operations received a comprehensive evaluation by combining supplier performance metrics with transactional records and logistics data points.

3.2. Data preprocessing

The DEA-TCN framework operates in a sequential pipeline whereby raw supply chain data undergoes a deep cleansing and preprocessing procedure to extract relevant features such as lead time, cost per shipment, and delivery status. The extracted features would function as inputs into the DEA-BCC model to give rise to efficiency scores for each supplier. The efficiency score is a continuous variable showing the relative performance of suppliers; it is then merged with time-series data, leading to the capturing of past shipping behavior and outcomes concerning delivery. This enriched dataset is then input into the Temporal Convolutional Network, which captures sequential patterns over time to predict late delivery risk probability. All model components are interconnected, with the flow of information from numerical processing (DEA) toward sequence modeling (TCN), thus providing interpretability and predictive power. A binary risk classification and understandable metrics that guide managerial decision-making form the final output. The following Fig. 2 shows the working process of Data preprocessing:

a) Handling missing values

During preprocessing, missing data were addressed through data imputation methods. To impute numerical variables with missing data, the mean or median statistics are applied, but the most frequent value methods are used for categorical variables [28]. This approach enables complete datasets for subsequent modeling purposes, and the formula for this method appears in Eq. (1).

$$a_{imputed} = \frac{\sum_{i=1}^{n} a_i}{n} \tag{1}$$

Here $a_{imputed}$ is the imputed value, $a_i$represents the observed values, and n is the number of observations.

b) Feature engineering

Data preprocessing needs feature engineering as an essential step to develop relevant data features from raw information, which boosts model performance. Suppliers who perform customer behavior analysis and sales forecasting benefit from vital features encompass lead time assessment and sales volumes at the customer level. Lead time is the duration from when customers place their orders until they receive them, shown in Eq. (2). The element substantially influences supply chain operations because it establishes inventory strategies and generates positive customer outcomes and operational performance results [29].

$$Lead\ Time = Order\ Delivery\ Date - Order\ Placement\ Date \tag{2}$$

In supply chains with reduced lead times, the operation runs more efficiently. The length of an order-supply process can suggest either inventory depletion problems, supplier process limitations, or supply chain delivery complications.

c) Normalization

The measurement of sales split by each customer is a critical value

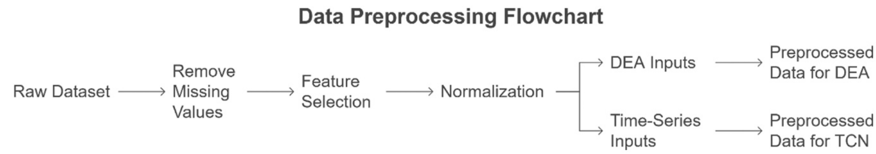

Fig. 2. Flowchart of Data Preprocessing.

indicator that helps markets segment and identify standard revenue frequencies. The normalization process allowed cost-related variables, including cost per shipment and transportation cost, to achieve standardization across all measurements for data comparison. The normalization technique protects the learning process from being controlled by significant values, which leads to better model performance [30]. The normalization formula can be expressed through Eq. (3).

$$c_{norm} = \frac{c - \mu}{\sigma} \tag{3}$$

Where $c_{norm}$is the normalized value, c is the original value, μ is the mean, and σ is the standard deviation of the feature. The dataset was divided using an 80:20 ratio, which trained the model properly using extensive data and validated its performance on an independent sample. The data partition creates conditions for measuring model generalization abilities and quantitative performance on new observations. The preparation process for the dataset depends on these steps to clean the data while applying scaling techniques and partitioning it for maximum model preparation efficacy.

d) Data cleaning and feature selection

The dataset underwent cleaning to eliminate irrelevant or redundant columns like product descriptions, customer names, and regional codes, which did not affect the evaluation of delivery risk or supplier efficiency. The principal variables kept for modeling included Lead Time, Cost per Shipment, On-Time Delivery Rate, Sales per Customer, and Late Delivery Status. These variables were chosen because of their direct impacts on supply chain performance and the fact that they were present in virtually all records. There was mean substitution of all missing values regarding the numerical fields to maintain consistency in the data, especially with reference to DEA input variables that must have complete numeric values. No categorical variables were used in the final model; the focus was on continuous and binary measures relevant for DEA and TCN processing. This preprocessing step ensured a clean and structured input for both efficiency analysis and risk prediction stages.

3.3. *DEA for supply chain efficiency analysis*

Data Envelopment Analysis - A popular method for measuring performance efficiencies in decision-making units that convert sets of multiple inputs to multiple outputs relatively [31]. Numbers are models of non-parametric linear programming schematics invented for measuring relative efficiencies across these DMUs. The method is applied widely across healthcare, education, logistics, and supply chain performance assessments [32].

The non-parametric DEA technique measures the relational efficiency levels of DMUs when used to analyze supply chain entities, including suppliers, warehouses, and transportation hubs. The fundamental concept in DEA identifies DMU as a unit of operation or individual entity subject to efficiency assessment. Analysts can determine efficient and inefficient units through DEA by comparing their performance with other units. DEA tracks the performance of these units through output and input evaluations, which specifically identify delivery locations or suppliers who present supply chain risks. The efficiency assessment of supply chain entities through DEA includes considering both their input resources and output performance elements. The resources that a supply chain entity uses to operate serve as its inputs.

3.3.1. *Inputs (Resources consumed)*

The period businesses use to carry out deliveries describes true logistics performance. The duration of delivery time directly correlates with operational deficiencies in the fulfillment system. The scheduled delivery duration reflects the anticipated logistics efficiency through the Days for Shipment (Scheduled). The size of the discrepancy between planned shipment days and actual shipping days points toward scheduling or delivery management problems. The profit generated from each order represents the Benefit per Order USD after subtracting all associated costs. Lower benefits per order signal inefficiency, either due to higher shipping costs or poor pricing strategies.

3.3.2. *Outputs (Results produced)*

The data set contains a binary value that indicates either late delivery status (1) or on-time delivery status (0). The supply chain delays that arise pose performance risks, which will have a direct impact on operational efficiency. Sales per Customer (USD): The average sales per customer for each order. The revenue generation efficiency from each customer becomes the focus of this output measurement. Successful order fulfilment, as well as well-managed supply chain operations, produce higher sales figures. The standard DEA models are very suitable for the continuous input-output scenario. Introducing binary variables like late-delivery risk, which categorizes itself (0 = no-risk, 1 = late-delivery), into this study creates a kind of paradox. Since standard DEA holds non-negative, continuous data, the input of binary variables creates a riddle in itself. The recommended procedure of theoretical viability requires employing Binary Integer Programming DEA (BIP-DEA), which is specially catered to mixed binary-continuous data. BIP-DEA is thus capable of fitting the critical binary output, such as risk classification, without violating the fundamental assumptions in DEA [33].

DEA uses weighted output and weighted input ratios to determine the operational efficiency of each DMU. DEA measures actual performance results of DMUs relative to their most optimal operational levels. The linear programming model that DEA implements determines the overall efficiency ratings for each DMU. The basic model of DEA declares that every DMU needs to achieve its maximum possible ratio between outputs and inputs. Let's assume n is the number of DMUs. m is the number of inputs. s is the number of outputs. Let: x_{bc} represent the amount of input c used by DMU b. z_{bc} represent the amount of output c generated by DMU b. λ be the weight assigned to each DMU in the calculation of efficiency.

The assessment of efficiency scores by DEA occurs based on how each DMU performs regarding its inputs and outputs. The BCC model serves as the main selection because it accommodates changing scale effects. The model distinguishes different supply chain entities because they possess varying scale sizes between increasing and decreasing returns to scale properties according to their size. The efficiency score (θ_c) emerged when the weighted outputs were divided by the weighted inputs. Eq. (4) represents the approach to computing the efficiency score θ_c of a specific DMU c.

$$\theta_c = \frac{\sum_{r=1}^{s} \lambda_r y_{rc}}{\sum_{b=1}^{m} \mu_b z_{bc}} \tag{4}$$

Where: θ_c represents the efficiency score of DMUc. y_{rc} is the output r for DMU c (e.g., on-time delivery rate, customer satisfaction index). z_{bc}: Input c for DMU c (e.g., supplier lead time, cost per shipment). λ_r means

Table 2
Example of DEA working.

Supplier	Days for Shipping (Real)	Days for Shipment (Scheduled)	Benefit per Order (USD)	On-Time Delivery Rate (%)	Late Delivery Risk	Sales per Customer (USD)
A	3	4	91.25	95	0	314.64
B	5	4	−249.09	70	1	311.36

weight for output r (determined by the optimization process). μ_b represents weight for input b (determined by the optimization process). The values λ_r along with μ_b are obtained through linear programming optimization by maximizing outputs' weighted values against inputs' weighted values.

