Baltic J. Modern Computing, Vol. 13 (2025), No. 2, 453-485
https://doi.org/10.22364/bjmc.2025.13.2.09

Supply Chain Analytics: A Performance Evaluation of Machine Learning, Statistical, and Time-Series Models

Vimal DWIVEDI[1], Karuna KADIAN[2], Bryan GARDINER[1], Rachana PANDEY[2], Anjali LATHWAL[2]

[1] Intelligent Systems Research Centre, Ulster University, Northern Ireland, UK
[2] Indira Gandhi Delhi Technical University for Women, India

v.dwivedi@ulster.ac.uk, karunakadian@igdtuw.ac.in, b.gardiner@ulster.ac.uk, rachana033mtcse21@igdtuw.ac.in, anjalilathwal@igdtuw.ac.in

ORCID 0000-0001-9177-8341, ORCID 0009-0004-8929-5999, ORCID 0000-0001-5642-6850, ORCID 0000-0002-9404-2548, ORCID 0000-0001-9810-8893

Abstract. Supply chain analytics is pivotal in enhancing supply chain performance through data-driven decision-making. This study evaluates the effectiveness of various analytical models in improving supply chain performance using the DataCo Smart Supply Chain dataset. This research comprehensively evaluates machine learning, statistical, and time-series models for forecasting accuracy, demand prediction, late delivery risk, shipment duration, and route mapping optimisation. The study methodically compares and contrasts the outcomes of various modelling techniques, providing valuable insights into the most suitable approaches for optimising supply chain operations. The results indicate that decision tree and random forest models excel in supply chain forecasting and sales prediction. Similarly, Random Forest, XGBoost, and Gradient Boost models accurately predict late delivery risk, while Exponential Smoothing, SARIMA, and ARIMA models effectively predict shipment duration. To validate these findings, rigorous statistical testing, cross-validation, and alignment with industry standards were employed, ensuring the reliability and applicability of the results. This research contributes significantly to supply chain analytics, offering practitioners and researchers guidance on selecting appropriate methodologies for enhanced supply chain performance.

Keywords: Supply chain management, Machine Learning, Simulations

1 Introduction

The supply chain is a complex network consisting of various organisations and enterprises involved in the production and delivery of goods and services to consumers, as illustrated in Figure 1, adapted from (Anitha and Patil., 2018). Efficient management of the supply chain is essential for the success of any company, as it requires the identification of inefficiencies and the implementation of strategies to optimise the network (Samir, 2023). However, this task is complicated by the participation of multiple stakeholders and the need to manage a wide range of resources.

In recent years, advancements in big data and analytics have provided organisations with new tools to enhance their supply chain management practices (Raman et al., 2018), Supply chain analytics, which leverages data analytics and quantitative methodologies, play a crucial role in optimising supply chains by applying sophisticated algorithms and statistical models to large datasets (Surie and Reuter, 2014). This approach enables businesses to improve efficiency, enhance customer satisfaction, and reduce operational costs.

This study seeks to systematically evaluate the performance and accuracy of ma- chine learning, time series, and statistical models across four specific use cases in supply chain analytics utilising the "DataCo supply chain" dataset (Constante et al., 2021). These use cases include supply chain forecasting, late delivery risk, shipment duration prediction, and route mapping. The selection of these particular four use cases was guided by the practical strengths and constraints of the "DataCo Supply Chain" dataset we considered in our study. Dataset, though, is quite rich in information but has significant gaps in customer-level or product-level fields, namely "Customer Demographics", "Product Description", etc. Such limitations restrict the exploration to other areas, like customer behavior analysis or location-based trends. Our use cases mainly deal with the operational and feasibility significance based on the available dataset. Through a comprehensive comparison of these models, this research aims to identify the most effective models for each scenario, thereby supporting informed decision-making in supply chain management. To ensure a thorough evaluation, the assessment criteria are aligned with the ISO 25010 standard, incorporating essential quality attributes such as functionality, usability, efficiency, reliability, maintainability and portability as defined by the standard.

This paper presents an experimental research study complemented by survey elements. The experimental aspect focuses on evaluating and comparing various models within supply chain analytics, while the survey offers a comprehensive review of the current state of knowledge, setting the context for the proposed investigations. The integration of a survey within an experimental framework provides a holistic understanding of the field and introduces new empirical insights.

To achieve the best prediction models, data pre-processing, exploratory data analysis (EDA), feature selection, and algorithm selection was performed. The models were evaluated based on r2 score, root mean square error (RMSE), and mean square error (MSE). The findings from this research can assist businesses in selecting appropriate models for supply chain analytics.

The remainder of the paper is structured as follows: The rest of this Section introduces the study, outlines its objectives, and the significance of supply chain analytics. Section 2 reviews previous research in supply chain analytics and identifies gaps in the existing literature. Section 3 details the research methodology, including data pre-processing, EDA, feature selection, algorithm selection, prediction models, and evaluation metrics. Section 4 presents the findings for each use case, supported by tables

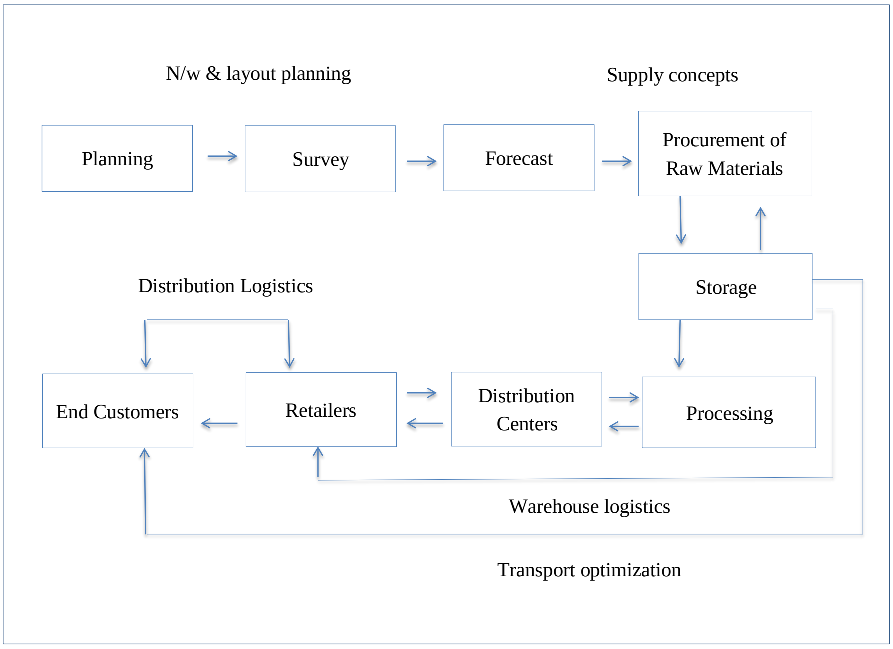

Fig. 1. Process of Supply Chain Management

and graphs. Section 6 summarise the findings, discuss critical implications for supply chain management, highlight limitations, and propose future research directions.

1.1 Background and motivation for the study:

Advanced analytics in supply chain management has become increasingly important due to its ability to enhance both efficacy and efficiency (Zekhnini et al., 2020). Supply chain analytics leverages statistical and computational tools to analyze data, thereby improving decision making in areas such as inventory management, transportation, and procurement.

Machine learning, statistical, and time-series models are commonly employed in supply chain management to predict key performance indicators (KPIs) such as sales forecasting, delivery timeframes, and transportation costs (Hahn, 2019). However, there is a lack of comprehensive research evaluating the performance of these models across different supply chain use cases.

The study focuses on determining the accuracy and effectiveness of various models within the four specified supply chain analytics use cases. The goal is to identify models that not only perform well but also meet industry standards for accuracy and efficacy.

1.2 Research questions and objectives:

– This study aims to address the following research questions:

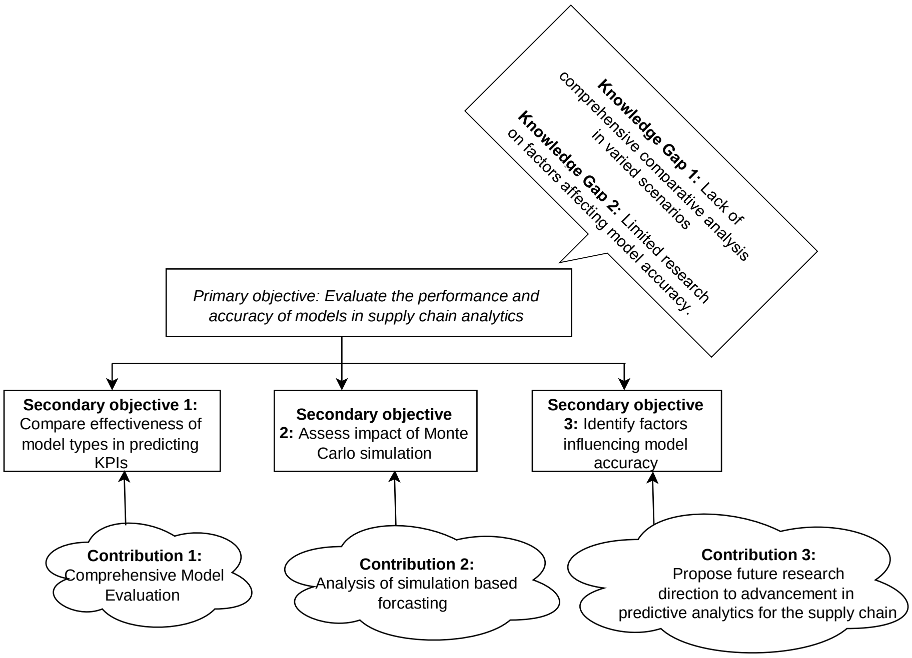

Fig. 2. Goal tree of research objectives and their corresponding contributions

1. Which type of model (machine learning, time series or statistical) performs best in each of the four use cases of supply chain analytics?
2. How does the performance of these models change when incorporating Monte Carlo simulation?
3. What are the key factors that most significantly impact the accuracy of these models in each use case?

