Applied Mathematics & Information Sciences
An International Journal

Analysis of predicting the diversity regional logistics demand based on SVR: the case of Sichuan in China

Si Chen[1], *Mi Gan*[1] *and Yinying Tang*[1]

[1]School of Transportation and Logistics, Southwest Jiaotong University, Sichuan 610031 P. R. China

Received: 12 Sep. 2012; Revised 9 Nov. 2012; Accepted 29 Nov. 2012
Published online: 1 Mar. 2013

Abstract: There has been worldwide concern for the regional logistic planning, and the prediction of logistic demand is the basic and important parts in the logistic planning. And in this paper the prediction of logistic demand has been divided into three parts, which are agricultural logistic demand, manufacturing logistic demand and business logistic demand. In this way the prediction would show the different influence factors of industry constructions in different area. In regional logistic demand prediction, the method of Support Vector Regression (SVR) is considered as a useful technique for data prediction. And the analyses of different industries are built up on the system of influence factors with different industries. And by the studying the traffic volumes and several influence factors of Sichuan from 2001 to 2010, which verified the suggestions of logistic prediction in regional logistic planning.

Keywords: Logistics demand, SVR, agricultural industry, manufacturing industry, Influence factors

1. Introduction

As a new field in the economic growth of China, Logistics system plays an increasing important role in economic society. Since healthy logistics system can provide high performance to support the whole society, logistics system have been considered as an indispensable role in the regional economic. Regional logistics demands are keys to regional economy and regional logistics system, so they are important parts of regional economic society.

By the reason of the diversity features of regional logistic demand, logistic demand prediction process may change with the different features. Recently most of researchers make their focus on the prediction methods[1]. And the features are shown different influence on the logistic demand and can be sorted out into many groups. The groups can be classified by time, different types products, different level of business, and different activities of logistics process. In short, all the features of logistic demand can be classified into the groups of different industries.

Regional logistics predictions seek to predict the generation, distribution and the quantity of logistics demand. However the quantity of logistics is not defined precise, most of researchers just can use the freight traffic as the key indicator instead of the real quantity of logistic demand[2]. Currently prediction methods of regional logistic demand usually make use of freight traffic and freight ton-kilometer to represent the actual logistics volume[3]. But in the real world the logistic volume is the accumulating of all the activities in the logistic process. So a transfer model which makes use of the freight traffic to represent logistic volume is needed to be built.

Support Vector Regression (SVR) is considered as a useful technique for data prediction. It produced by support vector regression only depends on a subset of the training data. And SVR is the most common application of SVM (Support Vector Machine). SVM is a new learning machine for two-group classification problems. It was first introduced by Vaonik and Cortes in 1995[4, 5]. The method of Support Vector Regression (SVR) can minimize the observed training error, and attempt to minimize the generalized error bound so as to achieve generalized performance[6, 7].

* Corresponding author: e-mail: cdchensi@gmail.com

2. The influence factors of different industries

In this section the whole regional logistic demand has been divided into three industries logistic demand. The three industries are agricultural industry, manufacturing industry and business. And the logistic demand volume will develop with the features of industries.

2.1. Agricultural Industry

The agricultural industry here includes farming, pasturage, forest industry and fishery. When the total output value of agricultural industry is large enough in a regional economic, it is better to analyze the influence factors of agricultural industry to logistic volume separately. The output and development of this industry almost depend on nature. And the products of agriculture have some unique features, such as seasonal nature, regional characteristics and etc. . In the mean time, the whole industry is kind of a basic need to a human beings society.