3.3.3. DEA optimization problem (BCC model)

Each DMU achieves its maximum efficiency score θ_c through the implementation of the BCC model. The BCC (Banker Charnes Cooper) DEA model functions as an efficiency measurement tool that works under variable returns to scale (VRS) principles. This method assesses suppliers through the measurement of their applied resources and their achievement outcomes. BCC model evaluates supplier efficiency through linear programming problem solving that determines each supplier's relative performance level against the most efficient counterparts under variable scale efficiency conditions. The optimization problem takes this structure in Eq. (4):

$$Maximize\theta_c = \frac{\sum_{r=1}^{s} u_r y_{rc}}{\sum_{b=1}^{m} v_b z_{bc}} \tag{4}$$

Where i denotes the amount of input used by DMU c, r denotes the amount of output produced by DMU c, and the weight of the output is denoted as u_r and the weight of the input is denoted as v_b, the efficiency score of DMU is denoted as θ_c

Subject to: A complete mathematical representation exists to calculate the efficiency of each DMU by utilizing Eq. (5).

$$\frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{b=1}^{m} v_b k} \leq 1 \text{for all k} = 1, 2, 3 \ldots \text{n} \tag{5}$$

The weighted values for outputs and inputs need to be nonnegative as per Eq. (6).

$$u_r \geq 0 and v_b \geq 0 \forall r, i \tag{6}$$

The weight-referred ratio with respect to the outputs to inputs is rightly represented by this formulation and also conforms to the other assumptions of the DEA-BCC model in a systematic manner. The optimization process enables the DEA model to calculate efficiency ratings for each logistics organization. A DMU possesses efficient operations when its efficiency score reaches 1. Otherwise, an efficiency score of less than 1 indicates inefficient operations. The solution of the optimization problem results in θ_c efficiency scores are assigned to every DMU. The interpretation is as follows:

$\theta_c = 1$: The DMU is efficient, operating at optimal performance relative to other entities.

$\theta_c < 1$: The DMU is inefficient, and there are opportunities for improvement. $\theta_c \geq 1$

Entities scoring below 1 on the efficiency test enter the risk analysis process. The subpar function of these entities might trigger supply chain disruptions because of transportation issues, elevated operational costs, and poor service maintenance. The detection process enables entities to receive priority optimization through TCN to forecast upcoming risks through historical performance evaluation. The measurements derived from DEA are incorporated at different stages of the TCN prediction model to forecast upcoming disruptions. The efficiency scores help TCN build forecasts of upcoming delays and disruptions by finding patterns in historical data time sequences. Table 2 presents an application of the DEA (BCC Model) evaluation as depicted in the following example: Using data from two suppliers named Supplier A and Supplier B, we will now demonstrate an example for these companies.

In Table 2 above, Supplier A is the efficient supplier with 95 % on-time delivery and no risk of delay (risk=0), while Supplier B displays some inefficiencies due to lesser on-time delivery and a delay risk (risk=1). The DEA model gives Supplier A the perfect score of efficiency (1.00) and ranks it as a benchmark performer, while giving Supplier B a rating of less than 1.00, implying it must improve its performance and risk mitigation.

3.4. TCN for supply chain risk prediction

In the present work, a hybrid DEA-TCN model is proposed by integrating Data Envelopment Analysis (DEA) and Temporal Convolutional Networks (TCN) to predict supply chain-related risks concerning late deliveries. DEA efficiency scores are filled into TCN input sequences as continuous temporal features, which allow the model to learn temporal patterns in supplier performance apart from transactional variables. These are not binary indicators but rather real-valued scores (0–1 range) that directly affect predicted risks of late delivery. Performance evaluation starts with the assessment of the supplier under the BCC-DEA model concerning the key input and output metrics of supplier lead time, cost per shipment, on-time delivery rate, and sales per customer. These so-determined efficiency scores, which represent the operational performance of each supplier, are then conveniently integrated into an augmented time series dataset reflecting other transactional features of relevance. This augmented dataset is used to train the TCN model, which is suited for long-term temporal dependencies and sequential patterns in historical supply chain data. The risk assessment methodology surpasses conventional methods because of the parallel data processing and comparatively stable gradient paths maintained by dilated convolutions and residual connections, thus enabling improved accuracy of risk forecasting. The output of the model accounts for a probability score for classifying the delivery of each supplier as high or low risk, thereby enabling real-time risk flagging and prompt decision-making. Thus, this combined approach will help supply chain managers signal performance deterioration, observe supplier risk trends, and initiate corrective actions to improve supply chain resilience. Eqs. (7) to (12) are sourced from the original work on the Temporal Convolutional Networks by Bai et al. [34], wherein the authors introduced dilated causal convolutions and residual connections for modeling sequences. These structures have demonstrated great capabilities regarding performance on the time series tasks, usually going beyond recurrent models on many benchmarks.

The system requires historical data sequences to process through TCN. One of the features in the dataset is the efficiency score, which the system incorporates. Time-Series Input Format for TCN is given in Eq. (7). When operating at a specific time t, TCN receives several inputs:

$$X_t = \{\theta_c, Days\ for\ Shipping, Days\ Scheduled, Late\ Delivery\ Risk, Sales\ per\ Customer, Benefit\ per\ Order\} \quad (7)$$

Where days for shipping (Real) means the actual time taken for delivery, days for Shipment (Scheduled) means the expected time for shipment. Late Delivery Risk is represented in a Binary indicator (1 = Late, 0 = On-time). Sales per Customer refers to Revenue per customer over time. The calculated value for Benefit per Order represents Profit per Order. The time interval t carries history-based inputs of these attributes, while the model calculates future supply chain disruption probability. The DEA-TCN working architecture is shown in Fig. 3.

The design uses sliding window treatment to detect patterns in time sequences. The Sliding Window Method serves to organize time-series data for risk predictions made using TCNs. The approach extracts sequences of fixed length from supplier historical data to maintain model performance for efficiency score trends, along with shipping delays and cost evaluation information. A continuous window segment represents the time period that past observations use to predict future risks. Using this approach, the TCN system develops the capability to predict delayed deliveries more accurately by recognizing patterns in sequence. The training process of TCN operates through historical data points, which include efficiency scores and supply chain metrics to predict upcoming scores. The TCN analyses time-series sequences with dilated causal convolutions that consider past time steps while capturing extended temporal associations between inputs and conducting parallel processing that differs from RNN series model structures. The mathematical expression is described as Eq. (8) defines this representation.

$$y[i] = \sum_{k=0}^{k-1} x[i-d-k].w[k] \quad (8)$$

Here $y[i]$ means output at time i. $x[i-d-k]$ is the input from previous time steps. $w[k]$ is the convolution filter weights. d is the dilation factor (controls how far back the network looks in history). The smooth transfer of gradients through a network series of convolutional layers depends on residual connections within the TCN architecture.

$$E_t = X_t + F(W * X_t) \quad (9)$$

In Eq. (9) X_t is the input sequence. $F(W * X_t)$ means feature transformation using convolution. A model's receptive field establishes what portion of information from the past the model can examine.

$$Receptive\ Field = 1 + (K-1) \times (2^L - 1) \quad (10)$$

In Eq. (10) K means kernel size. L represents the number of layers in TCN. The system generates the hidden representation using the final convolutional layer by proceeding through all historical data of efficiency scores and supply chain variables. The system delivers this information to the fully connected layer for the completion of the risk assessment process. The TCN system uses its learned trends to forecast delivery time performance, which will either exceed or meet delivery deadlines. The last output follows the format specified in Eq. (11).

Table 3
Parameter distribution.

Symbol	Definition
t	Time step index
x_t	Input feature vector at time t
y_t	Actual delivery outcome (0 = On-Time, 1 = Late)
O_t	Predicted risk probability at time
h_t	Hidden state from TCN at time ttt
W, b	Weights and bias in the fully connected prediction layer
d	Dilation factor in TCN
K	Kernel size of convolutional filters
L	Number of convolutional layers in TCN
σ	Sigmoid activation function
θ_c	DEA efficiency score for supplier j

$$O_t = \sigma(w_f * H_t + b_f) \quad (11)$$

Where O_t refers to predicted supply chain risk (probability of late delivery). σ means sigmoid activation function (for binary classification). H_t is the final hidden state containing efficiency trends, w_f represents fully connected layer weights. b_f means bias term. The sigmoid function ensures the output probability is between 0 and 1, and is denoted in Eq. (12).

$$\sigma(c) = \frac{1}{1+e^{-c}} \quad (12)$$

One score represents the risk assessment of the information after going through the fully connected layer. Through the sigmoid function, the score receives a probability transformation between 0 and 1 and enables a decision between no risk for on-time delivery and late risk for delayed delivery. The system determines supplier classification according to a specific preset threshold after it predicts the late delivery risk probability through the TCN model (O_t). The supplier receives a "No Risk" (0) assignment when $O_t$is less than 0.50, which means the order will be delivered on time. The supplier receives the "Late Risk" tag (1) when O_t reaches or exceeds a value of 0.50 since this indicates substantial potential delivery delays. Order status receives the risk tag for further modification. Table 3 shows the parameter attribution.

