



(12)发明专利申请

(10)申请公布号 CN 107064054 A

(43)申请公布日 2017. 08. 18

(21)申请号 201710113438.4

(22)申请日 2017.02.28

(71)申请人 浙江大学

地址 310058 浙江省杭州市西湖区余杭塘路388号

(72)发明人 卢建刚 蒋昊

(74)专利代理机构 浙江杭州金通专利事务有限公司 33100

代理人 刘晓春

(51)Int.Cl.

G01N 21/359(2014.01)

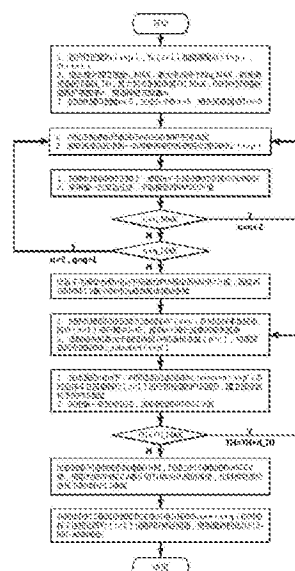
权利要求书2页 说明书6页 附图4页

(54)发明名称

一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法

(57)摘要

本发明属于光谱分析领域,尤其是涉及一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法。所述方法结合了相关系数法(CC)、偏最小二乘法(PLS)与径向基函数神经网络(RBFNN),并对重要参数进行了优化;包括:利用三阶Savitzky-Golay卷积平滑滤波及一阶导数校正对原始光谱进行预处理;在全波长段建立PLS模型并优化选取窗口宽度和PLS提取的主成分个数;计算每个波长变量的相关系数,截取相关系数大于设定阈值的波长变量参与建模并优化选取阈值大小;利用优化选取的窗口宽度、主成分个数和波长变量得到优化PLS模型;用优化PLS模型提取的主成分得分和性质矩阵训练RBF神经网络,获得最终的CC-PLS-RBFNN优化模型。本方法能显著提高近红外光谱分析的鲁棒性与精度。



1. 一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,其特征在于,所述基于CC-PLS-RBFNN优化模型的近红外光谱分析方法包括:利用三阶Savitzky-Golay卷积平滑滤波及一阶导数校正对原始光谱进行预处理;在全波长段建立PLS模型并优化选取平滑滤波的窗口宽度和PLS模型中提取的主成分个数;计算每个波长变量的相关系数,截取相关系数大于设定阈值的波长变量参与建模并优化选取阈值大小;利用优化选取的窗口宽度、主成分个数和波长变量得到优化PLS模型;用优化PLS模型提取的主成分得分矩阵和对应的样本性质矩阵训练RBF神经网络,获得最终的CC-PLS-RBFNN优化模型;具体步骤包括:

步骤(1):参数初始化:将样本划分为校正集和预测集;校正集的光谱矩阵为 $X_s (s \times p)$,具有 s 个样本,全波长点有 p 个,校正集的样本性质矩阵为 $Y_s (s \times 1)$;预测集的光谱矩阵为 $X_t (t \times p)$,具有 t 个样本,全波长点有 p 个,预测集的样本性质矩阵为 $Y_t (t \times 1)$;定义 n 为平滑滤波的窗口宽度, q 为主成分个数, TH 为相关系数阈值;设定最大窗口宽度 n_MAX ,最大主成分个数 q_MAX ,相关系数阈值遍历间隔 d_TH ,最大相关系数阈值 TH_MAX ,以及设定RBF神经网络核函数扩展因子 σ ,网络神经元数量 m ;窗口宽度 n 初始化为5,主成分个数 q 初始化为5,相关系数阈值 TH 初始化为0;

步骤(2):光谱预处理:对校正集原始光谱矩阵 $X_s (s \times p)$ 逐行进行 n 点三阶Savitzky-Golay卷积平滑滤波,并进行一阶求导完成光谱预处理,得到校正集预处理后光谱矩阵 $X_s' (s \times p)$;

步骤(3):PLS初始建模:对预处理后的 s 个校正集的样本利用PLS模型提取 q 个主成分,对预处理后光谱矩阵 $X_s' (s \times p)$ 及对应的样本性质矩阵 $Y_s (s \times 1)$ 进行回归建模并得到全波长PLS模型,并采取留一交叉验证的方法,得到相应的RMSECV值;

步骤(4):选择模型参数:依次遍历不同的 n 和 q ,重复步骤(2)到步骤(3),直至 $n=n_MAX$, $q=q_MAX$,得到不同模型参数下的RMSECV值; n 和 q 的遍历间隔分别为2和1,选取使得RMSECV最小的 n 和 q 作为优化选取下的模型参数;

