Send Orders for Reprints to reprints@benthamscience.ae

The Open Automation and Control Systems Journal, 2015, *7*, 809-815 809

Open Access

Agricultural Information Level Evaluation and Prediction Research Based on Supporting Vector Regression

Xia Zhang[1,*], Suzhen Wang[1], Lin Wang[2] and Qi Wang[1]

[1]*Hebei University of Economics and Business, Shijiazhuang, Hebei, 050061, China*

[2]*Shijiazhuang Agriculture and husbandry bureau, Shijiazhuang, Hebei, 050000, China*

Abstract: There are the following problems on research of the agricultural information level evaluation and prediction in China and other countries. (1) People do not find a generally accepted evaluation index system for the agricultural information. (2) The methods to determine the index weight are usually based on subjective judgment, cannot objectively reflect the correlation of index. (3) Most evaluation and prediction methods of agricultural information level study on social information from the economic viewing point and are based on linear transformation, they cannot exactly reflect the characteristics of nonlinear fitting between agricultural information index system and evaluation results. In order to solve the above mentioned problems, this paper first constructs a set of agricultural information index system, then gets the weight of each index according to the method of entropy and does evaluation and prediction of the agriculture information level with the support vector regression method. This paper, based on objective view and nonlinear method, is feasible, effective, has the value of application and popularization proven by empirical research.

Keywords: Agriculture information level, Evaluation and prediction, Index system, Support vector regression.

1. INTRODUCTION

The concept of agricultural information can be discussed from various aspects, with the development of information technology, information evaluation index are also renovated timely and leads to different methods of information level measurement, international agricultural information evaluation methods can be roughly divided into the following four types: (1). Information index method. It is a kind of information level evaluation system through the calculation of relative value, containing 4 categories such as information, rate of equipment, main body levels, information coefficient, but many indexes of this method are too old to keep up with the steps of new information technology. (2). Porat's method. American economist Marc U.Porat systematically put forward a method of information measurement in 1977. From the economy viewing point, the theory evaluates the information level through contribution rate of the information industry to the national economy. (3). Information society index method. It does the weighted calculation of three types of index such as social infrastructure, information infrastructure and computer infrastructure, the calculation is more convenient, but the method is lack of the indexes reflecting the of information flow and is short of the description of the social economic structure. (4). The IUP model method. It was launched by educational, scientific and cultural organization of United Nations as the largest and most comprehensive information evaluation model till now, it contains 230 indexes totally, but it has so many parameters and all they cross each other, so requirements for statistical data is very high, cannot be extensively used [1-8].

Compared with developed countries, Chinese agriculture information work starts rather late, the research on agricultural information level evaluation is poorer, so leads to not only the randomness of agriculture information construction in some areas and departments, but also the lack of master and scientific plan. Therefore, through the research on personality problems existing in specific region and the common problems of agriculture information in China, it will be very important and necessary to establish a set of scientific construction evaluation index system of agriculture information to study and evaluate the construction results from qualitative to quantitative, also good for decision making. The typical achievements of information evaluating methods got by current Chinese scholars is as follows: 1. Informatization with comprehensive index method brought by Y.X. Zhong (CIIC). The method is a new one considering the integrity and unity of factors to measure the information level based on Machlup's method, Porat's method and information index method. 2. Improved information index model brought by Liang Haili et. The index system of improved model is more reasonable but still has some problems, for example it does not give enough attention to such factors as information technology and network etc. 3. Method as 5-level index which was written by Wang Linong, it can evaluate the agriculture information situation in general but some indexes are crossed, difficult to quantify, data collection is difficult [9-12].

