doi:10.6041/j.issn.1000-1298.2015.04.040

基于支持向量机的叶绿素荧光预测光能利用效率研究*

任　顺[1]　于海业[1]　周丽娜[2]

(1. 吉林大学生物与农业工程学院, 长春 130022; 2. 吉林农业大学工程技术学院, 长春 130118)

摘要: 基于激光诱导叶绿素荧光光谱分析技术,提出了一种基于支持向量机理论的光能利用效率预测方法。同步采集黄瓜叶片的叶绿素荧光光谱、净光合速率和光合有效辐射,选取 500 ~ 800 nm 波段的叶绿素荧光光谱作为研究对象,首先对原光谱进行 SG - FDT 预处理;其次对预处理的光谱采用 PCA 方法提取特征值,根据累计贡献率选取前 10 个主成分代替原光谱信息;最后通过支持向量机建立光能利用效率预测模型。通过对惩罚系数 C 和核函数参数 g 的大量测试,最终确定 C 为 0.031 25、g 为 1,并利用 60 个训练样本对模型进行训练。10 个测试样本的预测结果表明,测试样本的平均误差为 8.94%,具有很好的预测效果。

关键词: 叶绿素荧光光谱　光能利用效率　主成分分析　支持向量机

中图分类号: S123; O657.3　　**文献标识码:** A　　**文章编号:** 1000-1298(2015)04-0273-04

Predicting Light Use Efficiency with Chlorophyll Fluorescence Spectra Based on SVM

Ren Shun[1]　Yu Haiye[1]　Zhou Li'na[2]

(1. *School of Biological and Agricultural Engineering, Jilin University, Changchun* 130022, *China*
2. *College of Engineering and Technology, Jilin Agricultural University, Changchun* 130118, *China*)

Abstract: Light use efficiency is an important parameter of plant productivity model. It is an evaluation index for plant to turn the solar energy into dry matter efficiency. Taking cucumbers as the study objects, a method for light use efficiency prediction was proposed with the help of analysis technique of laser-induced chlorophyll fluorescence spectra based on the theory of support vector machine (SVM). Chlorophyll fluorescence spectra, net photosynthetic rate and photosynthetic active radiation of cucumber leaves were synchronously acquired, and the 500 ~ 800 nm band of chlorophyll fluorescence spectrum was selected as study objects. Firstly, the original spectra was pretreated by SG - FDT method. Secondly, the characteristic values of pretreated spectra were extracted by using principal component analysis (PCA) method, the first ten principal components whose cumulative contribution rate was 93.49% were selected instead of the original spectral information in the study. Finally, the prediction model of light use efficiency was established through the SVM with the radial basis function. The penalty parameter C and kernel function parameter g were ultimately determined as $C = 0.031\ 25$, $g = 1$ by carrying out a large number of tests, and then 60 training samples were combined to train the model. Ten testing samples were used to test the established model, and the results showed that the average error of the testing samples was 8.94%, which indicated a good predictive power.

Key words: Chlorophyll fluorescence spectra　Light use efficiency　Principal component analysis　Support vector machine

收稿日期: 2014 - 11 - 09　修回日期: 2014 - 12 - 27

* 国家高技术研究发展计划(863 计划)资助项目(2012AA10A506 - 4、2013AA103005 - 04)

作者简介: 任顺,博士生,主要从事农业设施调控与节能技术研究,E-mail: 535884264@qq.com

通讯作者: 于海业,教授,博士生导师,主要从事农业设施调控与节能技术研究,E-mail: haiye@jlu.edu.cn

引言

利用叶绿素荧光光谱检测技术监测植株生长及环境逐渐受到人们的关注[1]。叶绿素荧光光谱分析技术作为一种快速、准确、无损植物生理信息采集方法，它为评价植物在环境中的生长状况提供了理论依据与技术支持[2-4]。

光能利用效率是植株生产力模型的重要参数，它是评价植株将太阳能转化成干物质效率的指标，其受各种因素限制，如温湿度、光照、叶绿素含量等[5-6]。植株在光合过程中，部分能量用于 CO_2 的固定，绝大多数能量通过热的形式耗散掉，还有一小部分能量以荧光的方式释放掉。荧光虽然仅占吸收总能量的 2% ~5%，却被誉为光合功能的理想“探针”，足以作为植株生长状态诊断的依据[7-11]。

