

# 基于激光诱导击穿光谱技术寻优定量 分析土壤中 Mn 元素

沙文<sup>1</sup>, 李江涛<sup>1</sup>, 鲁翠萍<sup>2\*</sup>

<sup>1</sup>安徽大学电气工程与自动化学院, 安徽 合肥 230061;

<sup>2</sup>合肥学院先进制造工程学院, 安徽 合肥 230601

**摘要** 激光诱导击穿光谱技术(LIBS)与支持向量机(SVM)相结合用于分析土壤中 Mn 元素含量。44 个土壤样品采集于安徽淮北市,采用 Kennard-Stone(K-S)方法将样品划分为训练集(34 个)和测试集(10 个),分别使用多元线性回归(MIR)、网格搜索法(GSM)、遗传算法(GA)、粒子群优化(PSO)和最小二乘法(LS)建立定量分析模型。结果表明:MIR、GSM 和 PSO 模型所得到的训练集相关系数  $R_{\text{trn}}^2$  只有 0.861、0.866 和 0.862,测试集相关系数  $R_{\text{t}}^2$  低于 0.9,相对误差大于 8.6%,误差较大;GA 模型的  $R_{\text{trn}}^2$  大于 0.93,  $R_{\text{t}}^2$  小于 0.9,训练时间较长,需减少训练时间和提高测试集相关性;LS 模型寻优效果较好,  $R_{\text{trn}}^2$  提高到 0.998,  $R_{\text{t}}^2$  提高到 0.967,相对误差小,训练时间同比大幅度缩短,相关性好,泛化能力强,更适合用于土壤中 Mn 元素的快速检测。

**关键词** 激光光学;支持向量机;激光诱导击穿光谱技术;土壤

中图分类号 O657.3

文献标志码 A

doi: 10.3788/CJL202047.0511001

## Quantitative Analysis of Mn in Soil Based on Laser-Induced Breakdown Spectroscopy Optimization

Sha Wen<sup>1</sup>, Li Jiangtao<sup>1</sup>, Lu Cuiping<sup>2\*</sup>

<sup>1</sup>School of Electrical Engineering and Automation, Anhui University, Hefei, Anhui 230061, China;

<sup>2</sup>School of Advanced Manufacturing Engineering, Hefei University, Hefei, Anhui 230601, China

**Abstract** This paper uses laser-induced breakdown spectroscopy and support vector machine to analyze the content of Mn in soil. Forty-four soil samples were collected in Huaibei, Anhui. The samples were divided into training set (34 samples) and test set (10 samples) using Kennard-Stone (K-S) method. Multiple linear regression (MIR), grid search method (GSM), genetic algorithm (GA), particle swarm optimization (PSO), and least squares method (LS) were used to establish quantitative analysis models. The results show that the correlation coefficients  $R_{\text{trn}}^2$  of the training set of the MIR, GSM, and PSO models are only 0.861, 0.866, and 0.862, respectively. The correlation coefficients  $R_{\text{t}}^2$  of the test set of corresponding models are lower than 0.9, the relative error is greater than 8.6%, and the error is larger. The  $R_{\text{trn}}^2$  of the GA model is greater than 0.93, and  $R_{\text{t}}^2$  is less than 0.9. The training time of the GA model is long, so the training time must be reduced, and the correlation of the test set must be improved. The LS model works well with  $R_{\text{trn}}^2$  0.998 and  $R_{\text{t}}^2$  0.967, and the relative error is small. The training time is greatly shortened year-on-year, correlation is good, and generalization ability is strong. The LS model is more suitable for the rapid detection of the Mn element in soil.

**Key words** laser optics; support vector machine; laser induced breakdown spectroscopy; soil

**OCIS codes** 140.3210; 140.3460; 140.3580

## 1 引言

锰元素是植物生长必需的微量元素,其供给直

接关系到植物生长发育,而土壤是植物中锰元素的主要来源。有研究表明,土壤中锰元素的氧化物和氢氧化物对植物的生长具有重要的意义<sup>[1]</sup>,如:在缺

收稿日期:2019-12-11;修回日期:2020-01-07;录用日期:2020-01-19

基金项目:国家自然科学基金(61505001)、中国科学院 STS 项目(KFJ-SW-STS-144)

\* E-mail: 632790346@qq.com

锰的土壤中施用锰肥后农作物增产较为明显,豆科植物单株荚数、果仁重量显著增加,空秕数大大减少<sup>[2]</sup>,故对土壤中锰元素进行快速准确检测可达到“精准施肥”效果。检测土壤元素多采用化学分析法:国家标准方法,电感耦合等离子原子发射光谱法(ICP-AES),原子荧光光谱法(AFS),X射线荧光光谱法(XRFS)<sup>[3]</sup>,近红外光谱法<sup>[4-5]</sup>等。这些方法因需对样本进行前期处理且需要专业人员操作,成本较高,耗时较长,误差较大,利用其很难实现对土壤中元素的快速检测。

激光诱导击穿光谱技术(LIBS)是一种较为理想的激光高温烧蚀光谱技术,因具有快速(毫秒量级)、原位(样品量少)、无损、安全、环保(无二次污染)等优点而被广泛应用于土壤和录井岩石分类,以及水体重金属和复合肥检测等各大领域<sup>[6-13]</sup>,区别于传统的化学检测方法,采用 LIBS 结合数据分析模型检测土壤元素含量近年来已成为该领域的研究热点。

