A Novel Prescriptive Model to Improve Supply Chain Networks by Monitoring the Relationship Between Influential Variables Across the Network

Mohmmad Khajehzadeh[a], Farhad Pazhuheian[a], Farima Seifi[b], Golshan Madraki[c], Alireza Ghorbani[d]

[a] Corresponding Author- Industrial Engineering Department, Iran University of Science and Technology, Tehran, Iran, E-mail: Khajezadeh.sina@gmail.com

E-mail: Pazhuheian: farhad.pzh@gmail.com

[b] Management Department, Anderson School of Management, University of California, Los Angeles, CA 90095, USA
E-mail: Farima.seifi.2025@anderson.ucla.edu

[c] Reh School of Business, Clarkson University, Potsdam, NY, 13699, USA.
E-mail:

[d] Department of Statistics, Shahid Beheshti University, Tehran, Iran
E-mail: ghorbanialireza@outlook.com
Tel: +98 (910) 676 8616
ORCID: https://orcid.org/0009-0002-8037-7977

A Novel Prescriptive Model to Improve Supply Chain Networks by Monitoring the Relationship Between Influential Variables Across the Network

Abstract

The goal of Supply Chain Monitoring is to provide an efficient tracking system for ensuring a secure flow of goods and services throughout the supply chain. Supply chain monitoring helps identify and address unexpected events early. There are five main components in supply chain networks including Manufacturing, Warehousing, Procurement, Logistic/transportation and Demand. Numerous factors in each of the five components of the supply chain have an impact on sales and production. This paper presents a comprehensive method to monitor and analyze these factors' impact on both sales and production, ultimately aiming to identify areas for cost reduction and improvement. To achieve this goal, the sales and production are modeled and evaluated. Then, products with out-of-control behavior are simultaneously identified. Finally, to optimize out-of-control products, we considered the most influential factors affecting sales and production. The optimal values for out-of-control products were selected to minimize operating costs while simultaneously maximizing operating profits within the supply chain. A case study in the Personal Care industry shows that the method increases the operation profit rate for out-of-control products.

Keywords: Supply Chin Monitoring, Generalized Estimating Equation (GEE), Hotelling T^2 Control Chart, Joint Optimization Plot

1 Introduction

A supply chain network consists of five main sections: Procurement and suppliers, manufacturing units, warehouses, logistic/transportation such as distribution centers and retailers, and demand/customers. These units are responsible for acquiring raw materials, producing goods, and delivering them to end-users. Extensive research on supply chain networks highlights the pivotal role of a streamlined network in determining a business's overall economic success (Wang et al., 2020). Accordingly, supply chain operations come across various types of risks, such as delays, poor quality from suppliers, procurement failures, imprecise forecasts, uncertain consumer demands, and potential supply chain disruption like natural disasters (Akkermans and Wassenhove, 2018). Clearly, in the absence of a well-organized supply chain strategy, these risks and vulnerabilities might result in financial losses or even a complete collapse of the supply chain network (Wang et al., 2020). Therefore, creating a measurement system that facilitates a coordination approach for joint decision-making in all components of supply chain is so crucial (Kim & Oh, 2005).

Due to the complex nature of processes and activities in supply chain, managing risks is so essential to avoid disruption (Ethirajan et al., 2021). A supply chain must overcome a disruption before it can leverage events as an opportunity for growth (Nikookar et al., 2024). Accordingly, supply chain operations come across various types of risks, such as delays, poor quality from suppliers, procurement failures, imprecise forecasts, uncertain consumer demands, and potential supply chain disruption like natural disasters (Chopra & Sodhi, 2006, Tang, 2006). Clearly, in the absence of a well-organized supply chain strategy, these risks and vulnerabilities might result in financial losses or even a complete collapse of the supply chain network (Wang et al., 2020). Moreover, supply chain disruptions have become more frequent and severe (Manhart et al., 2020) which requires a rethinking of supply chain management (SCM) for practice to cope with extreme situations, present and in the future, whether due to pandemics, war, climate change, or biodiversity collapse (Sodhi & Tang, 2021). Therefore, creating a measurement system that facilitates a coordination approach for joint decision-making in all components of supply chain is so crucial (Kim & Oh, 2005).

The operational complexity is increased as we extend to downstream supply chain sections connecting producers to consumers (Gómez & Lee, 2023). This system should be aligned with the goals of the supply chain network's independent elements, coordinate their efforts, and ultimately enhance the overall performance of the entire chain (Wang, 2010). Therefore, monitoring supply chain operations by timely detecting abnormal operations is crucial for the effective functioning and economic viability of a supply chain system.

Thus, *supply chain monitoring* (*SCM*o) is defined to promptly detect the network and provide early warnings of abnormal operations for effective *supply chain management* (*SCM*). SCMo has become an integral component of Supply Chain management. The primary objective of SCMo is to improve decision-making by characterizing the normal operating conditions of a supply chain, revealing discrepancies between planning and execution, warnings for abnormal situations, identifying possible root causes, and providing recommendations for mitigation (Wang et al., 2023).

During the last decades, the mainstream focus of supply chain management was on cost efficiencies by using just-in-time methods and avoiding holding excess inventory (Kovács and Falagara Sigala, 2021). Most studies in literature often analyze only one out of the five components in supply chain, e.g., solely focused on optimizing manufacturing or distribution. Consequently, they use methods that might be sufficient for limited number of variables such as statistical methods based on cross-sectional data, which can only provide one measurement for each response at a specific time point. The primary contribution of this paper is to address these shortcomings through the follows unique features:

1- Our proposed analysis focuses on all five components of supply chain functions over time to reduce operating costs and increase supply chain efficiency.

2- Our method is comprehensive and novel that combines established techniques and can create an efficient approach for handling numerous variables and correlated characteristics in the data when monitoring the supply chain network across its five functions over time. This method is rooted in a robust foundation grounded in principles of multivariate analysis, focusing on longitudinal analysis and optimization techniques, aiming to optimize supply chain networks.

3- Our proposed method can be applied and extended to larger and more complex supply chain networks, where multiple elements exist within each of the five sections of the supply chain network.

4- Finally, a novel case study on the supply chain of a personal care company in the Middle East is presented, and the application and performance of the proposed method is assessed over this real-world case study.

2 Literature Review and Background

Supply chain management now relies more than ever on data to capture cost and performance trends, monitor inventory, support process control and improvement, as well as optimize production. To gain insights and make informed decisions about all sections of supply chain management, it is important to understand the value of data analytics and its effective use in supply chain management (Sukha & Prabhu, 2022). This application of advanced data analytics techniques to supply chain management is called *Supply Chain Analytics* (*SCA*) (Gilvan, 2014). The SCA techniques have a significant role in SCMo and can be classified into three main types:

1) Descriptive analytics extracts valuable insights from the network data to describe "what is happening". For instance, real-time information about the location and quantities of goods in the supply chain network equips managers with the necessary tools to make adjustments regarding delivery schedules, replenishment orders, emergency orders, and transportation modes, etc.

2) Predictive analytics derives demand forecasts from historical data and predicts "what will be happening" in the future.

3) Prescriptive analytics generates decision recommendations by combining descriptive and predictive analytics models with mathematical optimization techniques. It addresses the question of "what should be happening" and guides decision-makers towards optimized strategies and solutions (Gilvan, 2014). Notably, prescriptive analytics receives significant attention in academic research, software development, and practical application within the domain of SCA.

--Insert FIGURE 1 Approximately Here--

Nguyen et al, (2017) proposed a visual framework FIGURE 1, which helps the understanding of the relationships between different supply chain layers/components and the role of data analytics studies in assessing their performance. The first layer represents the five main key functions of Supply Chain Management, while supply chain data analytics is associated with the second and third layers in this taxonomy, referring to the aforementioned types of SCA studies.

To explore the diverse usage of data analytics approaches in SCM, the following recent studies can be discussed as examples:

As for procurement, Jain et al. (2014) conducted a study employing a data mining approach to uncover the hidden relationships between data used for suppliers' selection and their overall rating based on prior performance. This approach significantly aids in optimizing the supplier selection process. Similarly, Choi et al., (2016) introduced a novel data analytics approach using Fuzzy Cognitive Maps to enhance decision-making in IT service procurement for the public sector. This unique method combines data analytics with intuitive qualitative techniques to create decision models, and its efficacy was validated through a case study, demonstrating its value in facilitating robust public decision-making. Mori et al., (2012) utilized Support Vector Machine and Logistic Regression to build a prediction model for customer-supplier relationships, which can help to identify potential business partners.

As for the manufacturing, Zhang et al., (2017) proposed an overall architecture called data-based analytics for product lifecycle. This architecture leverages data analytics and service-driven patterns and through a practical application scenario, it demonstrates impactful benefits for customers, manufacturers, the environment, and all stages of product lifecycle management.

In studies on inventory and warehousing, Chiang et al., (2011) introduced an association index and proposed a data mining-based storage assignment approach that enhances the efficiency of order picking. Khurana & Kumar, (2017) conducted a practical usage of data analytics in inventory management and implemented linear discriminant analysis on a large data set to find the dependencies. Chen, (2021) addressed an inventory control problem with active exploration in the inventory through lost sales in a shifting demand environment through historical data analysis. Suwignjo et al., (2023) applied gradient boosting model for solving the inventory status (overstock, understock) by considering demand forecast in an FMCG company.

In terms of logistics/transportation, Zhao et al., (2016) used the upper and lower limits of uncertain parameters from historical data to redesign a green supply chain network. Li et al., (2015) employed Lasso Granger causality models to select the most relevant data to build a traffic prediction model. Subhakanta & Mohanty (2018) proposed a deterministic linear programming model to address a transportation problem where the unit cost of supplies, transportation and demands are uncertain. They minimized the expected value of an uncertain objective function with respect to some constraints under certain confidence level. Amellal et al., (2023) addressed the lack of accurate lead time for meeting customer demand by developing a hybrid model combining Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) architectures.

As for demand management, Salehan and Kim, (2015) used a sentiment mining approach to study predictors of online consumer review performance. Chong et al., (2016) employed a neural network to examine the impact of different variables such as online reviews, promotion strategies, and sentiments on product sales. Additionally, Mohan et al., (2021) proposed a demand forecasting and route optimization approach for delivering products on time and meeting customer's growing expectations. Nguyen, T, (2023) reviewed the artificial intelligence models such as recurrent neural networks for demand forecasting in supply chain over various industries.