– The primary objectives and sub-objectives of this research are listed below and visually depicted in Figure 2:

1. To assess the performance and accuracy of machine learning, statistical and timeseries models in predicting key performance indicators (KPIs) across various supply chain management scenarios.
2. To compare these models in four critical supply chain analytics use cases, ensuring alignment with the ISO25010 quality standards.
3. To analyse the impact of Monte Carlo simulation on model performance.
4. To identify critical factors influencing model accuracy in each use case, thereby

providing a comprehensive evaluation of model efficacy in supply chain analytics.

1.3 Research Methodology:

The research methodology employed in this study involves a systematic process that includes a literature survey, data collection, data pre-processing, exploratory data analysis (EDA), feature selection, algorithm selection, prediction model development and model evaluation using appropriate metrics (see Figure 3).

1.3.1 Literature Survey: A comprehensive literature survey was conducted, covering leading journals and conference proceedings in the field of supply chain management, analytics, artificial intelligence (AI), and machine learning (ML). This survey informed the research objectives, guided the framework for model selection and evaluation, and highlighted the role of Monte Carlo simulation in enhancing forecast accuracy. The literature survey was carried out to understand the current state of the art of the research in supply chain analytics and to locate gaps or areas that are not well explored. It also helped us to select the most relevant use cases and the respective methods, and ensured that our study not only offered something new but also built on the foundation of what already exists.

1.3.2 Data Collection: The primary dataset used in this study is the "DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS" which provides comprehensive information relevant to supply chain operations. The dataset was acquired and prepared for analysis, with careful attention to maintaining data quality and integrity.

1.3.3 Model Development: The model development process involved data preprocessing, exploratory data analysis (EDA), and feature selection to prepare the dataset for modeling.

1.3.4 Algorithm Selection: Drawing on insights from the literature review and the specific requirements of each use case, a comprehensive set of models and algorithms were selected for evaluation.

1.3.5 Prediction Model Development: Prediction models were developed for each use case using the selected models and algorithms. These models were trained on the pre-processed dataset, with techniques and parameters tailored to each model type to ensure accuracy and reliability.

1.3.6 Model Evaluation: The performance of prediction models was evaluated using standard assessment measures such as the accuracy, R2 score, RMSE, and MSE. These metrics were chosen based on the type of task and model. For regression problems such as forecasting and shipment duration prediction, we used metrics such as RMSE and MAE to measure the performance of the model. Late delivery risk and other classification-type tasks' performance are measured by the accuracy metric across balanced and imbalanced data. All these metrics are standard, interpretable, and are the scientifically validated way to assess the performance of classification and regression models. The evaluation process involved comparing the performance of various models within each use case to identify those with the highest accuracy and reliability.

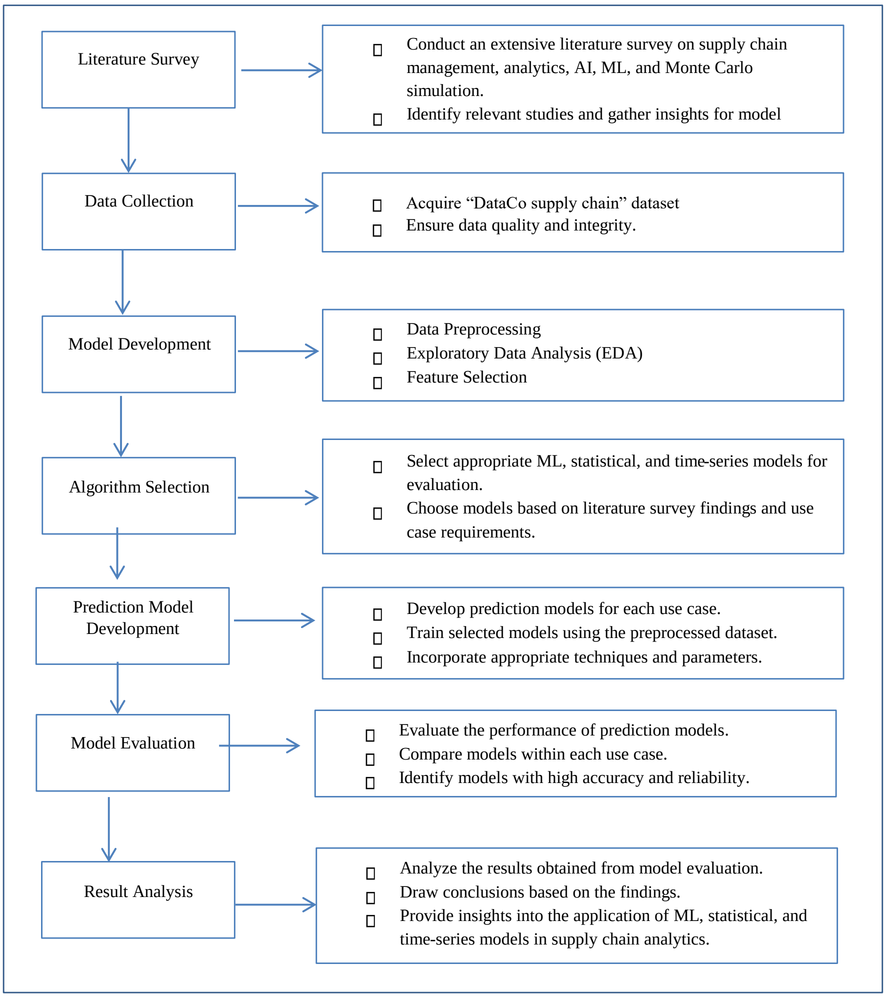

Fig. 3. Steps used in research methodology

The way this study is done makes sure we look carefully at how well time-series, statistical, and machine-learning models work for supply chain analysis. It also discusses the research goals and gives valuable advice using these models.

1.4 Brief overview of the data and the four use cases selected:

In this study, we utilised a dataset of "DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS" (Constante et al., 2021), which encompasses various supply chain components such as sales, transportation, and inventory. Developed by DataCo Global, this dataset supports an intelligent supply chain solution designed for extensive data analysis, particularly in sectors such as clothing, sports equipment, and electronics. The dataset, named DataCoSupplyChainDataset.csv, enables users to analyse and process structured data related to Provisioning, Production, Sales, and Commercial Distribution activities using machine learning methods and R software. Additionally, the system can correlate structured and unstructured data from Clickstream logs, tokenized in a file called tokenized access logs.csv, enabling the identification of patterns and trends within the supply chain.

The DataCoSupplyChain.csv file provides a comprehensive explanation of each variable in the DataCoSupplyChainDataset.csv. This system is classified under Data Mining, Management of Supply Chain, Big Data Analytics and ML. The latest version, Version 5, was published on March 13, 2019. By leveraging DataCo's intelligent supply chain technology, businesses can gain a competitive advantage by optimising their supply chain processes and making data-driven decisions, ultimately fostering business grow.

Four use cases in supply chain analytics were selected from this dataset:

1. Supply Chain Forecasting Sales Prediction: In this use case, we predict sales for a specific period using historical sales data.
2. Late Delivery Risk: Here, we estimate the probability of a shipment being delayed, considering factors such as weather conditions, transportation mode and delivery distance.
3. Predicting Shipment Duration: This use case involves predicting the time required to deliver a shipment based on variables such as transportation mode, distance and weather conditions. Relevent studies on predicting shipment duration can be found in (Mariappan et al., 2022) and (Sahoo et al., 2021).
4. Route Mapping: Route mapping: This use case focuses on identifying the optimal shipment route based on transportation mode, delivery distance, and road conditions.

For each of these use cases, statistical, machine learning and time-series models were applied to forecast the relevant Key Performance Indicators (KPIs) (Estampe et al., 2013). Additionally, Monte Carlo simulations can be performed to evaluate the accuracy of these models under varying conditions. We also analyse the key factors influencing the accuracy of the models for each use case.

1.5 Selection of use cases for evaluation:

This study examines the effective application of machine learning, time-series and statistics models to enhance the understanding of supply chain dynamics. Four key areas were selected for evaluation: sales prediction, late delivery risk assessment, shipment duration forecasting, and route optimisation. These use cases were chosen due to their significant impact on such situations on supply chain management.

1.5.1 Supply Chain Forecasting - Sales Prediction: Accurate sales forecasting is critical in supply chain management as it directly influences business planning, production, inventory management, and logistics. Effective forecasting ensures the customer demands are met while optimising costs and resource allocation. Therefore, it is essential to evaluate the effectiveness of different models in predicting sales within the supply chain.

1.5.2 Late Delivery Risk: The risk of late deliveries is a major concern in supply chain management, as it can lead to customer dissatisfaction, increased costs, and scheduling disruptions By assessing the accuracy of models in predicting delivery delays, businesses can proactively address issues, ensuring timely deliveries and enhancing customer satisfaction.

1.5.3 Predicting Shipment Duration: Accurately predicting shipment duration is essential for effective supply chain management, scheduling, and resources optimisation. Evaluating the precision of models in forecasting delivery times allows businesses to improve customer service, optimise inventory management and streamline transportation processes.

1.5.4 Route Mapping: Optimising delivery routes is essential for essential for efficient logistics, cost reduction, and overall supply chain performance. By assessing the effectiveness of models in identifying optimal routes considering various factors, the study aims to determine their potential in improving transportation efficiency and enhancing the overall supply chain.

