Available online at www.sciencedirect.com

ScienceDirect

Expert Systems with Applications 34 (2008) 1983–1994

Expert Systems with Applications

www.elsevier.com/locate/eswa

A case study: The prediction of Taiwan's export of polyester fiber using small-data-set learning methods

Der-Chiang Li *, Chun-Wu Yeh, Zheng-Yuan Li

Department of Industrial and Information Management, National Chen Kung University, Tainan, Taiwan

Abstract

During the past four decades, the textile industry has been the industry earning the largest amount of foreign exchange in Taiwan. Notably, polyester fibers are one of the most outstanding industries in Taiwan on the global economic stage. The productivity of the polyester fiber industry in Taiwan has remained steady since the 1980s, and the midstream and downstream industry is also tied in with this development. However, starting from 2000, Taiwan's export of polyester fibers has changed dramatically owing to the rapid economic rise of China. Since this sudden change occurred only 5 years ago, it is hard for researchers to predict the amount of future exports accurately using the trends of historical data over the past 20 years. This research adopts the methodology GIKDE (General Intervalized Kernel Density Estimator), which is a newly developed method for small-data-set prediction, to predict the amount of future exports and expects to obtain a more accurate estimation to use as a reference to help managers make plans for products, capacity and markets.

Keywords: GIKDE; Polyester fiber; Prediction; Textile industry

1. Introduction

During the past four decades, the textile industry has been the industry earning the largest amount of foreign exchange in Taiwan. Notably, polyester fibers are one of the most outstanding industries in Taiwan on the global economic stage. It has not only created new commercial opportunities, but has also developed the growth of Taiwan's economy. However, with fast changes in the market and customer demands, the environment is harsher and harsher for the polyester fiber textile industry. Every element included in the supply chain of the textile industry, such as material supply, polymerization, spinning, weaving, dyeing, and even transportation, is challenging the traditional management paradigm.

1.1. The history of fiber industry in Taiwan

The beginning of the chemical fiber industry in Taiwan can be traced back to the production of rayon by the China Man-made Fiber Corporation (CMFC) in 1955. The polyester fiber industry started in 1964, and Hua-Lon Corporation introduced the techniques from Japan and established the factory at Toufen. This event was 15 years after DuPont had set up similar factories. The introduction of polyester was regarded as the main method for the development of chemical fibers in Taiwan. Therefore, Nan-Ya Plastics Corporation set up its factory in Tai-Shan, Taipei. Shin-Kong Synthetic Fibers Corporation, Oriental Chemical Fibers Corporation, Hong-Zhou Chemical Fibers Corporation and many small factories were built in 1969. This can be considered as the first stage of the development of chemical fibers, and the scale of every factory was small. From 1970 to 1975, Tuntex Distinct Corporation, Da-Xing Chemical Fibers Corporation, and Tainan Spinning Corporation also joined the production and expanded their

* Corresponding author. Tel.: +886 911 874507; fax: +886 6 2374252.
E-mail address: lidc@mail.ncku.edu.tw (D.-C. Li).

0957-4174/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2007.02.007

production scope. This was the second stage of the development of chemical fibers in Taiwan, and soon the country's productivity ranked fourth in the world, behind America, Japan, and West Germany. From 1976 to 1986, the textile industry cooperated with the domestic petrochemical industry and became self-sufficient in material supplies. Although the second world oil crisis took place during this period of time, and the resulting recession made it hard for producers, the industry made efforts to adjust production and marketing, and also reformed and merged to make the whole industry more secure. At this time, Da-Xing Chemical Fibers Corporation was incorporated into Chung-Shing Textile Corporation; Oriental Chemical Fibers Corporation and Far Eastern Textile Corporation were amalgamated, Hua-Lon Corporation merged five small factories, and Yu-He Chemical Fiber Corporation and Da-Ming Chemical Fiber Corporation authorized other factories to manufacture their products. With the return of low price of up-stream petrochemical materials, the expansion in production of polyester was faster than for other chemical fibers. The productivity of polyester fibers in Taiwan exceeded Japan and West Germany, reaching 990,000 tons annually in 1986, ranking the second in the world. The textile industry in Taiwan, from the industries producing the raw materials to the midstream and downstream industries that process the semi-products, has become a huge and integrated production and marketing system, capital-, technology-, and labor-intensive.

1.2. The emerging challenge

The development of the textile industry in Taiwan has been a decades-long process, and the productivity of the industry grew steadily from 1987 to 1996, as shown in Table 1. Meanwhile, global production reached 18,900,000 tons in 2000, with most of it in the countries of East Asia, such as Taiwan, Japan, Korea, and China. The amount of polyester fiber produced in these countries and in Southeast Asia reached a total of 12,370,000 tons in 2000, accounting for about 65% of the global market. Since some countries have reduced planting cotton and increased using synthetic fibers, the polyester fiber industry is set to keep growing.

Industrial evolution is a long-term and complicated procedure, combining competition and the global division of labor. Since 1987, the polyester fiber industry in Taiwan has entered its fourth stage of internationalization. The unstable supply of the raw materials of PTA and the restrictions and obstructions caused by high costs has limited investment and growth in Taiwan. Large companies have invested overseas using the experience gained and techniques mastered over the years in the domestic polyester fiber industry. The productivity of these companies, such as Nan-Ya Plastics Corporation, Tuntex Distinct Corporation, The Far Eastern Group, and Hua-Lon Corporation, all rank among the top ten related factories in the world (as shown in Table 2). Nan-Ya has set up a factory in Southern Carolina, Far Eastern Textile Corporation absorbed some businesses in the

Table 1
The output value of the textile industry in Taiwan (1987–1997) (unit: 100 million New Taiwan Dollars; about 3 million US Dollars)

Year	Artificial fiber textile industry	Textile industry	Ready-made clothes textile industry	Total
1987	875	3827	956	5564
1988	947	3432	881	5564
1989	1024	3578	945	5564
1990	1003	3374	919	5564
1991	953	3825	832	4817
1992	933	3698	726	4817
1993	823	3338	656	4817
1994	1049	3706	575	5330
1995	1520	3804	545	5769
1996	1206	3776	553	5535
1997	1255	3940	548	5743

Table 2
The top ten firms by total productivity of polyester yarn and polyester staple fiber worldwide

Polyester yarn	Corporation	Nationality	Productivity	Polyester staple fiber	Corporation	Nationality	Productivity
1	Hua-Lon	Taiwan	549	1	Wellman	U.S.A.	557
2	Nan-Ya	Taiwan	541	2	Nan-Ya	Taiwan	507
3	DuPont	U.S.A.	440	3	Hovis	Korea	437
4	Tuntex	Taiwan	411	4	Tuntex	Taiwan	436
5	TEIJIN	Japan	374	5	DuPont	U.S.A.	400
6	Reliance	India	365	6	TEIJIN	Japan	344
7	Far Eastern	Taiwan	365	7	Yizheng	China	344
8	Hovis	Korea	338	8	Kosa	U.S.A.	340
9	Kosa	U.S.A.	298	9	Reliance	India	330
10	KCC	Korea	274	10	Far Eastern	Taiwan	274
Total productivity of ten corporations			3955	Total productivity of ten corporations			3969
Total productivity of polyester yarn worldwide			12,881	Total productivity of polyester staple fiber worldwide			9359
The percentage of productivity of 10 corporations			31%	The percentage of productivity of ten corporations			42%

Philippines, Hua-Lon Corporation made an investment in Malaysia, Tuntex Distinct Corporation set up factories of polyester staple fiber and polyester yarn in Thailand, and Shin-Kong Synthetic Fibers Corporation established a PET factory with annual productivity of 120,000 tons in Thailand.

