

基于污染种类判别特征的作物Pb、Cu污染种类判别

尚相春, 金倩, 杨可明, 高伟, 吴兵

引用本文:

尚相春, 金倩, 杨可明, 高伟, 吴兵. 基于污染种类判别特征的作物Pb、Cu污染种类判别[J]. 农业环境科学学报, 2023, 42(1): 46–54.

在线阅读 View online: <https://doi.org/10.11654/jaes.2022-0444>

您可能感兴趣的其他文章

Articles you may be interested in

基于SD-SVD-Burg的玉米叶片铜铅污染甄别与程度诊断

韩倩倩, 杨可明, 高伟, 李艳茹, 张建红

农业环境科学学报. 2021, 40(1): 35–43 <https://doi.org/10.11654/jaes.2020-0860>

铜胁迫下玉米光谱变化的奇异性诊断指数与污染甄别

李燕, 杨可明, 王敏, 程凤, 高鹏, 张超

农业环境科学学报. 2019, 38(1): 14–21 <https://doi.org/10.11654/jaes.2018-0501>

基于SD-SWT的铜胁迫下玉米光谱奇异性甄别与污染监测

李艳茹, 杨可明, 韩倩倩, 高伟, 张建红

农业环境科学学报. 2020, 39(9): 1869–1877 <https://doi.org/10.11654/jaes.2020-0468>

玉米/大豆间作的镉累积规律初探

李涵, 黄道友, 黄山, 汤彬, 郭欢乐, 曹钟洋, 陈松林, 陈志辉

农业环境科学学报. 2020, 39(9): 1900–1907 <https://doi.org/10.11654/jaes.2020-0401>

铜离子胁迫下玉米叶片污染信息的SM-DMFD探测模型

张伟, 杨可明, 孙彤彤, 王晓峰, 程龙

农业环境科学学报. 2017, 36(9): 1753–1761 <https://doi.org/10.11654/jaes.2017-0536>



关注微信公众号, 获得更多资讯信息

尚相春, 金倩, 杨可明, 等. 基于污染种类判别特征的作物 Pb、Cu 污染种类判别[J]. 农业环境科学学报, 2023, 42(1): 46–54.

SHANG X C, JIN Q, YANG K M, et al. Discrimination of Pb and Cu pollution types in crops based on the discrimination feature of lead–copper pollution type[J]. *Journal of Agro-Environment Science*, 2023, 42(1): 46–54.



开放科学 OSID

基于污染种类判别特征的作物 Pb、Cu 污染种类判别

尚相春¹, 金倩^{2*}, 杨可明³, 高伟³, 吴兵³

(1. 淮北矿业股份有限公司孙疃煤矿, 安徽 淮北 235000; 2. 河北省矿产资源与生态环境监测重点实验室, 河北 保定 071051; 3. 中国矿业大学(北京)地球科学与测绘工程学院, 北京 100083)

摘要:为判别作物中的重金属污染种类,开展了 Pb、Cu 胁迫下的典型作物(玉米)培育实验,获取了作物叶片的高光谱数据与 Pb、Cu 污染信息。首先对光谱进行分数阶、整数阶导数变换,而后利用差值比光谱指数构建特征参量以组成 Pb、Cu 污染种类判别特征(DFLCPT),最终基于 DFLCPT 数据构建了用于作物 Pb、Cu 污染种类判别的随机森林分类(RFC)、K-最邻近分类(KNNC)、支持向量机分类(SVC)、高斯过程分类(GPC)模型。结果表明:在依托多种导数光谱构建的差值比光谱指数(DRSI)中,以 0.9 阶导数光谱为基准的 DRSI[2 412, 1 223, 636]与样本 Pb、Cu 污染种类的相关系数绝对值最大,为 0.764 1;在依托多维度 DFLCPT (DFLCPT_m)建立的 SVC、RFC、KNNC、GPC 玉米 Pb、Cu 污染种类判别模型中,RFC 模型的效果优于 SVC、KNNC、GPC 模型,其在训练集与验证集中取得的最高正确率均为 100%,精度较好,稳定性较强。研究表明,依托 DFLCPT 的模型在 Pb、Cu 污染种类判别中达到了预期效果,可为大规模的作物重金属污染种类判别提供一定的技术支撑。