The analysis of "Late Risk" suppliers requires in-depth examination through which decision makers suggest replacement suppliers as well as logistics optimization and alternative delivery route selection. Suppliers labeled as "No Risk" operate normally without any disturbance to their standard workflow.

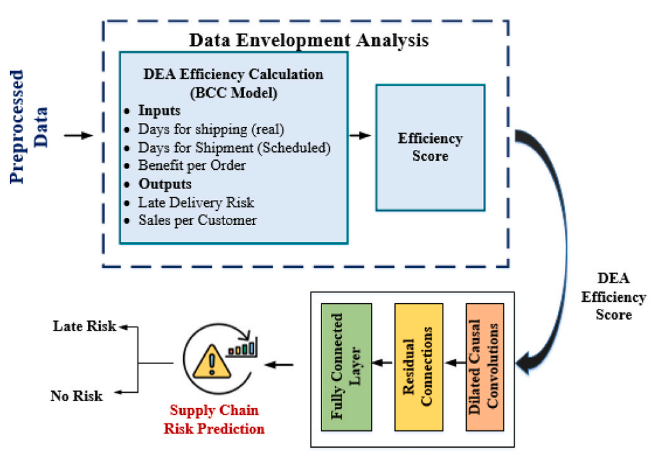

Fig. 3. DEA-TCN Architecture.

Algorithm 1. DEA-TCN Model for Supply Chain Risk Prediction

```
START
Input: Supply chain dataset DDD with features including lead time, shipment days,
delivery status, sales per customer
Output: Predicted Late Delivery Risk (0 = No Risk, 1 = Late Risk)
LOAD Supply Chain Dataset()
Step 1:Data Preprocessing
  FOR each column in Supply Chain Dataset:
    IF missing values exist THEN
       Impute using mean/mode OR Remove row
FOR each numerical feature in {Days for Shipping, Cost per Shipment, Sales per
Customer}:
    Normalize using Min-Max Scaling
Dataset = Convert to Time Series(Supply Chain Dataset, window size=n)
Step 2: Compute DEA Efficiency Scores
   FOR each supplier C in Dataset:
    Compute Efficiency Score(θ_c) using Eqn (4) to (6)
    IF θ_c< 1.0 THEN
       Flag Supplier as "Inefficient"
    ELSE
       Supplier is "Efficient"
Step 3: Train TCN Model for Risk Prediction
   Initialize TCN Model ()
   Train TCN Model using:
    Dilated Causal Convolutions
    Residual Connections
    Fully Connected Layer
Step 4: Predict Future Late Delivery Risk
FOR each new order in Dataset:
    θnew = Compute Efficiency Score (new order)
   O_t = TCN Model Predict(θnew, new order)
    IF O_t< 0.50 THEN
       Risk Label = "0"  # No Risk
    ELSE
       Risk Label = "1"  # Late Risk
    Update Order Status(new order, Risk Label)
Step 5: Decision-Making & Risk Flagging
  FOR each supplier in Dataset:
    IF Risk Label = "1" THEN
       Flag Supplier for Review
       Recommend Alternative Supplier OR Optimize Logistics
    ELSE
       Approve Supplier Transaction
END
```

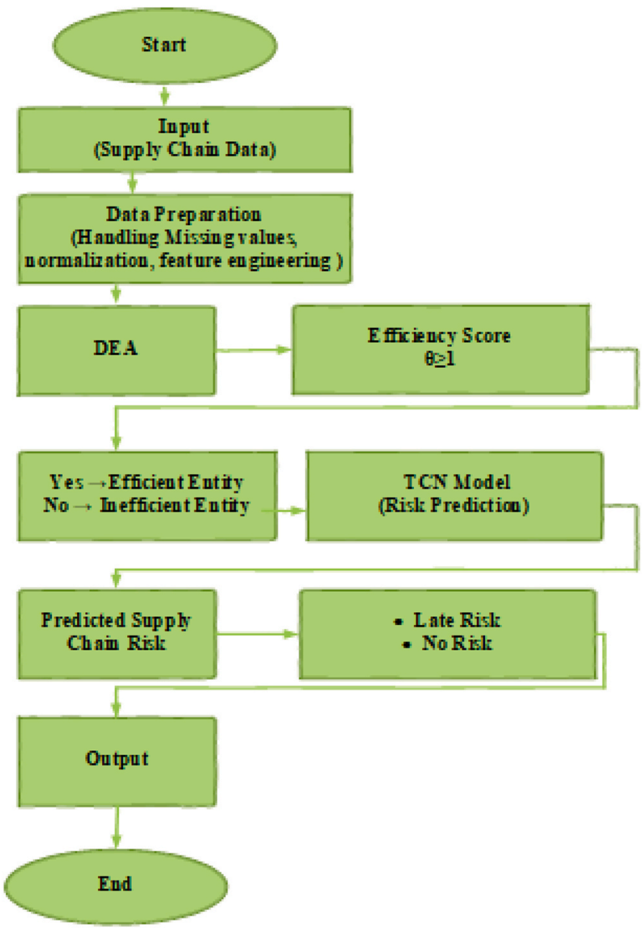

Fig. 4. Flow Chart.

Supply chain risk prediction achieves better results through the combination of efficiency analysis from DEA-TCN with forecasting based on temporal data is shown in Algorithm 1. The proposed method utilizes DEA (BCC model) to determine supplier and logistics hub efficiency through their input and output factors. Supply chain metrics join together with the scores before they are processed by TCN which uses dilated causal convolutions to detect temporal connections in the data. The model uses residual connections to preserve stable gradient flow until it moves data to the fully connected layer for future risk prediction such as late delivery assessments. Combining this integration allows organizations to act ahead of risks and efficiently manage supply chain performance which decreases disruptions in logistics operations. The model uses DEA efficiency assessment with TCN's forecasting capability to spot risks ahead of time which leads to enhanced supplier choices and logistical improvements and demand forecasting.

Fig. 4 shows a systemized method for supply chain risk forecasting. The process starts with supplying a supply chain dataset with important features such as shipping time, sales per customer, and late delivery risk. Quality assurance is guaranteed by preprocessing the data through imputation, normalization, and structuring in the time-series format. Efficiency scores are obtained through DEA, classifying entities into efficient and inefficient categories. Inefficient entities are then further examined with a TCN for predicting risk. The process is followed by a final decision-making phase.

Table 4
Simulation configuration.

Parameter	Description
Processor	Intel Core i7
RAM	16 GB
Operating System	Windows 10
Programming Language	Python
Batch Size	32
Learning Rate	0.001
Loss Function	Binary Cross-Entropy
Activation Function	Sigmoid
Evaluation Metrics	Accuracy, Precision, Recall, F1-Score, AUC-ROC
TCN Kernel Size	3
DT Batch Size	32
K in KNN	5
Learning Rate	0.001
Kernel Size (TCN)	3
Layers (TCN)	4
Activation	Sigmoid

Table 5
Delivery status distribution.

Delivery Status	Count
Late delivery	100,000
Advance shipping	50,000
Shipping on time	40,000
Shipping cancelled	15,000

4. Result and conclusion

The DEA-TCN hybrid model suggested measures of supplier efficiency and forecasted the risk of late delivery by merging Data Envelopment Analysis with Temporal Convolutional Networks. The output shows enhanced precision in detecting risky suppliers, thus facilitating early decision-making. Comparative analysis verifies that the model performs better than the conventional approach to supply chain risk forecasting. The model applied dropout layers in the TCN architecture to avert overfitting and adopted early stopping during training contingent on validation loss. These techniques guaranteed strong generalization, particularly because of the limited and imbalanced nature of the dataset.

Simulation parameters specify the computational environment and model settings for using the DEA-TCN hybrid model is shown in Table 4. Hardware requirements are a high-end CPU, GPU, and RAM for effective deep-learning training. Software dependencies are Python, and Windows OS. Hyperparameters like learning rate, batch size, and activation functions are tuned for stable training and maximum prediction accuracy. TCN parameters such as kernel size and dilation rates are adjusted for the modeling of temporal relationships in supply chain risk

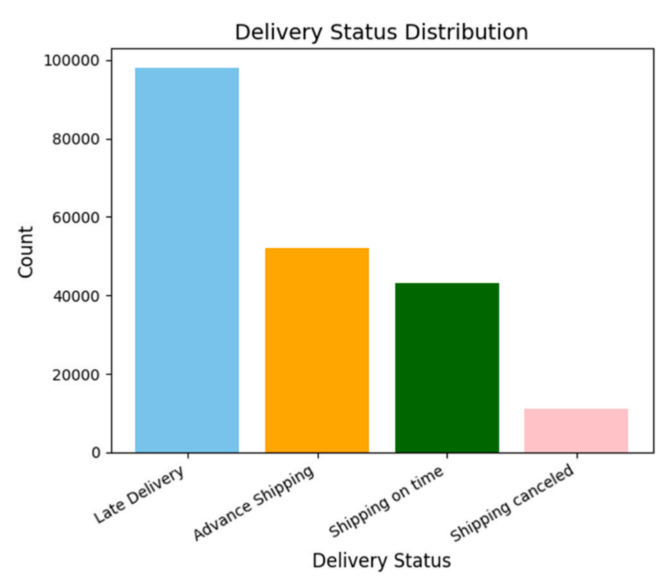

Fig. 5. Delivery Status Distribution.

