Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 11 (12): 2207-2213, 2011
ISSN 1812-5654 / DOI: 10.3923/jas.2011.2207.2213

Supply Planning Improvement: A Causal Forecasting Approach

Adnan Mukattash and Murad Samhouri
Department of Industrial Engineering, Hashemite University, Zarka, Jordan

Abstract: This study presents a methodology for building supply planning using a casual methodology (i.e., multivariate regression forecasting model) in order to model the effects and relationships between the finished goods and raw materials, then develop a forecasting model and use the forecast results to build a procurement plan. This work shows a novel application of multivariate regression forecasting model in supply chain as a link between finished goods and raw materials without a need to have information about the Bill of Material (BOM). This model minimizes the cumulative forecasting error in determining the raw materials that are needed to be purchased (instead of using the BOM to determine the demands on the raw materials). This model makes the planning system more secured i.e., the planner in the organization does not require to have an access on BOM.

Key words: Causal models, multivariate regression model, supply planning, var vector auto regression, demand forecasting

INTRODUCTION

Most manufacturers resort to supply planning to balance the uncertainty in demand. Supply planning helps the manufacturers to ensure that they get a complete picture of anticipated customer demand on a finished product level as well as supply. The procurement policy is one of the industrial competitive advantages. An over procurement can affect not only on the inventory and storage area but also budget spending on unnecessary procurement works. On the other hand, under procurement can make lack of raw materials which could result in interrupting a continuous process until the product cannot be finished on time or may cause customers' and money loss Armstrong and Collopy (2001).

Forecasting is a base of every part of the business plans including distribution and sales plan. In the field of logistics planning (i.e., sales, production and purchasing-planning), it is necessary to analyze a structured and detailed data according to items, markets and time behavior. Recently, many researchers have investigated the use of different forecasting models in several engineering and industrial applications. Kim (2007) suggested an optimal forecasting of bivariate Vector Autoregressive (VAR) process when one of predicted variables is available. Armstrong and Collopy (2001) used causal forces for the identification of asymmetric prediction interval. Shrinkage estimators for damping X12, ARIMA seasonals were developed by Miller and Williams (2004). MacGregor (2001) introduced a sound approach on using decomposition techniques for judgmental forecasting and estimation. Zareipour *et al.* (2006) used a Multivariate Adaptive Regression Splines (MARS) technique to forecast the Hourly Ontario Energy Price (HOEP). This technique was more accurate than the available forecasts for HOEP, demonstrating the MARS capability for electricity market price forecasting. Fu *et al.* (2004) built up a multivariate auto-regression model using the weather data of Fujin City. The model has good effects on fitting and forecasting. Figure 1 shows the application of forecasting in production planning. Forecasting is mainly used to forecast the following: (1) Raw material, (2) Components and (3) Finished goods.

Forecasting future customers' demands and determining the raw materials that are needed to be procured is quite easy if there is just one product and one customer. However, in reality, demand planning and procurement plan comprises hundreds or even thousands of individual finished goods, raw materials and customers. In some cases, it is even impossible to list all the finished products (e.g., in case of configurable products and finished goods), or to know all the customers (Simchi-Levi *et al.*, 2000).

One of the practical problems with raw materials procurement planning in some industrial organizations occurs when the raw materials are used to produce large number of finished goods (e.g., hundreds of products), then the problem appears by how the organization will determine the demand of the raw materials. Actually, there are many ways to determine the raw materials needed to be procured. One of the most popular and reliable

Corresponding Author: Murad Samhouri, Department of Industrial Engineering, Hashemite University, Zarka, Jordan

