

· 研究论文 ·

近红外光谱技术结合支持向量回归法 测定嘧菌酯的含量

吴厚斌^{1,2}, 刘苹苹², 吕 宁², 穆 兰², 刘丰茂¹,
张延秋², 江树人¹, 闵顺耕^{*1}

(1. 中国农业大学 理学院, 北京 100193; 2. 农业部 农药检定所, 北京 100125)

摘 要:采用近红外光谱(NIR)技术结合支持向量回归法(SVR)建立了乳油中嘧菌酯的定量分析方法。通过向25%苯醚甲环唑乳油中加入嘧菌酯原药和乙酸乙酯,配制不同浓度的校正集,采用SVR法建立了嘧菌酯的定量分析模型,得模型的决定系数(R^2)、校正集均方根误差(RMSEC)、检验集均方根误差(RMSEV)、预测集均方根误差(RMSEP)分别为0.999 7、0.000 9、0.000 9和0.001 7。结果表明,近红外光谱技术结合支持向量回归法可以准确地定量分析乳油中嘧菌酯的含量,方法简单、快捷,在农药市场质量快速监测中具有实际应用价值。

关键词:近红外光谱;支持向量回归;嘧菌酯;定量;分析

DOI:10.3969/j.issn.1008-7303.2011.06.09

中图分类号:O657.33

文献标志码:A

文章编号:1008-7303(2011)06-0608-05

Determination of the content of azoxystrobin by near-infrared spectroscopy and support vector regression

WU Hou-bin^{1,2}, LIU Ping-ping², LÜ Ning², MU Lan², LIU Feng-mao¹,
ZHANG Yan-qiu², JIANG Shu-ren¹, MIN Shun-geng^{*1}

(1. College of Science, China Agricultural University, Beijing 100193, China;

2. Institute for the Control of Agrochemicals, Ministry of Agriculture, Beijing 100125, China)

Abstract: The method of rapid determination of the content of azoxystrobin in EC was established by near-infrared spectroscopy (NIR), combined with support vector regression (SVR). The calibration set was composed of samples with different content of azoxystrobin by adding azoxystrobin TC and ethyl acetate to difenoconazole 250 EC. The determination coefficient (R^2), root mean squared error of calibration (RMSEC), root mean squared error of validation (RMSEV), and root mean squared error of prediction (RMSEP) for the model established by SVR for the content of active ingredient azoxystrobin were 0.999 7, 0.000 9, 0.000 9 and 0.001 7 respectively. The results showed that the method of NIR combined with SVR is convenient and quickly and can be used to accurately quantitative analysis of the content of azoxystrobin in EC.

Key words: NIR; SVR; azoxystrobin; quantitative; analysis

收稿日期:2011-06-04;修回日期:2011-11-01.

作者简介: 吴厚斌(1972-),男,浙江义乌人,博士研究生,电话:010-59194076, **E-mail:** dwhbin@sohu.com; * 通讯作者 (Author for correspondence): 闵顺耕(1963-),男,福建人,博士,教授,主要从事农产品安全及品质检测等研究工作,电话:010-62733091, **E-mail:** minsng@263.net

基金项目:国家自然科学基金资助项目(20575076).

嘧菌酯(azoxystrobin)^[1]是先正达公司成功开发的第一个甲氧基丙烯酸酯类杀菌剂,具有高效、广谱、抗菌活性高等优点,在国内外农药市场上广受好评。但也有一些不法农药生产企业为了提高其农药制剂的杀菌活性,采用了在其他杀菌剂中非法添加嘧菌酯的做法,已明显危害到市场秩序及农业生产和农产品质量安全,因此建立快速有效的农药制剂中嘧菌酯含量的分析方法具有重要意义。目前有关乳油中嘧菌酯含量的分析方法主要为高效液相色谱法^[2]和气相色谱法^[3],这些方法虽然结果准确,但是其前处理复杂、分析速度慢、成本高。笔者利用近红外光谱技术(NIR)结合支持向量回归法(SVR)^[4-7],建立了苯醚甲环唑乳油中嘧菌酯的定量分析方法,该方法不需要经过前处理步骤,分析速度快,成本低,具有较好的实际应用价值。