There are several features of agriculture logistic volume should be analyzed here. First, the logistic flow of agriculture is large and it covers widely. And because of the unique features of the products, the products are necessary as well as where people live. Second, the products of agriculture need particular equipments for transportation and distribution. Most of foods are easily spoiled after harvest, so they need special equipments, warehouses and technologies to keep the foods fresh during the logistic process. Thus it is harder to transport agricultural products than other industry products. The third, the logistic process of agricultural industry has high requirement on the timeliness. The production materials, such as seeds, farm chemicals and other necessary equipments for production, have to follow the schedule of nature. So time of delivery too early or too late may disorganize the whole production planning. And the products, neither the primary products nor the products which have been worked, all have time limit to consume. The fourth, the logistic flows of agriculture are not a balance flows. There are a lot of influence factors to the logistic demand under the agricultural industry. And here all the influence factors have been divided into four parts. First, economic influence factor is a basic one. The development of economic is the foundation of agricultural industry. All the industries in the regional will be influenced by the regional economic. When the agricultural industry is influenced, the agricultural logistics will be impacted too. And the logistic environment is also an important influence factor to the agricultural logistics too. The perfect logistic infrastructures, the high level logistic professionals and the quantity of agricultural logistic in the former years are positive influence factors to the future agricultural logistics. Since the logistic demand is a ramification of agricultural industry, the type of the agriculture, the amount of the production and the technology of agriculture all are the important factors.

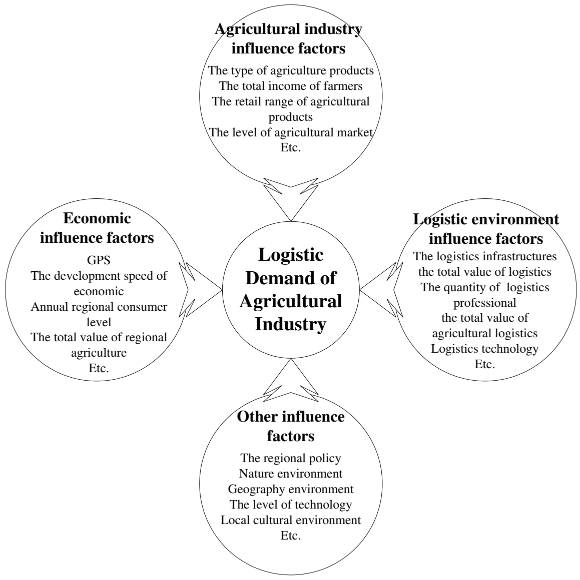

Figure 2.1 The influence factors of agricultural logistic demand

Here six representative indexes have been choose to model the influence of agricultural logistics. They are regional GDP; the total value of regional agriculture; the regional population; the total value of agricultural logistics; annual regional consumer level; the total value of logistics.

2.2. Manufacturing Industry

The whole manufacturing industry can be divided into two parts. One is light manufactory industry, which includes textile, food processing, another is heavy manufactory, which includes chemical industry, metallurgy industry and etc.. First, the logistic demand of manufacturing industry is huge. The purchasing of production materials, the process of producing and the delivery, storage and transportation of the products and half-finished products, all the steps of industry production can not continue without logistics service. Second, almost all the manufacturing logistics have a plan to follow, so the order quantity and batches are comparatively stable. And since every manufactory tends to build a long cooperation relationship between customers and suppliers, the quantity of materials and products can be predicted. Thus all the logistic flows of manufacturing industry are smooth. Third, in a supply chain there is always a core factory to control the speed of that supply chain. And most of the core factory is the manufactory. There are a lot of influence factors to the logistic demand under the manufacturing industry. The influence factors have been divided into five parts. First, economic influence factor is a basic one. The manufacturing industry is a result of development economic, so every changes in economic will effect the change in manufacturing industry. They are regional GDP, industrial added value, industrial fixed assets investment, annual regional consumer level and etc.. The second part is the logistic environment, which is also an important influence factor to the manufacturing logistics too. Different structure means different proportion between light and heavy industry, and means the different consumers and consumption areas. So the quantity and range of logistic may change with the structure. The regional environment also can influence the manufacturing logistics.

There are six representative indexes had been choose to model the influence of agricultural logistics. They are regional GDP; the total value of regional manufacturing industry; the number of employees in manufacturing industry; the total value of logistics; industrial added value.