支持向量机(SVM)是一种基于统计学习理论的机器学习方法,利用间隔最大化进行建模训练,并将原本空间中难以处理的问题映射到高维空间,寻求间隔最大化用于计算最优线性超平面。张成龙等<sup>[14]</sup>基于 SVM 定量分析鄂尔多斯盆地中 C 元素的含量,成功预测出局部富有机质层。邹孝恒等<sup>[15]</sup>基于 LIBS 分别采用遗传算法(GA)和偏最小二乘法(PLS)定量分析模型,以土壤中 Mn 元素为例,结果显示 PLS 模型误差更小。大量研究表明,可通过建立 SVM 模型对土壤中元素进行检测,但尚未有统一的 SVM 参数寻优方法,较为常用的寻优方法有网格搜索法-SVM(GSM-SVM)、遗传算法-SVM(GA-SVM)、粒子群优化-SVM(PSO-SVM)和最小二乘法-SVM(LS-SVM),本文分别采用多元线性回归和以上 4 种寻优方法对土壤中 Mn 元素含量进行定量分析。

## 2 实 验

### 2.1 实验装置

土壤 LIBS 光谱探测实验装置如图 1 所示。所用实验仪器如下:激发光源采用调 Q 开关 Nd:YAG 脉冲激光器(ICE450, BigSky, 1064 nm, 6 ns, 1 Hz),单脉冲能量为 100 mJ;光纤光谱仪(AvaSpec-ULS2048-USB2, Avantes, 荷兰),此光谱仪包含四通道,分辨率为 0.1 nm,覆盖波长为 190~510 nm 和 690~890 nm;样品台;激光聚焦装置;氩

气提供装置;光谱收集装置和数据处理系统。激光聚焦装置主要包括聚焦镜(直径 2.54 cm, 焦距 4.5 cm)和反射镜片,样品台包括旋转电机和载物台,光谱收集装置包括透镜(焦距 3.5 cm)和光纤,数据处理系统为计算机。脉冲激光器通过调 Q 开关发出沿水平方向的高能激光,并聚焦在样本表面,通过数字延迟器控制延迟时间(1.28  $\mu$ s)和积分时间(1.05 ms, 光谱仪最小积分时间)<sup>[16]</sup>,高温等离子体发出的特征光谱信号经光纤探头送至光谱仪进行分光检测,基于原子发射光谱展宽机制,Plasus Specline 软件利用洛伦兹线型的轮廓函数对光谱进行拟合<sup>[17]</sup>。为降低样品测量的不稳定性,将样品放置在步进电机旋转台上,手动调节步进电机,将 20 个激光脉冲作为一次测量,并进行 10 次重复测量。将采集于安徽淮北地区的 44 种土壤样品粉碎过 100 目细筛,每个样品取 3 g,利用直径为 25 mm 的模具在 8 MPa 压强下将样品压制为厚度大约 2 mm 的圆饼。因空气中灰尘和样品内部成分不均会导致测量结果误差较大,故选择在氩气环境下进行实验。

### 2.2 元素特征谱线选取

由于土壤体系庞大、结构复杂,大量基础元素发射谱线间的相互干扰较为严重,在选择 Mn 元素特征谱线时应尽可能挑选谱线较为清晰、未饱和且信噪比较高的特征谱线。图 2 为 1 号土壤样本在 257~408 nm 波段的激光诱导光谱图,因土壤中的 Mn 元素多与 Fe 元素以氧化物结合态存在,故光谱图中 Fe 元素谱线丰富。根据美国国家标准与技术研究院谱线库, Mn 元素特征谱线中强度较大的波长有 222.1, 257.6, 279.4, 280.1, 403.1, 403.3, 403.4 nm 等,图 2 中 257.6, 403.1, 403.3 nm 处的谱线较清晰,满足特征谱线标准,故选择此三条作为 Mn 元素特征谱线进行分析,并选取 Fe 元素的 404.6 nm 和 407.6 nm 特征谱线作为相关特征谱线进行分析。Mn 元素浓度值采用原子吸收光谱法(AAS)测得。使用 Kennard-Stone(K-S)将 44 种土壤样本划分为两种:训练集 34 个,测试集 10 个。土壤中 Mn 元素质量分数的统计数据如表 1 所示。

### 2.3 定量模型建立

对处于局部热平衡且不考虑自吸收效应的激发态离子,其 LIBS 强度可表示为

$$I_{k,i} = FC_s \frac{A_{k,i} g_k}{U_s(T)} \exp(-E_k/k_B T), \quad (1)$$

令

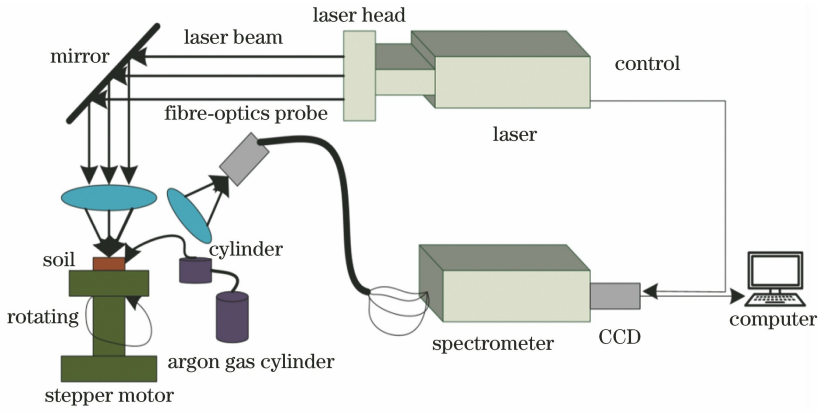


图 1 LIBS 系统示意图

Fig. 1 Schematic of LIBS system

表 1 Mn 元素实际质量分数

Table 1 Actual mass fraction of Mn element

Training set number	Actual mass fraction / ( $10^{-4} \%$ )	Training set number	Actual mass fraction / ( $10^{-4} \%$ )	Training set number	Actual mass fraction / ( $10^{-4} \%$ )	Test set number	Actual mass fraction / ( $10^{-4} \%$ )
1	1134.6	13	1157.5	25	1229.2	1	533.4
2	432.3	14	622.3	26	788.8	2	540.6
3	468.9	15	1220.5	27	769.2	3	581.1
4	997.8	16	1036.9	28	1333.1	4	613.6
5	905.8	17	498.8	29	728.2	5	622.2
6	650.6	18	510.9	30	523.6	6	666.4
7	1184.7	19	906.9	31	619.5	7	717.9
8	1470.3	20	499.0	32	633.1	8	751.4
9	744.2	21	643.9	33	789.0	9	1025.0
10	703.5	22	973.0	34	827.5	10	1258.7
11	715.9	23	730.5				
12	631.1	24	888.1				

