Research Journal of Engineering Technology and Management (ISSN: 2582-0028), Volume 02, Issue 04, December-2019
Available at www.rjetm.in/

Warehouse Sales Forecasting using Ensemble Techniques

SONAR PRASHANT SHANTILAL[1], MANISH GANGIL[2]
[1]M.Tech Student, [2]Professor
[1,2] Department of Mechanical Engineering
Sri Satya Sai College of Engineering, RKDF University, Bhopal.

[2]gangilmanish@gmail.com

* Corresponding Author: Manish Gangil Manuscript Received: Manuscript Accepted:

Abstract

Supply Chain Management (SCM) is one of the new concepts put into practice in the commercial sector in Bangladesh since the late 1990s. At the beginning, Multinational Companies (MNCs) incorporated the supply chain into their structures, then other private conglomerates and local people defended these concepts. From the beginning, the main functions of SCM were the management of purchases and purchases, but subsequently SCM took the integrated form i.e. consists of sourcing, materials management, production support and sales management. Given the highly competitive market scenario, supply chain management is becoming the most important functional area of the business. Direct, indirect supply chain management processes and source services (as a raw material) to the end customer as a final product. Demand forecasting models is designed to estimate the quantities of more products that customers will buy in the future. As inventory and product numbers increase, it is extremely difficult to predict demand. Most demand forecasting models are based on a single classifier or a simple combination of these models. To improve the accuracy of demand forecasting, this research proposed a methodology for combining various forecasting models to help the set of vector machines support demand forecasting. The proposed methodology is tested against single classifiers and classifier ensemble models using a real dataset of alcohol warehouse. Experiments indicate that the proposed methodology outperforms better by approx. 4% better than existing techniques.

Keywords: Warehouse, Sale Forecasting, Supply Chain Management, Ensemble Support Vector Machine.

1 INTRODUCTION

Modern companies face conflicting problems in a difficult environment. Successful companies are more adaptable and immediately follow the updated or revised corporate governance concepts. Apply these detailed techniques in the features. Supply Chain Management (SCM) is one of the new concepts implemented in Bangladesh since the late 1990s. First, multinationals have integrated supply chain management. Providing in their facilities, then other people and local conglomerates have adopted the concepts. These include purchases, materials management, production support and sales management. In the highly competitive market scenario, supply chain management becomes the most important functional area of the company. Supply Chain Management manages direct, indirect and source services (as basic material) for the end customer as a final product [1].

A supply chain is a network of supply chain partners, such as suppliers, manufacturers, distributors, dealers, couriers, etc. Forecasts are by far the beginning of all supply chain management activities that launch all other supply chain management actions. However, forecasts play an important role both inside and outside the company [2].

Forecasting is the main driver for planning and decision making at the supply chain and enterprise management level. Companies are truly professional and rely heavily on the actual numerical value of forecasts to make key decisions such as capacity building, resource allocation, expansion and integration, upstream and downstream, and so on. The exploratory study focuses on the following objectives:

- Understand the practice, management and application of forecasts in the three growing Bangladesh sectors, such as the life-saving industry, the retail chain and consumer products.
- Limit and resolve demand forecasts.
- Propose a forecast management model in supply chain management.

Supply chain is defined as a set of entities directly involved in the activities associated with the upstream and downstream flows of products, services, finances, and/or information from a source to a customer [2]. Supply chains can be categorized into three groups such as Direct Supply Chain, Extended Supply Chain and Ultimate Supply Chain [3]. In our problem, we focus on Direct Supply Chain which contains some manufacturers, warehouses and customers. In this type of supply chain, products of manufacturer are transported to warehouses and customers reach these products through warehouses. Considering all given definitions above, supply chain management (SCM) can be thought as a process which deals with the total flow of materials from suppliers through end users [4].