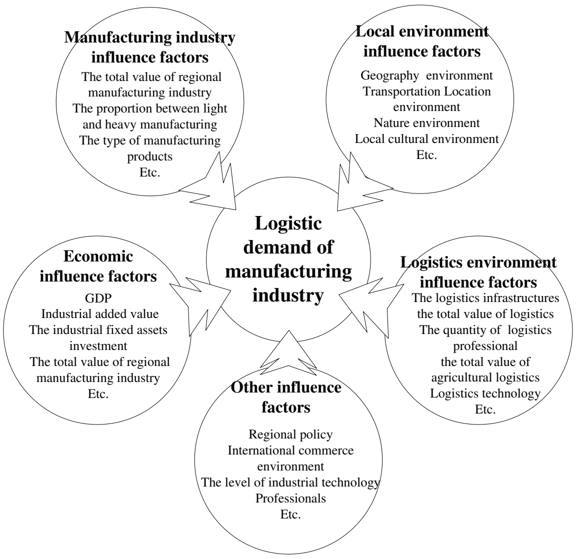

Figure 2.2 The influence factors of manufacturing logistic demand

2.3. Business

Business is an industry which can connect the customers and manufacturers. And in the operation process of business the value of products will be achieved. The features of business logistics are shown below. First of all, business has a lot of kinds customers, which include companies, factories schools, personals and etc.. And in the process of business there are all types of commodities and many kind of requires. All of this makes the whole business logistics become complicated. Finally, the level of logistic standardization is normally lower than other industries logistics. One reason is that the products have different shapes, characteristics and different requires to the logistic system. Another reason is that different place has different logistics infrastructures, which make it much harder to improve the level of logistic standardization.

There are a lot of influence factors to the logistic demand under the business. And the influence parts also have been divided into six parts. First, the economic influence factors are the basic in all factors. Second, the structure of business is also an influence factors to the business logistics. The whole process of business is also the process of business logistics. The third part of influence is the logistic environment. The influence of logistics environment is the same as it mentioned in section 2.1 and section 2.2. The culture environment is the forth influence factor. Business is kind of industry communicate with person, so person with different culture background may have different business behaviors. And the nature environment is also an important influence factors to the business logistics. There are many other factors, such as policy, population in the regional, the location and so on.

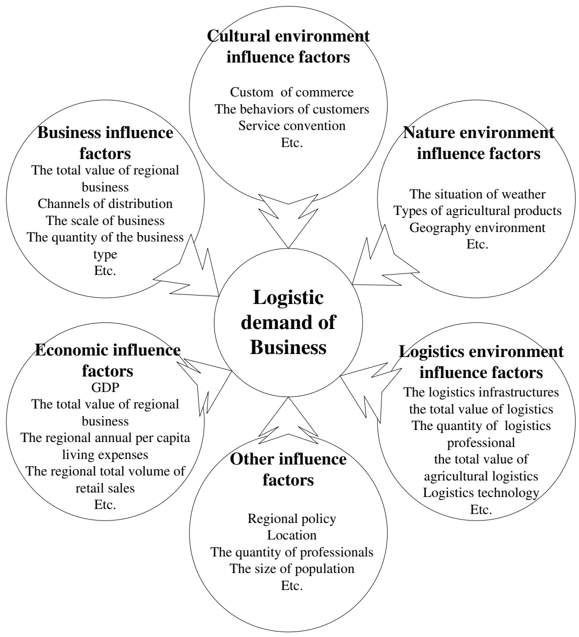

Figure 2.3 The influence factors of business logistic demand

There are six representative indexes which are used to model the influence of agricultural logistics. They are regional GDP; the total value of regional business; the total value of business logistics; the total value of logistics in the regional; the regional annual per capita living expenses; the regional total volume of retail sales.

