

IWD-Hankel-SVD 模型下玉米叶片光谱铜污染信息预测

张建红, 杨可明*, 韩倩倩, 李艳茹, 高伟

中国矿业大学(北京)地球科学与测绘工程学院, 北京 100083

摘要 近年来在工业化和城镇化快速发展的地区,由重金属污染导致的环境问题尤为突出,特别是农业重金属污染更为社会所关注,因此,探索快速便捷的重金属污染甄别与监测方法极为重要。高光谱遥感作为新兴的重金属污染监测技术已有了深入研究。提出了固有波长尺度分解(IWD)概念和方法,并结合 Hankel 矩阵和奇异值分解(SVD)等建立了植被重金属污染程度预测的 IWD-Hankel-SVD 模型,该模型分为单变量模型和多变量模型。单变量模型主要是通过重金属污染的植被光谱 IWD 处理来获取光谱信息固有旋转分量(PRC)以提取最佳 PRC 的有效特征波段;在对各特征波段所构建的 Hankel 矩阵进行奇异值分解(SVD)基础上,依据获得该模型的奇异熵实现重金属污染信息预测。多变量模型是以植物叶绿素浓度相对值、单变量模型奇异熵作为参数实现重金属污染的信息预测。根据不同重金属 Cu^{2+} 胁迫梯度下玉米植株污染的叶片光谱和叶绿素浓度以及叶片中 Cu^{2+} 含量测定的数据,首先对不同浓度 Cu^{2+} 胁迫下玉米叶片光谱进行 IWD 分析,获得能够较好保留原始输入光谱信息的最佳 PRC,并从中提取到有效特征波段 553~680, 681~780, 1 266~1 429, 1 430~1 631, 1 836~1 913 和 1 914~2 111 nm;然后对每一个特征波段构造其 Hankel 矩阵并进行 SVD 处理,以求取单变量的 IWD-Hankel-SVD 模型奇异熵;最后通过各特征波段所对应模型奇异熵与玉米叶片中 Cu^{2+} 含量的相关分析,得到依据 1 266~1 429 和 1 836~1 913 nm 特征波段计算出奇异熵与玉米叶片中 Cu^{2+} 含量的决定系数 R^2 均高达 0.9 左右,说明这两个特征波段用于 IWD-Hankel-SVD 模型的 Cu 污染程度预测更具优越性和解释能力。同时,再把玉米叶片中叶绿素浓度相对值、1 266~1 429 和 1 836~1 913 nm 特征波段相应模型奇异熵作为参数,采用偏最小二乘回归分析,得出多变量 IWD-Hankel-SVD 模型的玉米叶片 Cu 污染程度预测能力更强,决定系数 R^2 达到 0.9476,证明了多变量模型更具有鲁棒性和稳健性。

关键词 光谱分析;玉米叶片;重金属铜污染;固有波长尺度分解;预测模型

中图分类号: TP7 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)05-1505-08

引言

近些年来环境污染,特别是重金属污染在工业化和城市化快速发展地区尤为严重,因此重金属污染监测与治理技术等已成为当今环境保护领域的研究热点,其中农业重金属污染更是备受关注^[1]。农业重金属污染主要来源于农药和近农田区域采矿活动等^[2]。农田土壤受重金属污染程度不易被察觉,而重金属离子可以被农作物吸收并富集,进而能造成粮食减产、毒害食物链、威胁食品安全、危害人体健康。

如何实时快捷有效识别重金属污染、更加准确甄别重金属污染的类型与程度是当前亟待研究与突破的关键技术。传

统的重金属污染监测方法通常是基于野外取样后,在室内采用无焰原子吸收分光光度法(FAAS)、原子吸收光谱法(AAS)等进行测定,虽然检测精度高,但也具有耗费大量人力、物力和财力,损害被测对象,所检测仪器昂贵、维护成本高、操作程序复杂和检测时间长等缺点^[3];而今高光谱遥感技术具有图谱合一、光谱连续,高效、无损、低成本等优势,正在被广泛应用于重金属污染甄别与监测的探索研究中^[4]。Liu 等^[5]通过提取植被红边光谱曲线,构造归一化红边指数和红边叶绿素指数,用于识别植被受重金属镉(Cd)的胁迫;乔晓英等^[6]以矿区周边玉米、苦菜等作为研究对象,构建植物中重金属 Cd、铅(Pb)质量比与植物光谱红边、蓝谷等特征参数的关系模型,可间接反演植物中重金属的含量;

收稿日期:2020-04-27,修订日期:2020-08-15

基金项目:国家自然科学基金项目(41971401),中央高校基本科研业务费专项资金项目(2020YJSDC02)资助

作者简介:张建红,女,1988年生,中国矿业大学(北京)地球科学与测绘工程学院博士研究生 e-mail: Lenova20@163.com

* 通讯作者 e-mail: ykm69@163.com

付萍杰等^[7]对叶片光谱进行经验模态分解并结合自相关函数、多重分形理论构建 Cu、Pb 含量反演的多元线性回归模型,成功识别了玉米叶片中重金属类别及含量;朱叶青等^[8]通过获取不同生育期、不同 Cu 污染下的植被叶片光谱信息,采用 7 个特征波段结合光谱角方法研究 Cu 污染下的叶片光谱特征,表明 Cu 污染叶片与健康叶片的光谱存在明显差异,同时证明了 Cu 污染后的叶片内部结构更加紊乱。