In summary, now there are the following insufficient aspects mainly exist in the domestic and foreign research on

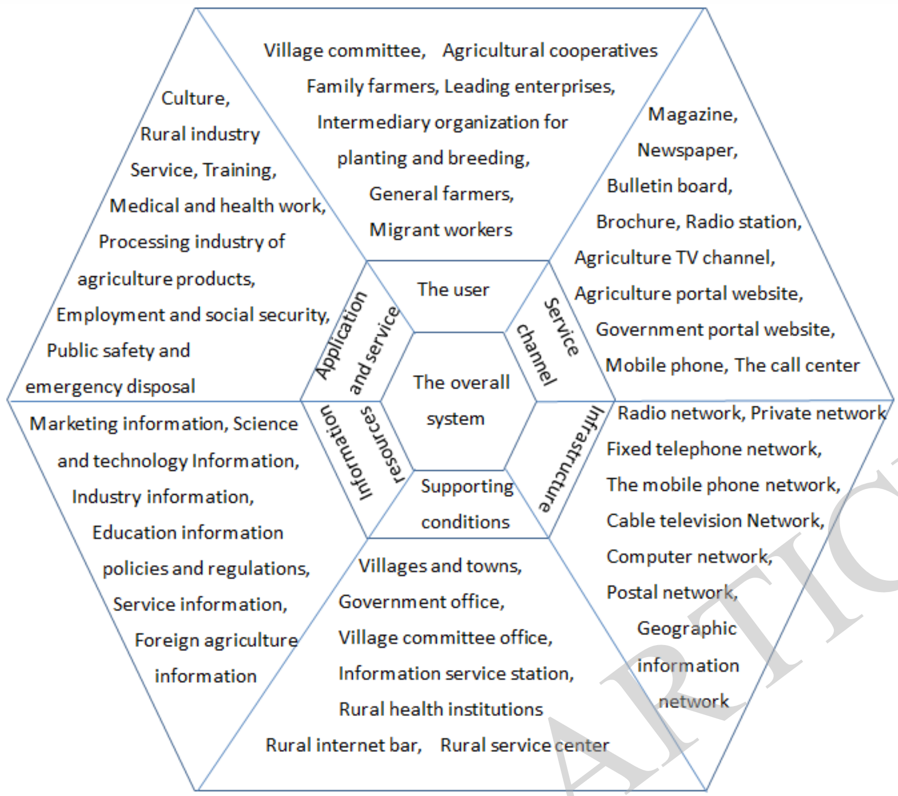

Fig. (1). The overall system of agriculture information.

evaluation of agricultural information level: (1) There is not an evaluation index system on agricultural information that is generally accepted. (2) Expert scoring method and AHP method are usually used to determine the weights of the indexes but all these methods are based on subjective judgment, it is difficult to guarantee the objectivity and operability. (3) Commonly used evaluation methods is Porat's method, principal component analysis (PCA) and comprehensive index method (CIM). The disadvantage of Porat's method is from the economic point and it cannot exactly reflect the measure of agricultural information; PCA and CIM are the statistical analysis methods that work on multiple indexes generally through linear transformation based on the agricultural information construction of a rational, index system. However, the agricultural information is an arduous and complex mission, as the technology of agricultural modernization changes rapidly by each passing day, agricultural information index also shows the characteristics of diversity, complexity and development, linear transformation evaluation method cannot fully reflect the nonlinear fitting characteristics between the agricultural information index system and evaluation results [13-15].

To solve above problems, first, this paper formulated the entire architecture of agriculture information as a set of structured, multi-level, scalable index evaluation system combined with the actual situation of the rural economic and social development; second, made a comprehensive analysis of current situation on agricultural information of the different counties and cities, determining the weight of each index by using the entropy method, this method can avoid the shortcomings of subjective as other existing methods ; third, used the agricultural information data in different counties as training set, established nonlinear regression model of agricultural information by the support vector machine method, then evaluated the agriculture information level and predicted the agricultural information level; finally, analyzed status of agriculture information level combined with the research data and results of empirical studies and gave some suggestions on it [16, 17].

