Contents lists available at GrowingScience

Decision Science Letters

homepage: www.GrowingScience.com/dsl

Adopting new ARIMA double layering technique in make-to-stock production policy

Ismail I. Almaraj[a,b*] and Mohammed S. Al Ghamdi[a]

[a]*Industrial and Systems Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia*
[b] *Interdisciplinary Center of Smart Mobility and Logistics, King Fahd University of Petroleum and Minerals, P.O. Box 5067, Dhahran 31261, Saudi Arabia*

C H R O N I C L E

Article history:
Received: October 10, 2023
Received in revised format: November 20, 2023
Accepted: February 4, 2024
Available online: February 4, 2024

Keywords:
Time series methods
ARIMA
ML-based models
Make-to-Stock production

A B S T R A C T

In supply chain management, the employed forecasting algorithm plays a very vital and crucial role in how a particular business performs well in the market. This study proposes a time series based forecasting model to help in designing an effective make-to-stock supply chain mechanism for tackling the large number of suppliers (suppliers diversion) and products (items diversion). The proposed technique extends the traditional ARIMA model to ARIMA Double Layering Technique (DLT) where the algorithm is implemented in two layers: first one is to obtain the number of orders level, and the other layer is implemented at the quantity level. The systematic implementation procedure of our approach is illustrated through a real case study. The accuracy evaluation and consistency of the results show that ARIMA -DLT provides superior forecasting when compared with ML-based models which provides insightful managerial views to other similar forecasting problems.

1. Introduction and literature review

1.1 Introduction

Demand forecasting in many supply chain contexts and application works as a major step for decisions related to operating, business, and production planning. In this regard, a variety of approaches are used to predict the demand that would maintain low inventory cost, high customer satisfaction, as well as effective production plan. However, the generation of accurate supply chain (SC) demand forecasts still represents a challenging task for both practitioners and researchers. The challenges of the SC forecasting demand are resulted from different aspects including order amplification or bullwhip effect (Jaipuria & Mahapatra, 2014), sales forecasts shows exogenous or nonlinear feature inputs (Gonçalves et al., 2021), uncertainties existing with demand, lead time and manufacturing process (Dhahri & Chabchoub, 2007). Other factors such as promotion, weather, market trends, and season are presented by Gilliland (2010). Consequently, a large number of studies have proposed various methods to reduce the volatility of demand and obtain better forecast results. Among these techniques, time series and data-driven forecasting has been the focus of extensive research studies across several fields of supply chain applications. The various models that assume demand follows times series pattern include autoregressive (AR), moving average (MA), exponential moving average (EMA), exponentially weighted moving average (EWMA), autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA). Although there is no single model that performs well for all different types of demand series, the appropriate forecasting model can be identified based on time series data features. Accordingly, Abolghasemi et al. (2020) scanned some papers discussing recommendations regarding appropriate forecasting approaches for time series which are recommended by researchers in different applications.

* Corresponding author.
E-mail address: almaraj@kfupm.edu.sa (I. I. Almaraj)

doi: 10.5267/dsl.2024.2.005

Exponential smoothing and ARIMA models are commonly used in different domains more than other time series forecasting models. When the description of trend and seasonality in the data are provided, then exponential smoothing is an effective approach. On the other hand, ARIMA models which were introduced by Box et al. (1970) exhibit unique features by involving the autocorrelations in the data as well. That is, it consists of an iterative three-stage process of model selection, parameter estimation and model checking. Despite its applicability limitation on linear modelling where it hardly captures the non-linearity inherent in time series data, it is still widely used time series models for forecasting purposes. Recently, other methods have been developed to tackle the forecast quality, namely machine learning (ML)-based models, and artificial neural network (ANN). Although ANN becomes popular in practice as it addresses the non-linearity issue of ARIMA, its practicality is worth taking if big data is available which may not be the case for many forecasting problems. The demand forecast quality in SC is always being treated as crucial, and the models used to improve this accuracy are still outstanding questions:

- ARIMA provides good accuracy in forecasting relatively stationary time series data, but how is it further improved when the future data values are neither stationary nor linearly dependent on the current and past data values?
- Compared with other similar approaches such as ML-based models, how does ARIMA perform? And at which approach is the optimal solution attained?
- Do ARIMA models provide a robust performance under a large number of suppliers (suppliers' diversion) and products (items diversion) in which the policy of a company is make-to-stock rather than make-to-order?

This paper is structured as follows. In the following section, a literature overview on supply chain demand forecasting models is provided. Section 2 provides a detailed description of the problem followed by section 3 which introduces the methodology used in the paper as well as the algorithm, notation, and model derivation. Section 4 provides a real case study demonstrating the model features and its output analysis. Finally, section 5 concludes the paper.

1.2 Literature Review:

When order amplification or bullwhip effect (BWE) was first recognized and analyzed by Forrester (1961), demand forecasting was identified as one of the major manifestations of this phenomena. A variety of forecasting techniques have been developed and extensively discussed to provide better market demand estimation in both upstream and downstream of the SC and management. The primary focus of this overview review is on the recent studies that employ ARIMA and ML methods. The forecasting performance using aforementioned techniques is always a major concern to the decision-makers. Ramos et al., (2015) study consumer retail sales forecasting using performance of state space and ARIMA models. The forecasting performance is demonstrated through a case study of retail sales, and the results show that based on different error measurements both methods perform well and are quite similar. In the environmental management field, Hikichi et al. (2017) utilize ARIMA models to forecast the number of ISO 14001 certifications in the Americas. The ARIMA models show a decreasing trend in the number of certifications in the Americas between 2016 and 2017. One obvious application of ARIMA models in water supply chain management is drought prediction. Karimi et al. (2019) focus on investigating and predicting the meteorological drought in Karkheh Basin using SPI index and ARIMA model. The findings are consistent with the previous studies which show a good agreement between the actual data and the forecasted data. Another domain of application is in the healthcare supply chain, particularly, blood supply chain in which an effective mechanism for tackling the COVID-19 pandemic is developed (Ayyildiz et al., 2021). In this study, the Multi-Layer Perceptron (MLP) approach is proposed and then compared with other traditional models such as ARIMA. The proposed approach shows better performance in forecasting the number of people recovered from COVID-19. Also, in the context of blood supply chain, Fanoodi et al. (2019) conduct a study to predict blood platelet demands based on ANNs and (ARIMA) models in order to reduce the uncertainty in the supply chain. The demand on a daily basis between 2013 and 2018 is collected from treatment centers and hospitals. The findings show that ANNs and ARIMA models are more accurate in predicting the uncertainties in demand than the baseline model adopted by Zahedan Blood Transfusion Center.

