Enhancing Supply Chain Efficiency with Predictive Analytics: A Machine Learning Approach to Dynamic Lead Time Prediction in a Chemical Industry

MALLIDIS IOANNIS[1], STAMELOU AFRODITI[1], GEORGIA AYFANTOPOULOU[1], ELIAS KANAKIS[2]
[1]Hellenic Institute of Transport/ Centre for Research and Technology Hellas,
6th km Charilaou, Thermis, 57001,
GREECE

[2]Entersoft S.A.
Leoforos Siggrou 362, Kallithea, 17674,
GREECE

Abstract: - We develop and employ a four-step methodological approach for predicting the lead time delays in all echelons of a supply chain (SC). The first step of the methodological approach involves a critical synthesis of academic research efforts for identifying the main sources of delays in all echelons of a supply chain. The second step involves the development of questionnaires for validating the findings of the research through workshops with industry stakeholders. The third step involves the development of a suite of machine learning (ML) models, namely, Random Forest Regression, Decision Tree Regression, and Linear Regression. These models were selected based on their prevalence in the recent literature and their ability to handle linear and nonlinear relationships between multiple variables. The final fourth step involves the implementation of the suite of machine learning models in the real case of a Hellenic chemical manufacturing supply chain. The implementation results reveal that Random Forest Regression exhibits the highest predictive accuracy throughout all stages of the supply chain, achieving the lowest Mean Absolute Percentage Errors (MAPE), ranging from 0.5 to 7% in the examined supply chain echelons.

Key-Words: - lead time prediction, machine learning, supply chain echelons, Random Forest Regression, Decision Tree Regression, and Linear Regression.

Received: May 5, 2024. Revised: September 21, 2024. Accepted: October 22, 2024. Published: November 14, 2024.

1 Introduction

Lead time variability constitutes one of the main challenges that companies face when optimizing their inventory and production planning decisions. Lead times with high variability, promote the maintenance of higher stocks of raw materials and final products to compensate for lead time disruption events, [1]. It is therefore intuitively sound that through accurate lead time predictions, businesses can optimize their inventory levels and streamline their production schedules, [2].

Machine learning (ML) algorithms have emerged as tools that could dynamically collect historical data and provide real-time predictions of order lead times. These dynamic adaptation capabilities allow for an improved understanding of the complex relationships between multiple independent variables involved in supply chains, [3].

Under this context, the purpose of this paper is to provide a holistic methodological approach for the dynamic prediction of order lead times in all echelons of a supply chain through a suite of ML algorithms. Moreover, we additionally provide practical technical guidelines for developing a dynamic prediction service for a real-world chemical manufacturer's SC in Greece.

The rest of the paper is organized as follows. Section 2 provides a literature review of academic research efforts that employ ML models for dynamic supply chain management, while Section 3, a literature review-based mapping of the main delay sources in all echelons of a manufacturing SC, along with the description of a workshop implementation for validating the literature review findings. Section 4 provides a technical analysis of the dynamic lead time prediction service developed based on the real case of a Hellenic chemical manufacturer, while Section 5 describes the consumption process of the service. Finally, Section 6 summarizes the findings of the research.

2 Literature Review

Due to their ability to train in dynamic data settings, and thus adapt to new dynamic environments, ML models have emerged as useful tools for addressing the complexity of continuously evolving supply chains, [4]. The effectiveness of these models in supply chain management processes is initially demonstrated by the work of [5] who assessed the predictive accuracy of ML models in a flow-shop environment within the optics industry. The predictions of the ML models were then incorporated in a digital data twin to facilitate real-time mapping of the entire flow-shop operations.

The necessity for the dynamic predictions of supply chains is also highlighted by the work of [6] who promote the use of AutoML in empowering SMEs to enhance dynamic data preprocessing and feature engineering and thus, improve the supply chains' dynamic predictive capabilities.

Similarly, [7] employed regression-based ML models in semiconductor manufacturing processes, finding that ML methods excel over traditional models in terms of predictive accuracy and times, particularly when handling complex production variables. [8], employed ensemble ML algorithms enhanced by fuzzy clustering, to predict lead times in a fashion SC, while [9] employed logistic regression and Random Forests in an oil and gas SC.

[10], introduced a dynamic forecasting framework in a make-to-order supply chain, optimizing shipment consolidations effectively.