By focussing on these four scenarios, the study aims to provide a comprehensive assessment of model performance acrossvarious aspects of supply chain analytics. Each scenario addresses a distinct supply chain management challenge, highlighting the flexibility of the models. The outcomes of these evaluations will identify the strengths and potential limitations of each model, guiding the selection of the the most appropriate tools for specific supply chain tasks.

While this study focused on evaluating the effectiveness of machine learning, statistical, and time-series models in four specific deliver chain analytics use cases, there remains several other potential use cases that warrant further exploration. Future research should investigate those additional areas to provide a more comprehensive evaluation supply chain analytics. Potential use cases include demand planning, inventory management, production, transportation planning, risk management, and compliance.

Each of these domains present unique challenges and opportunities for utilising analytical models to enhance supply chain efficiency and effectiveness. By broadening the scope of analysis to include these additional use cases will enable academics and practitioners to gain a deeper understanding of the capabilities and limitations of various modelling approaches in addressing the complexities of supply chain management. This broader investigation will contribute to advancement and refinement of supply chain analytics practices, ultimately leading to improved decision-making and performance in supply chain operations.

1.6 Motivation for Model Selection:

Several key factors influenced the selection of models and algorithms in this research paper. Firstly, the study aimed to evaluate the performance of machine learning (ML), statistical and time-series models, given their proven effectiveness across various domains. By comparing these diverse modeling techniques within the context of supply chain management, the research sought to gain insights into their effectiveness and applicability.

Secondly, a diverse set of models was incldued to explore the strengths and weaknesses of each approach. Machine learning models such as Linear Regression, Lasso Regressor, Ridge Regressor, KNN Classifier, Gaussian Naive Bayes, SVM Classifier, Gradient Boost, Decision Tree, Random Forest, LSTM Model, and Gated Recurrent Unit are well-known for their ability to capture complex patterns and relationships in data. Statistical models, including the Chi-square test, OLS, T-test, Multiple Linear Regression, and Kruskal-Wallis test, provide rigorous methods for analysing correlations and drawing statistically significant inferences. Time-series models such as the Exponential Smoothing Model, SARIMA, and ARIMA are specifically tailored to accommodate temporal data and capture time-dependent trends.

Thirdly, the literature review revealed that machine learning models are commonly employed in supply chain analytics research. However, this study sought to explore the performance of alternative modelling approaches, such as statistical and time-series models, to provide a more comprehensive evaluation of their suitability for supply chain analytics.

Lastly, the selected use cases of supply chain forecasting, late delivery risk, predicting shipment duration, and route mapping were carefully chosen to cover a broad spectrum of supply chain management challenges. By evaluating multiple models across these diverse use cases, the study aimed to understand how different modelling strategies perform under distinct supply chain scenarios.

The objective was to conduct a comprehensive evaluation of machine learning, statistical, and time-series approaches in supply chain analytics, compare their performance, and assess their suitability for addressing real-world supply chain challenges. By incorporating a wide range of models, the research aimed to provide valuable insights into their strengths and limitations, assisting researchers and practitioners in making informed decisions regarding model selection in supply chain management.

2 Literature review

2.1 Literature Review Methodology

2.1.1 Study Identification To ensure a comprehensive understanding of existing research in supply chain analytics, a structured literature review approach was undertaken. The review focused on identifying studies related to the application of machine learning, statistical, and time-series models in supply chain forecasting, risk prediction, shipment duration estimation, and route optimization.

Search Keywords and Queries: The following search strings were used, combining key concepts with Boolean operators:

(”supply chain analytics” OR ”supply chain forecasting” OR ”supply chain prediction”) AND (”machine learning” OR ”time series” OR ”statistical models” OR

”predictive analytics”) AND (”model evaluation” OR ”model comparison” OR ”performance assessment”)

Search Sources: The literature search was conducted across Scopus, Web of Science, IEEE Xplore, and Google Scholar. Table 1 mentions the number of articles across the respective sources considered as part of the review. These databases were selected for their extensive coverage of peer-reviewed journals and conference proceedings relevant to supply chain management and predictive analytics.

Table 1. Study sources and identified paper count.

Search Engine	Databases / Coverage	Article Count
Scopus*	Peer-reviewed journals, conference proceedings	104
Web of Science*	Peer-reviewed journals, conference proceedings	90
IEEE Xplore*	Conference papers, technical magazines	65
Google Scholar*+	Grey literature, technical reports, working papers	482
	Total	741
	Total (after duplicate removal)	452

*Includes journals, proceedings, and book chapters
+Includes technical reports, white papers, and working papers

2.2 Study Selection

To further refine the articles identified in the initial search, a two-phase screening process was applied: the inclusion phase and the exclusion phase. This process ensured that only relevant and high-quality articles aligned with the study’s objectives were retained.

A two-stage screening process was applied:

Table 2. Inclusion Criteria.

Criteria ID	Inclusion Criteria
IC1	The article applies machine learning, statistical, or time-series models for supply chain forecasting, risk prediction, shipment duration estimation, or route optimisation.
IC2	The study includes empirical validation of predictive models using real-world or benchmark datasets.
IC3	The article provides performance evaluation metrics (e.g., accuracy, RMSE, F1-score) for model comparison.
IC4	The article is published in peer-reviewed journals, conference proceedings, or technical reports.
IC5	The publication is within the date range of January 2013 to December 2024.

Table 3. Exclusion Criteria.

Criteria ID	Exclusion Criteria
EC1	The article focuses solely on supply chain strategy, policy, or qualitative frameworks without predictive modelling.
EC2	The article lacks empirical results or does not report performance evaluation of predictive models.
EC3	The study focuses on unrelated domains (e.g., manufacturing processes, IoT hardware) without supply chain forecasting context.
EC4	Non-English language publications.
EC5	Duplicate publications or extended versions of already included papers.

– *Title and Abstract Screening:* Articles unrelated to predictive modelling or lacking empirical validation were excluded, reducing the pool to 97 articles.
– *Full-Text Review*: Studies were further assessed for relevance based on methodological alignment (use of ML/statistical/time-series models), application to supply chain forecasting, risk, duration, or routing, and availability of performance evaluation results.

Final Sample Size: 32 articles as specified in the Table 4 were selected as the final operational sample, which formed the basis of this study's literature review and comparative analysis.

Following this systematic search and selection process with specific inclusion and exclusion criteria outlined in Table 2 and 3, the literature was thematically analyzed to provide a comprehensive understanding of current methodologies, identified gaps, and opportunities for further research. The subsequent sections summarize key findings from the reviewed studies, beginning with general approaches to supply chain forecasting, followed by examining specific modeling techniques and simulation methods.

Table 4. Number of articles selected from each selection phase.

Search Stages	Identified Articles
Study Identification Phase	741 452 *(after duplicate removal)*
Study Selection Phase (Title & Abstract Screening)	97 Primary Studies
Full-Text Review Phase	32 Selected Studies (Operational Sample)

2.3 Previous research on supply chain management and forecasting

Supply chain management involves optimising and coordination of all processes in the production and delivery of goods and services. Forecasting plays a crucial role in predicting demand, determining production needs, managing inventory levels, and setting delivery schedules. Although various methods have been proposed to improve prediction accuracy in supply chain management, a comprehensive comparison across a broad spectrum of models (machine learning, statistical, and time-series) across different use cases remains underexplored.

Time-series models, such as ARIMA, have been extensively employed due to their ability to identify trends and seasonal patterns. For instance, (Fattah et al., 2018) utilised ARIMA to forecast demand in the food industry. However, such studies, while effective, often lack a comparative analysis across different model types in a unified framework, which is the focus of this study

Machine learning algorithms are continuing to gain traction in supply chain forecasting. Studies such as (Guanghui et al., 2012) and (Kilimci et al., 2019) investigate the use of algorithms such as Support Vector Regression and deep learning methods. Despite these advancements, there remain a significant gap in the literature regarding a comprehensive evaluation of these models against traditional statistical and time-series models across multiple use cases - a gap that this study aims to address.

Recent research, including the work of (Terrada et al., 2022), has introduced hybrid forecasting models that combine different methodologies. However, these studies often do not provide a detailed comparative analysis of the effectiveness of machine learning, statistical, and time-series models in various supply chain scenarios. Addressing this gap is crucial to answering the research questions posed in this study.

The impact of Monte Carlo simulation on model performance is another area that has been insufficiently explored in existing literature, including research by (Terrada et al., 2022) and others. This research seeks to fill this gap by investigating how Monte Carlo simulation influences the performance of various model types in supply chain analytics.

While previous studies has often focused on evaluating model performance under specific conditions, it has not thoroughly explored the key factors influencing model accuracy across different supply chain management use cases. This study aims to address this significant gap in the existing body of research by identifying and analysing these critical factors.

2.4 Summary of the study's machine learning models

According to (Terrada et al., 2022), machine learning has become a vital tool for evaluating and modelling complex data in various domains. In the literature, supervised learning models such as SVM, linear regression, random forests, decision trees, and ANN have been employed yo address issues including demand forecasting, transportation planning, and inventory control. (Dash et al., 2019) highlight the importance of a well-functioning supply chain and the benefits of using Artificial Intelligence (AI) to optimise inventory forecasting and customer demand prediction. The authors argue that AI can enhance asset utilisation, increase revenue, and reduce costs by providing highly accurate predictions, optimising R & D and manufacturing processes, improving promotional strategies, and enhancing the customer experience. For instance, (Wan, 2021) proposes a hybrid model that combines the XGBoost algorithm with Random Forest to detect product fraud. This model can assist businesses better align their strategies with market demands and boost revenue. The hybrid model outperforms traditional machine learning techniques in terms of the F1 score, achieving increases of 0.49, 0.49, and 27.9 points over Gaussian Naive Bayes, SVM, and Logistic Regression. The study tested the proposed model using DataCo's innovative supply chain dataset.