In the textile and clothing supply chain, Arora and Majumdar (2022) systematically reviews 312 research articles published recently between 2000-2020 where novel techniques such as Machine learning (ML) and soft computing (SC) significantly contributed in process and quality control, and automation of operations of the textile and clothing supply chain.

Time series methods including ARIMA in its traditional forms do not provide superior forecasting when compared with ML-based models and Neural Networks. This proposes a new research direction in which a hybrid ARIMA and ANNs model is utilized to improve the demand forecast quality. The latter model can recognize non-linear patterns and improve forecast accuracy. In this regard, Díaz-Robles et al. (2008) employ the hybrid approach combining ARIMA and ANNs to forecast particulate matter in urban areas. That is, the above model is applied to an area suffering from bad air quality resulting from residential wood burning. The hybrid model is able to capture 100% and 80% of alert and pre-emergency episodes, respectively. Another study using the hybrid ARIMA–ANN model is conducted by Koutroumanidis et al. (2009). Specifically, ARIMA models ANN and a hybrid model are used to predict the future selling prices of the fuelwood (from broadleaved and coniferous species) produced by Greek state forest farms. The findings show that the hybrid models provide superior forecasting results.

The above presented studies indicate that ARIMA and ML-based models side by side significantly improve the forecast results quality, and enable the decision makers to proceed with the production. To the best of our knowledge, the superior forecasting results attained using ARIMA Double Layering Technique (ADLT) when compared with ML-based models has not been shown. Thus, our contributions can be characterized by presenting the ARIMA- ADLT features as follows:

- In cases in which two mega factors play a major role in a business (i.e., Quantity & Orders frequency), ARIMA-ADLT capability to handle both of them simultaneously.
- ARIMA- DLT considers the complex relation between two mega factors and relate them as "Forecasted output rate"
- ARIMA- DLT enables the decision maker to optimize the variable of interest forecasting by conducting one back-step forecasting layer of highly correlated variables to the one of interest.
- ARIMA- DLT can improve the variable of interest forecasting by knowing the latest error factor found using the first-layer to forecast the latest known period, and then utilizing this error factor as a multiplier of the second forecasting layer.
- The above features are validated through a real practical case which reveal valuable managerial results.

2. Problem Definition

Make-to-order, which is known as a pull method of production, requires that a customer order beforehand. If a company moves toward shortening the lead times and makes the products available in stock in local distribution centers, then a make-to-order approach may result in lower efficiency and higher production costs. Therefore, a make-to-stock type of manufacturing process is adopted to overcome the above disadvantages. In make-to-stock, forecasted customer demand is a key to determine production planning and scheduling, and resources and capacity planning as well. Our main objective here is to transform the business model of a corporation from 'make to order' to 'make to stock', which absolutely includes more risks when dealing with large scale supply chain systems. In other words, such a transformation can boost expectations to several levels, but at the same time, it can ruin the efforts of several years. Along with this transformation, more effective forecasting techniques play a significant role in achieving the high quality production plan and makes this transformation successful. Prior to introducing ARIMA- DLT approach, next we briefly describe the challenges associated with applying the aforementioned approach.

2.1 Items diversity

Supplier companies have higher diversity. Generally speaking, when the number of item types being forecasted is numerous, the traditional time series and ML-based methods algorithms may not provide the degree of desirable accuracy. The underlying latter case is resulted from the high variation of item types as well as increasing variability of orders up the supply chain (the bullwhip effect). Hence, increasing diversity of items is a prime determinant of the forecasting method to be used, allowing the forecaster to trade off cost against the value of accuracy in choosing a technique.

2.2 Suppliers' diversity

The variety of products to be forecasted is not only the main challenge associated with accuracy, but it also depends on the number of suppliers and their characteristics. The difficulty of the matter increases if the same product type can be ordered from multiple suppliers. To illustrate the above serious pitfalls, consider three supplier characteristics which can be used to characterize any supplier namely: response time, material quality, and stock availability. If 100 suppliers are considered with only these three aspects, then a total of 5.15×10^{47} combinations result which prevent manufactures from accomplishing good forecasting and thereby realizing sound, and strategic planning.

2.3 Orders duplications

In special cases, some clients request a next particular order from the same supplier within the leading time of the first order. This indicates that delivery time is doubled as well as the transportation costs, making the business model infeasible financially and timely inefficient. In fact, this issue does not only affect the time and the transportation costs, but it also decreases the ordering frequency from the client to the supplier. To clarify, when the clients notice a delay on the order delivery and extra costs, they may engage more suppliers for their orders, in other words, the reliance on the main supplier may be weakened.

3. Proposed Methodology

To cope with these challenges, we propose a new approach of forecasting in which the traditional ARIMA model is extended to ARIMA Double Layering Technique (DLT). The algorithm is implemented in two layers: first one is to obtain the number of orders level, and the other layer is implemented at the quantity level which is generated from the first layer. Furthermore, this approach employs family level forecasting (aggregated forecasting), which is based on a predetermined group of suppliers that contributes to the bulk of the business (Fig. 1).