In the realm of Just-in-Time (JIT) systems, [11] showed how ML could improve operational efficiencies through improved lead time predictions, while [12] and [13] expanded the application of deep learning techniques to provide deeper insights and improve predictive performance in complex manufacturing settings. From the literature reviewed, two primary research gaps have been identified, which present opportunities for substantial academic contributions.

- **Firstly,** there appears to be a lack of research efforts that extend lead-time prediction analyses to encompass all echelons of a supply chain. Existing research predominantly focuses on specific segments or processes within the supply chain, such as manufacturing lead times or delivery processes, without integrating these segments into a holistic supply chain perspective.
- **Secondly,** there seems to be a lack of comprehensive studies that detail the components and processes necessary for developing and implementing a lead time prediction service throughout all echelons of the supply chain in a real-world setting.

This paper addresses the above research gaps through:

1. **Holistic Approach to Lead Time Prediction Across All Echelons**: our paper systematically maps the root causes of delays across multiple levels of the supply chain. By integrating a suite of machine learning algorithms, the employed methodology extends beyond isolated segments of the supply chain to provide a comprehensive prediction model that spans from the supply of raw materials to production, then to the wholesaler, and all the way to the end retailer.
2. **Development and Deployment of a Comprehensive Lead Time Prediction Service**: Our paper details the development of a lead time prediction service, including the crucial components necessary for its success. It provides a clear framework for integrating machine learning algorithms with existing supply chain data systems, which is vital for the practical deployment of predictive analytics. To this end, the paper outlines the steps involved in setting up the prediction service, from data collection and preprocessing to model selection and training.

3 Main Factors Causing Delays in Various Stages of the Supply Chain

3.1 Delay Identifications

A critical step in creating a highly responsive supply chain is to break down the overall delays into delays specific to each stage (such as design, procurement, production, storage, sales - both wholesale and retail - and transportation), as well as by the type of supply chain (including cold chain, bulk and conventional goods, and hazardous materials). An exhaustive detailed review of the literature and analysis of delay factors has been made by the authors, [14], [15], [16], [17], [18].

The primary sources of lead time delays identified in the literature review were examined using the Supply Chain Operations Reference (SCOR) model [19], which classifies sources of delay by the main SC processes namely, design, procurement, production and transportation.

Design: Inability to synchronize stages of SCs, Catastrophic events (wars, earthquakes, floods)/political instability, Delay due to extreme weather, Cyber-attack incidents, Inflexible legal, regulatory, and bureaucratic procedures, and environmental constraints (yes/no), Shortages of SC management personnel.

Procurement: Delay in supply of raw materials, Stop (yes/no) the delay of production processes/ Delay due to non-existence of disruption recovery plan, Catastrophic events (wars, earthquakes, floods)/political instability, Delay due to extreme weather, Cyber-attack incident, Non-satisfaction of the order due to returns of defective and/or damaged products, Shortages of SC management personnel, Delay in the replenishment of raw material orders, Delayed demand satisfaction, Border delay, Delay in the management and sorting of raw materials/products, Delay in clearing the payment of the order by the customer, Delay in customer payments, Delay in stacking and storage of products which require compliance with safety standards.

Production: Delay in processing the order, Stop (yes/no) the delay of production processes, Delay due to non-existence of disruption recovery plan, Catastrophic events (wars, earthquakes, floods)/political instability, Delay due to extreme weather, Cyber-attack incident Non-satisfaction of the order due to returns of defective and/or damaged products, Shortages of SC management personnel, Delay in the replenishment of raw material orders, Delayed demand satisfaction, Border delay, Delay in the management and sorting of raw materials/products, Delay in clearing the payment of the order by the customer, Delay in customer payments, Delay in stacking and storage of products which require compliance with safety standards.

Transportation: Delay in issuing transport documents, Delay in processing the order, Catastrophic events (wars, earthquakes, floods)/political instability, Delay due to extreme weather, Delay in customs clearance of products, Cyber-attack incidents, Inflexible legal, regulatory and bureaucratic procedures and environmental constraints (yes/no), Shortages of SC management personnel, Border delay, Delay in stacking and storage of products which require compliance with safety standards.

The detailed documentation of the delay causes mentioned earlier served as the basis for the participatory processes carried out during the project. These causes were assessed through questionnaire surveys and co-creation workshops involving various companies engaged in different activities at multiple levels of the supply chain.