作为植被生长所必需的微量元素 Cu,也属于一种重金属,当浓度较低时有利于植被生长,浓度过高就会抑制植物的光合作用,对植被生长产生胁迫作用,轻则引起植物代谢紊乱,生长受阻,重则导致植物死亡。已有报道当植被受到重金属污染时,多数会长势矮小,叶片失去绿色等症状^[9]。本工作提出固有波长尺度分解(intrinsic wavelength-scale decomposition, IWD)的方法并运用于甄别农作物重金属污染的高光谱信息处理与分析中。依据受 Cu 污染的玉米叶片光谱数据,基于 Hankel 矩阵构建、奇异值分解(singular value decomposition, SVD)等建立玉米叶片 Cu 污染信息预测的 IWD-Hankel-SVD 模型,该模型可有效预测玉米叶片中铜离子(Cu^{2+})含量,能为植物 Cu 污染信息预测、甄别提供一种新方法。

1 理论与算法

1.1 固有波长尺度分解

固有波长尺度分解(IWD)引自于固有时间尺度分解(intrinsic time-scale decomposition, ITD)。ITD 是由 Mark G Frei 与 Ivan Osorio 首次提出^[10],相对于经验模态分解(empirical mode decomposition, EMD)和小波变换等,ITD 是一种新的时频分析方法^[11]。经过 ITD 分解后的信号分量具有相对较为完整的时频信息,能够反映信号频率变化等优点。本工作用波长(wavelength, W)替代 ITD 中的时间(time, T),探索形成光谱信号的 IWD 处理技术,并引入到高光谱数据的信息处理与分析中。IWD 可以将波长序列(w)自适应分解为几个固有旋转分量(proper rotation component, PRC)和一个趋势分量(r),每个旋转分量只需要通过迭代一次就可以得到。对于一条准备分解的光谱曲线 X_w ,有极值点 X_k ,其基线提取算子定义为 L ,可将基线分量 L_w 与准备分解的光谱曲线 X_w 分开,再定义一个固有旋转分离算子为 H ,于是有

$$X_w = LX_w + (1-L)X_w = L_w + H_w \quad (1)$$

其中, $L_w = LX_w$ 为基线分量, $H_w = (1-L)X_w = HX_w$ 为 PRC 分量。

对于给定的光谱曲线, IWD 的处理步骤如下:

(1)提取光谱曲线 X_w 上的极值点 X_k ($k=1, 2, \dots, N$, N 表示极值点的个数),令 τ_k 是 X_w 局部极值点所在的波长集合,定义 $\tau_0=0$ 作为 X_w 的起始波长,并且使 X_w 和 L_w 在 τ_k 处的值分别为 X_{τ_k} 和 L_{τ_k} ,为了表示方便,将 X_{τ_k} 和 L_{τ_k} 分别定义为 X_k 和 L_k 。

假设 L_w 和 H_w 被定义在波段 $[\tau_0, \tau_k]$ 上, X_w 被定义在波段 $[\tau_0, \tau_{k+2}]$ 上,并且,假设 L_w 是 X_w 在区间 $(\tau_k, \tau_{k+1}]$ 的

仿射线性逼近,即 $L_w = mX_w + n_w$, $w \in (\tau_k, \tau_{k+1}]$ 。因为 L_w 必须满足局部极值点的上述边界条件,所以 L_w 被表示为以下形式

$$L_w = L_k + \frac{L_{k+1} - L_k}{X_{k+1} - X_k}(X_w - X_k) \quad (2)$$

假设光谱曲线的基本趋势足够平滑,则可以忽略各局部极值点间的变化,于是可以得出

$$X_{k+1} = X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}(X_{k+2} - X_k) \quad (3)$$

因为 $L_{k+1} = X_{k+1}$,于是有

$$L_{k+1} = \alpha X_{k+1} + (1-\alpha)X_{k+1} =$$

$$\alpha \left[X_k + \left(\frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k} \right) (X_{k+2} - X_k) \right] + (1-\alpha)X_{k+1} \quad (4)$$

式(4)中: α 为线性缩放因子,可用于调整提取到的 PRC 幅度, $\alpha \in [0, 1]$,一般情况下,经验值取值 0.5。

(2)将式(2)计算得出的基线分量 $L_w = LX_w$,按照式(1)计算提取固有旋转分量 H_w ,即 PRC _{i} 。

(3)将基线分量 L_w 作为一个输入光谱再进行下一次分解,重复第(1)和(2)步的分解,当基线分量 L_w 变为单调或小于某个预设值的趋势分量(r)时,分解终止。

IWD 的总分解过程为

$$X_w = H_w + L_w = HX_w + LX_w = HX_w + (H+L)LX_w =$$

$$HX_w + HLX_w + L^2X_w = \left(H \sum_{k=0}^{p-1} L^k + L^p \right) X_w =$$

$$H_w^1 + H_w^2 + H_w^3 + \dots + H_w^k + L_w^p \quad (5)$$

式(5)中, p 为 IWD 分解结束时分解出的固有旋转分量个数, H_w^k 为第 k 个固有旋转分量, L_w^p 为趋势分量。

经过 IWD 分解后,原始输入的光谱曲线 X_w 即被分解为多个固有旋转分量(PRC _{i})和一个单调趋势分量(r)。该方法计算复杂度低,分解精度高,鲁棒性强,有效的避免了波形叠加、模态混叠现象。

1.2 Hankel 矩阵

汉克尔(Hankel)矩阵具有很多优良的性能和特殊的性质,因此,作为一种优秀的数学工具被广泛应用于诸多领域的数值模拟计算中。Hankel 矩阵的每一条逆对角线上的元素都相等。可选择经 IWD 处理后较好保留了原始输入光谱信息的 PRC,构造 Hankel 矩阵 Z_a

$$Z_a = \begin{bmatrix} f_{a,1} & f_{a,2} & \dots & f_{a,n} \\ f_{a,2} & f_{a,3} & \dots & f_{a,n+1} \\ \vdots & \vdots & \dots & \vdots \\ f_{a,m} & f_{a,m+1} & \dots & f_{a,N} \end{bmatrix} \quad (6)$$