In 1990, China attracted large chemical fiber factories with its lure of a huge market and its view of the industry as a key to economic development. Tuntex Distinct Corporation firstly started the investment, and set up a polyester fiber factory in Xiamen. Shin-Kong Synthetic Fibers Corporation then set up the factory of polyester yarn with annual productivity of 50,000 tons in Zhejiang. Far Eastern Textile Corporation set up the factory of polyester with annual productivity of 250,000 tons in Jiangsu. Its main products are PET, polyester staple fiber, and polyester poy (pre-oriented yarn). The manufacturing cost in the past five years has consistently fallen due to the large supply of cheap labors and the vast resources and land in China and Southeast Asia. China itself has enough capacity to supply the demands of its internal markets. Local manufacturers in China have thus replaced the role of providers from Taiwan. Furthermore, owing to recession and high petroleum prices, there is a lower utilization of manufacturing capacity in the polyester fiber factories in Taiwan, and the situation has become progressively worse. The utilization of factory capacity in Taiwan has generally dropped to 60% or 70%, and some factories have dropped to less than 50%. Hence this research hopes to provide an effective prediction of the future of Taiwan's export of polyester fiber.

1.3. The study's goal and methods

The productivity of the polyester fiber industry in Taiwan has remained steady since the 1980s, and the midstream and downstream industry is also tied in with this development. However, starting from 2000, Taiwan's export of polyester fibers has changed dramatically owing to the rapid economic rise of the People's Republic of China (PRC). Since this sudden change occurred only 5 years ago, it is hard for researchers to predict the amount of future exports accurately using the trends of historical data over the past 20 years. This research adopts the methodology GIKDE (General Intervalized Kernel Density Estimator), which is developed uniquely for small-data-set prediction, to predict the amount of future exports and expects to obtain a more accurate estimation to use as a reference to help managers make plans for products, capacity and markets.

The characteristics of GIKDE are as follows. First, it requires only a small amount of data to make predictions. Secondly, GIKDE combines the concepts of time series and randomization, and it can deal with data of different types, such as independent, dependent, numerical, and nominal. Therefore, GIKDE is applied in this research for the prediction of Taiwan's polyester fiber exports.

2. Applying GIKDE to export prediction

Due to the influence of joining WTO and the rapid economic development of China, the exports of the polyester fiber textile industry in Taiwan has changed abruptly. Therefore, the available data concerning exports are too few to reliably predict market demand. This research applies GIKDE, which is uniquely designed for small data learning, to predict the future exports of the polyester fiber textile industry in Taiwan. The detailed description of GIKDE is in the following sections.

2.1. The basic concept of GIKDE

To estimate the distribution density of a set of independent data, IKDE (Li & Lin, 2006, in press), the predecessor model, uses many kernels and location parameters to display the model. As shown in (1), the mathematic form of IKDE concerns the summation of m kernel functions, each K_i accompanied with its location parameter μ_i, intervalized smoothing parameter h_i, and the number of covered data n_i.

$$\text{IKDE}: \quad \hat{f} = \sum_{i=1}^{m} \frac{1}{n_i}\frac{1}{h_i} K_i\left(\frac{x-\mu_i}{h_i}\right) \tag{1}$$

However, the export data of polyester will be continuously collected, and the new data is likely to be affected by the past information. This means the new data are not independent of the past ones, so that the occurrence series of data needs to be considered. In addition, it is difficult to formulate the density for the next input data by prioritizing the kernels or re-setting the location parameters. GIKDE removes the assumption of independence in IKDE, and employs time series theories to make the application more general.

2.2. Kernel location functions of GIKDE based on time series data

Assume that $\{X_1,\ldots,X_t\}$ is a sequential data that comes from t respective populations, and each of them corresponds to a kernel. Under the assumption of time dependency, not all of the collected data share the same importance in capturing the trend of the next incoming data. Since each datum has been assigned to a particular kernel, this research introduces an inspection window, d, to assist in the determination of the number of kernels. The inspection window implies that using GIKDE for predicting the next incoming data should trace back the referred information by d stages. That is, the prediction of $f(x_{t+1})$ through GIKDE will take the last and the most recent d kernels of $\{X_1,\ldots,X_t\}$ into consideration. Since each kernel covers one datum, d indicates that the prediction of $f(x_{t+1})$ only considers $\{X_{t-d+1},\ldots,X_{t-1},X_t\}$. Furthermore, this research introduces $G(X_1,\ldots,X_t)$ to modify the argument of kernel functions, since time series theory is a useful tool for capturing the path changing of certain sequential data, and the path that we mention here is equivalent to the sample path in stochastic processes. In other words, the location parameter μ_t is replaced by $G(X_1,\ldots,X_t)$, and it can be set using the conditional expected value of X_{t+1}.

$G(X_1,\ldots,X_t) \equiv E[X_{t+1}|X_1,\ldots,X_t]$. A linear combination of $\{X_1,\ldots,X_t\}$ can also represent this conditional mean, that is

$$E[X_{t+1}|X_1,\ldots,X_t] = \beta_1 X_1 + \cdots + \beta_t X_t$$

where β_i is the coefficient of X_i.

Furthermore, the argument of kernel functions needs be modified. Thus, we introduce $G(X_1,\ldots,X_t)$, the location function of the tth kernel, to replace the location parameter, μ_t. When using time series theory for modeling kernel locations, the tth location, $G(X_1,\ldots,X_t)$, can be set using the conditional expected value of X_{t+1}; that is, $G(X_1,\ldots,X_t) \equiv E[X_{t+1}|X_1,\ldots,X_t]$. A linear combination of $\{X_1,\ldots,X_t\}$ can also represent this conditional mean, that is

$$E[X_{t+1}|X_1,\ldots,X_t] = \alpha_0 + \beta_1 X_1 + \cdots + \beta_t X_t$$

where β_i means the coefficient of X_i and α_0 is the constant term.

This is analogous to an autoregressive (AR) model of order t, AR(t), defined as

$$X_{t+1} = \alpha_0 + \beta_1 X_1 + \cdots + \beta_t X_t + \varepsilon_{t+1}$$

where ε_{t+1} follows a normal distribution ($\varepsilon_{t+1} \sim N(0,\sigma^2)$) as a random error of the tth stage. AR models help summarize the past collected data and the order t will include all of them. The value of t is small in the early stages, hence every collected datum at different phases from the initial period may have more or less influence on the next (or present) data in the early stages. In other words, $G(X_1,\ldots,X_t)$ has a general form equaling AR(t).

2.3. Weighted kernels series of GIKDE for showing the time trend

GIKDE is a weighted summation of kernel functions that employs $\{w_d, d \geqslant 1\}$ as weights in the coefficient sequence to replace all $1/n_i h_i'$s in IKDE, where d is the inspection window, since there are d kernels concerned. In addition, it is required that the sum of all w_ds equals one (i.e. $\sum_d w_d = 1$), so that $\{w_d,\ d \geqslant 1\}$ would satisfy $\int \hat{f}_{t+1} \mathrm{d}x = 1$. Note that when $\int \hat{f}_{t+1} \mathrm{d}x \neq 1$, we need to divide every w_d by the amount of $\int \hat{f}_{t+1} \mathrm{d}x$. We use $\{w_d,\ d \geqslant 1\}$ to represent the intensity of time, and it is an arbitrary sequence from the initial stage to the last. $\{w_d,\ d \geqslant 1\}$ represents the concept that the degree of influence of the past data diminishes over time. It also partially stresses the Markovian property even though it is not a stochastic process. That is, we enable the sequence of w_d to provide flexible adjustment of time intensity to the model.