支持向量机(Support vector machine, SVM)是基于统计学习理论的小样本学习方法，采用结构风险最小化和 VC 维理论构建学习机，具有很好的泛化性能，其在小样本、非线性以及高维模式识别中表现出特有的优势[12-15]。支持向量机技术在农业研究中具有良好的效果[16-18]。

本文从叶绿素荧光光谱分析的角度研究叶片光能利用效率与荧光光谱之间的关系。以嫁接黄瓜苗作为研究对象，以黄瓜叶片的叶绿素荧光光谱和光能利用效率为测试数据，提取叶绿素荧光光谱特征并通过支持向量机建立相应的光能利用效率预测模型。

1　材料与方法

试验于 2014 年 6 月 20 日—6 月 21 日 10:00 在吉林大学生物与农业工程学院现代化温室外进行，试验品种为嫁接黄瓜苗。挑选长势均匀、健康的黄瓜叶片为测量对象，同步采集黄瓜叶片的叶绿素荧光光谱和光合作用参数。

1.1　荧光光谱的测定

试验使用荷兰产 AvaSpec-2048-USB2 型光纤光谱仪采集系统，可测光谱波段为 360 ~1 100 nm，分辨率 2.1 nm。激发源使用中国科学院长春光学精密机械与物理研究所产的 MBL-Ⅲ-473 nm 型固体激光器[19]。为防止自然光对荧光光谱产生影响，将待测叶片置于黑色垫板之上，固体激光器和荧光探测器内置于黑色长方体内，激发光与荧光探测器呈 45°角且贴近叶片表面进行叶绿素荧光光谱采集，如图 1 所示。

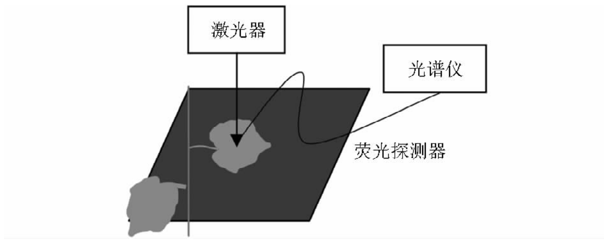

图 1　光谱采集示意图

Fig. 1　Schematic diagram of spectral acquisition

1.2　荧光光谱预处理

为避免噪声等对光谱的影响，首先采用 AvaSoft 7.2 for AvaSpec-USB2 软件以消除仪器等因素对光谱数据的影响。并选取荧光光谱 500 ~800 nm 波段进行分析，同时有研究发现光谱的一阶导数适合于荧光光谱分析，利用 Origin 9 软件对荧光光谱进行 Savitzky-Golay 平滑法(SG)和一阶导数变换(FDT)预处理，结果如图 2 所示。

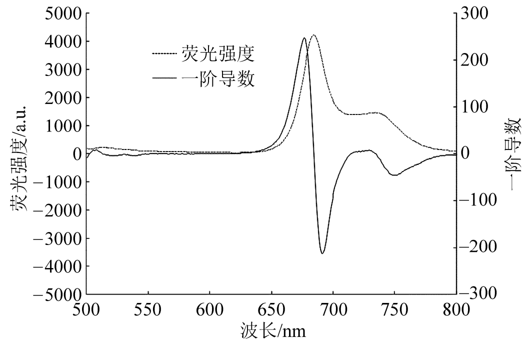

图 2　荧光光谱及 SG-FDT 预处理

Fig. 2　Fluorescence spectrum and fluorescence spectrum with SG-FDT pretreatment

1.3　光能利用效率的测定

试验使用美国产 LI-6400 型便携式光合作用仪对黄瓜叶片进行测定，测量时打开叶室，夹好待测量的黄瓜叶片，待参数基本稳定后，记录数据，每片叶片重复测量 3 次取平均值。光能利用效率为

$$L_{ue} = P_n / P_{ar} \times 100\% \quad (1)$$

式中　L_{ue}——光能利用效率

P_n——叶片净光合速率

P_{ar}——光合有效辐射

2　试验结果与分析

2.1　光谱特征信息提取

由于 500 ~800 nm 荧光光谱波段数据冗杂，采用主成分分析方法对预处理过的叶绿素荧光光谱进行降维，以提取光谱特征信息。主成分提取采用 Matlab 2010a 软件中 princomp 函数进行。表 1 为叶绿素荧光光谱的前 10 个主成分累计贡献率。

