APPLICATION OF PROJECTION PURSUIT MODEL IN SOIL EVALUATION OF CONSERVATION TILLAGE

Junjing Yuan, Hongwen Li [*]

College of Engineering, China Agricultural University, Beijing, China, 100083;

[*] *Corresponding author, Address: P. O. Box 46, College of Engineering, China Agricultural University, 17 Tsinghua East Road, Beijing, 100083, P. R. China, Tel:+86-10-62737631, Fax:+86-10-62737300, Email: lhwen@cau.edu.cn*

Abstract: This paper established a conservation tillage evaluation model based on projection pursuit to evaluate the soil composite achievement of conservation tillage. After optimizing the project direction, the multi-dimension data of the seven evaluation indices are synthesized to one dimension, and the author could evaluate each item with the projection data easily, which avoided the jamming of weight matrix. The results of the evaluation mode in Linfen, Shouyang and Linghai are accordant with the production, which indicated that the model was available and provided a new method or thought to evaluate the composite achievement of conservation tillage.

Key words: conservation tillage, projection pursuit, evaluation index

1. INTRODUCTION

In 1974, Friedman and Tukey put forward the term projection pursuit and proposed the first algorithm, and the projection pursuit method is well known since its first formulation. They considered “projection pursuit will automatically tend to avoid projections involving those measurement variables that do not contribute to the data structure” (Friedman et al., 1974). Many researchers (Huber, 1990; Eslava et al., 1999; Simone et al., 2002; Choulakian, 2006; Daniela, 2007) have studied the projection pursuit models and introduced them to overcome the problem of an excessive number of parameters in multivariate analysis. The application of PPC model in irrigation schemes obtained better results and provided a new thought for the evaluation of irrigation schemes (Jiang et al., 2006). Projection pursuit learning networks (PPLNs) were used in image processing and were successful at developing an inverse blur filter to enhance blurry images (Lesa et al., 2000). Compared with the time-consuming and expensive traditional approaches, it is quicker, easier and lower-cost to apply the projection pursuit regression (PPR) method in the study of shower cooling tower (Qi et al., 2008).

Conservation Tillage began in the 60's and the systematic study began in the 90's in the 20th Century in China. The contents of study consist of yield of crops (Peng et al., 2006; Zhao et al., 2006; Li et al., 2000), production cost (Wang et al.,2007; Li et al., 2000), soil properties and texture (Kay, 1990), water utilization (Liu et al., 2000; Zhang et al., 2002), greenhouse gas emission (Liu et al., 2006; Li et al., 2003; Lar et al., 2004) and etc.. And the methods of study consist of pairing comparison of captive test (Peng et al., 2006; Wang et al.,2007; Li et al., 2000), cost-benefit method (Zhao et al., 2006; Li et al., 2000), Delphi (Wang et al.2004), AHP (Wang et al., 2004; Tian et al., 2006) and so on. Pairing comparison of captive test is suitable for one single index or fewer indices when comparing the achievement between Conservation Tillage and traditional tillage. Generally, the cost-benefit method utilized fixed price to calculate the cost, benefit and static pay back period. Although the method is easy and simple, it is difficult to compare the achievement among different areas and different farming models. The methods of Delphi and AHP took into account the experts' experiences and intuition sufficiently, on the one hand, they solve the quantification of qualitative indices, on the other hand, they make use of the original information insufficiently and much jamming is unavoidable when put up weights (Wu et al. 2004).

It is from these considerations that this paper tries to apply projection pursuit to evaluate the soil achievement of Conservation Tillage in different areas.

2. PROJECTION PURSUIT MODEL

Projection pursuit is a recent statistical method and a new technique for nonparametric multivariate regression which can overcome the curse of dimensionality. The projection pursuit method is not affected by the presence of noisy variables and it is aimed at solving the difficult problem of identifying structure in high dimensional data. It does this by projecting the high-dimensional data onto low-dimensional subspace and clustered the observations into homogeneous groups (Angela, 2001). Compared with the methods of Delphi, AHP and Eigenvector, the projection pursuit method is an external evaluation method based on the characteristics and structure of original data (Yang et al., 2005).