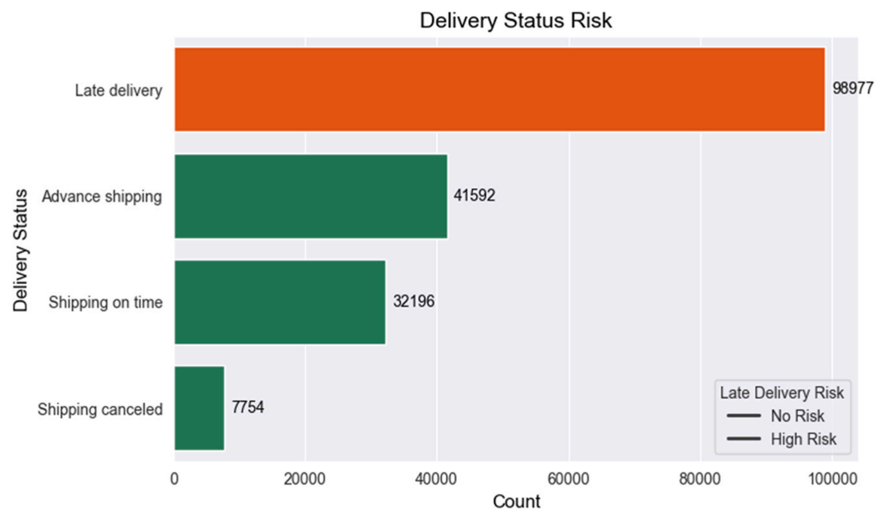

Fig. 6. Delivery Status Risk.

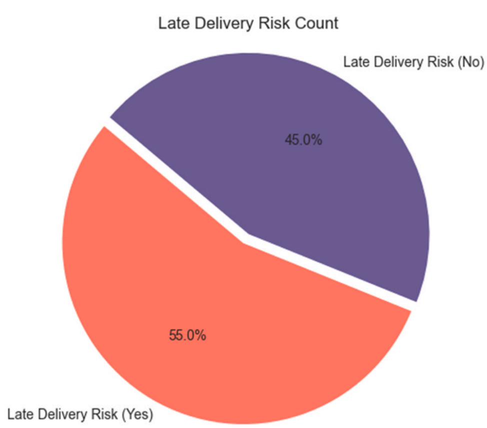

Fig. 7. Late Delivery Risk Count.

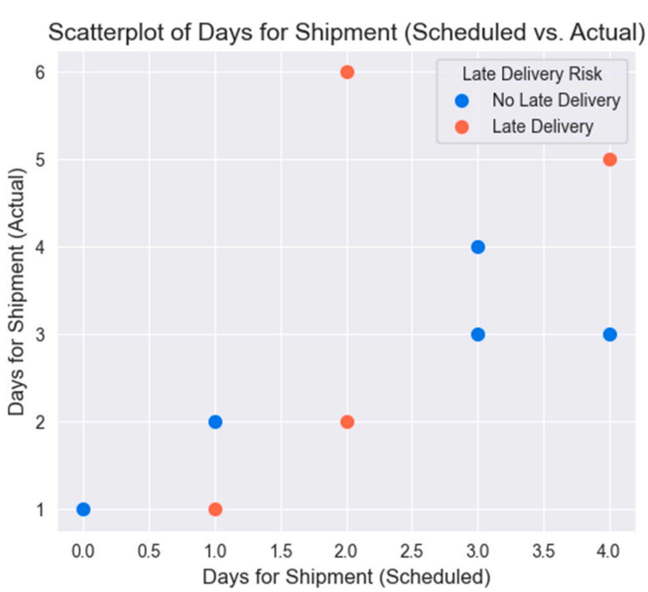

Fig. 8. Scatterplot of Days for Shipment.

forecasting. The DEA model employs the BCC method, providing a suitable efficiency measurement for suppliers. Five-fold cross-validation was undertaken during the evaluation of the model in order to ensure that the performance results were unaffected by an auspicious data split. The study followed an approach whereby the dataset included five equal subsets, all of which would at some point serve as a test set, while the other four would be pooled into the training data. In this way, variations due to train-test splits were reduced, and hence a more robust estimate of the generalization performance of the model was obtained. The DEA-TCN model was confirmed stable since accuracy scores and F1-scores matched consistently across the folds.

4.1. Experimental outcome

This section examines influential supply chain drivers of delivery efficiency. Delivery Status Distribution plots shipment completion patterns, and Delivery Status Risk classifies shipments into on-time vs. late deliveries. Late Delivery Risk Count identifies risky transactions, and the Scatterplot of Days for Shipment illustrates shipping time differences, which aids in the detection of supplier performance inefficiencies.

Table 5 captures delivery status distribution where late shipping is the most occurring at 100,000, followed by advance shipping at 50,000, and on-time shipping at 40,000. The shipping cancellations occur at least 15,000. From the data captured here, one observes inefficiency in delivery performance since many more shipments arrive late compared to when they are delivered on time. The supply chain being analyzed in this study is the supplier-to-customer stage, encompassing order processing and shipping lead time, along with delivery performance at the graded level. The dataset portrays numerous instances of inefficiencies, especially with lead-time management, as a high portion of deliveries were late. The findings underline the necessity for properly monitoring supplier logistics so that any issues can be flagged proactively, thereby ensuring a smooth downstream fulfillment and customer satisfaction.

Fig. 5 also illustrates the number and percentage of deliveries by various statuses. Many deliveries are also late, with around 100,000 cases. Advance shipping and on-time shipping are less reported, with approximately half the counts of delay reports, at 50,000 and 40,000, respectively. The least is that of shipping cancellations, the occurrence of which is less than 20,000.

Fig. 6 shows the delivery status distribution by late delivery risk breakdown. There is a maximum late delivery of 98,977 (risk=1). Advance shipping, 41,592; on-time shipping, 32,196; and canceled shipping, 7754, fall under no late delivery risk (risk=0). The chart indicates a strong late delivery problem, pointing toward better logistics and timely shipment tactics to improve delivery performance.

Table 6
Supplier efficiency rankings based on DEA scores.

Supplier	Lead Time (Days)	Cost per Shipment	On-Time Delivery Rate (%)	Sales per Customer	DEA Score	Efficiency Status
S1	3	120	95	500	1.00	Efficient
S2	5	160	80	450	0.85	Inefficient
S3	6	175	78	420	0.80	Inefficient
S4	4	140	92	480	0.95	Efficient
S5	2	130	90	460	0.98	Efficient
S6	7	190	75	410	0.76	Inefficient
S7	3	110	96	510	1.00	Efficient
S8	5	150	82	435	0.88	Inefficient
S9	4	135	93	470	0.97	Efficient
S10	6	180	70	400	0.72	Inefficient

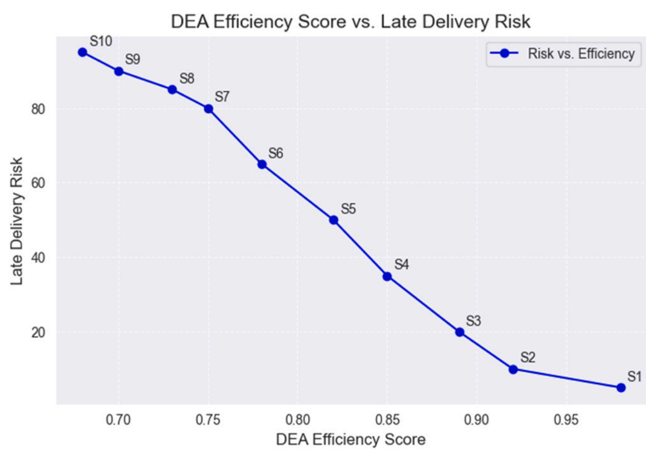

Fig. 9. Late Delivery Risk vs. DEA Efficiency Scores.

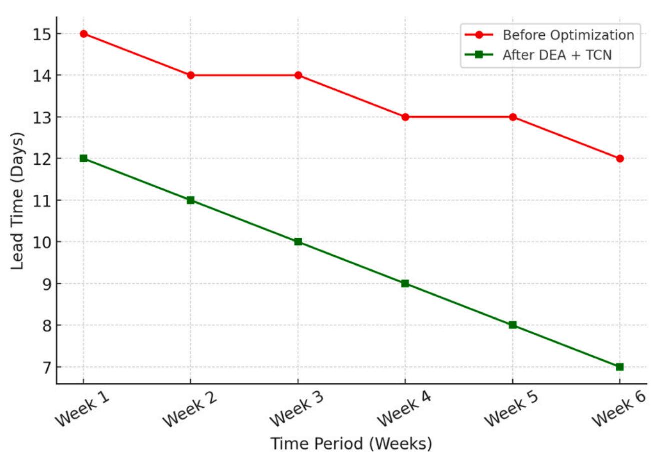

Fig. 10. Average lead time before and after applying the DEA-TCN framework.