1 材料与方法

1.1 药剂及主要试剂

97%的嘧菌酯(azoxystrobin)原药(江苏耕耘化工有限公司提供);25%苯醚甲环唑乳油(difenoconazole,250 EC)(市场购买的河北省石家庄三农化工有限公司产品);乙酸乙酯(分析纯)。

1.2 主要仪器及数据处理软件

MPA傅里叶变换近红外光谱仪,OPUS 5.0光谱采集软件(德国,Bruker公司);Spectrum Quant+V4.51化学计量学软件及Spectrum V5.01光谱处理软件(美国,Perkin Elmer公司),Matlab 6.5版本。

1.3 实验方法

1.3.1 样品配制

嘧菌酯标样溶液配制:称取一定量的嘧菌酯原药,用乙酸乙酯溶解并配制成标准储备液,使用时再用乙酸乙酯稀释。共配制37个标样溶液,嘧菌酯的质量浓度范围为0.019 4~0.388 0 g/mL。

样品溶液配制:准确量取25%苯醚甲环唑乳油,加入已配制好的嘧菌酯标样溶液,共配制成37个样品溶液,其中嘧菌酯的质量浓度范围为0.009 7~0.194 0 g/mL。

1.3.2 光谱采集

采集方式:透射测量,以空气为参比。在每个样品进行光谱测量前先用丙酮冲洗石英杯,并用待测样品冲洗石英杯3次。

采集条件:测量波数范围为10 000~4 000 cm⁻¹,石英样品池光程为2 mm。采用8 cm⁻¹分辨率,光

谱扫描次数为64次。

1.4 数据处理

从偏最小二乘法 and 主成分分析可知,嘧菌酯模型最佳预测效果区间为10 000~4 000 cm⁻¹和6 100~5 449 cm⁻¹。以6 100~5 449 cm⁻¹为例,该组合区间共包含680个光谱变量。由于SVR涉及到样品间距离的计算,为减少原始光谱随机噪声的影响,本研究中采用10点平均法对原始光谱进行预处理以提高信噪比,将光谱变量数从680个压缩为68个。

1.5 样品选择

样品总数为37个,选取其中序号为5、10、15、20、25、30、35的共7个样品为模型预测集;其他30个样品中,每3个选择1个作为模型的独立检验集,共10个样品;剩余20个样品作为校正集;建立模型。

1.6 模型评价方法

采用支持向量回归(SVR)法建立校正模型,用独立检验集对模型进行评价,评价指标为模型的校正集均方根误差(RMSEC)和检验集均方根误差(RMSEV)。

2 结果与讨论

2.1 嘧菌酯原药和制剂的近红外光谱测定结果

图1为嘧菌酯原药的近红外光谱图。8 815 cm⁻¹为芳烃C-H的二级倍频,8 561 cm⁻¹为甲基C-H的二级倍频,7 396、7 302和7 199 cm⁻¹为芳烃C-H的组合频,7 011和6 941 cm⁻¹为甲基C-H的组合频,5 978、6 003、6 014和6 107 cm⁻¹为芳烃C-H伸缩振动的一级倍频,5 819 cm⁻¹为甲基C-H伸缩振动的一级倍频,4 665 cm⁻¹为芳烃C-H伸缩振动的组合频。