3.2 Users' Opinion based on Questionnaire Surveys

From the causes of delay identified in the literature, it was critical to select those that best fit the potential customers of the service and cause the actual delays in the delivery time of their orders. For this reason, a questionnaire survey was carried out to the company's customers to identify from the participants the possible causes that cause delays in delivery time and to evaluate the degree to which they are found in the operations of their businesses.

To create the questionnaire, initially, the supply chain was divided into six typical stages that make it up (production, processing, storage, wholesale, retail, and transportation) and the potential sources of delay at each stage were classified. As expected, several sources of delay are common to the various stages of a supply chain.

The survey was carried out between March and May 2022 and a total of 47 responses were collected from representatives of the respective companies employed in the supply chain sector in various roles such as production units, retailers or wholesalers, carriers, etc.

The sources of delay affecting the delivery time of the final product were examined by the supply chain stage. Specifically, companies that selected a stage of the supply chain were redirected to questions about the corresponding sources of delay. In the most common case where they selected more than one stage, questions for all selected stages appeared. For the presentation of the results, they are divided into categories according to the logistics stage to which they refer. Regarding the production stage, it is observed that most of the respondents (79%) answered that the biggest delays appeared in the supply of raw materials and their delivery, either by the supplier or by the transport company. Half of the participants reported experiencing delays during order processing. Additionally, over 2 out of 5 participants (43%) indicated that significant delays occur both during the replenishment of raw material orders and due to insufficient communication between various levels of the supply chain, ultimately affecting the delivery time of the final product.

In contrast, cyber-attacks were not identified as a source of delays for any business. It is noteworthy that while some sources of delay are more prominent than others, almost all play a significant role in the final delivery of the product. At the supply chain processing stage, like the production stage, the primary reasons for delays in the final product's delivery time are related to the delivery of raw materials by suppliers and transport companies, with reported delays of 69% and 62%, respectively. Around half of the respondents mentioned that there are delays during order processing, while about 2 out of 5 encounter problems related to customer payments. Unlike the production stage, both the lack of communication between various supply chain levels and the replenishment process of raw materials are not considered major sources of delay, with only 23% of respondents identifying them as significant factors.

Moving to the storage stage of the supply chain, more than 1 in 2 respondents indicated that delays are most frequently caused by the delivery of products from suppliers and transport companies. It is important to note that there is a significant disparity between the primary source of delays and other contributing factors, as illustrated in the diagram above, where remaining delay factors are reported at much lower percentages (below 30%). Lastly, factors such as the absence of a disruption recovery plan and cyber-attacks, reported by 3% to 0% of respondents respectively, play a negligible role in product delays at the storage stage.

At the transportation stage, a more even distribution of delay sources is observed compared to previous stages of the supply chain. Specifically, the most significant issues arise during order processing and due to weather conditions, reported by 48% and 46% of respondents, respectively. Smaller, but still notable, delays were reported by 38% of respondents, who identified customs clearance during the import of raw materials or export of finished products as a cause of delays. Additionally, around 27% faced delays due to customer payments, and approximately 20% experienced delays when crossing borders. In general, it is observed that the reasons for delays in product delivery at this stage of the supply chain are varied. It is worth mentioning that most respondents' answers were related to the transportation stage, as many do not manage transportation themselves, resulting in various delays at this stage.

At the retail stage, more than 2 out of 3 respondents reported that the most significant delays occur during the delivery of the product by the supplier and the transport company. Most delay sources at this stage range between 26% and 37%, including issues such as customer payment problems and product sorting. Notably, the lack of communication between supply chain levels is only mentioned by 11% of respondents, the lowest percentage compared to other stages. Finally, like the retail stage, the wholesale stage also sees the biggest delays occurring during the delivery of the product.

3.3 Users' Opinion based on Co-Creation Workshops

Subsequently, two co-creation workshops were held to bring together supply chain experts and gather valuable insights into the causes of delays at various stages of the supply chain, resulting in the identification of the most critical causes and some quantitative estimates of these delays. The workshops were designed with a consistent structure and targeted both production and commercial companies. Specifically, the participants represented companies involved in one or more of the following stages: production and processing, storage, transportation, wholesale, and retail trade.

During the participatory activities, the causes of delay identified through the literature review were thoroughly explained. Participants then provided detailed insights into how these delays manifest in their own processes. This step was crucial for validating the delay causes to be used in forecasting project delivery times. Specifically, by examining the procedures where delays most frequently occur, the following conclusions can be drawn:

-Customer payments are one of the most significant factors that can delay the production, shipping, or distribution of an order. Without advance payment and subsequent payment of the total amount, the order cannot be sent to the final customer, leading to delays in production planning, as well as in the storage and transportation of the order.

