

Selecting Ordering Policy and Items Classification Based on Canonical Correlation and Cluster Analysis

Keisuke Nagasawa*, Takashi Irohara

Department of Information and Communication Sciences, Sophia University, Tokyo, Japan

Yosuke Matoba, Shuling Liu

Fairway Solutions Inc., Tokyo, Japan

(Received: February 10, 2012 / Revised: April 17, 2012 / Accepted: April 24, 2012)

ABSTRACT

It is difficult to find an appropriate ordering policy for a many types of items. One of the reasons for this difficulty is that each item has a different demand trend. We will classify items by shipment trend and then decide the ordering policy for each item category. In this study, we indicate that categorizing items from their statistical characteristics leads to an ordering policy suitable for that category. We analyze the ordering policy and shipment trend and propose a new method for selecting the ordering policy which is based on finding the strongest relation between the classification of the items and the ordering policy. In our numerical experiment, from actual shipment data of about 5,000 items over the past year, we calculated many statistics that represent the trend of each item. Next, we applied the canonical correlation analysis between the evaluations of ordering policies and the various statistics. Furthermore, we applied the cluster analysis on the statistics concerning the performance of ordering policies. Finally, we separate items into several categories and show that the appropriate ordering policies are different for each category.

Keywords: Inventory Management, Ordering Policy, Multivariable Analysis, Canonical Correlation Analysis, Cluster Analysis

* Corresponding Author, E-mail: nagasa-k@sophia.ac.jp

1. INTRODUCTION

Inventory management has been studied by many researchers and many practitioners. Tiacci and Saetta (2009) considered the relationship between forecasting and ordering policy with carrier capacity. Flores *et al.* (1992) provided a matrix-based methodology for multi-criteria ABC classification. Ramanathan (2006) proposed a weighted linear optimization model and classified inventory items by multiple criteria. Ernst and Cohen (1990) used cluster analysis to group similar items. Xaio *et al.* (2011) considered the importance of inventory items based on a loss rule. Berling and Marklund (2006) used

linear regression technique to obtain approximate values of the induced backorder cost in a one-warehouse multiple-retailer system.

Many researchers started studying the lost-sales model. In the case of stock-out, Gruen *et al.* (2002) revealed that only 15% of the customers who observe a stock out wait for the item to be on the shelves again. The other 85% of the customers decide to buy a different product, visit another store, or do not buy any product at all. According to Zipkin (2008), the cost deviations can run up to 30% when the lost-sales system is approximated by a backorder model. Janakiraman *et al.* (2007) compared the performance of optimal replenish-

ment policies in lost-sales and backorder models. Huh *et al.* (2009) compared the performance of base-stock policies in lost-sales and backorder models. Levi *et al.* (2008) proposed a dual-balancing policy, in which the risk of ordering and holding are balanced. The authors proved that the expected total cost of this policy is at most twice the expected cost of the optimal policy. Bijvank and Vis (2011) classified the models based on the characteristics of the inventory system and they reviewed their proposed replenishment policies. Van Donselaar and Broekmeulen (2011) showed that the assumption which the lost sales system can simply be approximated by a backordering system if the target fill rate is at least 95%, which may lead to serious approximate errors.

Forecasting is never separate from inventory management. Liao and Chang (2010) investigated the effects of five demand forecasting methods, two inventory policies, and three lead times on the total inventory cost of a 3-echelon serial supply chain system. Tanaka *et al.* (2012) proposed a decision support system that measures demand risks through a sales forecasting method, especially for new products. The system enables us to make the demand uncertainty controllable and gives the proper volume and timing guidance for daily reproduction of products.

One expects that if a single ordering policy is applied to many different items with different shipping trends, the inventory would be a mix of items which has frequent shortages and those having excess inventory. On the other hand, even in using multiple ordering policies, it is not easy to find an appropriate policy for many types of items, particularly when each has a different shipping trend. In addition, given a particular shipping trend, no method or criteria has been specifically developed for choosing the ordering policy which would be applied to these items. For these reasons, the proper relationship between different item trends and the appropriate ordering policies is still not clear.

Therefore, we propose a new algorithm for finding a reasonable ordering policy using the relationships between shipping trends. In particular, our proposed algorithm classifies items according to shipping trends in order to find a suitable ordering policy and in addition reveals the shipping trends which are most desirable. In this case, shipping trends refer to shipping statistics related to the criteria of the ordering policy, which are calculated from past shipping data.

In section 2, we introduce the proposed model. The model development includes the converting of actual shipping data to data matrices for analysis. The details of the experiment and results of analysis are reported in section 3.

2. PROPOSED MODEL

2.1 Problem Outline

This study considers the inventory management pro-

blem for items having different shipping trends. Here, the ordering policy is assumed to perform differently according to item trends. To achieve efficient inventory management, we would like to select a suitable ordering policy for each item. Hence, we want to know which item trends are suitable for determining the ordering policy, or to find out a reasonable method for grouping which items would be suitable and which is not for a given ordering policy.