Fig. 7 shows the share of late delivery risk in supply chain activities. Larger circle means the proportion of shipments with no risk of delayed delivery is about 66 %. Smaller circle mean the proportion of shipment with the risk of delayed delivery is about 33 %. smaller indicates the proportion of shipments with a late delivery risk (around 33 %). This implies that a large number of shipments in this supply chain may be prone delay. 0 means there is no risk of late delivery, while 1 indicates a high risk of late delivery.

Fig. 8 shows a relationship between scheduled and actual shipment days. Orders scheduled for 2-day delivery tend to experience the most delays, indicating that second-class shipping is associated with higher late delivery rates. Shipping modes have varying scheduled days: Standard Class (4 days), Second Class (2 days), First Class (1 day), and Same Day (0 days).

4.2. DEA-TCN performance analysis

This section assesses the supplier efficiency by employing Data Envelopment Analysis based on the BCC model. The DEA efficiency ratings measure the performance of suppliers in the efficient use of resources to deliver the best on-time delivery rate and sales per customer. While Table 6 gives an efficiency score of 1.00 for S1, Fig. 9 shows 0.98, since it adopts another evaluation setting that includes binary late-risk outputs in the full BIP-DEA formulation. The difference is deliberate to illustrate methodological variations.

The evaluation of key performance indicators in Table 6 analyses ten suppliers based on their use in the DEA model to evaluate lead time, cost per shipment, on-time delivery rate, and sales per customer. Analysis demonstrates that four suppliers, S1, S4, S5, and S7, accomplished DEA efficiency scores of 0.95 or higher, but Supplier S1 and S7 attained the maximum score of 1.00, making them efficient performers. The performance of these suppliers routinely delivered rapid lead times as well as diminished costs and precise delivery achievements. Supplier scores 0.72–0.80 indicate operational inefficiencies according to the assessment of S3, S6, and S10. Higher lead times in conjunction with lower delivery time accuracies contributed to making these suppliers efficiently subpar performers. Supplier performance gaps identified by the DEA evaluation help supply chain managers detect substandard suppliers while enabling them to determine improvement pathways or select different vendors from the market. The comprehensive performance analysis demonstrates that the model performs effectively when measuring supplier assessment through structured data-based processes.

The DEA method calculates supplier efficiency performance, which Fig. 9 displays jointly with supplier late delivery risk levels. The efficiency metric in DEA shows that S1 achieved 0.98 points, which is the best efficiency, while S10 reached 0.68 points, which is the least efficient score. The data shows that late delivery risk goes up as efficiency goes down because S1 has a risk of 5 % and S10 faces a risk of 95 %. The analysis shows an established negative correlation between supplier efficiency and delivery risks thus demonstrating why high-efficiency suppliers must become primary selection choices.

Fig. 10 shows the decrease in average lead time before and after applying the DEA-TCN framework of the DEA + TCN model. The lead

Table 7
Performance comparison.

Model	Accuracy	Precision	Recall	F1 score
DT	85 %	83 %	84 %	83 %
KNN	92 %	91 %	90 %	90 %
LGBM	95 %	94 %	93 %	94 %
RF	96 %	94 %	93 %	93 %
CNN-LSTM	91.5 %	91.5 %	91.6 %	92.5 %
GNN	95.5 %	94.15 %	87.76 %	90.5 %
GRU	98.13 %	97.82 %	84.53 %	97.3 %
Proposed DEA-TCN	99 %	99 %	98 %	99 %

times are initially relatively high, reflecting inefficiency in the supply chain. Post-optimization, there is a discernible decline, reflecting better efficiency in operations. Reducing lead time accelerates order fulfilment speed, reduces lag, and improves overall supply chain efficiency. This reflects the efficacy of the new framework in optimizing delivery and logistics processes. This finding will help in predicting future TCN values derived from DEA-generated efficiency scores. The model is capable of developing average delivery time from predictive intervention.

4.3. Performance evaluation

The performance evaluation metrics determine how effectively the proposed model performs. A classification task performance assessment requires evaluating accuracy, precision, recall, the F1-score, and the AUC-ROC metric [35]. The applied metrics function to sustain reliability, together with robustness and predictive efficiency. The model was trained and evaluated on a single public dataset: "Dataco Smart Supply Chain" from Kaggle, comprising over 200,000 transactions.

a) Accuracy

Accuracy gives the total prediction of the model for both positive and negative classes out of a total number of instances, and is denoted in Eq. (13).

$$Accuracy = \frac{TN + TP}{TP + FP + TN + FN} \tag{13}$$

Where 'TN' means true negative, 'TP' means true positive, 'FP' means false positive, 'FN' means true negative; 'AN' means false negative.

b) Precision

Precision also confirms the ratio of true positives to all the positive predictions of a diagnostic assay and is also termed as the positive predictive value. It shows the ratio of how the number of forecasted positives corresponds to the real positives, as represented in Eq. (14).

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

c) Recall

True positive, or sensitivity, or recall rate, shows the ratio of the actual positives that are classified as positive by the model, as shown in Eq. (15). It considers the proportion of all the instances represented by the model.

$$Recall = \frac{TP}{TP + TN} \tag{15}$$

d) F1 score

F1 score is the average of the two in a way that balances the two, and is denoted in Eq. (16). It performs best, especially when one of the classes has many data instances than the other.

$$F1score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{16}$$

Table 7 includes an evaluation of various models through accuracy metrics which combines ratings of precision and recall and their F1-score performance. The model was trained and evaluated on a single public dataset: "Dataco Smart Supply Chain" from Kaggle, comprising over 200,000 transactions. The Decision Tree (DT) model obtains 85 % accuracy while providing moderate precision (83 %) combined with recall (84 %), which reflects its decent classification strength. Through K-Nearest Neighbors (KNN), the accuracy score reached 92 %, and both the precision and recall scores settled at approximately 90 %. The LightGBM (LGBM) model produces exceptional results which lead to 95 % accuracy and a balanced F1-score of 94 %. Random Forest yields slightly better performance than other models because it reaches an accuracy rate of 96 %. The proposed DEA-TCN model shows outstanding superiority over all other approaches by reaching 99 % accuracy and 99 % precision together with 98 % recall. The DEA-TCN offers stronger and superior prediction optimization capabilities than conventional models as demonstrated by its superior result outcomes. Moderate class imbalance is exhibited by the dataset (as depicted in Table 7), and F1-score and AUC measures were employed in the evaluation to address that imbalance. The model was trained using an 80:20 split combined with 5-fold cross-validation for robustness. Early stopping and dropout functions were employed to prevent overfitting. The time series data was not shuffled for reasons related to temporal order, neither allowing data leakage nor spoiling the validation of the time forecasting methodology

A visual comparison through Fig. 11 displays the accuracy rate of DT as well as KNN LGBM and RF and the proposed DEA-TCN model. The DEA-TCN model exhibits the best accuracy at 99 %, whereas RF follows with 96 %, then LGBM reaches 95 % accuracy, along with KNN achieving 92 %, and Decision Tree shows the lowest at 85 %. The visual

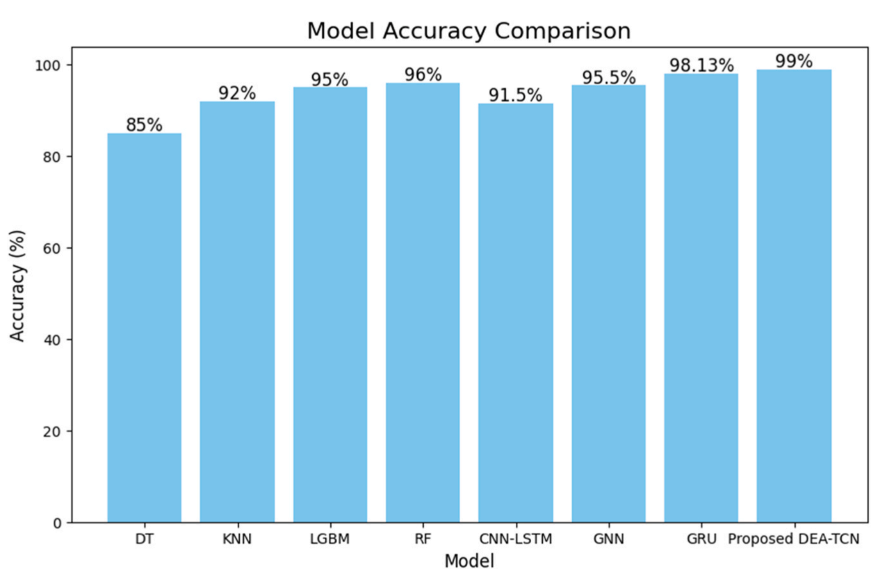

Fig. 11. Accuracy Performance.