式(6)中, $1 \leq a \leq J$,且 $1 < n < N$; N 为原始数据的采样点个数。当 $m = N - n + 1$ 时,矩阵 Z_a 即为 Hankel 矩阵^[12]。

1.3 奇异值分解与奇异嫡

奇异值分解(SVD)即对于任意一个维数的实矩阵 $A \in \mathbf{R}^{m \times n}$,都可分解为三个矩阵的乘积,即

$$A = USV^T \quad (7)$$

式(7)中, U 为单位正交矩阵,即其列向量是单位向量并且相互正交,是奇异值对应的特征向量^[13]; V 为单位正交矩阵,其行向量是单位向量并且相互正交,也是奇异值对应的特征向量; S 为对角矩阵, $S = [\text{diag}(\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_m), \mathbf{O}]$,其

中 \mathbf{O} 表示零矩阵, 它的对角线矩阵为奇异值。

奇异熵是用来度量奇异值所对应的信号分量信息量的多少, 奇异熵的计算公式是

$$E_k = \sum_{i=1}^k \Delta E_i \quad (k \leq q) \quad (8)$$

式(8)中, k 为奇异熵所对应的阶次; ΔE_i 为奇异熵在阶次 i 时得到的增量, 其计算公式为

$$\Delta E_i = -\left(\sigma_i / \sum_{i=1}^q \sigma_i\right) \ln\left(\sigma_i / \sum_{i=1}^q \sigma_i\right) \quad (9)$$

从式(9)可以看出, 随着 i 不断增大, $\sigma_i / \sum_{i=1}^q \sigma_i$ 逐渐减小, 得到的增量 ΔE_i 也在逐渐减小。随着阶次持续增加, 增量也逐渐趋于渐近值, 此时, 有用的信息量即达到饱和。

1.4 IWD-Hankel-SVD 模型

基于 IWD, Hankel 矩阵和 SVD 可构建 IWD-Hankel-SVD 模型, 用于光谱的信息分析与特征参量反演等。IWD-Hankel-SVD 模型的植物重金属污染下叶片光谱特征信息提取步骤为:

(1) IWD 分解。对需要甄别重金属污染的植物光谱数据进行 IWD 分解, 得到 PRC1, PRC2, ..., 以及 r 若干分量。

(2) 分量筛选。对经过 IWD 分解后得到的各 PRC 分量与原始输入光谱进行对比, 选择能够较好保留原始输入光谱信息的 PRC $_i$ 分量。

(3) 特征波段选择。在 PRC $_i$ 分量中选取与原始输入光谱相似的波段作为特征波段进行下一步的处理。

(4) 构造 Hankel 矩阵。利用从第(2)步得到的 PRC $_i$ 分量选择的特征波段, 构造 Hankel 矩阵 \mathbf{Z} 。

(5) SVD 分解。对第(4)步构造的 Hankel 矩阵进行 SVD 分解, 得到奇异值矩阵。

(6) 求取奇异熵。通过第(5)步得到的奇异值计算对应的奇异熵, 作为表征重金属污染下光谱信息分量的特征。

2 实验部分

(1) 实验准备。实验采用盆栽的玉米植株土培法, 玉米种子选用“中糯 301”。重金属 Cu^{2+} 胁迫试剂选用纯度较高、干扰杂质较少的分析纯级别 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 。光谱数据获取采用 SVC HR-1024i 便携式地物光谱仪, 其光谱测量范围为 350~2 500 nm, 其中 350~1 000 nm 光谱分辨率为 3.0 nm, 1 000~1 900 nm 光谱分辨率为 9.5 nm, 1 900~2 500 nm 光谱分辨率为 6.5 nm。叶绿素浓度相对值的测定采用 SPAD-502 手持便携式叶绿素测定仪。

(2) 植株培养。“中糯 301”玉米种植时采用底部有渗水孔的培土花盆, 将分析纯级别的 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 配置成相应梯度的 Cu^{2+} 溶液拌入培土作为所种植玉米生长的胁迫试剂。培植实验共设置三组平行实验, 每组设 7 个 Cu^{2+} 胁迫梯度, 即空白对照组(ck, 不添加胁迫试剂)和 100, 200, 400, 600, 800 和 1 000 $\mu\text{g} \cdot \text{g}^{-1}$ 等 6 级 Cu^{2+} 胁迫梯度, 分别标注为 ck(0), Cu(100), Cu(200), Cu(400), Cu(600), Cu(800) 和 Cu(1 000)。为了提高玉米培养效率, 事先对“中糯 301”玉米种

子进行催芽处理。植株培养期间保持水量充沛, 空气流通。2 个月后培植结束, 测定玉米叶片光谱信息、叶片中 Cu^{2+} 的含量和叶绿素浓度相对值。

(3) 光谱数据采集。在室内采用 SVC HR-1024i 光谱仪测定实验盆栽玉米在不同浓度 Cu^{2+} 胁迫下的叶片光谱数据, 选用 50W 卤素灯作为单一光源。光谱采集时分别测量每个平行组各胁迫浓度下玉米老(Old)、中(Middle)、新(New)叶片的光谱 3 次, 每个胁迫浓度获取 9 组数据, 共计 63 组光谱数据。最终获取的光谱数据是经由去除异常值, 求平均得到。