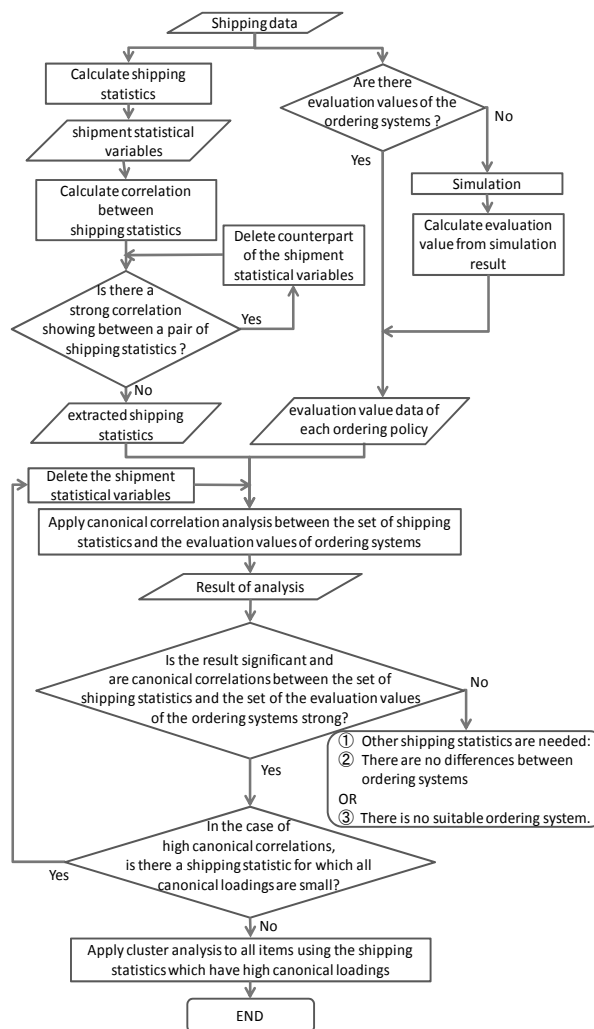


Figure 1. Flowchart of proposed algorithm.

This study assumes that there has not yet been revealed a difference in item shipping trends and also what the suitable ordering policy should be. For deciding an ordering policy, the ideal case would be when the suitable ordering policy could be determined using only a single statistic. However, such a statistic is not yet known. Thus, many researchers have sought to find such a statistic based on theoretical considerations. Similarly, practitioners are also struggling to find such a trend.

In this study, we calculate and use shipping statistics and evaluation values of ordering policies. A general algorithm is presented in the following section.

2.2 Proposed Algorithm

Figure 1 shows the flowchart of our proposed algorithm. First, we assume that there is shipping data, which informs shipping date and volume of items, like that shown in Table 1.

Table 1. Shipping data

Item no.	Date: T		
	Day: 1	...	Day: m
1	d_{11}	...	d_{1m}
⋮	⋮	...	⋮
n	d_{n1}	...	d_{nm}

In the following steps, we repeatedly use multivariate data matrices, as shown in Table 2. Before we proceed to explaining the algorithm flowchart, we will first explain the premise behind multivariate data matrices.

Let \mathbf{X} denote an $n \times v$ multivariate data matrix, which represents the shipping statistics. There are n items and v shipping criteria. x_{ij} gives the value for the j^{th} shipping criteria of the i^{th} item. For example, in Table 1, the daily shipping average is shipping criteria j , $x_{ij} = (d_{i1} + d_{i2} + \dots + d_{im})/m$. Let \mathbf{x}_j denote an $n \times 1$ multivariate data matrix. These are the values of all items for the j^{th} criteria.

Table 2. General variable of each item

Item no.	Shipping statistics : X		Evaluation value of ordering policy : Y			
	Statistic 1	Statistic v	Evaluation criteria 1 : Y ₁		Evaluation criteria z : Y _z	
			Ordering policy 1	Ordering policy q	Ordering policy 1	Ordering policy q
1	x_{11}	x_{1v}	y_{111}	y_{11q}	y_{1z1}	y_{1zq}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
i	x_{i1}	x_{iv}	y_{i11}	y_{i1q}	y_{iz1}	y_{izq}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	x_{n1}	x_{nv}	y_{n11}	y_{n1q}	y_{nz1}	y_{nzq}

Let \mathbf{Y} denote the multivariate data matrix, which represents evaluation values of ordering policies. \mathbf{Y} is similar to the previously mentioned \mathbf{X} . For example, y_{izp} gives the value for the z^{th} evaluation criteria of the p^{th} ordering policy of the i^{th} item. So, \mathbf{y}_{zp} denotes the $n \times 1$ variable matrix of the z^{th} evaluation criteria of the p^{th} ordering policy. Then, \mathbf{Y}_z denotes the $n \times q$ variable matrix of the z^{th} evaluation criteria.