The projection pursuit method does not establish direct relationship between the independent variables and the dependent variables, but conduct projections on high- dimensional independent variables; after a low dimensional projection variable is obtained, the relationship between the projection variable and independent variable can be determined (Qi, et al., 2008).

The steps of projection pursuit model for soil evaluation of conservation tillage are as follows (Zhang, 2007; Jin et al., 2005).

Step 1. Establishment of evaluation function. Given the matrix $X^*_{n\times p}$ present the index values of samples, where p is the number of observed variables, n is the number of units and x^*_{ij} denote the value of the index j of unit i.

Step 2. Method of dimensionless parameter. In order to make the dimension of all values coincident, dimensionless calculation for original values of evaluation indices is necessary. All the evaluation indices are divided into three kinds:

Evaluation indices are normalized to be maximizing ones, and after normalization, matrix $X^*_{n\times p}$ could be replaced by matrix $X_{n\times p}$.

$$x_{ij} = \frac{x^*_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \tag{1}$$

Evaluation indices are normalized to be minimizing ones,

$$x_{ij} = \frac{\max(x_j) - x^*_{ij}}{\max(x_j) - \min(x_j)} \tag{2}$$

Evaluation indices are normalized to be minimizing ones,

$$x_{ij} = \frac{\left|x^*_{ij} - mid(x_j)\right|}{mid(x_j)} \tag{3}$$

In the above-mentioned three formulas, $\max(x_j)$ means the maximum of the index j, $\min(x_j)$ means the minimum of the index j, and $mid(x_j)$ means the intermediatevalue of the index j.

Step 3. Analysis of linear projection. The aim of projection is to decrease the dimension and avoid the 'curse of dimensionality'. In this paper, the high-dimension values are projected to one-dimension. Assume unit vector a is the optimum projection direction and z_i is the projection eigenvalue of matrix X, the formulas are as follows.

$$a = \left\{a_1\ , a_2, \cdots a_p\right\}, (j = 1,2,3\ldots, p) \tag{4}$$

$$z_i = \sum_{j=1}^{p} a_j \cdot x_{ij}, (i = 1,2,3\ldots, n; j = 1,2,3\ldots, p) \tag{5}$$

Step 4. Establishment of objective function. According to the classification principles, the distribution of projection values should possess the following characteristics: (a) the dots of projection in the graphics should be as dense as possible in the local, namely it is better that most dots mass together; (b) the points mass in the graphics should be as sparse as possible in the whole. Therefore, it is better that the distances between classes and the densities within classes are the maximum at the same time. And the objective function could be denoted as product of the distances between classes and the density within classes.

$$\begin{aligned} &Q(a) = S_Z \cdot D_Z \\ &S_z = \sqrt{\frac{\sum_{i=1}^{n}\left(z_i - E(z)\right)^2}{n-1}} \\ &D_z = \sum_{i=1}^{n}\sum_{k=1}^{n}\left(R - r_{ik}\right)) \cdot f\left(R - r_{ik}\right) \end{aligned} \tag{6}$$

In the above functions, S_z means the Standard Deviation of z_i, and D_z means the partial density of z_i.

Step 5. Optimization of objective function. For the given indices values of the samples, the value of $Q(a)$ may change when changes take place in the projection direction a, and the optimal projection direction reflected certain structures of high-dimension furthest. Thus, the representation of optimal function is as follows.

$$\begin{aligned} &MaxQ(a) = S_z \cdot D_z \\ &sub.to: \sum_{j=1}^{p} a^2(j) = 1 \end{aligned} \tag{7}$$

Due to $a = (a_1, a_2, \cdots a_p)$ is as a variable for the complex nonlinear optimization function, it is difficult for the traditional optimization methods to deal with the problem, and real coding based accelerating genetic algorithm (RAGA) is effectual.