(4) 叶绿素浓度相对值测定。同步利用 SPAD-502 叶绿素仪分别测定采集光谱的叶片叶绿素浓度(重复 3 次), 最终通过求平均得到了各 Cu^{2+} 胁迫浓度下玉米叶片叶绿素浓度相对值。

(5) Cu^{2+} 含量的测定。在各类被测叶片光谱和叶绿素的数据采集结束后就即时进行叶片样本裁剪和保存处理, 为叶片中 Cu^{2+} 含量测定作准备。最终测定的叶片中 Cu^{2+} 含量是采用电感耦合等离子发射光谱仪(ICP-OES)进行分析得到。

表 1 不同浓度 Cu^{2+} 胁迫下玉米叶片中叶绿素浓度相对值及 Cu^{2+} 含量

Table 1 Relative value of chlorophyll concentration and content of Cu^{2+} in corn leaves under different concentrations of Cu^{2+}

胁迫梯度/ ($\mu\text{g} \cdot \text{g}^{-1}$)	叶绿素浓度 相对值	Cu^{2+} 含量/ ($\mu\text{g} \cdot \text{g}^{-1}$)
ck(0)	30.64	5.34
Cu(100)	29.84	6.56
Cu(200)	29.80	8.04
Cu(400)	27.69	9.19
Cu(600)	18.82	12.84
Cu(800)	19.00	14.17
Cu(1 000)	15.40	13.56

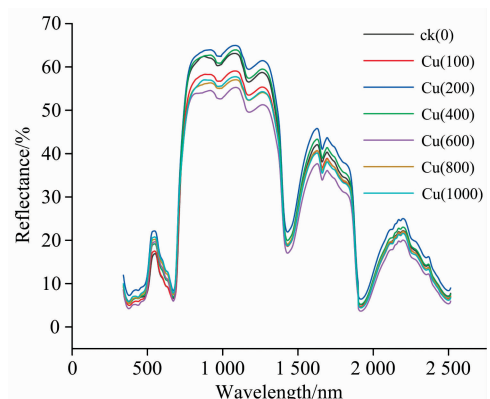


图 1 不同浓度 Cu^{2+} 胁迫下玉米叶片光谱曲线
Fig. 1 Spectral curves of corn leaves stressed by the different concentrations of Cu^{2+}

胁迫梯度以及对应的玉米叶片叶绿素浓度相对值、 Cu^{2+}

含量见表 1 所示。由表 1 可见,随着 Cu^{2+} 胁迫增大,叶绿素浓度相对值在逐渐减小,叶片中 Cu^{2+} 含量在逐渐增大。各胁迫梯度的叶片叶绿素浓度相对值与 Cu^{2+} 含量相关系数为 -0.953 , p 值为 $0.001 < 0.01$,可见两者呈负相关,负相关性高度显著,说明在重金属 Cu^{2+} 胁迫下,玉米叶片中叶绿素含量的积累受到影响^[14]。

3 结果与讨论

3.1 IWD 处理结果与分析

通过 SVC HR-1024i 光谱仪测定的不同 Cu^{2+} 胁迫梯度下玉米叶片光谱曲线如图 1 所示,受不同浓度的 Cu^{2+} 胁迫,玉米叶片光谱曲线与对照组 ck(0) 的健康玉米叶片光谱曲线有所不同,但是不能准确的甄别出受重金属 Cu^{2+} 胁迫的变化规律。为此,引入 IWD 对不同浓度 Cu^{2+} 胁迫下的玉米叶片光谱曲线进行变换,然后构造 Hankel 矩阵,进行 SVD 处

理,最后求得奇异熵,通过线性拟合建立的 IWD-Hankel-SVD 模型监测确定重金属 Cu^{2+} 胁迫下玉米叶片的污染程度。

采用 IWD 方法分解对照组 ck(0) 的健康叶片原始光谱,得到 PRC1 和 PRC2 的 2 个固有旋转分量和 1 个趋势分量 r ,如图 2 所示,其中并非每一个分量都包含有丰富的光谱特征信息,有些分量对于光谱特征信息的提取贡献很小。通过与原始输入光谱信息比较,发现 PRC1 分量较好的保留了原始输入光谱信息,并且对原始光谱中部分波段的反射率有所放大。同样,通过对 Cu(100), Cu(200), Cu(400), Cu(600), Cu(800) 和 Cu(1000) 胁迫梯度下原始叶片光谱进行 IWD 分解,发现不同胁迫浓度下的固有旋转分量中 PRC1 分量均较好地保留了原始输入光谱信息,如图 3(a—f) 所示。因此,根据 PRC1 分量的光谱保留信息,从中选择与原始输入光谱相似的 553~680, 681~780, 1 266~1 429, 1 430~1 631, 1 836~1 913 和 1 914~2 111 nm 波段区间作为有效特征波段。

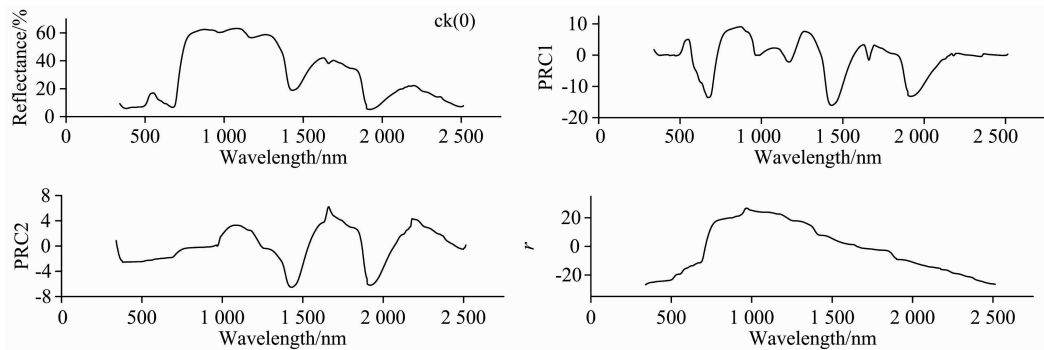


图 2 对照组 ck(0) 玉米叶片光谱及其 IWD 处理结果

Fig. 2 Spectrum and the results processed by IWD of corn leaves in ck(0) control group