An increasing sequence can be a good choice for deciding the type of a weight sequence. There are three types of increasing weight series considered in this study: noninformative, common increasing, and Martingale. The first and the last are degenerative cases of increasing series, they have decisive forms as the number of weight elements. For predicting the amount or distribution of the next incoming data, the Martingale type, $\{1, 0\cdots\}$, suggests being fully focused on the present data without considering the past information, since the next input will be affected by the current data only. Besides, the noninformative type of weight series that assumes all kernels occurring at the same time will share equal influence. Hence, the importance of each element is equal to the others, $w_1 =, \ldots, = w_d = 1/d$, in noninformative types. For example, one can use $w_1 = w_2 = w_3 = 1/3$ if explorers trace back three stages of data for establishing the GIKDE for $f(x_{t+1})$ with a noninformative weight series. Actually, this type of weight series always regards a sequence of time-dependent data as a set of random data. The common increasing type allows the weights to grow as time goes by, so that in the case of $d = 3$, $\{w_1 = \frac{1}{2},\ w_2 = \frac{1}{4},\ w_3 = \frac{1}{4}\}$ assigns half a percentage to the last kernel and the remaining of the percentage to the other two kernels. The Martingale type is actually a degenerative increasing sequence, even a monotonically increasing one, $(1, 0, \ldots, 0)$, with d terms assigning the weight of the last kernel to be equal to one and the remaining kernels zeros. This unit vector implies the standard Markovian property mentioned above. Table 3 shows these three weight series concerning the inspection window $d = 3$.

GIKDE emphasizes the assumption of time dependency so that it provides a density estimator for the next stage, $\hat{f}_{t+1}$. That is, $\hat{f}_{t+1}$ is defined as the weighted summation of different kernels coming at various times using weights $\{w_d,\ d \geqslant 1\}$. Since we have defined the number of kernels, proposed the location function for replacing the original location parameter of kernels, and set the weight sequence for the weighted summation, a GIKDE can thus be formed as

$$\hat{f}_{t+1} = \sum_{j=1}^{d} w_j K_j \left(\frac{x - G(X_1,\ldots,X_{t-d+j+1})}{h_j} \right) \tag{2}$$

Table 3
Three types of increasing weight series when inspection window $d = 3$

Types of increasing weight series	$d = 3$
Noninformative	$\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$
Common increasing	$\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\}$, etc.
Martingale	$\{1, 0, 0\}$

For a clearer demonstration, we use a sequence of $t = 5$, $\{X_1, \ldots, X_5\}$, to illustrate formulating the GIKDE of $f(x_6)$. The inspection window is user-defined according to researchers' demands, and should be less than or equal to 5 in this example, that is, $d \leqslant t$. Here, we set d equal to 3. From (2), we obtain,

$$\begin{aligned}\hat{f}(x_6) &= \sum_{j=1}^{3} w_j K_j\left(\frac{x - G(X_1, \ldots, X_{5-3+j+1})}{h_j}\right)\\ &= w_1 K_1\left(\frac{x - G(X_1, \ldots, X_4)}{h_1}\right) + w_2 K_2\left(\frac{x - G(X_1, X_2, X_3)}{h_2}\right)\\ &\quad + w_3 K_3\left(\frac{x - G(X_1, X_2)}{h_3}\right)\end{aligned}$$

The GIKDE of $f(x_6)$, $\hat{f}(x_6)$ includes three weighted kernels by summing them up. K_1 is the kernel function covering the last collected datum, X_5, and its location function takes all the data prior to the present, $\{X_1, \ldots, X_5\}$, into consideration. K_2 is the kernel function covering the second last collected data, X_4 and the corresponding location function considers only the data from the beginning up to X_4 (i.e. $G(X_1, \ldots, X_4)$). K_3 is formed likewise. GIKDE is actually an estimator of conditional densities when possessing the assumption of time dependency. $\hat{f}_{t+1}$ provides a very informative summary of the sequential data, $\{X_1, \ldots, X_t\}$, and it is a density estimator conditioned by the past data, that is,

$$\hat{f}_{t+1}(x) = \hat{f}(x_{t+1}) = \hat{f}(x|X_1, \ldots, X_t) \tag{3}$$

2.4. Monte Carlo methods for generating samples

To generate more data from the original small population based on GIKDE, we need a generation method that extracts a set of samples agreeing with the original data set and guarantees the quality of these generated data. Monte Carlo methods are the most common approaches to help extract the desired samples; therefore, this research employs these methods tied in with the inverse method for sample generation using the GIKDE.

3. Experimental study

It is practical to use a virtual sample generation (VSG) technique to improve the quality of the predictive model when the sample set acquired is small. Experimental results show that GIKDE helps to accelerate and stabilize the learning effectiveness. This research uses the export data of polyester fibers in Taiwan from year 2000 to 2005 (Table 4) as training samples to verify the prediction model.

3.1. The experimental procedure

To summarize the whole experimental process, the procedure is listed below.

Step 1: Decide $d, \{w_j\}, K_j, h_j$, and determine the central parameter $G(X_1, \ldots, X_{t-j+1})$.
For $j = 1$ to d

$$\begin{aligned}G(X_1, \ldots, X_{t-d+j-1}) &= E\left[X_{t-d+j}\middle|X_1, \ldots, X_{t-d+j-1}\right]\\ &= \alpha_0 + \beta_1 X_1 + \ldots + \beta_{t-d+j-1} X_{t-d+j-1}\end{aligned}$$

then GIKDE is set as $\hat{f}_{t+1} = \sum_{j=1}^{d} w_j K_j\left(\frac{x - G(X_1, \ldots, X_{t-d+j-1})}{h_j}\right)$ for each data set, we use GIKDE to build $\widehat{f}(X_{2005})$ based on $X_{2000}, \ldots, X_{2004}$.

Step 2: Generate virtual data sets with the desired size, based on $\hat{f}_{t+1}$, using Monte Carlo methods to generate virtual samples of $\hat{f}(X_{2005})$.

Step 3: Applying back-propagation neural networks (BPN) to the original data from year 2000 to 2004 to predict the exports for 2005 (that is $\widehat{X}_{2005}$) and calculating the mean square error (MSE) between the predicted values and the actual values. Furthermore, we also use BPN with expanded data sets that are composed of the original data and virtual samples to predict the exports for 2005 and calculate the MSE.

Step 4: Comparing both MSEs to verify the effectiveness of joining virtual samples.

Table 4
The statistics of the export of polyester fibers in Taiwan from 2000 to 2005

		2000	2001	2002	2003	2004	2005
1	Nylon fiber	12,297	9669	8928	6442	6359	6319
2	Polyester fiber	671,271	586,912	677,981	679,936	659,809	566,502
3	Acrylo nitrile fiber	75,554	86,338	98,436	110,994	125,009	114,246
4	Rayon fiber	75,574	52,775	102,984	95,058	81,603	67,504
5	Nylon draw texturing yarn	48,334	49,637	46,778	49,632	44,837	44,581
6	Nylon fiber yarn	997	1415	1250	2756	2881	1327
7	Nylon yarn	88,667	81,399	132,718	174,829	176,647	158,004
8	Polyester draw texturing yarn	306,885	296,746	325,139	290,460	299,959	257,407
9	Polyester fiber yarn	44,965	40,954	44,179	45,438	38,029	33,546
10	Polyester yarn	224,830	244,112	302,983	337,127	369,938	339,333
11	Acrylo nitrile fiber yarn	23,469	20,525	16,554	9726	13,015	14,798
12	Rayon yarn	2472	1905	1083	1181	1843	1529

Step 5: Repeating steps 1–3, to predict the exports for 2006 ($\widehat{X}_{2006}$) based on the information of $X_{2000}, \ldots, X_{2005}$.

For each material, this research shows the steps of building the predictive model in detail.