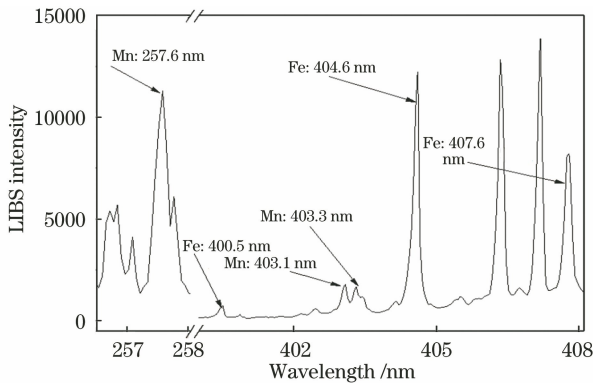


图 2 1 号土壤样品 Mn 元素光谱图

Fig. 2 Mn element spectrum of No.1 soil sample

$$A = \frac{U_s(T)}{FA_{k,i}g_k \exp(-E_k/k_B T)}, \quad (2)$$

式中： $k, i$  分别为跃迁谱线的上下能级； $F$  为与光收集装置效率有关，与波长无关，在同次实验测量中保持不变的实验参数； $C_s$  为特征谱线对应的原子或离

子的浓度； $A_{k,i}$  为特征谱线的跃迁概率； $g_k$  为上能级的简并度； $U_s(T)$  为发射元素  $s$  的配分函数； $E_k$  为跃迁谱线的上能级能量； $k_B$  为波尔兹曼常数； $T$  为温度，当等离子体满足局部热平衡时， $T$  为常数。将 (2) 式代入 (1) 式得

$$C_s = AI_{k,i}. \quad (3)$$

根据 (3) 式可计算出样品中待测元素的浓度。但受到基体效应等的影响， $A$  在实验中很难确定，则  $C_s$  和  $I_{k,i}$  的关系可表达为

$$C_s = aI_{k,i}^h, \quad (4)$$

式中： $a$  为比例系数； $h$  为自吸收系数，取值范围为  $0 \sim 1$ 。当元素浓度较大时，形成的厚等离子体发生自吸收现象，薄等离子体无自吸收现象。当等离子体满足局部热平衡且为薄等离子体时，无需考虑自吸收现象，此时  $h$  取值为 1<sup>[18-20]</sup>，令  $I = I_{k,i}^h$ ，得

$$C_s = \left\{ \sum_{i \in v} \partial_i k_{\text{libs}}(I_i, I) \right\} + b, \quad (5)$$

式中: $v$  为支持向量集; $\partial_i$  为拉格朗日乘子; $k_{\text{lib}}(I_i, I)$  为核函数; $b$  为常数。通过元素浓度和谐线强度之间的关系[(3)式和(4)式]得到激光诱导击穿光谱混合核函数为

$$k_{\text{lib}}(I_i, I) = cII_i + (1 - c) \exp\left(\frac{-\|I - I_i\|^2}{2g^2}\right), \quad (6)$$

式中: $c$  为惩罚系数; $g$  为核参数。混合核函数由两部分组成,前者为线性核函数  $cII_i$ , 后者为径向核函数(RBF)。SVM 核函数常用来解决数据的非线性映射问题,大量的实验和数据表明,RBF 具有较高的拟合和预测精度,常被选作核函数进行研究。因 SVM 中参数的调整很大程度上决定着回归效果,且所选择的 RBF 为内核函数,所以在土壤元素定量分析回归模型中主要对惩罚系数  $c$  和核参数  $g$  进行寻优。

受基体效应影响,采用单变量进行定量分析可能会降低测量准确性,因此采用多元素光谱信息进行定量分析,降低基体效应、提高准确性。多元线性回归(MIR)考虑因变量与自变量间的线性关系,记因变量为  $Y$ , 自变量为  $X_1, X_2, X_3, \dots, X_n$ , 则多元线性回归模型可表示为

$$Y = a_0 + a_1X_1 + \dots + a_nX_n + \epsilon, \quad (7)$$

式中: $\epsilon$  为随机误差,其期望值  $E(\epsilon) = 0$ 。经过  $p$  次独立测量获得  $Y$  样本 ( $Y_d, X_{d1}, X_{d2}, X_{d3}, \dots,$

$X_{dm}$ ), 则多次测量模型为

$$Y_d = a_0 + a_1X_{d1} + \dots + a_mX_{dm} + \epsilon_d, d = 1, 2, \dots, p. \quad (8)$$

则待测元素浓度的多元线性回归模型为

$$C_d = a_{0,d} + \sum_{j=1}^n a_{j,d}Q_j + \epsilon_d, \quad (9)$$

式中: $C_d$  为待测元素的浓度,个数为  $d$ ;  $Q_j$  为所选取的元素特征谱线强度,个数为  $n$ ;  $a_{j,d}$  为多元线性回归方程系数; $\epsilon_d$  为服从正态分布的随机误差。

土壤中 Mn 元素多与 Fe 元素以氧化物结合态存在,选取 Mn I: 257.6 nm、403.1 nm、403.3 nm 和 Fe I: 404.6 nm、407.6 nm 谱线强度为自变量,根据(9)式建立二元线性回归模型,采用(10)式计算 Mn 元素浓度:

$$C_{\text{Mn}} = a_0 + a_1Q_{\text{Mn}} + a_2Q_{\text{Fe}}. \quad (10)$$

### 3 定量分析结果

#### 3.1 多元线性回归

根据(10)式,采用 44 种土壤样品建立土壤中 Mn 元素多元线性回归模型,图 3 为土壤中 Mn 元素训练集和测试集的线性回归模型,横坐标为 Mn 元素实际浓度,纵坐标为预测浓度值。从图 3 可看出,训练集相关系数仅为 0.861,测试集相关系数为 0.841,相关性不强,耗时较长,达不到对土壤中 Mn 元素快速准确检测的要求。