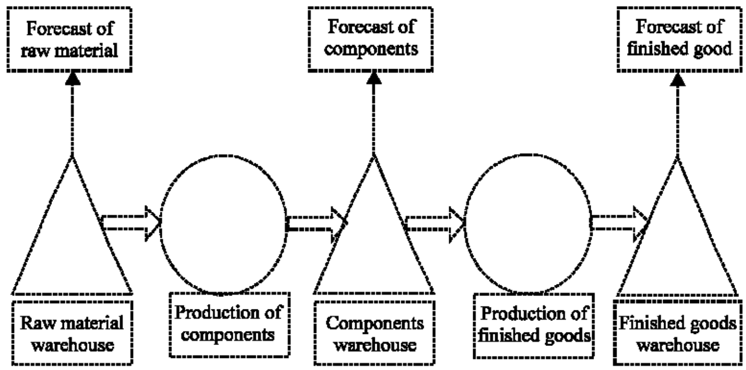

Fig. 1: The application of forecasting in production planning (Meindl, 2004)

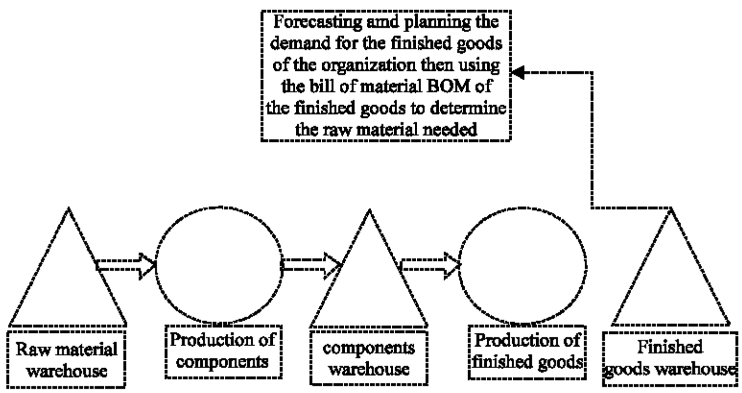

Fig. 2: Raw materials forecasting using BOM (Simchi-Levi *et al*., 2000)

methods to forecast raw materials and develop a proper procurement plan is by forecasting the demand for the finished goods using the Bill of Material (BOM) of the finished goods as it is shown in Fig. 2. In other words, using a univariate time series-forecasting model for the raw materials. But what if the organization has hundreds of these finished goods and there is one of the raw materials which is used to produce all of these finished goods, then it will be a must to forecast the demands of these hundreds of finished goods in order to determine the quantities needed of this raw material. Moreover, the use of BOM to determine the raw materials demand in this way, will generate a high cumulative forecasting error by times of the number of finished goods that have been forecasted (i.e., hundreds).

Another method to forecast raw materials and develop a procurement plan is using a direct forecasting of the raw material as it is shown in Fig. 3. In other words, to use the historical data of the raw materials consumption and build a univariate time series forecasting model (Fackler and Krieger, 1986). The main problem with this method is its direct way of forecasting the raw materials that does not represent the actual customer demand pattern which appears in the historical demand data of the finished goods.

To solve the previous problems, a practical and novel method that combines these two methods of forecasting is proposed in this study. Our new approach gives more accurate forecasts, minimizes the cumulative forecasting error and represents the actual customer demand patterns using a multivariate regression forecasting model as a new application of multivariate regression forecasting models in supply chain as a link between finished goods and raw materials. Figure 4 shows the new forecasting model. This model forecasts raw materials using the historical data of both raw materials and finished goods without a need to know or have access on the bill of materials of the finished goods.