In contrast, unsupervised learning models, such as clustering and association rule mining, are utilised to analyse data patterns and identify hidden structures within datasets. Additionally, statistical models, including the Chi-square test, Ordinary Least Squares (OLS), T-test, Multiple Linear Regression and Kruskal-Wallis test, have been used in literature to model and analyse supply chain data (Morgenthaler, 2009).

Time series models such as SARIMA, ARIMA, and exponential smoothing are commonly employed to analyse and predict time-series data in the supply chain. ARIMA, a popular time-series model, has been widely used to forecast various supply-chain factors, including demand, inventory levels, and delivery schedules. SARIMA extends ARIMA's capabilities to account for seasonality in the data. Exponential smoothing, another time-series technique, has been applied to demand forecast in the supply chain. For instance, (Nguyen et al., 2021) propose two data-driven strategies to enhance decision-making in supply chain management. The first strategy utilises an LSTM network-based approach to forecast multivariate time-series data by integrating internal and external business sources of data to improve performance. The second strategy involves detecting sales anomalies using a combination of a one-class SVM algorithm with an LSTM Autoencoder network-based technique. These methods were evaluated on real-world data from the fashion retail industry and benchmarking datasets, demonstrating superior performance compared to previous research. Additionally, (Shih et al., 2019) discusses two data-driven approaches for estimating the supply of blood components at blood centres to decrease blood wastage and shortage. The study compared time-series and machine learning techniques using five years of historical blood supply data from the Taiwan Blood Services Foundation (TBFS). The findings suggest that decision support systems for executives and pathologists at blood centres, hospitals and blood donation facilities can benefit from time-series forecasting techniques, particularly seasonal ESM and ARIMA models, which were found to outperform machine learning algorithms in terms of accuracy.

In addition to the aforementioned predictive models, simulation-based techniques have also been employed to enhance model robustness and manage uncertainties in supply chain analytics. The following subsection focuses specifically on the role of Monte Carlo simulation in this context.

2.5 Previous research on Monte Carlo simulation and its impact on prediction accuracy:

Monte Carlo simulation is a widely used technique in statistics and modelling, which leverages random number generation to simulate and analyse the behaviour of complex systems or processes. in supply chain analytics, Monte Carlo simulation is frequently employed to assess uncertainties related to demand forecasts, inventory levels, and overall supply chain performance. For example, (Schmitt et al., 2009) discusses a project aimed at evaluating the risk of supply chain disruptions and developing mitigation strategies. The authors constructed a simulation model utilising Arena and Monte Carlo simulations to evaluate the disruption risk at various supply chain nodes and the impact of these disruptions on customer service. This approach enabled a straightforward risk assessment and outcome estimation for adverse events. The initiative made several insightful points, such as no specific measures for mitigation at the strategic level. The project also highlighted the absence of specific strategy-level mitigation measures and the importance of historical distribution data in the database. The model's results indicate the need for operational changes within the network to maintain system performance, even under stable conditions. By employing Monte Carlo simulation, researchers can evaluate the effects of different scenarios on supply chain performance.

Researchers have discovered that integrating Monte Carlo simulation as part of a supply chain forecasting system, in combination with advanced modelling techniques such as time-series analysis and machine learning, can significantly enhance prediction accuracy. For instance, (Mohamed et al., 2020) explores the challenge of calculating the gradient of an expectation of a function which is particularly relevant in applied statistics, machine learning, and computational finance. The study presents three main estimators for computing gradients: score-function, measure-valued gradient and pathwise. It also provides guidance on selecting between these estimators based on problem characteristics, such as the number of parameters and whether gradients are decentralised, exhibit high variance, or do not exist (e.g., in cases lacking foundational knowledge in differential calculus). The study further addresses variance reduction techniques and recommends testing the unbiasedness of the gradient estimator. The report concludes with case examples and research suggestions.

Additionally, (Foerster et al., 2018) introduces DICE as a general method for constructing order gradient estimators suitable for stochastic computation graphs. DICE aims to address limitations in current methods for estimating higher-order gradient estimates. The study proposes a practical implementation of DICE for deep learning frameworks, testing its validity and utility in a multi-agent reinforcement learning scenario. The findings demonstrate the accuracy of DICE estimators, and the authors suggest that DICE could facilitate further exploration and adoption of higher-order learning approaches in metalearning, reinforcement learning, and other stochastic computation graph applications. Similarly, (Bousqaoui et al, 2017) compares the performance of several machine learning algorithms in supply chain management using Monte Carlo simulation, while (Thete, 2022) present a stochastic model for demand forecasting in Python, employing the Time-Series SARIMA model alongside Monte Carlo simulation. Overall, previous research has demonstrated that Monte Carlo simulation can enhance the accuracy and resilience and robustness of supply chain predictions, particularly when combined with advanced modelling techniques like machine learning and time-series analysis.

3 Methodology

The following section provides a detailed overview of the proposed methodology and the processes involved. The flowchart illustrating this method is depicted in Figure 4.

3.1 Data source and preprocessing

3.1.1 Data Source.

The “DataCo supply chain” dataset, which was publicly available on Kaggle, was used in this study. The dataset contains transactional data from an online retail organisation that has dealt with electronic products for over one year. The dataset has 180,519 rows and 53 columns. The data is structured and includes numerous features of a customer’s transaction. The methodology we adopted for this study is designed to evaluate and compare various supply chain analytics models systematically. This approach ensures a complete understanding of model performance in different scenarios within the supply chain domain.

3.1.2. Data Preprocessing

To ensure the accuracy, integrity, and readiness of the dataset for model development, a systematic preprocessing approach was undertaken. This phase encompassed data cleaning, transformation, and feature selection—each essential for ensuring analytical reliability and facilitating optimal model performance.

The following key steps were implemented:

1. **Handling Missing Values:** The dataset contained missing values across several attributes. The Product Description column, which lacked any usable entries, was removed entirely. Similarly, the Order Zipcode field exhibited over 85% missing data and was excluded due to insufficient representativeness. In contrast, the Customer Lname column had a moderate level of missing values, which were imputed using the placeholder value "Unknown" to preserve data completeness without introducing bias.
2. **Data Type Conversion:** Several fields were stored in inappropriate data types. For instance, Order Date and Shipping Date, originally in object format, were converted to datetime format. This transformation enabled the extraction of temporal features and supported subsequent time-series analysis.
3. **Feature Selection:** Features were selected based on a combination of domain relevance, statistical correlation, and model interpretability. This process ensured that only the most informative and non-redundant attributes were retained for training. By

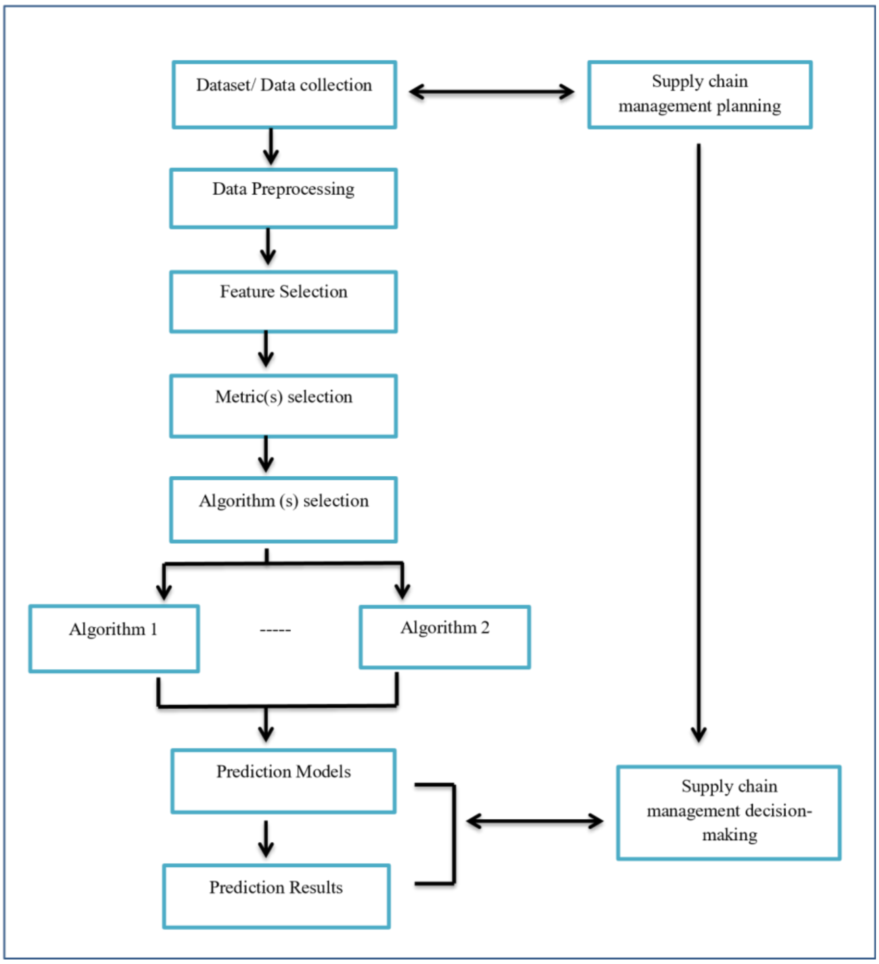

Fig. 4. Supply Chain Analytics - Performance Evaluation Flowchart

reducing dimensionality, the models were rendered more efficient, interpretable, and robust to overfitting.

