Evaluation Model of Winter Wheat Yield Based on Soil Properties

Wei Yang, Minzan Li, Lihua Zheng, Hong Sun

HAL Id: hal-01420282
https://hal.inria.fr/hal-01420282

Submitted on 20 Dec 2016

Evaluation model of winter wheat yield based on soil properties

Wei Yang[a], Minzan Li[b], Lihua Zheng[c], Hong Sun[d]

Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing, China)
[a]cauyw@cau.edu.cn, [b]gpac@cau.edu.cn, [c] zhenglh@cau.edu.cn, [d] sunhong@cau.edu.cn

Abstract. In order to realize precision management of winter wheat, two prediction models of winter wheat yield based on soil parameters were proposed and compared. The field tests were carried out in two years. The variety of the experimental winter wheat was Jingdong 12, and the test area was divided into 60 zones with 5m×5m grids. The sampling point was put in the center of the zone, and the depth of the sampling point was 5cm. Soil EC was measured by a DDB-307 EC meter, and the winter wheat yield data were provided by a CASE2366 grain harvester with GPS receiver. Gray theory were used to analyze the gray relation between soil EC value and each of other soil parameters, such as total nitrogen content, K^+、NO_3^- and pH of soil. Results showed that there were high gray relation between soil EC and three other indexes, total nitrogen content, K^+, pH of soil. Since soil organic horizons had high correlation with soil negative charge capacity, when soil had more organic horizons, there were more soil negative ions, and the soil EC was higher. Hence, the gray relation between K^+ and EC was high. Using nitrogen fertilizer could removal caution from soil, and increase the content of K^+，Na^+，$Ca^{2+}$and Mg^{2+},so that there was also high correlation between total nitrogen content and EC. The reason of high correlation between EC and soil pH was attributed to that the change of pH had influence on negative charge. After analyzing the correlation between winter wheat yield and soil EC, total nitrogen content, K^+, NO_3^- , pH of soil in different growth period respectively, two prediction algorithms of yield were proposed, Least Square-Support Vector Machine (LSSVM) and Fuzzy Least Square-Support Vector Machine (FLSSVM). LSSVM prediction model took soil EC, total nitrogen content and K^+ as the input factors and winter wheat yield as output. While FLSSVM prediction model took soil EC, nitrogen content, K^+ and gray relation as input factors and also winter wheat yield as output. Results showed that the prediction and validation R^2 of LSSVM model were 0.772 and 0.685 respectively. Prediction R^2 of FLSSVM was 0.8625, and validation R^2 of FLSSVM model was 0.8003. FLSSVM used Fuzz Similar Extent to fuzz input samples so that it could avoid over-training. Also because it was based on membership function, it had several advantages such as simple structure, efficient convergence, precise forecasting, and etc. FLSSVM had high accuracy prediction result and could be used in estimating yield and providing theory and technical support for precision management of crops.

Keywords: near infrared spectroscopy, waste cooking oil, support vector machine, parameters optimization

1   Introduction

Wheat is one of China's three major food crops, and most of wheat is winter wheat. Water and fertilizer management is very necessary to winter wheat. Soil electrical conductivity (EC) is a basic index to reflect soil electrochemical properties and fertility characteristics. Soil fertility, as the basic property of soil and the foundation of land productivity, is defined as the soil ability to supply the essential plant nutrients, as well as related to the nutrient supply capacity of various soil properties and state by Soil Science Society of USA. In recent years, combined with soil physical, chemical, biological and environmental conditions, some researches on comprehensive evaluation of soil fertility showed that, soil EC was closely related with the soil properties and could be used as a quantitative index of comprehensive evaluation of soil fertility. Since the measurement of soil EC, can effectively control the salt concentration, soil water status and other properties, and also can timely diagnose agricultural production problems, the soil EC in guiding agricultural production, precision cultivation, has a special status and important role [1-3]. Zhao Yong stated that the relationship between soil EC and yield of winter wheat was linear correlation, so that we can use soil EC value of winter wheat in heading period as evaluation index of winter wheat yield [4]. Min thought that high EC value could improve tomato's lycopene, glucose, fructose and soluble solids content [5]. Jiang Dongyan studied the effects of irrigation amount under field condition and nitrogen fertilizer on wheat yield and soil nitrate nitrogen content. The results showed that increasing the amount of nitrogen fertilizer, the grain yield, protein content and protein yield were significantly increased [6]. However, the above research mostly only analyzed the influence of some factors on the yield, the variety of the soil parameters were not considered. In this paper, winter wheat was chosen as the object of study, and the influences of soil parameters on the yield of winter wheat were analyzed. A forecasting model of winter wheat yield was established to realize the wheat yield prediction by monitoring the soil parameters in the growth process, and the theory and technology were provided to support for the realization of fine management in the field of winter wheat.