从表 1 可以看出前 5 个主成分可以表示出原变量 90.79% 的信息，为了使降维后主成分尽可能多的反映原变量信息，本文提取前 10 个主成分。降维后变量值通过计算 PC1 ~PC10 主成分的得分代替

原光谱信息，从而完成对光谱特征信息的提取。

表 1　前 10 个主成分方差及累计贡献率

Tab. 1　Variance and contribution of the first ten principal components

主成分	方差	累计贡献率/%	主成分	方差	累计贡献率/%
PC1	397.683	75.89	PC6	3.690	91.50
PC2	41.138	83.74	PC7	3.145	92.10
PC3	17.167	87.02	PC8	2.679	92.61
PC4	15.006	89.88	PC9	2.473	93.08
PC5	4.771	90.79	PC10	2.124	93.49

2.2　光能利用效率模型的建立

2.2.1　支持向量机算法

支持向量机是通过内积函数定义的非线性变换将输入空间变换到一个高维空间，在高维空间中寻找最优分类面，结构示意图如图 3 所示。

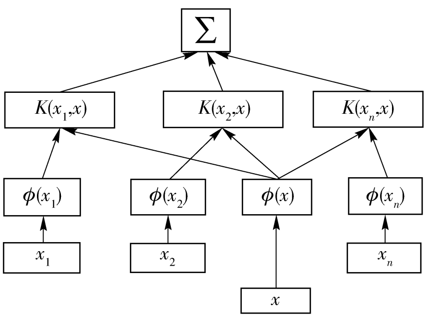

图 3　支持向量机结构图

Fig. 3　Sketch of support vector machines

不同的核函数可以构造实现输入空间中不同类型学习机器。其中常用的核函数有：

（1）多项式核

$$K(x,y)=(xy+1)^d \quad (d=1,2,\cdots,n)$$

（2）径向基核

$$K(x,y)=\exp\left(-\frac{|x-y|^2}{\sigma^2}\right)$$

（3）Sigmoid 核

$$K(x,y)=\text{th}(\phi(xy)+\theta)$$

2.2.2　模型的建立

假设有光能利用效率实例 (x_i,y_i) $(i=1,2,\cdots,k)$，$x_i\in\mathbf{R}^n$ 为影响光能利用效率的因素，$y_i\in\mathbf{R}$ 表示光能利用效率。支持向量机光能利用效率模型的建立，就是寻找 x_i、y_i 之间的关系

$$f:\ \mathbf{R}^n\rightarrow\mathbf{R} \quad y_i=f(x_i) \quad (i=1,2,\cdots,k) \tag{2}$$

根据支持向量机理论，光能利用效率预测模型的建立

$$f(x)=\sum_{i=1}^{k}(\alpha_i-\alpha_i^*)K(x,x_i)+b \tag{3}$$

式中　x——待预测样本的主成分得分

x_i——k 个样本中第 i 个样本主成分得分

$K(x,x_i)$——核函数

2.3　模型参数的选择及预测结果

选取 70 株黄瓜苗叶片的叶绿素荧光光谱，其中随机选取 60 个样本作为光能利用效率的预测模型学习样本，其余 10 个样本用于测试模型。

以径向基函数作为核函数建立预测模型，对各种惩罚参数 C 和核函数参数 g 的测试，确定 $C=0.031\,25$，$g=1$。利用式（3）对 10 个样本进行测试，测试结果见表 2。

表 2　支持向量机预测结果与实际值比较

Tab. 2　Comparison of SVM-predicted results with measured values

样本序号	真实值/%	预测值/%	相对误差/%
1	2.706 29	2.714 50	0.30
2	2.249 85	2.545 69	13.10
3	1.989 52	2.359 84	18.60
4	1.905 62	1.933 86	1.48
5	1.849 74	2.087 92	12.88
6	1.699 68	1.822 20	7.21
7	1.648 08	1.537 23	6.73
8	1.597 12	1.511 85	5.34
9	1.558 79	1.649 50	5.82
10	1.507 30	1.776 61	17.87