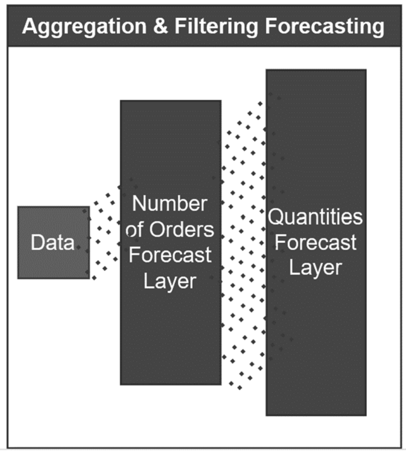

Fig. 1. The proposed study

3.1 DLT First layer

Basically, the first layer addresses the order duplication issue. It forecasts the number of orders of a specific product which are supposed to be received during the lead time. Therefore, if multiple orders are forecasted from a particular supplier, based on ARIMA DLT assumptions, all of them will be shipped together during the intended lead time rather than receiving them in individual shipments. Practically speaking, this leads to reducing the order frequency, and transportation costs. It is also assumed that during the lead time, no more orders can be requested from the same supplier until the previous orders are received. For example, if 7 orders are forecasted for a supplier s, then the customer (company) can not process the other orders before all 7 orders are received.

3.2 DLT Second layer

The second layer is concerned with forecasting the quantities that come from the orders received from first layer. Apart of make -to- stock policy, the significance of the second layer can be realized through the following aspects:

1- Stocking the products on the corporation warehouses to reduce the lead time and avoid any potential shortages.
2- Optimizing related operations such as the orders-quantity ratio per container based on total overall forecasted quantities

Notation

The parameters and decision variables are defined as follows:

- Indices

q: Quantity

o: Orders Number

- Sets

t: Set of time period $t \in T$

j: Set of quantity moving average parameters count $j \in Q_q$

i: Set of quantity autoregressive parameters count $i \in P_q$

k: Set of orders number autoregressive parameters count $k \in P_o$

m: Set of orders Number moving average parameters count $m \in Q_o$

- Parameters

$X_{t,q}$: Time series data of quantity q and at period t

Q_q: Total number of moving average parameters of quantity q

P_q: Total number of lag parameters of quantity q

D_q: Degree of differencing of quantity q

$\theta_{j,q}$: Moving average parameter of quantity q, and parameter number j

L_q: Lag Operator of quantity q

$\varepsilon_{t,q}$: Error terms of quantity q and at time t

$\varphi_{i,q}$: Autoregressive parameters of quantity q, and parameter i

δ_q: Model intercept of quantity q

$X_{t,o}$: time series data of order number o and at period t

Q_o: Total number of moving average parameters of order number o

P_o: Total number of lag parameters of order number o

D_o: Degree of differencing of order number o

$\theta_{k,o}$: Moving average parameter of order number o, and parameter k

L_o: Lag operator of order number o

$\varepsilon_{t,o}$: Error terms of order number o and at time t

$\varphi_{m,o}$: Autoregressive parameters of order number o, and parameter m

δ_o: Model intercept of order number o

- Variables

$X_{t+f,q}$: Forecasted quantity q in period t for future period f
$X_{t+f,o}$: Forecasted number of orders o in period t for future period f

Model Formulation

Definition: The general ARIMA model (P', Q) was developed by George Box and Gwilym Jenkins is as follows:

$$X_{t+f,q} = \frac{\left(1 + \sum_{j=1}^{Q_q} \theta_{j,q} L^j\right) \varepsilon_{t,q}}{\left(1 - \sum_{i=1}^{P'_q} \varphi_{i,q} L^i{}_q\right)} \tag{1}$$

The model has two parameters: P' which represents the order of the Auto Regressive part, and Q which represents the average of the Moving Average Part. Eq (1) presents the forecasted quantity in period t and for future period f where $L^i{}_q$ is the lag operator of quantity q, and $\varphi_{i,q}$ are the autoregressive part parameters. In addition the error term $\varepsilon_{t,q}$ is assumed to be independent and identically normally distributed variable. Also, the term $\left(1 - \sum_{i=1}^{P'_q} \varphi_{i,q} L^i{}_q\right)$ is assumed to have a unit root of multiplicity D_q. Correspondingly, it has a factor of $(1 - L^i)$, and can be expanded to:

$$\left(1 - \sum_{i=1}^{p'_q} \varphi_{i,q} L^i{}_q\right) = \left(1 - \sum_{i=1}^{P'_q - D_q} \varphi_{i,q} L^i_q\right)(1 - L)^{D_q} \tag{2}$$

Now, let P_q defined in the $ARIMA(P, D, Q)$, to be $P_q = P'_q - D_q$, in which D_q represents how many times we difference the data for the sake of making the pattern stationary. Therefore, the ARIMA model of the forecasted quantity is presented as follows:

$$X_{t+f,q} = \frac{\delta_q + (1 + \sum_{j=1}^{Q_q} \theta_{j,q} L^j)\varepsilon_{t,q}}{\left(1 - \sum_{i=1}^{P_q} \varphi_{i,q} L^i{}_q\right)(1 - L)^{D_q}} \tag{3}$$

The above derivation can also be applied to obtain the forecasted orders number model:

$$X_{t+f,o} = \frac{\delta_o + (1 + \sum_{k=1}^{Q_o} \theta_{k,q} L^k)\varepsilon_{t,o}}{\left(1 - \sum_{m=1}^{P_o} \varphi_{m,o} L^m{}_o\right)(1-L)^{D_o}} \tag{4}$$

Next, we develop our ARIMA -DLT as follows. We defined the rate of the two outputs, X_{t+f} provided in Eqs. (2-3) and can be presented in Eq. (5):

$$X_{t+f} = \frac{X_{t+f,q}}{X_{t+f,o}} \tag{5}$$

The corresponding model is then extended as given in Eq. (6).

$$X_{t+f} = \frac{(\delta_q + (1 + \sum_{j=1}^{Q_q} \theta_{j,q} L^j)\varepsilon_{t,q})(1-L)^{D_o}\left(1 - \sum_{m=1}^{P_o} \varphi_{m,o} L^m{}_o\right)}{\left(1 - \sum_{i=1}^{P_q} \varphi_{i,q} L^i{}_q\right)(1-L)^{D_q} * (\delta_o + (1 + \sum_{k=1}^{Q_o} \theta_{k,q} L^k)\varepsilon_{t,o})} \tag{6}$$

where X_t represents the average forecasted quantity per forecasted order. Our ARIMA -DLT model here is defined by dividing the ARIMA quantity model (Eq 3) by the ARIMA orders number model (Eq 4). This is because the quantity and number of orders can be related in a "rate" relation. The proposed algorithm can be presented as follows:

$Set\ X_{t,q}\ to\ X_{t,q}\ +\ 0.99999$
$Set\ Log\_X_{t,q}\ TO\ Log(X_{t,q})$
$Set\ X_{t,o}\ to\ X_{t,o}\ +\ 0.99999$
$Set\ Log\_X_{t,o}\ TO\ Log(X_{t,o})$
$\boldsymbol{Layer\ One}$:
$Input: ARIMA\ Model\ TO\ auto\_arima\left(Log_{X_{t,q}}, \max\_P = 9,, \max\_D = 2, \max\_Q = 9\right)^a$
$Return\ P, D, Q$
$then\ begin$
$Model\ TO\ ARIMA(Log\_X_{t,q}, order = (P, D, Q))$
$Set\ f\ TO\ Targeted\ Forecasting\ Period$
$For\ f \in \{1, \dots, F\}\ do$:
$PredictedValues_{t,q} \leftarrow Fitted\ values_f$
$Set\ X_{t+f,q} = e^{Fitted\ values_f} - 0.99999$
$Set\ RSS\ to \sum_{t+f=2}^{F} (X_{t+f,q} - X_{t,q})^2$
$\boldsymbol{Layer\ Two}$:
$Input: ARIMA\ Model\ TO\ auto\_arima\left(Log_{X_{t,o}}, \max\_P = 9,, \max\_D = 2, \max\_Q = 9\right)$
$Return\ P, D, Q$
$then\ begin$
$Model\ TO\ ARIMA(Log\_X_{t,o}, order = (P, D, Q))$
$Set\ f\ TO\ Targeted\ Forecasting\ Period$
$For\ f \in \{1, \dots, F\}\ do$:
$PredictedValues_{t,i} \leftarrow Fitted\ values_f$
$Set\ X_{t+f,o} = e^{Fitted\ values_f} - 0.99999$
$Set\ RSS\ to \sum_{t+f=2}^{F} (X_{t+f,o} - X_{t,o})^2$
Output: $X_{t+f}\ TO\ \frac{X_{t+f,q}}{X_{t+f,o}}$
End Procedure
[a] Auto_arima is a function that returns the optimized ARIMA parameters by minimizing the RSS value.

3.3 Case Study: Al Haitam Company

A case study is conducted at Al Haitam Company, which is an industrial supplier for products and services in Saudi Arabia. The headquarter of Al-Haitam is in Al Khobar city with an outreach to a discerning customer base spread across the kingdom. In this study 350 suppliers and more than 5300 item types are considered. Moreover, a big percentage of their clients order the same items multiple times in a period within the same leading time, which would significantly increase the ordering and transportation costs. The increasing variety and changing data streams would bring a real complexity to the current status.

4. Computational Results & Analysis

We start by data processing and cleaning in which the data is extracted, and the period is formatted as a data framework (i.e., in months). It is followed by setting the data in index oriented format after deleting the old period columns. Sample data translated into usable information is provided in Table 1.

Table 1
Sample of Log Qty data

Month	Log Quantity
12/1/2017	7.91
1/1/2018	6.6
2/1/2018	2.19
3/1/2018	3.65
4/1/2018	1.31
5/1/2018	1.27
6/1/2018	1.86
7/1/2018	1.03
8/1/2018	5.95

Table 2
Sample of Log Orders

Month	Log Orders #
12/1/2017	1.48
1/1/2018	1.4
2/1/2018	1.61
3/1/2018	1.86
4/1/2018	1.79
5/1/2018	1.6
6/1/2018	1.08
7/1/2018	1.85
8/1/2018	1.46

Notice that we are considering all items here. The Data shown in Table 1 represents the quantity ordered based on months. We applied the log-transformation to the data to minimize the data variation (to stabilize the variance). We also applied the same transformation for order number data in Table 2. One crucial aspect in the time series data is the stationarity. In this regard, the quantity data is tested through the plot provided in Fig. 2. The test results do not show any unique behavior or trend.

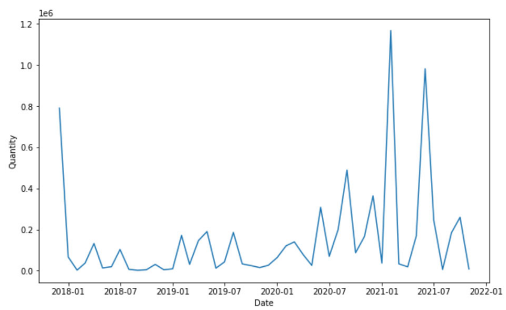

Fig. 2. The plot of log-transformed data

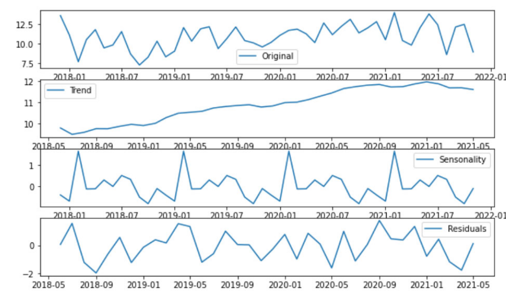

Fig. 3. Data decomposition

We can apply the log function to facilitate the data elements decomposition to uncover the hidden trend and seasonality of our data. Fig. 3 presents our data elements (Original, Trend, Seasonality and Residuals). Two methods in the literature are commonly used to test the stationarity of the data: (1) visualizing the trend by data parameters, (2) applying Dickey-Fuller (DF) test. The first method is a graphical technique where the stationarity of the data can be verified through finding the moving average and moving standard deviation. Then, the two plots along with the original data plot are provided for analysis, (Fig. 4). It can be concluded from Fig. 4 that the data is not stationary.