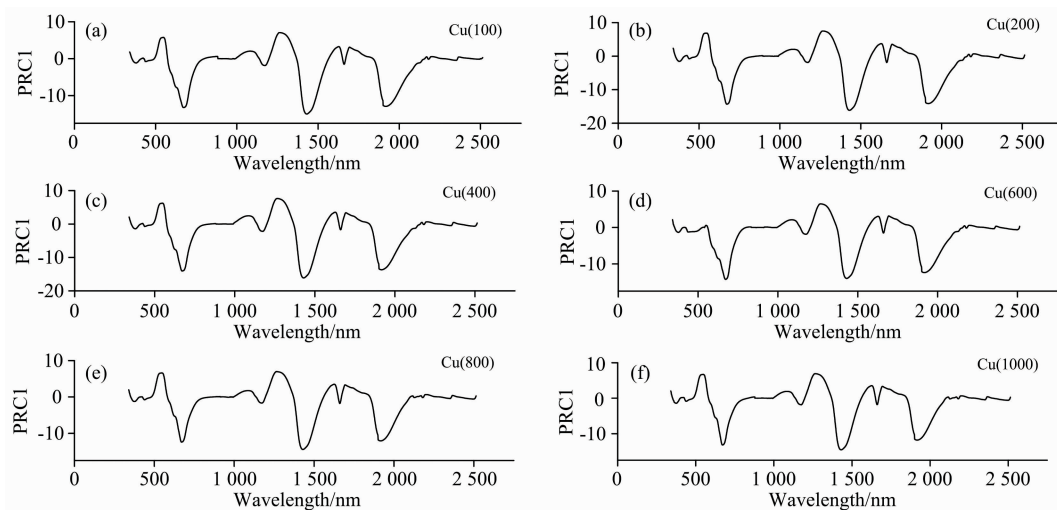


图 3 不同浓度 Cu^{2+} 胁迫下玉米叶片光谱 IWD 处理的 PRC1 分量

Fig. 3 Spectral PRC1 components of the corn leaves stressed by the different concentrations of Cu^{2+} based on the IWD processing

3.2 单变量的 IWD-Hankel-SVD 模型应用

通过对特征波段 553~680, 681~780, 1 266~1 429, 1 430~1 631, 1 836~1 913 和 1 914~2 111 nm 对应的 PRC1

分量分别构造 Hankel 矩阵,并对该矩阵进行 SVD 分解,实现 IWD-Hankel-SVD 模型的应用,最后求取模型的奇异熵用于分析监测效果。各特征波段所对应的奇异熵与玉米叶片中 Cu^{2+}

含量的相关关系见表 2。由表 2 可见, 通过特征波段 1 266~1 429 和 1 836~1 913 nm 计算求得的奇异熵与玉米叶片中 Cu^{2+} 含量的相关系数分别为 0.950 和 0.913, 说明经过 IWD 分解后的波段 1 266~1 429 和 1 836~1 913 nm 对应的 PRC1 分量对 Cu 污染较为敏感。选取相关系数大于 0.9 的特征波段计算得到的奇异熵分别与玉米叶片中 Cu^{2+} 含量进行线性拟合, 拟合结果如图 4、图 5 所示, 可见: (1) 根据特征波段 1 266~1 429 nm 计算得到的奇异熵与玉米叶片中 Cu^{2+} 含量线性拟合, 得到的模型决定系数 R^2 为 0.903, p 值小于 0.01, 相关性高度显著; (2) 根据特征波段 1 836~1 913 nm 计算得到的奇异熵与叶片中 Cu^{2+} 含量线性拟合, 得到的模型决定系数 R^2 为 0.834, p 值小于 0.01, 相关性高度显著。因此得出, 可通过 IWD-Hankel-SVD 模型来预测玉米植株受重金属 Cu 污染的程度。

表 2 特征波段的奇异熵与玉米叶片 Cu^{2+} 含量的相关关系

Table 2 The correlation between the singular entropy of characteristic bands and the content of Cu^{2+} in corn leaves

特征波段/nm	相关系数
553~680	0.678
681~780	-0.433
1 266~1 429	0.950
1 430~1 631	-0.406
1 836~1 913	0.913
1 914~2 111	-0.794

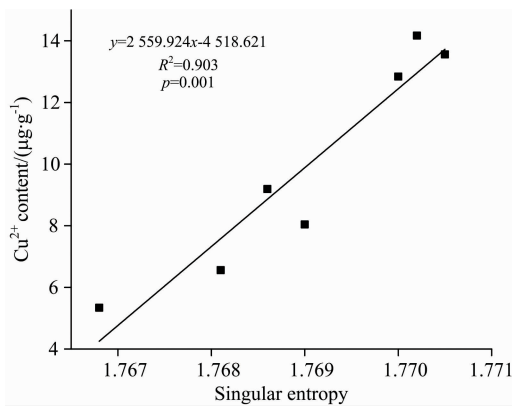


图 4 特征波段 1 266~1 429 nm 的奇异熵与玉米叶片中 Cu^{2+} 含量变化关系

Fig. 4 Relationship between the singular entropy of 1 266~1 429 nm characteristic band and the content of Cu^{2+} in corn leaves