2. IDEA OF CONSTRUCT THE EVALUATION AND PREDICTION SYSTEM ON AGRICULTURE INFORMATION LEVEL

The new agriculture information construction is a comprehensive system, so we should analyze, design and arrange all the new agriculture information elements systematically in order to effectively and scientifically reflect the inherent logic relation and its action mechanism, understanding the basic law of the new agriculture information development correctly, extracting the real observing point which can display the new agriculture information development level, then lay the foundation for the construction of evaluation index system. This paper shows the overall system of new agriculture information with Fig. (**1**) including basic information infrastructure, information resources, information applications and services, application supporting conditions, service transmission channel and the user [18].

Table 1. Index System of Agriculture Information Level.

Junior Index	Senior Index	Index Code
1. Agricultural information infrastructure index	1.1. Broadband in-home ratio	X_1
	1.2. Home popularity rate of fixed telephone	X_2
	1.3. Popularity rate of mobile phone	X_3
	1.4. Popularity of computer	X_4
	1.5. The proportion of administrative village with cable TV connected	X_5
	1.6. The proportion of village and township with broadband connected.	X_6
2. Human resources index of agriculture information	2.1. Rural technical personnel number per ten thousand people	X7
	2.2. Rural practitioners accounted for the proportion of rural population	X_8
	2.3. The average number of village-level informants of each county	X_9
3. Index of rural service channel and information spread	3.1. The county agriculture website	X_{10}
	3.2. The agriculture website hits per month	X_{11}
	3.3. The agriculture website article content updates per month	X_{12}
4. Index of rural application supporting conditions	4.1. The average number of township agricultural information service institutions in county-level	X_{13}
	4.2. The average village-level agricultural information service-shop in county-level	X_{14}
5. Index of Main environment of agriculture information	5.1. Per capita annual income of peasants	X_{15}
	5.2. The annual trade volume of agricultural products	X_{16}
	5.3. Special agricultural information funds	X_{17}

3. CONSTRUCTION OF THE AGRICULTURE INFORMATION INDEX SYSTEM

We carry on the reasonable design and arrangement of the constituent elements of the agricultural information systematically, and formulate the overall system of the agricultural information as a structured, hierarchical, open, extensible evaluation index system, it can be improved and optimized according to the change of the agriculture development status, change trend, statistics channel, data acquisition channel, as shown in Table **1**.

4. AGRICULTURAL INFORMATION INDEX WEIGHT CALCULATION

4.1. The Characteristics of Entropy

In 1948, Mr. Shannon proposed to introduce the concept of entropy in physics to information theory, used as a measure of the uncertainty information system structure. In information theory, entropy is a concept used to measure the amount of information, the more orderly a system is, the less the entropy is, conversely, the more disordered a system is, the higher the entropy is. Therefore, according to the characteristics of entropy, the weight of index can be judged by calculating the entropy. When the entropy value is bigger, the disorder degree is smaller, individual decision attribute value in data set is dominated, the greater influence will be put on comprehensive evaluation by this index, on the contrary, the smaller the entropy value, the smaller the influence of the index of comprehensive evaluation.

4.2. Steps to Determine Weight by Entropy Method

Depends on the characteristics of entropy, the steps to determine weight by entropy method as follows:

Standardized matrix: do the standardized process to the agricultural information raw data matrix $X=(x_{ij})_{m\times n}$, we can get standardization matrix as $P=(p_{ij})_{m\times n}$, among them: $p_{ij}=\frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}}(1\le i\le m,1\le j\le n)$.

Weights determined by the weight of entropy method: to get the entropy of the j index: $e_j=-k\sum_{i=1}^{m} p_{ij}\ln p_{ij}(1\le j\le n)$, the weight of the j index is: $\omega_j = \frac{g_j}{\sum_{j=1}^{n} g_j}$ $(1 \le j \le n)$, among them, $g_j = 1 - e_j$ $(1 \le j \le n)$. The bigger the index weight is, the greater the amount of agricultural information is.