3.2. Step 1

For all materials we here set $d = 3$, $h_1 = h_2 = h_3 = 1$ and use the common increasing weight series. The coefficients of $G(X_1, \ldots, X_{t-j+1})$ can be obtained from any statistical regression software, and MINITAB is used in this research. Take nylon fiber for example:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 0.998345X_{2004} = 6348.48\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.78104X_{2004} - 0.78654X_{2003} = 6258.74\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 1.37781X_{2004} - 0.24612X_{2003} - 0.62256X_{2002}\\
&= 4788.83
\end{aligned}
$$

$$
\begin{aligned}
\widehat{f}(X_{2005}) &= \sum_{j=1}^{3} W_j K_j\left(\frac{X_{2005} - G(X_{2000},\ldots,X_{2004-j+1})}{h_j}\right)\\
&= \sum_{j=1}^{3} W_j K_j(X_{2005} - \mathrm{AR}(j))\\
&= \sum_{j=1}^{3} W_j \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(j))^2} \quad \text{(Gaussian kernel function)}\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2}\\
&\quad + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2} \quad \text{(Common increasing)}
\end{aligned}
$$

Similarly, for polyester fiber:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 1.001493X_{2004} = 660793.9\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.750091X_{2004} - 0.75375X_{2003} = 642223.1\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 1.327158X_{2004} - 0.32195X_{2003}\\
&\quad - 0.64955X_{2002} = 654191\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For acrylo nitrile fiber:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 1.000507X_{2004} = 125072.4\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.952896X_{2004} - 0.95218X_{2003} = 138443.5\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 2.213496X_{2004} - 1.59313X_{2003}\\
&\quad - 0.379829X_{2002} = 137268.3\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For rayon fiber:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 0.998993X_{2004} = 81520.81\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 0.45548X_{2004} - 0.545198X_{2003} = 88993.92\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 0.252935X_{2004} - 0.187076X_{2003}\\
&\quad - 0.55995X_{2002} = 96089.19\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For nylon draw texturing yarn:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 1.001278X_{2004} = 44894.32\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.933202X_{2004} - 0.97667X_{2003} = 38204.86\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 2.055752X_{2004} - 1.24805X_{2003}\\
&\quad - 0.165435X_{2002} = 37969.23\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For nylon fiber yarn:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 1.001956X_{2004} = 2886.636\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 0.851216X_{2004} - 0.152705X_{2003} = 2873.207\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 0.807313X_{2004} - 1.002211X_{2003}\\
&\quad - 0.80762X_{2002} = 4078.436\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For nylon yarn:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 1.000746X_{2004} = 176778.8\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.46771X_{2004} - 0.46669X_{2003} = 177675.7\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 1.168734X_{2004} - 0.299833X_{2003}\\
&\quad - 0.46502X_{2002} = 197155.7\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For polyester draw texturing yarn:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 1.001456X_{2004} = 300395.8\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.831061X_{2004} - 0.83939X_{2003} = 305433.6\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 1.088988X_{2004} - 0.71805X_{2003}\\
&\quad - 0.86888X_{2002} = 252708.2\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For polyester fiber yarn:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 0.999109X_{2004} = 37995.11\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.903697X_{2004} - 0.9163X_{2003} = 30760.95\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 1.595932X_{2004} - 0.27608X_{2003}\\
&\quad - 0.36545X_{2002} = 32001.77\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For polyester yarn:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 1.000494X_{2004} = 370120.7\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.900126X_{2004} - 0.90132X_{2003}\\
&= 399069.5\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 1.523208X_{2004} - 0.05078X_{2003} - 0.48488X_{2002}\\
&= 399465.7\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

For acrylo nitrile fiber yarn:

$$
\begin{aligned}
G(X_{2000},\ldots,X_{2004}) &= \mathrm{AR}(1) = 0.998737X_{2004} = 12998.56\\
G(X_{2000},\ldots,X_{2003}) &= \mathrm{AR}(2) = 1.705139X_{2004} - 0.70709X_{2003} = 15315.22\\
G(X_{2000},\ldots,X_{2002}) &= \mathrm{AR}(3) = 1.908483X_{2004} - 1.41826X_{2003} - 0.509456X_{2002}\\
&= 19478.42\\
\widehat{f}(X_{2005}) &= \frac{1}{2\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}
\end{aligned}
$$

Table 5
The virtual sample sets of 12 materials of polyester fibers

Nylon fiber		Polyester fiber		Acrylo nitrile fiber		Rayon fiber	
COM10	COM20	COM10	COM20	COM10	COM20	COM10	COM20
6135.389	6256.03	662101.3	665078.3	128152.5	126966.3	84788.7	77970.47
6213.045	6265.69	659858.1	656827.1	125359.8	124922.8	81380.35	80155.49
6278.887	6559.431	655569.8	660383	124516.1	128135.8	83248.21	80123.38
6360.565	6589.919	661276.9	664741.5	127681.4	122820.4	80887.66	79779.33
6352.639	6200.659	659673.2	661521.7	124729.1	127220	80917.43	80356.72
5802.287	6136.209	642521.4	662683.2	136392.8	124440.4	90610.49	80801.27
6855.485	5896.958	645776	658283.8	137531.7	123851.2	89966.4	80798.73
6093.804	6373.91	644782.8	656585.8	137195.8	122558.9	90807.46	84825.16
4521.311	6168.591	651769.6	657724.3	132965.4	123651.7	99114.51	84243.9
5010.509	6537.53	660624.9	659623	140103.6	121624.1	97256.9	82146.39
	6122.881		644899.7		139973.3		80946.81
	6220.192		639270.3		136888		82161.17
	6421.422		639954.3		136460.2		79154.44
	6297.136		639625.5		136045		83230.86
	6310.305		640586.6		136935		79368.55
	4886.028		654939.6		138022.6		99590.77
	4921.328		651579.1		141045.7		93499.02
	5184.786		654564.5		138505.8		94717.74
	4976.402		655219.6		136343.3		95660.07
	4807.132		648053.4		136470.6		97913.25
Nylon draw texturing yarn		Nylon fiber yarn		Nylon yarn		Polyester draw texturing yarn	
44981.63	44548.65	2815.481	2955.846	181697.4	177609.7	302203.9	301287.4
44477.63	44940	3059.656	3004.905	170738.1	179940.9	299138.6	300804.9
45041.73	44669.49	2910.699	3006.288	178380.5	184694.1	300913.7	300579.5
45120.18	45039.1	2786.277	2714.334	177417.1	181228.4	300618.4	298889.3
45169.2	44689.44	3038.245	2761.626	179334.8	177597.4	301609.1	300195.3
38217.42	44844.74	2786.025	2973.583	177025.9	176600.1	307058.9	297630.2
37896.28	44602.96	2915.682	2937.252	181884.8	171575.4	306580.3	300521.8
38276.44	44556.49	2977.787	2785.379	183075.7	175234.9	303687.4	302696.7
37880.17	45128.12	4105.756	2725.92	190748.6	176702.1	252786.6	303810.4
37338.11	45436.52	3968.719	2908.698	193066.9	173326.6	254566.9	303334
	38201.72		2982.006		184213.5		304806.4
	38265.51		2819.944		185620.7		305990.7
	38481.35		3009.293		173037.3		305339.3
	38045.46		2840.418		184571.5		304664.1
	37845.42		2848.509		180932.5		251775.3
	38021.62		4106.946		202767.6		252529.4
	38299.37		4176.783		181023.5		252005
	38322.66		4201.614		201330.8		251366.1
	38030.74		4163.804		196151.7		251888.9
	37880.62		4120.451		185583.6		252624.4
Polyester fiber yarn		Polyester yarn		Acrylo nitrile fiber yarn		Rayon yarn	
37616.22	37600.27	376523.6	364090.6	13903.26	12295.18	1821.577	1899.577
37826.15	38123.25	365430.8	368805	13271.78	13094.7	1864.727	1909.035
38493.35	38011.2	359547.7	377514.7	12487.52	13492.52	1821.597	1870.718
38256.7	37903.85	364330.2	371156.6	12346.06	12405.33	1877.056	1798.755
38127.75	37306.89	377094	366914.6	13076.91	12223.46	1828.238	1972.651
30838.66	37819.85	411145.2	367568.8	14198.5	12780.97	2329.385	1890.857
30662.46	38329.18	393765.6	371708.1	15537.78	12998.65	2450.427	1864.758
31048.57	38428.17	389578.2	374539.8	14797.11	13334.11	2321.922	1891.415
32000.67	37877.96	406354	375630.4	18760.39	12567.21	2101.066	2234.136
32188.58	38118.11	394529.9	367170.5	19974.21	12192.91	2278.38	2166.813
	30754.17		402503.9		19778.21		2175.517
	30876.89		393863.8		18900.56		2076.318
	30660.51		390044.2		20282.69		2098.503
	30591.05		396666.9		19625.51		2258.873
	31004.42		403883.5		18804.19		2295.935
	31935.95		385469.7		14449		2271.195
	31513.78		395493.4		16238.88		2330.206
	31727.4		399848.9		14769.67		2312.338
	31756.21		406858.1		15765.42		2356.135
	32357.63		405359.8		14907.87		2273.502