For analysis, since there are differences in the scale of numbers, we should standardize the evaluation value of each evaluation criteria and item in order to compare the ordering policies. If y_{ikp} is the value for the k^{th}

evaluation criteria of the p^{th} ordering policy of the i^{th} item, then let y'_{ikp} be the standardized evaluation value for this same combination, where the y'_{ikp} are calculated using Eq. (1) and take values in $[0, 1]$.

$$y'_{ikp} = \frac{y_{ikp} - \min(y_{ikp})}{\max(y_{ikp}) - \min(y_{ikp})}, \quad \forall i, k, p \quad (1)$$

In order to evaluate multiple evaluation values, we apply a weighted linear mixed evaluate criteria. For example, if the criteria under evaluation are criterion e , criterion f , ... and the corresponding weights are w_e, w_f, \dots , then y_{lpi} , the value of the weighted linear mixed evaluation of criterion l , ordering policy p , and item i , is calculated by $y_{lpi} = w_e y_{iep} + w_f y_{ifp} + \dots$. Similarly, $\mathbf{y}_{lp} = w_e \mathbf{y}_{ep} + w_f \mathbf{y}_{fp} + \dots$. In addition, in order to find differences in ordering policy tends, we applied another weighted linear mixed evaluate criteria. Notably, the second evaluation value was calculated from the difference between evaluation values. For example, to analyze the trend difference for criteria c between policy f and policy g , naming the evaluation criterion d , it is calculated $\mathbf{y}_d = \mathbf{y}_{cf} - \mathbf{y}_{cg}$.

All the notations above, shown in Tables 1 and 2, are used in the sections that follow. The two large sets of variables, \mathbf{X} and \mathbf{Y} , are standard to, for example, the canonical correlation analysis and cluster analysis.

2.2.1 Canonical correlation analysis

The simplest known measure of relationships is the simple correlation coefficient between two variables. When interested in the relation between the set of variables \mathbf{X} and a variable \mathbf{y} , the multiple linear regression analysis can be used. However, in this study, we are interested in the correlation between the set of variables \mathbf{X} and the set of variables \mathbf{Y} , that is, each object for comparison has multiple variables. Canonical correlation analysis is one method for determining the relationship between sets of variables. According to Basu and Mandal(2010), this was initially developed by Hotelling (1936).

The canonical correlation analysis focuses on the correlation between the i^{th} linear combination of the set of variables \mathbf{X} (say \mathbf{U}_i) and the i^{th} linear combination of the set of variables \mathbf{Y} (say \mathbf{V}_i). It determines the coefficients of the linear combination which make the correlation between \mathbf{U}_i and \mathbf{V}_i the highest possible, whereas the correlation between both \mathbf{U}_i and \mathbf{V}_i and each of $\mathbf{U}_{i-1}, \dots, \mathbf{U}_1, \mathbf{V}_{i-1}, \dots, \mathbf{V}_1$ is zero (i.e., uncorrelatedness of other combinations is a constraint on the optimization). The pairs of linear combinations are called the canonical variables, and their correlation coefficients are called the canonical correlation coefficients. In addition, the correlations between a linear combination of the set of variables and the set of variables themselves are called canonical loadings. The canonical correlation analysis is also useful for determining how many dimensions are

needed to account for the relationship.

In the present study, we apply the canonical correlation analysis to the set of shipping statistics \mathbf{X} and the evaluation values \mathbf{Y} . In addition, we expect to find relationships between the sets of linear combinations of variables. This would allow us to sort shipping statistics into those which are significant and those which are not.

2.2.2 Cluster analysis

The cluster analysis is an exploratory technique. Cluster analysis techniques themselves can be broadly grouped into hierarchical clustering and non-hierarchical clustering. We here apply hierarchical clustering using the Ward method, which is a frequently used method.

In this study, we separate items into some number of clusters from x_s , the shipping statistics, according to their relationship to Y_i , the evaluation value of the criteria i which the practitioner is focused on. Then we search for trends in shipping statistics and evaluation values of ordering policies.

3. NUMERICAL EXPERIMENTS

Now, we calculate shipping statistics from actual shipping data and generate evaluation values of ordering policies from the simulation results. Then, we apply canonical correlation analysis to the set of shipping statistics and the evaluation values of the ordering policies. From these results, we can obtain the shipping statistics which exhibit a relatively strong relation with evaluation values of ordering policies. Clustering the items by these shipping statistics, we can group items according to a particular trend in the evaluation values of ordering policies.

In this study, we used SPSS ver. 19.0 (IBM Co., New York, NY, USA), a data mining and statistical analysis software package, for the canonical correlation analysis and cluster analysis, as well as the multiple linear regression analysis. The experimental environment is an Intel® Core™ 2 Duo CPU, E8400 at 3.00 GHz, 4.00 GB RAM; we did not need long computational times for our algorithm.

3.1 Ordering Policy and Experiment

We used two ordering policies for this study: the so-called (s, Q) policy and the (R, s, S) policy. There are two reasons why we choose these two ordering policies. First, these policies are the most commonly used ordering policies for inventory control. Second, the company that provided actual shipment data for us set the parameters and applied these ordering policies. So, we focused on these policies.

For the (s, Q) policy, when the inventory position declines to or below the reorder point s , a batch quantity of size Q is ordered. In this case, we never go beyond the position $s + Q$.

For the (R, s, S) policy, at the end of each time period of length R , if the inventory position declines to or below the reorder point s , order an amount equal to the difference between the order-up-to level S and the current inventory position. In the special case in which $s = S - 1$, we call this the (R, S) policy.