4. **Data Cleaning:** The dataset was further refined by eliminating duplicate records, correcting inconsistencies, and addressing outliers. Duplicate rows were identified by checking for identical entries across all fields and were subsequently removed. Outliers were detected using z-score normalisation ($|z| > 3$) and visually confirmed via box plots. Depending on their contextual validity, they were either removed or capped. Additional data inconsistencies were resolved using domain knowledge to maintain data integrity.
5. **Data Integration:** Relevant information distributed across multiple tables was

consolidated into a unified dataset using unique identifiers such as Order ID and Order Item ID. This integration facilitated coherent analysis across the various supply chain scenarios addressed in the study

The resulting structured and cleaned dataset formed the analytical foundation for the four use cases examined in this research: *Supply Chain Forecasting – Sales Prediction, Late Delivery Risk, Predicting Shipment Duration, and Route Mapping.*

3.2 Exploratory Data Analysis

Exploratory Data Analysis (EDA) is an essential step in gaining insights into the dataset, helping to identify patterns, irregularities and relationships essential for the modelling process. This phase informs data preparation and model selection, ensuring that the analysis is grounded in a thorough understanding of the underlying data characteristics.

EDA is particularly important in supply chain analysis, as highlighted by (Colicchia et al., 2018) and (Dohale et al., 2021). It aids in understanding the data, detecting outliers, and perform descriptive statistics. EDA is necessary according to (Morgenthaler, 2009) and (Li Vigni et al., 2013) as it helps to understand the significant aspects of a dataset such as its distributions, patterns, outliers, correlations and missing values. EDA gives data insights and helps make informed decisions about data preparation, feature engineering, and modelling, resulting in improved results. This, in turn, enhances the overall modelling process and results.

In this section, the dataset from DataCo is explored in detail, focusing on its properties and characteristics. The dataset contains 180,519 records with 53 attributes, which includes a mix of categorical, numerical and date-time data. The dataset contains information on customers, orders, products and shipping details, and is analysed through four supply chain use cases.

The EDA process begins with an investigation of the summary statistics of the numerical columns, providing an overview of the central tendencies, variances and other key metrics. The distribution of these variables is then visualised using histograms, box plots or density plots, focusing on metrics such as shipping, sales per customer, and benefit per order. For categorical variables such as Delivery Status, delivery, Customer Segment, etc., bar plots are used to display the frequency of each category.

Relationships between variables are analysed using scatter plots, correlation matrices or heat maps. For instance, the relationship between Sales per customer and Order Item Total is explored to determine any correlation. Additionally, potential links between Late delivery risk and Delivery Status are investigated.

Outliers, missing values and inconsistencies are also identified during EDA. Outliers are detected using box or scatter plots, while missing values are highlighted through summary statistics or visualisation techniques such as heat maps.

Finally, the distribution of data across different categories is examined. For example, the distribution of shipping days across various customer segments or order regions is analysed to uncover patterns or differences. Similarly, the distribution of Late delivery risk across different market segments or shipping modes is examined to identify trends or patterns.

Overall, EDA provides a comprehensive understanding of the dataset, helping to identify patterns, trends, and guiding the selection of appropriate models for addressing the research questions.

3.3 Feature Selection and Engineering

The primary goal of feature selection and engineering in this study was to refine the dataset to enhance model accuracy, reduce computational complexity, and increase interpretability. This process involved identifying the most impactful features and transforming the data into formats more suitable for predictive modelling. According to (Akbar et al., 2021) and (Brintrup et al., 2019), effective feature selection and engineering are crucial in supply chain analytics, where data is often heterogeneous, multidimensional, and context-dependent.

This section outlines the techniques used for selecting and engineering features, as well as the rationale behind the inclusion and exclusion of specific attributes.

3.3.1 Feature Selection Techniques

1. **Correlation Analysis:** Pearson correlation coefficients were computed to assess linear relationships between independent features and target variables. Features with very low correlation ($|r| < 0.1$) were considered weak predictors and excluded from the final model pipeline. This helped reduce noise and avoid redundancy in the feature space, particularly in regression-based tasks such as sales forecasting and shipment duration.
2. **Recursive Feature Elimination (RFE):** RFE was applied to systematically remove the least impactful features by recursively training models and ranking features based on importance. This method was utilised for both classification and regression tasks. For each use case, the top 10 features were retained based on performance scores across cross-validation folds, ensuring generalisability and relevance.
   For transparency, the top 10 features selected by the RFE process for each use case are summarised below. These features consistently demonstrated strong predictive capacity across multiple cross-validation folds:

– **Sales Prediction:** Order Quantity, Product Price, Order Item Total, Shipping Mode, Customer Segment, Order Priority, Product Category, Profit Margin, Order Region, Shipping Duration.
– **Late Delivery Risk:** Shipping Mode, Distance, Weather Conditions, Delivery Segment, Product Category, Order Priority, Carrier Name, Shipment Date, Customer City, Delivery Status.
– **Shipment Duration:** Distance, Weather, Shipment Date, Shipping Mode, Carrier Name, Order Priority, Product Category, Customer Location, Delivery Segment, Day of the Week.
– **Route Mapping:** Distance, Road Conditions, Transport Mode, Order Region, Order

ID, Delivery Zip Code, Traffic Pattern, Carrier Name, Delivery Segment, Customer Latitude/Longitude

The feature ranking was determined individually for each model and averaged across repetitions to ensure consistency.

3. **Principal Component Analysis (PCA):** To manage high-dimensional data and improve computational efficiency, PCA was used to reduce the number of variables while retaining the variance in the dataset. Though interpretability was reduced, PCA was particularly useful in shipment duration and route mapping tasks, where interaction effects were more complex. This aligns with approaches in supply chain dimensionality reduction discussed in (How and Lam, 2018) and (Wang et al., 2020).
4. **Feature Importance (Tree-based Models**): Feature importance scores from Random Forests were used to rank variables based on their contribution to prediction accuracy. This was especially effective for categorical and nonlinear interactions in late delivery risk and route optimisation models. Features consistently ranked low were excluded from further processing to reduce overfitting and improve model robustness.

Features were excluded if they exhibited high multicollinearity, contained excessive missing data (e.g., *Product Description*), had low variance, or lacked contextual relevance as judged by domain knowledge. This curation step ensured that only meaningful predictors were retained, enabling cleaner, more interpretable models.

3.3.2 Feature Engineering Techniques

1. **Date Features:** Calendar-based features such as day of the week, day of the month, and month were extracted from Order Date and Shipping Date fields. These features were critical in capturing seasonal trends and lead time patterns.
2. **Categorical Features:** Categorical attributes like Market, Delivery Segment, and Customer City were encoded using one-hot encoding. This allowed models to effectively leverage the categorical structure of the dataset without introducing ordinal assumptions.
3. **Numerical Features:** Numerical variables were standardised using the StandardScaler to bring them onto a comparable scale. This step was essential to ensure balanced contribution to distance-based models and neural networks.
4. **Text Features:** The Product Description field was sparsely populated and thus removed from modelling, though initial experiments involved tokenisation, stopword removal, and stemming to evaluate its utility. Future studies could revisit this attribute with a more complete dataset.

3.4.3 Key Performance Indicators (KPIs) The KPIs selected for this study were chosen based on their significance in supply chain operations and alignment with business-critical outcomes. Each KPI corresponds to one of the four use cases and reflects a distinct operational objective:

– *Sales Forecasting (Demand Prediction):* Vital for inventory control, procurement planning, and reducing stockouts or overstocking.
– *Late Delivery Risk:* Linked directly to customer satisfaction and operational efficiency; enables proactive intervention in logistics.
– *Shipment Duration*: Important for scheduling, cost estimation, and service-level management.
– *Route Mapping:* Supports optimisation of transportation paths, contributing to cost reduction and timely delivery.

These KPIs are widely recognised in both academic and industry literature as fundamental to effective supply chain decision-making. Their inclusion in this study ensures that the modelling outputs are not only technically sound but also practically applicable in real-world supply chain contexts.

3.4 Selection and Implementation of ML, time series and statistical model:

Various models were slected based on their ability to address different supply chain analytics challenges. The right approach is used in the implementation process to guarantee that each model is set up and evaluated correctly for the particular use case. This systematic approach is critical for making a fair and useful comparison of model performances.

This study compares and contrasts several time-series, statistical, and machine learning algorithms for supply chain analytics. The selection and application of these statistical and machine learning models are detailed in this section. The experimental results for the four use cases - supply chain forecasting (sales prediction), late delivery risk, shipment duration prediction - are all discussed.

3.4.1 Selection of ML, Statistical and Time series Model: In this research, various machine learning, statistical and time-series models were evaluated to assess their performance in supply chain analytics. The machine learning models considered include Linear Regression, Lasso Regressor, Ridge Regressor, KNN Classifier, Gaussian Naive Bayes, SVM Classifier, Gradient Boost, Decision Tree, Random Forest, LSTM Model, and Gated Recurrent Unit. The statistical models considered include the chi-square test, OLS, T-test, Multiple Linear Regression and Kruskal-Wallis test. The Exponential Smoothing Model, SRIMA, and ARIMA are the time-series models considered.

The choice of a particular algorithm was guided by the specific type of use case and the datatype employed. For shipment duration and demand forecasting tasks, time-series models, e.g., ARIMA, SARIMA, and Exponential Smoothing, are suitable as they can capture temporal patterns. Classification tasks like late delivery risk are modelled using SVM, Decision Trees, and Random Forests, as these are quite robust with imbalanced and categorical data. Linear, Lasso, and Ridge regressors are suitable for predicting continuous outcomes. Deep neural networks, namely, LSTM and GRU, were used due to their ability to model sequential dependencies, particularly in time-series contexts. Various statistical tests were conducted to validate assumptions and relationships within the data for the enhanced robustness of our models. In summary, each selected method was chosen based on its compatibility with the problem complexity, data type, and interpretability needs.