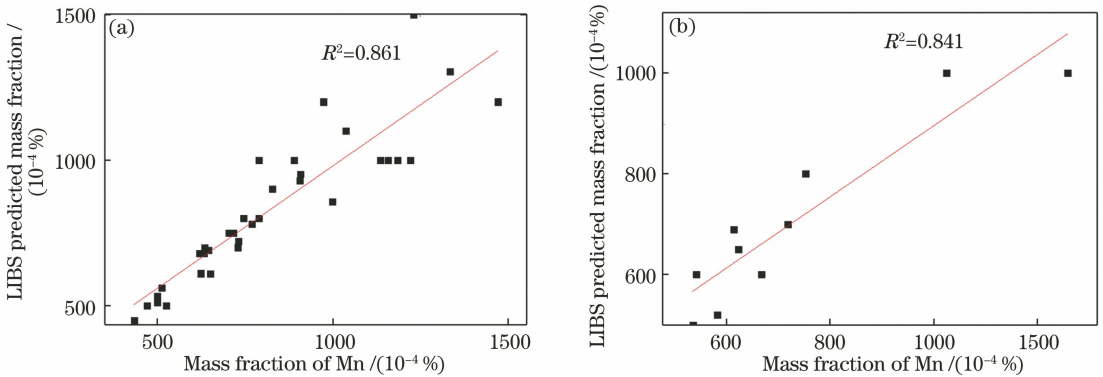


图 3 考虑 Fe 元素影响时 Mn 元素的定量模型。(a)训练集;(b)测试集

Fig. 3 Quantitative models of Mn element considering the influence of Fe element. (a) Training set; (b) test set

#### 3.2 网格搜索法

刘祥楼等<sup>[21]</sup>提出一种网格搜索法,在一定的空间范围中将待搜索的  $c$  (惩罚系数) 和  $g$  (核参数) 划分成均等网格,每一结点即代表一组参数。有研究表明,参数较多会大大增加寻优计算量,当网格足够小、搜索空间足够大时,会找到最优参数,但耗时往往较长且易出现过度学习。图 4 所示为基于 GSM-

SVM 寻优方法得到的 Mn 元素训练集和测试集的寻优效果图。表 2 为 GSM-SVM 寻优方法的结果(表中  $R_e$  为相对误差),寻优时间为 1.8313 s,训练集相关系数( $R_{\text{tra}}^2$ )为 0.866,测试集相关系数( $R_{\text{t}}^2$ )为 0.897,模型效果较差、泛化能力较弱,相关系数过低且均方根误差较大( $9.6459 \times 10^{-3}$ ),需要进一步提高预测浓度与实际浓度的相关性。



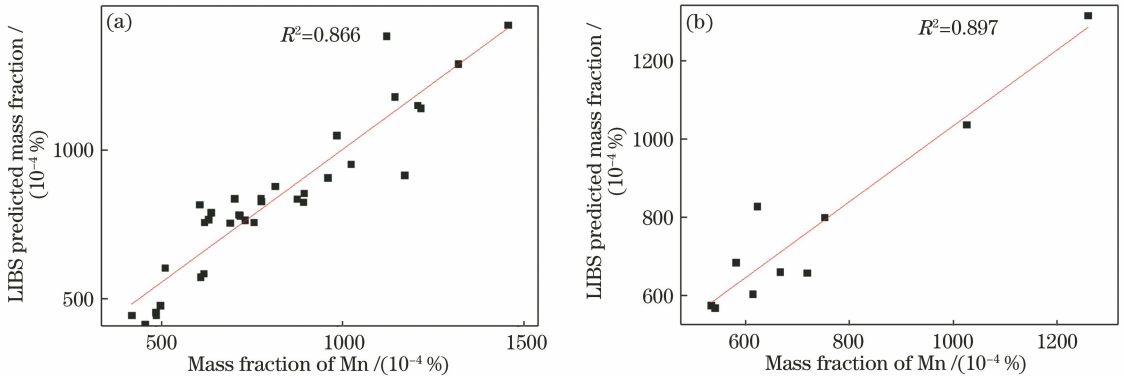


图 4 GSM-SVM 方法的效果图。(a)训练集;(b)测试集

Fig. 4 Renderings of GSM-SVM method. (a) Training set; (b) test set

表 2 GSM-SVM 寻优方法的结果

Table 2 Results of GSM-SVM optimization method

Parameter	$t/s$	$R_{tra}^2$	$R_t^2$	Average $R_e/\%$
Result	1.8313	0.866	0.897	8.7

### 3.3 遗传算法

Chen 等<sup>[22]</sup>提出使用一种基于达尔文生物进化过程的遗传算法,模拟自然界中适者生存的进化现象,将问题解决空间映射到遗传空间,将解决方案编

码到染色体中,利用遗传(选择、交叉和变异)方法计算每个染色体的适应度值,选择最佳染色体,得到最优解。图 5 为基于 GA-SVM 寻优模型得到的 Mn 元素训练集和测试集的寻优效果图,表 3 所示为对应寻优结果。由表 2 和表 3 可知,较 GSM-SVM 寻优,GA-SVM 相关性  $R_{tra}^2$  有所增加, $R_t^2$  有所减少,仅为 0.894,相对误差较大超过实际应用可接受范围,训练寻优耗时较长,为 3.8475 s,不适合用于土壤中 Mn 元素的定量分析。