5. EVALUATION AND PREDICTION OF AGRICULTURAL INFORMATION LEVEL

5.1. Support Vector Regression

Support vector machine (Support Vector Machine) was first proposed by Mr. Corinna Cortes and Mr. Vapnik in 1995 and has been widely used since that. The method shows many unique advantages in solving the small sample, nonlinear and high dimensional pattern recognition problem. If the support vector machine is applied to the function fitting problem, it will be the support vector regression (Support Vector Regression). The basic idea of support vector regression will be described as follows: set the training sample $S = \{(x_1, y_1), ..., (x_l, y_l) \mid x_i \in R^n,\ y_i \in R\}$, at any given precision condition ε, constructing the hyperplane $f(x) = < w, x > + b$ to conform to $|y_i - f(x_i)| \le \varepsilon$, $\forall (x_i, y_i) \in S$, so $f(x)$ will be the support vector regression model. Support vector regression is for resolving convex quadratic optimization problem, it can guarantee that extremal solution is global optimal solution.

Considering the allowable, introducing slack variables ξ_i, ξ_i^*, support vector regression problem will be optimized as:

$$\min \tfrac{1}{2} \| w \|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*)$$

$$s.t. \begin{cases} y_i - < w, x_i > - b \le \varepsilon + \xi_i \\ < w, x_i > + b - y_i \le \varepsilon + \xi_i^*,\ i = 1,2,...,l \\ \xi_i, \xi_i^* \ge 0 \end{cases}$$

Among them, the constant $C > 0$ shows the degree of punishment beyond the error ε.

For nonlinear problem, we need to find a kernel function $K(x_i, x_j) = < \phi(x_i), \phi(x_j) >$, instead of the calculation of inner product in the original feature space, the low-dimensional sample set will be mapped to high-dimensional feature space through nonlinear mapping to achieve the nonlinear regression in original space by linear regression in high dimensional space. The dual form of optimization problem will be obtained by construction of Lagrange equation,:

$$\min \frac{1}{2} \sum_{i,j=1}^{l} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(x_i, x_j) - \sum_{i=1}^{l} y_i (\alpha_i^* - \alpha_i) + \sum_{i=1}^{l} \varepsilon (\alpha_i^* + \alpha_i)$$

$$s.t. \quad \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) = 0, \quad 0 \le \alpha_i, \alpha_i^* \le C, \quad i = 1,2,...,l$$

Evaluate α_i, α_i^* and calculate the Constant deviation b by using KKT (Karush-Kuhn-Tucker) condition.

$$b = \begin{cases} y_j + \varepsilon - \sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*) K(x_j, x_i), \alpha_i \in (0, C) \\ y_j - \varepsilon - \sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*) K(x_j, x_i), \alpha_i^* \in (0, C) \end{cases}$$

Finally, construct the nonlinear function as:

$$f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) K(x_j, x) + b, \quad x_i \in R^n,\ b \in R \cdot$$

5.2. Evaluation and Prediction on Agriculture Information Level

(1) SVM model: set as ε-SVR

(2) Kernel function: RBF (Radial Basis Function) is selected in this paper.

$$K(x_i, x_j) = \exp(-\lambda \|x_i - x_j\|^2,$$

(3) Construction of SVR agriculture information level evaluation model.

The construction of agricultural information evaluation model is mainly the determination of two parameters: the fault-tolerant penalty parameter C and kernel function parameter λ. SVR regression model parameters selection can be determined by the grid search, respectively to set the maximum value and the minimum value of these two parameters, then divides their domain value space into a grid by a certain step, determines the optimal parameter at each grid point through testing one by one.

(4) Evaluation and prediction on agriculture information level with the model above.

6. EMPIRICAL RESEARCH

6.1. Agriculture Information Data of Different Counties in Hebei Province

The empirical research data depends on the data of Hebei Province Economic Statistical Yearbook 2008~2013 and the research data that counties reported to the Hebei provincial agriculture department. For example, the counties of 2013 Hebei Province agricultural information data shown in the Table **2**:

6.2. The Weight of Agricultural Information Index in Hebei Province

The calculation for the standard matrix of agricultural information data of counties and cities in Hebei Province to

Table 2. Agriculture Informational Data of Different Cities and Counties in Hebei Province, 2013.