步骤(5):筛选光谱矩阵:对校正集预处理后光谱矩阵 $X_s' (s \times p)$ 及对应样本性质矩阵 $Y_s (s \times 1)$ 展开相关分析,得到 p 个波长变量的相关系数,构成相关系数矩阵 $C (1 \times p)$;设相关系数大于设定的相关系数阈值 TH 的波长变量共有 p' 个,选取这部分波长变量对应的预处理后光谱矩阵 $X_s' (s \times p)$ 的子集,构成筛选后光谱矩阵 $X_{s_selected} (s \times p')$;

步骤(6):部分波长PLS建模:在步骤(4)优化选取的模型参数下,对筛选后光谱矩阵 $X_{s_selected} (s \times p')$ 及对应样本性质矩阵 $Y_s (s \times 1)$ 进行预处理和PLS回归建模,得到部分波长下的PLS模型,并采取留一交叉验证的方法,得到相应的RMSECV值;

步骤(7):选择相关系数阈值 TH :以 d_TH 为相关系数阈值遍历间隔,依次遍历不同的 TH ,重复步骤(5)至步骤(6),直至 $TH=TH_MAX$,得到不同的RMSECV值;选取使得RMSECV最小的 TH 作为优化选取的相关系数阈值,建立部分波长下的优化PLS模型;

步骤(8):训练RBF神经网络:将步骤(7)所得的优化模型所提取的主成分得分矩阵 $Score (s \times q)$ 作为输入变量,对应的样本性质矩阵 $Y_s (s \times 1)$ 作为目标变量,进行RBF神经网络的训练,得到最终的CC-PLS-RBFNN优化模型。

2. 根据权利要求1所述的一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,其特征在于,所述步骤(4)中,所述最大窗口宽度 n_MAX 的取值范围优选15到45之间。

3. 根据权利要求1所述的一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,其

特征在于,所述步骤(4)中,所述最大主成分个数 q_{MAX} 的取值范围优选10到50之间。

4.根据权利要求1所述的一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,其特征在于,所述步骤(7)中,所述相关系数阈值遍历间隔 d_{TH} 的取值范围优选0.001到0.1之间。

5.根据权利要求1所述的一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,其特征在于,所述步骤(7)中,所述最大相关系数阈值 TH 的取值范围优选0.4到0.6之间。

6.根据权利要求1所述的一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,其特征在于,所述步骤(8)中,所述RBF神经网络所用径向基函数为高斯核函数,形式为 $k(||x-x_c||)=\exp[-||x-x_c||^2/(2\sigma^2)]$,式中 x_c 是核函数中心; σ 为扩展因子,优选为0.1至2之间;网络的神经元数量 m 优选 $s/3$ 到 s 之间,其中 s 为训练网络的校正集样本数量。

一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法

技术领域

[0001] 本发明涉及近红外光谱分析领域,尤其是涉及一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法。

背景技术

[0002] 近红外光谱分析技术作为一种快速无损的定量分析方法,已成功应用于农业、食品、化工和生物科学等多个领域,创造了可观的经济、社会效益。近红外光谱中包含了丰富的样品基团信息,然而近红外谱区内光谱吸收带较宽并且存在严重的光谱重叠,对此通常引入统计学方法建立合适的多元校正模型,从而实现光谱数据与预测目标数据间的关联。

[0003] 目前,关于近红外光谱分析中校正模型的研究,典型的线性建模方法有多元线性回归、主成分回归以及偏最小二乘法等,非线性校正方法主要包含支持向量机和神经网络。多元线性回归的局限性在于,建模过程中包含对光谱矩阵或样本性质矩阵求逆的步骤,该步骤在矩阵奇异时无法完成,同时多元线性回归模型自身不具备去除数据噪声的能力;主成分回归通过对光谱矩阵或样本性质矩阵进行分解从而确保矩阵非奇异,但进行分解时未考虑到光谱数据和样本性质数据间的关联性,因此不适用于分析目标信息在光谱信息中含量微弱的情况;偏最小二乘法在主成分回归的基础上考虑了光谱数据和样本性质数据的相关性,但作为一种线性校正方法,其建立的模型始终存在一定的非线性误差;基于支持向量机的回归模型有较好的非线性拟合能力,但在建模过程中借助二次规划来求解支持向量,涉及样本数量阶次的矩阵计算,在样本数量偏大时耗费大量的机器内存和计算时间;神经网络模型可以映射任意复杂的非线性关系,并且学习规则简单便于计算机实现,但对数据的充分性有很高的要求,当数据量不足时容易过拟合,降低模型鲁棒性,上述现有方法的局限性,影响了模型的鲁棒性和精度。

发明内容

[0004] 本发明的目的在于,针对现有技术的不足,提出将相关系数法(correlation coefficient method,简称CC)、偏最小二乘法(Partial Least Squares,简称PLS)、径向基函数神经网络(Radial Basis Function Neural Network,简称RBFNN)相结合并优化选取重要参数,提供一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,以提升近红外光谱建模的鲁棒性和精度。

[0005] 为此,本发明采用如下解决方案:

[0006] 一种基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,所述基于CC-PLS-RBFNN优化模型的近红外光谱分析方法包括:利用三阶Savitzky-Golay卷积平滑滤波及一阶导数校正对原始光谱进行预处理;在全波长段建立PLS模型并优化选取平滑滤波的窗口宽度和PLS提取的主成分个数;计算每个波长变量的相关系数,截取相关系数大于设定阈值的波长变量参与建模并优化选取阈值大小;利用优化选取的窗口宽度、主成分个数和波长变量得到优化PLS模型;用优化PLS模型提取的主成分得分和性质矩阵训练RBF神经网络,获得最终

的CC-PLS-RBFNN优化模型;具体步骤包括:

[0007] 步骤(1):参数初始化:将样本划分为校正集和预测集;校正集的光谱矩阵为 $X_s (s \times p)$,具有 s 个样本,全波长点有 p 个,校正集的样本性质矩阵为 $Y_s (s \times 1)$;预测集的光谱矩阵为 $X_t (t \times p)$,具有 t 个样本,全波长点有 p 个,预测集的样本性质矩阵为 $Y_t (t \times 1)$;定义 n 为平滑滤波的窗口宽度, q 为主成分个数, TH 为相关系数阈值;设定最大窗口宽度 n_MAX ,最大主成分个数 q_MAX ,相关系数阈值遍历间隔 d_TH ,最大相关系数阈值 TH_MAX ,以及设定RBF神经网络核函数扩展因子 σ ,网络神经元数量 m ;窗口宽度 n 初始化为5,主成分个数 q 初始化为5,相关系数阈值 TH 初始化为0;

[0008] 步骤(2):光谱预处理:对校正集原始光谱矩阵 $X_s (s \times p)$ 逐行进行 n 点三阶Savitzky-Golay卷积平滑滤波,并进行一阶求导完成光谱预处理,得到校正集预处理后光谱矩阵 $X_s' (s \times p)$;

[0009] 步骤(3):PLS初始建模:对预处理后的 s 个校正集的样本利用PLS模型提取 q 个主成分,对预处理后光谱矩阵 $X_s' (s \times p)$ 及对应的样本性质矩阵 $Y_s (s \times 1)$ 进行回归建模并得到全波长PLS模型,并采取留一交叉验证的方法,得到相应的RMSECV值;

[0010] 步骤(4):选择模型参数:依次遍历不同的 n 和 q ,重复步骤(2)到步骤(3),直至 $n = n_MAX$, $q = q_MAX$,得到不同模型参数下的RMSECV值; n 和 q 的遍历间隔分别为2和1,选取使得RMSECV最小的 n 和 q 作为优化选取下的模型参数;

[0011] 步骤(5):筛选光谱矩阵:对校正集预处理后光谱矩阵 $X_s' (s \times p)$ 及对应样本性质矩阵 $Y_s (s \times 1)$ 展开相关分析,得到 p 个波长变量的相关系数,构成相关系数矩阵 $C (1 \times p)$;设相关系数大于设定的相关系数阈值 TH 的波长变量共有 p' 个,选取这部分波长变量对应的预处理后光谱矩阵 $X_s' (s \times p)$ 的子集,构成筛选后光谱矩阵 $X_{s_selected} (s \times p')$;

[0012] 步骤(6):部分波长PLS建模:在步骤(4)优化选取的模型参数下,对筛选后光谱矩阵 $X_{s_selected} (s \times p')$ 及对应样本性质矩阵 $Y_s (s \times 1)$ 进行预处理和PLS回归建模,得到部分波长下的PLS模型,并采取留一交叉验证的方法,得到相应的RMSECV值;

[0013] 步骤(7):选择相关系数阈值 TH :以 d_TH 为相关系数阈值遍历间隔,依次遍历不同的 TH ,重复步骤(5)至步骤(6),直至 $TH = TH_MAX$,得到不同的RMSECV值;选取使得RMSECV最小的 TH 作为优化选取的相关系数阈值,建立部分波长下的优化PLS模型;

[0014] 步骤(8):训练RBF神经网络:将步骤(7)所得的优化模型所提取的主成分得分矩阵 $Score (s \times q)$ 作为输入变量,对应的样本性质矩阵 $Y_s (s \times 1)$ 作为目标变量,进行RBF神经网络的训练,得到最终的CC-PLS-RBFNN优化模型。

[0015] 在采用上述技术方案的同时,本发明还可以采用或者组合采用以下进一步的技术方案:

[0016] 所述步骤(4)中,所述最大窗口宽度 n_MAX 的取值范围优选15到45之间。

[0017] 所述步骤(4)中,所述最大主成分个数 q_MAX 的取值范围优选10到50之间。

[0018] 所述步骤(7)中,所述相关系数阈值遍历间隔 d_TH 的取值范围优选0.001到0.1之间。

[0019] 所述步骤(7)中,所述最大相关系数阈值 TH_MAX 的取值范围优选0.4到0.6之间。

[0020] 所述步骤(8)中,所述步骤(8)中,所述RBF神经网络所用径向基函数为高斯核函数,形式为 $k(|x-x_c|) = \exp[-|x-x_c|^2/(2\sigma^2)]$,式中 x_c 是核函数中心; σ 为扩展因子,优选