Table 6
The original training data set of nylon fiber

(X_{2000}, X_{2001})	(12297,9669)
(X_{2001}, X_{2002})	(9669,8928)
(X_{2002}, X_{2003})	(8928,6442)
(X_{2003}, X_{2004})	(6442,6359)

Table 7
The training data set of nylon fiber produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	12297	9669	12297	9669
(X_{2001}, X_{2002})	9669	8928	9669	8928
(X_{2002}, X_{2003})	8928	6442	8928	6442
(X_{2003}, X_{2004})	6442	6359	6442	6359
(X_{2003}, X_{2004})	6442	6135.389	6442	6256.03
(X_{2003}, X_{2004})	6442	6213.045	6442	6265.69
(X_{2003}, X_{2004})	6442	6278.887	6442	6559.431
(X_{2003}, X_{2004})	6442	6360.565	6442	6589.919
(X_{2003}, X_{2004})	6442	6352.639	6442	6200.659
(X_{2003}, X_{2004})	6442	5802.287	6442	6136.209
(X_{2003}, X_{2004})	6442	6855.485	6442	5896.958
(X_{2003}, X_{2004})	6442	6093.804	6442	6373.91
(X_{2003}, X_{2004})	6442	4521.311	6442	6168.591
(X_{2003}, X_{2004})	6442	5010.509	6442	6537.53
(X_{2003}, X_{2004})			6442	6122.881
(X_{2003}, X_{2004})			6442	6220.192
(X_{2003}, X_{2004})			6442	6421.422
(X_{2003}, X_{2004})			6442	6297.136
(X_{2003}, X_{2004})			6442	6310.305
(X_{2003}, X_{2004})			6442	4886.028
(X_{2003}, X_{2004})			6442	4921.328
(X_{2003}, X_{2004})			6442	5184.786
(X_{2003}, X_{2004})			6442	4976.402
(X_{2003}, X_{2004})			6442	4807.132

For rayon yarn:

$$G(X_{2000},\ldots,X_{2004}) = \mathrm{AR}(1) = 0.999147X_{2004} = 1841.428$$
$$G(X_{2000},\ldots,X_{2003}) = \mathrm{AR}(2) = 1.75282X_{2004} - 0.75632X_{2003} = 2337.234$$
$$G(X_{2000},\ldots,X_{2002}) = \mathrm{AR}(3) = 1.696571X_{2004} - 1.64866X_{2003} - 0.95282X_{2002} = 2211.613$$
$$\hat{f}(X_{2005}) = \frac{1}{2\sqrt{2\pi}}e^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(1))^2} + \frac{1}{4\sqrt{2\pi}}e^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(2))^2} + \frac{1}{4\sqrt{2\pi}}e^{-\frac{1}{2}(X_{2005}-\mathrm{AR}(3))^2}$$

(Common increasing)

3.3. Step 2

This research uses Monte Carlo methods that assist in generating random numbers, and these random numbers are virtual samples. VSG generates virtual samples based on GIKDE to help predict the value of the next period, $\hat{f}(x_{t+1})$, and it can be implemented in all kinds of materials. Determination of the number of virtual samples needs meticulous mathematical inference concerning learning structure and accuracy, and trial-and-error is an alternative to decision-making. This research uses $\hat{f}(C_{t+1})$ to assist in estimating the scale of each material at the stage of $t+1$. Through Monte Carlo methods, we generate different numbers of virtual samples that include 10 and 20 virtual data (that is $n = 10, 20$) for different $\hat{f}(X_{2005})$. Take nylon fiber for instance, we generate two virtual sample sets, COM10$(n = 10)$ and COM20$(n = 20)$, that have 10 and 20 virtual data, respectively. The virtual sample sets of

Table 8
The training data set of polyester fiber produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	671271	586912	671271	586912
(X_{2001}, X_{2002})	586912	677981	586912	677981
(X_{2002}, X_{2003})	677981	679936	677981	679936
(X_{2003}, X_{2004})	679936	659809	679936	659809
(X_{2003}, X_{2004})	679936	662101.3	679936	665078.3
(X_{2003}, X_{2004})	679936	659858.1	679936	656827.1
(X_{2003}, X_{2004})	679936	655569.8	679936	660383
(X_{2003}, X_{2004})	679936	661276.9	679936	664741.5
(X_{2003}, X_{2004})	679936	659673.2	679936	661521.7
(X_{2003}, X_{2004})	679936	642521.4	679936	662683.2
(X_{2003}, X_{2004})	679936	645776	679936	658283.8
(X_{2003}, X_{2004})	679936	644782.8	679936	656585.8
(X_{2003}, X_{2004})	679936	651769.6	679936	657724.3
(X_{2003}, X_{2004})	679936	660624.9	679936	659623
(X_{2003}, X_{2004})			679936	644899.7
(X_{2003}, X_{2004})			679936	639270.3
(X_{2003}, X_{2004})			679936	639954.3
(X_{2003}, X_{2004})			679936	639625.5
(X_{2003}, X_{2004})			679936	640586.6
(X_{2003}, X_{2004})			679936	654939.6
(X_{2003}, X_{2004})			679936	651579.1
(X_{2003}, X_{2004})			679936	654564.5
(X_{2003}, X_{2004})			679936	655219.6
(X_{2003}, X_{2004})			679936	648053.4

Table 9
The training data set of acrylo nitrile fiber produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	75554	86338	75554	86338
(X_{2001}, X_{2002})	86338	98436	86338	98436
(X_{2002}, X_{2003})	98436	110994	98436	110994
(X_{2003}, X_{2004})	110994	125009	110994	125009
(X_{2003}, X_{2004})	110994	128152.5	110994	126966.3
(X_{2003}, X_{2004})	110994	125359.8	110994	124922.8
(X_{2003}, X_{2004})	110994	124516.1	110994	128135.8
(X_{2003}, X_{2004})	110994	127681.4	110994	122820.4
(X_{2003}, X_{2004})	110994	124729.1	110994	127220
(X_{2003}, X_{2004})	110994	136392.8	110994	124440.4
(X_{2003}, X_{2004})	110994	137531.7	110994	123851.2
(X_{2003}, X_{2004})	110994	137195.8	110994	122558.9
(X_{2003}, X_{2004})	110994	132965.4	110994	123651.7
(X_{2003}, X_{2004})	110994	140103.6	110994	121624.1
(X_{2003}, X_{2004})			110994	139973.3
(X_{2003}, X_{2004})			110994	136888
(X_{2003}, X_{2004})			110994	136460.2
(X_{2003}, X_{2004})			110994	136045
(X_{2003}, X_{2004})			110994	136935
(X_{2003}, X_{2004})			110994	138022.6
(X_{2003}, X_{2004})			110994	141045.7
(X_{2003}, X_{2004})			110994	138505.8
(X_{2003}, X_{2004})			110994	136343.3
(X_{2003}, X_{2004})			110994	136470.6

