International Journal of Computer Applications (0975 – 8887)
Volume 183 – No. 52, February 2022

Supply Chain Forecasting Employing Auto-Regressive Integrated Moving Average Model

Jagriti Singh
Department of Computer Science and Engineering
SDBCT, Indore
India

Rahul Maheshwari
Department of Computer Science and Engineering
SDBCT, Indore
India

ABSTRACT
Supply chain Management has remained an active area of research due to its widespread applications in several domains of manufacturing business. Off late, supply chain forecasting has emerged as a very effective tool which is useful in streamlining production, logistics and manpower thereby critically affecting the profit margin. Supply chain forecasting predominantly deals with forecasting of demands of the products and goods based on previously available data. The estimation of demands directly impacts the production, which in turn influences the supply. The current supply has a critical impact on the future demands. Due to the enormity of data to be analyzed, supply chain forecasting is prone to errors in forecasting. Off late, machine learning based algorithms have been in the forefront for supply chain forecasting. This work presents an Auto-Regressive Integrated Moving Average Machine Learning Model for Supply Chain Forecasting. It is shown that the proposed work attains higher accuracy of forecasting compared to existing technique.

General Terms
Supply Chain Management, Machine Learning, ARIMA

Keywords
Machine Learning, Auto-Regressive Integrated Moving Average (ARIMA), Supply Chain Forecasting, Mean Absolute Percentage Error, Accuracy.

1. INTRODUCTION
Global Markets have encountered a lot of volatility in the last decade due to the following reasons:
1) Trade Wars
2) Formation of cartels
3) Outbreaks of new diseases such as Ebola and Covid-
4) Global economic slowdown etc.

Some specific products and services account for a large portion for the Gross Domestic Product (GDP) of countries. There is a large diversity in such products and services. For example, Switzerland and Denmark rely heavily in dairy based products, while China relies heavily on Electronic and industrial manufacturing [1]-[2]. Generally, the major export depends on the geographic conditions, manpower and natural resources of the country. Supply chain management plays a pivotal role for such industries in streamlining the processes and deciding the profits. Supply chain management can be defined as the management of the flow of goods and services including all processes which are intertwined with the transformation of raw materials into final products. The different domains affected by supply chain management are [10]:

1) Inventory Management
2) Warehousing and distribution
3) Logistics
4) Procurement
5) Revenue Management
6) Order Management
7) Revenue Generation and Management
8) Information Technology (IT) related to businesses.

Fig 1: Applications of Supply Chain Management

2. SUPPLY CHAIN FORECASTING
Supply chain forecasting can be defined as the prediction of demand metrics based on the previous demands and associated variables. It's a critical domain of supply chain management. Thus apart from increasing profits and streamlining businesses, supply chain management and supply chain forecasting can also have deep impacts on employment, food security, peace and political conditions in a country [5]-[6].Supply chain forecasting tries to find patterns in previously existing data and forecast future demands [7]. The sales or demands are generally modelled as a function of time and associated variables given by:

$$Demand = function(time, associated\ variables) \quad (1)$$

The associated variables can be:
- Current Economic Situation of the Country of export
- Global Economic condition
- Political Relations among countries
- Supply from other countries etc.

Accurate predictions or estimates need to be made considering a pervasive set of associated parameters.The genuine requirement for the demand or sales forecasting and the challenges faced in this domain are illustrated as under:
- Large scale manufacturing involves monetary investments.
- The associated variables exhibit severe volatility. It varies with several parameters such a GDP, Inflation,

market needs etc.

- The data size is generally highly uncorrelated and extremely random.
- It is challenging to achieve high accuracy.

3. MACHINE LEARING AND NEURAL NETWORKS FOR SUPPLY CHAIN FORECASTING

Machine learning has been extensively used for time series forecasting problems such as supply chain forecasting. Machine Learning, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. Other advantages include:

1. Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
2. Self-Organization: It can create its own organization or representation of the information it receives during learning time.
3. Real Time Operation: It computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability

One of the most effective machine learning approaches has been the Auto-Regressive Integrated Moving Average (ARIMA) model. In statistics and econometrics, and especially in time series analysis, an “Auto-Regressive Integrated Moving Average” (ARIMA) is a broader model than the Auto-Regressive Moving Average (ARMA). These models are used in time series to better understand the model or predict the future. These models are used where the data is non-stationary.In an autoregressive integrated moving average model, the future value of a variable is assumed to be a linear function of several past observations and random errors. That is, the underlying process that generate the time series has the form:

$$y_t = \theta_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_p y_{t-p} + \cdots . \theta_q \varepsilon_{t-q} \quad (2)$$