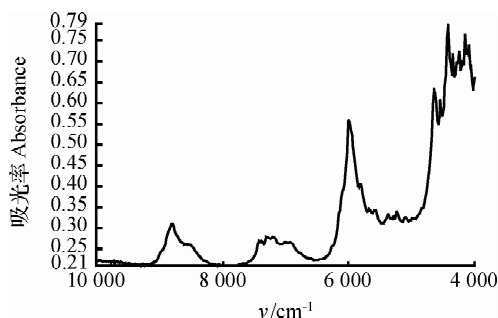


图1 嘧菌酯原药的近红外光谱图

Fig. 1 NIR spectrum of azoxystrobin TC

图2是苯醚甲环唑和嘧菌酯复配制剂的近红外光谱图,8 375 cm⁻¹为甲基C-H的二级倍频,

7 353 cm^{-1} 为甲基 C - H 的组合频,7 214 cm^{-1} 为亚甲基 C - H 伸缩振动的组合频,5 900 和 5 965 cm^{-1} 为甲基 C - H 的一级倍频,5 905、5 800 和 5 853 cm^{-1} 为亚甲基 C - H 伸缩振动的一级倍频,5 970 cm^{-1} 为与 C = O 相连甲基伸缩振动的组合频,5 666 和 5 618 cm^{-1} 为亚甲基 C - H 的一级倍频,5 160 cm^{-1}

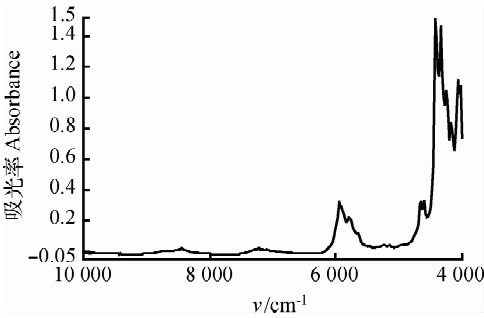


图 2 嘧菌酯制剂的近红外谱图

Fig. 2 NIR spectrum of azoxystrobin EC

为 C = O 的二级倍频。

2.2 嘧菌酯定量模型的建立及优化

利用 SVR 法建立嘧菌酯定量分析模型,通过调整 ε 不敏感损失函数、RBF 核函数的宽度 σ 和正则化系数 C 对模型进行优化。参数 ε 值控制 ε 不敏感区带的宽度,影响着构建模型支持向量的个数;该值过大,则校正精度低,支持向量数偏少,影响模型的预测性能。 σ 值与样本数据的置信度或信噪比相关,调整该值将影响非线性:随着 σ 值的增大,核变宽,模型趋向于线性。参数 C 表示模型复杂度(平坦度)与容许偏离量之间的平衡,一般情况下不宜太小,否则不能达到目标精度;也不宜过大(无限大),如过大则仅能最小化经验风险而不能考虑模型复杂度。选用不同的参数值对模型进行优化,初步的优化结果见表 1,其中 ε 为 0.000 1, $\lg \sigma^2$ 为 4, $\lg C$ 为 8 时, $RMSEV$ 为最小值 0.001 0,支持向量个数为 19,支持向量率为 95%。

表 1 初步优化后的 $RMSEV$ 值

Table 1 $RMSEV$ obtained with different parameters

$\lg \sigma^2$	ε	$\lg C$					
		3	4	5	6	7	8
0	0.01	0.006 5	0.006 5	0.006 5	0.006 5	0.006 5	0.006 5
1		0.006 5	0.006 5	0.006 5	0.006 5	0.006 5	0.006 5
2		0.010 5	0.006 6	0.006 6	0.006 6	0.006 6	0.006 6
3		0.042 8	0.010 5	0.006 6	0.006 6	0.006 6	0.006 6
4		0.051 1	0.042 8	0.010 5	0.006 6	0.006 6	0.006 6
0	0.001	0.001 7	0.001 7	0.001 7	0.001 7	0.001 7	0.001 7
1		0.001 2	0.001 7	0.001 8	0.001 8	0.001 8	0.001 8
2		0.002 1	0.001 2	0.001 2	0.001 8	0.001 8	0.001 8
3		0.043 7	0.002 1	0.001 1	0.001 0	0.001 0	0.001 0
4		0.050 6	0.043 7	0.002 1	0.001 1	0.001 0	0.001 0
0	0.000 5	0.001 6	0.001 4	0.001 9	0.001 9	0.001 9	0.001 9
1		0.001 1	0.001 6	0.002 0	0.002 0	0.002 0	0.002 0
2		0.002 0	0.001 0	0.001 0	0.001 9	0.001 9	0.001 9
3		0.043 7	0.002 0	0.001 0	0.001 0	0.001 1	0.001 0
4		0.050 6	0.043 7	0.002 0	0.001 0	0.001 0	0.001 0
0	0.000 1	0.001 5	0.001 5	0.004 3	0.004 4	0.004 4	0.004 4
1		0.001 2	0.001 6	0.001 7	0.003 5	0.004 0	0.004 0
2		0.002 1	0.001 1	0.001 3	0.001 7	0.001 9	0.001 9
3		0.043 7	0.002 1	0.001 1	0.001 1	0.001 1	0.001 1
4		0.050 6	0.043 7	0.002 1	0.001 1	0.001 0	0.001 0
0	0.000 01	0.001 6	0.001 6	0.004 5	0.005 0	0.005 0	0.005 0
1		0.001 2	0.001 6	0.001 7	0.003 6	0.004 5	0.004 5
2		0.002 3	0.001 1	0.001 3	0.001 7	0.002 0	0.002 0
3		0.043 7	0.002 3	0.001 1	0.001 2	0.001 1	0.001 1
4		0.050 6	0.043 7	0.002 3	0.001 1	0.001 0	0.001 0