IMPLEMENTING QUANTITATIVE FORECASTING

Forecasting Methods are divided into two categories as follows:

- Qualitative forecasting methods
- Quantitative forecasting methods

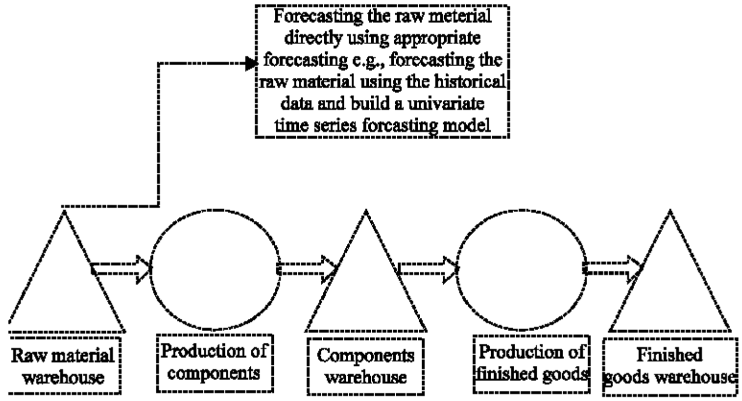

Fig. 3: Direct forecasting method of raw materials (Fackler and Krieger, 1986)

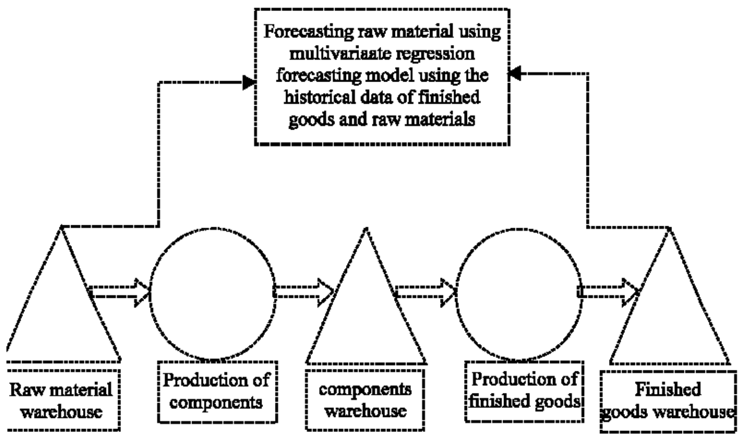

Fig. 4: The new multivariate regression forecasting model

Qualitative or subjective methods incorporate factors like the forecaster's intuition, emotions, personal experience and value system. They include: (1) Jury of executive opinion, (2) Sales force composites, (3) Delphi method and (4) Consumer market surveys. Quantitative or objective methods employ one or more mathematical models that rely on historical data and/or causal/indicator variables to forecast demand. Quantitative forecasting methods include: (1) Time series methods and (2) Causal models.

When selecting the forecasting method, it should be based on the following considerations:

- Orecasting horizon (i.e., validity of extrapolating past data)
- Availability and quality of data
- Lead times (i.e., time pressures)
- Cost of forecasting (i.e., understanding the value of forecasting accuracy)
- Forecasting flexibility (i.e., amenability of the model to revision; quite often, a trade-off between filtering out noise and the ability of the model to respond to abrupt and/or drastic changes)

The flow chart, shown in Fig. 5, gives the detail of process of implementing quantitative forecasting.

VECTOR AUTOREGRESSION (VAR) MODELS

Vector autoregression forecasting (VAR) model are used in the economic theory to find the relation between variables. The vector autoregressive (VAR) is commonly used for forecasting systems of interrelated time series and for analyzing the dynamic impact of random disturbances on the system of variables (Dickey and Fuller, 1979).

The VAR approach sidesteps the need for structural modeling by treating every variable as endogenous in the