3.4.2 Implementation of Machine Learning and Statistical Models The data preprocessing and exploratory data analysis (EDA) provided critical insights into the dataset, including the presence of missing values, data imbalance, outliers, and feature redundancies. These findings directly informed both the feature selection process and the subsequent model implementation strategy. Specifically, the correlation analysis, Recursive Feature Elimination (RFE), and Random Forest feature importance methods enabled the identification of high-impact attributes tailored to each predictive task. For each use case, models were chosen based on the type of target variable (classification or regression), the nature of the features (categorical, numerical, temporal), and the desired balance between interpretability and performance. The refined feature sets ensured alignment with domain knowledge and model compatibility. The selected models were then trained on the cleaned and feature-engineered dataset, and evaluated using appropriate metrics: Root Mean Square Error (RMSE), Mean Square Error (MSE), and R-squared (R2) for regression tasks, and accuracy for classification tasks.

– **Supply Chain Forecasting – Sales Prediction:**

This regression task aimed to forecast sales using structured transaction data. Key features selected via RFE included *Order Quantity, Product Price, Order Item Total, Shipping Mode, Customer Segment, Order Priority, Product Category, Profit Margin, Order Region, and Shipping Duration*. A variety of machine learning, statistical, and time-series models were applied, including Linear Regression, Lasso, Ridge, Decision Tree, Random Forest, and SARIMA. Among these, the Random Forest and Decision Tree models achieved the highest accuracies of 0.99984 and 0.9, respectively. SARIMA emerged as the most effective time-series model with an accuracy of 71.88%. Monte Carlo simulation further validated SARIMA under stochastic demand scenarios, maintaining consistent accuracy.

– **Late Delivery Risk:**

This classification task focused on predicting the probability of shipment delays due to variables such as transport issues or environmental disruptions. The most impactful features included *Shipping Mode, Distance, Weather Conditions, Delivery Segment, Product Category, Order Priority, Carrier Name, Shipment Date, Customer City, and Delivery Status.* Applied models included SVM, Random Forest, XGBoost, and Gaussian Naive Bayes, alongside statistical techniques like Chi-Square Test. Among them, Random Forest and XGBoost achieved top accuracies of 0.90443 and 0.90442, respectively. Time-series modelling with ARIMA yielded 98.28% accuracy, which remained robust (93.23%) under Monte Carlo simulations simulating delayed scenarios.

– **Predicting Shipment Duration:**

This regression-based use case estimated delivery times using features such as *Distance, Weather, Shipment Date, Shipping Mode, Carrier Name, Order Priority, Product Category, Customer Location, Delivery Segment, and Day of the Week.* Machine learning models including Random Forest and XGBoost were evaluated, with Random Forest achieving an accuracy of 0.81351. Statistical and time-series models such as Multiple

Linear Regression, Exponential Smoothing, and SARIMA were also assessed. The Exponential Smoothing model yielded the highest accuracy at 99.37%, a result that remained consistent during Monte Carlo simulations simulating transportation delays and weather impacts.

Route Mapping:
The final use case explored route optimisation using a classification approach informed by spatial and logistical data. The most relevant features included *Distance, Road Conditions, Transport Mode, Order Region, Delivery Zip Code, Traffic Pattern, Carrier Name, Delivery Segment, Customer Latitude, and Customer Longitude.* Gradient Boost, Random Forest, and Multilayer Perceptron models were employed, along with ARIMA for time-dependent routing decisions. The Random Forest model achieved the highest accuracy of 0.98205, while ARIMA reached 99.13% accuracy in timeseries forecasting. Monte Carlo simulations confirmed the robustness of ARIMA and its derivatives under variable routing conditions, such as fluctuating delivery windows and traffic disruptions.

Overall, this implementation strategy ensured that models were not only statistically valid but also operationally meaningful, enhancing their potential application in realworld supply chain management contexts.

3.5 Monte Carlo simulation methodology and implementation

Monte Carlo simulation are crucial in supply chain management for addressing uncertainty and testing model robustness across various scenarios. This method helps assess how well adapted and reliable our models represent variation among these obstacles. For a given variable or set of variables, to construct its probable distribution (if it has one), we use the computational procedure known as Monte Carlo simulation, which produces random samples. This technique involves generating random samples to simulate different scenarios and assess model performance under varying conditions. This section outlines the methodology and implementation of Monte Carlo simulations for the identified use cases.

– Steps in Monte Carlo simulation Methodology:

1. Data Preprocessing: Data cleaning is performed to remove missing value and irrelevant information, preparing the dataset for analysis.
2. Exploratory Data Analysis (EDA): EDA techniques are employed to identify patterns, relationships, and categories within the dataset, guiding the selection of models.
3. Feature Selection: Relevant features are selected based on their relationships and importance, helping to optimise model performance
4. Algorithm Selection: Appropriate machine learning, statistical, or time-series algorithms are selected based on the dataset’s characteristics and the problem at hand.
5. Model Development: After the algorithm is developed, models are trained on the training data and tested on a test dataset to evaluate their performance.
6. Monte Carlo Simulation: Once the model is built, the next step must be to perform a Monte Carlo simulation. Random samples are drawn from the probability distribution

of variables, and different scenarios are simulated to assess model performance.

7. Results Evaluation: The final step in the Monte Carlo simulation approach is to evaluate outcomes. The model's performance is measured using accuracy-based metrics such as root mean square error (RMSE), R2 and average (mean) values.

– Monte Carlo simulation methodology is implemented for the identified use cases, as explained below:

1. Data Preprocessing: The values of sales from the above dataset were converted into datetime, resampled to monthly frequency and summed for analysis.
2. Exploratory Data Analysis: EDA was conducted to uncover trends and patterns in the data.
3. Feature Selection: Relevant features for model development were selected through feature selection.
4. Algorithm Selection: For model development, the algorithms selected were ARIMA, SARIMA and Exponential Smoothing.
5. Model Development: Models were trained and tested, with performance evaluated using RMSE, R2 and error of mean square to gauge all models 'performance in terms of accuracy.
6. Monte Carlo Simulation: A Monte Carlo simulation was performed using random samples drawn from the probability distribution of a given variable.

Varying Simulation Conditions: The Monte Carlo simulations were designed to reflect uncertainty in key input variables such as demand volume, shipping lead times, and weather-related disruptions. By repeatedly sampling from probability distributions defined for these variables, we generated a range of realistic operational scenarios. This allowed us to assess how model performance metrics (e.g., RMSE, R2) responded to these fluctuations, thereby testing the robustness of each model under variable conditions reflective of real-world supply chain environments.

4 Results

In this research study, the performance of multiple models and algorithms were evaluated across four distinct use cases in supply chain analytics using the DataCo supply chain dataset (Constante et al., 2021). The analysis aimed to identify the most efficient models for each use case, with a focus on accuracy, performance and alignment with the ISO25010 quality model. To ensure internal validity of our experiments, variables and conditions were meticulously controlled during the modelling process, minimising biases and errors. External validity has been addressed by selecting a diverse and representative dataset, helping to generalise the findings to broader supply chain contexts. Criterion validity was established through alignment with the ISO25010 quality model, ensuring that the assessment criteria are industry-relevant and standardised.

For each use case, a range of machine learning, statistical, and time-series models were tested, both with and without Monte Carlo simulation. The primary objective was to ascertain the most efficient models and algorithms for each use case, focusing on accuracy, performance, and alignment with the ISO25010 quality model for product quality evaluation. This comprehensive approach ensures a thorough assessment that extends beyond basic performance metrics.

4.1 Supply Chain Forecasting - Sales Prediction

The initial focus was on predicting sales within the supply chain. Several machine learning models, statistical models, and time-series models were explored. The highes performing accuracy results for each category are presented in Table 5. The robustness of these results is supported by rigorous statistical testing and cross-validation techniques, ensuring their reliability and validity.

Table 5. Accuracy of Models for Supply Chain Forecasting - Sales Prediction

Model Category	Best Accuracy
Machine Learning	Decision Tree (0.99989)
Statistical	T-test (1)
Time Series	SARIMA (0.6962)
Time Series + Monte Carlo	SARIMA (0.7188)

From Table 5, it is evident that the model Decision Tree is achieving the highest accuracy among all machine learning models, the T-test yielded the best results among the statistical models, and SARIMA outperformed other time series models. Moreover, incorporating Monte Carlo simulation further improved the accuracy of the SARIMA model.

4.2 Late Delivery Risk

In the second use case, the focus was on predicting the risk of late delivery. Machine learning, statistical, and time-series models were applied and their performance evaluated. Table 6 summarises the highest accuracy results for each category.

Table 6. Accuracy of Models for Late Delivery Risk Prediction

Model Category	Best Accuracy
Machine Learning	Random Forest (0.90443)
Statistical	T-test (1)
Time Series	ARIMA (0.9828)
Time Series + Monte Carlo	ARIMA (0.9324)

According to Table 6, a similar trend in model performance to that demonstrated in Table 6 is presented. The Random Forest model demonstrates the highest accuracy among the machine learning models, the T-test achieved the best results among the statistical models, and the ARIMA model outperforms other time-series models. After we added Monte Carlo simulation, it further improved the ARIMA model's accuracy.

4.3 Predicting Shipment Duration

The third use case centered on predicting shipment duration. To assess accuracy, various machine learning, statistical and time-series models were employed. A summary of the accuracy results is provided in Table 7.