Wang et al., (2020) and Kapil et al., (2021) described that the data-driven optimization techniques are playing a significant role in enhancing SCM in uncertain environments. By integrating machine learning, data analytics, and robust optimization, the planning of supply chain network can be more efficient and accurate. These data-driven techniques display their potential in improving SCM under uncertainty. Nahum, Méndez-Sánchez (2023) utilized machine learning method in the context of supply chain management for predicting the electricity consumption in Turkey to achieve smart and sustainable processes in making decisions. Finally, Nitin et al.., (2023) employed bibliometric statistical analysis on supply chain analytics to provide a systematic analysis of this area for identifying key research themes and sub-themes for enhancing performance of supply chain management and business value. TABLE 1 summarized the details of aforementioned studies on SCM using data-driven techniques:

--Insert Table 1 Approximately Here--

A major gap in literature is that most of the current studies have focused solely on one specific function out of the five in the supply chain network i.e., Manufacturing, procurement, demand, warehousing, transportation. To the best of our knowledge, there are no comprehensive studies integrating the analysis of all five main functions of supply chain network through data analytics approaches. Accordingly, a comprehensive analysis through a data-driven statistical model across all five supply chain functions for reducing operating costs will be a compelling research opportunity in this field. Moreover, to monitor the supply chain network across its five functions over time, a statistical method is required to be used due to numerous variables to be considered and the correlated characteristics in the data. Longitudinal data is particularly suitable for handling such data, unlike cross-sectional data, which can only provide one measurement for each response at a specific time point. One of the main advantages of longitudinal data is that the correlation among observations within an experimental unit leads to more specific power level in longitudinal data compared to cross-sectional data. Consequently, a smaller number of experimental units is needed in the sample to achieve a specific power level. Achieving a specific power level can guarantee a high likelihood of detecting meaningful effects or relationships between characteristics (Diggle et al., 2002). In this regard, a Generalized Estimating Equation model is employed to create a control chart for detecting inefficient items in the problem and then the paper proposes a joint optimization approach to enhance these items and boost supply chain efficiency.

2.1 Background

The key terms and concepts that will be used in our method are reviewed in this section to aid in the understanding of the proposed solutions and their effectiveness in addressing the challenges presented in the case study.

2.1.1 Longitudinal Data

The primary goal of a longitudinal study is to characterize changes in responses over time and determine the factors that influence these changes. Thus, the main characteristic of longitudinal data is the repeated measurement of subjects over time. One important feature of longitudinal clustered data is that each cluster consists of observations from a single experimental unit at different time points. (Diggle et al., 2002 and Fitzmaurice et al., 2012). Finding a useful set requires understanding the sources of random changes in longitudinal data. These sources can be classified into the following three general categories (Fitzmaurice et al., 2012):

- *Random effects*: In a situation where a population is randomly sampled, various aspects of the sample members' behavior represent random variations between experimental units. Random effects are the variables that differ among subjects.
- *Sequential correlation*: At least a portion of each measured unit exhibits a time-dependent response within that unit. These random changes are caused by the correlation between the measured pairs within the same unit, and this correlation depends on the time difference between the measured pairs. Usually, the correlation decreases with increasing time interval.
- *Measurement error*: The measurement process within the experimental unit may lead to changes in the data.

In longitudinal data, y_{ij} represents the response variable of the i-th subject in the j-th measurement, and x_{ij} is a p-dimensional vector of explanatory variables at time t_{ij}, where $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, m$, where n is the number of subjects and m is the number of measurements. Most longitudinal analyses are based on a regression model, such as the following linear model:

$$y_{ij} = \beta_1 x_{ij1} + \beta_2 x_{ij2} + \cdots + \beta_p x_{ijp} + \epsilon_{ij} \quad (1)$$

This model can be expressed in matrix form as follows:

$$y = X_i \beta + \epsilon_i \quad (2)$$

where X_i is an $n_i \times p$ matrix of explanatory variables, β is the vector of unknown regression coefficients of dimension p, $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})$ is the random vector of errors, and $y_i = (y_{i1}, \ldots, y_{in_i})$ represents the repeated responses for the i-th subject (Diggle et al., 2002 and Fitzmaurice et al., 2012).

2.1.2 Marginal Models

Marginal models are one of the common methods for longitudinal data modeling that will be used in this study. In marginal models, the response variable is modeled on covariate variables apart from the within-subject correlation structure (Fitzmaurice et al, 2012). In this model, the marginal expectation of the response variable is expressed as a function of the explanatory variables. The term "*marginal expectation*" refers to the average response in a subpopulation with common values of X. A marginal model is characterized by the following three key features.

The marginal expectation of the response, $\mu_{ij} = E\left(y_{ij} \middle| X_{ij}\right)$, depends on the covariates with a certain link function Eq. (3):

$$g(\mu_{ij}) = \eta = X'_{ij}\beta \tag{3}$$

The marginal variance of the responses is related to the marginal mean as Eq. (4):

$$Var(y_{ij}) = \phi.v(\mu_{ij}) \tag{4}$$

where, $v(\mu_{ij})$ is a specified variance function; and, ϕ is a scale parameter that may need to be estimated. The correlation between within-subject observations is a function of the marginal mean and additional parameters α.

within-subject communication of the repeated responses vector is modeled as (5) by considering the correlation pattern of the first-order autoregressive model, where $0 \leq \alpha \leq 1$ (Diggle et al, 2002):

$$Corr\left(y_{ij}, y_{ik} \middle| X_{ij}, X_{ik}\right) = \alpha^{|k-j|} \tag{5}$$

Therefore, in the marginal models, the correlation in the longitudinal data is considered through the variance-covariance matrix. In a marginal model, the identical relationship between the response variable and the matrix of covariate variables is assumed to apply to all subjects in the sample. The key feature of marginal models is that they model the mean response and within-subject relationships separately. Consequently, the regression coefficients in this model are interpreted as population averages, meaning that changes in the mean response relative to the predictor variables are examined in the sub-population defined by these predictors (Carrière et al., 2002 and Fitzmaurice et al., 2012). As a result, when studying time-independent predictor variables, i.e., variables that do not change for each individual during the follow-up period, population-average interpretations are typically preferred (Wu et al., 2012).

In these models, the method of *Generalized Estimating Equations* (*GEE*) is employed to estimate the parameters (Fitzmaurice et al, 2012). GEE allows for the estimation of parameters while considering the correlation structure between the response variables, without requiring knowledge of their specific distribution. The correlation matrix derived from this structure is assumed to be identical for all subjects in the sample. The data consists of repeated measures of the response variable and covariate variables within a group of subjects. With this method, a suitable model is constructed for the mean of the response variable, incorporating separate observations and correlated variables (Fitzmaurice et al, 2012). In most cases, according to the type of response variable and specific design conditions, a generalized linear model such as (6) can be used to model grouped structures. In (6), y_i represents the value of the response variable for subject i, X_i is the correlated variable or covariate, and β is a vector of model parameters or independent coefficients of X_i. ε_i represents the random terms, and g is the link function, which maps the set of possible values of the response variable to a linear function of the X variable.

$$y_i = \mu_i + \varepsilon_i \quad , g(\mu_i) = f(x) = X_i\beta \qquad (6)$$

To estimate the parameters of the "generalized linear model" and perform inference, it is typically assumed that the error terms (ε) have the same distributions and are independent. However, this assumption often does not hold in practice. As an alternative, Generalized Estimating Equations (GEE) offer a non-parametric approach that does not rely on the normal distribution assumption for the error term.

2.1.3 Generalized Estimating Equations (GEE)

In GEE, instead of assuming a specific distribution for the data, the best estimate for β is generated using iterative calculation techniques and trial and error. This approach aims to be the most descriptive for capturing the relationship between the response and dependent variables (Ziegler et al., 1998).

The GEE estimator for β in the marginal model is obtained by minimizing the objective function (7):

$$\sum_{i=1}^{N} \{y_i - \mu_i(\beta)\}' V_i^{-1} \{y_i - \mu_i(\beta)\} \tag{7}$$

where V is not dependent on β and μ_i, and μ_i is a vector of the average response with the following components:

$$\mu_{ij} = \mu_{ij}(\beta) = g^{-1}\left(X_{ij}' \beta\right) \tag{8}$$

Using differential and integral calculus, we can demonstrate that the existence of the minimum function in Eq. (7) requires the solution of the following generalized estimating equations:

$$\sum_{i=1}^{N} D_i' V_i^{-1} (y_i - \mu_i) = 0 \tag{9}$$

in which V_i is called "working" covariance matrix and $D_i = \partial \mu_i / \partial \beta$ represents the gradient or derivative matrix. Since GEE relies on both β and α, the following iterative two-step estimation procedure is necessary:

1. According to the current estimates of α and ϕ, V_i is estimated and the updated estimate of β is obtained as generalized estimating equations resulting from Eq. (9).
2. According to the current estimate of β, the updated estimates of α and ϕ are obtained based on the standardized residuals Eq. (10):

$$e_{ij} = \frac{y_{ij} - \widehat{\mu_{ij}}}{\sqrt{v(\hat{\mu}_{ij})}} \tag{10}$$

Finally, in this two-step estimation method, the process is typically iterated between steps 1 and 2 to ensure convergence (Fitzmaurice et al., 2012).

2.1.4 Hotelling Multivariate Control Chart (T^2 Control Chart)

In many instances, monitoring multiple related quality characteristics simultaneously is crucial. It helps control these traits effectively and evaluate their potentially deceptive nature. To tackle such scenarios, specialized tools must be employed to detect, identify, and analyze the significant sources of variability in a given process. Among the various techniques, *Multivariate Control Charts* stand out since they can simultaneously monitor and control multiple characteristics that define the quality of a single production process. The *Hotelling* T^2 control chart holds significant recognition in the literature and comes highly recommended for processes involving multiple qualitative characteristics. Since these features are interconnected, monitoring them collectively is crucial. The T^2 test statistic is derived from the Eq. (11) (Montgomery, 2019).

$$T^2 = n(\bar{X} - \bar{\bar{X}})' + S^{-1}(\bar{X} - \bar{\bar{X}}) \tag{11}$$

In (11), $\bar{X}$ is the mean vector, and S represents the covariance matrix of the process. The application of multivariate control Hotelling T^2 chart is performed in two phases:

- *Phase I*: the chart's upper control bound is calculated using Eq. (12)

$$UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,P,mn-m-p+1} \tag{12}$$

In (12), p is the number of variables, m is the number of samples, n is the sample size, and α is the parameter of the F distribution degree (Bersimis et al., 2007 and Tracy et al., 1992).

- *Phase II*: the chart's upper control bound is expressed by Eq. (13)

$$UCL = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{\alpha,P,mn-m-p+1} \tag{13}$$

The lower control limit for both phases is equal to zero in the control chart (Bersimis et al., 2007 and Tracy et al., 1992).