为0.1至2之间;网络的神经元数量 m 优选 $s/3$ 到 s 之间,其中 s 为训练网络的校正集样本数量。

[0021] 本发明所提供的近红外光谱分析方法具有以下优点:

[0022] (1)通过卷积平滑滤波和求导校正对原始光谱进行预处理,降低光谱中随机噪声对建模的不利影响;

[0023] (2)采用相关系数法对光谱进行波长筛选,最大化甄别出有益波长参与建模;

[0024] (3)通过PLS模型和RBFNN模型的结合,在保留PLS提取光谱有效信息能力的同时,利用RBFNN的任意非线性拟合能力,降低主成分得分的非线性回归残差,提升了模型精度;

[0025] (4)采用留一交互验证,最优化选择建模过程中的关键参数从而优化所建模型,提升了模型的鲁棒性和精度。

附图说明

[0026] 图1为本发明基于CC-PLS-RBFNN优化模型的近红外光谱分析方法的流程图。

[0027] 图2为玉米籽粒的原始近红外光谱图。

[0028] 图3为滤波窗口宽度、PLS主成分个数与模型RMSECV的关系图,图中标出点为筛选出的最优参数点。

[0029] 图4为经滤波并求导后的近红外光谱图。

[0030] 图5为相关系数阈值与模型RMSECV的关系图,图中标出点是筛选出的最优阈值点。

[0031] 图6为玉米籽粒淀粉含量预测值与真实值的比较图。

具体实施方式

[0032] 本发明的具体实施方式,将结合附图和近红外光谱分析的具体实施例进行详述。

[0033] 图1为本发明提出的基于CC-PLS-RBFNN优化模型的近红外光谱分析方法。

[0034] 本具体实施例,采用玉米籽粒近红外光谱数据样本来验证本发明的方法。该数据集包含80个近红外光谱测量样本,淀粉的质量含量范围为0%~100% (w/w),考察近红外光谱与玉米籽粒中淀粉含量之间的关系。样本集中,光谱仪的扫描区间为1100~2498nm,扫描间隔为2nm,即每份光谱样本数据中包含700个采样波长。

[0035] 针对玉米籽粒近红外光谱数据样本,实施本发明提出的基于CC-PLS-RBFNN优化模型的近红外光谱分析方法,步骤如下:

[0036] 步骤(1):参数初始化:对于样本个数为80,波长个数为 $p=700$,原波长集为 $V=\{x_1, \dots, x_{p-1}, x_p\}=\{1100, 1102, \dots, 2496, 2498\}$ 的玉米籽粒近红外光谱样本集,按3:1随机划分校正集和预测集,得到校正集样本60个,光谱矩阵为 $X_s(60 \times 700)$,表征玉米籽粒中淀粉含量的性质矩阵为 $Y_s(60 \times 1)$;得到预测集样本20个,光谱矩阵为 $X_t(20 \times 700)$,表征玉米籽粒中淀粉含量的性质矩阵为 $Y_t(20 \times 1)$;所述最大窗口宽度 n_MAX 的取值范围优选15到45之间,所述最大主成分个数 q_MAX 的取值范围优选10到50之间,所述最大相关系数阈值 TH_MAX 的取值范围优选0.4到0.6之间, n_MAX 、 q_MAX 和 TH_MAX 的值越大,参数寻优所需的时间越长, n_MAX 、 q_MAX 和 TH_MAX 的值太小,则搜索范围可能不包含最优解,本具体实施例中设定 $n_MAX=39$, $q_MAX=40$, $TH_MAX=0.5$;所述相关系数阈值遍历间隔 d_TH 的取值范围优选0.001到0.1之间,决定相关系数阈值的搜索密度,太大会使得求解的参数不够逼近最优值,太小则寻优耗时过长,本具体实施例中 $d_TH=0.02$;所述扩展因子 $\sigma=1$;所述神经元数量 m

=30;窗口宽度n初始化为5,主成分个数q初始化为5,相关系数阈值TH初始化为0。

[0037] 步骤(2):光谱预处理:对校正样本集光谱矩阵 X_s (60×700) 逐行进行n点三阶Savitzky-Golay卷积平滑滤波,再求取一阶导数完成光谱的预处理,得到校正样本集预处理后的光谱矩阵 X_s' (60×700);