Table 7. Accuracy of Models for Predicting Shipment Duration

Model Category	Best Accuracy
Machine Learning	Random Forest (0.81351)
Statistical	Chi-square test (0.74826)
Time Series	Exponential Smoothing (0.99372)
Time Series + Monte Carlo	Exponential Smoothing (0.99372)

From Table 7, it is observed that the model Random Forest is achieving the highest accuracy among the machine learning models, the Chi-square test yielded the best results among the statistical models, and the Exponential Smoothing model outperformed other time-series models. The inclusion of Monte Carlo simulation did not significantly impact the accuracy in this particular use case.

4.4 Route Mapping

The fourth use case involved route mapping. To evaluate accuracy, machine learning models, statistical models, and time-series models were applied. The summarised accuracy results for each category are represented in Table 8.

Table 8. Accuracy of Models for Route Mapping

Model Category	Best Accuracy
Machine Learning	Random Forest (0.98205)
Statistical	Multilayer Perceptron (0.92686)
Time Series	ARIMA (0.99135)
Time Series + Monte Carlo	ARIMA (0.99135)

According to Table 8, the Random Forest model achieved the highest accuracy among the machine learning models, the Multilayer Perceptron model demonstrated the best results among the statistical models, and the ARIMA model outperformed other time-series models. The inclusion of Monte Carlo simulation also did not significantly impact the accuracy in this use case.

To visually compare performance across the four case studies, Figure 5 displays a grouped bar graph illustrating the accuracy of different model categories. In conclusion, the research study employed a range of models and algorithms to address various use cases in supply chain analysis. The analysis identified the best-performing models and algorithms based on accuracy, with Decision Tree, Random Forest, ARIMA, and Exponential Smoothing outperforming others across different scenarios. These results highlight the effectiveness of machine learning, time-series and statistical models in enhancing supply chain analytics and supporting informed decision-making processes. Note: For the code discussed in this paper and detailed results, and comparison table see the Appendix section.

While the dataset provides a comprehensive view of supply chain operations, the authors recognise limitations in its representativeness. To enhance the generalisability of the findings, future studies are encouraged to replicate the methodology across diverse datasets. To further validate the results, supplementary material are available on GitHub, including detailed experimental setups and additional analyses, promoting transparency and reproducibility of this research.

In conclusion, the study presents a detailed comparative analysis of various models in supply chain analytics, identifying the best-performing models based on accuracy and robustness. These findings, validated through rigorous experimental designs and alignment with industry standards, offer significant insights for enhancing decision-making processes in supply chain management. However, the scope for further validation on diverse datasets remains, inviting additional research to confirm and expand upon these conclusions.

5 Discussion

5.1 Factors Influencing Model Accuracy

Across the four use cases, several key factors were observed to influence model accuracy:

- **Data Quality and Completeness:** Features with missing or inconsistent data (e.g., Product Description, Order Zipcode) negatively impacted accuracy if not adequately handled.
- **Feature Relevance and Selection:** Models using domain-informed feature selection (e.g., RFE, correlation filtering) performed significantly better, particularly in shipment prediction and late delivery classification.

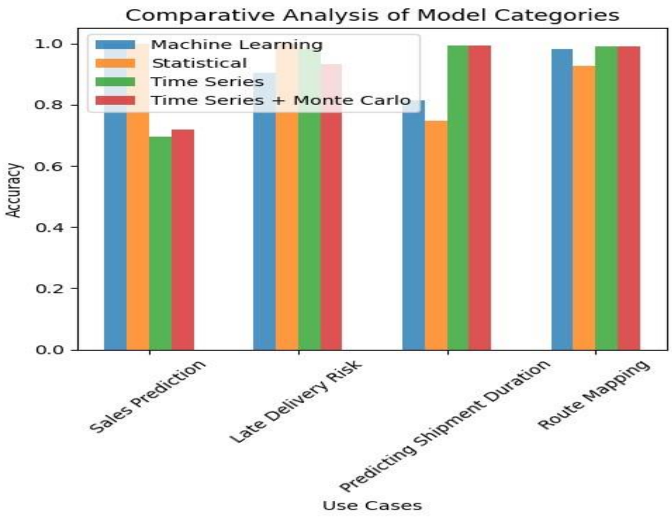

Fig. 5. Comparative Analysis of Model Categories for each Use Case

- **Model Complexity:** Ensemble models (e.g., Random Forest, XGBoost) consistently outperformed simpler linear models due to their ability to capture nonlinear relationships, especially under varied data distributions.
- **Simulation Conditions:** Model performance varied with changing assumptions in Monte Carlo simulations—models that were less sensitive to fluctuations in input distributions showed better reliability under operational uncertainty.

These factors highlight the importance of preprocessing, model selection, and simulation design in achieving robust and generalisable results in supply chain analytics.

5.2 Implications for Research and Practice

The results of this study provide meaningful contributions to both academic research and supply chain management practice.

Implications for Research: This study advances the scholarly understanding of predictive modelling in supply chain analytics by offering a systematic, multi-model evaluation across four critical use cases. Unlike prior research that often focuses on isolated techniques or single use-case analysis, this work offers a comparative framework that integrates machine learning, statistical, and time-series approaches within a common dataset. Furthermore, the inclusion of simulation-based testing (Monte Carlo simulations) under varying conditions enhances the methodological robustness, offering insights into model stability and generalisability—an area that remains underexplored in supply chain analytics literature.

In addition, the alignment of model evaluation with the ISO/IEC 25010 quality framework introduces a novel validation layer rarely employed in data analytics studies. This bridges the gap between performance-focused evaluation and system-level quality attributes such as reliability, usability, and efficiency, thereby setting a precedent for more structured model assessment in future analytics research.

Implications for Practice: From a practical standpoint, this research offers decisionmakers guidance on selecting suitable predictive models for specific supply chain objectives. For instance, the superior performance of Random Forest and XGBoost in tasks involving classification and regression suggests their applicability in operational scenarios such as delivery risk prediction and demand forecasting. Similarly, time-series models like SARIMA and Exponential Smoothing demonstrate robustness in handling historical trend-based forecasting.

The study also underscores the importance of feature selection and preprocessing in achieving high model performance. Supply chain professionals and data analysts can benefit from incorporating structured feature selection techniques and simulation-driven stress testing to improve model reliability in real-world applications

Furthermore, by highlighting the influence of data quality, task alignment, and model complexity on predictive accuracy, this research contributes to the development of more intelligent, data-driven supply chain management systems. It encourages a move toward adaptive analytics solutions that are not only accurate but also transparent and deployable within operational decision-support tools.

Overall, this study lays the groundwork for the practical implementation of AI and machine learning models in supply chain contexts, fostering a more evidence-based and resilient approach to supply chain optimisation.

5.3 Validation Framework

To ensure the robustness and applicability of our results, we adopted the ISO/IEC 25010 standard as a guiding framework for evaluating model quality. ISO/IEC 25010 is a widely recognised international standard that defines a quality model for software and system evaluation, encompassing eight quality characteristics: functionality, reliability, usability, efficiency, maintainability, portability, compatibility, and security.

In the context of supply chain analytics, where predictive models serve as decision support systems, this standard offers a structured and holistic approach to validate the performance and operational relevance of the developed models. For this study, we focused on a subset of ISO/IEC 25010 characteristics that are most applicable to data driven predictive models:

- **Functionality:** The ability of the model to fulfil its intended purpose—accurately predicting KPIs such as sales, shipment duration, or delivery risk.
- **Reliability:** Evaluated through statistical validation methods (e.g., cross-validation, RMSE, MSE), assessing whether the model consistently produces stable outputs under different data splits or simulation scenarios.
- **Efficiency:** Considered in terms of computational performance and resource usage during training and inference.
- **Usability:** Measured indirectly by the interpretability and ease of deployment of models, particularly those used in operational decision-making (e.g., tree-based models).

– **Maintainability and Portability**: Reflected in the modular pipeline design and compatibility with standard machine learning environments (e.g., Python, scikitlearn).

The inclusion of the ISO/IEC 25010 framework supports the generalisability and practical applicability of our findings, aligning the evaluation of analytical models with recognised software quality principles. This contributes to a more rigorous and standardised validation process, ensuring that the models not only perform well statistically but also meet broader expectations for operational deployment in supply chain contexts.

6 Conclusion

This research examined the effectiveness of various machine learning, time-series and statistical models in supply chain analytics, focusing on four use cases: Sales Prediction, Late Delivery Risk, Predicting Shipment Duration and Route Mapping. The analysis included comprehensive data preprocessing, exploratory data analysis, feature selection, selection of algorithm, and prediction model development. WThe models were evaluated using metrics such as r2 score, mean square error, and RMSE, all applied to the DataCo supply chain dataset.

In the first use case, Supply Chain Forecasting – Sales Prediction, the Gaussian Naive Bayes model exhibited the highest accuracy among all models, achieving an accuracy of 91.35%. Conversely, the Gradient Boost and LSTM models demonstrated the lowest accuracy, yielding negative scores. For the second use case, Late Delivery Risk, the Random Forest and XGBoost models performed exceptionally well, with accuracies of 90.44% and 90.43%, respectively. The Lasso Regressor and LSTM Models, however, showed the lowest accuracy with negative scores. In the third use case, Predicting Shipment Duration, the Exponential Smoothing Model demonstrated the highest accuracy of 99.372%, while the KNN Classifier and SVM Classifier recorded the lowest accuracies, both showing negative scores. Finally, for the fourth use case, Route Mapping, the Random Forest model achieved an accuracy of 39.74%, while the Lasso Regressor performed the least accurately with a negative score.