对模型进行进一步优化。选择 $\lg \sigma^2$ 为 4、 $\lg C$ 为 8,调整 ε 值,从结果(见图 3)可知,在 C 值和 $\lg \sigma^2$ 值固定时,随着 ε 值的减小, $RMSEV$ 值也减小,

到一定程度时达到稳定,此时的 ε 值为 0.000 5,但该点的支持向量数只有 75%,因此选择 ε 值为 0.000 1,支持向量数为 95%。选择 ε 值为 0.000 1,

$\lg C$ 值为 8, 调整 $\lg \sigma^2$ 的值, 从结果 (见图 4) 可知, 在 ε 和 $\lg C$ 值固定时, 随着 $\lg \sigma^2$ 值的增大, $RMSEV$ 值减小, 到一定程度时达到稳定, 此时的 $\lg \sigma^2$ 值为 3.5。选择 ε 为 0.000 1、 $\lg \sigma^2$ 为 3.5, 调整 $\lg C$ 值, 从结果 (见图 5) 可知, 在 ε 和 $\lg \sigma^2$ 值固定时, 随着 $\lg C$ 值的增大, $RMSEV$ 值减小, 到一定程度时达到稳定, 此时的 $\lg C$ 值为 7.4。

因此, 选定优化后的模型参数为: ε , 0.000 1; $\lg \sigma^2$, 3.5; $\lg C$, 7.4; R^2 , 0.999 7; $RMSEC$, 0.000 9;