2.1.5 Joint Optimization Plot

In an industrial experiment or decision-making system, there are several control factors (independent variables) denoted as $x_1, \dots, x_k$, multiple control responses (dependent variables) represented by $y_1, \dots, y_N$, and various target values $\tau_1, \dots, \tau_N$. When aiming to optimize such a system, conflicts may arise in the results while attempting to optimize the control factors individually. Consequently, a relative combination of the factors is necessary to bring the multiple responses as close as possible to the specified target values. The application of the Joint Optimization method enables us to achieve this objective. Joint optimization refers to the process of finding the optimal values for multiple variables or parameters simultaneously. It involves considering the trade-offs and compromises between different objectives or constraints. The strategy for simultaneously optimizing multiple responses is presented as follows (Kuhnt & Rudak, 2013 and Pignatiello & Joseph, 1993):

Consider an experiment with control factors $x_1, \ldots, x_k$ and N responses $y_1, \ldots, y_N$ with target values $\tau_1, \ldots, \tau_N$. The optimal settings for the control factors should be determined to ensure that the means of the responses are on target with minimal variances. This can be achieved by minimizing the expected loss of y with respect to x, which is referred to as the risk function and is defined as follows:

$$\begin{aligned} R(x) &= E\big(loss(y|x)\big) = E((y-\tau)^T C(y-\tau)|x) \\ &= trace\,\big(C\Sigma(x)\big) + (\mu(x)-\tau)^T C(\mu(x)-\tau) \end{aligned} \tag{14}$$

where $(y-\tau)^T C(y-\tau)$ represents the loss function, and C is the cost matrix, $\mu(x) = E(y|x)$ denoting the expected value of y given x, and $\Sigma(\mathrm{x})$ represents the covariance matrix of y given x. In the case of independent responses $y_1, \ldots, y_N$, both the covariance matrix $\Sigma(x)$ and C become diagonal matrices, so Eq (14) turns to Eq (15) where $\mathrm{c_i}$ represents the i th element of cost matrix.

$$R(x) = \sum_{i=1}^{N} c_i . (\sigma_i^2(x) + (\mu_i(x) - \tau_i)^2) \tag{15}$$

Minimizing the risk function, as described in Eq. (16), means adjusting the average value (mean) towards the desired goal while keeping the variability (variance) as low as possible (Pignatiello, 1993). In situations involving an unknown matrix C, this cost matrix is decomposed to Eq. (16) in which A is diagonal standardization matrix and W is diagonal weight matrix.

$$C = A^T W A \tag{16}$$

Diagonal elements of weight matrix W are specified through a slop vector $\mathrm{d}\epsilon\mathrm{R}^{\mathrm{N}}$ for N responses and a stretch value log(a) in the following form

$$log\ \omega = d.log\ a \tag{17}$$

where ω is diagonal of weight matrix W and $\{a_t\}_{t=1}^{N}$ is an increasing equidistance vector within the interval $[\log a_{low}, \log a_{high}]$. Standardization matrix A for k control factors is defined as

$$A_y = diag([\frac{1}{k}\sum_{k=1}^{K}\widehat{var}(y_i|x_k)]_{i=1,\dots,N}^{-1/2}) \tag{18}$$

Therefore, the estimated risk function in Eq. (14) is given by (19) where b_i denotes the inverse of i th element of standardization matrix A.

$$\hat{R}(x) = \sum_{i=1}^{N}\omega_i.\frac{(\widehat{var}(y_i|x) + (\hat{E}(y_i|x) - \tau_i)^2)}{b_i^2} \tag{19}$$

The sequence of weight matrices ensures an optimal solution, and a joint optimization plot displays the optimal parameter setting for every cost matrix $C_t = A^T W_t A$ in one plot and its corresponding predicted response in other plot (Pignatiello, 1993).

3 Proposed Method

We propose a comprehensive method to monitor and optimize variables across all five main sections of a supply chain network. The goal is to enhance the overall performance of any general supply chain network. The method requires data from the five main sections of the supply chain over a twelve-month interval and organizing it as a matrix. Due to the high correlation and longitudinal structure of the data, it is necessary to employ a method that does not require assuming normality for the error distribution in the regression model. This leads to the application of the GEE method for modeling the problem. Notably, such a comprehensive statistical method has not been employed in previous studies analyzing supply chain networks. This study explored changes in production & sales over time (using GEE analysis) and subsequently used a Hotelling T^2 multivariate control chart to monitor product performance and identify any supply chain issues. Recognizing the need for optimization, a joint optimization method considering interconnected variables was employed to simultaneously optimize costs and profits. This approach, rarely used in supply chain monitoring, offers a unique and effective solution for improving overall performance. These steps were outlined as follows and shown

FIGURE ***2***:

Step 1: The variables resulting from the problem were reshaped into a longitudinal form to make a set of explanatory and response variables for sales and production. The supply chain has essentially been decomposed for further analysis.

Step 2: In this longitudinal supply chain study, the time variable is introduced as a fixed effect. Additionally, variables associated with all five functions of the supply chain are considered as covariates to assess their influence on the response variables. The (GEE) method was utilized to analyze the data, considering the within-*Stock Keeping Unit* (*SKU*) effect as a latent variable. The modeling process involved utilizing the "*xtgee command*" in *Stata software*. Two separate analyses were carried out: one for sales and another for production.

Step 3: Using the Hotelling T^2 control chart, we use this control chart to monitor two variables derived from the fitted values in Step 2 of the Phase I control chart. In this step (3), we detect and optimize out-of-control products using the *Joint Optimization* (*JOP*) method. We performed this crucial step with the *MSQC* package in R-4.3.2 software.

Step 4: The variables related to products beyond the control chart should be optimized. We utilized the JOP model, considering the cost of goods sold and finance costs of the products as the response variables. The decision variables for the modeling process were internal factors within the business. This phase was executed using the JOP package in R-4.3.2 software.

--Insert FIGURE 2 Approximately Here--

In the following, the efficiency of the proposed method is demonstrated on a real case study in the supply chain of a personal care company.

4 A Real-World Case Study

In this paper, a real-world case study was presented, conducted on the supply chain network of a company in the personal care industry operating in the Middle East. Due to confidentiality agreement, the company's name and specific products cannot be disclosed.

The study focused on analyzing a Multi-Echelon Supply Chain network that involved 51 *Stock Keeping Unit* (*SKU*) products. Data were comprehensively collected across all five main sections of the supply chain: procurement, manufacturing, warehousing, logistics/transportation, and demand management.

The examination covered the entire supply chain network, starting from the plant to the distribution centers (depots), and finally, the presence of products on store shelves. Due to the company's reliance on imported raw materials from foreign countries, most of the independent variables are from the procurement section. Notably, three main raw materials are identified as the major portion of the total purchasing from suppliers. These materials are represented as the I, II, and III raw materials in the study. Three economic factors: the consumer price index, currency rate, and point-to-point inflation rates have direct impacts on consumer purchasing behavior (Khajehzadeh et al, 2022). Therefore, during the study, these significant factors were also collected and analyzed within the demand section. These factors were collected as part of the demand management section to understand their influence on the overall demand for products. Given that there are 51 SKUs and 21 variables derived from all five sections of the supply chain network, data collection extended over twelve months for all products. This resulted in a database represented as a matrix with $612 = 51^{\times}12$ rows, and 21 columns. The response variables and independent variables corresponding to each supply chain section were integrated into the dataset for analysis.

This case study provides valuable insights into the complexities of the supply chain network of our targeted personal care company. **Error! Reference source not found.** shows the variables of supply chain:

---Insert Table 2 Approximately Here---

4.1 Results

A random effects model with the identity link function was fitted to the data for each of the two variables of sales and production. The mixed-effects model for each response is as Eq.20 where β_0 is the intercept, u_i is the random effect of the i-th, β_{ij} is the effect of variable x_{ij}, and ε_{ij} is the random error.

$$y_{ij} = \beta_0 + u_i + \sum_{i=1}^{52} \sum_{j=1}^{12} \beta_{ij} x_{ij} + \varepsilon_{ij} \tag{20}$$

The results of fitting the GEE model on production and sales are shown in TABLE 3

RESULTS FOR PRODUCTIONand **Error! Reference source not found.**, respectively.

---Insert Table 3 Approximately Here---

Based on the results of TABLE 3

RESULTS FOR PRODUCTION, variables x_4 , x_5 , x_6, and x_{19} are statistically significant as their p-values are less than 0.05. This means they have a significant impact on production. With each unit increase in x_4, production increases by 0.22 units, and with each unit increase in "storage of raw material (II)" (x_5), production increases by 0.14 units. Furthermore, according to the results, each unit increase in "storage of raw material (III)" (x_6) leads to a significant increase in production by 1.89

units. However, variables “third raw material purchasing quantity” (x_{15}) has a negative coefficient and is statistically significant, suggesting that an increase in x6 leads to a decrease in production, each unit increase will result in a decrease in production by 0.012 and 0.002 units.

--Insert Table 4 Approximately Here--

Based on the results of TABLE 4

RESULTS FOR SALE, variables x_4, x_5 , x_6 , x_{18} , x_{19} and x_{21} are statistically significant as their p-values are less than 0.05. This means they have a significant impact on sales; accordingly, with each unit increase in “storage of raw material (I)” (x_4), sales increase by 0.19 units, and with each unit increase in “storage of raw material (II)” (x_5), sales decrease by 0.098 units. Additionally, with each unit increase in “shelf Price” x_{19} and “consumer price index” x_{21} , sales decrease by 0.001 and 0.02 units, respectively. On the other hand, each unit increase in x_{18} leads to a sales increase of 0.023 units. (p-value < 0.05). The result of monitoring the fitted values obtained by the Hotelling T^2 control chart is shown in FIGURE .

--Insert FIGURE 3 Approximately Here--

As shown in the chart, there are 14 samples in the company's supply chain network that are out of control (Sample 260, 285, 292, 316, 365, 426, 427, 460, 461, 464, 467, 468, 537, 538). By analyzing the data, we observed that three products were repeatedly identified as inefficient in the outputs. Therefore, these items were eliminated for further examination. The performance of the remaining eleven sample products should be improved.

The correlation chart of each of the two response variables versus the significant variables is shown in FIGURE . In this figure, there is a correlation between the explanatory variables and sale and production. Therefore, the existence of a relationship between the explanatory variables and Y_1 and Y_2

is intuitively confirmed. In the next step, we used this interpretation for applying Joint Optimization plot in rendering optimal values.

--Insert FIGURE 4 Approximately Here--

4.2 Applying Joint Optimization to Render Optimal Values

The analysis of the "Hotelling T^2 Control Chart" revealed that 11 products within the supply chain network were experiencing inefficiencies and were out-of-control. To gain a deeper understanding of these problematic products, an investigation was conducted using the *Profit and Loss (P&L)* statements. The P&L statements indicated that each of the 11 out-of-control products incurred losses in at least one month during this 12-month interval. Consequently, transforming these losses into profitable outcomes is crucial for improving the overall performance of the supply chain network.

Moreover, after a meticulous analysis of the data and the influence of variables on the supply chain during preceding stages, a clear revelation emerged: the storage of raw materials (I, II, and III) and the shelf price are identified as critical variables, bearing a substantial impact on the company's operations, performance, and overall success.