Here,
y_t is the value of the output variable at time ‘t’
ε is the prediction error
θ *and* φ are called the model parameters
p *and* q are called the orders of the model

The ARIMA model assumes that the prediction error has the following properties:

1) Expectation $\mu = 0$
2) Constant variance, $v = \sigma^2$, where σ is the standard deviation

Objective Function:
Typically for time series analysis, the errors ε_t can attain both positive and negative values. Hence, errors can cancel out in case normal average is computed. Hence, the squared average often termed as mean square error is computed as the objective function. The mean square error is computed as:

$$mse = \frac{e_1^2 + \cdots \ldots e_n^2}{n} \quad (3)$$

The error is computed in each iteration as:

$$e_t = predicted\ value_t - actual\ value_t \quad (4)$$

4. PROPOSED METHOD

One of the most complex tasks related to demand forecasting is the analysis of extremely random and volatile data. Thus the proposed approach employs data cleaning techniques such as the wavelet transform.
The mathematical formulation for the wavelet transform is given by the scaling and shifting approach of the wavelet function [7].
The scaling, shifting dependence can be defined as:

$$W\varphi(Sc, Sh) = ₩[\mathrm{x}, \mathrm{t}] \quad (5)$$

Here,
x is the space variable
t is the time variable
₩ is the transform
sc is the scaling factor
sh is the shifting factor

The wavelet transform is an effective tool for removal of local disturbances. Pharmaceutical demands show extremely random behavior and local disturbances. Hence conventional Fourier methods do not render good results for highly fluctuating data sets. The major advantage of the wavelet transform is the fact that it is capable of handling fluctuating natured data and also local disturbances. The DWT can be defined as [8]:

$$\mathrm{W}\Phi\ (\mathrm{Jo, k}) = \frac{1}{\sqrt{M}} \sum_n S(n).\Phi(n)_{jo'k} \quad (6)$$

The next critical aspect is the design of the ARIMA model and training it in a manner to attain least number of iterations for the cost function or performance function to stabilize.The important aspect about ARIMA is the fact that it repeatedly feeds the errors in every iteration to the network till the errors become constant or the maximum number of allowable iterations are over. This can be mathematically given by:

$if\ PF \neq constant$
$for\ (k = 1, k \leq k_{max}\ = constant, k = k + 1)$
{
$$W_{k+1} = f(X_k, W_k, e_k)$$
}
else
{
$W_{k+1} = W_k \&\&\ training\ stops$
}
Here,
X_k is the input to the kth iteration
W_k is the weight to the kth iteration
W_{k+1} is the weight to the (k+1)st iteration
e_k is the error to the kth iteration
k is the iteration number
PF is the performance function deciding the end of training
k_{max} is the maximum number of iterations

Thus if the error is within tolerance, which is generally not feasible to find beforehand in time series data, the training is stopped if the performance function (which can be the training error) becomes constant for multiple iterations or the maximum number of iterations are over. Now there are various ways in which the error can be minimized. However, the steepest fall of the error with respect to weights is envisaged. The concept is explained in figure 2.

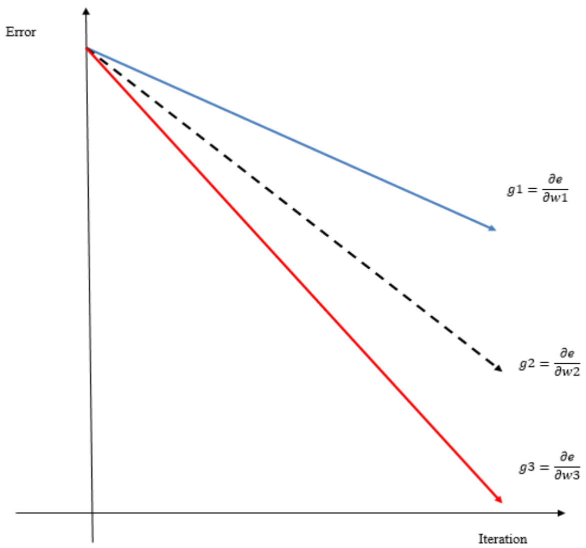

Fig 2: Minimizing the Performance Function as a function of iterations.