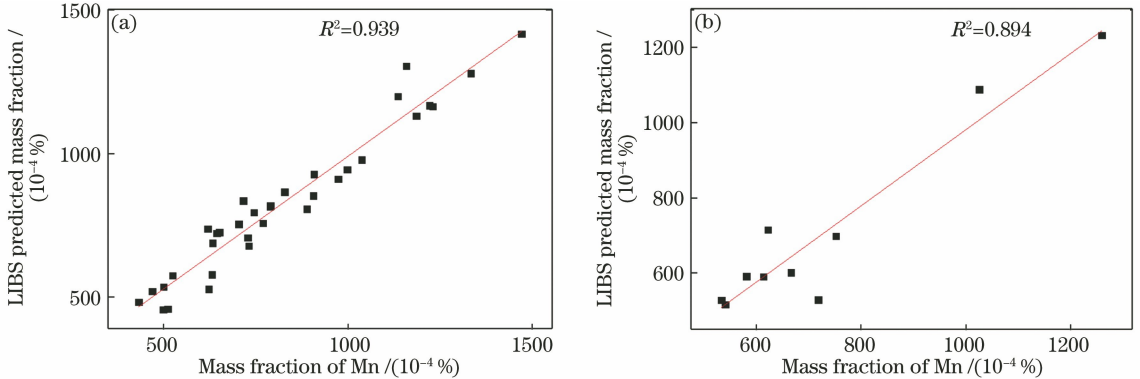


图 5 GA-SVM 方法的效果图。(a)训练集;(b)测试集

Fig. 5 Renderings of GA-SVM method. (a) Training set; (b) test set

表 3 基于 GA-SVM 寻优方法的结果

Table 3 Results of GA-SVM optimization method

Parameter	$t/s$	$R_{tra}^2$	$R_t^2$	Average $R_e/\%$
Result	3.8475	0.939	0.894	7.7

### 3.4 粒子群算法

由 Kennedy 和 Eberhart 等提出的 PSO 算法源于对鸟群捕食行为的研究<sup>[23]</sup>,将鸟群中每只鸟比作 PSO 算法中的一个粒子,即每个粒子类似寻找食物的鸟,鸟群通过向离食物最近的那一只鸟靠拢并搜索它周围区域,然后通过每只鸟之间的信息共享寻找到最优解。图 6 所示为基于 PSO-SVM 寻优模型

得到的训练集和测试集的寻优效果图,表 4 为对应的具体数据。经分析,PSO-SVM 的训练集相关系数在以上 4 种寻优方法中最小( $R_{tra}^2=0.862$ ),测试集相关系数较低,均方根误差最大,为  $9.8087 \times 10^{-3}$ 。训练时间较长,训练集和测试集相关性不强,模型的泛化能力弱和相对误差较大等导致将此模型用于土壤中 Mn 元素的定量分析时效果较差。

表 4 PSO-SVM 寻优方法的结果

Table 4 Results of PSO-SVM optimization method

Parameter	$t/s$	$R_{tra}^2$	$R_t^2$	Average $R_e/\%$
Result	3.5037	0.862	0.895	8.8

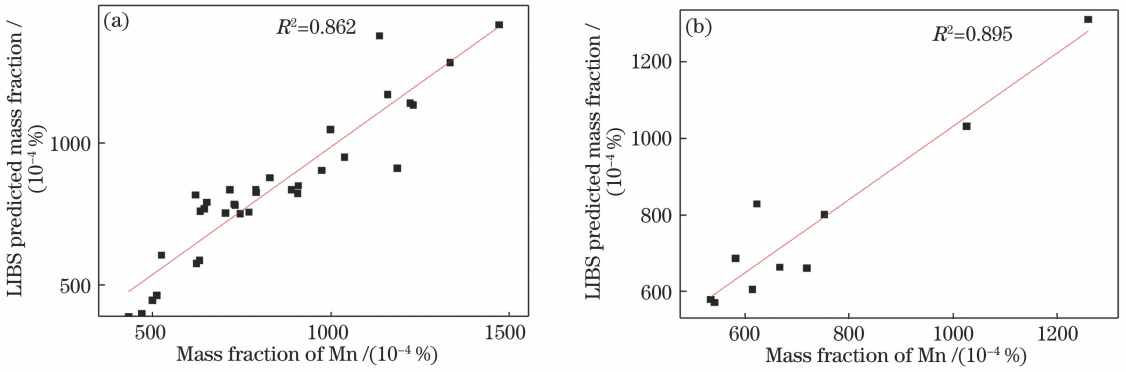


图 6 PSO-SVM 方法的效果图。(a)训练集;(b)测试集

Fig. 6 Renderings of PSO-SVM method. (a) Training set; (b) test set

### 3.5 最小二乘法

Castro-Garcia 等<sup>[24]</sup>通过在标准 SVM 目标函数中增加误差平方和项,提出了 LS-SVM 方法,该方法的一般计算形式如下:目标函数 =  $\sum (\text{观测值} - \text{理论值})^2$ ,其中观测值为样本值,理论值为假设拟合函数,目标函数取最小值时得到拟合函数模型。设函数  $H(x_1, x_2, \dots, x_n) = \theta_0 + \theta_1 x_1 + \dots + \theta_{n-1} x_{n-1}$  可表示为  $H(\mathbf{X}) = \mathbf{X}\boldsymbol{\theta}$ , 损失函数  $J(\boldsymbol{\theta}) = \frac{1}{2}(\mathbf{X}\boldsymbol{\theta} - \mathbf{R})^T(\mathbf{X}\boldsymbol{\theta} - \mathbf{R})$ , 两边求导取 0 得:  $\frac{\partial}{\partial \mathbf{x}}(\mathbf{X}^T \mathbf{X}) = 2\mathbf{X}; \nabla_{\mathbf{x}} f(\mathbf{A}\mathbf{X} + \mathbf{B}) = \mathbf{A}^T \nabla_{\mathbf{R}} f, \mathbf{R} = \mathbf{A}\mathbf{X} + \mathbf{B}, f(\mathbf{R})$  为标量,整理后得:  $\boldsymbol{\theta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{R}$ , 其中:  $\mathbf{X}$  为