Step 6. Classification and evaluation. Suppose $a*$ is the optimal projection direction and $z*(i)$ is the eigenvalue of the optimal projection direction, we can obtain $a*$ according to $MaxQ(a) = S_z \cdot D_z$, and also we can obtain $z*(i)$ according to $a*$ and function (5). Comparing $z_1^*, z_2^*, \ldots, z_n^*$, the maximum means its corresponding sample is the best and the minimum means its corresponding sample is the worst and we also could classify the samples in terms of the approximate degree of $z_1^*, z_2^*, \ldots, z_n^*$. Comparing $a_1^*, a_2^*, \ldots, a_p^*$, the result magnitude means its contribution degree or sensitivity degree of the corresponding index to the system.

3. APPLICATION

Conservation Tillage, a kind of technology of dry farming, is different from traditional agriculture, and the main contents include No Tillage, Reduced Tillage, Straw Mulching, and so on. Table 1 is the standard grade of soil confirmed by the experts, and there are seven evaluation indices and six samples (grades).

Table 1. soil standard grade

Grade	I	II	III	IV	V	VI
Ph	6.5~7.5	6.0~6.5 7.5~7.8	5.5~6.0 7.8~8	5.0~5.5 8~8.5	4.5~5 8.5~9	≤4.5 >9
Bulk density g/c m3	1.25~1.35	1.2~1.25 1.35~1.40	1.15~1.20 1.40~1.45	1.10~1.15 1.45~1.50	1.00~1.10 1.50~1.60	>1.6 <1
Water storage %	14~16	16~18 12~14	18~20 10~12	20~22 8~10	22~24 6~8	<6 >24
Organic matter %	>4	3~4	2~3	1~2	0.6~1	<0.6
Total N %	>0.2000	0.15~0.20	0.10~0.15	0.075~0.10	0.05~0.075	<0.050
Available P mg/kg	>40	20~40	10~20	5~10	3~5	<3
Available K mg/kg	>200	150~200	100~150	50~100	30~50	<30

Adopt RAGA method to optimize the projection direction and adopt MATLAB7.0 to program, suppose the parameters are as follows, initial population size n=400, crossover probability p_c=0.8, mutation probability p_m=0.03, number of superior individualsis 20, α=0.05, times of acceleration are 11, and after running software we can obtain the optimum projection direction and projection eigenvalue, namely,

$$a_s^* = (0.1897, 0.3416, 0.4419, 0.4723, 0.3997, 0.4987, 0.4235)$$

$$z_s^* = (2.6138, 1.7216, 1.2366, 0.9895, 0.6585, 0.5165)$$

Fig. 1 is about the relationship between projection values and evaluation grades, consider y*represent the grade and the evaluation model of soil achievement based on projection pursuit is as follows.

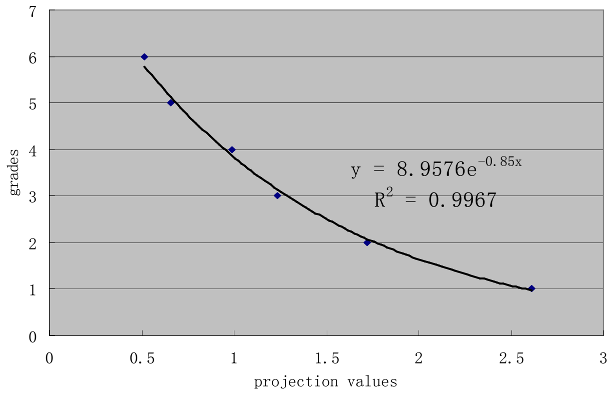

Fig. 1 scatter diagram of grade and projection value

The error analysis in table 2 indicate that the precision of evaluation model possesses a much higher precision and reliability

Table 2. Error Analysis

Empirical value	Projection value	calculated value	absolute error	Relative error
1	2.6138	0.9712.	-0.0288	-2.8805
2	1.7216	2.0733	0.0733	3.6651
3	1.2366	3.1311	0.1311	4.3708
4	0.9895	3.8629	-0.1371	-3.4266
5	0.6585	5.1181	0.1181	2.3614
6	0.5165	5.7746	-0.2254	-3.7559
Average			0.1190	3.41

Test data in Table 3 root in the conservation tillage research center of Agricultural Ministry.