3. The model of SVR

In this section we try to introduce the basic model and algorithm of SVR[3–6].

It supposes the training set is

$$S = ((x_1, y_1), \cdots, (x_i, y_i)) \tag{1}$$

In which $i = 1, \cdots, l$ and $x_i \in R^d, y_i \in R$. In the mean time try to find the optimal function $f(x) =< w \cdot x > +b$, and $w, x_i \in R^d, b \in R$. It makes the functional to minimum:

$$R(f) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l} L^{\varepsilon}(x_i, y_i, f) \tag{2}$$

In which $\|w\|^2$ means structural risk, C is the penalty factor, ε in the $L^{\varepsilon}(x_i, y_i, f)$ is the insensitive factor. When ε in the $f(x) =< w \cdot x > +b$, it becomes error precision. So the problem can be expressed in Formula 3, when it tries to find the minimum of w.

$$\begin{gathered} \min_{w,b} \tfrac{1}{2}\|w\|^2 \\ s.t. \begin{cases} [(w \cdot x_i) + b] - y_i \le \varepsilon \\ y_i - [(w \cdots x_i) + b] \le \varepsilon \\ i = 1, 2, \cdots, l \end{cases} \end{gathered} \tag{3}$$

Here we introduce slack variables ξ and ξ^*, the problem becomes into Formula 4

$$\begin{gathered} \min_{w,b} \tfrac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \\ s.t. \begin{cases} [(w \cdot x_i) + b] - y_i \le \varepsilon + \xi_i \\ y_i - [(w \cdot x_i) + b] \le \varepsilon + \xi_i^* \\ i = 1, 2, \cdots, l \quad \xi_i, \xi_i^* \ge 0 \end{cases} \end{gathered} \tag{4}$$

Then we introduce the lagrange function and the dual variables sets $\{\alpha, \alpha^*, \eta, \eta^*\}$, the function becomes into Formula 5.

$$\begin{aligned} & L(w, b, \zeta, \zeta^*, \alpha, \alpha^*, \eta, \eta^*) = \\ & C\sum_{i=1}^{l}(\zeta_i + \zeta_i^*) + \tfrac{1}{2}(w \cdot w) \\ & -\sum_{i=1}^{l}\alpha_i(\varepsilon + \zeta_i + y_i - (w \cdot x_i) - b) \\ & -\sum_{i=1}^{l}\alpha_i^*(\varepsilon + \zeta_i^* - y_i + (w \cdot x_i) + b) \\ & -\sum_{i=1}^{l}(\eta_i\zeta_i + \eta_i^*\zeta_i^*) \end{aligned} \tag{5}$$

And the dual problem is:

$$\begin{aligned} & \min_{\alpha,\alpha^*} \tfrac{1}{2}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i \cdot x_j) \\ & +\varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) - \sum_{i=1}^{l} y_i(\alpha_i^* - \alpha_i) \\ & s.t. \begin{cases} \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_j, \alpha_j^* \le C, i = 1, \cdots, l \end{cases} \end{aligned} \tag{6}$$

Since w can be expressed by the training set x_i, the function $f(x)$ and is shown as below:

$$\begin{aligned} f(x) &= \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)(x_i \cdot x) + b \\ b &= y_i - (w \cdot x_i) - \varepsilon \quad \alpha_i \in (0, C) \\ b &= y_i - (w \cdot x_i) + \varepsilon \quad \alpha_i^* \in (0, C) \end{aligned} \tag{7}$$

In the meantime, based on the optimality conditions of Karush-Kuhn-Tucker, the relational express is shown in Formula 8.

$$\begin{aligned} & \alpha_i[\varepsilon + \xi_i - y_i + (w \cdot x_i) + b] = 0 \\ & \alpha_i^*[\varepsilon + \xi_i^* - y_i + (w \cdot x_i) + b] = 0 \\ & (C - \alpha_i)\xi_i = 0 \\ & (C - \alpha_i^*)\xi_i^* = 0 \\ & \xi_i\xi_i^* = 0, \alpha_i\alpha_i^* = 0 \end{aligned} \tag{8}$$

The basic thought of the SVR is make use of a mapping to map the input to another space, and analyzing and calculate in the new space. And the mapping means a kernel function here. Different kernel function has different algorithm and maps different space. There is no certain method to determine the kernel function now. The general method is by the experience and try. Here four kernel functions are usually used [7].