In this study, we use (s, Q) and (R, S) policies. In particular, s , Q , and S of the i^{th} item in the t^{th} date period are denoted by $s_{i,t}$, $Q_{i,t}$, and $S_{i,t}$, which are calculated as follows:

$$\begin{aligned}
 Q_{i,t} &= \begin{cases} LT \times \bar{d}_{i,t} & \forall i \in I, \forall t \in C \\ Q_{i,t-1} & \forall i \in I, \forall t \in D \end{cases} \\
 s_{i,t} &= \begin{cases} LT \times \bar{d}_{i,t} + k \times ss_{i,t} \times \sqrt{LT} & \forall i \in I, \forall t \in C \\ s_{i,t-1} & \forall i \in I, \forall t \in D \end{cases} \\
 S_{i,t} &= \begin{cases} (LT + R) \times \bar{d}_{i,t} + k \times ss_{i,t} \times \sqrt{LT + R} & \forall i \in I, \forall t \in C \\ S_{i,t-1} & \forall i \in I, \forall t \in D \end{cases} \\
 \bar{d}_{i,t} &= \frac{\sum_{j=1}^{t-1} d_{i,j}}{SP} & \forall i \in I, \forall t \in T \\
 ss_{i,t} &= \sqrt{\frac{\sum_{j=1}^{t-1} (d_{i,j} - \bar{d}_{i,t})^2}{SP}} & \forall i \in I, \forall t \in T
 \end{aligned}$$

$Q_{i,t}$: ordering quantity of item i on the t^{th} day

$s_{i,t}$: reorder point for item i on the t^{th} day

$S_{i,t}$: order-up-to level of item i on the t^{th} day

k : safety coefficient

$\bar{d}_{i,t}$: shipment average of item i on the t^{th} day

$d_{i,t}$: shipments of item i on the t^{th} day

R : periodic review term

LT : lead time

SP : sampling period

I : the set of items

T : the set of days

C : the set of days that the parameter is recalculated ($C \subseteq T$)

D : the complement of C with respect to T

If practitioners would like to manage items precisely, it is necessary to shift ordering policy by some reason, like demand forecasting or seasonality. But, because we would like to find the relation between the shipping statistics values and the evaluation values of ordering policies, we did not shift ordering policies in the example applied.

3.2 Data Set

In numerical and simulation experiments, we used the actual data of a distribution company's shipments for the previous year. These items are categorized broadly into 11 varieties, for example, beer, Japanese sake, foods, and spices. In these categories, there are no obvious characteristic trends. So we selected about 5,000 items to be analyzed from about 7,000 items total. The other 2,000 items were less frequently shipped, and in smaller amounts. We applied our algorithm to the se-

lected 5,000 items, those considered analyzable.

We simulated (s, Q) and (R, S) policies from the shipping data, which were used to calculate shipping statistics. We set common parameters between the set of (s, Q) and (R, S) policies, as described below. Safety coefficient k was set to 1.96, sampling period SP was set to 2 months, and lead time LT , the time between when the items are ordered and when they are delivered, was set to 2 days.

We compared the performance between five ordering policies: the (s, Q) policy and the (R, S) policies for R equal to 1, 2, 3, or 4 week. For simulation, we used the first 2 months of shipping data for calculating initial values of $s_{i,t}$, $Q_{i,t}$, and $S_{i,t}$. For the (R, S) policies, $S_{i,t}$ were recalculated periodically, where as $s_{i,t}$ and $Q_{i,t}$ were fixed at the initial calculation values.

3.3 Evaluation Criteria

From the results of the five ordering policy simulations, we define \mathbf{Y}_π as the set of evaluation values, where \mathbf{Y}_π is calculated as $\mathbf{Y}_\pi = [\mathbf{y}_\pi(s, Q), \mathbf{y}_\pi(R = 1\text{week}, S), \dots, \mathbf{y}_\pi(R = 4\text{weeks}, S)]$. $\mathbf{y}_{\pi,p}$ is the set of linear combinations for ordering policy p , of the standardized evaluation of shortage time, $\mathbf{y}_{\beta,p}$, and the standardized evaluation of total daily stock, $\mathbf{y}_{\gamma,p}$. We define $\mathbf{y}_{\pi,p}$ as follows:

$$\mathbf{y}_{\pi,p} = w_\beta \mathbf{y}_{\beta,p} + w_\gamma \mathbf{y}_{\gamma,p} \quad \forall p \quad (2)$$

We take the weights of shortage time, w_β , as 0.5, and the weights of total daily stock, w_γ , as 0.5.

Similarly, we define \mathbf{Y}_δ as the difference between evaluation values for criteria π of policy f and another policy using $\mathbf{y}_{\delta g} = \mathbf{y}_{\pi f} - \mathbf{y}_{\pi g}$. If $\mathbf{y}_{\delta g}$ is positive, then ordering policy g is more suitable than f for item i . In the present study, we compare the evaluation value of criteria

between policy (s, Q) and (R, S) . Thus, \mathbf{Y}_δ are calculated as follows:

$$\mathbf{Y}_\delta = [\mathbf{y}_{\pi(s,Q)} - \mathbf{y}_{\pi(R=1\text{week},S)}, \dots, \mathbf{y}_{\pi(s,Q)} - \mathbf{y}_{\pi(R=4\text{week},S)}] \quad (3)$$

We then apply canonical correlation analysis to the set of shipping statistics \mathbf{X} and the set of evaluation criteria, using either \mathbf{Y}_π or \mathbf{Y}_δ . Next, cluster analysis is applied for all items using the shipping statistics, which are with respect to the ordering policy.

3.4 Experimental Results

The results of the canonical correlation analysis are listed in Table 3. We apply the canonical analysis to the sets (σ, π) and (σ, δ) . Here, σ is the set of shipping statistics and π is the set of evaluation values. The set δ is the set of differences in evaluation value. This notation

is also used in Tables 4-6.