$z \times u$  维矩阵,  $z$  代表样本个数,  $u$  代表样本的特征数。

图 7 所示为 LS-SVM 寻优模型得到的寻优效果图,表 5 为对应寻优结果。LS-SVM 模型对应的训练集 ( $R_{\text{tra}}^2 = 0.998$ ) 和测试集 ( $R_t^2 = 0.967$ ) 相关系数相较于 GSM-SVM ( $R_{\text{tra}}^2 = 0.866, R_t^2 = 0.897$ )、GA-SVM ( $R_{\text{tra}}^2 = 0.939, R_t^2 = 0.894$ )、PSO-SVM ( $R_{\text{tra}}^2 = 0.862, R_t^2 = 0.895$ ) 都得到很大的提高。训练集和测试集的相关系数高,此模型具有较强的泛化能力;测试集相关系数大于 0.96 表明预测结果较好;且训练时间短,误差较小。LS-SVM 寻优模型更适合土壤中 Mn 元素的定量分析。

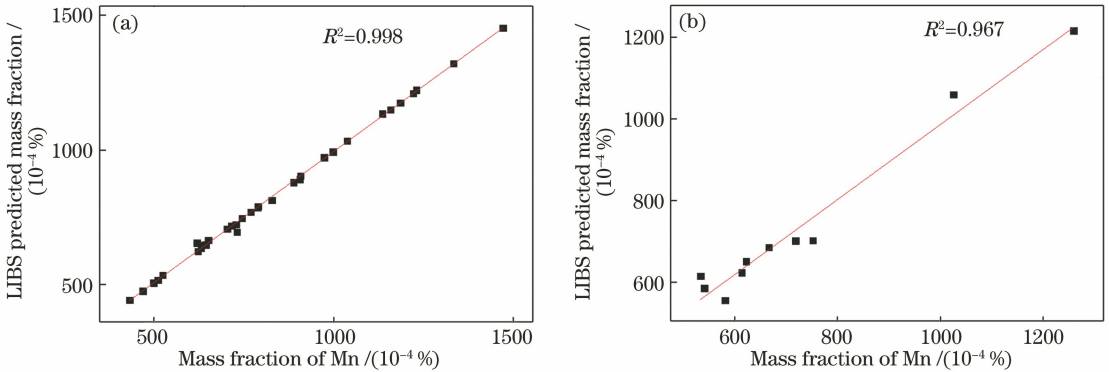


图 7 LS-SVM 方法的效果图。(a)训练集;(b)测试集

Fig. 7 Renderings of LS-SVM method. (a) Training set; (b) test set

表 5 LS-SVM 寻优方法的结果

Table 5 Results of LS-SVM optimization method

Parameter	$t/s$	$R_{\text{tra}}^2$	$R_t^2$	Average $R_e/\%$
Result	0.0182	0.998	0.967	5.4

## 4 实验讨论

为了表示模型预测值  $\gamma$  与实际值  $\mu$  之间的差值,衡量模型预测结果的准确性,在实验中引入绝对

误差  $A_e$  和相对误差  $R_e$  来表示二者离散程度,具体表达式为

$$A_e = |\gamma - \mu|, \quad (11)$$

$$R_e = \left| \frac{\gamma - \mu}{\mu} \right| \times 100\%. \quad (12)$$

图 8 为 Mn 元素训练集和测试集的绝对误差图,分析发现:基于多元线性回归和三种寻优方法 (GSM-SVM、GA-SVM、PSO-SVM) 得到的训练集和测试集的绝对误差最大值均大于  $150 \times 10^{-4} \%$ ,

最小绝对误差大于  $1.0 \times 10^{-3} \%$ ，且误差走势波动较大，预测效果不稳定；LS-SVM 寻优方法的训练集和测试集绝对误差均小于  $2.0 \times 10^{-3} \%$ ，均方根误差较前三种模型减小，曲线基本平滑，离散波动程度较小，寻优拟合精度较高，取得了较好的定量分析效果。表 6 所示为基于多元线性回归和 4 种寻优方式得到的

Mn 元素测试集质量分数的相对误差，分析发现：PSO-SVM 测试集相对误差最大为 33.4%，最小仅为 0.2%，但总体平均相对误差较大，达到 8.8%；LS-SVM 模型平均相对误差较小为 5.4%，相关性、训练时间、绝对误差和均方根误差都比前 4 种寻优方法效果更好，更适合土壤中 Mn 元素的定量分析。