关键词:高光谱;重金属污染;种类判别;玉米叶片

中图分类号:X71 文献标志码:A 文章编号:1672–2043(2023)01–0046–09 doi:10.11654/jaes.2022–0444

Discrimination of Pb and Cu pollution types in crops based on the discrimination feature of lead–copper pollution type

SHANG Xiangchun¹, JIN Qian^{2*}, YANG Keming³, GAO Wei³, WU Bing³

(1. Suntuan Coalmine, Huaibei Mining Co., Ltd., Huaibei 235000, China; 2. Key Laboratory of Mineral Resources and Ecological Environment Monitoring of Hebei Province, Baoding 071051, China; 3. College of Geoscience and Surveying Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China)

Abstract: To discriminate the heavy metal pollution types present in crops, a typical crop (maize) was cultivated under Pb and Cu stress, following which hyperspectral data containing Pb and Cu pollution information were obtained from the crop leaves. Fractional–order and integer–order derivative transformations were first applied to the spectra. Then, difference ratio spectral indexes (DRSIs) were used to construct feature covariates for forming the discrimination feature of lead–copper pollution type (DFLCPT). Finally, random forest classification (RFC), K–nearest neighbor classification (KNNC), support vector machine classification (SVC), and Gaussian process classification (GPC) models were constructed on the basis of the DFLCPT data for discrimination of the Pb and Cu pollution type in the crop. It was found that among the DRSIs constructed from a variety of derivative spectra, those [2 412, 1 223, 636] based on the 0.9

收稿日期:2022–05–03 录用日期:2022–06–09

作者简介:尚相春(1984—),男,高级工程师,主要从事矿山测量管理工作。E-mail:284575990@qq.com

*通信作者:金倩 E-mail:13373121110@163.com

基金项目:淮北矿业企业委托项目(2020–117);中央高校基本科研业务费专项(2022JCCXDC01, 2022YJSDC22);国家自然科学基金项目(41971401)

Project supported: Entrusted project of Huaibei Mining Enterprise (2020–117); The Fundamental Research Funds for the Central Universities (2022JCCXDC01, 2022YJSDC22); The National Natural Science Foundation of China (41971401)

derivative spectrum had the largest absolute value of the correlation coefficient between the Pb and Cu pollution types of the sample, which is 0.764 1. Among the four classification models established on the basis of multidimensional DFLCPT (DFLCPT_{nd}) data, the RFC model had a better effect than the SVC, KNNC, and GPC models. The highest accuracy of the RFC model for the training set and verification set was 100%, with good accuracy and strong stability. The results show that the DFLCPT-based model has achieved the expected effect in the discrimination of Pb and Cu pollution types and can provide technical support for the discrimination of heavy metal pollution types in crops grown on a large scale.

Keywords: hyperspectral; heavy metal pollution; pollution type discrimination; maize leave

近百年来,随着人类文明的发展和工业化进程的加速,矿山无序开采、矿石违规冶炼、超标废水废气排放、农药化肥过量使用等现象屡见不鲜,导致大量重金属流入到水土环境中,影响了重金属的全球循环^[1-3]。超过安全限值的重金属会破坏生态系统的正常功能,它们难以被生物降解,会通过根部及叶面气孔进入植物阻碍其生长^[4],且极有可能沿食物链向农作物与动物传导,进而在人类体内富集,从多方面对人体造成损害^[5-6]。由此可见,重金属污染严重威胁着全球粮食安全与人类健康,已经成为当今世界面临的重大生态环境问题之一。Pb和Cu是常见的重金属元素,但它们对于生命体的作用与危害差异极大。Pb属于生命体的非必需金属元素,较低含量便会对生物体表现出严重毒性^[7],环境中Pb超标则影响植物生长、产量和品质,重则导致植物死亡;而适量Cu对植物的生长有促进作用,只有过量Cu才会对生命体产生有害影响^[8];此外,Pb和Cu诱发的疾病也不尽相同,Pb中毒会造成正常红细胞性贫血等^[9-10]疾病,Cu中毒则会造成多种神经系统退行性疾病^[11-13]。由此可见,不同种类、不同程度的重金属污染产生的危害存在差异,而作物是环境中重金属沿食物链向人体转移的重要环节,因此鉴别作物的重金属污染种类尤为重要。