Table 3. Canonical loadings (shipping statistics-evaluation value, differences between policy evaluation values)

		$\varpi_{\pi 1}$	$\varpi_{\pi 2}$	$\varpi_{\pi 3}$	$\varpi_{\delta 1}$	$\varpi_{\delta 2}$
		Canonical variable (1st)	Canonical variable (2nd)	Canonical variable (3rd)	Canonical variable (1st)	Canonical variable (2nd)
X	σ_1 MinimumShippedAmount	-0.067	-0.151	-0.308	-0.041	0.341
	σ_2 ShippedAmountStdev /ShippedAmountAvg	-0.657	-0.354	0.097	-0.518	0.009
	σ_3 ProbabilityOfShippedAmount OverShippedAmountAvg	-0.592	-0.024	-0.535	-0.589	0.542
	σ_4 BDSVariety	-0.315	-0.058	0.547	-0.267	-0.500
	σ_5 MinimumBDS	0.482	-0.008	-0.258	0.458	0.232
	σ_6 BDSAvg	0.751	-0.151	-0.143	0.759	0.152
	σ_7 BDSStdev	0.589	-0.256	0.048	0.640	0.001
	σ_8 BDSStdev/BDSAvg	-0.048	-0.485	0.464	0.113	-0.323
	σ_9 BDSSkewness	-0.574	0.047	0.465	-0.538	-0.447
	σ_{10} BDSKurtosis	-0.376	0.073	0.266	-0.366	-0.268
	σ_{11} ProbabilityOfBDSShorter thanLeadTime	-0.674	0.459	-0.050	-0.774	-0.052
	σ_{12} ((ShippedAmount/BDS)Stdev) /((ShippedAmount/BDS)Avg)	-0.178	-0.479	0.254	-0.023	-0.119
	σ_{13} MonthlyShippedTimeStdev /MonthlyShippedTimeStdev	-0.695	0.177	0.068	-0.706	-0.095
	σ_{14} MonthlyShippedTimeStdev /MonthlyShippedTimeAvg	0.446	-0.648	-0.283	0.605	0.420
Y	π_1 EV(s,Q)	-0.558	-0.356	-0.749		
	π_2 EV(R=1week,S)	0.144	-0.807	0.038		
	π_3 EV(R=2week,S)	0.111	-0.733	0.176		
	π_4 EV(R=3week,S)	-0.106	-0.441	0.354		
	π_5 EV(R=4week,S)	-0.732	-0.183	0.636		
	δ_1 EV(s,Q) - EV(R=1week,S)				-0.654	0.645
	δ_2 EV(s,Q) - EV(R=2week,S)				-0.588	0.720
	δ_3 EV(s,Q) - EV(R=3week,S)				-0.430	0.825
δ_4 EV(s,Q) - EV(R=4week,S)				-0.033	0.990	
Canonical correlation coefficient		0.353	0.238	0.209	0.336	0.210

BDS: between the date of shipping, Avg: average, EV: evaluation value, Stdev: standard deviation.

Conducting the analysis shown in Table 3, we obtained the five canonical variables, $\varpi_{\pi 1}$, $\varpi_{\pi 2}$, $\varpi_{\pi 3}$, $\varpi_{\pi 4}$ and $\varpi_{\pi 5}$ from the canonical analysis of the sets (σ, π) and four canonical variables, $\varpi_{\delta 1}$, $\varpi_{\delta 2}$, $\varpi_{\delta 3}$ and $\varpi_{\delta 4}$, from the canonical analysis of the sets (σ, δ) . We extracted as canonical variables only those for which the canonical correlation coefficient was at least 0.20, the value at which variables can be said to have a weak relation in terms of one group being affected by the other groups. Here, in addition to the canonical correlation coefficients, we must consider the absolute value of canonical loadings. If these are large in absolute value, then the statistic indicates a strong effect on the other groups. For example, in Table 3, if $\varpi_{\pi i}$ in \mathbf{X} are listed in order of the strength of their relation to \mathbf{Y} , then the canonical loadings are $\sigma_6 = 0.751$, $\sigma_{13} = -0.695$, $\sigma_{11} \dots$. So, σ_6 , which is the average of between the date of shipping (BDS), BDSAvg is the most significant statistic for the first canonical variable, $\varpi_{\pi 1}$, of \mathbf{Y} . And, if $\varpi_{\pi i}$ in \mathbf{Y} are listed in order of their relation to \mathbf{X} , these are $\pi_5 = -0.732$, $= -0.558$, \dots . So, π_5 , which is evaluation value (EV; R = 4 week, S), EV of (R, S) policy when $R = 4$ weeks, and π_1 , which is EV (R = 1 week, S), EV of (R, S) policy when $R = 1$ week, were strongly affected by the first canonical variable, $\varpi_{\pi 1}$, of \mathbf{X} . Thus, either π_5 or π_1 and σ_6 , or possibly σ_{13} , are strongly related.