2   Materials and Methods

2.1   Experimental material

The test was conducted in two years in National Precision Agriculture Demonstration Base. The experimental area was 30m×60m and divided into 60 cells, each of which was a 5m×5m zone. Sampling point was located in the intersection position of the cell and sampling depth was 5cm.

2.2 Experimental Methods

A soil EC meter DDB-307 (Shanghai Precision Scientific Instrument Co. LTD, Shanghai, China) was used to measure standard soil EC value. Each sample was cleaned by deionizer water before each measurement, so as not to affect the measurement result. Winter wheat yield information was provided by a combine harvester CASE2366 (CASE IH, USA) with global positioning system (GPS). When the GPS receiver provided accurate position information, the whole plot of yield can be obtained based on the flow data of CASE2366. And then the relationship between soil EC and yield of wheat were analyzed [7].

3 Results and Discussion

3.1 Analysis of influence factors of soil EC

As the amount of charge in soil can affect soil EC, the influence of nitrogen content, K^+, NO_3^- and pH value of soil was analyzed. Gary relation analysis method was used, which was used to compare the geometric relationship of comparative sequence data. If two comparative sequences in each moment overlapped together, the association degree is 1. Otherwise, the association degree is smaller than 1. Specific methods are introduced as follows.

X_0 is reference sequence $X_0=\{x_0(k)|k=1,2,...,n\}$, X_i is comparative sequence $X_i=\{x_i(k)|k=1,2,...,n\}$, where n denotes the number of data of reference sequence and comparative sequence. Then correlation coefficient is calculated as:

$$\zeta_{i,0}(k)=\frac{\min\limits_{i}\min\limits_{k}|x_0(k)-x_i(k)|+\rho\max\limits_{i}\max\limits_{k}|x_0(k)-x_i(k)|}{|x_0(k)-x_i(k)|+\rho\max\limits_{i}\max\limits_{k}|x_0(k)-x_i(k)|} \tag{1}$$

ρ is called resolution coefficient, is a pre-fixed constants between zero to one, generally taking as 0.5. $\zeta_{i,0}(k)$ represents the relative difference of comparative sequence X_i and the reference sequence X_0 at the moment k, and reflects similarity between the different sequence and the similarity of the reference sequence in the same point. Because the correlation coefficient is comparative sequence and reference sequence at each time point (i.e. curve points) degree value, the values of the efficient are bigger than 1, and the information is too scattered to facilitate an overall comparison. This study using average correlation coefficient of each times represents the correlation relationship between comparative sequence and the reference sequence correlation formula is as follows:

$$r(x_i,x_0)=\frac{1}{n}\sum_{i=1}^{n}\zeta_{i,0}(k) \quad ; \quad r(x_i,x_0)\in(0,1] \tag{2}$$

According to formula (2), the gray relation between soil EC and one of the other soil parameters, nitrogen content, K^+, NO_3^-, and pH, were calculated, and all results are shown in Table1[8-10]. Form Table1 we can see that the gray relation results were K^+>nitrogen content >pH>water content >weight > NO_3^-. Soil organic matter and soil cation exchange capacity (CEC) had a high significant positive correlation, which meant that the soil had more organic matter and more soil cation. Anion adsorption on soil colloids is more, and the greater the soil EC is. The grey correlation degree between K^+ and soil EC was high. Application of nitrogen fertilizer can promote the dissociation of cation adsorption on soil colloids, increase K^+, Na^+, Ca^{2+}, Mg^{2+} content, so that the grey correlation degree of nitrogen content in soil and soil EC was relatively high[11-14]. Soil EC and pH had high correlation; it might be because the pH change affected the amount of charge soil colloid and soil EC. The gray relation between winter wheat yield and one of the soil parameters, soil EC , nitrogen content, K^+, NO_3^-, pH, were calculated. Form Table1 we can see that the gray relation results were soil EC>nitrogen content >pH>soil water content > NO_3^-> K^+>soil weight.