There are various sub-processes of supply chain management which are quite complicated and challenging such as demand forecasting. Demand forecasting can be summarized as an estimation of a supply chain constituent's (such as warehouse, end sale point etc.) expected sales during a specified future period. Forecasting demand correctly for different constituents provides planning all processes of supply chain effectively [5]. For instance, accurate demand forecasting prevents redundant shipping charges or storage costs. Thus, forecasting the demand of warehouses is an important task and it forms the motivation of our work. In this research, we study forecasting the demand of warehouses with low error rate problem [6].

The aim of the work is to propose a time series prediction model using artificial neural networks technique. Demonstrate the effectiveness of the proposed approach based on real data from a Alcohol warehouse. To make the output results more effective and to determine the future requirements of the business that will examine the requirements of last year.

II. RELATED WORK

HaixiaSang's et al. (2018) [1] this article presents a simulation approach to analyze the inventory problem of the rental housing chain. Unlike most products, the rental unit is a "traffic" type product and the inventory problem is complicated. This paper proposes a systematic and flexible process that effectively provides managers with critical decision-making tools to help them understand and validate inventory problems in the supply chain of rented accommodation. The proposed method takes inventory factors into account, such as the forecasting method, the time period, the initial inventory and the supply indicator. Furthermore, the procedure is applied to a real supply chain of rented accommodation to confirm its effectiveness. Furthermore, the proposed methodology has proved to be practical and effective in helping managers make ongoing decisions.

ChristophFlöthmann et al. (2018) [2] this study aims to improve understanding of the skills needs of supply chain planners and analysts (SCP & As) and to identify the different personal preferences of HR managers in relation to the skills profile. Candidates. Design / Methodology / Approach A total of 243 recruitment-led supply chain leaders participated in an election-based, adaptive experiment to discover the relative importance of six skill attributes, namely analytical and analytical skills. Problem solving, interpersonal skills, Supply Chain Management (SCM) skills and experience in the sector. Results Knowledge and analysis of supply chain management and problem-solving skills have been identified as the most important skills. They are three times larger than general management skills. Based on a convergent analysis of groups and sets, two types of recruitment managers have been identified. The first group is characterized by a strong preference for candidates with extensive skills in supply chain management. On the other hand, members of the second group prefer candidates with a more balanced skill profile. Originality / value The authors' conclusions help companies to better adapt to the person and the workplace, a determining factor for employee performance and satisfaction.

Shanika L. Wickramasuriya et al. (2018) [3] Major time series collections often have aggregation restrictions due to product groups or geographic areas. Forecasts for the most disaggregated series are usually needed to accurately add aggregated series forecasts, a constraint we call "consistency". Forecasting is the process of adjusting forecasts to make them consistent. We demonstrate that this matrix cannot be estimated in practice due to identifiable conditions. We propose a new approach to forecasting that includes information from a complete covariance matrix of forecasting errors to produce a set of consistent predictions. Our approach minimizes the average quadratic error of consistent predictions between time series and assumes fairness. The minimization problem has a closed form solution. We make this solution scalable by offering a computationally demanding yield. We evaluate the performance of the proposed method with respect to alternative methods using a series of simulation projects that take into account the different characteristics of the time series acquired. This is followed by the empirical application with data on Australian domestic tourism. The results show that the proposed method works well with artificial and real data.

Gokhan MertYagli et al. (2018) [4] Forecasts for photovoltaic production play an important role in the functioning of the electricity grid. The forecast is required in different geographical and temporal scales, which can be modeled as hierarchies. In a geographical hierarchy, global forecasts for the region can be obtained directly by forecasting the regional time series or by aggregating the individual forecasts for the sub-regions. This results in a problem known as total inconsistency, as it is probable that both sets of predictions differ due to the uncertainties of the model. Therefore, the practice is not optimal. A statistically optimal aggregation, called reconciliation, provides consistent predictions. The comparison helps system operators to have greater foresight at the regional level, eventually leading to effective system planning. This document focuses on improving the accuracy of correspondence. Furthermore, the impact of detailed and aggregated forecasts on final concerted forecasts was analyzed. In California, 318 simulated photovoltaic plants were used to establish a geographical hierarchy. Aggregate predictions based on the numerical prediction of time are obtained with model output statistics and models of artificial neural networks. In the agreed forecasts, significant improvements are observed without the use of exogenous information.