Table 4. Correlation between shipping statistics and evaluation value of ordering policy

		Y								
		π_1	π_2	π_3	π_4	π_5	δ_1	δ_2	δ_3	δ_4
X	S ₁	0.07	0.05	-0.01	-0.03	-0.02	0.04	0.07	0.08	0.07
	S ₂	0.14	0.02	0.06	0.08	0.20	0.12	0.09	0.08	0.00
	S ₃	0.20	-0.01	-0.05	-0.03	0.08	0.19	0.20	0.19	0.12
	S ₄	-0.02	0.04	-0.02	0.03	0.16	-0.04	-0.01	-0.03	-0.10
	S ₅	-0.06	-0.02	0.06	-0.02	-0.15	-0.04	-0.08	-0.04	0.04
	S ₆	-0.11	0.04	0.08	-0.02	-0.20	-0.13	-0.14	-0.09	0.02
	S ₇	-0.10	0.07	0.08	0.01	-0.13	-0.14	-0.13	-0.09	-0.01
	S ₈	-0.02	0.11	0.08	0.08	0.09	-0.09	-0.06	-0.06	-0.07
	S ₉	0.04	-0.04	0.00	0.05	0.21	0.06	0.03	0.00	-0.09
	S ₁₀	0.03	-0.04	-0.01	0.03	0.13	0.05	0.03	0.00	-0.05
	S ₁₁	0.10	-0.15	-0.08	-0.01	0.15	0.18	0.13	0.09	0.00
	S ₁₂	0.04	0.10	0.07	0.07	0.10	-0.03	-0.01	0.00	-0.03
	S ₁₃	0.11	-0.09	-0.04	0.04	0.18	0.16	0.12	0.07	-0.01
	S ₁₄	0.01	0.12	0.14	0.05	-0.13	-0.07	-0.06	-0.02	0.08

The results of correlation analysis are listed in Table 4. In the Table 4, (σ_6, π_2) represents the correlation of the average of BDS and EV of ordering policy (R, S) for $R = 1$ week, which is 0.04. It can seem from Table 3 that there is no relation between the shipping statistics and ordering policy.

The last result is interesting in light of a comparison of Tables 3 and 4. Table 4 gives simple correlations of shipping statistics and evaluation values of ordering policies. By comparing these Tables, it is possible to say that no individual variables and evaluation values of an ordering policy clearly relate to each other. But individual groups of statistics and groups of evaluation values do. A comparison of Tables 3 and 5 reveals similar aspects, where Table 5 shows the results of the multiple linear regression analysis. For example, in Table 5, all entries of the σ_{13} and σ_{11} rows indicate a lack of a strong relation, whereas Table 3 does indicate a strong relation.

Table 5. Standard partial regression coefficients of multiple linear regression analysis between the set of shipping statistics and each evaluation value

	π_1	π_2	π_3	π_4	π_5	δ_1	δ_2	δ_3	δ_4
σ_1	0.078	0.047	-0.009	-0.019	0.012	0.041	0.073	0.076	0.058
σ_2	0.051	0.045	0.078	0.073	0.114	0.018	0.003	0.005	-0.023
σ_3	0.170	0.030	-0.033	-0.054	-0.013	0.135	0.166	0.171	0.147
σ_4	-0.004	0.015	0.045	0.059	0.155	-0.013	-0.028	-0.034	-0.091
σ_5	0.118	-0.055	0.108	0.045	0.085	0.141	0.045	0.076	0.048
σ_6	-0.249	0.053	-0.073	-0.072	-0.203	-0.258	-0.178	-0.173	-0.089
σ_7	0.085	-0.082	0.053	0.044	0.106	0.128	0.046	0.049	0.010
σ_8	-0.037	0.083	-0.033	-0.020	0.009	-0.086	-0.015	-0.021	-0.036
σ_9	-0.062	0.011	0.131	0.105	0.118	-0.063	-0.124	-0.107	-0.118
σ_{10}	0.011	-0.019	-0.048	-0.048	-0.062	0.021	0.035	0.034	0.044
σ_{11}	0.009	-0.089	-0.018	0.005	0.113	0.064	0.017	0.005	-0.056
σ_{12}	0.062	0.049	0.067	0.029	0.009	0.025	0.018	0.037	0.046
σ_{13}	0.059	-0.072	-0.048	0.004	-0.003	0.098	0.077	0.048	0.050
σ_{14}	0.103	0.097	0.173	0.123	0.061	0.033	-0.003	0.024	0.050
Constant	0.238	0.210	0.304	0.428	0.458	0.028	-0.066	-0.189	-0.219
R^2	0.070	0.047	0.040	0.021	0.086	0.070	0.064	0.052	0.044

We, therefore, conjecture that searching for relationships between the set of shipping statistics and the ordering policies would produce results different than the other analyses.

By using cluster analysis, we separate all items into

two groups, also called clusters. In Table 6, the averages of each statistic of each separated group, columns $\chi_{\alpha 1}$ and $\chi_{\alpha 2}$, and the averages of each statistic of all items, column $\chi_{\alpha all}$, are shown. In the cluster analysis, we use the Ward method and calculate distance of each item from the statistics σ_6, σ_{11} , and σ_{13} , those for which canonical correlation analysis results showed a relationship between shipping statistics and evaluation values of ordering policies in the canonical variable $\pi_{\pi 1}$, the first canonical variable of canonical analysis to the sets (σ, π) . The signs “+” and “-” to the right of the values for columns $\chi_{\alpha 1}$ and $\chi_{\alpha 2}$ are relative to $\chi_{\alpha all}$. For example, the value of ($\sigma_6, \chi_{\alpha 1}$) is higher than ($\sigma_6, \chi_{\alpha all}$). Thus, the sign to the right, separated by a dashed, is “+.” From Table 6, we can observe that there are trends within each cluster.