Given that these factors can be controlled by the company, there is an opportunity for their enhancement. To do so, the Joint Optimization modeling was employed to derive new decision variables that boost the company's performance in generating profits.

The simultaneous optimization plot illustrates the results of optimizing multiple responses graphically. In this plot, the optimization of control factors (variables) is shown in one graph, and the corresponding estimated responses aligned with the desired optimum are displayed in another graph (Kuhnt, 2004). As an example, consider the Product 1 (FIGURE) which incurred losses in the fifth month of sales. By applying the Joint Optimization model, careful adjustments and fine-tuning were made to the values of critical variables throughout each twelve-month interval. These optimized decision variables were precisely customized to ensure profitability rather than loss.

As a result, effectively managing the storage of raw materials and setting appropriate shelf prices can lead to greater cost efficiency, improved revenue streams, and overall success in the competitive market landscape. FIGURE shows results of the joint optimization plot:

--Insert FIGURE 5 Approximately Here--

Optimal results for x_4 , x_5 , x_6 , and x_{19} were obtained, and the operational profit was recalculated to assess improvements since the last performance evaluation of these 11 products in the supply chain network. The **Error! Reference source not found.** summarizes the performance of these products based on the new optimal values for the determined variables, comparing their previous operational profit to the optimized one:

--Insert Table 5 Approximately Here--

Regarding the outputs, we can conclude that effectively managing the storage of raw materials and setting appropriate shelf prices can lead to cost efficiency, improved revenue streams, and overall success in the competitive market landscape. Emphasizing this approach as part of the company's supply chain strategy is crucial for sustainable growth and continued success.

5 Conclusion

The supply chain analytics can enhance supply visibility and improve forecasting, lead to lower inventory levels and cost savings, and increase overall efficiency. The primary contribution of this paper is to enhance multiple facets of the supply chain, encompassing storage levels and optimal shelf prices, to maximize profits. This is achieved through the utilization of statistical methodologies that have not been extensively applied before in the literature. Unlike previous studies focusing on variables from one or two sections of supply chain networks, this study implements a comprehensive analysis across all five sections of the supply chain network. By examining variables across all sections, the study identifies the most significant ones to determine optimal values for these important variables, thereby enhancing the overall performance of the supply chain network. The method involves: 1) extracting data from the five main sections of the supply chain over a twelve-month interval, 2) utilizing the GEE method along with Hotelling T^2 control chart to monitor product performance, 3) detecting any unusual or out-of-control behavior in the supply chain, 4) applying Joint Optimization modeling to the products that exhibited out-of-control behavior during the supply chain monitoring, 5) optimizing relevant variables derived from previous stages to find the optimal cost values, including the cost of goods and operational expenses.

This study examined a real-world Multi-Echelon Supply Chain Network with 51 personal care products to provide prescriptive analytics for enhancing its performance. Data were collected across all five supply chain sections for a 12-month period based on 22 variables from the entire chain, from production to distribution centers and store shelves.

As a result, the products exhibiting inefficiencies within the supply chain network undertake substantial improvement through the optimization of prices and the selection of optimal storage levels for raw materials during manufacturing which ensures profitability and enhances the company's supply chain performance. This strategic method led to a profit in their financial statement within a period that had previously incurred losses.

For future studies, there is an opportunity to employ the Generalized Linear Model (GLM) in conjunction with a multivariate control chart capable of handling outliers and data with non-normal distributions, which are common characteristics in supply chain data.

References

Akkermans, H. and Van Wassenhove, L.N. (2018). Supply Chain Tsunamis: Research on Low-Probability, High-Impact Disruptions. *Journal of Supply Chain Management*, 54: 64-76. DOI: https://doi.org/10.1111/jscm.12162

Amellal, A., Amellal, I., & Seghiouer, H., & Ech-Charrat, M. (2023). Improving Lead Time Forecasting and Anomaly Detection for Automotive Spare Parts with A Combined CNN-LSTM Approach. *Operations and Supply Chain Management: An International Journal*, 16(2), 265-278. DOI: http://doi.org/10.31387/oscm0530388

Bersimis, S., Psarakis, S. and Panaretos, J. (2007). Multivariate Statistical Process Control Charts: An Overview. *Quality and Reliability Engineering International*, 23: 517-543. DOI: https://doi.org/10.1002/qre.829

Chen, B. (2021). Data-Driven Inventory Control with Shifting Demand. *Production and Operations Management*, 30: 1365-1385. DOI: https://doi.org/10.1111/poms.13326

Carrière I, Bouyer J. (2002). Choosing Marginal or Random-Effects Models for Longitudinal Binary Responses: Application to Self-Reported Disability Among Older Persons. *BMC Medical Research Methodology*. 2(1):15. DOI: http://dx.doi.org/10.1186/1471-2288-2-15

Chiang, D. M.-H.; Lin, C.-P.; Chen, M.-C. (2011). The Adaptive Approach for Storage Assignment by Mining Data of Warehouse Management System for Distribution Centers. *Enterprise Information Systems* 5, 219−234. DOI: http://dx.doi.org/10.1080/17517575.2010.537784

Choi, Y., Lee, H. and Irani, Z. (2016). Big Data-Driven Fuzzy Cognitive Map for Prioritizing IT Service Procurement in the Public Sector. *Annals of Operations Research*, *Springer US*, Vol. 243 No. 1–2, pp. 1–30. DOI: http://dx.doi.org/10.1007/s10479-016-2281-6

Chong, A. Y. L.; Li, B.; Ngai, E. W.T.; Ch'ng, E.; Lee, F. (2016). Predicting Online Product Sales via Online Reviews, Sentiments, and Promotion Strategies. *International Journal of Operations & Production Management 36*, 358−383. DOI: http://dx.doi.org/10.1108/IJOPM-03-2015-0151

Chopra, S.; Sodhi, M. S. (2004). Managing Risk to Avoid Supply-Chain Breakdown. *MIT Sloan Management Review 46*, 53−62. DOI: https://api.semanticscholar.org/CorpusID:150827016

Diggle P, Heagerty P, Liang K-Y, Zeger S. (2002). Analysis of Longitudinal Data: *Oxford University Press*. DOI: http://dx.doi.org/10.2307/2291352

Ethirajan, M., Arasu, M. T., Kandasamy, J., K.E.K., V., Nadeem, S. P., & Kumar, A. (2021). Analysing the Risks of Adopting Circular Economy Initiatives in Manufacturing Supply Chains. *Business Strategy and the Environment, Wiley Blackwell*, Vol. 30(1). 204–236 DOI: https://doi.org/10.1002/bse.2617

Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis: John Wiley & Sons; 2012.

Gómez, M. I., & Lee, D. (2023). Transforming Food Supply Chains for Sustainability. *Journal of Supply Chain Management*, 59(4), 79–92. DOI: https://doi.org/10.1111/jscm.12310

Gilvan C. Souza, (2014). Supply Chain Analytics, *Business Horizons*, Volume 57, Issue 5, Pages 595–605, ISSN 0007–6813, DOI: https://doi.org/10.1016/j.bushor.2014.06.004.

Jain, R.; Singh, A. R.; Yadav, H. C.; Mishra, P. K. (214). Using Data Mining Synergies for Evaluating Criteria at Pre-qualification Stage of Supplier Selection. *Journal of Intelligent Manufacturing 25*, 165−175. DOI: http://dx.doi.org/10.1007/s10845-012-0684-z

Wang, J., Swartz, C. L. E., Corbett, B., & Huang, K. (2020). Supply Chain Monitoring Using Principal Component Analysis. *Industrial & Engineering Chemistry Research*, 59(27), 12487-12503. DOI: https://doi.org/10.1021/acs.iecr.0c01038

Wang, J., Swartz, C. L. E., & Huang, K. (2023). Data-Driven Supply Chain Monitoring Using Canonical Variate Analysis. *Computers & Chemical Engineering*. Volume 174, 108228, ISSN 0098-1354. DOI: https://doi.org/10.1016/j.compchemeng.2023.108228

Kapil, M., Gumte., Priyanka, Devi, Pantula., Srinivas, Soumitri, Miriyala., Kishalay, Mitra. (2021). Data Driven Robust Optimization for Handling Uncertainty in Supply Chain Planning Models. *Chemical Engineering Science*, 246:116889-. DOI: https://doi.org/10.1016/j.ces.2021.116889

Khajehzadeh, M., Pazhuheian, F., Seifi, F., Noorossana, R., Ali, A., & Saeedi, N. (2022). Analysis of Factors Affecting Product Sales with an Outlook Toward Sale Forecasting in Cosmetic Industry Using Statistical Methods. *International Review of Management and Marketing*; Mersin Vol. 12, Iss. 6, (2022): 55-63. DOI: https://doi.org/10.32479/irmm.13337

Kim, B. and Oh, H. (2005). The Impact of Decision-Making Sharing Between Supplier and Manufacturer on Their Collaboration Performance. *Supply Chain Management: An International Journal*, Vol. 10, No. 3, pp.223–236. DOI: https://doi.org/10.1108/13598540510606287

Kovács, G. and Falagara Sigala, I. (2021). Lessons Learned from Humanitarian Logistics to Manage Supply Chain Disruptions. *Journal of Supply Chain Management*, 57: 41-49. DOI: https://doi.org/10.1111/jscm.12253

Kuhnt, S., & Erdbruegge, M. (2004). A Strategy of Robust Parameter Design for Multiple Responses. *Statistical Modelling*, 4(4), 249-264. DOI: https://doi.org/10.1191/1471082X04st078oa

Kuhnt, S., & Rudak, N. (2013). Simultaneous Optimization of Multiple Responses with the R Package JOP. *Journal of Statistical Software* 54.9: 1-23. DOI: https://doi.org/10.18637/jss.v054.i09

Li, L.; Su, X.; Wang, Y.; Lin, Y.; Li, Z.; Li, Y. (2015). Robust Causal Dependence Mining in Big Data Network and Its Application to Traffic Flow Predictions. *Transportation Research Part C: Emerging Technologies* 58, 292−307. DOI: https://doi.org/10.1016/j.trc.2015.03.003

Manhart, P., Summers, J.K. and Blackhurst, J. (2020). A Meta-Analytic Review of Supply Chain Risk Management: Assessing Buffering and Bridging Strategies and Firm Performance. *Journal of Supply Chain Management*, 56: 66-87. DOI: https://doi.org/10.1111/jscm.12219

Khurana, M., & Kumar, D. (2017). The Study of Data Analytics in Inventory Management. *International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS)*, Dubai, United Arab Emirates, 2017, Pp. 140-144. DOI: https://doi.org/10.1109/ICTUS.2017.8285993

Mohan, B, A., Harshavardhan, B., Karan, S., Mohammed, Jawaad, Shariff., Pranav, M, G. (2021). Demand Forecasting and Route Optimization in Supply Chain Industry Using Data Analytics. *Asian Conference on Innovation in Technology (ASIANCON)*, PUNE, India, 2021, pp. 1-7, DOI: https://doi.org/10.1109/ASIANCON51346.2021.9544942

Montgomery, D. C. (2019). Introduction to Statistical Quality Control (8th ed.). *John Wiley & Sons*. ISBN: 978-1-119-39930-8.