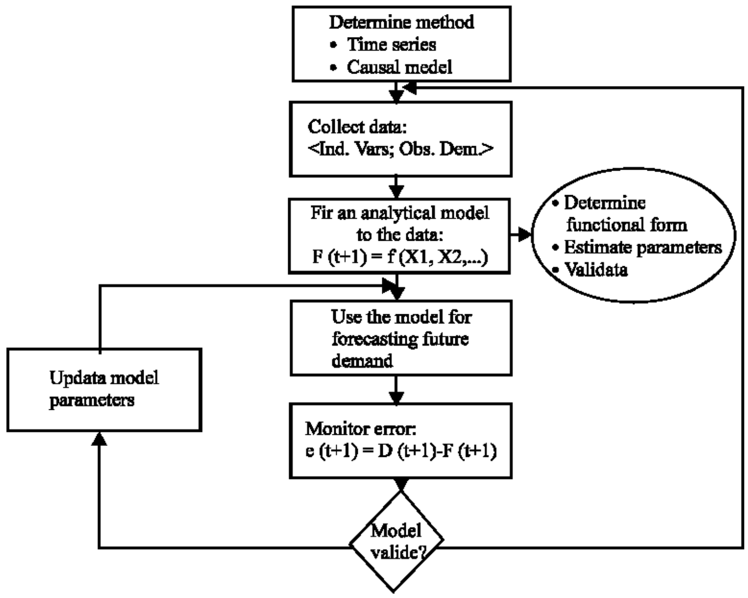

Fig. 5: A flow chart for implementing quantitative forecasting (Engle and Brown, 1986)

system as a function of the lagged values of all endogenous variables in the system. The term autoregressive is due to the appearance of the lagged values of the dependent variable on the right-hand side and the term vector is due to the fact that a vector of two (or more) variables is included in the system model.

The mathematical representation of a VAR system is given in Eq. 1:

$$[Y]_t = [A][Y]_{t-1} + \ldots + [A'][Y]_{t-k} + [e]_t \qquad (1)$$

where, p is the number of variables be considered in the system, k is the number of lags be considered in the system, $[Y]_t$, $[Y]_{t-1}$, ... $[Y]_{t-k}$ are the $1\times p$ vector of variables, and the [A], ... and [A'] are the $p\times p$ matrices of coefficients to be estimated, $[e]_t$ is a $1\times p$ vector of innovations that may be contemporaneously correlated but are uncorrelated with their own lagged values and uncorrelated with all of the right-hand side variables.

Since there are only lagged values of the endogenous variables appearing on the right-hand side of the equations, simultaneity is not an issue and OLS yields consistent estimates. Moreover, even though the innovations may be contemporaneously correlated, OLS is efficient and equivalent to GLS since all equations have identical regressors.

Before estimating the VAR, it is a must to decide the maximum lag lengths, k to generate the white noise of error terms. This is an empirical question. The maximum lag lengths in the VAR model is based on the smallest value of Akaike (AIC) or Schwarz (SIC) of the VAR to determine the appropriate lags.

Basically, the advantages of VAR could be summarized as follows:

- The method is simple; one does not have to worry about determining which variables are endogenous and which ones exogenous. All variables in VAR are endogenous
- Estimation is simple; the usual ordinary least square OLS method can be applied to each equation separately
- The forecasts obtained by this method are in many cases better than those obtained from the more complex simultaneous-equation models

The regression model is an explicitly multivariate model, in which variables are explained and forecasted on the basis of their own history and the histories of other related variables. Exploiting such cross-variables linkages may lead to good and intuitive forecasting model and to better forecasts than those obtained from univariate models.

Some of the practical problems (e.g., if the organization produces many products (e.g., n finished goods) and one of the raw materials procured (e.g., x) is a component for all n finished goods or many of the n products. One method to determine the raw material demand and a procurement plan is determined by forecasting the n finished product and using Bill Of Materials (BOM) of the n finished goods. This method makes a cumulative forecasting error (n forecasting errors related to the number of finished goods).

Using multivariate regression forecasting models (e.g., VAR vector autoregression forecasting model) minimizes the forecasting errors and gives a more accurate raw material procurement plan. This is performed by only determining the historical data (i.e., historical time series) for a part of the n finished goods) i.e., 1 to n) and the organization's demand for the raw material x.