Counties	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}
Bao Ding(BD)	0.08	0.50	0.68	0.39	0.61	0.94	9.7	1.77	98	14	625	201	10	5	46775	860000	125
Cang Zhou(CZ)	0.06	0.65	0.69	0.24	0.77	0.95	19.5	0.57	297	47	396	21	10	97	67708	1400556	164
Zhang Jia Kou(ZJK)	0.17	0.37	0.59	0.25	0.58	1.00	12.9	0.58	102	13	5918	30	12	7	85472	392871	59
Han Dan(HD)	0.27	0.39	0.57	0.41	0.86	1.00	8.3	1.05	96	21	546	42	11	5	41616	1023703	22
Chen De(CD)	0.11	0.71	0.61	0.33	0.86	0.95	13.9	0.60	245	6	349	56	22	79	21767	263352	95
Qin Huang Dao(QHD)	0.45	0.72	0.70	0.37	0.87	0.94	8.5	0.56	122	42	733	77	17	10	27671	403478	27
Tan Shan(TS)	0.44	0.50	0.73	0.47	0.94	1.00	12.7	0.64	410	108	18629	203	16	152	478101	2273183	535
Shi Jia Zhuang(SJZ)	0.31	0.46	0.79	0.50	0.86	0.95	37.4	0.53	362	432	68808	345	11	79	100369	1660699	292
Xing Tai(XT)	0.09	0.43	0.67	0.38	0.51	1.00	15.0	0.04	207	40	3150	25	6	35	59133	70800	19
Lang Fang(LF)	0.17	0.72	0.83	0.50	0.60	1.00	2.2	0.67	177	7	3216	139	10	22	64595	287383	38
Heng Shui(HS)	0.22	0.55	0.69	0.38	0.75	0.97	14.01	0.70	211.60	73.00	10237.00	113.90	12.50	49.10	99320.70	863602.50	137.60

Table 3. Index Weight of Agriculture Information in Hebei Province.

Junior Index	**Infrastructure Index of Agriculture Information**						**Human Resources Index of Agriculture Information**			**Index of Rural Service Channel and Information Spread**			**Index of Rural Application Supporting Conditions**		**Index of Main Environment of Agriculture Information**		
Sinor index	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}
Senior index weight	0.045	0.019	0.015	0.018	0.017	0.015	0.041	0.041	0.034	0.144	0.194	0.068	0.023	0.084	0.091	0.061	0.090
Junior index weight	0.129						0.116			0.406			0.107		0.242		

determine the weight of each index shown as in Table **3** by entropy value method.

6.3 Evaluation and Prediction on Agriculture Information Level in Heibei Province

Set the ag ricultural information data 2008 ~ 2012 of counties in Hebei province as the training data, make a data standardization. Get the SVR agricultural information level model with the training set. Based on application software package LIBSVM developed and designed by Doctor Lin zhiren who is from the national Taiwan University, all the experiment of this article uses cross validation parameter optimization method between 70% training set and 30% test set. Parameter ε is determined as 0.01, through the grid search, finally we get C = 9.1896, λ = 0.011842, shown as in Fig. (**2**). Evaluate the agricultural information data 2013 of counties in Hebei province by the model. Fig. (**3**) is the curve of BP (Back Propagation) model and SVR (Supporting Vector Regression) model fitting respectively to the result of WAM(Weighted Aggregation Method).

Set the data of agricultural information 2008~2013 of counties in Hebei province as training data, constructs the regression function model of them separately, forecast the agricultural information level data during 2014~2018 and shown as in Fig. (**4**).

6.4. Analysis

(1) Comprehensive analysis of various agricultural information present situation, we can draw the following conclusion: the evaluation of agricultural information level in Hebei province can be roughly divided into three levels. First level: SJZ, TS, QHD, CZ, their information levels are higher; secondary level: LF, BD, CD, HS, the information levels are medium; third level: XT, HD, ZJK, the information levels are low.