(1) Linear kernel function:

$$K(x_i, x_j) =< x_i, x_j > \tag{9}$$

(2) Polynomial kernel function:

$$K(x_i, x_j) = (\gamma < x_i, x_j > +r)^d, \gamma > 0 \tag{10}$$

(3)Radial basis function (RBF) kernel function:

$$K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2), \gamma > 0 \& \gamma = -\frac{1}{2\alpha^2} \tag{11}$$

(4)Sigmoid kernel function:

$$K(x_i, x_j) = \tanh(\gamma < x_i, x_j > +r) \tag{12}$$

There are four steps to calculate based on the SVR:

Step 1. Try to divide all the data of sets into two groups. One is the training group, another is the test group. And the amount of training sets should larger than test sets.

Step 2. Select a appropriate kernel function $K(x_i, x_j)$, estimate the parameters. This is the key step.

Step 3. Constructing and solving Formula 6. And get the parameters optimum solution of
$\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_1^*, \cdots, \bar{\alpha}_i, \bar{\alpha}_i^*)^T$.

Step 4. Make use of the optimum solution of $\bar{\alpha}$ to get the optimal function $f(x)$, and predict with $f(x) = \sum_{i=1}^{l}(\bar{\alpha}_i^* - \bar{\alpha}_i)K(x_i \cdot x) + \bar{b}$ finally.

4. A case of Sichuan

Based on the analysis in the above sections, we choose the data of Sichuan province in China to verify the feasibility of the method of SVR. And we predict the logistic demand based on the three industries, which constitute the whole regional economic. Since different industry has different influence factors, the prediction of three industries may be more precise than the prediction of the whole regional.

The source of the data comes from China Logistics Yearbook (2002-2011) [8] and Sichuan Statistical Yearbook (1986-2011) [9]. The China Logistics Yearbook started from year 2002, so the data about logistics we can find is start from year 2000 to year 2010. There are only 11 sets of data can be used in the three industries prediction. We set 9 sets of data as training sets, the other 2 sets as the testing sets.

We process every data to reduce the level of calculation. Actually we set $x_{ij} = x_{ij}/x_{\max j}$, and x_{ij} is the jth data of the ith index. And $x_{\max j}$ is the maximal value of the ith index. And we also set the $x_{\max j} = 0.999$. Then the results of data processing are show in Table 4.1, 4.2 and 4.3.

Table 4.1 The influence factors data processing results of agricultural logistic demand

Indexes Years	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	Q_1
2000	0.229	0.363	0.934	0.182	0.312	0.143	0.422
2001	0.250	0.376	0.937	0.174	0.331	0.151	0.417
2002	0.275	0.405	0.941	0.192	0.356	0.171	0.440
2003	0.310	0.437	0.948	0.233	0.392	0.186	0.487
2004	0.371	0.552	0.955	0.239	0.447	0.216	0.495
2005	0.430	0.602	0.960	0.296	0.505	0.292	0.505
2006	0.506	0.637	0.969	0.431	0.550	0.425	0.509
2007	0.615	0.826	0.979	0.465	0.643	0.447	0.568
2008	0.733	0.903	0.990	0.718	0.742	0.694	0.871
2009	0.823	0.904	0.998	0.920	0.839	0.811	0.880
2010	0.999	0.999	0.999	0.999	0.999	0.999	0.999

x_{11} is regional GDP; x_{12} is the total value of regional agriculture; x_{13} is the regional population; x_{14} is the total value of agricultural logistics; x_{15} is annual regional consumer level; x_{16} is the total value of logistics; Q_1 is the traffic volume of agricultural industry.