从表 2 中 10 个测试样本得到的相对误差可知，基于支持向量机建立的预测模型平均相对误差为 8.94%。

2.4　模型参数对预测结果的影响

在模型的建立过程中，考虑不同支持向量机参数对预测结果的影响，首先利用 libsvm 工具箱中的 gridregression.py 搜索得到一个较优的初始参数 C 和 g，再在这两个值附近采用单因素法和不同 C、g 组合对参数 C、g 进行优化[20]。表 3 所示为在光能利用效率预测模型中不同 C、g 组合对平均相对误差的影响，图 4、5 所示为在光能利用效率预测模型中 C 和 g 单因素对平均相对误差的影响。

表 3　不同 C、g 组合对预测结果的影响

Tab. 3　Influence of C and g on the result of prediction

序号	C	g	相对误差/%
1	0.031 25	0.125	9.0
2	0.031 25	0.5	9.2
3	0.031 25	1.0	8.9
4	0.125	0.125	14.7
5	0.125	0.5	14.0
6	0.125	1.0	9.9
7	0.5	0.125	17.5
8	0.5	0.5	15.1
9	0.5	1.0	10.1

从表3可知,当 $C=0.031\,25$,$g=1$ 时,测试样本得到的平均相对误差最小,为8.9%。为了进一步优化参数 C 和 g,对 $C=0.031\,25$ 时,不同 g 的取值;以及 $g=1$ 时,不同 C 的取值进行了单因素的分析,结果分别如图4、图5所示。

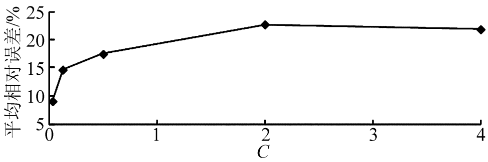

图4 C 值对预测结果的影响($g=1$)

Fig.4 Influence of C on result of prediction ($g=1$)

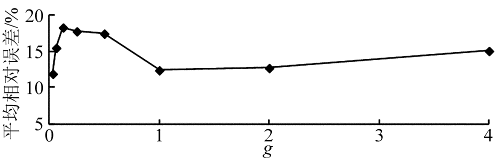

图5 g 值对预测结果的影响($C=0.031\,25$)

Fig.5 Influence of g on result of prediction ($C=0.031\,25$)

从图4、图5可知平均相对误差最小时 C 和 g 的值都是0.031 25和1,综合比较,模型参数的取值为 $C=0.031\,25$、$g=1$。

3 结束语

研究了SVM在光能利用效率预测中的应用,采用激光诱导叶绿素荧光光谱分析技术对黄瓜光能利用效率预测模型进行了研究。在保证原光谱90%以上信息的情况下,根据主成分累计贡献率选取PCA处理后前10个主成分的得分作为新的输入值,以光能利用效率作为回归值,建立了支持向量机光能利用效率模型。文中对惩罚系数 C 和核函数参数 g 进行了大量测试,最终确定 C 为0.031 25、g 为1,并结合60个训练样本得到了光能利用效率预测模型,通过10个测试样本的预测结果可知,测试样本的平均相对误差为8.94%,该模型的预测效果较好。

参 考 文 献

1 Ndao A S, Konté A, Biaye M, et al. Analysis of chlorophyll fluorescence spectra in some tropical plants[J]. Journal of Fluorescence, 2005, 15(2):123-129.

2 Bolhar-Nordenkampe H R, Long S P, Baker N R, et al. Chlorophyll fluorescence as a probe of the photosynthetic competence of leaves in the field: a review of current instrumentation[J]. Functional Ecology, 1989, 3(4): 497-514.

3 Class B. Variability and application of the chlorophyll fluorescence emission ratio red/far-red of leaves[J]. Photosynthesis Research, 2007, 92(2):261-271.

4 Sayed O H. Chlorophyll fluorescence as a tool in cereal crop research[J]. Photosynthetica, 2003, 41(3): 321-330.

5 Green D S, Erickson J E, Kruger E L. Foliar morphology and canopy nitrogen as predictors of light-use efficiency in terrestrial vegetation[J]. Agricultural and Forest Meteorology, 2003, 115(3): 163-171.

6 van Oijen M, Dreccer M F, Firsching K H, et al. Simple equations for dynamic models of the effects of CO_2 and O_3 on light-use efficiency and growth of crops[J]. Ecological Modelling, 2004, 179(1): 39-60.