-Imports and exports of raw materials and finished products, particularly when customs procedures are involved, frequently cause delays and incur additional costs in production and transportation. These delays also affect customer satisfaction and the overall reliability of the business.

-It is observed that most causes of delay are common across several stages of the supply chain and impact multiple departments. Interrelationships between processes and communication between departments are crucial for building more resilient and sustainable supply chains.

-Quantitative estimates of the analysed delays vary depending on the cause. Even a few days of delay can disrupt the supply chain and create significant problems.

-Delays caused by extreme events and social, political, and economic disruptions also contribute to supply chain disruptions, but they do not allow for accurate estimations of lag time.

Given the importance of this analysis, the delivery time prediction algorithm will be based on the main causes of delay identified through both the literature review and the insights gained from the questionnaire survey and expert discussions during the workshops.

4 Development of the Escalator Lead Time Prediction Service

As part of the ESCALATOR project, a suite of machine learning algorithms was developed, building on a critical review of relevant literature. This suite of algorithms was then implemented to predict order lead times at three echelons of the SC, namely, from the supply of raw materials to production, then from production to the wholesaler's warehouse, and from the wholesaler's warehouse to the retailer.

The algorithms integrate seamlessly with Entersoft's Warehouse Management System (WMS) through the "Entersoft Escalator Connector," which facilitates the dynamic transmission of data between Entersoft's systems and the predictive models via RESTful APIs. Figure 1 and Figure 2 present the dynamic data transmission.

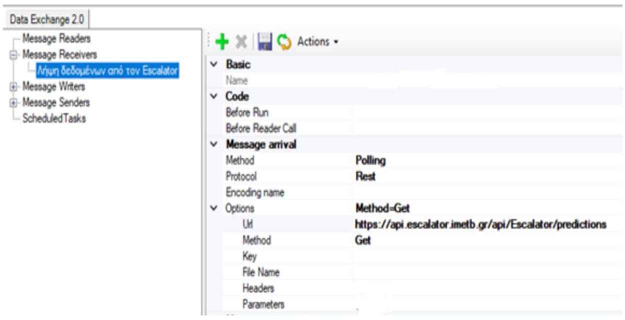

Fig. 1: Get data from the database

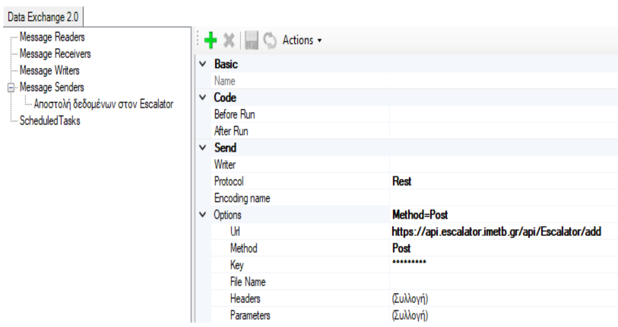

Fig. 2: Post data from the database

This system allows users to send historical data and receive predictions of lead time delays. The independent variables or predictors of lead time delays in each one of the three echelons of the SC, along with the corresponding database schema of each SC echelon are summarized in Table 1, Table 2 and Table 3 respectively.

Table 1. Predictors Raw Material Supply-Production

Column Name in Database	Description
RowID	Registration key
TenantID	Cloud SaaS customer
AMSPLRPaymentDate	Date of payment of the raw materials supplier
AMLSDate	Date of loading and shipment of raw materials
AMAgreedDDate	Agreed date of delivery of raw materials
AMDDate	Date of delivery of raw materials
AMLoadingArea	Raw materials loading area
AMLoadingLongitude	Longitude of raw materials' loading area
AMLoadingLatitude	Latitude of raw materials' loading area
ProductionUnitArea	Production unit area
ProductionUnitLongitude	Longitude of production unit
ProductionUnitLatitude	Latitude of production unit
AMProductionUnitETA	Expected time of raw materials' order arrival at the producer