(2) Relative application infrastructure, human resources and support conditions index, rural service channels and information transmission index has more and more influence to the agricultural information level. We can see from the analyzed data in Hebei province the rural service channels and information transmission index weight value is 0.406, it

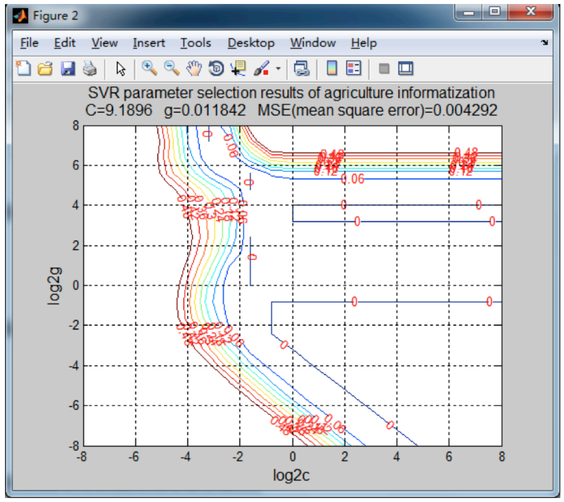

Fig. (2). Parameters Selection Results.

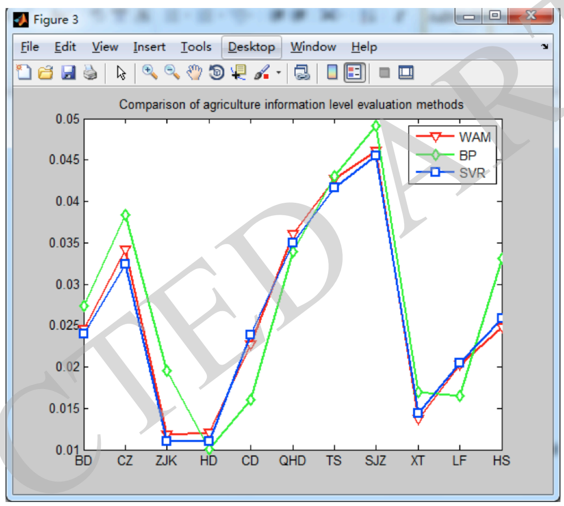

Fig. (3). Comparison of Evaluation Methods of Agriculture Information Level.

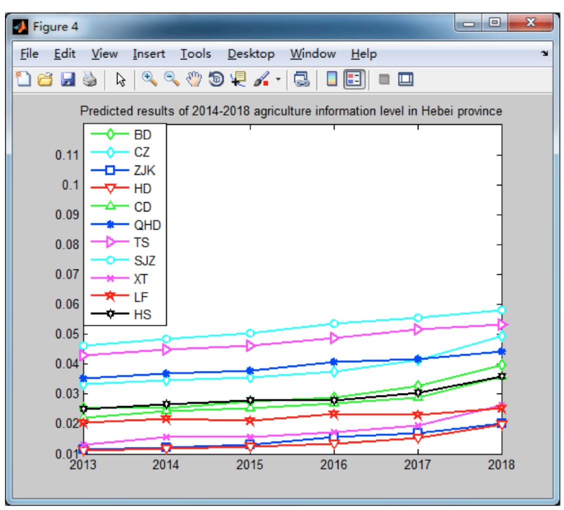

Fig. (4). Predicted Results of 2014-2018 Agriculture Information Level in Hebei Province.

is larger than the infrastructure, human resources and application support index (0.129, 0.116, 0.107), weight value, it shows a new trend in the development of information technology of new rural construction in Hebei province. In recent years, the popularity rate of fixed telephone, mobile phone, computer, connected broadband in the rural areas was stable; coverage rate of township agricultural information service agencies, village-level agricultural information service shops tended to be stable after years of construction in the rural areas ; Agricultural websites, Internet information platform of agricultural products and other new type of rural service channels and information communicating way has played a more and more important role in the process of agricultural information.