It can be observed from figure 2 that although the error in training keeps plummeting in all the three cases of gradient descent, the gradient 3 or g3 attains the maximum negative descent resulting in the quickest training among all the approaches and hence the least time complexity. This would be inferred from the number of iterations which are required to stop training. Thus the number of iterations would be a function of the gradient with which the error falls.Auto regression tries to find out the relation among a variable and its delayed version. Mathematically:

$$C_{Auto} = \int_0^T x(k)x(t-k)dk \qquad (7)$$

Here,
C_{Auto} is the auto correlation of the variable
x is the variable
t is the time delay
k is the variable of integration
The auto correlation tells us about the changing nature of a variable.

1) Moving Average:

As data is fed to a neural network for pattern recognition, the weights keep updating. However, it has been found that in case of time series problems, the latest data sample have the maximum impact on the latest output. Hence it is logical to calculate a moving average of latest (previous) data and apply it to the neural network. This is also called a moving average. Mathematically,

$$I_k = X_{1,k}, Mean(X)_{k,k-n}, Y_k \qquad (8)$$

Here,
I_k is the kth input sample to the neural network
$X_{1,k}$ are the data samples from the first to the kth sample
$Mean(X)_{k,k-n}$ is the mean of the data samples form k-n to k, i.e. it is a moving average depending on the value of k
Y_k is the target.

The training is stopped based on the mean square error or mse given by:

$$mse = \frac{\sum_{i=1}^{n} e_i^2}{n} \qquad (9)$$

The final computation of the performance metric is the mean absolute percentage error given by:

$$MAPE = \frac{100}{M}\sum_{i=1}^{N}\frac{E-E_i|}{i} \qquad (10)$$

Here,
n is the number of errors
i is the iteration number
E is the actual value
E_i is the predicted value

5. RESULTS AND DISCUSSIONS

The dataset used in this study is extracted from Kaggle with 12 training features. The data is collected from 2015-2020 over a tenure of 5 years. The system used has a configuration of RAM of 8GB, Intel Core i7 processor with a clock speed of 2.4GHz. The experimental results obtained in this work are cited here. The system has been designed on Matlab. The data set used is extracted from Kaggle [1]. The variable used in the dataset govern the actual demand of the forecasted value. The parameters used in the dataset are:

1) Date
2) Month
3) Generic Name
4) Brand Name
5) Medical Use
6) Shipped to Country
7) Sold in
8) Delivery Plant
9) External Agent Assigned/Not-Assigned
10) Price
11) Revenue
12) Commission
13) Demand/Sales