Table 1. Gray relation between wheat yield and all indexes

Name	Correlation with EC	Gray relation with yield
Weight of soil	0.6891	0.2376
Water content of soil	0.7352	0.4783
Nitrogen content	0.8215	0.6421
K^+	0.8358	0.4206
NO_3^-	0.6548	0.4529
pH	0.7361	0.5632
Soil EC	1.0000	0.5839

In order to better understand the relationship between soil EC and nitrogen content, K^+, NO_3^-, pH and yield, the coefficients of correlation were calculated and results are shown in Table 2.

Table 2 Correlation between wheat yield and all indexes in different growth time

Name	Correlation coefficient (heading stage)	Correlation coefficient (filling stage)	Correlation coefficient (milky stage)
Soil EC	-0.3887	-0.5385	-0.6274
Nitrogen content	0.5360	0.5842	0.5925
K^+	0.4863	0.4247	0.4014
NO_3^-	-0.2992	-0.2674	-0.2258
pH	-0.1462	-0.0979	-0.1036

From Table 2 we can find that soil EC, nitrogen content and K^+ had notable correlation with yield in the heading stage, grain filling stage and milk stage. NO_3^- and pH had low correlation with yield; it was because although the ammonium

nitrogen and nitrate had the same value as nitrogen source of plants, if the two kinds of nitrogen source could be chosen, the relative absorption rate of different plants still had obvious difference. Generally, tobacco, cotton crop response to nitrogen is better, while wheat is more sensitive to ammonium nitrogen. The correlation analysis between soil EC and yield showed there was a linear correlation between them. Since there were more soil nutrients in the early stage of growth of winter wheat, soil contribution to the winter wheat was larger, and there were more obvious changes in the value of soil EC. Since winter wheat photosynthesis was strong in the late stage of the growth, winter wheat got less nutrition from soil, and the soil EC was stable, and yield significantly associated with soil EC.

3.2 Prediction model of winter wheat yield

After analyzing the correlation between soil EC, nitrogen content, K^+ and yield, soil EC, nitrogen content and K^+ were chosen as input, and yield was used as output, LSSVM and FLSSVM were used to build winter wheat yield prediction model.

(1) LS-SVM

The core of LS-SVM (least square-support vector machine) is using the kernel function to change the nonlinear problem into a linear problem, and get better data classification or regression results in latent feature space operations. At present, the common kernel function of SVM is linear kernel, polynomial kernel function (ploy), radial basis function (RBF) kernel function and sigmoid kernel function. Among them, the RBF kernel function with fewer parameters and high regression accuracy is adopted in this study. For RBF kernel function, the optimal parameters include the penalty parameter "C", "γ" parameters and loss parameters "ε". "C" is the regularization parameter, and is used to handle the complexity and the training error of the model and to obtain good generalization; Parameter "γ" reflects the data distribution or range property, which determines the width of the local neighborhood. Because "C" and "γ" have larger effect on the model, the two parameters are the key point to establish the high precision model based on SVM. As shown in Figure2, using SVM model to predict yield, and predict R^2 is 0.772, verification R^2 is 0.685.