高光谱遥感技术是一种新兴的科学观测手段,相较于传统的化验分析方法,其获取的数据更加丰富多样,监测能力与效率更强,同时也更加绿色环保^[14]。近年来,高光谱数据凭借其独特优势,在土壤成分监测、植物要素监测、水土环境监测等方面发挥了重要作用。在作物重金属污染监测中,较多研究聚焦于重金属含量的反演预测,如RATHOD等^[15]在研究过程中使用了水分胁迫指数,计算其与叶片砷含量之间的相关系数,探讨植物生长状态,建立回归模型,为监测植物健康状况提供了一种新思路。HEDE等^[16]探究了植被受重金属污染的影响程度,通过与常规方法进行比较,结果表明所提出的植被指数对重金属胁迫的灵

敏度更高。HAN等^[17]对作物叶片光谱进行了奇异值分解,而后将获得的能量值作为作物自适应模糊神经网络模型的输入层,进行了叶片Pb含量的反演。BANNERJEE等^[18]通过构建植被健康指数,分析植被健康与其受重金属污染的关系。MIRZAEI等^[19]以采集的葡萄叶片光谱数据为研究对象,采用偏最小二乘法、多元线性回归以及支持向量机方法进行建模,预测叶片内的重金属含量并取得一定效果。

随着作物重金属污染种类鉴别意义的凸显,该问题引发了部分研究人员的关注,如LI等^[20]将变分模态分解应用于光谱曲线分解,联合主成分分析算法构建二维分类平面,利用支持向量机中的超分类线对玉米Pb、Cu污染进行了可视化分析,并取得了较好的效果。FU等^[21]绘制了重金属污染环境作物叶片的功率谱密度曲线,进而依据曲线形态区分了重金属污染种类。LI等^[22]构建了区分Pb、Cu污染的光谱指数,通过朴素贝叶斯算法制定区分玉米叶片Pb、Cu污染的判别规则。上述研究侧重于从视觉角度进行种类鉴别,虽然结果清晰,但不适用于海量数据,不能转换为计算机语言进行结果统计,考虑到作物重金属污染种类判别的最终目标是进行区域级的应用,故应探寻具备大面积应用潜力的作物重金属污染种类判别方法。

本研究拟依托于Pb、Cu污染下玉米叶片光谱数据及对应的重金属污染种类信息,通过分数阶、整数阶的导数(Derivative, D)进行光谱处理,借助差值比光谱指数(Difference ratio spectral index, DRSI)结构生产指数参量,进而基于优势指数参量构建多维Pb、Cu污染种类判别特征(DFLCPT),最终以DFLCPT为驱动,结合随机森林分类(Random forest classification, RFC)、K-最邻近分类(K-nearest neighborhood classification, KNNC)、支持向量机分类(Support vector classification, SVC)、高斯过程分类(Gaussian process classification, GPC)原理,开展作物Pb、Cu污染种类判别研究,以期为工业级的重金属污染种类判别实践提供技术支持或参考。

1 材料与方法

1.1 试验与数据

1.1.1 植株培育

为满足作物 Pb、Cu 污染高光谱甄别研究的数据需求,在温室大棚内开展重金属胁迫下的玉米培育实验。实验从重金属胁迫种类、重金属胁迫程度两个角度出发,进行玉米植株差异化培育。通过对穗期玉米叶片的测量、检验,获取不同种类、不同程度重金属胁迫下的玉米叶片高光谱数据及 Pb、Cu 含量数据,用于后续研究。其中,Pb、Cu 污染源为 $\text{Pb}(\text{NO}_3)_2$ 、 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$,胁迫程度依据《土壤环境质量标准》(GB 15618—1995) 设置为 50、100、150、200、300、400、600、800 $\text{mg} \cdot \text{kg}^{-1}$,对应标记为 Pb(50)、Cu(50)、Pb(100)、Cu(100)……Pb(800)、Cu(800)。同一污染种类下各胁迫程度的土壤均分为三盆,每盆中培育一组作物,对应标记为 Pb(50)a、Pb(50)b、Pb(50)c、Cu(50)a、Cu(50)b、Cu(50)c……Cu(800)c,共计 48 组。培育过程中各植株的水、气和养分条件适宜且一致。