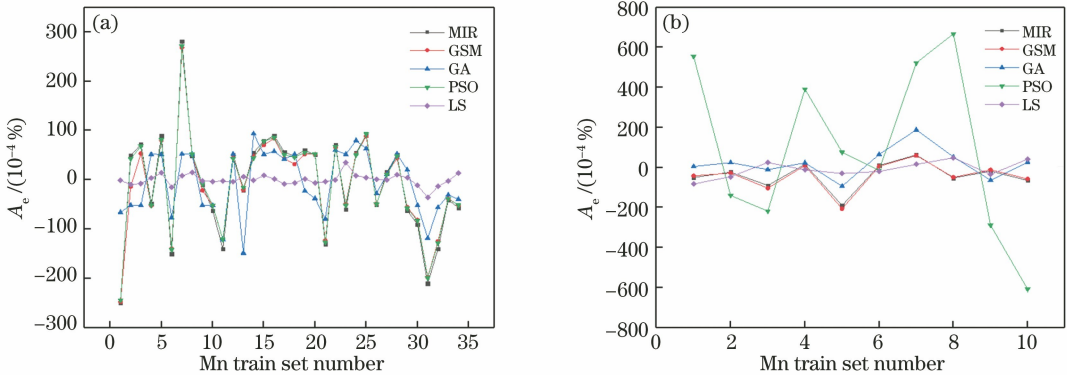


图 8 Mn 元素的绝对误差图。(a)训练集;(b)测试集

Fig. 8 Absolute error diagrams of Mn element. (a) Training set; (b) test set

表 6 Mn 元素测试集质量分数相对误差

Table 6 Relative error of mass fraction of Mn element test set

No.		1	2	3	4	5	6	7	8	9	10	Average
Actual /		533.6	540.6	518.1	613.6	622.2	666.2	718.0	751.4	1025.0	1258.7	724.7
( $10^{-4} \%$ )												
Predicted /	MIR	583.7	564.1	671.4	598.0	812.5	656.9	655.5	803.7	1043.6	1323.9	771.3
	GSM	575.5	568.6	685.0	604.6	828.5	661.0	658.6	800.7	1037.1	1316.3	773.6
	GA	529.1	516.8	592.2	590.9	715.8	601.7	529.9	699.0	1088.9	1232.3	709.7
	PSO	580.4	572.9	687.8	606.9	830.2	664.7	662.1	802.9	1033.2	1311.9	775.3
	LS	616.6	587.0	556.7	625.3	652.6	686.6	702.7	703.4	1060.6	1316.8	750.8
$R_e / \%$	MIR	9.4	4.3	15.5	2.5	30.6	1.4	8.7	7.0	1.8	5.2	8.6
	GSM	7.9	5.2	17.9	1.5	33.2	0.8	8.3	6.6	1.2	4.6	8.7
	GA	0.8	4.4	1.9	3.7	15.0	9.7	26.2	7.0	6.2	2.1	7.7
	PSO	8.8	6.0	18.4	1.1	33.4	0.2	7.8	6.9	0.8	4.2	8.8
	LS	15.6	8.6	4.2	1.9	4.9	3.1	2.1	6.4	3.5	3.3	5.4

## 5 结 论

本文对 44 种土壤样本进行了 Mn 元素定量分析，将土壤样本进行简单预处理，通过 K-S 法将样本划分 34 个训练集和 10 个测试集。分别采用多元线性回归、GSM-SVM、GA-SVM、PSO-SVM 和 LS-SVM 共 5 种寻优方法对 Mn 元素进行定量分析对比研究，结果表明：前 4 种寻优方式得到的训练集相关系数小于 0.93，测试集相关系数小于 0.90，4 者泛化能力较弱，模型准确度较低。GSM-SVM 寻优方法的时间较短且相对误差较大，训练时间最长，测试结果较差，且相对误差超过接受范围；PSO-SVM 模型寻优效果最差，泛化能力最弱，相对误差最大，不

适合土壤中 Mn 元素定量分析。LS-SVM 模型取得较好寻优结果，训练集相关系数大于 0.99，测试集相关系数大于 0.96（泛化能力强、误差小），相对误差较小，训练时间较短，更适用于土壤中 Mn 元素快速准确定量分析。

## 参 考 文 献

[1] Gao F, Jia J Y, Mei X Q, et al. The environmental meaning of trace element manganese and its chemical compound in soil[J]. Trace Elements Science, 2004, 11(4): 10-13.  
高凡, 贾建业, 梅雪清, 等. 土壤中微量元素锰及其化合物的环境意义[J]. 广东微量元素科学, 2004, 11(4): 10-13.