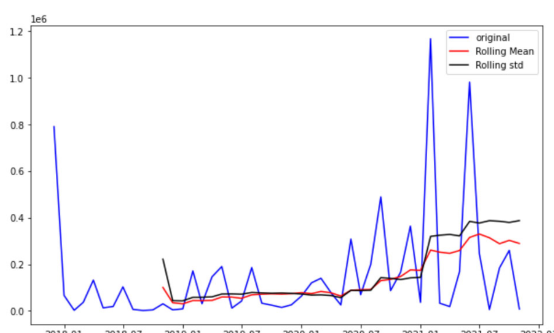

Fig. 4. Rolling mean and STD

Table 3
Dickey-Fuller test results on original data

Attribute	Value
Test Stats	-1.8
P-Value	0.38
# of Lags Used	8
# of Obs.	39
Critical Value (1%)	-3.61
Critical Value (5%)	-2.94
Critical Value (10%)	-2.6

The second method is a statistical method and based on Dickey and Fuller (1979) work which created and popularized (DF) test .In this test, the null hypothesis states that the AR model has a unit root, which implies that the data series is not stable. While the alternative hypothesis states that the data shows a stationary trend. However, it may vary based on the testing method used. The results of the DF test with attributes generated are shown in Table 3. The less p-value the more stationary is the data. However, to come up with a definite conclusion, one may ensure that all critical values should be greater than the test statistic. Accordingly, we can conclude that the data is not stationary. To address the non-stationarity of the data, we shift it for one period, and then subtract it from itself. This process is so-called data differencing, and it determines the parameter D. For example, if the data is differenced once and test result shows stationarity, then the value of D is 1. In other words, the differencing value depends on how many times the data is differenced. Considering $D = 1$, the transformed data becomes as in Fig. 5.

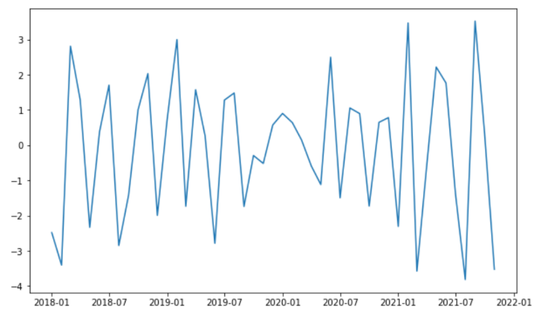

Fig. 5. Log subtracted shifting data behavior By appling the DF test again

Table 4
Dickey-Fuller test results on Log subtracted shifting data

Attribute	Value
Test Stats	-1.166
P-Value	0.00005
# of Lags Used	1
# of Obs.	4.5
Critical Value (1%)	-3.58
Critical Value (5%)	-2.93
Critical Value (10%)	-2.6

We can notice that this method is far the best, p-value is almost zero and the critical values are significantly greater than the test statistics. The above analysis of stationarity concludes that the data is a different stationary process, or integrated when it is differenced one. Thus, the original data is transformed into a weakly-stationary process by applying the integration of order $D = 1$. The other two ARIMA-DLT model parameters, namely, P and Q are determined next. Initially, the underlying parameters can be found from the autocorrelation and partial autocorrelation graphs, which are provided in Fig. 6 and Fig. 7. The upper and lower dashed limits represent the 95% confidence intervals. Finding the initial value of P can be done by identifying where the autocorrelation line provided in (Fig. 6) exactly cuts the upper confidence interval limit. The same technique can be used to find the initial value of Q in which the partial autocorrelation line cuts the upper confidence interval limit. It should be noted that both P and Q are supposed to be integers and thus they are rounded up, Fig. 6 and Fig. 7, respectively.

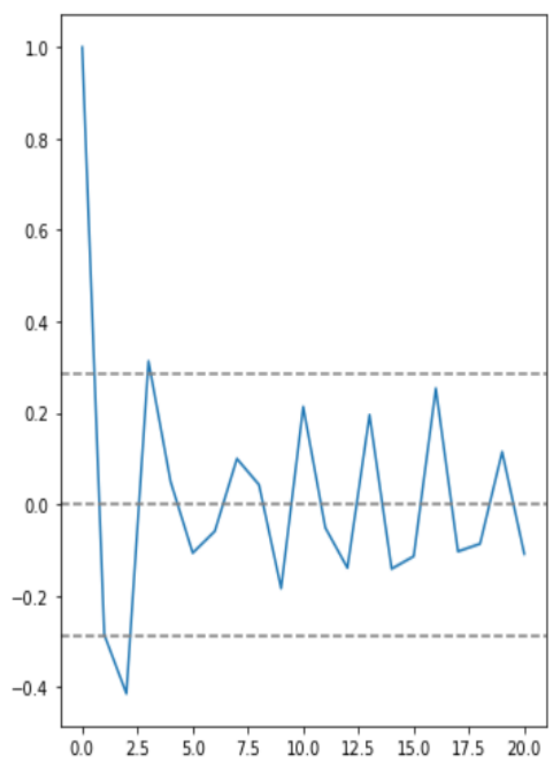

Fig. 6. Autocorrelation Graph

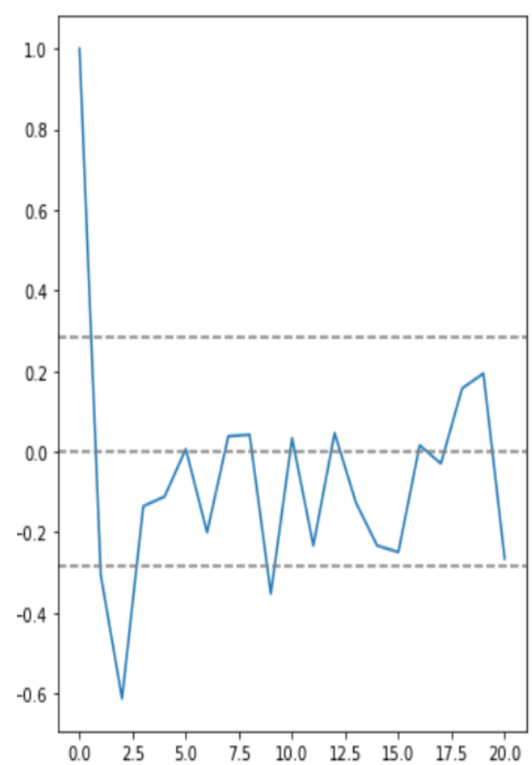

Fig. 7. Partial Autocorrelation Graph

After the model is fitted and evaluated by employing Residuals Sum of Squares metric (RSS), we tweak the P and Q values for the sake of optimizing the forecasting accuracy. The final determined parameters are $P = 4$, $D = 1$ and $Q = 5$. The transformed (stationary) data version is shown in Fig. 8.