Table 10
The training data set of rayon fiber produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	75574	52775	75574	52775
(X_{2001}, X_{2002})	52775	102984	52775	102984
(X_{2002}, X_{2003})	102984	95058	102984	95058
(X_{2003}, X_{2004})	95058	81603	95058	81603
(X_{2003}, X_{2004})	95058	84788.7	95058	77970.47
(X_{2003}, X_{2004})	95058	81380.35	95058	80155.49
(X_{2003}, X_{2004})	95058	83248.21	95058	80123.38
(X_{2003}, X_{2004})	95058	80887.66	95058	79779.33
(X_{2003}, X_{2004})	95058	80917.43	95058	80356.72
(X_{2003}, X_{2004})	95058	90610.49	95058	80801.27
(X_{2003}, X_{2004})	95058	89966.4	95058	80798.73
(X_{2003}, X_{2004})	95058	90807.46	95058	84825.16
(X_{2003}, X_{2004})	95058	99114.51	95058	84243.9
(X_{2003}, X_{2004})	95058	97256.9	95058	82146.39
(X_{2003}, X_{2004})			95058	80946.81
(X_{2003}, X_{2004})			95058	82161.17
(X_{2003}, X_{2004})			95058	79154.44
(X_{2003}, X_{2004})			95058	83230.86
(X_{2003}, X_{2004})			95058	79368.55
(X_{2003}, X_{2004})			95058	99590.77
(X_{2003}, X_{2004})			95058	93499.02
(X_{2003}, X_{2004})			95058	94717.74
(X_{2003}, X_{2004})			95058	95660.07
(X_{2003}, X_{2004})			95058	97913.25

Table 11
The training data set of nylon draw texturing yarn produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	48334	49637	48334	49637
(X_{2001}, X_{2002})	49637	46778	49637	46778
(X_{2002}, X_{2003})	46778	49632	46778	49632
(X_{2003}, X_{2004})	49632	44837	49632	44837
(X_{2003}, X_{2004})	49632	44981.63	49632	44548.65
(X_{2003}, X_{2004})	49632	44477.63	49632	44940
(X_{2003}, X_{2004})	49632	45041.73	49632	44669.49
(X_{2003}, X_{2004})	49632	45120.18	49632	45039.1
(X_{2003}, X_{2004})	49632	45169.2	49632	44689.44
(X_{2003}, X_{2004})	49632	38217.42	49632	44844.74
(X_{2003}, X_{2004})	49632	37896.28	49632	44602.96
(X_{2003}, X_{2004})	49632	38276.44	49632	44556.49
(X_{2003}, X_{2004})	49632	37880.17	49632	45128.12
(X_{2003}, X_{2004})	49632	37338.11	49632	45436.52
(X_{2003}, X_{2004})			49632	38201.72
(X_{2003}, X_{2004})			49632	38265.51
(X_{2003}, X_{2004})			49632	38481.35
(X_{2003}, X_{2004})			49632	38045.46
(X_{2003}, X_{2004})			49632	37845.42
(X_{2003}, X_{2004})			49632	38021.62
(X_{2003}, X_{2004})			49632	38299.37
(X_{2003}, X_{2004})			49632	38322.66
(X_{2003}, X_{2004})			49632	38030.74
(X_{2003}, X_{2004})			49632	37880.62

polyester fiber, acrylo nitrile fiber, rayon fiber, nylon draw texturing yarn, nylon fiber yarn, nylon yarn, polyester draw texturing yarn, polyester fiber yarn, polyester yarn, acrylo nitrile fiber yarn, and rayon yarn are shown in Table 5.

Table 12
The training data set of nylon fiber yarn produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	997	1415	997	1415
(X_{2001}, X_{2002})	1415	1250	1415	1250
(X_{2002}, X_{2003})	1250	2756	1250	2756
(X_{2003}, X_{2004})	2756	2881	2756	2881
(X_{2003}, X_{2004})	2756	2815.481	2756	2955.846
(X_{2003}, X_{2004})	2756	3059.656	2756	3004.905
(X_{2003}, X_{2004})	2756	2910.699	2756	3006.288
(X_{2003}, X_{2004})	2756	2786.277	2756	2714.334
(X_{2003}, X_{2004})	2756	3038.245	2756	2761.626
(X_{2003}, X_{2004})	2756	2786.025	2756	2973.583
(X_{2003}, X_{2004})	2756	2915.682	2756	2937.252
(X_{2003}, X_{2004})	2756	2977.787	2756	2785.379
(X_{2003}, X_{2004})	2756	4105.756	2756	2725.92
(X_{2003}, X_{2004})	2756	3968.719	2756	2908.698
(X_{2003}, X_{2004})			2756	2982.006
(X_{2003}, X_{2004})			2756	2819.944
(X_{2003}, X_{2004})			2756	3009.293
(X_{2003}, X_{2004})			2756	2840.418
(X_{2003}, X_{2004})			2756	2848.509
(X_{2003}, X_{2004})			2756	4106.946
(X_{2003}, X_{2004})			2756	4176.783
(X_{2003}, X_{2004})			2756	4201.614
(X_{2003}, X_{2004})			2756	4163.804
(X_{2003}, X_{2004})			2756	4120.451

Table 13
The training data set of nylon yarn produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	88667	81399	88667	81399
(X_{2001}, X_{2002})	81399	132718	81399	132718
(X_{2002}, X_{2003})	132718	174829	132718	174829
(X_{2003}, X_{2004})	174829	176647	174829	176647
(X_{2003}, X_{2004})	174829	181697.4	174829	177609.7
(X_{2003}, X_{2004})	174829	170738.1	174829	179940.9
(X_{2003}, X_{2004})	174829	178380.5	174829	184694.1
(X_{2003}, X_{2004})	174829	177417.1	174829	181228.4
(X_{2003}, X_{2004})	174829	179334.8	174829	177597.4
(X_{2003}, X_{2004})	174829	177025.9	174829	176600.1
(X_{2003}, X_{2004})	174829	181884.8	174829	171575.4
(X_{2003}, X_{2004})	174829	183075.7	174829	175234.9
(X_{2003}, X_{2004})	174829	190748.6	174829	176702.1
(X_{2003}, X_{2004})	174829	193066.9	174829	173326.6
(X_{2003}, X_{2004})			174829	184213.5
(X_{2003}, X_{2004})			174829	185620.7
(X_{2003}, X_{2004})			174829	173037.3
(X_{2003}, X_{2004})			174829	184571.5
(X_{2003}, X_{2004})			174829	180932.5
(X_{2003}, X_{2004})			174829	202767.6
(X_{2003}, X_{2004})			174829	181023.5
(X_{2003}, X_{2004})			174829	201330.8
(X_{2003}, X_{2004})			174829	196151.7
(X_{2003}, X_{2004})			174829	185583.6

3.4. Step 3

Back-propagation neural networks are applied in this research as the learning machine for prediction, and we