CONCLUSION

This paper first constructs a set of agricultural information index system, then gets the weight of each index according to the method of entropy and does evaluation and prediction of the agricultural information level with the support vector regression method based on all these previous work. Based on objective view and nonlinear method, this paper is feasible, effective, has the value of application and popularization proven by empirical research.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

This work is supported by the the soft science research plan project of Hebei Province, China (No. 14457419D).

REFERENCES

[1] M.U. Porat, “The information economy: Definition and measurement,” *Office of Telecommunications Special Publication*, vol. 58, pp. 77-12, 1977.

[2] H.L. Liang, “Chinese informatization level index measure research,” *Intelligencei Information Work*, vol. 4, pp. 4-8, 1999.

[3] L.N. Wang, “Research of Chinese rural informatization measurement and development,” *Central University of Finance and Economics*, 2009.

[4] T.W. Shultz, “The value of ability to deal with disequilibria,” *Journal of Economic Literature*, vol. 13, pp. 827-996, 1975.

[5] S.J. Yao, “Determinants of grain production and technical efficiency in china,” *Journal of Agriculture Economics*, vol. 49, pp. 171-184, 1998.

[6] K.P.A. Kalirajan, “Decomposition of total factor productivity growth: The case of China’s agriculture growth before and after reform American,” *Journal of Agriculture Economics*, vol. 78, pp. 331-338, 1996.

[7] S.G. Fan, “Effects of technological change and institution reform on production growth in Chinese agriculture,” *American Journal of Agriculture Economics*, vol. 72, pp. 266-275, 1991.

[8] F.Q. Mei, “Chinese low cost rural informatization development pattern choice,” *China Information Industry*, vol. 5, pp. 11-16, 2007.

[9] Y.X. Zhong, “comprehensive information based methodology for natural language understanding,” *Journal of Beijing University of Posts and Telecommunications*, vol. 27, pp. 1-12, 2004.

[10] W. Chen, and J.H. Xia “ An optimal weights combination method considering both subjective and objective weight information,” *Journal of Mathematics in Practice and Theory*, vol. 1, p. 17-22, 2007.

[11] L. Zhao, Y.M. Zhu, M.C. Fu, P. Zhang, and Y. Cao, “Comparative study on intensive use of rural residential land based on principal component analysis and entropy method,” *Transactions of the Chinese Society of Agricultural Engineering*, vol. 7, pp. 235-242, 2012.

[12] S.F. Ding, B.J. Qi, and H.Y. Tan, “An Overview on Theory and Algorithm of Support Vector Machines,” *Journal of University of Electronic Science and Technology of China*, vol. 1, pp. 2-10, 2011.

[13] Y.S. Zhu, and G.H. Yan, “Parameters selection method for support vector machine regression,” *Computer Engineering*, vol. 7, pp. 218-220, 2009.

[14] W.P. Bai, “Study on the Effect Evaluation Index System of New Rural Informationization Construction,” *Journal of Anhui Agriculture and Science*, vol. 31, pp. 13910-13912, 2008.

[15] G.M. Du, Y. Zhou, G. Liu, *et al.*, “The evaluation of Heilongjiang agricultural modernization,” *Chinese Journal of Agricultural Resources and Regional Planning*, vol. 5, pp. 55-64, 2013.

[16] Hebei people’s government. “Hebei people government Hebei economy yearbook,” *China Statistics Press*, 2013.

[17] Z.L. Xie, and G.C. Lin, “Research on degree of informatization index system in the new rural construction,” *China Collective Economy*, vol. 5, pp. 144-145, 2008.

[18] F.S. Song, and B.H. Zhao, “Agricultural information development in Hebei province,” *Chinese Agricultural Science and Technology Press*, 2009.

Received: July 02, 2015 Revised: August 10, 2015 Accepted: September 21, 2015