Disease Medical Use	Invoice date	Company code	ship to country	Sold-to party Country Full Name	Delivery Plant	External Agent	Sales quantity	Price TC /Kg	Revenue	External commissions	Months
Psychosis; depression	09-11-2015	5704	Mexico	Mexico	8370	Not assigne	450	$ 204.0	$ 91,800.0	$ -	Nov
Pain	30-06-2015	5704	Thailand	Thailand	8370	Not assigne	374	$ 472.0	$1,76,528.0	$ -	Jun
Pain	05-12-2015	5704	Vietnam	Japan	8370	Assigned	138	$ 472.0	$ 65,136.0	$ 4,186.0	Dec
Pain	09-05-2015	5704	Australia	Australia	8370	Not assigne	227	$ 472.0	$1,07,144.0	$ -	May
Pain	09-07-2015	5704	Singapore	Singapore	8370	Not assigne	394	$ 472.0	$1,85,968.0	$ -	Jul
Pain	25-02-2015	5704	Singapore	Singapore	8370	Not assigne	261	$ 472.0	$1,23,192.0	$ -	Feb
Pain	09-04-2015	5704	Thailand	Thailand	8370	Assigned	129	$ 472.0	$ 60,888.0	$ 817.0	Apr
Pain	30-03-2015	5704	Thailand	Thailand	8370	Assigned	310	$ 472.0	$1,46,320.0	$ 5,063.3	Mar
Pain	09-02-2015	5704	Thailand	Thailand	8370	Assigned	111	$ 472.0	$ 52,392.0	$ 1,184.0	Feb
Pain	21-04-2015	5704	Egypt	Egypt	8370	Not assigne	109	$ 472.0	$ 51,448.0	$ -	Apr
Neuropathic pain	16-11-2015	5887	India	India	8095	Not assigne	46	$ 500.0	$ 23,000.0	$ -	Nov
Neuropathic pain	01-09-2015	5887	India	India	8095	Not assigne	376	$ 500.0	$1,88,000.0	$ -	Sep
Neuropathic pain	09-07-2015	5887	India	India	8095	Not assigne	135	$ 500.0	$ 67,500.0	$ -	Jul
Neuropathic pain	22-07-2015	5887	India	India	8095	Not assigne	379	$ 500.0	$1,89,500.0	$ -	Jul
Neuropathic pain	12-10-2015	5887	India	India	8095	Not assigne	104	$ 500.0	$ 52,000.0	$ -	Oct
Neuropathic pain	23-02-2015	5887	India	India	8095	Assigned	468	$ 500.0	$2,34,000.0	$ 10,296.0	Feb
Neuropathic pain	04-05-2015	5887	Thailand	Thailand	8095	Not assigne	365	$ 500.0	$1,82,500.0	$ -	May
Neuropathic pain	12-06-2015	5887	Thailand	Thailand	8095	Not assigne	119	$ 500.0	$ 59,500.0	$ -	Jun
Neuropathic pain	13-05-2015	5887	Thailand	Thailand	8095	Not assigne	127	$ 500.0	$ 63,500.0	$ -	May
Neuropathic pain	27-07-2015	5887	India	India	8095	Not assigne	254	$ 500.0	$1,27,000.0	$ -	Jul
Neuropathic pain	09-12-2015	5887	India	India	8095	Not assigne	171	$ 500.0	$ 85,500.0	$ -	Dec
Neuropathic pain	09-02-2015	5887	South Afric	South Afric	8095	Assigned	250	$ 500.0	$1,25,000.0	$ 7,833.3	Feb
Neuropathic pain	20-07-2015	5887	India	India	8095	Not assigne	240	$ 500.0	$1,20,000.0	$ -	Jul

Fig 3: Dataset Used in the experiment

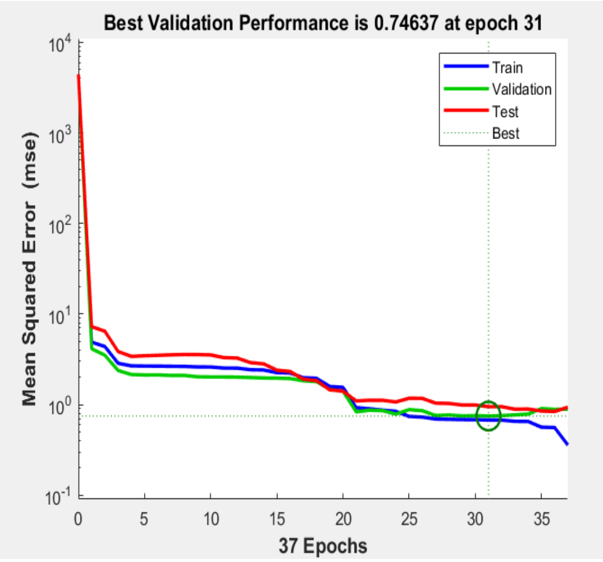

Fig 4:Variation in Performance function

Figure 4 depicts the changes or variations in the cost function which in this case is chosen as the mean squared error. It can be observed that the mean squared error attains a convergence at 31 iterations.