In order to assure a strong causality relationship between the variables in the forecasting model, it is recommended that x has a high percentage in the formulation of product 1 and 2 compared with other n finished goods. This will have the highest demand across the other finished products to assure a strong relationship between variables and consequently give more accurate raw material procurement plan. In addition, it is recommended when building such a model that the finished goods that are chosen to represent the model represent the majority of all finished goods patterns with the same trend (i.e., seasonality and cycles).

VECTOR ERROR CORRECTION MODEL (VEC)

The concept of cointegration enriches the kinds of dynamic models. If Y_t and X_t, given in Eq. 1 are not cointegrated, we might estimate a dynamic model in first differences. As an example, consider Eq. 2:

$$\Delta Y_t = \beta_0 + \sum_{j=1}^{k} \beta_j \Delta X_{1t-j} + \sum_{j=1}^{h} \alpha_j \Delta Y_{t-j} + e_t \qquad (2)$$

where, it has zero mean given $\Delta Y_{t-1}, \ldots, \Delta Y_{t-h}, \Delta X_t, \Delta X_{t-1}, \ldots, \Delta X_{t-k}$. If one views this as a rational distributed lag model, the impact propensity, long run propensity and lag distribution for ΔY could be found as distributed lag in ΔX. If Y and X are cointegrated, then the obtained estimated error term must be stationary (i.e., I (0)). The lagged estimated error term is included as given in Eq. 3:

$$\Delta Y_t = \beta_0 + \sum_{j=1}^{k} \beta_j \Delta X_{1t-j} + \sum_{j=1}^{h} \alpha_j \Delta Y_{t-j} + \delta\, Z_{t-1} + \varepsilon_t \qquad (3)$$

Where:

$$Z_t = \hat{e}_t = Y_t - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j \Delta X_{1t-j} - \sum_{j=1}^{h} \hat{\alpha}_j \Delta Y_{t-j}$$

is the one-period lagged value of the estimated error of the cointegrating regression obtained from OLS estimation. This term is called the error correction term, ΔY is the difference in the model output variables, β_0 is the initial short-run dynamic coefficient of the input variables, β_j are the short-run dynamic coefficients of the input variables, ΔX is the difference in the model input variables, α_j are the short-run dynamic coefficients of the output variables, k is the number of time periods for input model variables, h is the number of time periods for output model variables, δ is the long-run coefficient, Z_t or $\hat{e}_t$ is the one-period lagged value of the estimated error, ε_t is the error in the model, Y_t are the output variables, $\hat{\beta}_0$ is the one-period lagged value short-run initial coefficient of the input variables, $\hat{\beta}_j$ are the one-period lagged value short-run coefficients of the input variables and $\hat{\alpha}_j$ are the one-period lagged value short-run coefficients of the output variables The principle behind this model is that there often exists a long run equilibrium relationship between two economic variables. In the short run, however, there may be disequilibrium with the error correction mechanism, a proportion of the disequilibrium is corrected in the next period. The error correction process is thus a means to reconcile short-run and long run behavior. Therefore, in the error correction model, the right hand side contains the short-run dynamic coefficients (i.e., α_i, β_i) as well as the long-run coefficient (i.e., δ). The absolute value of δ decides how quickly the equilibrium is restored.

The error correction model of the consumption function becomes as given in Eq. 4:

$$\Delta C_t = \beta_0 + \sum_{j=1}^{k} \beta_j \Delta Y_{t-1} + \sum_{j=1}^{h} \alpha_j \Delta C_{t-j} + \delta Z_{t-1} + e_t \qquad (4)$$

The error correction term:

$$Z_t = C_t - \sum_{i=0}^{a} \gamma_i Y_{t-i} - \sum_{i=1}^{b} \vartheta_i C_{t-i}$$

is obtained from the OLS regression, where, ΔC_t is the error in consumption function, C_t is the consumption function, a is the number of time periods for the output variables, b is the number of time periods for the consumptions, γ_i are the short-run dynamic coefficients of the output variables and ϑ_i are the short-run dynamic coefficients of the consumption function.