Table 2. Predictors Production – Wholesaler Warehouse

Column Name in Database	Description
AMQCStartDate	Start of the quality control of raw materials
FaultyAMIWDate	Date of identification and withdrawal of damaged end products
AMPutawayDate	Date of raw materials deposit for production
ProductionStartDate	Production start date
ProducerFPQCDate	Date of quality control of end products
FaultyFPIWDate	Date of identification and withdrawal of damaged end products
FPPutawayDate	Date of deposit of final products for delivery to the wholesaler's warehouse
WSLRPaymentCLRDate	Payment clearance date by a wholesaler
WSLROrderLSDate	Date of deposit of final products for transfer to the wholesaler's warehouse
WSLROrderDDate	Delivery date of the order
WSLRAgreedOrderDDate	The agreed delivery date of the order
WSLRWarehouseArea	Wholesale warehouse area
WSLRWarehouseLongitude	Longitude wholesale warehouse
WSLRWarehouseLatitude	Latitude wholesale warehouse
ProductionUnitWSLRETA	Expected arrival time of the end products at the wholesaler's warehouse

The developed service allows users to: (i) collect historical data on the lead time predictors of each supply chain echelon and send them anonymized to Entersoft's database, through the Entersoft Escalator Connector for real-time forecasting (ii) Pre preprocess the historical data of the predictors and dynamically predict lead time delays in each echelon of the SC (Figure 3 and Figure 4) (iii) Issue alerts in the cases where the forecasted lead time delays exceed the expected arrival times of the orders in each echelon of the SC, allowing users to take proactive steps to address this expected disruption. (iv) Evaluate the supplier's performance based on the percentages of orders arriving on time and thus seek, if required, new reliable suppliers.

The proposed methodological framework was applied to the supply chain of a Hellenic chemical manufacturing supply chain.

Table 3. Predictors Wholesaler Warehouse-Retailers

Column Name in Database	Description
WSLRFPQCDate	Date of quality control of the finished products in the warehouse
DamagedFPIWDate	Date of identification and withdrawal of damaged end products
PutawayDate	Date of deposit of end products
PickingDate	Date of collection of end products
SortingDate	Date of sorting of end products
PackingDate	Date of palletizing end products
RTLRPaymentCLRDate	Payment clearance date by retailer
RTLROrderLSDate	Order loading and shipping date
RTLROrderDDate	Date of delivery of the order to the retailer
RTLRAgreedOrderDDate	The agreed date of delivery of the order to the retailer
RTLRArea	Retailer area
RTLRLongitude	Longitude of the retailer area
RTLRLatitude	Latitude of the retailer area
WSLRRTLRETA	Estimated time of arrival of the order of end products at the retailer's area

5 Consumption of the Service

Following the historical data exchange of the predictors, the first step of data preprocessing involved handling the dataset's time stamps of order arrivals in each echelon of the examined supply chain. Missing values (NaN) of the variables were then filled. Specifically, for categorical variables, missing values were replaced by the most frequently occurring categorical data values as proposed by [20]. Regarding the continuous variables, the missing variables were replaced by the mean values of the variable's values of their column, using the mean substitution method, [21]. After filling in all missing values, categorical values were converted into arrays of binary (0-1) variables.

The transformed dataset was then split into a train and test set using the 80%/20% rule. The developed suite of machine learning models was then trained on the train set and tested on the test set by inputting the values of the independent variables from the test set.

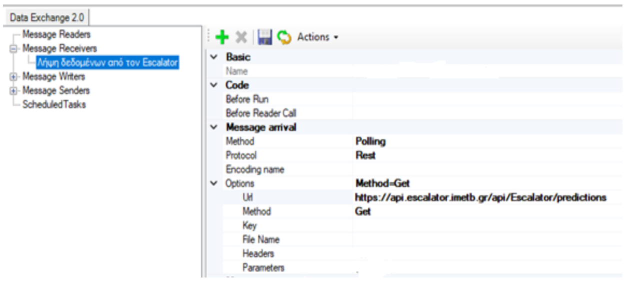

Fig. 3: Get forecasts from the service

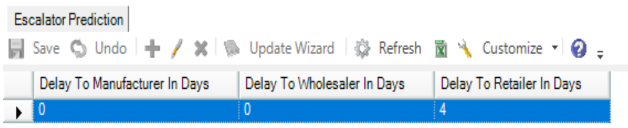

Fig. 4: Overview of forecasted data

The predictive accuracy of the suite's developed machine learning algorithms was evaluated using the mean absolute percentage error (MAPE) of the predictions at each echelon of the examined SC.

The results of the suite's implementation are summarized in Table 4 and Table 5 and reveal that Random Forest Regression demonstrates the highest predictive accuracy, with the lowest MAPE across all stages ranging from 0.5% to 7%, followed by Decision Tree Regression ranging from 0.8% to 12%.