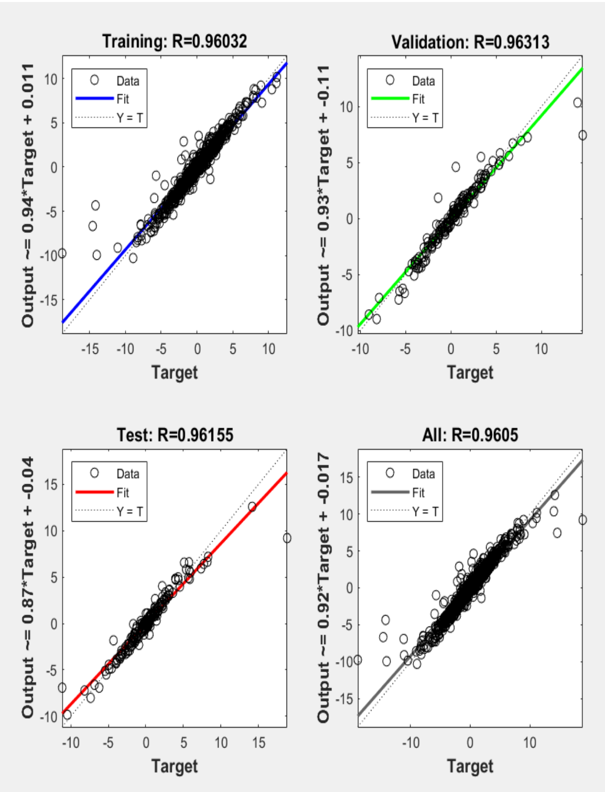

Fig 5: Forecasting Results

The regression analysis of the proposed model has been conducted for training, testing, validation and overall cases. The proposed system attains an average regression of 0.9605.A high value of the regression can be attributed to the accuracy of the system in predicting the sales or demand values based on the independent features of the data.

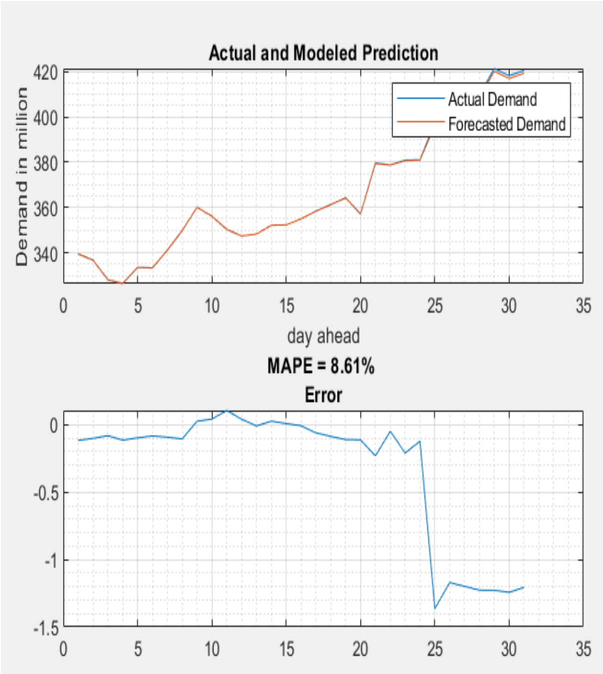

Fig 6: Forecasting Results

The performance function that decides the culmination of training is the mean squared error in this case given mathematically by equation 9. Figure 6 depicts the actual and forecasted values of the demand (in millions). The mean absolute percentage error (MAPE) has also been shown in table 1. The mean absolute percentage error (MAPE) is computed using equation (10). The variation of the error (actual value) is depicted in figure 6. To alleviate the possibility of cancelling of the positive and negative polarities of the errors, the mean absolute percentage error is computed which in turn renders the percentage accuracy.The training parameters are depicted in table 1.

Table 1. Training Parameters

S.No.	Parameter	Value
1.	Cost Function	mse
2.	Mse at convergence	0.74637
3.	Regression (Training)	0.96032
4.	Regression (Testing)	0.96155
5.	Regression (Validation)	0.96313
6.	Regression (Overall)	0.9605

Table 2. Summary of Results

Parameter	Value
Data Source	https://www.kaggle.com/mnshsh07/pharmaceutical-business-dataset
Machine Learning Model	ARIMA
Training Algorithm	Gradient Descent
Iterations	37
Regression	0.9605 (OVERALL)
MAPE (Proposed Work)	8.65% (Wavelet+ARIMA)
MAPE of Previous Work [1]	14% (Neural Network+Fuzzy)

A summary of the results obtained is tabulated in table 2. The summary. It can be observed that the proposed work attains a higher value of accuracy compared to existing work even at low/moderate number of iterations. In this case, a dual approach comprising of the DWT (acting as a filter) and the ARIMA model acting as the pattern analysis achieves convergence of training in 37 iterations thereby indicating that the time complexity of the proposed work is relatively low making it suitable even for resource constrained applications.

Future directions of work can be employing hybrid neural models or deep learning. Moreover, clustering algorithms can also be used to facilitate the training process.