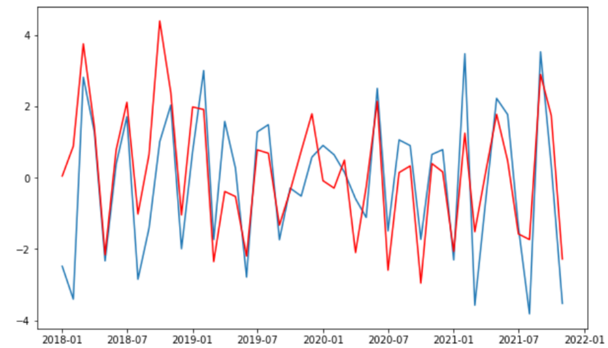

Fig. 8. Resulted model on the transformed (stationary) data

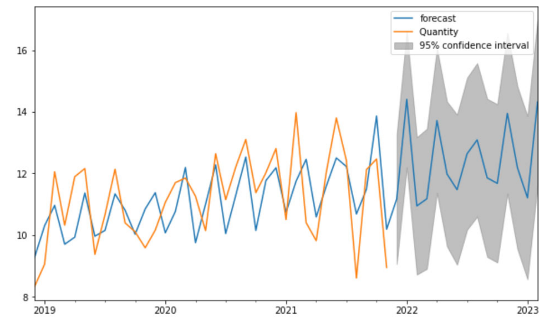

Fig. 9. Forecasted quantity integrated with 95% confidence intervals

The red and blue lines in Fig. 8 indicate the forecasted and original quantities, respectively. In addition, the forecasted quantity integrated with 95% confidence intervals are shown in Fig. 9. It should be noted that the forecasting period here is 12 months (i.e., 2022 to 2023). One may argue that the above forecasted results are not up to the desirable accuracy level due to the following model data complexities:

1. Hundreds of suppliers are considered, and each has a unique behavior.
2. Thousands of items are considered and divided in multiple groups, each with a unique trend.
3. Multiple orders for a particular product are requested from the same supplier in less than the lead time.

The first two issues are resolved as the DLT technique decomposes the forecasted data into two filters: supplier filter and item group filter. Using these two filters, the data chosen to be forecasted will be specified into a specific supplier and item group. Also, the above third issue should not exist along with our DLT which can be used to find the Forecasted Output Rate. The order number can be related to the quantities as Quantity Per Order, which would help in optimizing ordering time and cost.

4.1 Comparative analysis

In this section, the proposed ARIMA-DLT based forecasting methodology is compared with different regular ML-forecasting methods to demonstrate its robustness. The results are compared according to two KPIs: : accuracy and consistency (Section *IV.3.4*). For this purpose, the input features considered in the study are related to the following aspects:

- Item
- Supplier
- Ordered Quantity (Target)
- UOM (unit type)
- PriceUSD
- Ship-to Location

The data type associated with the above features is shown in Table 5.

Table 5
Data types associated with data labels

Data Label	Data Type
Item	Object
Supplier	Object
Ordered Quantity	Float
PriceUSD	Float
UOM	Object
Ship-to Location	Object

As shown from Table 5, the majority of the data types is (Object), and one-hot encoding is utilized in this regard to convert them into dummy variables which means that each unique value inside them will be converted to an independent feature with a value 0 or 1. However, one-hot encoding brings a challenge related to computational complexity in which the original data frame is converted into a much bigger matrix. The matter gets worse when dealing with 352 unique suppliers and 2166 different items. Practically speaking, this issue is resolved by conducting several meetings with the management. After doing so, we identified that only 5 suppliers are considered to be the most important contributors, and 5 family groups which can be aggregated together. All the other suppliers and items are labeled as (other suppliers) and (other items), respectively.

4.2 Fitting Supervised ML Models

Various algorithms and computation techniques are used in supervised machine learning processes. Below are some of the most commonly used learning methods:

- Lasso Regression
- Ridge Regression
- Elastic Net Regression
- Linear Regression
- SVR
- Decision Tree Regressor
- Random Forest Regressor
- Gradient Boosting Regressor

The brief explanation of each one of them is presented in the next sub-sections, and there are no assumptions conflicting with each other's. In addition, we highlight some considerations before fitting:

1. To discover overfitting and inconsistency of the prediction accuracies, Stratified K Folds methods are used through Cross Validation technique. This approach takes a portion of the data as a testing sample from multiple locations to ensure that the accuracy value is representative.
2. In case that the accuracy values are found below the target value, or they are inconsistent, some improving techniques are applied to the data. Those techniques include Data Augmentation, Data Scaling, and finally Data Transformation.

The next sub-sections give a brief explanation followed by the results obtained from the methods employed in this study.

4.2.1 Lasso Regression:

Lasso regression is a type of regularization which uses shrinkage. In this approach, data values are shrunk towards a central point known as the mean. It is a popular penalization linear regression model where the regularization term is added to the error function to control the over-fitting. This form of regression is ideal when it is preferred to automate elements of the model selection process, such as variable selection and parameter removal. That is, it does feature selection, and is useful when there are unrelated input columns. The point of using this model in our case is to automate variable elimination and feature selection as we are having a big number of dummy variables. The Lasso regression mathematical model is linear, and uses L1 regularization technique where its cost function is as follows:

$$\sum_{i=1}^{M}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{M}(y_i - \sum_{j=0}^{p} w_j \times x_{ij})^2 \quad (7)$$