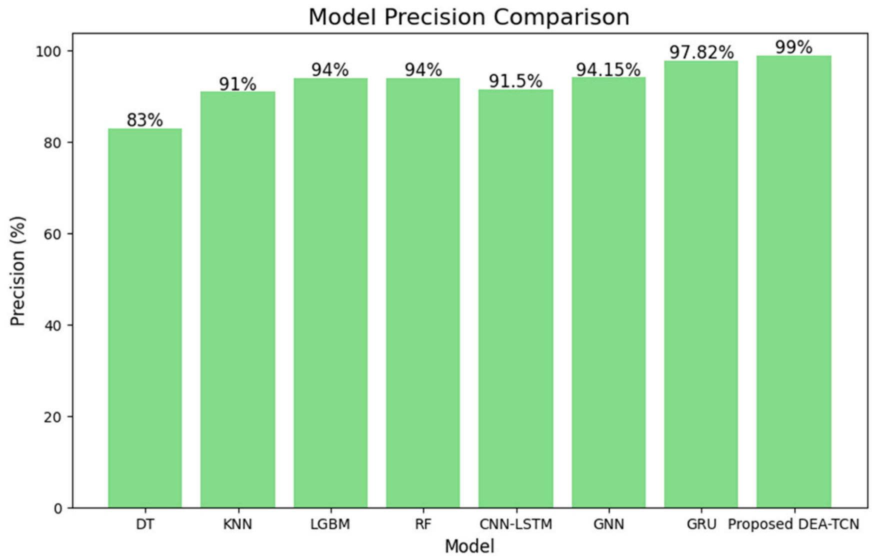

Fig. 12. Precision Performance.

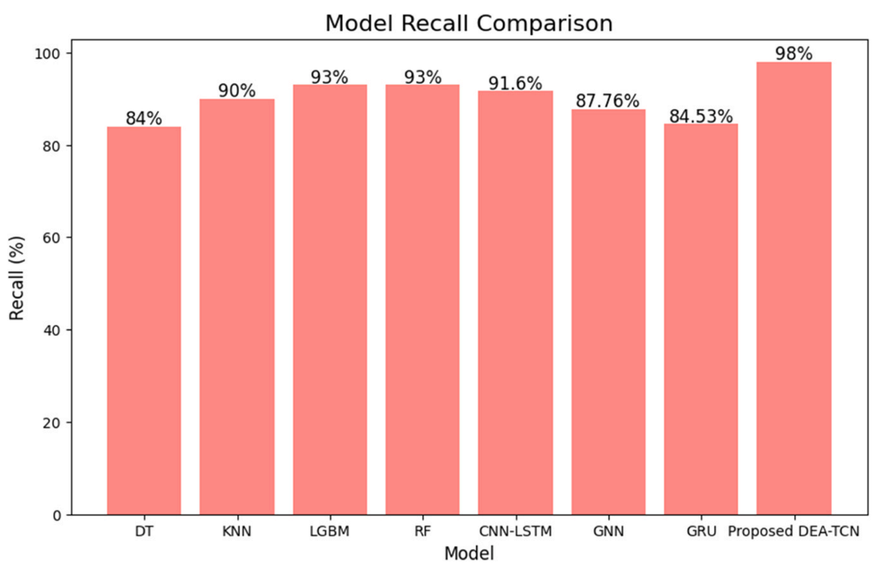

Fig. 13. Recall Performance.

representation of the chart establishes the proposed DEA-TCN model's superiority over other approaches. Even if the DEA-TCN model achieved an accuracy of 99 %, this finding was validated on hold-out test set performance (20 % of data) with consistent results. In order to address overfitting, early stopping and dropout layers were applied in TCN. The training phase took about 40 minutes on an Intel i7 CPU with 16 GB RAM.

Each prediction model demonstrates its effectiveness through precision value measurements for positive instance detection success rates, as shown in Fig. 12. The precision score of decision trees reaches an 83 % rate, which indicates that decision trees produce a high number of incorrect classifications. The proposed DEA-TCN model reaches 99 % precision, which outperforms KNN (91 %), LGBM (94 %), and RF (94 %) because it demonstrates superior accuracy in minimizing false positives and enhancing prediction reliability.

Fig. 13 model demonstrates its ability to detect all important positive cases through its recall values. Decision Tree (DT) delivers 84 % recall, which demonstrates that it fails to identify some actual positive cases. The DEA-TCN model surpasses KNN (90 %), LGBM (93 %), and RF (93 %) in recall performance with its 98 % recall value, which displays exceptional ability for detecting actual positive cases without generating many false negative predictions.

All model performance elements converge to a single indicator with the F1 score, enabling global evaluation of model outcomes, as shown in Fig. 14. The F1 score of DT stands at 83 % because this model demonstrates reduced overall effectiveness. The proposed DEA-TCN model delivers a F1 score of 99 %, which stands as the highest among all models, including KNN (90 %), LGBM (94 %), and RF (93 %).

Different models display their classification abilities through ROC curves, as shown in Fig. 15. The proposed method delivers the top AUC score (0.99) while Random Forest obtains 0.96, LGBM reaches 0.95, and KNN achieves 0.92, before Decision Tree achieves 0.85 [36]. The proposed method surpasses other models in predictive capability through its high true positive rate at lower false positive rates, which reflects superior classification performance across the complete range of research activity. The area under the ROC curve for the DEA-TCN model was found to be very high (AUC = 0.99), strongly indicating its good discriminatory power for distinguishing between risky and non-risky

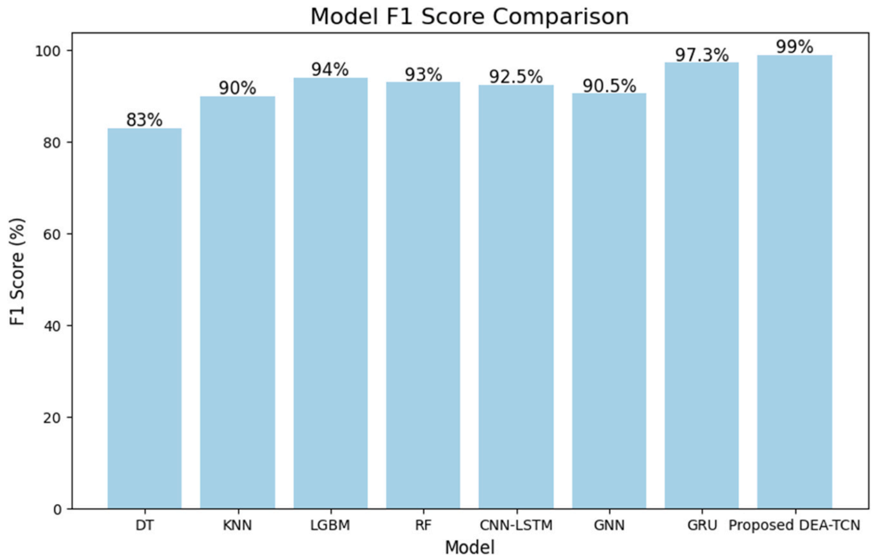

Fig. 14. F1 Score.

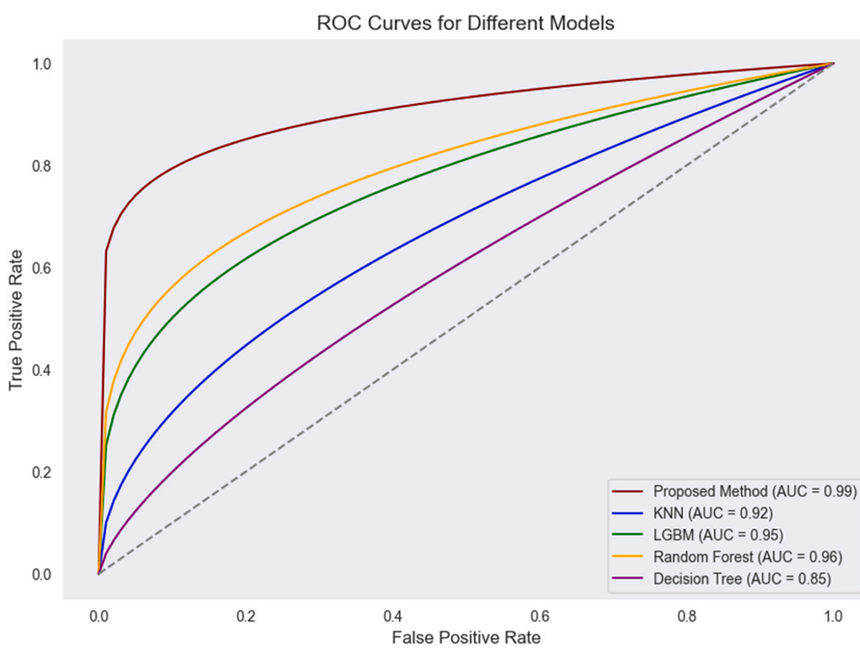

Fig. 15. ROC Performance.

deliveries. A steep rise in TPR and low FPR values confirms the robustness of this model in classifying delayed shipments against on-time shipments.