7 Hendrickson L, Furbank R T, Chow W S. A simple alternative approach to assessing the fate of absorbed light energy using chlorophyll fluorescence[J]. Photosynthesis Research, 2004, 82(1): 73-81.

8 Allen J F. How does protein phosphorylation regulate photosynthesis[J]. Trends in Biochemical Sciences, 1992, 7(1): 12-17.

9 周丽娜,于海业,张蕾,等. 基于叶绿素荧光光谱分析的稻瘟病害预测模型[J]. 光谱学与光谱分析,2014,34(4): 1003-1006.
Zhou Li'na, Yu Haiye, Zhang Lei, et al. Rice blast prediction model based on analysis of chlorophyll fluorescence spectrum[J]. Spectroscopy and Spectral Analysis, 2014, 34(4): 1003-1006. (in Chinese)

10 杨昊谕,于海业,张蕾,等. 基于激光诱导荧光光谱分析的黄瓜叶片叶绿素含量检测[J]. 农业机械学报,2009,40(10): 169-172.
Yang Haoyu, Yu Haiye, Zhang Lei, et al. Detecting of chlorophyll content of cucumber leaves based on laser-induced fluorescence spectrum analysis technique[J]. Transactions of the Chinese Society for Agricultural Machinery,2009, 40(10): 169-172. (in Chinese)

11 周丽娜, 于海业, 于连军,等. 基于叶绿素荧光光谱分析的光能利用效率研究[J]. 农业机械学报,2014,45(7): 255-259.
Zhou Li'na, Yu Haiye, Yu Lianjun, et al. Evaluation of light use efficiency using the chlorophyll fluorescence spectra[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(7): 255-259. (in Chinese)

12 Cortes C, Vapnik V. Support vector networks[J]. Machine Learning, 1995, 20(3): 273-297.

13 Flake G W, Lawrence S. Efficient SVM regression training with SMO[J]. Machine Learning, 2002, 41(1): 271-290.

14 David M, Friedrich L, Kurt H. The support vector machine under test[J]. Neurocomputing-IJON, 2003, 55(1): 169-186.

15 Zhang Haoran, Wang Xiaodong, Zhang Changjing, et al. A new support vector machine for data mining[J]. Advanced Data Mining and Applications, 2005, 3584:256-266.

(下转第251页)

参　考　文　献

1 Jermini M, Geiges O, Schmidt-Lorenz W. Detection, isolation and identification of osmotolerant yeasts from high-sugar products [J]. Journal of Food Protection, 1987, 50: 468 – 472.

2 Deak T, Beuchat L. Yeasts associated with fruit juice concentrates[J]. Journal of Food Protection, 1993, 56(9): 777 – 782.

3 Thomas D. Yeasts as spoilage organisms in beverages[M]//Rose A, Harrison J. The yeasts. Vol. 5: yeast technology. 2nd ed. London: Academic Press, 1993: 517 – 561.

4 Tudor E A, Board R. Food-spoilage yeasts[M]//Rose A, Harrison J. The yeasts. Vol. 5: yeast technology. 2nd ed. London: Academic Press, 1993:435 – 516.

5 Worobo R W, Splittstoesser D F. Microbiology of fruit products[M]//Barret D M, Somogyi L, Ramaswamy H. Processing fruit. 2nd ed. Boca Raton: CRC Press, 2005: 161 – 284.

6 Combina M, Daguerre C, Massera A, et al. Yeast identification in grape juice concentrates from Argentina[J]. Letters in Appled Microbiology, 2008, 46(2):192 – 197.

7 Rojo M, Arroyo López F, Lerena M, et al. Effects of pH and sugar concentration in *Zygosaccharomyces rouxii* growth and time for spoilage in concentrated grape juice at isothermal and non-isothermal conditions[J]. Food Microbiology, 2014, 38: 143 – 150.

8 Koh T Y. Studies on the ‘osmophilic’ yeast *Saccharomyces rouxii* and an obligate to osmophilic mutant[J]. Microbiology, 1975, 88(1):101 – 114.

9 Steels H, James S, Roberts I, et al. *Zygosaccharomyces lentus*: a significant new osmophilic, preservative-resistant spoilage yeast, capable of growth at low temperature[J]. Journal of Applied Microbiology, 1999, 87(4):520 – 527.

10 Barnett J A, Payne R W, Yarrow D. Yeasts: characteristics and identification[M]. London: Cambridge University Press, 1983.