THE MULTIVARIATE REGRESSION FORECASTING MODEL

The development of our new approach aimed to create supply plan using causal forecasting methodology to ensure a profitable match of demand and supply. In order to maximize the return on assets for organizations that produce many finished goods and uses one of the raw materials in these finished products, the following procedure for a multivariate regression-forecasting model was adopted:

- **Step 1:** Determine the planning horizon (forecasting horizon) and the one step ahead forecast

The forecasting horizon is defined as the number of periods between today and the date of the forecasting is made, where the step ahead forecast depends on the observation of frequency data (e.g., month, quarter, half year and year).

- **Step 2:** Determine the raw material needed to be forecasted in the procurement plan and used to produce all of the finished goods

In this step, the reference to determine where the raw material is used is the bill of material of finished goods. This step provides the exogenous variable (i.e., determines the related variables that explain and forecast the raw material)

- **Step 3:** Build the multiple linear regression forecasting model with the exogenous variables for all finished goods and without any endogenous variables and a constant term as given in Eq. 5, then run the Least Square (LS) regression

$$D = b_0 + b_1.\ X_1 + ...\ b_k\ X_k + e \tag{5}$$

where, X_i, I = 1… k, are the model exogenous variables (finished goods), b_i, I = 0,…, k, are unknown model parameters, e is a random variable following a normal distribution with zero mean and some unknown variance s^2. D follows a normal distribution where it is the raw material.

- **Step 4:** Minimize the number of exogenous variables of the estimated model using the t-statistics (i.e., this provides a test of hypothesis of irrelevance e.g. with a 5% probability of incorrect rejection), in order to check whether zero is outside the 95% confidence interval for the parameter and whether t-statistics absolute value is greater than 2, so to test the hypothesis of irrelevance for all exogenous variables and select the forecasting model with less exogenous variables

Another way to build the model is using Pareto analyses (i.e., 80% of the effects come from 20% of the causes) to determine where the major demand of the raw material is used in producing finished goods (i.e., which of the finished goods take the majority of the raw material demand). This makes the model simple parsimony principle where other thing being the same. Simple models are preferable to complex models and using Pareto analyses, the exogenous variables can be determined directly and this model can represent a forecasting model using a Pareto analyses as a model selection criteria. But Pareto analyses must be determined for each h-step-ahead forecast because the major demand of the raw material is used in production finished goods may change from finished good to another.

- **Step 5:** Build a VAR using vector autoregression forecasting model with exogenous variables (i.e., explanatory variables are the finished goods that were previously determined) that have been selected in step 3 using any of the two ways (i.e., building the multivariate regression model or using Pareto analyses) and with endogenous variables which are the lag of the raw material and finished goods with order p using model selection criteria such as AIC, the planning horizon for the raw material can be forecasted
- **Step 6:** Determine the h-step-ahead forecasts for the raw material using vector autoregression forecasting model, but here encounter the right-hand-side variable problem (i.e., it is a must to insert the optimal value of the h-step-ahead forecast for the finished goods of vector autoregression forecasting model before forecasting the h-step-ahead for raw material (e.g., fit a univariate forecasting model for each exogenous variable of the finished goods)

Instead of steps 5 and 6 (i.e., forecasting the finished goods to forecast the raw material needed to be procured), a VEC long-run relationship model can be built to model the relationships between the raw material and finished goods with an endogenous variables just not exogenous variables with the lag operator of finished goods that are determined using Pareto analyses.

The model introduced in this study is a novel usage of forecasting methods in supply planning. It is a new approach that can be used with or without the demand planning process. This model was compared to previous studies and models in this field and it has overcome several disadvantages of these models. The model of (Kim, 2007) introduced an optimal forecasting of bivariate Vector Autoregressive (VAR) but with one of the predicted variables is available and this was a major disadvantage of this model when compared to our multi-variable (i.e., multivariate) regression forecasting model.