(2) Wheat yield forecast based on fuzzy least squares support vector machines

Least squares support vector machine (LSSVM) is sensitive to outliers, and it will bring over-fitting problems. Thus, this paper tried to use the concept of fuzzy membership LSSVM. A fuzzy least squares support vector machine support vector domain description (FLSSVM) was proposed in the paper. Data samples were mapped into a high dimensional space and the smallest enclosing ball was built, and then the membership values based on the sample from the heart to the ball were determined, while the isolated points outside the centralized ball was given a small positive number to reduce the impact. In this way it can improve noise immunity and fitting results of FLSSVM. FLSSVM used fuzzy membership as fuzzy input samples, and gave samples different degree of membership values based on their importance,

and then trained samples. In this paper, input vectors of sample were soil EC, nitrogen content and K^+. Furthermore, introducing the concept of fuzzy membership degree of the input vector, which was called gray correlation weights, the new input vectors now were soil EC, nitrogen content, K^+ and gray correlation. The gray correlation degree represented the weight vector between the variables and yield. The gray correlation between soil EC, nitrogen content, K^+ and yield are shown in Table 2. Different samples corresponding gray correlation as the membership value would make the forecast result more accurate. Specific algorithm of FLSSVM is shown below:

The LSSVM optimization problem can be expressed as:

$$y_i = \omega\phi(x_i) + b + \xi_i$$

$$Q = \min\left\{\frac{1}{2}\|\omega\|^2 + \frac{c}{2}\sum_{i=1}^{n}\xi_i^2\right\} \tag{3}$$

ω is weight vector, b is offset, ξ_i is error term, c is penalty parameter. The index ψ_i was used to distinguish the importance of different data in the training process in FLSSVM, which was called fuzzy membership and the formula above can be changed as follow:

$$Q = \min\left\{\frac{1}{2}\|\omega\|^2 + \frac{c}{2}\sum_{i=1}^{n}\psi_i\xi_i^2\right\} \tag{4}$$

The steps of FLSSVM are shown below：

(a) Normalizing the data; (b) Forming the training samples and the testing samples; (c) Establishing objective function of Q with the training sample; (d) Solving the objective function; (e) Building prediction model for wheat yield using samples. The prediction results are shown in Figure 2, R_v^2 and R_c^2 are 0.8625 and 0.8003. The result shows that, soil EC, nitrogen content and K^+ can be used to predict winter wheat yield. Based on the three soil parameters, increasing the degree of grey incidence as a weight vector to be fourth input vector and then building the FLSSVM winter wheat yield prediction model has a high forecasting precision.

FLSSVM used fuzzy membership to fuzzy input samples, and gave different samples with different degrees of membership values, and then trained model. This method can solve the problem of small sample learning, and avoid the defects of over learning. AS using the membership data for the weight, the method has high precision and is easy to realize. It is obtained that using FLSSVM to predict winter wheat yield has higher prediction accuracy and is practical.

4 Conclusions

Gray theory was used to analyze the gray relation between soil EC value and each of other soil parameters, total nitrogen content, K^+, NO_3^- and pH of soil. Results showed that there were high gray relations between soil EC and total nitrogen content, K^+, pH

of soil Correlation coefficient data between yield and soil EC, nitrogen content, K^+, NO_3^-, pH value of soil were analyzed in different growth periods of winter wheat. The EC value of soil, the nitrogen content and K^+ has significant correlation with yield in the heading stage, filling stage and milk stage respectively. Therefore, taking the soil EC value, nitrogen content and K^+ as the model input, and yield as the output of the model, LS-SVM model was built to predict the yield of wheat. The prediction R_c^2 of the model is 0.772, and the validation R_v^2 is 0.685. The fuzzy membership concept into LSSVM is proposed to support fuzzy least square support vector domain description of the support vector machine (FLSSVM). The results show that, we using soil EC value, nitrogen content, K^+, grey correlation degree as the input, and the wheat yield as output, the prediction R_c^2 of the FLSSVM model is 0.8625, and the validation R_v^2 of the model is 0.8003. It was concluded that FLSSVM prediction model with high precision can be used to estimate crop yield, to provide theoretical and technical support for precision agriculture management.
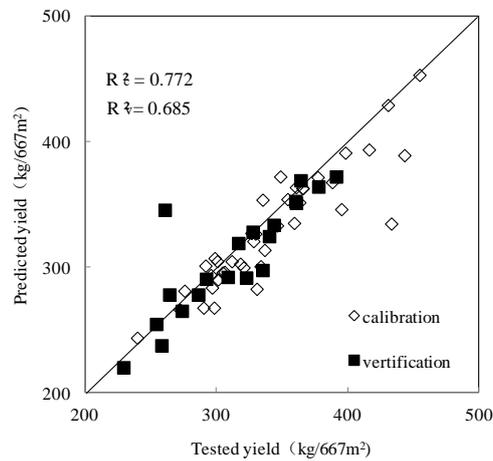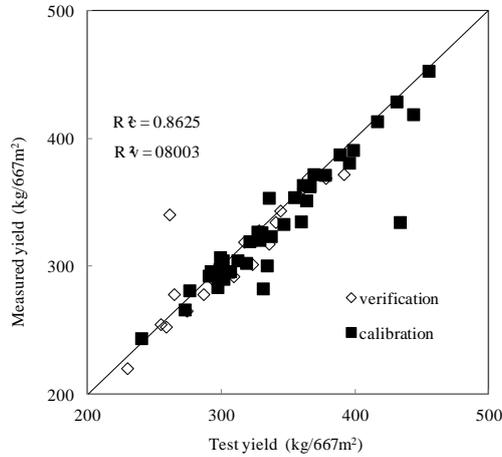