为了验证依据所提取特征波段以及 IWD-Hankel-SVD 模型预测植物重金属污染程度的优越性, 选择传统的蓝边最大值、红肩最大值、绿峰高度等光谱特征参数(见表 3)监测方法以及采用连续投影变换(SPA)选择的特征波段如图 6、图 7 所示, 进行应用结果对比分析。通过蓝边最大值、红肩最大值、绿峰高度的监测结果与玉米叶片中 Cu^{2+} 含量的线性拟合,

得到拟合效果如图 8、图 9、图 10 所示, 通过连续投影变换(SPA)选择的特征波长与玉米叶片中 Cu^{2+} 含量的线性关系见表 4, 对比结果表明, IWD-Hankel-SVD 模型的监测能力明显要优于其他传统光谱特征等参数所建立的模型。

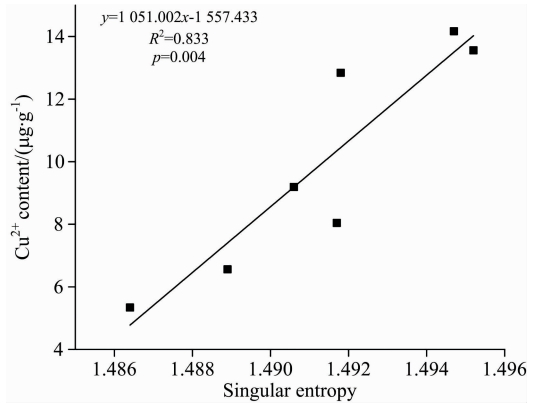


图 5 特征波段 1 836~1 913 nm 奇异熵与玉米叶片中 Cu^{2+} 含量变化关系

Fig. 5 Change relationship between the singular entropy of 1 836~1 913 nm characteristic band and the content of Cu^{2+} in corn leaves

表 3 玉米叶片的光谱特征参数名称与定义

Table 3 Names and definitions of spectral characteristic parameters of corn leaves

参数名称	定义
蓝边最大值	反射率光谱一阶导数在 450~550 nm 的最大值
红肩最大值	反射率光谱一阶导数在 670~780 nm 的最大值
绿峰高度	反射率光谱在 500~600 nm 的最大值

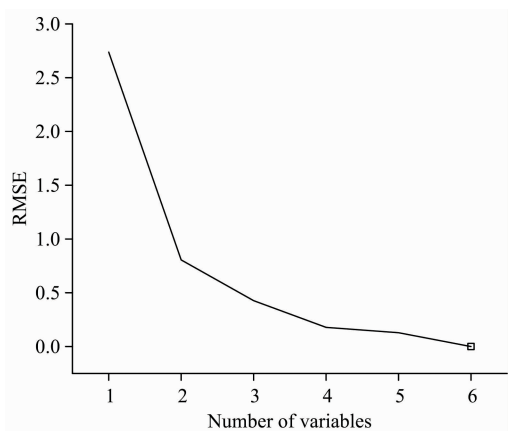


图 6 连续投影变换选择特征变量的个数

Fig. 6 Number of characteristic bands selected by SPA

3.3 多变量的 IWD-Hankel-SVD 模型应用

植被中的叶绿素含量是植物生长健康状况的重要生物指标, 决定着绿色植物的光合作用效率。随着 Cu^{2+} 胁迫浓度的增大, 玉米叶片中的 Cu^{2+} 含量在逐渐增大, 而叶绿素浓度相对值在逐渐减少。这是因为重金属胁迫对植物的生长具有负向影响, 抑制植物进行光合作用, 造成叶绿素水解增加, 含量降

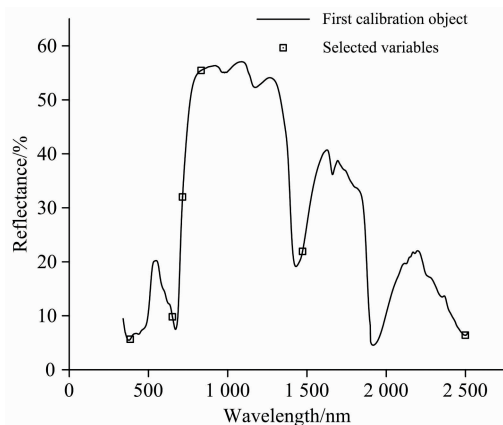


图 7 连续投影变换提取特征波长

Fig. 7 Characteristic bands extraction by SPA

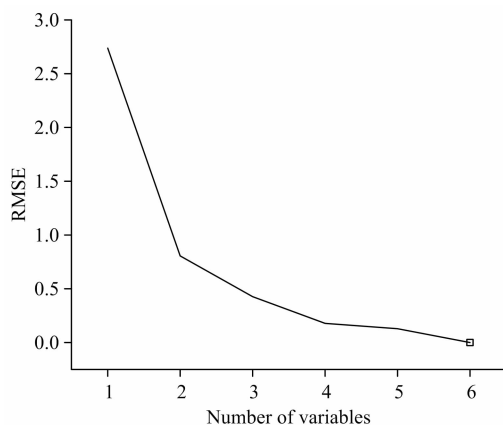


图 8 蓝边最大值方法监测结果与叶片中 Cu²⁺ 含量的变化关系

Fig. 8 Relation between monitoring results based on the maximum of blue edge and Cu²⁺ content in corn leaves