1.1.2 数据获取

在密闭暗室中,以 50 W 卤素灯为光源,利用 ASD FieldSpec 4 型光谱仪采集叶片光谱,光纤视场角为 25° ,采集时与叶片间距为 5 cm。每盆植株均选择三枚不同位置的叶片用于光谱采集,且每枚叶片均重复测量三次光谱,以单一植株多个叶片多次光谱的均值表示该植株的叶片光谱。其中,各叶片对应标记为 Pb(50)a1、Pb(50)a2、Pb(50)a3、Pb(50)b1、Pb(50)b2、Pb(50)b3……Cu(800)c1、Cu(800)c2、Cu(800)c3。

对叶片进行化学消解等预处理,而后利用电感耦合等离子发射光谱仪(ICP-OES)测定叶片中的 Pb、Cu 含量。植物的生长培育环境除栽培土污染种类与程度不同外,其余条件均一致,故在重金属污染种类与程度相同的前提下,植株叶片同种重金属含量应在同一范围内。依据四分位距法,筛查异常叶片,基于 Pb 含量的筛查结果如图 1,基于 Cu 含量的筛查结果见图 2。依据异常值筛查结果,剔除异常叶片,更新用于研究的相关数据。不含异常叶片的 48 组植株样本光谱见图 3,不含异常叶片的叶片 Pb、Cu 含量见表 1,基于剔除异常叶片后的数据集开展后续研究。