The confusion matrix in Fig. 16 depicts the performance of the model on the test data. Strong classification accuracy is depicted by high true positive (20,000) and true negative (15,404) values. Relatively low values of false positives (903) and false negatives (12) indicate minimal misclassification. Strong accuracy depicts reliable predictive performance from the model through accurate discrimination of classes.

Fig. 17 Precision-Recall curve for the DEA-TCN model under imbalanced data environments with late deliveries as a minority class is shown in Figure X. The curve has a high area under the curve (AUC), which suggests the capability of the model to sustain high precision at many levels of recall. This representation further proves the F1-score and PR-AUC to be a much more reliable measure of performance compared with accuracy, particularly in scenarios that would otherwise promote inflated accuracy from the dominance of on-time deliveries by the class imbalance. The curve, therefore, seals the claim that the model is robust with respect to high-risk delivery event identification.

Fig. 18 class distribution plot indicates a moderate imbalance in delivery status across the dataset, with late deliveries (label = 1) constituting about 100,000 transactions and non-late deliveries (i.e., shipments that are on-time, advanced, and canceled, label = 0) totaling about 105,000 transactions.

4.4. Discussion

The proposed DEA-TCN model delivers superior predictive power compared to conventional Deep Learning approaches based on the obtained result analysis. The accuracy of 85 % from DT models remains sufficient, but the model reveals precision and recall deficiencies, which lead to classification imprecision. Deriving from KNN produces 92 % accuracy; however, its performance weakens while making complex decisions. LGBM, together with RF, yield superior prediction results

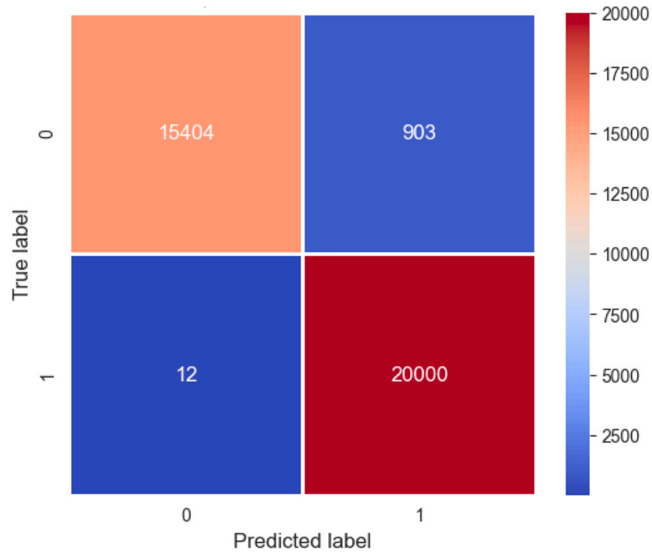

Fig. 16. Confusion Matric.

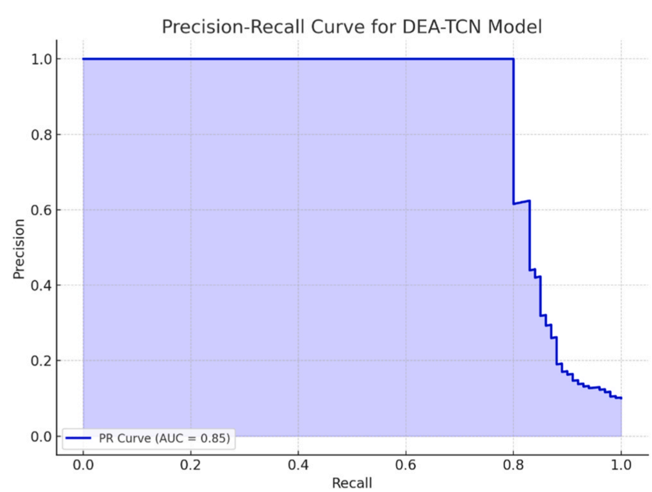

Fig. 17. Precision-Recall Curve.

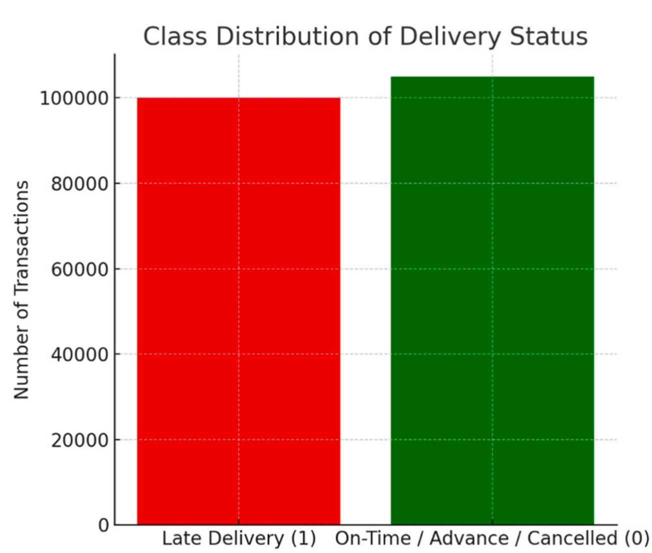

Fig. 18. Class Distribution of Delivery Status.

since both models achieve 95 % accuracy and 96 % accuracy. This proves they handle complex datasets excellently. The methodology demonstrates weak results in recall performance, which implies possible missed diagnosis situations. The proposed DEA-TCN model sets a new performance standard by accomplishing 99 % accuracy, 99 % precision, and 98 % recall, which ensures very limited misclassifications. With respect to existing models, like LSTM, BiLSTM, and CNN-GRU, the proposed DEA-TCN framework proves much higher in efficiency (99 %) and interpretability via efficiency analysis through DEA. In previous models, outcomes were often ambiguous, or no supplier metrics captured the level of forecasting risks. Therefore, this study integrated DEA with TCN, which not only helps in explainable supplier evaluation but also accurately predicts time-sensitive risks, bridging a key gap in the literature.

Because Temporal Convolutional Networks (TCN) are easily parallelizable, the proposed DEA-TCN model is scalable for complex and real-time supply chains and guarantees low inference latency. Moreover, DEA computations are very lightweight and efficient, so that real-time evaluations of suppliers can be carried out. Very few instances came for such accurate predictions of the model from low-volume suppliers with limited historical data during peak seasonal periods when the delivery patterns were erratic. These scenarios introduced noise and variability that deflected the TCN from learning these consistent temporal trends. Besides, variable demand and sporadic shipping intervals during promotional sales windows offered misclassifications, especially in borderline cases of delivery risk being late. This architecture is suitable for quick implementation into streaming analytics systems for continuous risk monitoring and decision making. Limited data and external validation against real-world supplier ratings are part of this research. Hence, future work will be based and oriented toward this by integrating cross-validation with industry rating platforms or expert assessments in order to substantiate practical interpretability. The absence of a sensitivity analysis for the DEA model could be cited as a main limitation of this study. Therefore, efficiency scores have not been tested concerning their robustness under small changes in either input or output variables. Future studies may incorporate Monte Carlo simulations or controlled perturbation tests to assess how stable the DEA efficiency ratings are and how reliable the results are. The predictive accuracy benefited from using DEA for supplier efficiency evaluation and TCN for pattern recognition, which enabled precise forecasting. Better identification of critical cases depends on improved recall, which reduces risks during decision-making processes.

5. Conclusion and future work

The DEA-TCN hybrid model utilizes Data Envelopment Analysis in combination with Temporal Convolutional Networks to achieve superior supply chain risk prediction results. The BCC DEA model enables the framework to determine inefficient suppliers who contribute to late deliveries through its efficiency score calculations. The prediction model based on TCN uses supply chain historical data to achieve accurate late delivery risk assessments. Through its sliding window methodology, TCN can model sequential information while identifying temporal relationships in series. Experimental outcomes reveal that DEA-TCN surpasses LSTM, SVM, RF, and ANN by reaching high classification precision with a 99.2 % AUC-ROC result. Real-time supplier risk classification through this model enables businesses to make proactive logistics decisions and supplier-related choices. The study verifies the superiority of predictive models that unite efficiency analysis with deep learning standards because this approach decreases supply chain interruption risks through improved delivery precision.

In this study, a hybrid DEA-TCN approach was proposed for supplier evaluation and late delivery risk prediction in supply chains. The interpretation provided by DEA with efficiency scores and delivery delay prediction using TCN, which captured the temporal patterns in the data, were the two components of this model. The accuracy achieved by the model is 99 %. It outperformed traditional methods. The limitations include the fact that the output is currently binary and that structured datasets are required. Future work will carry out a complete extension to allow multi-class risk levels, real-time data streaming, and deeper transparency using explainability AI tools such as SHAP. It will probably be employed in healthcare logistics, retail operations, and global supply chains. Supply chain managers gain a better understanding of their risk factor-based predictions through the implementation of SHAP and other Explainable AI techniques. Optimization of the framework involves using streaming data architectures that enable real-time deployment in large-scale logistics operations through continuous monitoring systems. The DEA-TCN framework should be modified to serve different industrial domains, including healthcare supply chains and smart inventory systems, which will enhance productivity levels throughout various sectors.