[0038] 步骤(3):PLS初始建模:对预处理后的校正集光谱矩阵 X_s' (60×700) 以及对应样本性质矩阵 Y_s (60×1),按公式(1)至公式(3)建立主成分个数为q的PLS模型;该模型的数学表达式如下:

$$[0039] \quad X_s = TP^T + E \quad (1)$$

$$[0040] \quad Y_s = UQ^T + F \quad (2)$$

$$[0041] \quad T = XW(P^TW)^{-1} \quad (3)$$

[0042] 其中校正光谱矩阵 X_s 的得分矩阵是T,载荷矩阵是P,权重矩阵是W,残差光谱是E;校正样本性质矩阵 Y_s 的得分矩阵是U,载荷矩阵是Q,残差光谱是F。

[0043] 采取留一交叉验证的方法,得到所建立PLS模型的交互验证均方根误差RMSECV。

[0044] 步骤(4):选择模型参数:依次遍历不同的n和q,重复步骤(2)到步骤(3),直至 $n = n_MAX$, $q = q_MAX$,得到不同模型参数n和q下的RMSECV值,n和q的遍历间隔分别为2和1。具体遍历过程为,先保持q值为初始值5,从n的初始值5开始,每次循环后若 $n < n_MAX$,则执行 $n = n + 2$ 并继续循环过程;若 $n = n_MAX$ 则再判断若 $q < q_MAX$ 则令n回归初始值5,执行 $q = q + 1$ 并继续循环过程,当 $q = q_MAX$ 时完成遍历终止循环。比较所有的RMSECV值,选取使得对应RMSECV最小的n和q作为优化选取的模型参数;

[0045] 步骤(5):筛选光谱矩阵:对校正集预处理后光谱矩阵 X_s' (60×700) 及对应样本性质矩阵 Y_s (60×1) 展开相关分析,得到700个波长变量的相关系数,构成相关系数矩阵C (1×700)。第p个波长变量上的相关系数计算公式如下

$$[0046] \quad c_p = \frac{\sum_{i=1}^n (x_{i,p} - \bar{x}_p)(y_{i,p} - \bar{y}_p)}{\sqrt{\sum_{i=1}^n (x_{i,p} - \bar{x}_p)^2 \sum_{i=1}^n (y_{i,p} - \bar{y}_p)^2}} \quad (4)$$

[0047] 式中 $x_{i,p}$ 是预处理后光谱矩阵 X_s' (60×700) 中第i个样本在第p个波长变量上的值, $y_{i,p}$ 为样本性质矩阵 Y_s (60×1) 中第i个样本的值,其中 $\bar{x}_p = \frac{\sum_{i=1}^n x_{i,p}}{n}$, $\bar{y}_p = \frac{\sum_{i=1}^n y_{i,p}}{n}$ 。

[0048] 选取相关系数大于设定阈值TH的总计 p' 个波长变量,该 p' 个波长变量对应的预处理后光谱矩阵 X_s' (60×700) 的子集,构成筛选后光谱矩阵 $X_{s_selected}$ ($60 \times p'$);

[0049] 步骤(6):部分波长PLS建模:在步骤(4)优化选取的模型参数n和q下,对筛选后光谱矩阵 $X_{s_selected}$ ($60 \times p'$) 及对应样本性质矩阵 Y_s (60×1) 进行预处理和PLS回归建模,得到部分波长下的PLS模型,采取留一交叉验证的方法,得到相应的RMSECV值;

[0050] 步骤(7):选择相关系数阈值TH:以 d_TH 为相关系数阈值遍历间隔,依次遍历不同的TH,重复步骤(5)至步骤(6),直至 $TH = TH_MAX$,得到不同的RMSECV值。具体遍历过程为,从TH的初始值0开始,每次循环后若 $TH < TH_MAX$,则执行 $TH = TH + d_TH$ 并继续循环过程;若 $TH = TH_MAX$,则完成遍历终止循环。最后选取使得RMSECV最小的TH作为优化选取的相关系数阈值,建立部分波长下的优化PLS模型;

[0051] 步骤(8):训练RBF神经网络:将步骤(7)所得的优化PLS模型所提取的主成分得分

矩阵Score ($60 \times q$) 作为输入变量,对应的样本性质矩阵 Y_s (60×1) 作为目标变量,进行RBF神经网络的训练,得到最终的CC-PLS-RBFNN优化模型。相比于更常用的反向传播(backpropagation,简称BP)神经网络而言,RBF神经网络具有唯一最佳逼近,避免陷入局部最优的能力。此处RBF神经网络选用高斯核函数为径向基函数,形式为 $k(||x-x_c||)=\exp[-||x-x_c||^2/(2\sigma^2)]$,式中 σ 为扩展因子,其值过大会引起大的网络逼近误差,过小则容易造成过拟合现象,此处取 $\sigma=1$;网络的神经元数量根据样本数选择为30。

[0052] 图2为校正样本集的原始近红外光谱谱图;

[0053] 图3为滤波窗口宽度 n 、PLS主成分个数 q 与全波长PLS模型交互验证均方根误差RMSECV的关系图,图中标出RMSECV取得最小值的点,对应最优模型参数 $n=19$, $q=30$;