4.2.2 Ridge Regression

Ridge Regression is similar to Lasso regression, but it uses L2 norm regularization instead of L1 norm. By limiting the coefficient and maintaining all the variables, L2 regularization can cope with multicollinearity (independent variables are highly correlated) difficulties. It also shrinks the estimated values, and is useful when there are correlated input columns. L2 regression may be used to determine the relevance of predictors and punish unimportant predictors depending on the results. The model is applied to our case because we assumed that our case has many large parameters of about the same value. The Ridge regression mathematical model is linear, and its cost function is as follows:

$$\sum_{i=1}^{M}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{M}(y_i - \sum_{j=0}^{p} w_j \times x_{ij})^2 + \lambda \sum_{j=0}^{p} w_j^2 \quad (8)$$

4.2.3 Elastic Net Regression

Elastic Net arose from criticism of Lasso, whose variable selection might be overly reliant on data, making it unstable. The best of both approaches may be acquired by combining the penalties of ridge regression and Lasso (i.e., its algorithm uses a weighted combination of L1 and L2 regularization). The goal of Elastic Net is to reduce the following loss function:

$$\frac{\sum_{i=1}^{n}(y_i - x_i^J \hat{\beta})^2}{2n} + \lambda \left(\frac{1-\alpha}{2} \sum_{j=1}^{m} \hat{\beta}_j^2 + \alpha \sum_{j=1}^{m} \left| \hat{\beta}_j \right| \right) \quad (9)$$

4.2.4 Linear Regression

Linear Regression is one in which the input variables, x and the single output variable, y are assumed to have a linear relationship. The dependent variable can be determined using a linear combination of the input variables. The procedure is known as simple linear regression when there is only one input variable, x. When there are several input variables, the procedure is referred to as multiple linear regression in statistics literature. It can be expressed simply as:

$$Y_i = f(X_i, \beta) + e_i \tag{10}$$

The model is used because it is a long-established statistical procedure and its properties are well understood and can be trained very quickly.

4.2.5 SVR

The supervised learning technique Support Vector Regression (SVR) is used to predict discrete values. SVMs and Support Vector Regression are both based on the same proposition. SVR's main objective is to locate the best-fitting line which is a hyperplane with the greatest number of points.

The SVR, unlike other regression models, aims to fit the best line inside a threshold value, rather than minimizing the error between the real and projected value. The distance between the hyperplane and the boundary line is the threshold value. We applied SVR to our case because its computational complexity does not depend on the dimensionality of the input space as we are having a big number of dummy variables.

4.2.6 Decision Tree Regressor

The purpose of this technique is to develop a model that predicts the value of a target variable, and the decision tree solves the problem by using the tree representation, where the leaf node corresponds to a class label and characteristics are represented on the internal node of the tree. It is used in our case because it is very fast, straightforward and effective

4.2.7 Random Forest Regressor

Random Forest Regression is a supervised learning approach for a regression that uses the ensemble learning method. The ensemble learning approach combines predictions from several machine learning algorithms to get a more accurate forecast than a single model. The structure of a Random Forest is seen in Fig. 10 .It should be noted that the trees go in a straight line with no contact. During training, a Random Forest develops many decision trees. The average of the outputs refers to the forecast of the classes.

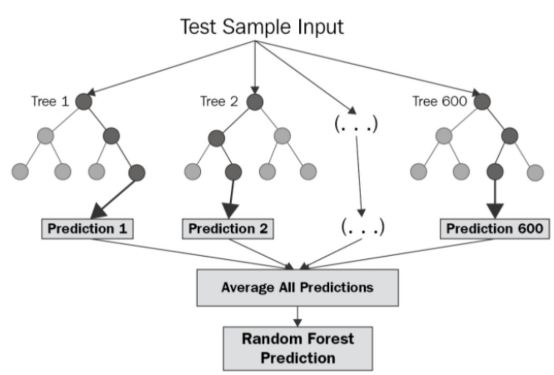

Fig.10. Random Forest Regressor illustration graph

4.2.8 Gradient Boosting Regressor

Gradient Boosting Regressor a is a machine learning ensemble strategy problem that generates output from a group of weak learners, particularly decision trees. Its premise is based on creating multiple weak models and then combining them to get better performance as a whole. The model is used in our case because it grants better accuracy, higher flexibility in terms of loss functions and parameters used.

4.3 Improving the Data Inputs for the ML Models

To boost the output accuracy and make it more consistent, Three techniques are applied: Data Augmentation, Scaling, and Transforming.

In Data Augmentation, the number of data rows is increased based on uniform distribution as well as the natural variability among the features themselves. For continuous variables, the following approach is applied:

$$FSTD_i = STD(F_i) \quad \rightarrow \quad standard\ deviation\ of\ the\ feature\ i$$

$$NewF_i = \quad F_i + Uniform(1\ , FSTD_i) \quad \rightarrow \quad incrementing\ feature\ i\ based\ on\ its\ STD$$

By adopting this techniques, the number of records is duplicated, which helps the models learn more about the specifications inside the data. For the case of dummy variables, they are replicated instead because they lack the natural variability (i.e., 0 and 1). In the other two techniques, namely, Scaling and Transforming, the Standard Scaler technique are applied. As a result, the continuous variables range resides between [-1 1] and is assumed normally distributed.

4.4 Results and Discussion

The summary of the ML model accuracies compared with ARIMA-DLT are presented in Table 6. As noted from Table 6, 5 trials are done and represent the cross validation method where the data is divided into multiple groups, and then applied to the R^2 metric. The R^2 metric here returns a value from 0 to 1, unless the data does not apply minimum assumptions of the model, in which in this case it returns negative numbers. Two criteria are used to evaluate each technique: accuracy and consistency (variability between obtained forecasted results). Table 6 presents these evaluations scores. Note that we define the accuracy which is above 80% to be excellent, from 50% to 79% is somewhat good, from 20% to 49% is somewhat poor and the rest range is considered to be extremely poor. Regarding consistency evaluation, standard deviation between trials is used to determine its score. That is, if the standard deviation is less than 0.05 is considered to be consistent, from 0.051 to 0.1 is considered to average, while from 0.11 and above is considered to be extremely inconsistent. As indicated by Table 6, ARIMA -DLT performs well according to the results evaluation and consistency, and thus it provides superior forecasting when compared with ML-based models.