Lindsay R. Berry et al. (2018) [5] we introduce a new Bayesian methodology for consumer sales forecasts. Focusing on multi-step forecasting of daily sales of many supermarket items, we adapt dynamic counter models to predict each customer's transactions and introduce new cascaded dynamic models to predict the number of items per transaction. These transaction and sales models can include predictors of time, season, price, action, randomness, and other points of sale for individual items. Sequential Bayesian analysis involves rapid parallel filtering of decoupled element groups and can be adapted to elements that may have very different properties. A multi-scale approach allows the exchange of information on elements with correlated models over time to improve forecasts while maintaining the scalability of many elements. A motivating case study of multi-period sales forecasts and multi-stage pre-shopping in supermarkets provides examples that demonstrate better accuracy of forecasts in different metrics and illustrate the benefits of complete probabilistic models for assessing and comparing the accuracy of forecasts. Keywords: Bayesian forecast; decoupling / torque; dynamic binary waterfall; Calibration prediction; intermittent question; multi-scale forecast; Forecast of rare events; Revenue per transaction Forecast of supermarket sales

Christoph Flöthmann et al. (2018) [7] this exploratory study analyzes the careers of 307 supply chain managers (SCEs). Motivated by career theory, our ideas create new insights into the educational background and careers that lead to SCE positions. Based on an optimal matching analysis, we can distinguish six career models for SCEs. They differ in terms of previous work experience, training and time needed to reach a position of leadership. By characterizing the antecedents and career paths of CEMs, we show that Supply Chain Management (SCM) is a truly cross-functional profession. Our findings suggest that the former employee responsibility appears to be a more important recruitment criterion than the full experience of supply chain management. While 56% of managers had ex-employee responsibilities, only 12% of the accumulated careers were actually spent in the SCM function.

Islek et al. (2017) [8] discussed demand forecasting is the process of constructing forecasting models to estimate the quantities of several products that customers will purchase in the future. When the warehouse and the number of products grow, forecasting the demand becomes dramatically hard. Most of the demand forecasting models rely on a single classifier or a simple combination of these models. In order to improve demand forecasting accuracy, author investigated several different classifiers such as MLP, Bayesian Network, Linear Regression and SVM analyzing their accuracy and performance. Moreover, we also studied some classifier combination techniques by approaching from demand forecasting perspective. In this paper, we propose a methodology to combine various forecasting models using neural networks rather for supporting demand forecasting. The proposed methodology is tested against single classifiers and classifier ensemble models using a real dataset.

III.METHODOLOGY

Demand and sales forecasts are one of the key characteristics of manufacturers, resellers and retailers. By balancing supply and demand, they reduce inventory and excess inventory and improve profitability. If the producer wants to satisfy the overestimated demand, overproduction leads to further storage linking the surplus stock. On the other hand, an undervalued demand leads to unrealized orders, lost sales opportunities and reduced service levels. Both scenarios lead to an inefficient supply chain. Therefore, accurately predicting the demands of participants in the supply chain is a challenge.

The ability to predict the future from past data is an important tool for supporting individual and organizational decisions. TSF (Time Series Forecasting) aims in particular to predict the behavior of complex systems looking only at the past trends of the same phenomenon. The forecast is an integral part of supply chain management. Traditional forecasting methods are subject to severe limitations that affect forecasting errors. Artificial Intelligence have proven to be useful techniques for forecasting demand due to their ability to support non-linear data.