11 Kurtzman C P, Robnett C J. Three new insect-associated species of the yeast genus *Candida*[J]. Canadian Journal of Microbiology, 1998, 44(10): 965 – 973.

12 Baleiras-Couto M M, Gomes A S, Casal M, et al. Survey of yeast diversity during wine bottling processes using restriction analysis of 26S ribosomal DNA (rDNA)[J]. Australian Journal of Grape and Wine Research,2012, 18(1): 39 – 42.

13 Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL – X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools[J]. Nucleic Acids Research, 1997, 25(24):4876 – 4882.

14 Tamura K, Dudley J, Nei M, et al. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0[J]. Molecular Biology and Evolution,2007, 24(8):1596 – 1599.

15 谢勇,张建新,余清谋. 浓缩苹果汁中耐高渗透压酵母菌的分离与培养条件优化研究[J]. 中国酿造, 2007, 26(4): 4 – 9.
Xie Yong, Zhang Jianxin, Yu Qingmou. Isolation of osmophilic yeasts from concentrated apple juice and optimization of their incubation conditions[J]. China Brewing, 2007, 26(4): 4 – 9. (in Chinese)

16 李国良. 浓缩苹果汁中耐高渗酵母的检测与控制技术[D]. 杨凌:西北农林科技大学,2012.
Li Guoliang. Control and detection technology of osmphilic yeast in apple juice concentrate[D]. Yangling: Northwest A&F University, 2012. (in Chinese)

17 Tokuoka K. Sugar-and salt-tolerant yeasts[J]. Journal of Applied Microbiology, 1993, 74(2):101 – 110.

18 Deák T. Yeasts in specific types of foods[M]//Deák T. Handbook of food spoilage yeasts. 2nd ed. Boca Raton: CRC Press, 2008:117 – 201.

19 胡贻春,岳田利,袁亚宏,等. 浓缩苹果汁车间空气中嗜酸耐热菌的分离鉴定研究[J]. 农产品加工·学刊,2007(3):11 – 13,21.
Hu Yichun, Yue Tianli, Yuan Yahong, et al. Research on thermoacidiphilic bacteria isolated from the air of AJC production workshop[J]. Academic Periodical of Farm Products Processing,2007(3): 11 – 13,21. (in Chinese)

(上接第 276 页)

16 孙俊, 王艳, 金夏明, 等. 基于 MSCPSO 混合核 SVM 参数优化的生菜品质检测[J]. 农业机械学报, 2013, 44(9):209 – 213.
Sun Jun, Wang Yan, Jin Xiaming, et al. Detection of lettuce quality based on parameters optimization of MSCPSO mixed kernel SVM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(9):209 – 213. (in Chinese)

17 李淼, 方德洲, 张建, 等. 一种基于支持向量机的模糊施肥预测方法研究[J]. 计算机应用与软件, 2009, 26(2): 50 – 52.
Li Miao, Fang Dezhou, Zhang Jian, et al. A new approach for fuzzy fertilization forecast based on support vector mechanism[J]. Computer Applications and Software, 2009, 26(2): 50 – 52. (in Chinese)

18 隋媛媛, 于海业, 张蕾, 等. 基于叶绿素荧光光谱分析的黄瓜霜霉病害预测模型[J]. 光谱学与光谱分析, 2011, 31(11): 2987 – 2990.
Sui Yuanyuan, Yu Haiye, Zhang Lei, et al. Cucumber downy mildew prediction model based on analysis of chlorophyll fluorescence spectrum[J]. Spectroscopy and Spectral Analysis, 2011, 31(11): 2987 – 2990. (in Chinese)

19 于海业, 杨昊谕, 张蕾. 激光诱导式叶绿素荧光强度与激光强度关系[J]. 农业工程学报, 2009, 25(增刊 2): 245 – 249.
Yu Haiye, Yang Haoyu, Zhang Lei. Relationship between laser-induced chlorophyll fluorescence intensity and laser power[J]. Transactions of the CSAE, 2009, 25(Supp.2): 245 – 249. (in Chinese)

20 Huang Cheng-lung, Wang Chieh-jen. A GA-based feature selection and parameters optimization for support vector machines[J]. Expert Systems with Applications-ESWA, 2006, 31(2): 231 – 240.