Mori, J.; Kajikawa, Y.; Kashima, H.; Sakata, I. (2012). Machine Learning Approach for Finding Business Partners and Building Reciprocal Relationships. *Expert Systems with Applications*, 39, 10402−10407. DOI: https://doi.org/10.1016/j.eswa.2012.01.202

Nahum, Méndez-Sánchez. (2023). Machine Learning and Data Analytics, in Smart and Sustainable Operations and Supply Chain Management in Industry 4.0. *Taylor & Francis CRC Pres*, 307-321. DOI: 10.1201/9781003180302-15

Nikookar, E., Stevenson, M., & Varsei, M. (2024). Building an Antifragile Supply Chain: A Capability Blueprint for Resilience and Post-disruption Growth. *Journal of Supply Chain Management*, 60(1), 13–31. DOI: https://doi.org/10.1111/jscm.12313

Nitin, K., Singh, K.-h., Lai, K.-h., & Zhang, J. (2023). Intellectual Core in Supply Chain Analytics: Bibliometric Analysis and Research Agenda. *International Journal of Information Technology and Decision Making*, 1-29. Doi: https://doi.org/10.1142/S0219622023300021

Nguyen, T. (2023). Applications of Artificial Intelligence for Demand Forecasting. *Operations and Supply Chain Management: An International Journal*, 16(4), 424-434. http://doi.org/10.31387/oscm0550401

Sukha, C. S., & Prabhu, S. (2023). Framework in Supply Chain Using Inventory Management, Big Data and Inventory Analytics. *IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT),* Bhopal, India, 2023, pp. 549-553, DOI: https://doi.org/10.1109/CSNT57126.2023.10134627

Pignatiello Jr, J. J. (1993). Strategies for Robust Multiresponse Quality Engineering. *IIE transactions*, 25(3), 5-15, DOI: https://doi.org/10.1080/07408179308964286

Salehan, M.; Kim, D. J. (2015). Predicting the Performance of Online Consumer Reviews: A Sentiment Mining Approach to Big Data Analytics. *Decision Support Systems*, 81, 30−40. DOI: https://doi.org/10.1016/j.dss.2015.10.006

Sodhi, M.S. and Tang, C.S. (2021). Supply Chain Management for Extreme Conditions: Research Opportunities. *Journal of Supply Chain Management*, 57: 7-16. DOI: https://doi.org/10.1111/jscm.12255

Subhakanta, Dash., S., P., Mohanty. (2018). Uncertain Transportation Model with Rough Unit Cost, Demand and Supply. *Operations Research*, 55(1):1-13. DOI: 10.1007/S12597-017-0317-6

Suwignjo, P., Panjaitan, L., & Baihaqy, A., & Rusdiansyah, A. (2023). Predictive Analytics to Improve Inventory Performance: A Case Study of an FMCG Company. *Operations and Supply Chain Management: An International Journal*, 16(2), 293–310. DOI: http://doi.org/10.31387/oscm0530390

Tang, C. S. (2006). Perspectives in Supply Chain Risk Management. *International Journal of Production Economics*, 103, 451−488. DOI: http://dx.doi.org/10.1016/j.ijpe.2005.12.006

Tracy, N. D.; Young J. C.; Mason, R. L. (1992). Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, v. 24, n. 2, p. 88-95. DOI: https://api.semanticscholar.org/CorpusID:123054505

Nguyen, T., Zhou, L., Spiegler, V., Ieromonachou, P., & Lin, Y. (2018). Big Data Analytics in Supply Chain Management: A State-of-the-Art Literature Review. *Computers & Operations Research*, 98, 254-264. DOI: https://doi.org/10.1016/j.cor.2017.07.004

Wang, Y. (2010). Imprecise Probabilities Based on Generalized Intervals for System Reliability Assessment. *International Journal of Reliability & Safety*, Vol. 4, No. 4, pp.319–342. DOI: http://dx.doi.org/10.1504/IJRS.2010.035572

Wu, L., Liu, W., Yi, G. Y., & Huang, Y. (2012). Analysis of Longitudinal and Survival Data: Joint Modeling, Inference Methods, and Issues. *Journal of Probability and Statistics*, Volume 2012, 640153. DOI: https://doi.org/10.1155/2012/640153

Zhang, Y., Ren, S., Liu, Y., & Si, S. (2017). A Big Data Analytics Architecture for Cleaner Manufacturing and Maintenance Processes of Complex Products. *Journal of Cleaner Production*, 142(2), 626-641. DOI: https://doi.org/10.1016/j.jclepro.2016.07.123

Zhao, R.; Liu, Y.; Zhang, N.; Huang, T. (2016). An Optimization Model for Green Supply Chain Management by Using a Big Data Analytic Approach. *Journal of Cleaner Production*, 142, 1085−1097. DOI: http://dx.doi.org/10.1016/j.jclepro.2016.03.006

Ziegler, A., Kastner, C., & Blettner, M. (1998). The Generalized Estimating Equations: An Annotated Bibliography. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 40(2), 115–139. DOI: https://doi.org/10.1002/(SICI)1521-4036(199806)40:2%3C115::AID-BIMJ115%3E3.0.CO;2-6

Appendices (Tables and Figures):

TABLE 1

RELEVANT LITERATURE REVIEWS

Reference	Method	Results
Chiang et al., 2011	Data mining-based storage assignment approach	Enhancing the efficiency of order picking through an association index in warehousing.
Mori et al., 2012	Support vector machine and logistic regression	Building a prediction model for customer-supplier relationships to identify potential business partners.
Jain et al., 2014	Data mining approach	Uncovering hidden relationships for suppliers' selection and optimizing the supplier selection process.
Li et al., 2015	Lasso Granger causality models	Building a traffic prediction model with a balance between complexity and performance.
Choi et al., 2016	Fuzzy cognitive maps (FCM)	Enhancing decision-making in IT service procurement for the public sector through data analytics approach.
Salehan and Kim, 2015	Sentiment mining approach	Studying predictors of online consumer review performance.
Chong et al., 2016	Neural network approach	Exploring significant variables influencing product sales using a neural network.
Zhang et al., 2017	Data-based analytics for Product Lifecycle	Leveraging data analytics and service-driven patterns for product lifecycle management.
Khurana & Kumar, 2017	Linear discriminant analysis	Implementing data analytics in inventory management to find dependencies.
Zhao et al., 2016	Extracting uncertain parameters for green supply chain	Redesigning a green supply chain using historical data.
Subhakanta & Mohanty, 2018	Deterministic linear programming model	Addressing a transportation problem with uncertain costs and demands using deterministic linear programming.
Jing Wang et al., 2020 and Kapil et al., 2021	Data-driven optimization techniques	Enhancing SCM in uncertain environments through machine learning and data analytics.
Chen, 2021	Lost-sales inventory control problem with active exploration	Addressing an inventory control problem with active exploration and uncertainty in demand distributions.
Mohan et al., 2021	Time series analysis and simulated annealing algorithm	Devising strategies for demand forecasting and route optimization in SCM
Singh et al, 2023	Bibliometric statistical analysis on SCA	Identifying key research themes for SCM and providing a foundation for future studies in this area
Nahum, Méndez-Sánchez, 2023	Machine Learning for predicting electricity consumption	Demonstrating the applicability of data analytics method in the context of SCM for accurate prediction
Nguyen, T, 2023	AI-based techniques for demand forecasting in SCM: A review of 2013-2023 literature.	Demonstrating prominent trend in application of hybrid methods for demand forecasting in SCM and challenges of AI-based methods in selecting suitable inputs.
Suwignjo , P et al., 2023	Gradient boosting model	Predicting the occurrence of inventory overstock and understock by using classification and regression models.
Amellal et al, 2023	a hybrid model combining LSTM and CNN	Overcoming the issue of lead time information in SCM for the logistics and distribution network section.

TABLE 2

VARIABLES OF THE SUPPLY CHAIN

Variables	Indexes	Layers (components)
Consumption of first raw material (per unit of kg)	x_1	Manufacturing
Consumption of second raw material (per unit of kg)	x_2	
Consumption of third raw material (per unit of kg)	x_3	
Storage of raw material (I)	x_4	Warehousing
Storage of raw material (II)	x_5	
Storage of raw material (III)	x_6	
First raw material purchasing price	x_7	Procurement
Second raw material purchasing price	x_8	
Third raw material purchasing price	x_9	
First raw material inputs subsidy	x_{10}	
Second raw material inputs subsidy	x_{11}	
Third raw material inputs subsidy	x_{12}	
Firs raw material Purchasing Quantity	x_{13}	
Second raw material Purchasing Quantity	x_{14}	
Third raw material Purchasing Quantity	x_{15}	
Freight Charges (suppliers to factory)	x_{16}	Logistic/transportation
Transportation Cost (factory to distributors)	x_{17}	
Order Size (distributors)	x_{18}	
Shelf Price	x_{19}	Demand
Currency exchange	x_{20}	
Consumer price index	x_{21}	
Inflation Rate	x_{22}	

TABLE 3

RESULTS FOR PRODUCTION

Production	Coef.	Std. Err.	z	$P > \|z\|$	[95% Conf. Interval]	
x_1	97593.59	222695	0.44	0.661	-338881	534067 .7
x_2	-237101.3	524465.7	-0.45	0.651	-1265035	790832.6
x_3	-1534.438	3919.006	-0.39	0.695	-9215.55	6146.673
x_4	0.221107	0.012325	17.9	0	0.196951	0.245263
x_5	0.14155	0.01231	11.5	0	0.117423	0.165678
x_6	1.896452	0.137458	13.8	0	1.62704	2.165865
x_7	0.014487	0.033645	0.43	0.667	-0.05146	0.08043
x_8	0.006707	0.015286	0.44	0.661	-0.02325	0.036667
x_9	0.050431	0.114288	0.44	0.659	-0.17357	0.274432
x_{10}	2.629571	5.923298	0.44	0.657	-8.97988	14.23902
x_{11}	-2.462099	5.551389	-0.44	0.657	-13.3426	8.418423
x_{12}	-0.265546	0.613631	-0.43	0.665	-1.46824	0.937148
x_{13}	0.001009	0.001019	0.99	0.322	-0.00099	0.003006
x_{14}	0.000255	0.000204	1.25	0.211	-0.00015	0.000655
x_{15}	-0.012652	0.005155	-2.45	0.014	-0.02276	-0.00255
x_{16}	-239563.6	545887	-0.44	0.661	-1309482	830355.3
x_{17}	400218.4	897608.2	0.45	0.656	-1359061	2159498
x_{18}	0	(omitted)				
x_{19}	-0.00213	0.000782	-2.72	0.006	-0.00366	-0.0006
x_{20}	0	(omitted)				
x_{21}	0	(omitted)				
x_{22}	0	(omitted)				
_cons	-33881.01	103491	-0.33	0.743	-236720	168957.7