x_{21} is regional GDP; x_{22} is the total value of regional manufacturing industry; x_{23} is the total value of regional manufacturing logistics; x_{24} is the number of employees in manufacturing industry; x_{25} is the total value of logistics; x_{25} is industrial added value; Q_2 is the traffic volume of manufacturing industry.

x_{31} is regional GDP; x_{32} is the total value of regional business; x_{33} is the total value of business logistics; x_{34} is the total value of logistics in the regional; x_{35} is the regional annual per capita living expenses; x_{36} is the regional total volume of retail sales; Q_3 is the traffic volume of business industry.

Table 4.2 The influence factors data processing results of manufacturing logistic demand

Indexes Years	x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	Q_2
2000	0.229	0.155	0.102	0.155	0.143	0.138	0.249
2001	0.250	0.169	0.107	0.169	0.151	0.184	0.269
2002	0.275	0.185	0.120	0.185	0.171	0.236	0.292
2003	0.310	0.216	0.136	0.216	0.186	0.294	0.298
2004	0.371	0.271	0.166	0.271	0.216	0.359	0.360
2005	0.430	0.340	0.242	0.340	0.292	0.417	0.433
2006	0.506	0.423	0.373	0.423	0.425	0.505	0.457
2007	0.615	0.528	0.402	0.528	0.447	0.642	0.516
2008	0.733	0.667	0.668	0.667	0.694	0.819	0.845
2009	0.823	0.764	0.787	0.764	0.811	0.860	0.868
2010	0.999	0.999	0.999	0.999	0.999	0.999	0.999

Table 4.3 The influence factors data processing results of business logistic demand

Indexes Years	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	Q_3
2000	0.229	0.273	0.264	0.143	0.312	0.245	0.640
2001	0.250	0.307	0.293	0.151	0.331	0.276	0.729
2002	0.275	0.342	0.338	0.171	0.356	0.304	0.769
2003	0.310	0.386	0.336	0.186	0.392	0.337	0.731
2004	0.371	0.442	0.382	0.216	0.447	0.384	0.824
2005	0.430	0.500	0.466	0.292	0.505	0.441	0.824
2006	0.506	0.585	0.607	0.425	0.550	0.510	0.735
2007	0.615	0.684	0.594	0.447	0.643	0.590	0.754
2008	0.733	0.804	0.771	0.694	0.742	0.726	0.999
2009	0.823	0.916	0.824	0.811	0.839	0.846	0.960
2010	0.999	0.999	0.999	0.999	0.999	0.999	0.959

The simulation environment is MATLAB 2010a. And we select radial basis function (RBF) to be the kernel function of this simulation. Parameters are achieved with the method of Grid search. Then the kernel function of logistic demand of agricultural industry is $K(x_i,x_j) = \exp(-\gamma\|x_i - x_j\|^2), \gamma > 0$, and $C = 3.1$, $g = 0.32$. The test results are shown in Table 4.4.

Table 4.4 The traffic volume prediction results of agriculture industry

Year	True value (Billion ton kilometer)	Predicted value (Billion ton kilometer)	Error value (Billion ton kilometer)	Error rate (%)
2009	197.24	196.7	0.54	0.273
2010	224.01	219.9	4.11	1.83

And the kernel function of logistic demand of manufacturing industry is $K(x_i,x_j) = \exp(-\gamma\|x_i - x_j\|^2), \gamma > 0$, and $C = 2.5$, $g = 0.25$. The test results are shown in Table 4.5.

And the kernel function of logistic demand of business is $K(x_i,x_j) = \exp(-\gamma\|x_i - x_j\|^2), \gamma > 0$, and $C = 1.6$, $g = 0.14$. The test results are shown in Table 4.6.