Table 6
The summary of the ML model accuracies compared with ARIMA-DLT

Model	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Evaluation	Variability Between Obtained Forecasted Results
Lasso Regression	-4.79	-1.97	-5.17	-1.64	-1.34	Extremely poor	Extremely Inconsistent
Ridge Regression	0.28	0.33	0.23	0.41	0.26	Somewhat Poor	Average
Elastic Net Reg.	-4.8	-1.97	-5.17	-1.64	-1.34	Extremely poor	Extremely Inconsistent
Linear Regression	0.28	0.33	0.22	0.41	0.26	Somewhat Poor	Average
SVM Regressor	0.28	0.27	0.35	0.32	0.19	Somewhat Poor	Average
Decision Tree Reg.	0.41	0.87	0.52	0.6	0.33	Excellent	Average
Random Forest Reg.	0.46	0.77	0.59	0.5	0.32	Somewhat good	Average
Gradient Boost Reg.	0.48	0.71	0.54	0.6	0.3	somewhat good	Average
ARIMA DLT	0.81	0.81	0.81	0.81	0.81	Excellent	Excellent

5. Conclusion

This study introduces a new ARIMA model to the literature which offers the opportunity to enhance the forecasting accuracy in the supply chain. It has a different point of view than other forecasting studies employing ARIMA models where the algorithm is implemented in two layers. The first layer aims to obtain the number of orders level, while the second layer is implemented at the quantity level, and no similar study has been found in the literature to the best of the authors' knowledge.The practicality of the proposed forecasting approach is validated through a real case study of Al Haitam Company. The company is transforming from make-to-order to make -to-stock policy which may raise three challenges: (1) Item diversity which represents the variations between item types themselves, (2) Supplier diversity which represents the variations on the supplier level, and finally (3) orders duplication which indicates the duplication of clients orders that can be avoided to reduce times and costs. The comparative study with different ML-forecasting methods shows that ARIMA -DLT provides superior forecasting results with respect to the evaluation and consistency criteria.

Future work may integrate the approach adopted in this paper with other forecasting methods to obtain even better forecasting results. This includes different ANN-based forecasting models. The proposed forecasting approach could eventually help the decision-makers to significantly improve the forecasting accuracy in different domains including supply chain, energy, healthcare, and so on.

Acknowledgements

The authors acknowledge the support provided by King Fahd University of Petroleum and Minerals and Al Haitham Company for conducting this research.

References

Abolghasemi, M., Beh, E., Tarr, G., & Gerlach, R. (2020). Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion. *Computers and Industrial Engineering, 142*(July 2019), 106380. https://doi.org/10.1016/j.cie.2020.106380

Arora, S., & Majumdar, A. (2022). Machine Learning and Soft Computing Applications in Textile and Clothing Supply Chain: Bibliometric and Network Analyses to Delineate Future Research Agenda. *Expert Systems with Applications, 200*(June 2021), 117000. https://doi.org/10.1016/j.eswa.2022.117000

Ayyildiz, E., Erdogan, M., & Taskin, A. (2021). Forecasting COVID-19 recovered cases with Artificial Neural Networks to enable designing an effective blood supply chain. *Computers in Biology and Medicine, 139*(August), 105029. https://doi.org/10.1016/j.compbiomed.2021.105029

Box, G., Jenkins, G., & Reinsel, G. (2008). *Time Series Analysis.* 4th ed., Wiley.

Dhahri, I., & Chabchoub, H. (2007). Nonlinear goal programming models quantifying the bullwhip effect in supply chain based on ARIMA parameters. *European Journal of Operational Research, 177*(3), 1800–1810. https://doi.org/10.1016/j.ejor.2005.10.065

Díaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., & Moncada-Herrera, J. A. (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment, 42*(35), 8331–8340. https://doi.org/10.1016/j.atmosenv.2008.07.020

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), 427-431. https://doi.org/10.2307/2286348.

Fanoodi, B., Malmir, B., & Jahantigh, F. F. (2019). Reducing demand uncertainty in the platelet supply chain through artificial neural networks and ARIMA models. *Computers in Biology and Medicine, 113*(January), 103415. https://doi.org/10.1016/j.compbiomed.2019.103415

Gilliland, M. (2010). *The Business Forecasting Deal: Exposing Myths, Eliminating Bad Practices, Providing Practical Solutions* (Vol. 27). John Wiley & Sons.

Gonçalves, J. N. C., Cortez, P., Carvalho, M. S., & Frazão, N. M. (2021). A multivariate approach for multi-step demand forecasting in assembly industries: Empirical evidence from an automotive supply chain. *Decision Support Systems, 142*(March 2020). https://doi.org/10.1016/j.dss.2020.113452

Hikichi, S. E., Salgado, E. G., & Beijo, L. A. (2017). Forecasting number of ISO 14001 certifications in the Americas using ARIMA models. *Journal of Cleaner Production, 147*, 242–253. https://doi.org/10.1016/j.jclepro.2017.01.084

Jaipuria, S., & Mahapatra, S. S. (2014). An improved demand forecasting method to reduce bullwhip effect in supply chains. *Expert Systems with Applications, 41*(5), 2395–2408. https://doi.org/10.1016/j.eswa.2013.09.038

Karimi, M., Melesse, A. M., Khosravi, K., Mamuye, M., & Zhang, J. (2019). Analysis and prediction of meteorological drought using SPI index and ARIMA model in the Karkheh River Basin, Iran. In Extreme Hydrology and Climate Variability: Monitoring, Modelling, Adaptation and Mitigation. Elsevier Inc. https://doi.org/10.1016/B978-0-12-815998-9.00026-9

Koutroumanidis, T., Ioannou, K., & Arabatzis, G. (2009). Predicting fuelwood prices in Greece with the use of ARIMA models, artificial neural networks and a hybrid ARIMA-ANN model. *Energy Policy, 37*(9), 3627–3634. https://doi.org/10.1016/j.enpol.2009.04.024

Ramos, P., Santos, N., & Rebelo, R. (2015). Performance of state space and ARIMA models for consumer retail sales forecasting. *Robotics and Computer-Integrated Manufacturing, 34*, 151–163. https://doi.org/10.1016/j.rcim.2014.12.015