TABLE 4

RESULTS FOR SALE

Sale	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
x_1	252562.3	225174.2	-1.12	0.262	-693896	188770.9
x_2	-605025	530300.7	1.14	0.254	-434345	1644396
x_3	-4668.31	3962.623	1.18	0.239	-3098.28	12434.91
x_4	0.192825	0.012238	-15.8	0	-0.21681	-0.16884
x_5	0.098772	0.012161	-8.12	0	-0.12261	-0.07494
x_6	1.73389	0.134552	-12.9	0	-1.99761	-1.47017
x_7	0.037457	0.03402	-1.1	0.271	-0.10414	0.029221
x_8	0.017453	0.015456	-1.13	0.259	-0.04775	0.012841
x_9	0.131243	0.115561	-1.14	0.256	-0.35774	0.095253
x_{10}	6.742258	5.989259	-1.13	0.26	-18.481	4.996474
x_{11}	-6.31508	5.613214	1.13	0.261	-4.68662	17.31677
x_{12}	-0.69629	0.620463	1.12	0.262	-0.51979	1.912379
x_{13}	-0.00048	0.001027	0.47	0.638	-0.00153	0.002494
x_{14}	0	(omitted)				
x_{15}	0	(omitted)				
x_{16}	-610217	551967.5	1.11	0.269	-471620	1692053
x_{17}	1028013	907600.1	-1.13	0.257	-2806877	750850.1
x_{18}	0.000448	0.000206	-2.18	0.03	-0.00085	-4.4E-05
x_{19}	-0.00143	0.000738	1.94	0.043	-1.7E-05	0.002876
x_{20}	0	(omitted)				
x_{21}	-0.02321	0.005192	4.47	0	0.013034	0.033385
x_{22}	0	(omitted)				
_cons	-99128.2	104646.4	0.95	0.344	-105975	304231.3

TABLE 5

OPERATIONAL PROFIT FOR INEFFICIENT PRODUCTS

Product Number	Previous Operation Profit (rial)	Optimized Operation Profit (rial)	Deviation (rial)
1	(373,381)	25,619	399,000
2	(31,877)	45,723	77,600
3	(70,080)	2,640	72,720
4	(127,859)	182,141	310,000
5	(458,744)	6,256	465,000
6	(15,948)	216,552	232,500
7	(178,771)	13,529	192,300
8	(134,055)	28,245	162,300
9	(217,158)	118,856	336,014
10	(130,906)	32,994	163,900
11	(96,140)	88,000	184,140

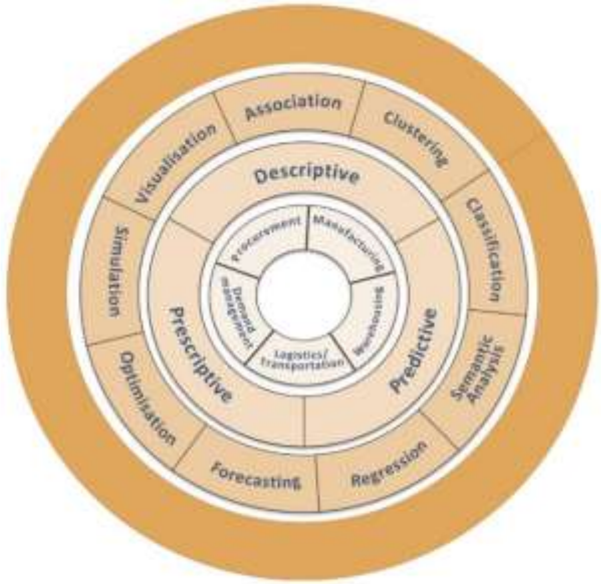


FIGURE 1

STUDIES CLASSIFICATION FRAMEWORK (Nguyen et al, 2017)

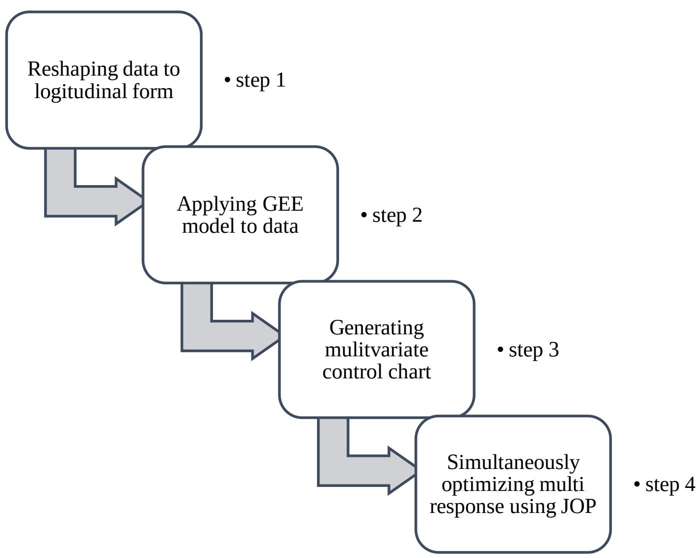


FIGURE 2

STEPS OF THE METHOD

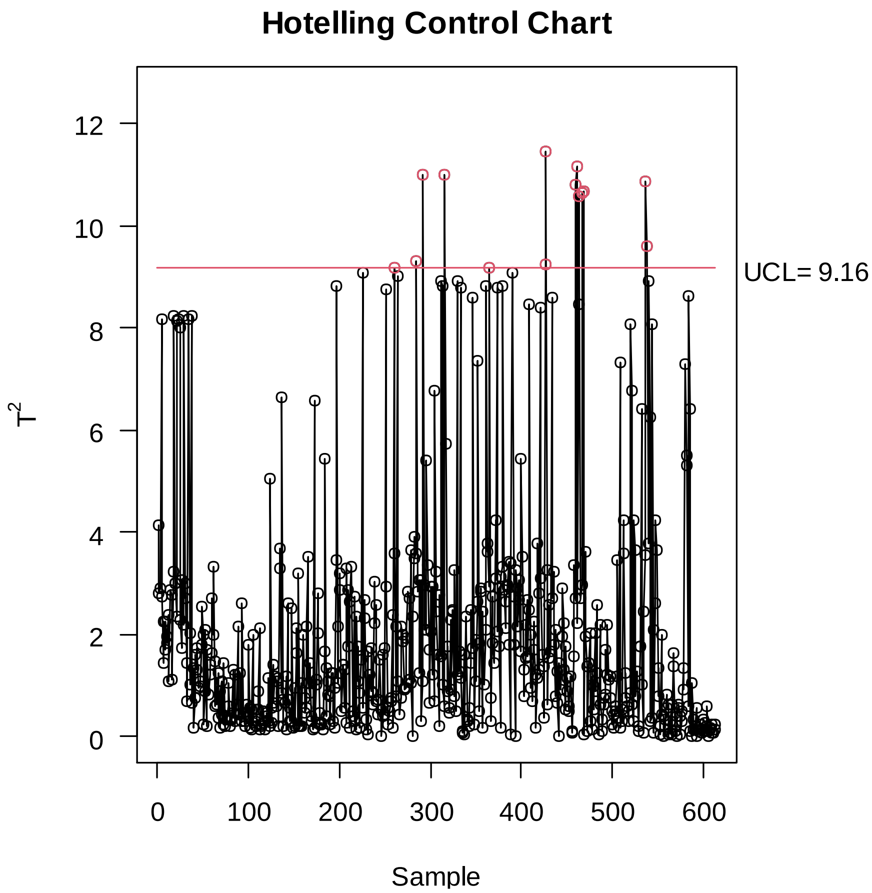


FIGURE 3

THE HOTELLING CONTROL CHART

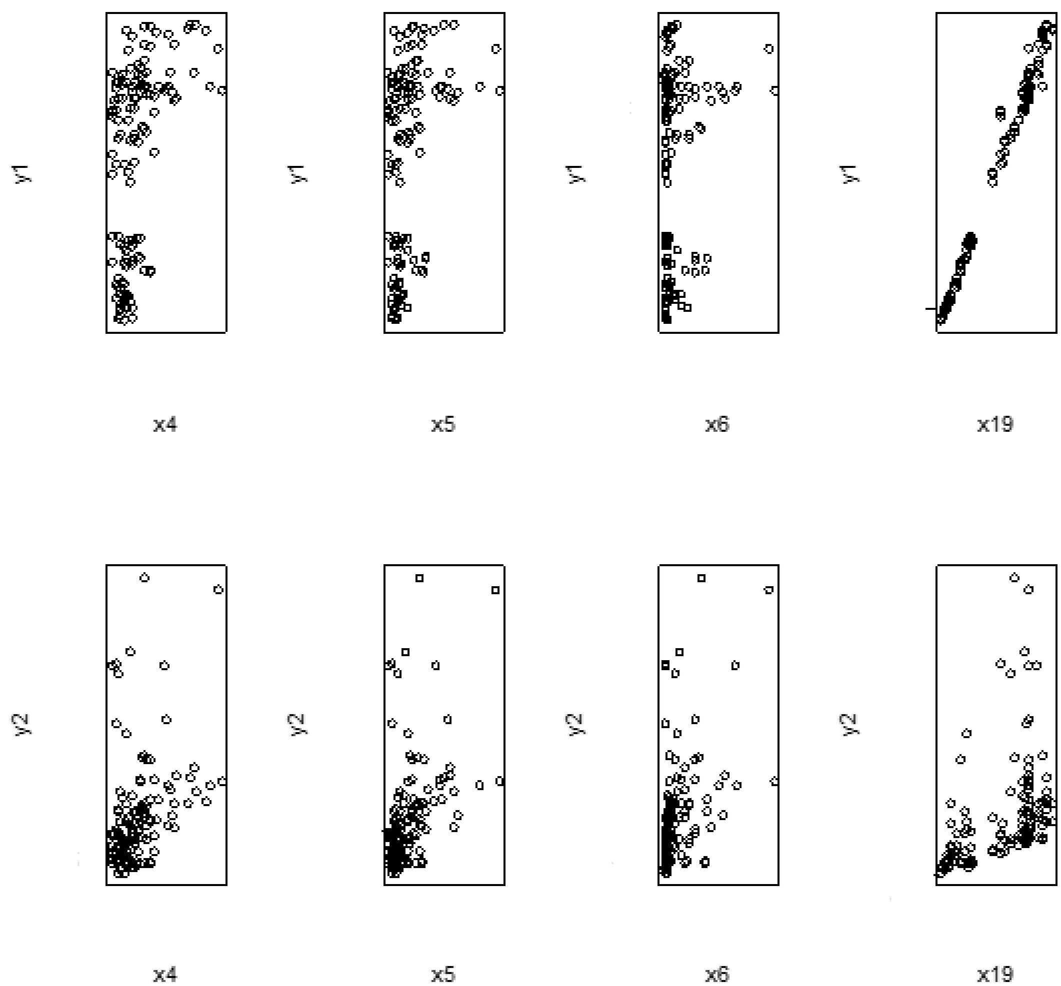
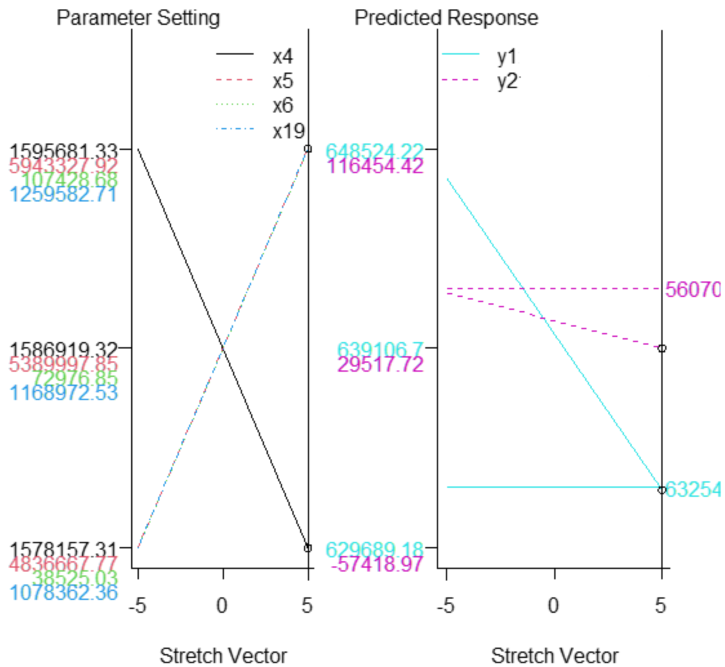