Artificial intelligence forecasting techniques have been receiving much attention lately in order to solve problems that are hardly solved by the use of traditional methods. They have been cited to have the ability to learn like humans, by accumulating knowledge through repetitive learning activities. Therefore, the objective of the research is to propose new forecasting techniques via the artificial approaches to manage demand in a fluctuating environment. In this study, a comparative analysis based on neural techniques is presented for customer demands in a multi-level supply chain structure. The artificial techniques used in this study are explained below.

In the current scenario, an algorithm is proposed which provide a way to predict the warehouse and retail sales forecasting in supply chain management.

Figure 1 shows the overall architecture for prediction/forecast the demand or warehouse sales. The proposed work is designed for forecasting the sales/demand using ensemble support vector machine. The result is performed over short as well as long term forecasting. Following diagram describes flow of demand forecasting System.

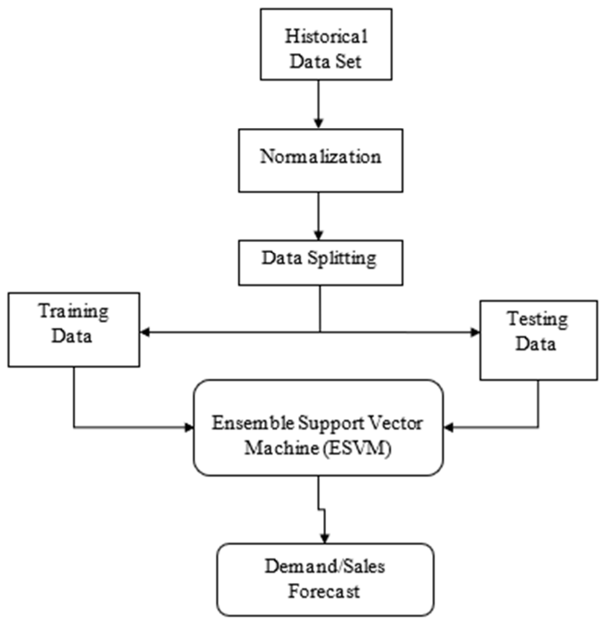

Figure 1: Flow Diagram of Demand Forecasting

Input: D {Historical Warehouse Sales Data};
Output: Sales/Demand Forecast Weekly as well as Monthly
Step1: Normalization of data, D
Step2: For each entity in D, do
Find feature vector (V) from D
Step 3: For each V do
Predict using Ensemble Support Vector Machine (ESVM)
Step 4: Determine the error
Find Performance Parameters i.e. RMSE, MAPE
end for

A. Historical Data Collection

This Dataset is for Warehouse and Retail Sales monthly data from January, 2017 to March, 2018. Total Number of Data is 128355. Warehouse and Retail Sales dataset contains a list of monthly sales of Alcoholic drink and non-alcoholic drinks [30].

B. Ensemble Support Vector Machine

The SVM has two drawbacks. First, since it is originally a model for the binary-class classi6cation, we should use a combination of SVMs for the multi-class classi6cation. Methods for combining SVMs for the multi-class classi6cation is also there but the performance does not seem to improve as much as in the binary classi6cation. Second, since learning of the SVM is a very time consuming for a large scale of data, we should use some approximate algorithms.

Using the approximate algorithms can reduce the computation time, but degrade the classi6cation performance. To overcome the above drawbacks, we propose to use the SVM ensemble. We expect that the SVM ensemble can improve the classi6cation performance greatly than using a single SVM by the following fact. Each individual SVM has been trained independently from the randomly chosen training samples and the correctly classi6ed area in the space of data samples of each SVM becomes limited to a certain area. We can imagine that a combination of several SVMs will expand the correctly classi6ed area incrementally. This implies the improvement of classi6cation performance by using the SVM ensemble. Likewise, we also expect that the SVM ensemble will improve the classi6cation performance in case of the multi-class classi6cation.