Another good model was suggested by Zareipour *et al.* (2006). This model is based on a Multivariate Adaptive Regression Spline (MARS) and achieved highly-accurate forecasting results. But this model has a limitation of using a direct forecasting technique to predict the output. In fact, our suggested model uses an approach that gives more accurate forecasts without a need for direct relationship beween inputs and outputs.

The model built by Simchi-Levi *et al.* (2000) comprised a univariate time series-forecasting model for the raw materials. The major setback of this model was the use of Bill of Materials (BOM) to determine the raw materials demands depending on the finished products forecasts, so they needed to forecast all the finished products demands regardless their large number. At the same time, our novel model, suggested in this paper, forecasts raw materials using only the historical data of both raw materials and finished products and does not need the BOM information or finished products forecasts. The forecasting model of Fackler and Kreiger (1986) suggested a method to forecast raw materials and develop a procurement plan using a direct forecasting data of the raw materials. The main problem of this method is its direct way of forecasting that that does not represent the actual customer demand pattern of the finished goods. On the other hand, our model in this work represents the actual customer demand pattern using a multivariate regression forecasting model in supply chain as a link between finished goods and raw materials.

CONCLUSIONS

Demand planning serves as the basis of every planning activity in a demand-supply network and ultimately determines the effectiveness of manufacturing and logistics planning. This paper presented a novel usage of forecasting methods in supply planning. The newly generated model can be used with or without the demand planning process.

Causal forecasting methodology for building supply planning and modeling the effects and relationships between the finished goods and raw materials used the forecast results to build a procurement plan in the supply planning.

VAR models are considered reduced forms which are compatible with many structural forms which are descriptions of how the economy works. These models can be used in supply planning and a specific version of the VAR-models is the VEC-model for variables that contain stochastic trends which have levels that are connected (i.e., cointegrated) in the long run which also can be used to model the relationship between finished goods and raw materials and to generate forecasts in special cases in supply planning. Using these models depends on the data patterns available i.e., historical data of raw materials and finished goods and the relationships between them.

REFERENCES

Armstrong, J. and F. Collopy, 2001. Identification of a symmetric prediction intervals through causal forces. J. Forecasting, 20: 273-283.

Dickey, D.A. and W.A. Fuller, 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc., 74: 427-431.

Engle, R.F. and S.J. Brown, 1986. Model selection for forecasting. Applied Math. Comput., 20: 313-327.

Fackler, J.S. and S.C. Krieger, 1986. An application of vector time series techniques to macroeconomic forecasting. J. Bus. Econ. Statist., 4: 71-80.

Fu, Q., Y. Ma and M. Young, 2004. Applying multivariate auto-regression model to forecast the water requirements of well irrigation rice in sanjiang plain. Nat. Sci., 2: 8-14.

Kim, Y., 2007. Optimal forecasting using bivariate recursive structural VAR model. Korean J. Econ., 14: 220-229.

MacGregor, D., 2001. Decomposition for Judgmental Forecasting and Estimation. In: Principles of Forecasting, Armstrong, J.S. (Ed.). Kluwer Academic Publishers, Massachusetts, Norwell, MA, pp: 107-123.

Meindl, S.P., 2004. Supply Chain Management, Strategy, Planning and Operations. Pearson Prentice Hall, New Jersey.

Miller, D. and D. Williams, 2004. Shrinkage estimators for damping X12, ARIMA seasonals. Int. J. Forecasting, 20: 529-549.

Simchi-Levi, D., P. Kaminsky and E. Simchi-Levi, 2000. Designing and Managing the Supply Chain. McGraw Hill, Boston.

Zareipour, H., K. Bhattacharya and C.A. Canizares, 2006. Forecasting the hourly ontario energy price by multivariate adaptive regression splines. Proceedings of the IEEE Power Engineering Society General Meeting, (PESG'06), Montreal, Que, pp: 7-7.