1.2 理论与方法

1.2.1 导数

导数变换是一项应用广泛的光谱处理技术,在消除噪声、获取隐藏信息等方面发挥着重要作用^[23]。传

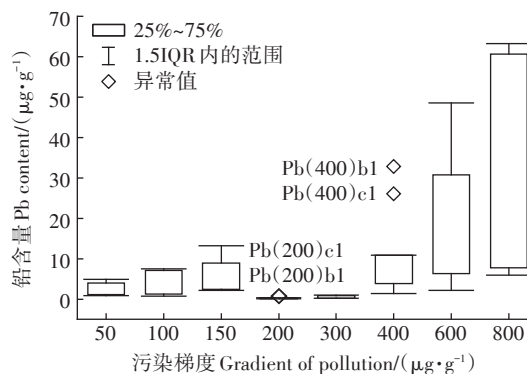


图1 基于Pb含量的异常叶片筛查结果

Figure 1 Abnormal leaf screening results based on Pb content

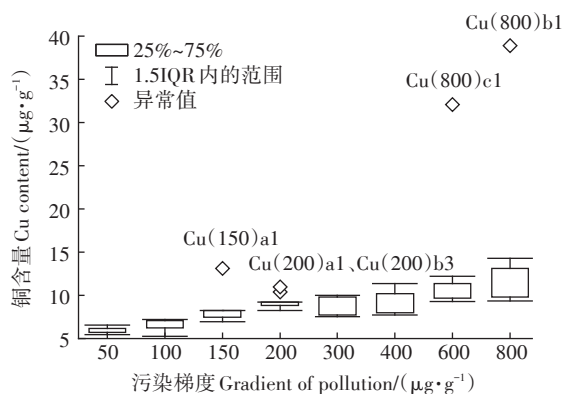


图2 基于Cu含量的异常叶片筛查结果

Figure 2 Abnormal leaf screening results based on Cu content

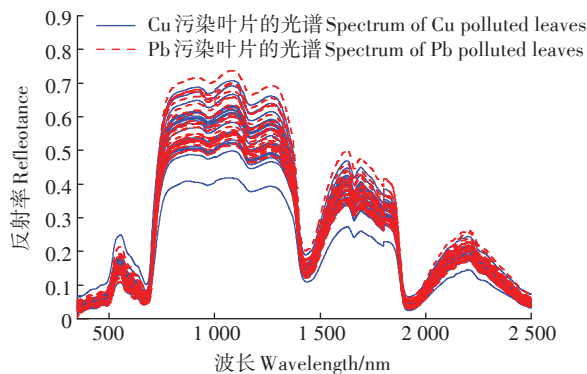


图3 剔除异常样本后的叶片光谱

Figure 3 Leaf spectra after the removal of abnormal samples

统导数的阶次一般为整数,而分数阶导数是传统导数的拓展,其在光谱信号中常依据 Grünwald-Letnikov (G-L)法^[24-25]实现。由 G-L 法可知,光谱 $f(x)$ 在波长 $[b_{\text{start}}, b_{\text{end}}]$ 间的 v 阶导数为:

表1 剔除异常样本后的叶片Pb、Cu含量(mg·kg⁻¹)Table 1 Pb and Cu contents in leaves after the removal of abnormal samples(mg·kg⁻¹)

污染程度 Pollution level	Cu 含量 Cu content	Pb 含量 Pb content
0	5.34±0.92	0.74±0.37
50	5.92±0.34	2.28±1.63
100	6.56±0.60	3.60±2.71
150	7.70±0.50	5.70±3.96
200	8.85±0.37	0.15±0.07
300	8.76±0.95	0.46±0.30
400	9.19±1.30	5.96±3.39
600	10.43±0.97	17.36±17.07
800	11.08±1.63	32.16±22.59

$$d^v f(x) = \lim_{h \rightarrow 0} \frac{1}{h^v} \sum_{m=0}^{INT[(b_{\text{end}} - b_{\text{start}})/h]} (-1)^m \frac{\Gamma(v+1)}{m!(v-m+1)!} f(x-mh) \quad (1)$$

其中Gamma变换为:

$$\Gamma(z) = \int_0^\infty \exp(-u) u^{z-1} du = (z-1)! \quad (2)$$

式中: v 为阶次,且大于0; b_{start} 为光谱的起始波长; b_{end} 为光谱的终止波长; $INT()$ 为取整运算; Γ 为Gamma函数; x 为指定波段的光谱值; h 为光谱的采样间隔; z 为自变量; u 为积分变量。

故而,式(1)可以表示为:

$$\begin{aligned} \frac{d^v f(x)}{dx^v} &\approx f(x) + (-v)f(x-1) + \frac{(-v)(-v+1)}{2} f(x-2) + \\ &\cdots + \frac{\Gamma(v+1)}{m!(v-m+1)!} f(x-m) \end{aligned} \quad (3)$$

1.2.2 Pb、Cu污染种类判别特征(DFLCPT)

为开展Pb、Cu污染种类判别,设计了一种多个要素(E)构成的特征集以表征样本的种类属性,即DFLCPT。依据包含要素的数量,可分为不同维度的DFLCPT。

n 维尺度的DFLCPT(DFLCPT_{nd})可表示为:

$$\text{DFLCPT}_{nd} = (E_1, E_2 \cdots E_n) \quad (4)$$

式中: $E_1, E_2, \cdots, E_n$ 为DFLCPT_{nd}构成要素。

适宜的光谱指数能够有效利用多波段信息,进一步提升对目标的表征能力,故以光谱指数作为DFLCPT_{nd}的构成要素。差值和比值是两种常见的光谱指数形式,相关性分析能够实现面向属性的信息提取,故结合两种指数的结构优势,加入相关性最优波,构建了差值比光谱指数(DRSI)结构。DRSI的具体参数受基准光谱(Basic spectrum, BS)、构成波段的影响,因此将DRSI记作DRSI[$b1, b2, br$]BS, DRSI[$b1, b2, br$]

BS的计算方法为:

$$\text{DRSI}[b1, b2, br] \text{BS} = (R_{br} - R_{b1}) / (R_{br} - R_{b2}) \quad (5)$$

式中: br 为相关性最强波段波长, R_{br} 为波长 br 处的光谱值。

1.2.3 判别模型原理

(1)随机森林分类(RFC)