The study reveals that the performance of the models varied across different use cases. In some scenarios, machine learning models outperformed statistical and timeseries models, whereas in other cases, statistical and time-series models outperformed machine learning models. Notably, the Exponential Smoothing model consistently demonstrated high accuracy across all use cases, making it a reliable choice for supply chain analytics.

In conclusion, thi study underscores the importance of selecting the appropriate model for each use case in supply chain analytics. Understanding the strengths and weaknesses of different models can assist supply chain management professionals in making informed decisions and optimising their supply chain operations. Further research could explore the performance of additional models and incorporate real-time data to enhance the accuracy and applicability of supply chain analytics.

6.1 Limitations and direction of future research

While the research provides valuable insights into the performance of various models in supply chain analytics, it is essential to acknowledge certain limitations and suggest directions for future exploration.

6.1.1 Limitation

- *Dataset Limitations:* The dataset, though it includes over 100’000 records, may still be small for supply chain applications, typically needing larger, more diverse datasets. The dataset also lacks diversity in product types, customer profiles, and geographic coverage. We understand that these issues can impact the ability of the model to generalize and exhibit good performance in real-world cases, but the clean, structured, and accessible nature of the data on shipments, orders, and deliveries is good for validating the initial hypothesis and prototyping models. However, less diversity in the dataset, specifically in product, customer, and geographic information, might hamper the ability of the model to generalize in real-world environments. This can be addressed in future work by integrating a larger number of diverse datasets, spanning different customer segments, locations, and product types.
- Scope of Use Cases: Our research focused on four specific use cases in supply chain analytics. While we chose these because of their relevance and applicability, other use cases in supply chain management should have been covered in this study. This limitation might be affecting the generalizability of our findings to these unexplored areas.

6.1.2 Future Research Direction

To fix the weak points and improve the range, we provide the study:

- *Growing the Use Cases and Dataset:* This means that we can use the big datasets to cover all supply chains, which can be done in future studies. This will provide more correct results and improve the model performance. Also, studying use cases other than the four this study looked at would give a bigger view of how useful and correct the model is.
- *Combining Models and Approaches:* Employing a combination of various models and approaches could potentially boost model accuracy. Integrating different modeling techniques might offer more nuanced insights and improved predictive capabilities. The use of hybrid models—such as combining time-series techniques with machine learning—offers strong potential for improving prediction accuracy and capturing both temporal patterns and complex nonlinear relationships in certain use cases, like forecasting and shipment duration prediction.
- *Exploring Additional Performance Metrics:* Investigating several performance variables, such as recall, precision, and F1-score, could provide a more rounded assessment of model effectiveness. These metrics would offer additional dimensions to model evaluation, especially in scenarios where accuracy alone might not be the sole indicator of model performance.

– *Innovative Strategies in Supply Chain Optimization*: There is scope for exploring other strategies in optimizing supply chain activities. Future research could dive into the potential of emerging techniques like reinforcement learning and transfer learning. These advanced methods could offer novel solutions to complex supply chain challenges.

This research lays the groundwork for future studies in supply chain analytics. By acknowledging the limitations of our study and proposing these future research directions, we aim to encourage a more expansive and thorough exploration of this field. The continued investigation and improvement of models and methodologies will undoubtedly contribute to more robust and effective supply chain management strategies.

References

Anitha, P., Patil, M. M. (2018). A review on data analytics for supply chain management: a case study. *International Journal of Information Engineering and Electronic Business*, 14(5), 30.

Babaee Tirkolaee, E., Sadeghi, S., Mooseloo, F. M., Vandchali, H. R., Aeini, S. (2021). Application of machine learning in supply chain management: A comprehensive overview of the main areas, *Math. Probl. Eng*. 2021, 1–14.

Bousqaoui, H., Achchab, S., Tikito, K. (2017). Machine learning applications in supply chains: An emphasis on neural network applications, *2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech)*, Rabat, Morocco, 2017, pp. 1-7.

Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., McFarlane, D. (2019). Supply chain data analytics for predicting supplier disruptions: A case study in complex asset manufacturing, *Int. J. Prod. Res.* **58**, 3330–3341.

Colicchia, C., Creazza, A., Menachof, D. A. (2018). Managing cyber and information risks in supply chains: Insights from an exploratory analysis, Supply Chain Manag.: Int. J. **24**, Dec. 2018.

Constante, F., Silva, F., Pereira, A. (n.d.). Dataco smart supply chain for big data analysis, Data.Mendeley.com, 5.

Dash, R., McMurtrey, M., Rebman, C., Kar, U. K. (2019). Application of artificial intelligence in automation of supply chain management, *J. Strateg. Innov. Sustainab.* **14**, Jul. 2019.

Dohale, V., Ambilkar, P., Gunasekaran, A., Verma, P. (2021). Supply chain risk mitigation strategies during COVID-19: Exploratory cases of "make-to-order" handloom saree apparel industries, *Int. J. Phys. Distrib. Logist. Manag*., ahead-of-print, Apr. 2021.

Estampe, D., Lamouri, S., Paris, J.-L., Brahim-Djelloul, S. (2013). A framework for analysing supply chain performance evaluation models, *Int. J. Prod. Econ.* **142**, 247–258.

Fattah, J., Ezzine, L., Aman, Z., El Moussami, H., Lachhab, A. (2018). Forecasting of demand using ARIMA model, *Int. J. Eng. Bus. Manag*. **10**, ID 184797901880867.

Ghanadian, S. A., Ghanbartehrani, S. (2021). Evaluating supply chain network designs: An approach based on SNA metrics and random forest feature selection, *Univ. J. Oper. Manag.* **1**, 15–35.

Guanghui, W. (2012). Demand forecasting of supply chain based on support vector regression method, *Procedia Eng.* **29**, 280–284.

Hahn, G. J. (2019). Industry 4.0: A supply chain innovation perspective, *Int. J. Prod. Res*. **58**, 1425–1441.

How, B. S., Lam, H. L**.** (2018). Sustainability evaluation for biomass supply chain synthesis: Novel principal component analysis (PCA) aided optimisation approach, *J. Clean. Prod.* **189**, 941–961.

ICML (2018). DiCE: The infinitely differentiable Monte Carlo estimator, *Proc. Int. Conf. Mach. Learn.*, 2018.

Kilimci, Z. H., Akyuz, A. O., Uysal, M., Akyokus, S., Uysal, M. O., Bulbul, B. A., Ekmis, M. A. (2019). An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain, Complexity 2019, 1–15.

Li Vigni, M., Durante, C., Cocchi, M. (2013). Exploratory data analysis, in *Data Handling in Science and Technology,* pp. 55–126.

Mariappan, M. B., Devi, K., Venkataraman, Y., Lim, M. K., Theivendren, P. (2022). Using AI and ML to predict shipment times of therapeutics, diagnostics and vaccines in e-pharmacy supply chains during COVID-19 pandemic, *Int. J. Logist. Manag*., Jan. 2022.

Mohamed, S., Rosca, M., Figurnov, M., Mnih, A. (2020). Monte Carlo gradient estimation in machine learning, *J. Mach. Learn. Res*. **21**, 1–62.

Morgenthaler, S. (2009). Exploratory data analysis, *Wiley Interdiscip. Rev. Comput. Stat*. **1**, 33–44.

Nguyen, H. D., Tran, K. P., Thomassey, S., Hamad, M. (2021). Forecasting and anomaly detection approaches using LSTM and LSTM autoencoder techniques, *Int. J. Inf. Manag.* **57**, 102282.

Raman, S., Patwa, N., Niranjan, I., Ranjan, U., Moorthy, K., Mehta, A. (2018). Impact of big data on supply chain management, *Int. J. Logist. Res. Appl*. **21**, 579–596.

Saci, S. (2023). What is a supply chain digital twin?, Mar. 2023.

Sahoo, R., Pasayat, A. K., Bhowmick, B., Fernandes, K., Tiwari, M. K. (2021). A hybrid ensemble learning-based prediction model to minimise delay in air cargo transport, *Int. J. Prod. Res.*, 1–17.

Shih, H., Rajendran, S. (2019). Comparison of time series methods and machine learning algorithms for forecasting Taiwan's blood supply, *J. Healthc. Eng*. 2019, 1–6.

Surie, C., Reuter, B. (2014). Supply chain analysis, in *Springer Texts in Business and Economics,* pp. 29–54.

Terrada, L., El Khaili, M., Ouajji, H. (2022). Demand forecasting model using deep learning methods for supply chain management 4.0, *Int. J. Adv. Comput. Sci. Appl.* **13**, 2022.

Thete, J. (2022). A stochastic model for demand forecasting in Python, Mar. 2022.

Wan, F. (2021). XGBoost based supply chain fraud detection model, Mar. 2021.

Wang, J., Swartz, C. L. E., Corbett, B., Huang, K. (2020). Supply chain monitoring using principal component analysis, *Comput. Chem. Eng*. **59**, 12487–12503.

Winter Simulation Conference (WSC) (2009). Quantifying supply chain disruption risk using Monte Carlo and discrete-event simulation, *Proc. Winter Simul. Conf*., 2009.

Zekhnini, K., Cherrafi, A., Bouhaddou, I., Benghabrit, Y., Garza-Reyes, J. A. (2020). Supply chain management 4.0: A literature review and research framework, *Benchmarking: Int. J.*, **28**(2).

Appendices

All the source code discussed in this paper for the implementation of the proposed methodology in the supply chain management of DataCo Smart supply chain Dataset are available under the following Github repository:
https://github.com/Rachana-pandey11/Supply-Chain-Analytics

Received August 20, 2024, revised May 26, 2025, accepted May 26, 2025