Table 4.5 The traffic volume prediction results of manufacturing industry

Year	True value (Billion ton kilometer)	Predicted value (Billion ton kilometer)	Error value (Billion ton kilometer)	Error rate (%)
2009	1015.26	1002.62	12.64	1.25
2010	1169.64	1147.1	22.54	1.93

Table 4.6 The traffic volume prediction results of business

Year	True value (Billion ton kilometer)	Predicted value (Billion ton kilometer)	Error value (Billion ton kilometer)	Error rate (%)
2009	316.5	296.24	20.26	6.4
2010	316.35	309.7	6.65	2.1

And the regional traffic volume prediction results can be achieved by accumulating all the industries results. And the results are shown in Table 4.7.

Table 4.7 The regional traffic volume prediction results (the sum of all the industries)

Year	True value (Billion ton kilometer)	Predicted value (Billion ton kilometer)	Error value (Billion ton kilometer)	Error rate (%)
2009	1529	1495.56	33.44	2.19
2010	1710	1676.7	33.3	1.95

And here we try to compare the results of the different industries prediction and the whole regional prediction. So we eliminate the influence factors of industry logistics, and use the same method to find the parameters of the kernel function. There are 7 influence factors which are choose to predict the trend of development. And since all of the factors are from Sichuan Statistical Yearbook, there are 26 sets of data which can be utilized. We choose 21 sets of data from year 1985 to 2005 as training sets and choose the rest five sets of data from year 2006 to 2010 as testing sets. Then the results of data processing are show in Table 4.8.

And according to the order of indexes in the table 4.8, the meanings are regional GDP; the total value of regional agriculture; the regional population; the regional annual per capita living expenses; total value of regional manufacturing industry; the total value of regional business; the regional total volume of retail sales. And the last one is the regional traffic volume.

And the kernel function of regional logistic demand is $K(x_i,x_j)=\exp(-\gamma\|x_i-x_j\|^2)$, $\gamma>0$ and $C=32$, $g=0.58$. The test results are shown in Table 4.9.

The error rates of different industries predictions are around 2%, which is really low. Although we just have 11 sets of data to predict, the results are compared accurate with the regional prediction, some of which error rates are larger than 30%.

Table 4.8 The influence factors data processing results of regional logistic demand

Indexes / Years	x_{41}	x_{42}	x_{43}	x_{44}	x_{45}	x_{46}	x_{47}	Q_4
1985	0.025	0.058	0.824	0.042	0.017	0.018	0.040	0.215
1986	0.027	0.062	0.835	0.046	0.019	0.021	0.046	0.216
1987	0.031	0.072	0.846	0.053	0.022	0.025	0.055	0.273
1988	0.038	0.088	0.857	0.065	0.028	0.032	0.070	0.290
1989	0.043	0.098	0.867	0.075	0.031	0.038	0.078	0.284
1990	0.052	0.119	0.877	0.086	0.037	0.045	0.051	0.343
1991	0.059	0.126	0.883	0.093	0.045	0.053	0.059	0.412
1992	0.069	0.139	0.888	0.102	0.051	0.064	0.069	0.532
1993	0.086	0.162	0.893	0.117	0.067	0.080	0.084	0.570
1994	0.116	0.228	0.900	0.167	0.090	0.109	0.109	0.458
1995	0.142	0.273	0.907	0.201	0.112	0.141	0.141	0.506
1996	0.167	0.312	0.913	0.230	0.132	0.167	0.167	0.380
1997	0.189	0.342	0.918	0.254	0.142	0.193	0.190	0.405
1998	0.202	0.357	0.924	0.274	0.145	0.218	0.207	0.332
1999	0.212	0.354	0.929	0.287	0.148	0.242	0.223	0.336
2000	0.229	0.363	0.934	0.312	0.155	0.273	0.245	0.349
2001	0.250	0.376	0.937	0.331	0.169	0.307	0.276	0.379
2002	0.275	0.405	0.941	0.356	0.185	0.342	0.304	0.412
2003	0.310	0.437	0.948	0.392	0.216	0.386	0.337	0.409
2004	0.371	0.552	0.955	0.447	0.271	0.442	0.384	0.470
2005	0.430	0.602	0.960	0.505	0.340	0.500	0.441	0.521
2006	0.506	0.637	0.969	0.550	0.423	0.585	0.510	0.521
2007	0.615	0.826	0.979	0.643	0.528	0.684	0.590	0.573
2008	0.733	0.903	0.990	0.742	0.667	0.804	0.726	0.885
2009	0.823	0.904	0.998	0.839	0.764	0.916	0.846	0.894
2010	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999