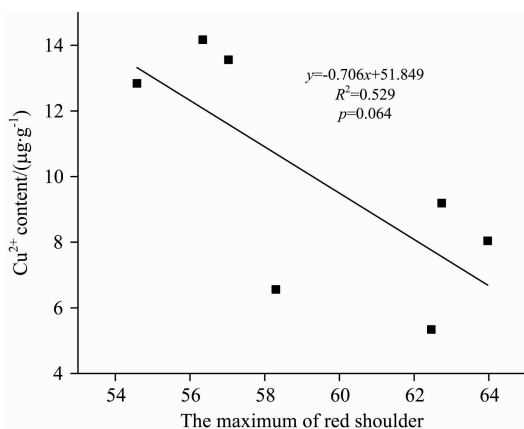


图 9 红肩最大值方法监测结果与叶片中 Cu²⁺ 含量的变化关系

Fig. 9 Relation between monitoring results based on the maximum of red shoulder and Cu²⁺ content in corn leaves

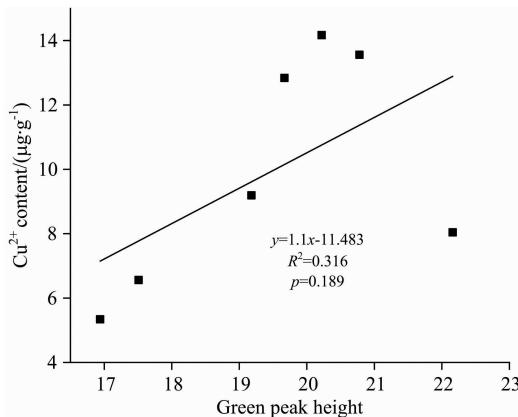


图 10 绿峰高度方法监测结果与玉米叶片中 Cu²⁺ 含量的变化关系

Fig. 10 Relation between monitoring results based on the green peak height and Cu²⁺ content in corn leaves

表 4 SPA 选择的特征波长与玉米叶片中 Cu²⁺ 含量的预测表达式

Table 4 The prediction expression of Cu²⁺ content in corn leaves and characteristic wavelengths selected by SPA

特征波长/nm	预测表达式	R ²	p 值
385	$y = -0.780x + 14.409$	0.041	0.663
651	$y = 1.643x - 4.889$	0.311	0.193
715	$y = 0.378x - 2.238$	0.040	0.666
832	$y = -0.737x + 52.950$	0.546	0.058
1 472	$y = -0.698x + 25.276$	0.103	0.482
2 499	$y = -2.260x + 25.112$	0.360	0.154

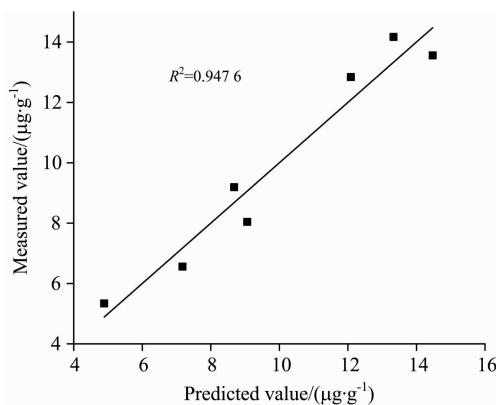


图 11 玉米叶片中 Cu²⁺ 含量的多变量模型预测精度

Fig. 11 Accuracy of the multivariable model on predicting Cu²⁺ content in corn leaves

低, 改变细胞结构。植物叶片中叶绿素含量的变化会引起可见光和近红外波段的反射率变化^[15]。选择数据采集时测得的叶绿素浓度相对值、特征波段 1 266~1 429 和 1 836~1 913 nm 分别计算得到的模型奇异熵作为参数, 结合偏最小二乘回归分析预测玉米叶片受重金属污染下的 Cu²⁺ 含量。经验证可得预测值与实测值的相关性更强, 如图 11 所示, 其

决定系数 R^2 为 0.947 6。与单变量的模型应用结果相比, 多变量的 IWD-Hankel-SVD 模型具有较强的鲁棒性、稳健性。

4 结 论

重金属 Cu^{2+} 胁迫引起的玉米叶片光谱特征变化难以依据光谱曲线形态甄别其污染程度。提出了 IWD 概念和方法, 并通过构造 IWD-Hankel-SVD 模型对玉米叶片高光谱数据的 Cu 污染信息进行反演, 甄别和预测玉米叶片受 Cu 污染的程度, 认为: (1) 对 Cu 污染的玉米叶片光谱数据进行 IWD 处理, 分解提取到的固有旋转分量 PRC1 能较好地保留了原始输入光谱信息, 有效避免了模态混叠现象; (2) 通过特征

波段 1 266~1 429 和 1 836~1 913 nm 求得的模型奇异熵与玉米叶片中的 Cu^{2+} 含量存在很强的相关性, 模型预测的决定系数 R^2 均高达 0.9 左右。通过一些传统预测方法和由 SPA 选择的特征波段的预测结果比较, 验证了依据所提取特征波段这单一变量形式的 IWD-Hankel-SVD 模型在玉米叶片 Cu 污染信息和污染程度方面具有一定的优越性和较好的解释能力; (3) 选择玉米叶片中叶绿素浓度相对值以及叶片光谱特征波段 1 266~1 429 和 1 836~1 913 nm 分别计算得到的模型奇异熵作为参数, 形成多变量的模型预测技术, 可得到叶片中 Cu 含量的预测值与实测值之间相关性更强, 其决定系数 R^2 达到 0.9476。与单变量模型相比, 多变量模型具有较强的鲁棒性和稳健性。