Table 14
The training data set of polyester draw texturing yarn produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	306885	296746	306885	296746
(X_{2001}, X_{2002})	296746	325139	296746	325139
(X_{2002}, X_{2003})	325139	290460	325139	290460
(X_{2003}, X_{2004})	290460	299959	290460	299959
(X_{2003}, X_{2004})	290460	302203.9	290460	301287.4
(X_{2003}, X_{2004})	290460	299138.6	290460	300804.9
(X_{2003}, X_{2004})	290460	300913.7	290460	300579.5
(X_{2003}, X_{2004})	290460	300618.4	290460	298889.3
(X_{2003}, X_{2004})	290460	301609.1	290460	300195.3
(X_{2003}, X_{2004})	290460	307058.9	290460	297630.2
(X_{2003}, X_{2004})	290460	306580.3	290460	300521.8
(X_{2003}, X_{2004})	290460	303687.4	290460	302696.7
(X_{2003}, X_{2004})	290460	252786.6	290460	303810.4
(X_{2003}, X_{2004})	290460	254566.9	290460	303334
(X_{2003}, X_{2004})			290460	304806.4
(X_{2003}, X_{2004})			290460	305990.7
(X_{2003}, X_{2004})			290460	305339.3
(X_{2003}, X_{2004})			290460	304664.1
(X_{2003}, X_{2004})			290460	251775.3
(X_{2003}, X_{2004})			290460	252529.4
(X_{2003}, X_{2004})			290460	252005
(X_{2003}, X_{2004})			290460	251366.1
(X_{2003}, X_{2004})			290460	251888.9
(X_{2003}, X_{2004})			290460	252624.4

Table 15
The training data set of polyester fiber yarn produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	44965	40954	44965	40954
(X_{2001}, X_{2002})	40954	44179	40954	44179
(X_{2002}, X_{2003})	44179	45438	44179	45438
(X_{2003}, X_{2004})	45438	38029	45438	38029
(X_{2003}, X_{2004})	45438	37616.22	45438	37600.27
(X_{2003}, X_{2004})	45438	37826.15	45438	38123.25
(X_{2003}, X_{2004})	45438	38493.35	45438	38011.2
(X_{2003}, X_{2004})	45438	38256.7	45438	37903.85
(X_{2003}, X_{2004})	45438	38127.75	45438	37306.89
(X_{2003}, X_{2004})	45438	30838.66	45438	37819.85
(X_{2003}, X_{2004})	45438	30662.46	45438	38329.18
(X_{2003}, X_{2004})	45438	31048.57	45438	38428.17
(X_{2003}, X_{2004})	45438	32000.67	45438	37877.96
(X_{2003}, X_{2004})	45438	32188.58	45438	38118.11
(X_{2003}, X_{2004})			45438	30754.17
(X_{2003}, X_{2004})			45438	30876.89
(X_{2003}, X_{2004})			45438	30660.51
(X_{2003}, X_{2004})			45438	30591.05
(X_{2003}, X_{2004})			45438	31004.42
(X_{2003}, X_{2004})			45438	31935.95
(X_{2003}, X_{2004})			45438	31513.78
(X_{2003}, X_{2004})			45438	31727.4
(X_{2003}, X_{2004})			45438	31756.21
(X_{2003}, X_{2004})			45438	32357.63

Table 16
The training data set of polyester yarn produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	224830	244112	224830	244112
(X_{2001}, X_{2002})	244112	302983	244112	302983
(X_{2002}, X_{2003})	302983	337127	302983	337127
(X_{2003}, X_{2004})	337127	369938	337127	369938
(X_{2003}, X_{2004})	337127	376523.6	337127	364090.6
(X_{2003}, X_{2004})	337127	365430.8	337127	368805
(X_{2003}, X_{2004})	337127	359547.7	337127	377514.7
(X_{2003}, X_{2004})	337127	364330.2	337127	371156.6
(X_{2003}, X_{2004})	337127	377094	337127	366914.6
(X_{2003}, X_{2004})	337127	411145.2	337127	367568.8
(X_{2003}, X_{2004})	337127	393765.6	337127	371708.1
(X_{2003}, X_{2004})	337127	389578.2	337127	374539.8
(X_{2003}, X_{2004})	337127	406354	337127	375630.4
(X_{2003}, X_{2004})	337127	394529.9	337127	367170.5
(X_{2003}, X_{2004})			337127	402503.9
(X_{2003}, X_{2004})			337127	393863.8
(X_{2003}, X_{2004})			337127	390044.2
(X_{2003}, X_{2004})			337127	396666.9
(X_{2003}, X_{2004})			337127	403883.5
(X_{2003}, X_{2004})			337127	385469.7
(X_{2003}, X_{2004})			337127	395493.4
(X_{2003}, X_{2004})			337127	399848.9
(X_{2003}, X_{2004})			337127	406858.1
(X_{2003}, X_{2004})			337127	405359.8

Table 17
The training data set of acrylo nitrile fiber yarn produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	23469	20525	23469	20525
(X_{2001}, X_{2002})	20525	16554	20525	16554
(X_{2002}, X_{2003})	16554	9726	16554	9726
(X_{2003}, X_{2004})	9726	13015	9726	13015
(X_{2003}, X_{2004})	9726	13903.26	9726	12295.18
(X_{2003}, X_{2004})	9726	13271.78	9726	13094.7
(X_{2003}, X_{2004})	9726	12487.52	9726	13492.52
(X_{2003}, X_{2004})	9726	12346.06	9726	12405.33
(X_{2003}, X_{2004})	9726	13076.91	9726	12223.46
(X_{2003}, X_{2004})	9726	14198.5	9726	12780.97
(X_{2003}, X_{2004})	9726	15537.78	9726	12998.65
(X_{2003}, X_{2004})	9726	14797.11	9726	13334.11
(X_{2003}, X_{2004})	9726	18760.39	9726	12567.21
(X_{2003}, X_{2004})	9726	19974.21	9726	12192.91
(X_{2003}, X_{2004})			9726	19778.21
(X_{2003}, X_{2004})			9726	18900.56
(X_{2003}, X_{2004})			9726	20282.69
(X_{2003}, X_{2004})			9726	19625.51
(X_{2003}, X_{2004})			9726	18804.19
(X_{2003}, X_{2004})			9726	14449
(X_{2003}, X_{2004})			9726	16238.88
(X_{2003}, X_{2004})			9726	14769.67
(X_{2003}, X_{2004})			9726	15765.42
(X_{2003}, X_{2004})			9726	14907.87

use Pythia (Runtime Software) to implement the learning task. The topology of the neural network is 1-2-1, and the training data are virtual samples with three different weighted kernels for the prediction of the exports for 2005. For the first five data of nylon fiber, we reform them as four training items, as shown in Table 17. The first datum of each training item is the input variable of the BPN, and the second datum is the output variable. That is why we choose a 1-2-1 structure as the network topology. After the learning of four training items, $(X_{2004}, X_{2005}) = (6359, 6391)$ is the testing item to identify the difference between the actual value of 2005 (X_{2005}) and the predicted

Table 18
The training data set of rayon yarn produced by joining common-increasing virtual samples