Table 6. Statistical average of all items and two groups separated by the ward method (cluster analysis)

		$\chi_{\alpha 1}$	$\chi_{\alpha 2}$	$\chi_{\alpha all}$
		Average (cluster1)	Average (cluster2)	Average (all items)
X	σ_1 MinimumShippedAmount	2.293 +	1.479 -	2.152
	σ_2 ShippedAmountStdev / ShippedAmountAvg	0.441 -	0.813 +	0.506
	σ_3 ProbabilityOfShippedAmount OverShippedAmountAvg	0.244 -	0.379 +	0.267
	σ_4 BDSVariety	10.263 +	6.085 -	9.538
	σ_5 MinimumBDS	6.139 +	1.000 -	5.247
	σ_6 BDSAvg	23.599 +	2.183 -	19.881
	σ_7 BDSStdev	20.895 +	2.130 -	17.637
	σ_8 BDSStdev/BDSAvg	0.890 +	0.706 -	0.858
	σ_9 BDSskewness	1.176 -	2.301 +	1.371
	σ_{10} BDSKurtosis	1.730 -	7.984 +	2.815
	σ_{11} ProbabilityOfBDSShorter thanLeadTime	0.088 -	0.605 +	0.178
	σ_{12} ((ShippedAmount/BDS)Stdev) / ((ShippedAmount/BDS)Avg)	1.080 +	0.935 -	1.055
	σ_{13} MonthlyShippedTimeStdev	1.502 -	3.802 +	1.901
	σ_{14} MonthlyShippedTimeStdev / MonthlyShippedTimeAvg	1.069 +	0.468 -	0.965
Y	π_1 EV(s,Q)	0.331 -	0.370 +	0.337
	π_2 EV(R=1week,S)	0.278 +	0.218 -	0.267
	π_3 EV(R=2week,S)	0.383 +	0.356 -	0.378
	π_4 EV(R=3week,S)	0.484 +	0.484 -	0.484
	π_5 EV(R=4week,S)	0.554 -	0.598 +	0.562
	δ_1 EV(s,Q) - EV(R=1week,S)	0.053 -	0.151 +	0.070
	δ_2 EV(s,Q) - EV(R=2week,S)	-0.052 -	0.014 +	-0.041
	δ_3 EV(s,Q) - EV(R=3week,S)	-0.154 -	-0.114 +	-0.147
δ_4 EV(s,Q) - EV(R=4week,S)	-0.224 +	-0.228 -	-0.225	

BDS: between the date of shipping, Avg: average, EV: evaluation value, Stdev: standard deviation.

From the results shown in Table 3 and the σ rows in Table 6, it is possible to make predictions. For example, from looking at the $\pi_{\pi 1}$ column of Table 3, let us consider σ_6, σ_{11} , and σ_{13} , which were the top three in terms of the strength of their relation in X to Y of column $\pi_{\pi 1}$. In Table 3, the values of those σ 's, ($\sigma_6, \pi_{\pi 1}$), ($\sigma_{11}, \pi_{\pi 1}$), and ($\sigma_{13}, \pi_{\pi 1}$), are greater than zero, less than zero, and less than zero, respectively; thus, ($\sigma_6, \pi_{\pi 1}$), ($\sigma_{11}, \pi_{\pi 1}$), and ($\sigma_{13}, \pi_{\pi 1}$) correspond to (+, -, -). As another example, in the $\chi_{\alpha 1}$ column of Table 6, the

signs of (σ_6, χ_{a1}) , (σ_{11}, χ_{a1}) , and (σ_{13}, χ_{a1}) are (+, -, -). Therefore, we can predict that (π_1, χ_{a1}) and (π_5, χ_{a1}) in Table 6 will follow the same trends as that represented in the ϖ_{π_1} column of Table 3, that is, they will be less than the average.

In a similar way, in the ϖ_{π_2} column of Table 3, let us consider σ_8 , σ_{12} , and σ_{14} , which were the top three in terms of the strength of their relation in \mathbf{X} to \mathbf{Y} of column ϖ_{π_2} . In Table 3, all three values were less than zero. In the χ_{a2} column of Table 6, (σ_8, χ_{a2}) , (σ_{12}, χ_{a2}) , and (σ_{14}, χ_{a2}) correspond to (-, -, -). From these signs, we can predict that (π_2, χ_{a2}) and (π_3, χ_{a2}) , and possibly (π_4, χ_{a2}) and (π_1, χ_{a2}) in Table 6, will follow the same trend as that represented in the ϖ_{π_2} column of Table 3, which means they will be less than the average.

Next, in the ϖ_{π_1} column of Table 3, $(\sigma_6, \varpi_{\pi_1})$, $(\sigma_{11}, \varpi_{\pi_1})$ and $(\sigma_{13}, \varpi_{\pi_1})$ have signs (+, -, -); in the χ_{a1} column of Table 6, (σ_6, χ_{a1}) , (σ_{11}, χ_{a1}) , and (σ_{13}, χ_{a1}) have the same signs, (+, -, -). Thus, as we would predict, (δ_1, χ_{a1}) , (δ_2, χ_{a1}) , and (δ_3, χ_{a1}) in Table 6 would follow the same trend as that shown in the ϖ_{π_1} column of Table 3, meaning they would be less than the average.