The SVM has been known to show a good generalization performance and is easy to learn exact parameters for the global optimum [2]. Because of these advantages, their ensemble may not be considered as a method for improving the classi6cation performance greatly. However, since the practical SVM has been implemented using the approximated algorithms in order to reduce the computation complexity of time and space, a single SVM may not learn exact parameters for the global optimum. Sometimes, the support vectors obtained from the learning is not sufficient to classify all unknown test examples completely. So, we cannot guarantee that a single SVM always provides the global optimal classi6cation performance over all test examples.

To overcome this limitation, we propose to use an ensemble of support vector machines. Similar arguments mentioned above about the general ensemble of classi6ers can also be applied to the ensemble of support vector machines. Fig. 2 shows a general architecture of the proposed SVM ensemble. During the training phase, each individual SVM is trained independently by its own replicated training data set via a bootstrap method.

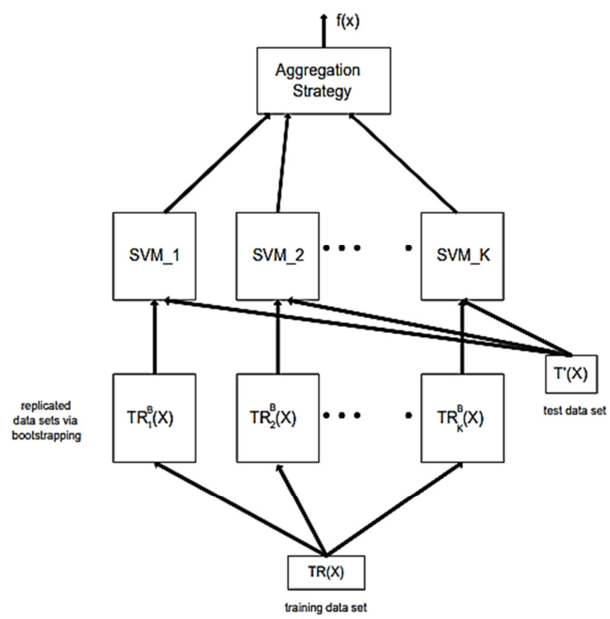

Figure 2: Architecture of the Ensemble SVM

Many methods for constructing an ensemble of classi6ers have been developed. The most important thing in constructing the SVM ensemble is that each individual SVM becomes different with another SVM as much as possible.

This requirement can be met by using different training sets for different SVMs. Some methods for selecting the training samples are bagging, boosting, randomization, stacking and bagging. Among them, we put focus on the representative methods such as bagging.

Bagging First, we explain a bagging technique to construct the SVM ensemble. In bagging, several SVMs are trained independently via a bootstrap method and then they are aggregated via an appropriate combination technique. Usually, we have a single training set TR = $\{(x_i; y_i)|i = 1; 2,\ldots\ldots\ldots\ldots\ldots,l\}$.

But we need K training samples sets to construct the SVM ensemble with K independent SVMs. From the statistical fact, we need to make the training sample sets different as much as possible in order to obtain higher improvement of the aggregation result. Each example x_i in the given training set TR may appear repeated times or not at all in any particular replicate training data set. Each replicate training set will be used to train a certain SVM.

C. Performance Evaluation Parameters

Root Mean Square Error (RMSE): RMSE is a parameter that determines the difference in squares between the output and the input.

$$RMSE = \sqrt{MSE} \tag{1}$$

Where, MSE= Mean Square Error

MSE of any estimator (classifier) measures the average squares of errors or deviations, i,e. the difference between the estimator and what is estimated. MSE is a risk function corresponding to the expected value of the squared error loss.

$$MSE = \frac{1}{N}(Target_{value} - Obtained_{value}) \tag{2}$$

Mean Absolute Percentage Error (MAPE): The mean absolute percentage error (MAPE) is a measure of the predictive accuracy of a forecasting method in statistics, for example in estimating the trend. It usually expresses the precision in percentage and is defined by the formula:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{Target_{value} - Obtained_{value}}{Target_{value}} \tag{3}$$

IV.RESULT ANALYSIS

The MATLAB platform is used to implement proposed model for the functional approximation of demand forecasting. Input for the demand forecasting model are :

1. Previous weekly warehouse sale (2017-2018)

2. Average moving sales of the last 2 weeks (may- June)

3. Moving average of the last 4 weekly sales

The output of the model corresponds to the expected demand for the next weekly sale or monthly sale. M-file programs are designed to predict demand with ensembled support vector machine technique. The proposed methodology is analyzed weekly as well as monthly.