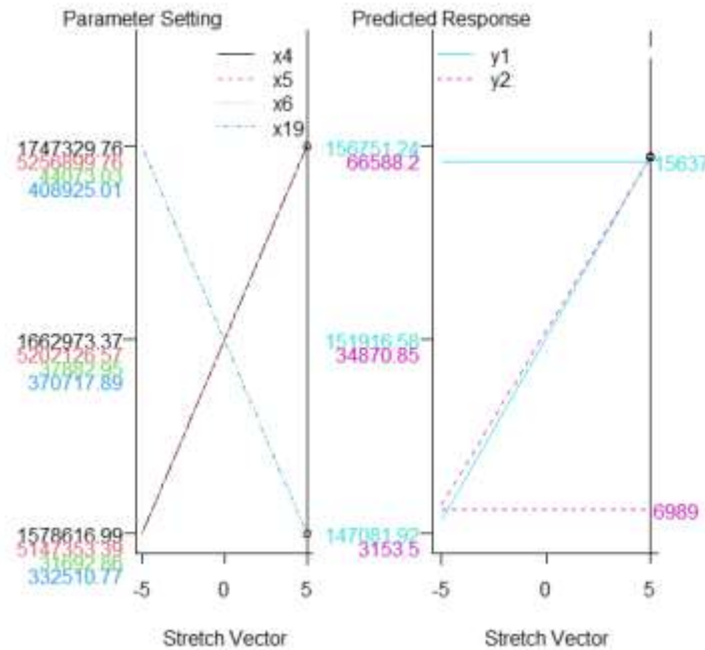
FIGURE 4

CORRELATION CHART OF SIGNIFICANT VARIABLES VERSUS RESPONSE VARIABLES

Product 1



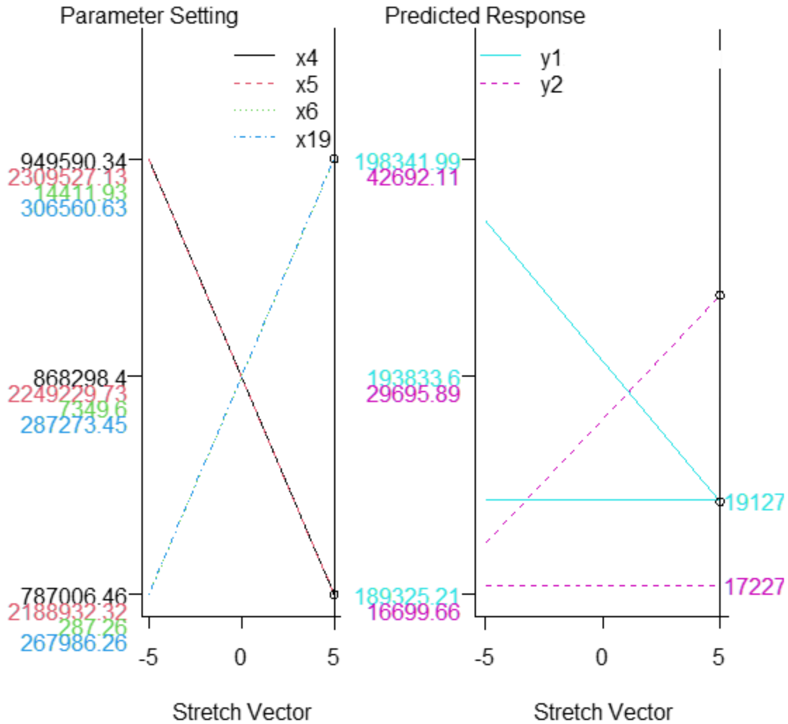
Product 2



Optimal Values of Variables				Optimal Values of Responses	
X_4	X_5	X_6	X_{19}	Y_1	Y_2
1578157.3	5943327.9	107428.7	1259582.7	632451.62	29517.72

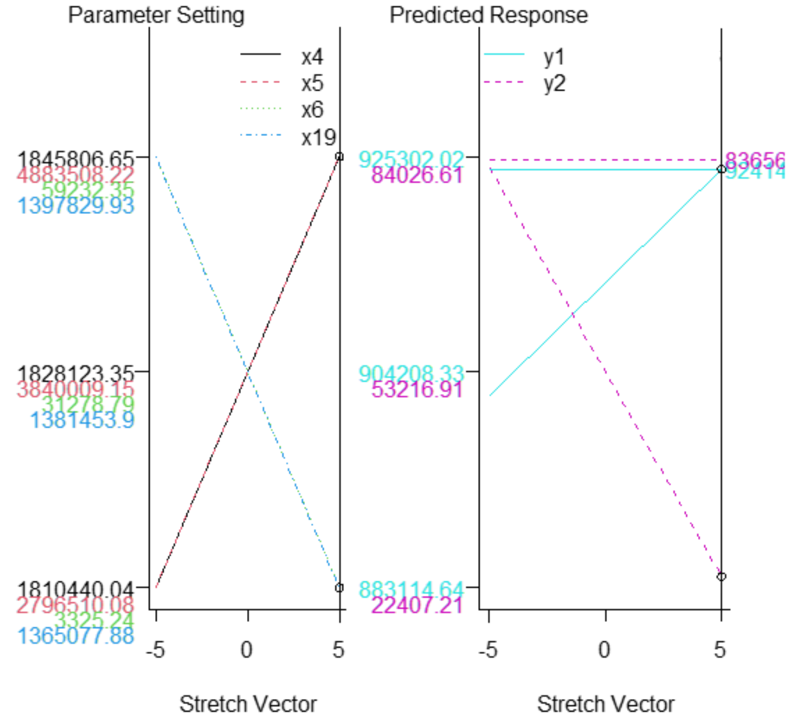
Optimal Values of Variables				Optimal Values of Responses	
X_4	X_5	X_6	X_{19}	Y_1	Y_2
1747329.76	5256899.76	31692.86	332510.77	156518.69	64864.55

Product 3



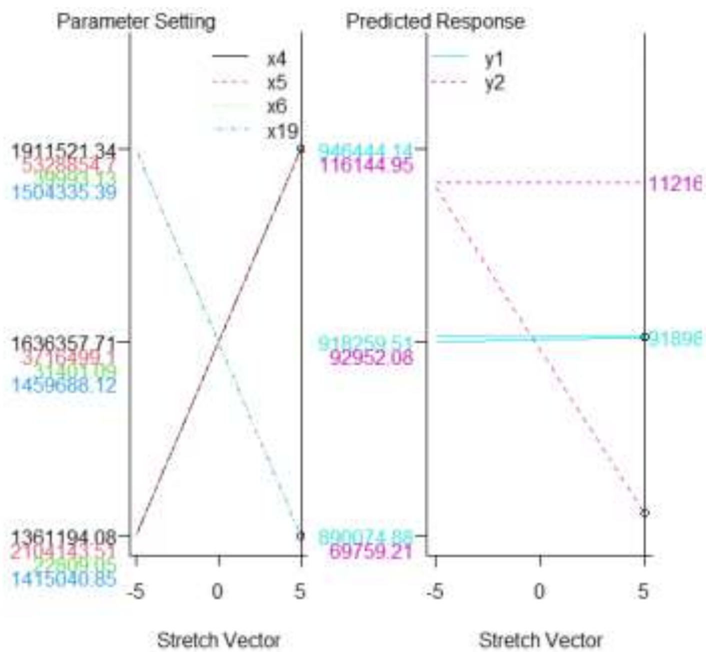
Optimal Values of Variables				Optimal Values of Responses	
X_4	X_5	X_6	X_{19}	Y_1	Y_2
787006.46	2188932.32	14411.93	306560.63	191246.33	34566.37

Product 4

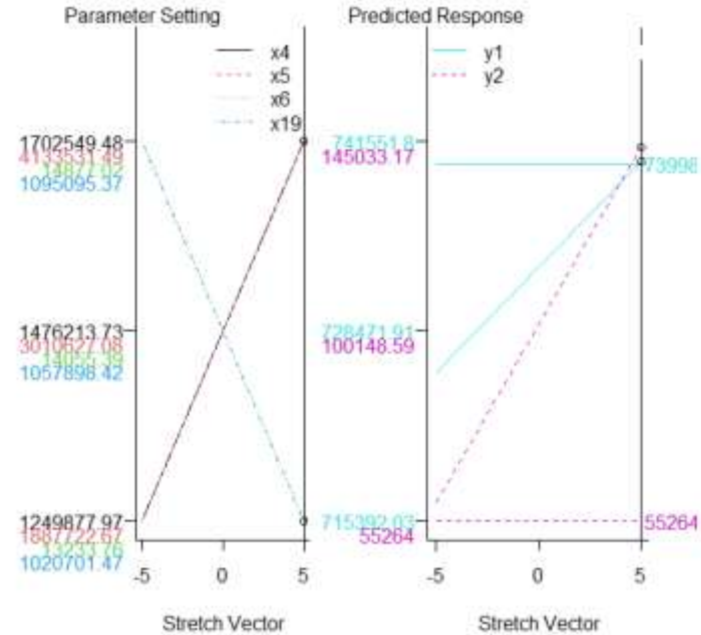


Optimal Values of Variables				Optimal Values of Responses	
X_4	X_5	X_6	X_{19}	Y_1	Y_2
1845806.651	4883508.222	3325.239	1365077.877	924085.20	24034.37