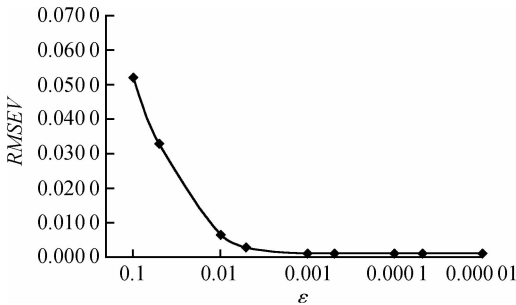


图 3 $RMSEV$ 值随 ε 值变化图

Fig. 3 $RMSEV$ values at various level of ε

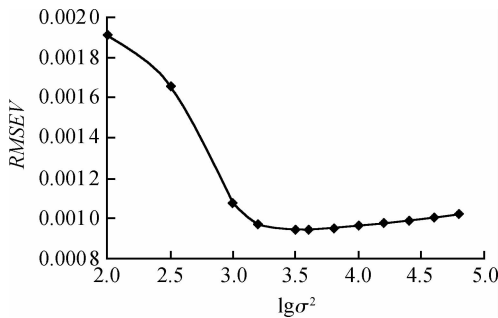


图 4 $RMSEV$ 值随 $\lg \sigma^2$ 值变化图

Fig. 4 $RMSEV$ values at various level of $\lg \sigma^2$

$RMSEV$, 0.000 9; 支持向量个数为 19; 支持向量率为 95%。模型校正集预测值与化学值的关系如图 6 所示。

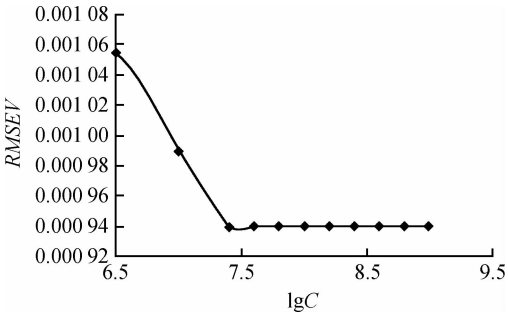


图 5 $RMSEV$ 值随 $\lg C$ 值变化图

Fig. 5 $RMSEV$ values at various level of $\lg C$

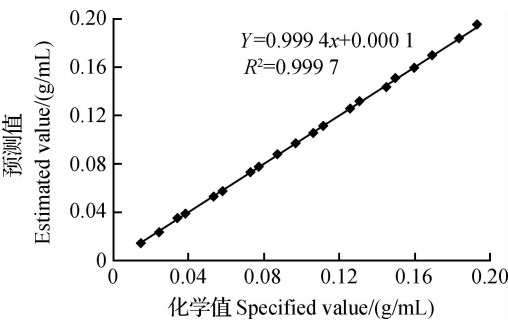


图 6 校正集预测值与化学值关系图

Fig. 6 Estimated value vs specified value of calibration set

2.3 对噻菌酯定量模型的检验

对 10 个检验集样品, 选择最优模型进行预测, 其预测结果 (见表 2) 表明, 检验集的 R^2 为 1.004 5, $RMSEV$ 为 0.000 9, 相对标准偏差 (RSD) 为 1.08%, 相对分析误差 (RPD) 为 55.38。检验集样品预测值

表 2 检验集的预测结果

Table 2 Estimated results of validation set

样品序号 NO.	化学值 Specified value/(g/mL)	预测值 Estimated value/(g/mL)	绝对误差 Absolute error/(g/mL)	相对误差 Relative error/%
3	0.174 6	0.174 0	0.000 6	-0.34
7	0.155 2	0.156 7	-0.001 5	0.93
11	0.135 8	0.136 9	-0.001 1	0.80
14	0.121 3	0.121 8	-0.000 4	0.37
18	0.101 9	0.103 4	-0.001 5	1.48
22	0.082 5	0.082 3	0.000 2	-0.29
26	0.063 1	0.061 7	0.001 4	-2.21
29	0.048 5	0.049 1	-0.000 6	1.28
33	0.029 1	0.029 6	-0.000 5	1.79
37	0.009 7	0.009 5	0.000 2	-1.82

与化学值的关系见图 7,模型稳定性和预测性能较好。10 个样品的相对误差均小于 3%,低于联合国粮农组织(FAO)允许的误差范围,可满足实际分析的要求。

2.4 噁菌酯定量模型的预测

对 7 个预测集样品,选择最优模型进行预测,

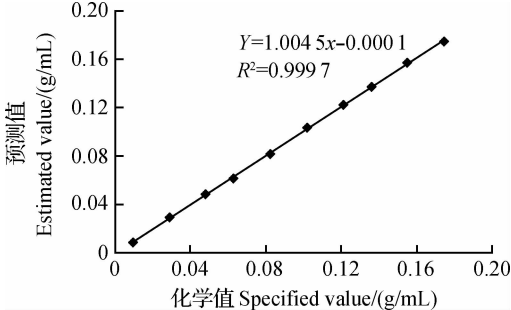


图 7 检验集预测值与化学值关系图

Fig. 7 Estimated value vs specified value of validation set

结果(见表 3)表明,预测集的 R^2 为 0.999 1, $RMSEP$ 为 0.001 7, RSD 为 2.03%, RPD 为 37.17。预测集样品预测值与化学值的关系见图 8。除 2 个样品因化学值较小造成相对误差较大外,其余 5 个样品的相对误差均小于 3%,低于 FAO 允许的误差范围,可满足实际分析的要求。