Table 3. Test Data

	pH	bulk density g/c m3	water storage %	organic matter %	Total N /%	available P mg/kg	available K mg/kg
A1	8.08	1.30	13.51	1.45	0.0816	48.91	268.63
A2	8.02	1.41	14.43	1.72	0.0877	36.77	362.16
A3	7.98	1.39	14.49	1.91	0.1287	39.42	350.46
B1	7.97	1.26	12.51	1.46	0.0702	21.93	190.69
B2	8.03	1.45	14.21	2.28	0.0736	38.63	405.02
B3	8.04	1.37	15.04	2.60	0.1001	46.04	412.82
C1	7.02	1. 27	19.9	2.47	0.0765	21.05	152.69
C2	7.01	1.31	21.9	2.50	0.0772	20.46	234.53
C3	7.02	1.32	23.4	2.51	0.0782	19.95	227.71

There are seven indices and nine samples in this system, and tests of class A, B and C are processing respectively in Linfen, Shouyang and Linghai. And the detail significations of the samples are in table 4.

Table 4. Significations of the sample

Sample	Signification
A1, B1, C1	Traditional tillage
A2,B2	Shallow tillage and straw cover
A3, B3	No tillage and straw cover
C2	No tillage and standing stubble
C3	No tillage and break straw cover

With the same method to deal with the test data, the projection eigenvalue and the soil grade with different farming models are as table 5.

Table 5. Soil grade of the sample

Sample	Projection Values	Grades
A1	0.8592	4.3154
A2	1.2408	3.1200
A3	1.4512	2.6090
B1	0.8369	4.3979
B2	1.2378	3.1279
B3	1.4758	2.5551
C1	1.1701	3.3132
C2	1.2311	3.1458
C3	1.2384	3.1263

It is obviously that these nine samples are classified three clusters, A1 and B1 belong to the grade Ⅳ; A2, B2, C1, C2 and C3 belong to t the grade Ⅲ; A3 and B3 belong to the grade Ⅱ.

4. RESULTS AND DISCUSSION

The principal contributions of this article are the following:

(1) The model based on projection pursuit avoided the jamming of weight matrix, and take full advantage of the inter-structure of original data.

(2) In Linghai, conservation tillage has been applied 2 years, and the soil grades of C2, C3 are promoted a little compared to C1. It is showing that the effect of short-term conservation tillage is not significant in improving soil properties.

(3) In Linfen and Shouyang, conservation tillage has been applied 15 years, and the grade of soil promoted from 4 to 3 and 2. Therefore, the effect of long-term conservation tillage is significant, which is accordance with productions.

REFERENCES

Angela Montanari, Laura Lizzani. A projection pursuit approach to variable selection. Computational Statistics & Data Analysis 35 (2001) 463~473

Daniela G. Calò. Gaussian mixture model classification: A projection pursuit approach. Computational Statistics & Data Analysis 52 (2007) 471 – 482

Friedman, J.H., Tukey, J.W., 1974. A projection pursuit algorithm for exploratory data analysis. IEEE Trans. Comput C 23, 881~890.

G. Eslava, F.H.C. Marriott. Projection indices and multimodality for mixtures of normal distributions. Statistics & Probability Letters 43 (1999) 289~297

Huber, P.J., 1990. Data analysis and projection pursuit. Technical Report PJH-90-1. MIT, Dept. of Mathematics.