Product 5



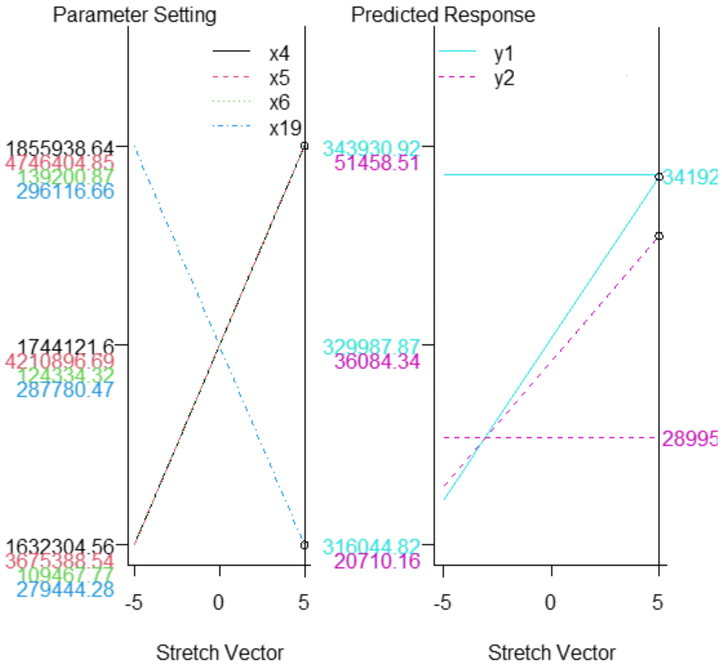
Product 6



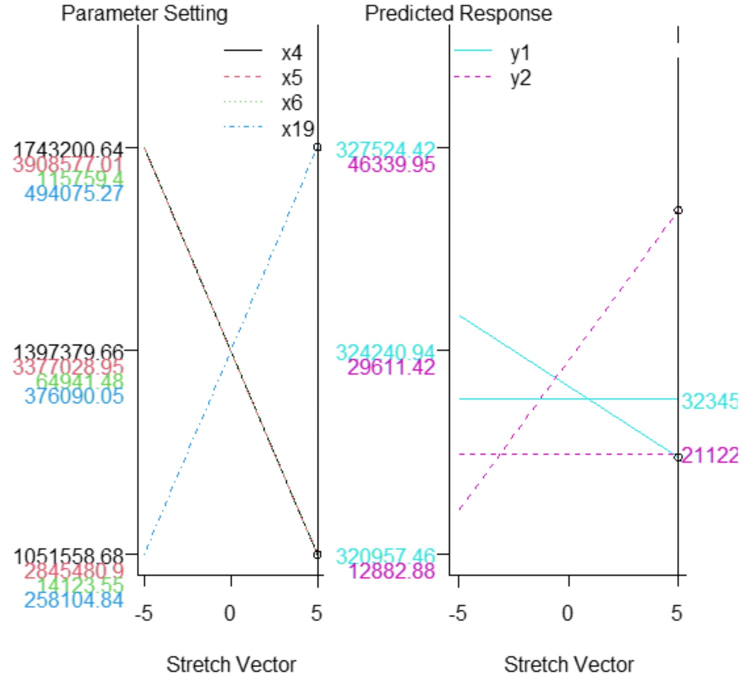
Optimal Values of Variables				Optimal Values of Responses	
X_4	X_5	X_6	X_{19}	Y_1	Y_2
1911521.34	5328854.70	22809.05	1415040.85	919001.22	72457.67

Optimal Values of Variables				Optimal Values of Responses	
X_4	X_5	X_6	X_{19}	Y_1	Y_2
1702549.48	4133531.49	13233.76	1020701.47	740124.7	143583.7

Product 7



Product 8



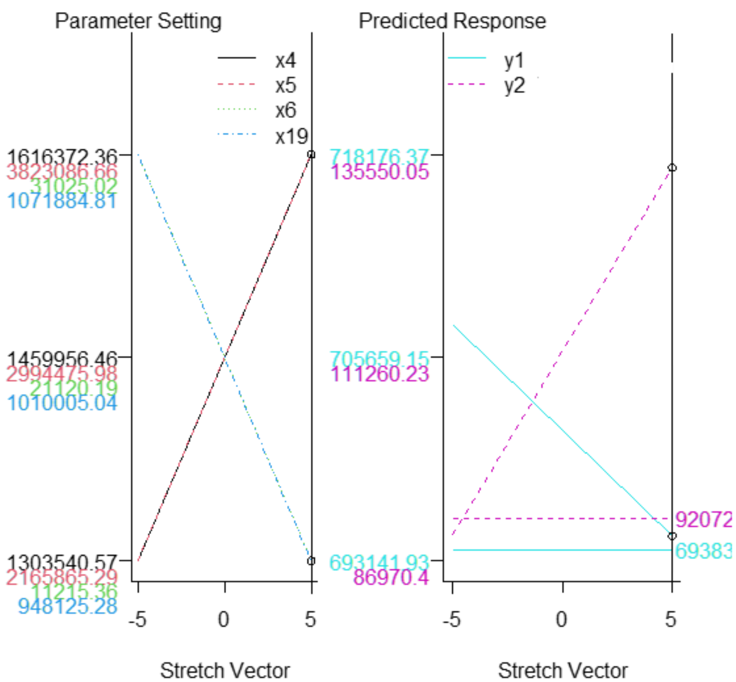
Product 7: Optimal Values of Variables / Optimal Values of Responses

X_4	X_5	X_6	X_{19}	Y_1	Y_2
1855938.6	4746404.8	139200.9	279444.3	341750.14	44521.19

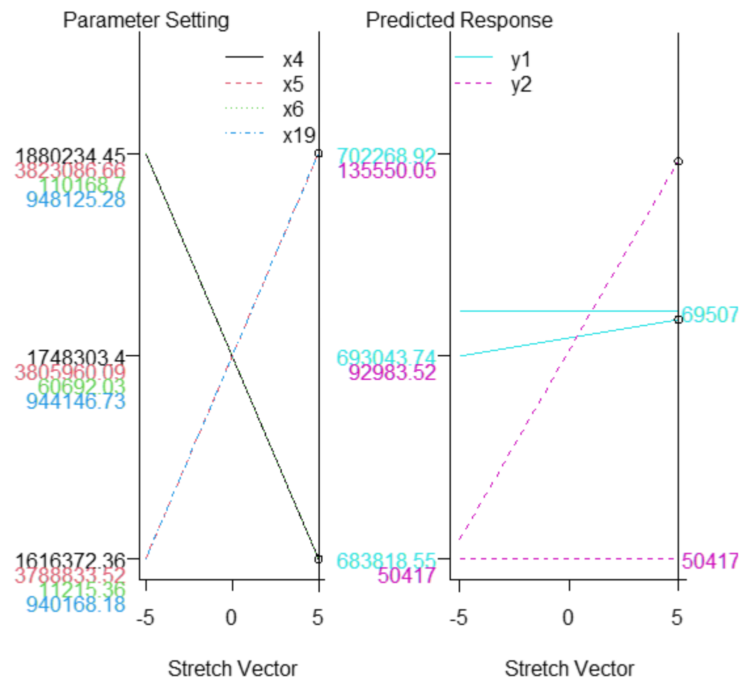
Product 8: Optimal Values of Variables / Optimal Values of Responses

X_4	X_5	X_6	X_{19}	Y_1	Y_2
1051558.68	2845480.90	14123.55	494075.27	494075.27	41172.41

Product 9



Product 10



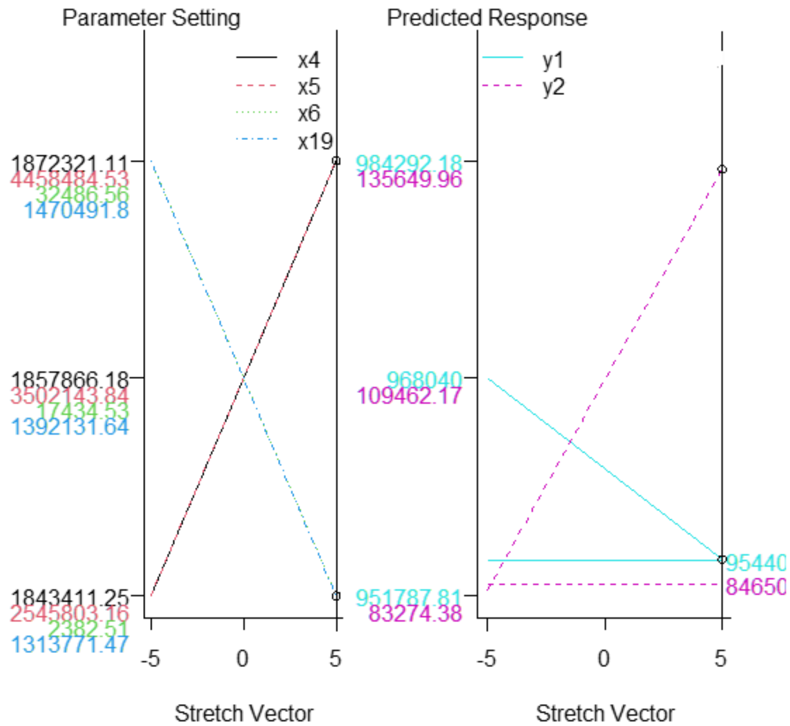
Optimal Values of Variables / Optimal Values of Responses

X_4	X_5	X_6	X_{19}	Y_1	Y_2
1616372.36	3823086.66	11215.36	948125.28	694692.1	134070.2

Optimal Values of Variables / Optimal Values of Responses

X_4	X_5	X_6	X_{19}	Y_1	Y_2
1616372.36	3823086.66	11215.36	948125.28	694692.1	134070.2

Product 11



Optimal Values of Variables Optimal Values of Responses

X_4	X_5	X_6	X_{19}	Y_1	Y_2
1872321.108	4458484.526	2382.511	1313771.473	954517.5	134657.9

FIGURE 5

JOINT OPTIMIZATION PLOTS