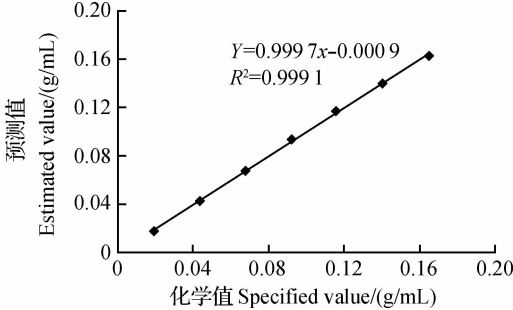


图 8 预测集预测值与化学值关系图

Fig. 8 Estimated value vs specified value of prediction set

表 3 预测集的预测结果
Table 3 Estimated results of prediction set

样品序号 NO.	化学值 Specified value/(g/mL)	预测值 Estimated value/(g/mL)	绝对误差 Absolute error/(g/mL)	相对误差 Relative error/%
5	0.164 9	0.161 9	0.003 0	-1.82
10	0.140 7	0.139 6	0.001 1	-0.78
15	0.116 4	0.116 8	-0.000 4	0.34
20	0.092 2	0.093 8	-0.001 6	1.74
25	0.067 9	0.067 5	0.000 4	-0.59
30	0.043 7	0.042 2	0.001 5	-3.43
35	0.019 4	0.017 0	0.002 4	-12.37

3 结论

利用近红外光谱技术结合支持向量回归法,建立了噁菌酯含量的快速分析方法。结果表明,在苯醚甲环唑乳油中加入噁菌酯乙酸乙酯溶液,配制不同浓度的校正集,采用支持向量回归法建立噁菌酯的定量分析模型,样品的预测值和化学值之间具有很好的相关性,模型的决定系数(R^2)、校正集均方根误差($RMSEC$)、检验集均方根误差($RMSECV$)和预测集均方根误差($RMSEP$)分别为 0.999 7、0.000 9、0.000 9 和 0.001 7,该模型可以准确地定量分析苯醚甲环唑乳油中噁菌酯的含量。利用近红外光谱技术和支持向量回归法建立模型的分析方法简单、快捷,在农药市场质量快速监测中具有较好的实际应用价值。

参考文献:

[1] YE Ji-ming(叶纪明),CHEN Tie-chun(陈铁春),WU Hou-bin

(吴厚斌), *et al.* Handbook of Harmonization Code for Pesticide in China(中国农药进出口商品编码实用手册)[M]. Beijing(北京):China Agriculture Press(中国农业出版社),2006:446.
[2] YU Kang-ping(于康平),LI Ze-fang(李泽方),XU Shao-kang(徐韶康), *et al.* 噁菌酯的高效液相色谱分析[J]. *Pesticide*(农药),2008,47(4):275-276.
[3] LI Yan(李岩),YU Rong(于荣),JIANG Yi-fei(姜宜飞), *et al.* 苯甲·噁菌酯 325 克/升悬浮剂气相色谱分析方法研究[J]. *Pestic Sci Manage*(农药科学与管理),2010,31(11):44-46.
[4] VAPNIK V. The Nature of Statistical Learning Theory[M]. New York,USA:Springer-Verlag New York,Inc,1995.
[5] CORTES C,VAPNIK V. Support-vector networks[J]. *Machine Learning*,1995,20(3):273-297.
[6] AN Wen-sen(安文森),SUN Yan-guang(孙彦广). 一种新的支持向量回归核函数构建方法[J]. *Information and Control*(信息与控制),2006,35(3):378-381.
[7] CHAUCHARD F,COGDILL R,ROUSSEL S, *et al.* Application of LS-SVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes[J]. *Chemometr Intell Lab Syst*,2004,71:141-150.

(责任编辑:唐 静)