For the case of comparing the evaluation value of a policy, if the average performance of ordering policies were not significantly different and there are complementarities, such as π_1 and π_3 , then a difference in evaluation values of ordering policies implies which ordering policy should be applied. For example, from Table 6, applying the original policy of the evaluation value π_1 , which is a (s, Q) policy, for group χ_{a1} , and applying the origin policy of the evaluation value π_3 for group χ_{a2} is superior to using either one alone.

4. CONCLUSION

In this paper, our focus was on an approach to how to categorize items for a suitable ordering policy based on shipping statistics according to an evaluation of the ordering policy.

The practical point of view for real business of our proposed method is that practitioners can find, easily, significant shipping statistics values for selecting an ordering policy and can select an ordering policy for each item based on those significant statistics values efficiently. In the phase of finding significant shipping statistics values, we can look at the relation between the set of shipping statistics and the set of evaluation values of ordering policies.

In the numerical experiments, actual shipping data were used for calculating shipping data and simulating ordering policies. To evaluate the ordering policy, we considered the simulation results of (s, Q) and (R, S) policies for each item. Then, we applied canonical correlation analysis between the set of shipping statistics and the set of evaluation values of inventory policies.

Using the results, we showed the effectiveness of grouping items by shipping statistics. Because groups,

which are separated by cluster analysis, have different trends in evaluation values, it is possible to see how to select an ordering policy from shipping statistics. For example, by selecting ordering policies appropriately, we can reduce both shortage time and total daily stock.

Furthermore, we conjecture that the results of applying our method would vary according to the demand forecasting method, demand uncertainty, parameter settings for ordering policies, evaluation criteria settings and criteria of shifting ordering policy. By analyzing the interactions of these factors, we would like to extend the model beyond inventories to consider the global optimization of the performance of the entire supply chain management system.

REFERENCES

- Basu, A. and Mandal, A. (2010), Canonical correlation. In Peterson, P., Baker, E., and McGaw, B. (eds), *International Encyclopedia of Education* (Amsterdam: Elsevier), 52-57.
- Berling, P. and Marklund, J. (2006), Heuristic coordination of decentralized inventory systems using induced backorder costs, *Production and Operations Management*, **15**, 294-310.
- Bijvank, M. and Vis, I. F. A. (2011), Lost-sales inventory theory: a review, *European Journal of Operational Research*, **215**, 1-13.
- Ernst, R. and Cohen, M. A. (1990), Operations related groups (ORGs): a clustering procedure for production/inventory systems, *Journal of Operations Management*, **9**, 574-598.
- Flores, B. E., Olson, D. L., and Dorai, V. K. (1992), Management of multicriteria inventory classification, *Mathematical and Computer Modelling*, **16**, 71-82.
- Gruen, T. W., Corsten, D. S., and Bharadwaj, S. (2002), *Retain Out-of-stocks: A Worldwide Examination of Extent, Causes and Consumer Responses*, Grocery Manufacturers of American, Washington, DC.
- Hotelling, H. (1936), Relations between two sets of variates, *Biometrika*, **28**, 321-377.
- Huh, W. T., Janakiraman, G., Muckstadt, J. A., and Rusmevichientong, P. (2009), Asymptotic optimality of order-up-to policies in lost sales inventory systems, *Management Science*, **55**, 404-420.
- Janakiraman, G., Seshadri, S., and Shanthikumar, J. G. (2007), A comparison of the optimal costs of two canonical inventory systems, *Operations Research*, **55**, 866-875.
- Levi, R., Janakiraman, G., and Nagarajan, M. (2008), A 2-approximation algorithm for stochastic inventory control models with lost sales, *Mathematics of Operations Research*, **33**, 351-374.

- Liao, T. W. and Chang, P. C. (2010), Impacts of forecast, inventory policy, and lead time on supply chain inventory: a numerical study, *International Journal of Production Economics*, **128**, 527-537.
- Ramanathan, R. (2006), ABC inventory classification with multiple-criteria using weighted linear optimization, *Computers and Operations Research*, **33**, 695-700.
- Tanaka, K., Akimoto, H., and Inoue, M. (2012), Production risk management system with demand probability distribution, *Advanced Engineering Informatics*, **26**, 46-54.
- Tiacci, L. and Saetta, S. (2009), An approach to evaluate the impact of interaction between demand forecasting method and stock control policy on the inventory system performances, *International Journal of Production Economics*, **118**, 63-71.
- Van Donselaar, K. H. and Broekmeulen, R. A. C. M. (2011), Determination of safety stocks in a lost sales inventory system with periodic review, positive lead-time, lot-sizing and a target fill rate, *International Journal of Production Economics*, in press. <http://dx.doi.org/10.1016/j.ijpe.2011.05.020>.
- Xiao, Y.-Y., Zhang, R.-Q., and Kaku, I. (2011), A new approach of inventory classification based on loss profit, *Expert Systems with Applications*, **38**, 9382-9391.
- Zipkin, P. (2008), Old and new methods for lost-sales inventory systems, *Operations Research*, **56**, 1256-1263.