A. Weekly Result Analysis

The performance of the proposed methodology for warehouse and retail sales forecasting is performed weekly. The weekly demand forecasting is also termed as short term prediction. The performance of the ensemble support vector machine is showed by parameters such as MAPE as well as RMSE.

Table I: Result analysis for Weekly Demand Forecasting

Week	Actual Demand	Forecasted Demand	MAPE	RMSE
1ST WEEK	480	480	0.0746577	0.0484564
2ND WEEK	778	778	0.0205923	0.0467737
3RD WEEK	817	816	0.0105534	0.0576028
4TH	871	871	0.0059323	0.0551642

WEEK				

B. Monthly Result Analysis

The performance of the proposed methodology for warehouse and retail sales forecasting is performed monthly from April-2017 to February-2018. The weekly demand forecasting is also termed as long-term prediction. The performance of the ensemble support vector machine is showed by parameters such as MAPE as well as RMSE.

Table II: Monthly Demand Forecasting of Warehouse Sales

Month	**Actual Demand**	**Forecasted Demand**	**MAPE**	**RMSE**
April-17	299	299	1.70815e-07	0.0487264
May-17	367707	367715	1.53885e-08	0.043823
Jun-17	379457	379467	2.08708e-08	0.044035
Jul-17	444	444	1.03895	0.0415644
Aug-17	757291	757302	5.63645e-09	0.043233
Sep-17	533680	533682	1.35378e-08	0.0456359
Oct-17	1741060	1741080	7.90376e-10	0.0439157
Nov-17	341446	341446	1.78961e-08	0.0431957
Dec-17	2391250	2391280	3.984e-10	0.0436662
Jan-18	254532	254540	2.16573e-08	0.0437386
Feb-18	2910390	2910420	2.48537e-10	0.0436792
Average Value			0.09445002	0.0441102

C. Comparative Analysis

In [8], author proposed an approach which combines multiple machine learning models in order to forecast demands of warehouses. The methodology consists of two levels termed as Stacked MLP with SMO, Bayesian Network and Linear Regression. Experiments are performed on three-year real sales data of a national dried fruits and nuts company from Turkey. The experimental results show 12% MAPE for Stacked MLP with SMO, Bayesian Network and Linear Regression than best result of stand-alone models. Table III and figure 3 shows the performance analysis.

Table III: Comparative Performance Evaluation for Demand Forecasting

Techniques	**MAPE**
Proposed Methodology	9%
Stacked MLP with SMO, Bayesian Network and Linear Regression [8]	12.7%
Stacked Generalization [8]	21.9%
Bayesian Network [8]	17.5%

Stacked Generalization with Bayesian Network and Linear Regression [8]	20.5%
Stacked Generalization with Bayesian Network, MLP and Linear Regression [8]	20.3%