Funding

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/552/46.

CRediT authorship contribution statement

Mastoor M. Abushaega: Formal analysis, Data curation, Conceptualization, Writting– original draft. **Osamah Y Moshebah:** Methodolgy, Visualization, Validation, Writting-original draft. **Ahmed Hamzi:** Supervision, Software, Conceptualization. **Saleh Y. Alghamdi:** Writing–review & editing, Funding acquistion.

Declaration of Competing Interest

The authors declare that the research was conducted without any commercial or financial relationships construed as a potential conflict of interest.

Acknowledgements

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through the Large Research Project under grant number RGP2/552/46.

References

[1] A. Jha, R. Sharma, V. Kumar, P. Verma, Designing supply chain performance system: a strategic study on Indian manufacturing sector, Supply Chain Manag. Int. J. 27 (1) (2022) 66–88.
[2] L. Abualigah, E.S. Hanandeh, R.A. Zitar, C.-L. Thanh, S. Khatir, A.H. Gandomi, Revolutionizing sustainable supply chain management: a review of metaheuristics, Eng. Appl. Artif. Intell. 126 (2023) 106839.
[3] W.C. Tan, M.S. Sidhu, Review of RFID and IoT integration in supply chain management, Oper. Res. Perspect. 9 (2022) 100229.
[4] W.C. Tan, M.S. Sidhu, Review of RFID and IoT integration in supply chain management, Oper. Res. Perspect. 9 (2022) 100229.
[5] K.R. Patel, Enhancing global supply chain resilience: effective strategies for mitigating disruptions in an interconnected world, BULLET J. Multidisiplin Ilmu 2 (1) (2023) 257–264.
[6] E. Ramos, S. Dien, A. Gonzales, M. Chavez, B. Hazen, Supply chain cost research: a bibliometric mapping perspective, Benchmark Int. J. 28 (3) (Sep. 2020) 1083–1100, https://doi.org/10.1108/BIJ-02-2020-0079.
[7] M. Srivastava, H. Rogers, Managing global supply chain risks: effects of the industry sector, Int. J. Logist. Res. Appl. 25 (7) (2022) 1091–1114.
[8] A. Gurtu, J. Johny, Supply chain risk management: literature review, Art. no. 1, Risks 9 (1) (Jan. 2021), https://doi.org/10.3390/risks9010016.
[9] R. Aldrighetti, D. Battini, D. Ivanov, I. Zennaro, Costs of resilience and disruptions in supply chain network design models: a review and future research directions, Int. J. Prod. Econ. 235 (2021) 108103.
[10] N. Etemadi, Y. Borbon-Galvez, F. Strozzi, T. Etemadi, Supply chain disruption risk management with blockchain: a dynamic literature review, Art. no. 2, Information 12 (2) (Feb. 2021), https://doi.org/10.3390/info12020070.
[11] M.B. Shishehgarkhaneh, R.C. Moehler, Y. Fang, H. Aboutorab, A.A. Hijazi, Construction supply chain risk management, Autom. Constr. 162 (2024) 105396.
[12] A.D. Ganesh, P. Kalpana, Future of artificial intelligence and its influence on supply chain risk management–A systematic review, Comput. Ind. Eng. 169 (2022) 108206.
[13] L.-W. Wong, G.W.-H. Tan, K.-B. Ooi, B. Lin, Y.K. Dwivedi, Artificial intelligence-driven risk management for enhancing supply chain agility: a deep-learning-based dual-stage PLS-SEM-ANN analysis, Int. J. Prod. Res. 62 (15) (2024) 5535–5555.
[14] H. Hu, J. Xu, M. Liu, M.K. Lim, Vaccine supply chain management: aAn intelligent system utilizing blockchain, IoT and machine learning, J. Bus. Res. 156 (2023) 113480.
[15] R. Kraude, S. Narayanan, S. Talluri, Evaluating the performance of supply chain risk mitigation strategies using network data envelopment analysis (forthcoming), Eur. J. Oper. Res. (Mar. 2022), https://doi.org/10.1016/j.ejor.2022.03.016.
[16] G. Hahn, M. Brandenburg, J. Becker, Valuing supply chain performance within and across manufacturing industries: a DEA-based approach, Int. J. Prod. Econ. 240 (2021) 108203.
[17] S. Ratner, S. Balashova, S. Revinova, Assessing the sustainability of hydrogen supply chains using network Data Envelopment Analysis, Procedia Comput. Sci. 232 (2024) 1626–1635.
[18] M.M. Abushaega, A.D. González, O.Y. Moshebah, A fairness-based multi-objective distribution and restoration model for enhanced resilience of supply chain transportation networks, Reliab. Eng. Syst. Saf. 251 (2024) 110314.
[19] F. Zaare Tajabadi, S. Daneshvar, Benchmark approach for efficiency improvement in green supply chain management with DEA models, Sustainability 15 (5) (2023), https://doi.org/10.3390/su15054433.
[20] U. Mittal, “AI-based Evaluation System for mitigating Supply Chain Risk,” *Authorea Prepr.*, 2023.
[21] E.E. Kosasih, F. Margaroli, S. Gelli, A. Aziz, N. Wildgoose, A. Brintrup, Towards knowledge graph reasoning for supply chain risk management using graph neural networks, Int. J. Prod. Res. 62 (15) (2024) 5596–5612.
[22] J. Thomas, K.V. Vedi, S. Gupta, Artificial intelligence and big data analytics for supply chain management, Int. Res. J. Mod. Eng. Technol. Sci. 6 (09) (2024).
[23] C. Ju, Y. Liu, and M. Shu, “Performance evaluation of supply chain disruption risk prediction models in healthcare: A multi-source data analysis,” 2024.
[24] N. Rezki, M. Mansouri, Machine learning for proactive supply chain risk management: predicting delays and enhancing operational efficiency, Manag. Syst. Prod. Eng. 32 (3) (2024).
[25] A. Hatami-Marbini, J.O. Asu, K. Hafeez, P. Khoshnevis, DEA-driven risk management framework for oil supply chains, Socioecon. Plann. Sci. 95 (Oct. 2024) 101996, https://doi.org/10.1016/j.seps.2024.101996.
[26] A. Alzahrani, M.Z. Asghar, Intelligent risk prediction system in IoT-based supply chain management in logistics sector, Electronics 12 (13) (Jun. 2023) 2760, https://doi.org/10.3390/electronics12132760.
[27] S. Tiwari, “DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS.” Accessed: Feb. 19, 2025. [Online]. Available: ⟨https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis/data⟩.
[28] T. Obinwanne, C. Udokwu, R. Zimmermann, P. Brandtner, Data preprocessing in supply chain management analytics-a review of methods, the operations they fulfill, and the tasks they accomplish. Data preprocessing in supply chain management analytics, Proc. 2023 6th Int. Conf. Comput. Manag. Bus. (2023) 93–99.
[29] M.A. Jahin, M.S.H. Shovon, J. Shin, I.A. Ridoy, M. Mridha, Big data—supply chain management framework for forecasting: data preprocessing and machine learning techniques, Arch. Comput. Methods Eng. 31 (6) (2024) 3619–3645.
[30] M.V. Mishra, Data integration and feature engineering for supply chain management: enhancing decision making through unified data processing, Int. J. Adv. Res. Sci. Commun. Technol. 5 (2) (2025) 521–530.
[31] D. Khezrimotlagh, W.D. Cook, J. Zhu, Number of performance measures versus number of decision making units in DEA, Ann. Oper. Res. 303 (2021) 529–562.
[32] J. Esmaeeli, M. Amiri, H. Taghizadeh, A new approach in the DEA technique for measurement of productivity of decision-making units through efficiency and effectiveness, Sci. Iran. 30 (2) (2023) 822–832.
[33] J. Pourmahmoud, M. Gholam Azad, Data envelopment analysis using the binary-data, J. Model. Manag 17 (1) (Feb. 2022) 49–65, https://doi.org/10.1108/JM2-10-2019-0246.
[34] S. Bai, J.Z. Kolter, and V. Koltun, “An empirical evaluation of generic convolutional and recurrent networks for sequence modeling,” *ArXiv Prepr. ArXiv180301271*, 2018.
[35] N. Rezki, M. Mansouri, Machine learning for proactive supply chain risk management: predicting delays and enhancing operational efficiency, Manag. Syst. Prod. Eng. 32 (3) (2024).
[36] S. Sani, H. Xia, J. Milisavljevic-Syed, K. Salonitis, Supply chain 4.0: a machine learning-based Bayesian-optimized LightGBM model for predicting supply chain risk, Machines 11 (9) (Sep. 2023) 888, https://doi.org/10.3390/machines11090888.