References

- [1] YANG Ke-ming, ZHANG Wei, FU Ping-jie, et al(杨可明, 张伟, 付萍杰, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(7): 2228.
- [2] Toth G, Hermann T, Da Silva MR, et al. Environment International. 2016, 88: 299.
- [3] Shi T, Wang J, Chen Y, et al. International Journal of Applied Earth Observation and Geoinformation, 2016, 52: 95.
- [4] LI Jing, YANG Chao-yuan, YIN Shou-qiang, et al(李晶, 杨超元, 殷守强, 等). Journal of China Coal Society(煤炭学报), 2019, 44(12): 3676.
- [5] Liu M L, Wang T J, Skidmore A K, et al. Science of the Total Environment, 2018, 637: 18.
- [6] QIAO Xiao-ying, MA Shao-yang, HOU Hui-fang, et al(乔晓英, 马少阳, 候会芳, 等). Journal of Safety and Environment(安全与环境学报), 2018, 18(1): 335.
- [7] FU Ping-jie, YANG Ke-ming, WANG Xiao-feng, et al(付萍杰, 杨可明, 王晓峰, 等). Science Technology and Engineering(科学技术与工程), 2018, 18(23): 134.
- [8] ZHU Ye-qing, QU Yong-hua, LIU Su-hong, et al(朱叶青, 屈永华, 刘素红, 等). Journal of Remote Sensing(遥感学报), 2014, 18(2): 335.
- [9] Maisto G, Santorufo L, Arena C. Journal of Plant Nutrition and Soil Science, 2013, 176: 776.
- [10] Frei M G, Osorio I. Proceedings of the Royal Society A-Mathematical Physical and Engineering Sciences, 2007, 463(2078): 321.
- [11] WANG Zhi-hong, FAN Yu-gang, HUANG Guo-yong(王之宏, 范玉刚, 黄国勇). Journal of Yunnan University(云南大学学报·自然科学版), 2018, 40(2): 228.
- [12] ZHANG Bo, LI Jian-jun(张波, 李健君). Journal of Vibration and Shock(振动与冲击), 2009, 28(2): 162.
- [13] LONG Ying, SU Yan-chen, LI Yan-ping, et al(龙莹, 苏燕辰, 李艳萍, 等). China Measurement & Test(中国测试), 2018, 44(5): 24.
- [14] Liu S, Liu X, Hou J, et al. Science in China Series E-Technological Sciences, 2008, 51: 202.
- [15] Delegido J, Verrelst J, Alonso L, et al. Sensors, 2011, 11: 7063.

Predicting the Copper Pollution Information of Corn Leaves Spectral Based on an IWD-Hankel-SVD Model

ZHANG Jian-hong, YANG Ke-ming*, HAN Qian-qian, LI Yan-ru, GAO Wei

College of Geoscience and Surveying Engineering, China University of Mining and Technology-Beijing, Beijing 100083, China

Abstract The environmental problems caused by heavy metal pollution have been particularly prominent in the regions with the rapid industrialization and urbanization development, especially the heavy metal pollution in agriculture is more concerned by the society. Therefore, it is very important to explore some fast and convenient methods on screening and monitoring heavy metal pollution. As a new technology of monitoring heavy metal pollution, the hyperspectral remote sensing has been paid attention and researched deeply by many scholars. A concept and method of inherent wavelength-scale decomposition (IWD) was proposed in the paper, and an IWD-Hankel-SVD model was established for predicting heavy metal pollution degree of vegetation combined with the Hankel matrix and the singular value decomposition (SVD), here the model was divided into single-variable model and multi-variable model. The single-variable model was mainly used to obtain the intrinsic rotation components (PRC) of spectral information of vegetation polluted by heavy metal through IWD processing and to extract the effective characteristic bands of the best PRC, then it could be realized to predict the heavy metal pollution according to the singular entropy of the model acquired by using SVD to decompose the Hankel matrix constructed based on each characteristic band. But the multi-variable model was used to realize the prediction of heavy metal pollution information by taking the relative values of vegetation chlorophyll concentration and the singular entropy acquired by the single-variable model as parameters. According to the data of leaf spectra, measured chlorophyll concentrations and Cu^{2+} contents in corn leaves polluted by heavy metal Cu^{2+} under different stress gradients, firstly the spectra of corn leaves stressed by the different Cu^{2+} concentrations were analyzed by IWD, the best PRC was obtained which could well retain the original spectral information, and some effective characteristic bands were extracted to be 553~680, 681~780, 1 266~1 429, 1 430~1 631, 1 836~1 913 and 1 914~2 111 nm from the PRC, then the Hankel matrix of each characteristic band was constructed and processed by SVD to obtain the singular entropy of the single-variable model, finally through the correlation analysis between the singular entropy of the model corresponding to each characteristic band and the Cu^{2+} contents in corn leaves, it was found that the determination coefficients R^2 of the singular entropy and the Cu^{2+} contents in the leaves were all about 0.9 computed based on the 1 266~1 429 and 1 836~1 913 nm characteristic bands, the result shows that the two characteristic bands had more advantageous and interpretable for the IWD-Hankel-SVD model on predicting the Cu pollution degrees. At the same time, it was concluded that the multi-variable IWD-Hankel-SVD-model had stronger prediction ability of Cu pollution degrees in corn leaves by using the partial least square regression analysis based on taking the relative values of chlorophyll concentration in corn leaves and the singular entropy of the single-variable model corresponding to 1 266~1 429, 1 836~1 913 nm characteristic bands as parameters, and the determination coefficient R^2 reached 0.947 6, so the multi-variable model was proved to be more robustness and steadiness.

Keywords Spectral analysis; Corn leaf; Heavy metal copper pollution; Intrinsic wavelength-scale decomposition; Prediction model

(Received Apr. 27, 2020; accepted Aug. 15, 2020)

* Corresponding author