Jiang Guoyong, Cao Hualiang. Application of PPC model for comprehensive evaluation of irrigation schemes. Research of soil and water conservation, 2006, vol.13, No.5: 87~89

Jin Juliang, Hong Tianqiu, LI Ruzhong. Objective weighted model for evaluating dynamic eco-environmental comprehensive quality of Chaohu lake basin. Journal of Hydroelectric Engineering, 2005,24(5):99~103

Kay, .B. D. Rates of change of soil structure under different cropping systems. Advance in Soil Science. 1990,12:1～2

Lal R, Griffin M, Apt J. Managing soil carbon [J]. Science, 2004,304(4):393

Lesa M. Kennedy, Mitra Basu. Application of projection pursuit learning to boundary detection and deblurring in images. Pattern Recognition 33 (2000) 2019~2031

Li Changsheng, Xiao Xiangming, Frolking S,Moore B,etc. GREENHOUSE GAS EMISSIONS FROM CROPLANDS OF CHINA. Quaternary Sciences, 2003, 23(5):493-503.

Li Hong-wen, Gao Huan-wen, Zhou Xing-xiang, Mao Ning. Economic analysis of maize production in dryland with conservation tillage. Agricultural research in the arid areas, 2000,9,18(3):44~4

Liu Jianmin1, Hu Lifeng, Zhang Aijun. Research Progress in Greenhouse Effect Caused by Conservation Tillage. Chinese Agricultural Science Bulletin, 2006,8,22(8):246～249

Liu Li-jing, Gao Huan-wen, Li Hong-wen, Jin Shuang-yi. Experimental study on influencing factors of evaporation under conservation tillage. Journal of China Agricultural University 2006,11(4): 78～82

Peng Wen-ying- , Zhang Ya-bin. Review of impacts of no-tillage on crop yield and economic benefit. Agricultural Research in the Arid Areas, 2006,7,24(4):113～118

Simone Borra, Agostino Di Ciaccio. Improving nonparametric regression methods by bagging and boosting. Computational Statistics & Data Analysis 38 (2002) 407 – 420

Tian Shu-min, Liu Rui-han, HOU Fu-qiang. Study on Evaluation Indicator System of Conservation Tillage Technology Implementary Benefit in Beijing Rural Areas. Journal of Beijing Agricultural College, 2006,10,21(4):50～53

V. Choulakian. L1-norm projection pursuit principal component Analysis. Computational Statistics & Data Analysis 50 (2006) 1441~1451

WANG Xu-mei, ZHANG Ying, AN Xiao-guang. Evaluating indexes system and weight about agricultural eco-system. Journal of Northeast Agricultural University , 2004,35(3):386～388

Wang Yu-zhen, Discussion of the economic benefit and ecological benefit of conservation tillage. Agricultural economy,2007,5:30

Wu Dianting, Li Dongfang. SHORTCOMINGS OF ANALYTICAL HIERARCHY PROCESS AND THE PATH TO IMPROVE THE METHOD. Journal of Beijing Normal University(Natural Science, 2004, 40(2):264～268

Xiaoni Qi , Zhenyan Liu, Dandan Li. Prediction of the performance of a shower cooling tower based on projection pursuit regression. Applied Thermal Engineering 28 (2008) 1031~1038

Yang Xiao-hua, YANG Zhi-feng, SHEN Zhen-yao, LI Jian-qiang. A Multi-Objective Decision-Making Ideal Interval Method for Integrated Assessment of Urban Environmental Quality. Systems Engineering-theory & Practice, 2005,(8):119~123

Zhang Bin. Application of PPC Model to Grasses of Sloping Field for Soil and Water Conservation. Research of Soil and Water Conservation, 2007,14(2):299~301

Zhang Hailin , Chen Fu , Qin Yaodong , Zhu Wenshan. Water Consumption Characteristics for Summer Corn Under Under No-Tillage With Mulch. Transactions of The Chinese Society of Agricultural Engineering, 2002,18(2):36-40

Zhao Hai-tao, Shi Shuai-bing. Benefit analysis on the technology of mechanized conservation tillage. Journal of Agricultural Mechanization Research, 2006, 8:74~76