6. CONCLUSION

This paper presents a mechanism for supply chain forecasting. The ARIMA model along with discrete wavelet transform is used to implement machine learning. The two fold approach has been used keeping in mind the random fluctuations in the data which inhibits the pattern recognition mechanism for ARIMA. The wavelet transform has been used as the smoothening filter. The ARIMA model has been specifically chosen keeping in mind the volatility of the pharmaceutical industry and the necessity of any training model to be fed with recent means or averages of the data samples. The mean absolute percentage error has been chosen as the metric for performance evaluation along with the regression values. The proposed system has been shown to attain higher accuracy compared to existing systems.

The future directions of the work can be employing clustering algorithms such as K-means and employing other pre-processing tools such as the maximum overlap discrete wavelet transform (MODWT) so as to improve upon the existingwork.

7. REFERENCES

[1] Oglu A.R.B., Kizi I.I.T. (2021) A Method for Forecasting the Demand for Pharmaceutical Products in a Distributed Pharmacy Network Based on an Integrated Approach Using Fuzzy Logic and Neural Networks. In: Kahraman C., Cevik Onar S., Oztaysi B., Sari I., Cebi S., Tolga A. (eds) Intelligent and Fuzzy Techniques: Smart and Innovative Solutions. INFUS 2020. Advances in Intelligent Systems and Computing, vol 1197. Springer.

[2] Amalnick, M.S., Habibifar, N., Hamid, M. et al (2019) An intelligent algorithm for final product demand forecasting in pharmaceutical units. Int J Syst Assur Eng Manag 11, 481–493 Springer.

[3] Zadeh, Lotfi A. (1994):Fuzzy Logic, Neural Networks, and Soft Computing», Communications of the ACM, vol. 37, 77-84.

[4] Aliev R.A., Bijan Fazlollahi, Aliev R.R. (2004): Soft Computing and its Applications in Business and Economics, Springer-Verlag, Berlin Heidelberg.

[5] How to Use ABC Analysis for Inventory Management (and the Added Value of XYZ Analysis). https://www.eazystock.com/blog/how-to-use-abc-analysis-for-inventory-management, last accessed on 26.03.2020.

[6] Takagi T., Sugeno M. (1985) Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man and Cybernetics(15), pp.116—132.

[7] Beale M.H., Hagan M.T., Demuth H.B.: Neural Network Toolbox. User's Guide, Natick: Math Works, Inc.(2014).

[8] Balashirin A.R., Akbar I.K. (2019) Neural Network Modeling and Estimation of the Effectiveness of the Financing Policy Impact on the Socio-Economic Development of the SocioEducational System. In: Aliev R., Kacprzyk J., Pedrycz W., Jamshidi M., Sadikoglu F. (eds) 13th International Conference on Theory and Application of Fuzzy Systems and Soft Computing,ICAFS-2018,Advances in Intelligent Systems and Computing, vol. 896, pp. 754-759, Springer, Cham.

[9] Balashirin A.R., Tarlan I.I. (2018), Application of Neural Networks for Segmentation of Catering Services Market Within the Overall System of Consumer Market on the Model of Restaurant Business with the Aim to Advance the Efficiency of Marketing Policy. In: Aliev R., Kacprzyk J., Pedrycz W., Jamshidi M., Sadikoglu F. (eds) 13th International Conference on Theory and Application of Fuzzy Systems and Soft Computing,ICAFS-2018,Advances in Intelligent Systems and Computing, vol. 896, pp. 905-913, Springer, Cham.

[10] Fildes R, Goodwin G, (2021) Stability in the inefficient use of forecasting systems: A case study in a supply chain company, International Journal of Forecasting, Elsevier Volume 37, issue 2, 2021, pp. 1031-1046.

[11] Chopra S, Meindl P, Kalra DV, 2016 "Supply Chain Management: Strategy, Planning and Operation", 6th Edition, Pearson Publications.

[12] Bowersox DJ, Closs D, Cooper MB, Bowersox JC, 2018 "Supply Chain Logistics Management", 4th Edition, McGraw Hill Publications.

[13] Taghipour A, 2020 "Demand Forecasting and Order Planning in Supply Chains and Humanitarian Logistics", IGI Global Publications.

[14] Szidarovszky, F., Gershon, E.M., Lucien, D.,1986. Techniques for Multi-objective Decision Making in Systems Management. 1st ed. vol. 2. Elsevier, West Lafayette, UN

[15] Hillermeier, C., 2001. Generalized homotopy approach to multi-objective optimization. Journal of Optimization Theory and Applications 110, 557–583.