Moreover, predictive accuracy improves as we move upstream of the supply chain, indicating improved performance in the later stages of the supply chain.

Table 4. Predicted lead time delays per ML model and stage of the SC (Days)

Model	Supply-Production	Production – Wholesaler Warehouse	Wholesaler Warehouse-Retailers
Decision Tree Regression	0.8	1.4	5.2
Random Forest Regression	0.5	0.9	4.1
Multiple Linear Regression	1.8	2.3	7.8

Table 5. MAPE per ML model and stage of the SC

Model	Supply-Production	Production – Wholesaler Warehouse	Wholesaler Warehouse-Retailers
Decision Tree Regression	12%	4%	0.8%
Random Forest Regression	7%	3%	0.5%
Multiple Linear Regression	17%	11%	7%

6 Summary and Conclusions

This study develops and employs a holistic methodological approach for predicting lead time delays in all echelons of an SC. The implementation of the developed methodology was evaluated in the real-world case of a Chemical manufacturing supply chain in Athens Greece. A critical synthesis of academic research efforts revealed critical delays in three echelons of a SC namely, from the raw material supplier to the manufacturer, from the manufacturer to the wholesaler's warehouse, and from the wholesaler's warehouse to the retailer. A suite of machine learning models was then employed to predict the lead time delays in each one of the SC echelons. The suite of ML models included Random Forest Regression, Decision Tree Regression, and Linear Regression.

The selection of Random Forest regression is based on its ability to process a large variety of input features and detect complex nonlinear relationships between variables. Moreover, and due to its ensemble method, which aggregates multiple decision trees, Random forests can make more accurate predictions, [22]. Similarly, the Decision tree's hierarchical structure allows it to model complex decision-making processes by breaking down a dataset into smaller subsets. This provides a straightforward visual representation of decision paths, making it easier to understand how variables affect outcomes, [23]. Finally, multiple linear regression constitutes a fast and simplistic tool for identifying the relationships between variables, while providing timely forecasts. The model can therefore serve as a baseline model for predicting lead times, providing initial insights that can be refined with more complex models, [24]. The results derived from the implementation of the suite of ML models in our real-world Hellenic chemical manufacturing SC reveal that Random Forest Regression demonstrates the highest predictive accuracy, with the lowest MAPE across all stages, followed by Decision Tree Regression. Moreover, predictive accuracy improves as we move upstream the supply chain, indicating improved performance in the later stages of the supply chain.

References:

[1] Simchi-Levi, D., Kaminsky, P., & Simchi-Levi, E. (2007). *Designing and managing the supply chain: Concepts, strategies, and case studies*. McGraw-Hill.

[2] Christopher, M. (2016). *Logistics and Supply Chain Management* (5th ed.). Pearson Education.

[3] Ivanov, D., Dolgui, A., & Sokolov, B. (2018). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. *International Journal of Production Research*, vol. 57(3), pp.829–846. https://doi.org/10.1080/00207543.2018.1488086.

[4] Gang Wang, Angappa Gunasekaran, Eric W.T. Ngai, Thanos Papadopoulos, Big data analytics in logistics and supply chain management: Certain investigations for research and applications, *International Journal of Production Economics*, Vol. 176, 2016, pp. 98-110. https://doi.org/10.1016/j.ijpe.2016.03.014.

[5] Gyulai, D., Pfeiffer, A., Nick, G., Gallina, V., Sihn, W., & Monostori, L. (2018). "Lead time prediction in a flow-shop environment with analytical and machine learning approaches." *IFAC PapersOnLine*, 51-11, pp.1029–1034. DOI: 10.1016/j.ifacol.2018.08.472.

[6] Bender, J., Trata, M., & Ovtcharova, J. (2022). "Benchmarking AutoML-Supported Lead Time Prediction." *Procedia Computer Science*, 200, pp.482-494. DOI: 10.1016/j.procs.2022.01.246.

[7] Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W., & Monostori, L. (2018). "Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer." *Procedia CIRP*, 72, pp.1051–1056. DOI: 10.1016/j.procir.2018.03.148.

[8] Barzizza, E., Biasetton, N., Disegna, M., Molena, A., & Salmaso, L. (2024). "An ensemble Machine Learning algorithm for Lead Time prediction." *Journal of Machine Intelligence and Data Science (JMIDS)*, Vol. 5. DOI: 10.11159/jmids.2024.003.