	COM10		COM20	
(X_{2000}, X_{2001})	2472	1905	2472	1905
(X_{2001}, X_{2002})	1905	1083	1905	1083
(X_{2002}, X_{2003})	1083	1181	1083	1181
(X_{2003}, X_{2004})	1181	1843	1181	1843
(X_{2003}, X_{2004})	1181	1821.577	1181	1899.577
(X_{2003}, X_{2004})	1181	1864.727	1181	1909.035
(X_{2003}, X_{2004})	1181	1821.597	1181	1870.718
(X_{2003}, X_{2004})	1181	1877.056	1181	1798.755
(X_{2003}, X_{2004})	1181	1828.238	1181	1972.651
(X_{2003}, X_{2004})	1181	2329.385	1181	1890.857
(X_{2003}, X_{2004})	1181	2450.427	1181	1864.758
(X_{2003}, X_{2004})	1181	2321.922	1181	1891.415
(X_{2003}, X_{2004})	1181	2101.066	1181	2234.136
(X_{2003}, X_{2004})	1181	2278.38	1181	2166.813
(X_{2003}, X_{2004})			1181	2175.517
(X_{2003}, X_{2004})			1181	2076.318
(X_{2003}, X_{2004})			1181	2098.503
(X_{2003}, X_{2004})			1181	2258.873
(X_{2003}, X_{2004})			1181	2295.935
(X_{2003}, X_{2004})			1181	2271.195
(X_{2003}, X_{2004})			1181	2330.206
(X_{2003}, X_{2004})			1181	2312.338
(X_{2003}, X_{2004})			1181	2356.135
(X_{2003}, X_{2004})			1181	2273.502

Table 19
The MSE between original data sets and expanded data sets with common increasing weight series concerning different materials

Training data set	Original (4 data)	COM10	COM20
Nylon fiber	0.753845	0.429762	0.499988
Polyester fiber	0.914378	0.209018	0.458741
Acrylo nitrile fiber	0.365828	0.252599	0.103462
Rayon fiber	0.932466	0.572551	0.603341
Nylon draw texturing yarn	0.791495	0.10101	0.275776
Nylon fiber yarn	1.167906	0.533143	0.145837
Nylon yarn	0.258549	0.242562	0.632689
Polyester draw texturing yarn	0.160151	0.302801	0.07005
Polyester fiber yarn	0.22779	0.528879	0.137599
Polyester yarn	0.506568	0.182162	0.038905
Acrylo nitrile fiber yarn	0.361231	0.448004	0.277095
Rayon yarn	0.207452	0.131435	0.069439

value of the network ($\widehat{X}_{2005}$). We repeat this experimental process ten and twenty times respectively, and calculate the MSE of each virtual sample set. As for the virtual samples based on GIKDE, this research regards them all as X_{2004}. For example, the number of training items will be 14 after we add the virtual sample set COM10($n = 10$). Table 6 shows the training data set by joining common increasing virtual samples that include (X_{2000}, X_{2001}), (X_{2001}, X_{2002}), (X_{2002}, X_{2003}), (X_{2003}, X_{2004}), (X_{2003}, X_{2004}), $\ldots, (X_{2003}, X_{2004})$, 14 data in all. The latter 10 data (X_{2003}, X_{2004}) are virtual samples generated from GIKDE. If we join the virtual sample set COM20($n = 20$), the number of training items becomes 24 and the latter 20 data (X_{2003}, X_{2004}) are virtual samples generated from GIKDE. The training items of nylon fiber made by joining the virtual sample sets COM10 and COM20 are shown in Table 18. In the same way, the training items of polyester fiber, acrylo nitrile fiber, rayon fiber, nylon draw texturing yarn, nylon fiber yarn, nylon yarn, polyester draw texturing yarn, polyester fiber yarn, polyester yarn, acrylo nitrile fiber yarn, and rayon yarn made by joining virtual sample sets COM10 and COM20 are shown in Tables 7–18, respectively.

3.5. Step 4

Table 19 shows MSEs of both the original data sets and expanded data sets (that is, the original data and virtual data). Comparing all kinds of raw materials with their MSEs, the results show that MSE decreases when we join virtual sample sets. Take nylon fiber for example, the MSE of the original data set (4 data) is 0.753845 and the MSE of the expanded data set is 0.429762 (COM10), and 0.499988 (COM20). In other words, adding virtual samples helps reduce the MSE of predicted values.

3.6. Step 5

For each data set, we use GIKDE to establish $\hat{f}(X_{2006})$ based on $X_{2000}, \ldots, X_{2005}$. This research generates virtual samples for $\hat{f}(X_{2006})$ through Monte Carlo methods. To combine original data with virtual samples, we use the

Table 20
The prediction of exports of polyester fibers in Taiwan for 2006 (unit: tons)

Product		2000	2001	2002	2003	2004	2005	Prediction for 2006
1	Nylon fiber	12,297	9669	8928	6442	6359	6319	6182
2	Polyester fiber	671,271	586,912	677,981	679,936	659,809	566,502	599,224
3	Acrylo nitrile fiber	75,554	86,338	98,436	110,994	125,009	114,246	117,006
4	Rayon fiber	75,574	52,775	102,984	95,058	81,603	67,504	85,336
5	Nylon draw texturing yarn	48,334	49,637	46,778	49,632	44,837	44,581	47,752
6	Nylon fiber yarn	997	1415	1250	2756	2881	1327	2455
7	Nylon yarn	88,667	81,399	132,718	174,829	176,647	158,004	167,334
8	Polyester draw texturing yarn	306,885	296,746	325,139	290,460	299,959	257,407	286,422
9	Polyester fiber yarn	44,965	40,954	44,179	45,438	38,029	33,546	37,981
10	Polyester yarn	224,830	244,112	302,983	337,127	369,938	339,333	351,228
11	Acrylo nitrile fiber yarn	23,469	20,525	16,554	9,726	13,015	14,798	14,990
12	Rayon yarn	2472	1905	1083	1181	1843	1529	1447

learning procedure of BPN with these available data to predict the exports for 2006. Table 20 shows the predicted values using the virtual sample sets COM20.

4. Conclusions and discussion

The operating environment that Taiwan's polyester fiber textile industry confronts is unstable due to increased competition. The industry needs to find a way to forecast and handle future demands more precisely under such conditions. If such predictions are incorrect, the industry will suffer from the waste of capitalized cost (the condition that the supply exceeds the demand) and the loss of customers (the condition that the demand exceeds the supply).

The periodic characteristic of data influences the accuracy of prediction, and the data of some related products have regional and seasonal changes that make accurate forecasting quite difficult. Moreover, data trends also make the result of prediction inconsistent. Notably, the export data of polyester in this research are affected not only by the weather, but also by other causes such as the output of cotton, the price of petroleum, and market demand. The GIKDE acquires only a small amount of data for prediction, since it combines the concepts of time series and randomization. Furthermore, it can deal with data of different types, such as independent, dependent, numerical, or nominal. GIKDE thus creates extra information to raise the prediction accuracy.

We believe that it is essential to accurately anticipate the demand for exports to further improve the bullwhip effect of the imbalance between supply and demand. The results for the prediction of polyester exports in this research can help factory managers make more complete plans for the supply chain and marketing. In other words, the results of this study offer manufacturers crucial information for decision making concerning inventory management, marketing channels, modulation of product structure, and budget planning.

Owing to limitations in data collection, this research focuses only on Taiwanese exports and uses a uni-variate process to implement the prediction. We suggest researchers who are interested in this topic collect the export data of other countries and use multi-variate methods for prediction to make the analysis more complete. Moreover, it should be noted that the export of polyester products is influenced by the demand, the price of raw materials, natural disasters (such as earthquakes), and political issues (such as intervention from mainland China). Further research could include these factors into the proposed model to make the prediction more accurate. Regarding the amount of virtual samples, the results show that more virtual samples do not ensure a better prediction. This may have a relationship with the data structure, and is worth further research efforts.

References

Li, D. C., & Lin, Y. S. (2006). Using virtual sample generation to build up management knowledge in the early manufacturing stages. *European Journal of Operational Research, 175*, 413–434.

Li, D. C., & Lin, Y. S. (in press). Learning management knowledge for manufacturing systems in the early stages using time series data. *European Journal of Operational Research*.