随机森林是一种机器学习技术,决策树是其基分类器,该算法将随机因素引入到决策树的训练中,其主要思路:首先将预测空间分为多个样本空间,然后再从各个决策树中选取最优特征分割点,选取时从每一子空间随机选取部分特征,构建一个特征数据集,最终汇总获取结果。由于决策树的训练集及候选集均为独立选取,因此其预测结果具有独立性。随机森林的分类结果是由所有决策树联合表决确定的,与单一决策树相比,其具有不过度拟合、所需参数少、泛化能力较强等特点,因此随机森林的性能更佳。

(2)K-最邻近分类(KNNC)

最邻近分类算法是基于样本的一种分类算法,其主要研究重心为提高模型的分类精度,其主要操作过程:确定待分类的对象,首先寻找训练样本与待分类对象距离较近的样本,然后计算相似度,并将相似度进行排名,在排列结果中,选择最相似的K个样本,K的取值在很大程度上依赖于训练样本的大小,然后计算出K个样本的所属类别,从而决定出该测试样本的类别。

(3)支持向量机分类(SVC)

支持向量机算法可应用于样本分类及回归等问题处理方面,是一种按照监督分类方式进行分类的机器学习方法,其基本原理是首先选定训练集与测试集,并确定核函数,然后选取样本的特征信息,映射到高维空间,在该空间寻找一个最优超平面,从而在数据特征之间创建边界,最终达到分类的目的。SVM算法的优点是可以应对样本较少的情况,具有对样本数据敏感的特点。

(4)高斯过程分类(GPC)

高斯过程分类与SVM分类操作过程类似,同样是基于核函数确立的算法模型,其属于贝叶斯机器学习模型,被用来解决二分类等问题研究,GPC的基本建立过程:首先依托训练样本,划分分类级次;其次进行训练学习,确定似然函数,求解最优参数;而后基于高斯理论及贝叶斯规则,综合先验知识,调整预测分布;最终得到分类结果。

2.4 随机搜索优化(RSO)

网格搜索是机器学习与深度学习领域重要的超

参数优化方法,但其在非关键参数维度消耗了大量的算力。RSO是一种设计更加便捷的参数搜索方法,其将原参数空间划分为规则的网格,而后从中均匀且独立地抽取,以生产新的参数训练集。RSO具备GS的所有优势(原理简单、易于实现、并行性小),并以低维搜索空间中效率的小幅降低换取了高维搜索空间中效率的大幅提高。

2 结果与讨论

2.1 光谱导数变换结果及分析

对叶片原始光谱(OS)进行不同阶次导数(D)处理,阶次分为整数阶与分数阶,使用的整数阶次为1~10,使用的分数阶次为0.1~0.9、1.1~1.9。不同阶次背景下生成的28种导数光谱,在本研究中记作“vD”光谱,v表示D的阶次。

分析28种导数光谱及OS与样本Pb、Cu污染种类的相关性,统计光谱各波段与样本Pb、Cu污染种类相关系数绝对值(|r|)的均值(r_{mean})与最大值(r_{max}),以及 r_{max} 对应的波长(b_r),见表2。从与样本Pb、Cu污染种类的 r_{mean} 来看,分数阶导数变换与整数阶导数变换均达到了预期效果,相较于OS凸显了更有价值的信息,整体上增强了数据与研究目标的相关程度;从与样本重金属污染种类的 r_{max} 来看,分数阶导数变换与整数阶导数变换同样均达到了预期效果,提取的最有价值信息相对于OS更具潜力,增强了最优数据与研究目标的相关程度,但随着导数阶次的增长,光谱对应的 r_{max} 并未一直增加。

2.2 DFLCPT构建结果

基于不同阶次的导数光谱与对应的 b_r 信息构建

DRSI,共得到28类BS不同的DRSI,同一BS下,共有4 624 650种组成波段不同的DRSI,即生产了129 490 200种DRSI。分析每种DRSI与样本Pb、Cu污染种类的相关性,以DRSI中 $b1$ 、 $b2$ 波长值为自变量,以对应的|r|为因变量进行制图,见图4。可见,相较于其他阶次导数光谱,0.5~1.2D光谱生产的DRSI与样本Pb、Cu污染种类相关系数整体上更高,对污染种类表征能力更强。