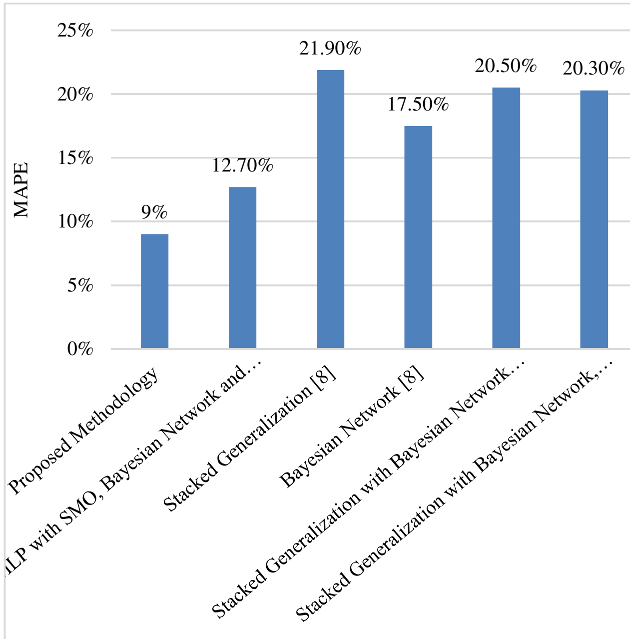

Figure 3: Comparative Performance Analysis

V. CONCLUSION

For the optimum balancing the supply chain in demand and sale better forecasting is important factor and by using Artificial Intelligence. Technique is easy and appropriate way to identify the realistic demand of future in various conditions by using previous data of demand and supply with inventory record by whole seller. The result of the evaluation shows that Support Vector Machine Ensemble Model is more efficient technique in Artificial Intelligence.

Following analysis is concluded from results obtained from proposed methodology:

- The Mean Absolute Percentage Error and Root Mean Square Error have been analyzed and it is identified that the MAPE is minimum as compared to existing work.
- You can achieve a broader perspective for your demand plan by producing a range of forecasts that you can re-calculate frequently to reflect market conditions, changing assumptions, and probabilities.
- Besides historical data, the demand sensing approach to get a more accurate, daily demand plan for the short-term forecast (4-6 weeks) so that it respond quickly to immediate changes on both the supply and demand side.
- Maximum MAPE has been calculated in the month Feb 2018 which is 2.48537e-10.
- To successfully estimate demand, the current and future marketing plans. This will help estimate any the increased the demand for any product. Where possible, review the competitors' promotions and if any of them had successes with something that you haven't done.

Now a days Artificial Intelligence is a best method to enhance the supply chain management and recent research shows that for the deep learning process is introduced for optimizing various conditions in supply chain management.

REFERENCES

[1] HaixiaSang's "A dynamic modeling simulation for supply chain management inventory service: a case study on a rental housing unit manufacturing and logistics company" Conference: the 2018 International Conference August 2018
[2] Christoph Flöthmann, Kai Hoberg "Career Patterns of Supply Chain Executives: An Optimal Matching Analysis" February 2018 DOI: 10.1111/jbl.12150

[3] Shanika L Wickramasuriya, George Athanasopoulos "Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization" March 2018 DOI: 10.1080/01621459.2018.1448825
[4] GokhanMertYagli, Dazhi Yang "Solar Forecast Reconciliation and Effects of Improved Base Forecasts" December 2018 DOI: 10.1109/PVSC.2018.8547846 Conference: 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC), At Waikoloa Village, HI, USA
[5] Lindsay R. Berry, Paul Helman "Probabilistic forecasting of heterogeneous consumer transaction-sales time series" August 2018
[6] Britta Gammelgaard , ChristophFlöthmann "Disentangling supply chain management competencies and their impact on performance: A knowledge-based view" April 2018 DOI: 10.1108/IJPDLM-02-2017-0120
[7] ChristophFlöthmann, Kai Hoberg "Competency requirements of supply chain planners and analysts and personal preferences of hiring managers" November 2018 DOI: 10.1108/SCM-03-2018-0101
[8] Irem Islek, Sule Gunduz Oguducu, "A Decision Support System for Demand Forecasting based on Classifier Ensemble", Communication papers of the Federated Conference on Computer Science and Information Systems, Vol. 13, 2017, pp. 35–41.
[9] Amit Sinha, William P. Millhiser "Matching supply with demand in supply chain management education" November 2016 DOI: 10.1108/IJLM-03-2015-0058
[10] Konstantinos Nikolopoulos, Konstantinos Bozos "Forecasting Supply Chain sporadic demand with Nearest Neighbor approaches" April 2016.