[0054] 图4为经19点三阶Savitzky-Golay卷积平滑滤波并一次求导后的预处理后校正样本集光谱谱图,对比图2显示原始光谱中的噪声和基线漂移得到显著改善;

[0055] 图5为相关系数阈值TH与部分波长PLS模型交互验证均方根误差RMSECV的关系图,图中标出RMSECV取最小值的点,对应最优相关系数阈值 $TH=0.28$;

[0056] 结果验证:图6显示了采用上述方法得到的CC-PLS-RBFNN优化模型对20个预测集样本淀粉含量的预测情况,通过回归分析得到预测值与真实值的关系。

[0057] 表1

	建模分析方法	RMSEP	回归相关系数
	基于 PLS 回归模型	0.0592	0.9982
[0058]	基于反向传播神经网络模型	0.1412	0.9911
	基于径向基函数神经网络模型	0.0750	0.9974
	基于偏最小二乘-反向传播神经网络模型	0.0766	0.9972
	本发明方法	0.0569	0.9987

[0059] 表1以预测均方根误差RMSEP(Root Mean Square Error of Prediction)和回归相关系数作为评价指标,比较了本发明方法与基于PLS回归模型、基于反向传播神经网络(BPNN)模型、基于径向基函数神经网络(RBFNN)模型以及基于偏最小二乘-反向传播神经网络(PLS-BPNN)模型的建模分析方法的预测能力。比较结果显示本发明方法预测误差最小,回归相关系数最大。

[0060] 表1的结果充分表明:本发明方法能够有效地加强模型的预测能力,提升近红外光谱分析的鲁棒性和精度。

[0061] 本发明上述具体实施方式中所涉及的近红外光谱分析对象为玉米籽粒的淀粉含量,仅为优选实施例,具体实施时,也可以用于农业、食品、化工和生物科学等领域各种近红外光谱分析对象的分析过程中。

[0062] 上述具体实施方式用来解释说明本发明,仅为本发明的优选实施例,而不是对本发明进行限制,在本发明的精神和权利要求的保护范围内,对本发明作出的任何修改、等同