[9] Sethi, F. (2020). "Using Machine Learning Methods to Predict Order Lead Times." *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, Vol. 54, No. 3, pp 87-96.

[10] Alnahhal, M., Ahrens, D., & Salah, B. (2021)."Dynamic Lead-Time Forecasting Using Machine Learning in a Make-to-Order Supply Chain." *Applied Sciences*, 11, 10105. DOI: 10.3390/app112110105.

[11] Singh, S., & Soni, U. (2019). "Predicting Order Lead Time for Just in Time Production System Using Various Machine *Learning Algorithms: A Case Study*." DOI: 10.1109/ACCESS.2019.2941278.

[12] Aslan, A., Vasantha, G., El-Raoui, H., Quigley, J., Hanson, J., Corney, J., & Sherlock, A. (2023). "Hierarchical Ensemble Deep Learning for Data-Driven Lead Time Prediction." *The International Journal of Advanced Manufacturing Technology*, 128, 4169–4188. DOI: 10.1007/s00170-023-12123-4.

[13] Flores-Huamán, K.-J., Escudero-Santana, A., Muñoz-Díaz, M.-L., & Cortés, P. (2024). "Lead-Time Prediction in Wind Tower Manufacturing: A Machine Learning-Based Approach." *Mathematics*, 12, 2347. DOI: 10.3390/math12152347.

[14] Vakharia, Asoo & Yenipazarli, Arda. (2009). Managing Supply Chain Disruptions. Foundations and Trends® in *Technology Information and Operations Management*. 2(4): pp.243-325. http://dx.doi.org/10.1561/0200000012.

[15] Abdellatif, H., Alshibani, A. (2019). Major Factors Causing Delay in the Delivery of Manufacturing and Building Projects in Saudi Arabia. *Buildings*, 9(1), 1–15. https://doi.org/10.3390/buildings9010015.

[16] Macdonald, J.R., Corsi, T.M. (2013). Supply Chain Disruption Management: Severe Events, Recovery, and Performance. *Journal of Business Logistics*, 34(4), 270-288. https://doi.org/10.1111/jbl.12026.

[17] Mohamed, A., Coutry, N. (2015). Analysis of Lead Time Delays in Supply Chain: A Case Study. World Academy of Science, Engineering and Technology, Open Science Index 102, *International Journal of Economics and Management Engineering*, Vol:9, No:6, 2015.

[18] Lian, M. and Erichsen, PM. (2019). An empirical investigation into the sources of supply chain disruptions. University of Agder, Faculty of Engineering and Science, School of Business and Law, [Online]. https://uia.brage.unit.no/uia-xmlui/bitstream/handle/11250/2626866/Thesis.pdf?sequence=4T (Accessed Date: October 12, 2024).

[19] Ntabe, EN., LeBel, L., Munson, AD., Santa-Eulalia, LA. (2015). A systematic literature review of the supply chain operations reference (SCOR) model application with special attention to environmental issues. *Int J Prod Econ*. 2015; 169:310–32.

[20] Quintero, M., & LeBoulluec, A. (2018). Missing Data Imputation for Ordinal Data. *International Journal of Computer Applications*, 181(5), 10–6.

[21] Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Wiley.

[22] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324.

[23] Loh, W.-Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1), 14-23. https://doi.org/10.1002/widm.8.

[24] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning with applications* in R. Springer. https://doi.org/10.1007/978-1-4614-7138-7.

Contribution of individual authors to the creation of a scientific article (ghostwriting policy)

Ioannis Mallidis and Afroditi Stamelou - Conceptualization

Georgia Ayfantopoulou and Ioannis Mallidis - Methodology

Afroditi Stamelou and Georgia Ayfantopoulou - Formal analysis

Ioannis Mallidis and Elias Kanakis - Data curation and algorithms development

Ioannis Mallidis and Afroditi Stamelou – Writing and original draft preparation Georgia Ayfantopoulou supervision.

All authors have read and agreed to the published version of the manuscript.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

The project «Development of an intelligent Cloud application for the dynamic prediction of order fulfilment time at supply chain level» (Project code: KMP6-0279494) was implemented under the framework of the Action «Investment Plans of Innovation» of the Operational Program «Central Macedonia 2014 2020», that is co-funded by the European Regional Development Fund and Greece.

Conflict of Interest

The authors have no conflicts of interest to declare.