Fig. 1. The result of LSSVM

Fig. 2. The result of FLSSVM

References

1. Liu Guangming, Yang Jinsong, Ju Maosen, et al. Technology of chronometry using electromagnetic induction and its application in agriculture. Soil, 2003, (3): 27-29.
2. Wang Qi, Li Minzan, Wang Maohua. Development of a portable detector for soil electrical conductivity. Journal of China Agriculture University, 2003, 8(4): 20-23.
3. Naiqian Zhang, Kyeong-Hwan Lee, et al. Simultaneous Measurement of Soil Water content and Method Salinity Using a Frequency-Response, 2004 CIGR International Conference, 11~14 October 2004, Beijing, China.
4. Zhao Yong, Li Minzan, Zhang Junning. Correlation between soil electrical conductivity and winter yield [J]. Transactions of the CSAE, 2009, 25(Sl): 34~37. (In Chinese with English abstract)
5. Min Wu, Chieri K. Effects of high electrical conductivity of nutrient solution and its application timing on lycopene, chlorophyll and sugar concentrations of hydroponic tomatoes during ripening[J]. Scientia Horticulturae, 2008, 116(2): 122-129.
6. Lin Yicheng, Ding Nengfei, Fu Qinglin, et al. The measurement of electric conductivity in soil solution and analysis of its correlative factors[J]. Acta Agriculturae Zhejiangensis, 2005, 17(2): 83-86.
7. Shi Yongchen, Sui Jidong, He Chuanqin. Fast determination of soil electrical conductivity [J]. Journal of Heilongjiang August First Land Reclamation University, 2000, 12(4): 15-18.

8. Xie Naiming, Li Sifeng. Discrete GM(1,1) and mechanism of grey forecasting model [J]. System engineering-Theory&practice, 2005, 25 (1): 93-99.

9. Chen Zhujun, Wang Yijun, Xu Anmin, et al. Effects of the application of different nitrogen fertilizers on the ion compositions in solution of the greenhouse soil. Plant nutrition and fertilizer science, 2008, 14(5): 907-913.

10.Ritter C，Dicke D, Weis M, et al. An on-farm approach to quantify yield variation and to derive decision rules for site-specific weed management [J]. Precision Agric., 2008, 9(3): 133-146．

11. Nadler A. Effect of soil structure on bulk soil electrical conductivity (Eca) using the TDR and 4P techniques [J]. Soil Sci., 1991, 152: 199-203.

12. Kitchen N R, Sudduth K A, Drummond S T. Soil electrical conductivity as a crop productivity measure for clay pan soils [J]. Prod Agric., 1999, 12(4): 607-617.

13. Rhoades J. D., Corwin D. L. Determining Soil Electrical Conductivity-depth Relations Using an Inductive Electromagnetic Conductivity Meter. Soil Sci. Soc. Am., 1992, 45: 255-260.

14. McBride R. A., Gordon A. M., and Shrive S. C. Estimating Forest Soil Quality from Terrain Measurements of Apparent Electrical Conductivity. Soil Science Society America Journal, 1990, 54: 290-293.