替换、改进等,都落入本发明的保护范围。

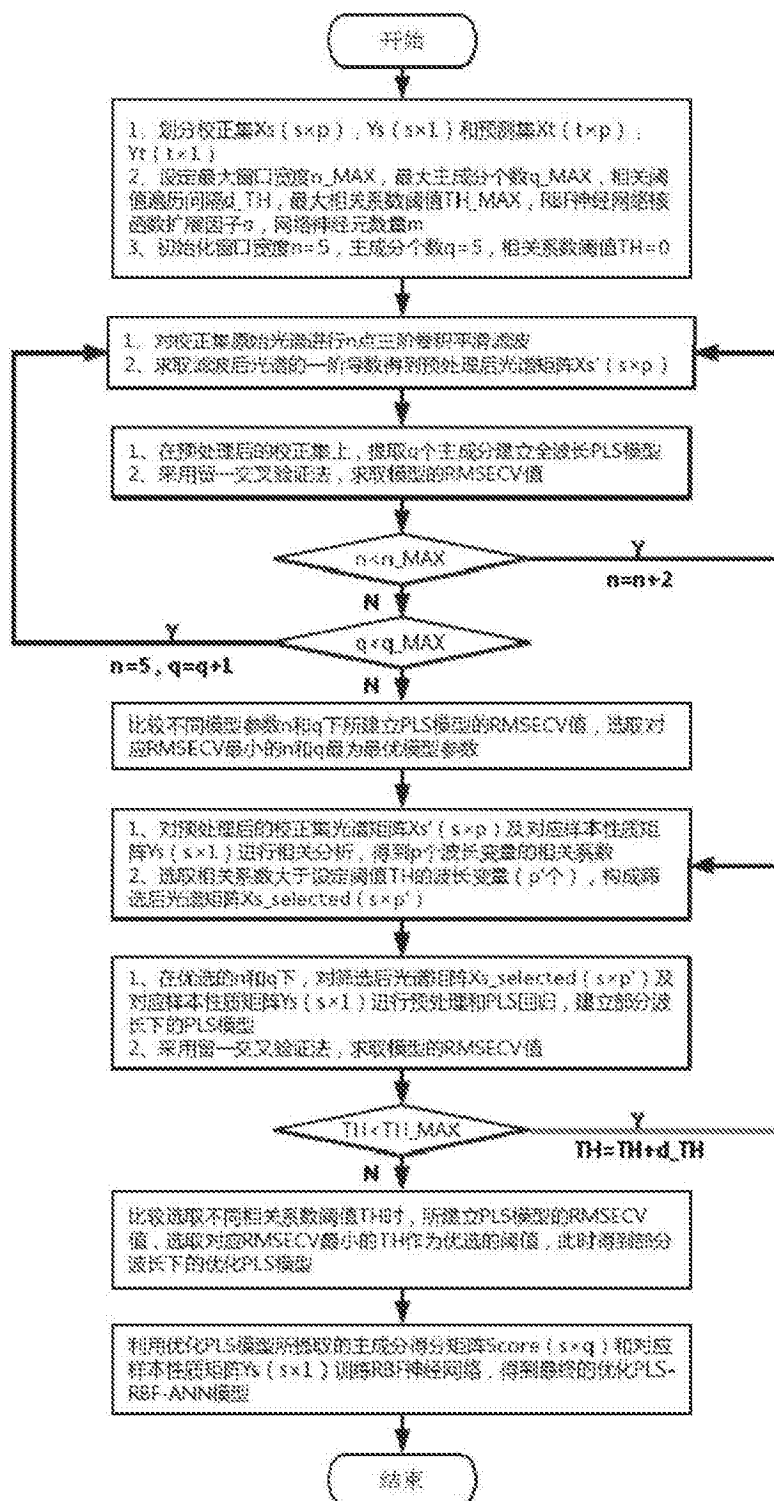


图1

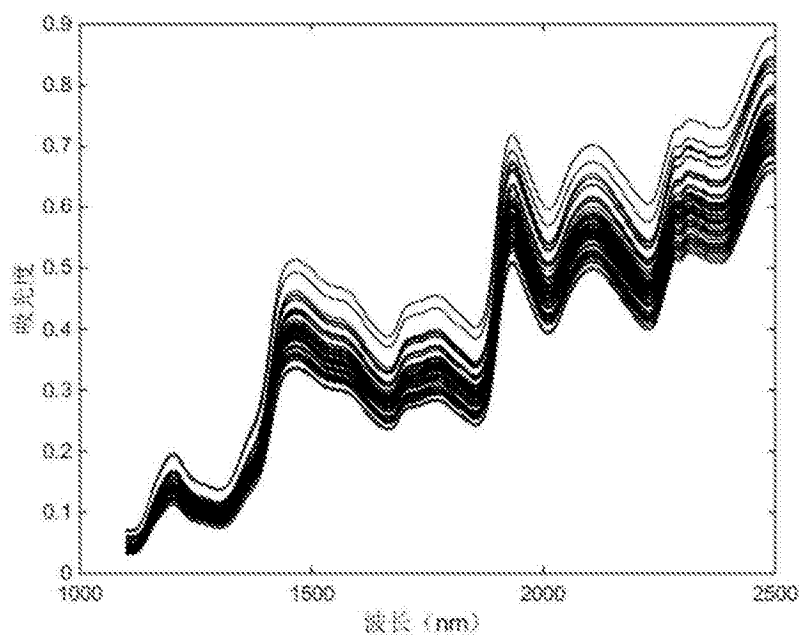


图2

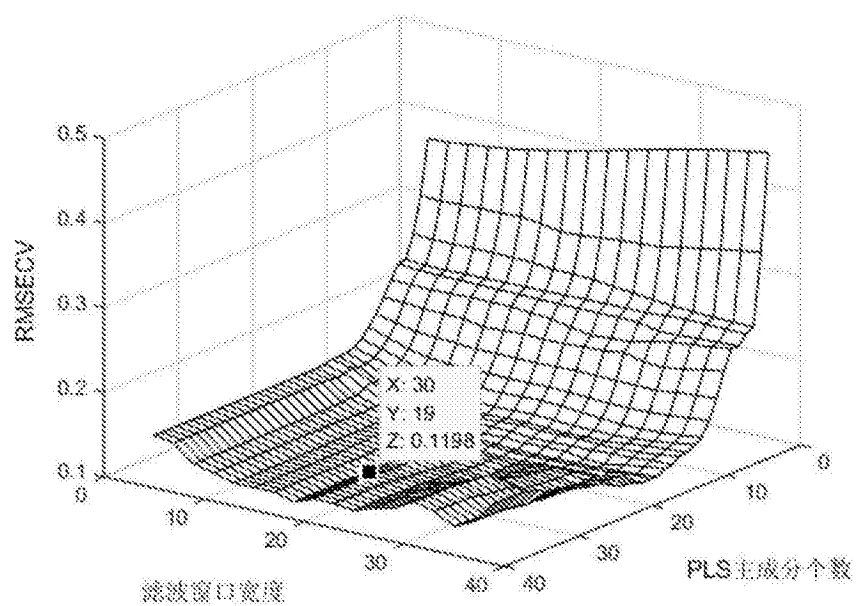