Table 4.9 The regional traffic volume prediction results

Year	True value (Billion ton kilometer)	Predicted value (Billion ton kilometer)	Error value (Billion ton kilometer)	Error rate (%)
2006	891	1123.3	-232.3	26.07
2007	979	1356.7	-377.7	38.58
2008	1513	1503.3	9.7	0.64
2009	1529	1501.6	27.4	1.72
2010	1710	1539	171	10

5. Conclusion

The basic suggestion of this paper is that the sum of every industry logistic demand prediction is more accurate than the result of regional logistic demand prediction without considering the difference development influence factors of industries. The prediction results of different industries prediction are much more precise than the traditional regional prediction.

And the prediction method of Support Vector Regression (SVR) is considered as a useful technique for data prediction. The results of different industries prediction is tested well with the method of SVR, and which is proved to be more accurate as compared to the prediction of whole regional without considering the difference of industries. In the future study, the more prediction methods should be tested, in order to find much more appropriate one. The Study on the division of industries in an area will focus on the new features of traditional industries and the new industries which are founded with the development of economic and technology. And the data mining is also very important to the study of prediction.

References

[1] George P. Boretos. IS model: A general model of forecasting and its applications in science and the economy. Technological Forecasting & Social Change 78,1016-1028(2011)

[2] Yue Wang. Study on the Forecasting Methods of Regional Logistics Demand. Logistics Technology, **28** ,168-171(2009)

[3] Corinna Cortes. Vladimir Vapnik, Support-Vector Networks,Machine Learning. **20**, 273-297 (1995)

[4] Nello Cristianini. John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Publishing House of Electronics Industry. BeiJing .(2004)

[5] Smola A.JScholkopf B. A tutorial on support vector regression.Statistics and Computing **14**,199-222(2004)

[6] Hongmin Xu, HaiyingWang, Jin Liang, Shuai Huang. The algorithmand application of support vector regression. Journalof Beijing Institute of Petrochemical Technology, **18**, 62-67(2010)

[7] Smola A.JScholkopf B. A tutorial on support vector regression.Statistics and Computing. **14**, 199 (2004)

[8] Sichuan Statistical Bureau. China Logistics Yearbook. China Statistics Press (2002-2011)

[9] China Federation of Logistics and Purchasing (CFLP) Sichuan Statistical Yearbook. China Logistics Publish House.(1986-2011).

Si Chen received the BS degree in logistic management from Southwest Jiaotong University in 2005. And she got MS degree in logistics engineering from Southwest Jiaotong University in 2008. And She now is a Ph.D. candidate in Transportation planning and management of Southwest Jiaotong University. She studied in University of Arizona from 2010 to 2011.And she is a member of TRB. Her research interests are logistics management, supply chain management, transportation network optimization and logistics system optimization.

Mi Gan received the PHD degree in Logistics enineering from Southwest Jiaotong University in 2012. And she is a teacher in Southwest Jiaotong University. Her research interests are supply chain management, logistics network, mathematical model, transportation planning.

Yinying Tang received the PHD degree in Transportation planning and management from Southwest Jiaotong University in 2000. Now she is a currently an associate professor in Southwest Jiaotong University. Her research interests are in the areas of Logistics management, Transportation planning.