- [2] Zeng Z H. The relationship between yield of crop and environment of soil [J]. *Hunan Geology*, 2000, 19(1): 25-29.  
曾昭华. 作物产量与土壤环境的关系[J]. *湖南地质*, 2000, 19(1): 25-29.
- [3] Ma F Y, Hu L, Zhao L. Determination of trace germanium in soils by hydride generation-atomic fluorescence spectrometry after alkali fusion [J]. *XianDai NongYe KeJi*, 2019(2): 134-135.  
马凤英, 胡兰, 赵玲. 碱熔-氢化物发生原子荧光光谱法测定土壤中的痕量锗[J]. *现代农业科技*, 2019(2): 134-135.
- [4] Wang L S, Lu C P, Wang R J, et al. Optimization for vis/NIRS prediction model of soil available nitrogen content [J]. *Chinese Journal of Luminescence*, 2018, 39(7): 1016-1023.  
汪六三, 鲁翠萍, 王儒敬, 等. 土壤碱解氮含量可见/近红外光谱预测模型优化[J]. *发光学报*, 2018, 39(7): 1016-1023.
- [5] Wang R J, Chen T, Wang Y B, et al. Soil near-infrared spectroscopy prediction model based on deep sparse learning [J]. *Chinese Journal of Luminescence*, 2017, 38(1): 109-116.  
王儒敬, 陈天娇, 汪玉冰, 等. 基于深度稀疏学习的土壤近红外光谱分析预测模型[J]. *发光学报*, 2017, 38(1): 109-116.
- [6] Fang L, Zhao N J, Ma M J, et al. Quantitative analysis of Pb, Cd, Cr and Cu in soil using standard addition method combined with laser-induced breakdown spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2017, 37(10): 3274-3280.  
方丽, 赵南京, 马明俊, 等. 激光诱导击穿光谱结合标准加入法定量分析土壤中铅、镉、铬和铜[J]. *光谱学与光谱分析*, 2017, 37(10): 3274-3280.
- [7] She M J, Fu H B, Jia J W, et al. Determination of terrestrial and marine sedimentary rocks by laser-induced breakdown spectroscopy[J]. *Acta Photonica Sinica*, 2018, 47(8): 0847009.  
余明军, 付洪波, 贾军伟, 等. 激光诱导击穿光谱判定岩石陆相和海相沉积[J]. *光子学报*, 2018, 47(8): 0847009.
- [8] Li Y, Zhai K H, Li Y L, et al. Detection and analysis of heavy metal content in urban soil based on laser induced breakdown spectroscopy [J]. *Natural Science Journal of Xiangtan University*, 2018, 40(3): 86-88, 128.  
李艳, 翟开华, 李艳丽, 等. 基于激光诱导击穿光谱技术的城市土壤重金属含量检测与分析[J]. *湘潭大学自然科学学报*, 2018, 40(3): 86-88, 128.
- [9] Sha W, Li J T, Xiao W B, et al. Quantitative analysis of elements in fertilizer using laser-induced breakdown spectroscopy coupled with support vector regression model[J]. *Sensors*, 2019, 19(15): 3277.
- [10] Andrade D F, Sperança M A, Pereira-Filho E R. Different sample preparation methods for the analysis of suspension fertilizers combining LIBS and liquid-to-solid matrix conversion: determination of essential and toxic elements[J]. *Analytical Methods*, 2017, 9(35): 5156-5164.
- [11] Nicolodelli G, Senesi G S, de Oliveira Perazzoli I L, et al. Double pulse laser induced breakdown spectroscopy: a potential tool for the analysis of contaminants and macro/micronutrients in organic mineral fertilizers [J]. *Science of the Total Environment*, 2016, 565: 1116-1123.
- [12] Li W P, Zhou W D. Comparative study of underwater single pulse and orthogonal double pulse laser-induced breakdown spectroscopy of barium element in solution [J]. *Chinese Journal of Lasers*, 2019, 46(9): 0911003.  
李文平, 周卫东. 溶液中 Ba 元素的水下单脉冲与正交双脉冲 LIBS 的比较研究[J]. *中国激光*, 2019, 46(9): 0911003.
- [13] Liu S M, Xiu J S, Liu Y Y. Rapid quantitative analysis of element content ratios in Cu(In, Ga)Se<sub>2</sub> thin films using laser-induced breakdown spectroscopy[J]. *Chinese Journal of Lasers*, 2019, 46(9): 0911001.  
刘世明, 修俊山, 刘云燕. 基于激光诱导击穿光谱技术的铜铟镓硒薄膜中元素含量比的快速定量分析方法[J]. *中国激光*, 2019, 46(9): 0911001.
- [14] Zhang C L, Tao S Z, Bai B, et al. Source rock TOC content prediction based on the support vector machine model: an application in Ordos Basin [J]. *Natural Gas Geoscience*, 2019, 30(5): 761-768.  
张成龙, 陶士振, 白斌, 等. 基于支持向量机模型的烃源岩有机碳含量预测: 以鄂尔多斯盆地为例[J]. *天然气地球科学*, 2019, 30(5): 761-768.
- [15] Zou X H, Hao Z Q, Yi R X, et al. Quantitative analysis of soil by laser-induced breakdown spectroscopy using genetic algorithm-partial least squares[J]. *Chinese Journal of Analytical Chemistry*, 2015, 43(2): 181-186.  
邹孝恒, 郝中骐, 易荣兴, 等. 基于遗传算法和偏最小二乘法的土壤激光诱导击穿光谱定量分析研究[J]. *分析化学*, 2015, 43(2): 181-186.
- [16] Sha W, Niu P G, Zhen C H, et al. Analysis of phosphorus in fertilizer using laser-induced breakdown spectroscopy [J]. *Journal of Applied Spectroscopy*, 2018, 85(4): 653-658.
- [17] Deng C F, Liao B, Zhang X, et al. Single peak fitting based on least squares during emission spectral



- diagnosis of nickel plasma produced by MEVVA ion source [J]. *Journal of Beijing Normal University (Natural Science)*, 2012, 48(3): 257-262.
- 邓春风, 廖斌, 张旭, 等. 基于最小二乘的 MEVVA 源产生的金属镍等离子发射谱线的单峰拟合 [J]. *北京师范大学学报(自然科学版)*, 2012, 48(3): 257-262.
- [18] Gao X, Shao Y, Du C, et al. Pre-ablation laser parameters effect on the spectral enhancement of double pulsed laser induced breakdown spectroscopy [J]. *Chinese Journal of Lasers*, 2013, 40(8): 0815003.
- 高勋, 邵妍, 杜闯, 等. 预烧蚀激光参数对双脉冲激光诱导击穿光谱增强的影响 [J]. *中国激光*, 2013, 40(8): 0815003.
- [19] Rusak D A, Clara M, Austin E E, et al. Investigation of the effect of target water content on a laser-induced plasma [J]. *Applied Spectroscopy*, 1997, 51(11): 1628-1631.
- [20] Lu C P, Liu W Q, Zhao N J, et al. Influence of humidity on characteristic of laser-induced soil plasmas [J]. *Spectroscopy and Spectral Analysis*, 2010, 30(11): 2885-2888.
- 鲁翠萍, 刘文清, 赵南京, 等. 湿度对激光诱导土壤等离子体特性的影响 [J]. *光谱学与光谱分析*, 2010, 30(11): 2885-2888.
- [21] Liu X L, Jia D X, Li H, et al. Research on kernel parameter optimization of support vector machine in speaker recognition [J]. *Science Technology and Engineering*, 2010, 10(7): 1669-1673.
- 刘祥楼, 贾东旭, 李辉, 等. 说话人识别中支持向量机核函数参数优化研究 [J]. *科学技术与工程*, 2010, 10(7): 1669-1673.
- [22] Chen P W, Wang J Y, Lee H M. Model selection of SVMs using GA approach [C] // 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), July 25-29, 2004, Budapest, Hungary. New York: IEEE, 2004: 2035-2040.
- [23] Lu J H. Research on the particle swarm optimization for solving the problem of cotton-blending [D]. Hangzhou: Zhejiang University, 2011.
- 陆俊虎. 求解自动配棉问题的粒子群算法研究 [D]. 杭州: 浙江大学, 2011.
- [24] Castro-Garcia R, Agudelo O M, Tiels K, et al. Hammerstein system identification using LS-SVM and steady state time response [C] // European Control Conference 2016 (ECC'16), June 29-July 1, 2016, Aalborg, Denmark. New York: IEEE, 2016: 1063-1068.