图3

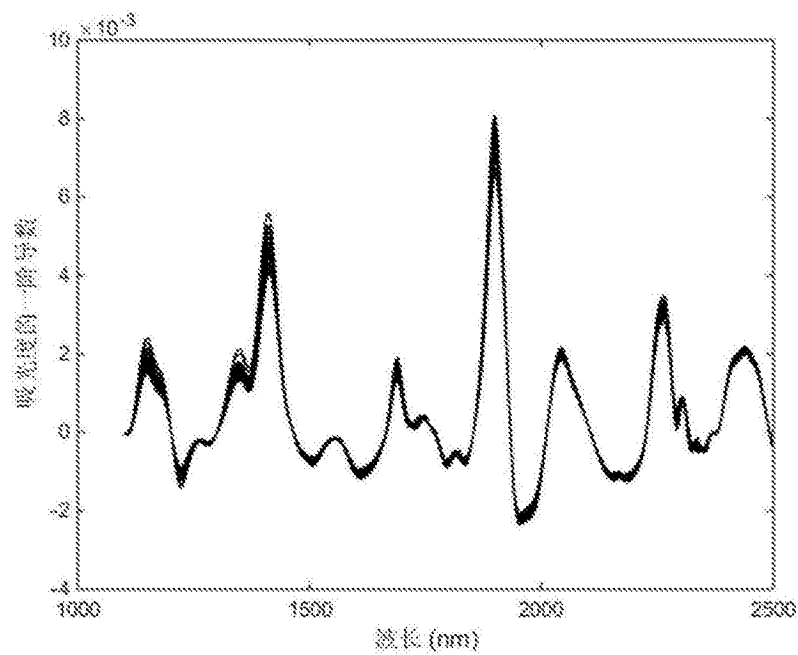


图4

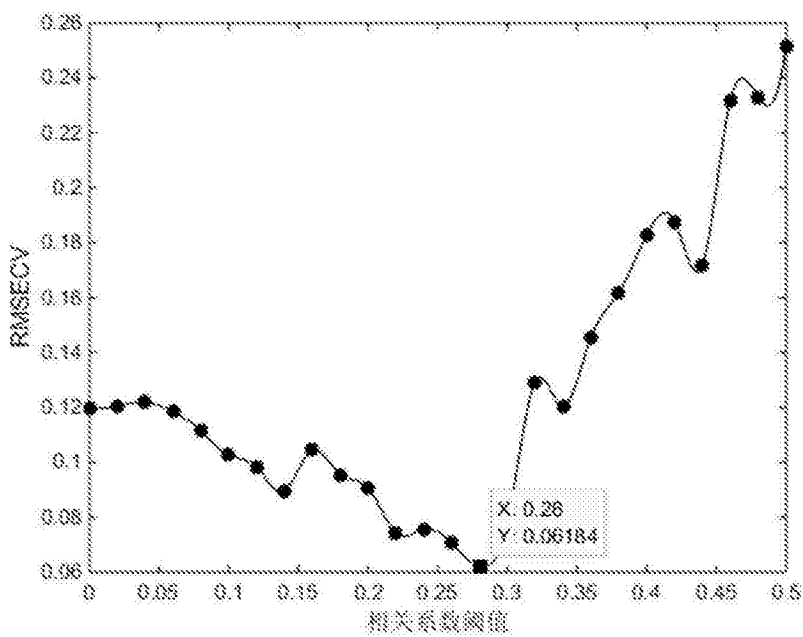


图5

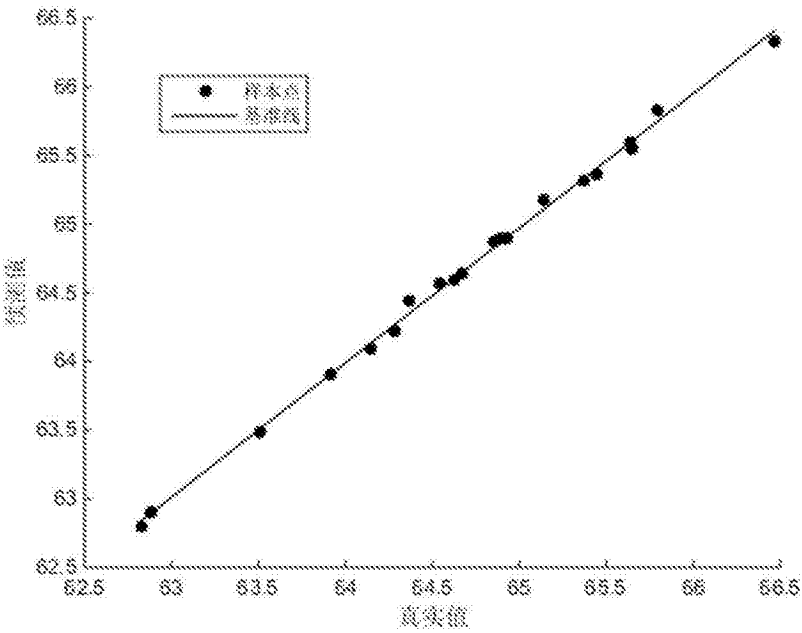


图6