为准确获取能力最强的DRSI,统计多种DRSI与样本Pb、Cu污染种类的|r|,最大的10个|r|及对应的DRSI构成参数见表3。对Pb、Cu污染种类的表征效果而言,基于分数阶导数光谱的DRSI相较于基于整数阶导数光谱的DRSI更好,说明导数阶次并非越高越好,阶次的分数化在适当的环境中能获得更好的效果。

统筹分析表3中相关数据,依据对研究目标表征能力的强弱,依两种指数构型下各指数对研究目标表征能力的强弱顺序,逐个纳为DFLCPT_{nd}的构成要素。

2.3 基于DFLCPT_{nd}甄别模型的建立与检验

以DFLCPT_{nd}数据为驱动,结合SVC、RFC、KNNC、GPC理论分别构建Pb、Cu污染甄别模型。DFLCPT_{nd}可细化为多种不同维度的特征数据,也意味驱动甄别模型的数据维度可能是不同的。随着DFLCPT_{nd}维度的上升,其各要素中有益信息的占比存在升高或降低的可能,进而对甄别模型的性能产生影响。因此,分析不同维度DFLCPT_{nd}数据驱动下的甄别模型性能是必要的,故分别选取多个维度的DFLCPT_{nd}为驱动数据,进行甄别模型构建,其中选取的DFLCPT_{nd}数据维度为DFLCPT_{1D}、DFLCPT_{11D}、DFLCPT_{21D}、DFLCPT_{31D}、DFL-

表2 不同类型光谱与样本Pb、Cu污染种类的相关程度表

Table 2 Correlation between multiple types of spectra and types of Pb and Cu pollution in samples

项目 Item	光谱名称 Name of the spectrum									
	OS	0.1D	0.2D	0.3D	0.4D	0.5D	0.6D	0.7D	0.8D	0.9D
r_{mean}	0.08	0.09	0.10	0.10	0.11	0.11	0.12	0.12	0.12	0.12
r_{max}	0.17	0.18	0.21	0.21	0.24	0.31	0.37	0.42	0.51	0.54
b_r/nm	699	1 801	1 801	1 801	614	645	642	636	636	636
	1.0D	1.1D	1.2D	1.3D	1.4D	1.5D	1.6D	1.7D	1.8D	1.9D
r_{mean}	0.11	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
r_{max}	0.55	0.54	0.52	0.49	0.49	0.49	0.49	0.48	0.48	0.47
b_r/nm	636	636	636	1 205	1 205	1 205	1 205	1 205	1 205	1 205
	2.0D	3.0D	4.0D	5.0D	6.0D	7.0D	8.0D	9.0D	10.0D	
r_{mean}	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	
r_{max}	0.46	0.47	0.45	0.45	0.44	0.46	0.48	0.49	0.50	
b_r/nm	728	1 206	729	549	771	772	772	773	773	

CPT_{41D}、DFLCPT_{51D}、DFLCPT_{61D}、DFLCPT_{71D}、DFLCPT_{81D}、DFLCPT_{91D}。部分建模算法中的参数对性能影响较大,同时考虑驱动数据维度的影响,故利用RSO算法求取各模型参数的动态最优解。

以训练集样本对应的数据为自变量,训练不同维度DFLCPT_{nd}数据为驱动下的甄别模型,获得模型在训练集中的正确率;而后结合验证集数据进行检验,获得模型在验证集中的正确率。由表4可知,RFC模型效果最优时,对应的训练集与验证集正确率均为100%,SVC模型效果最优时,对应的训练集、验证集正

确率均分别为89%、100%。GPC模型效果最优时,训练集和验证集正确率分别为92%、100%;KNNC模型效果最优时,训练集和验证集正确率为100%、90%或89%、100%。综上所述,基于DFLCPT_{nd}数据的RFC模型在Pb、Cu污染甄别中精度较好,且稳定性较强。

3 结论

(1)叶片光谱的分数阶导数变换、整数阶导数变换均增强了其与Pb、Cu污染种类的相关程度,但相关程度的峰值并未随着阶次的增长而增长。

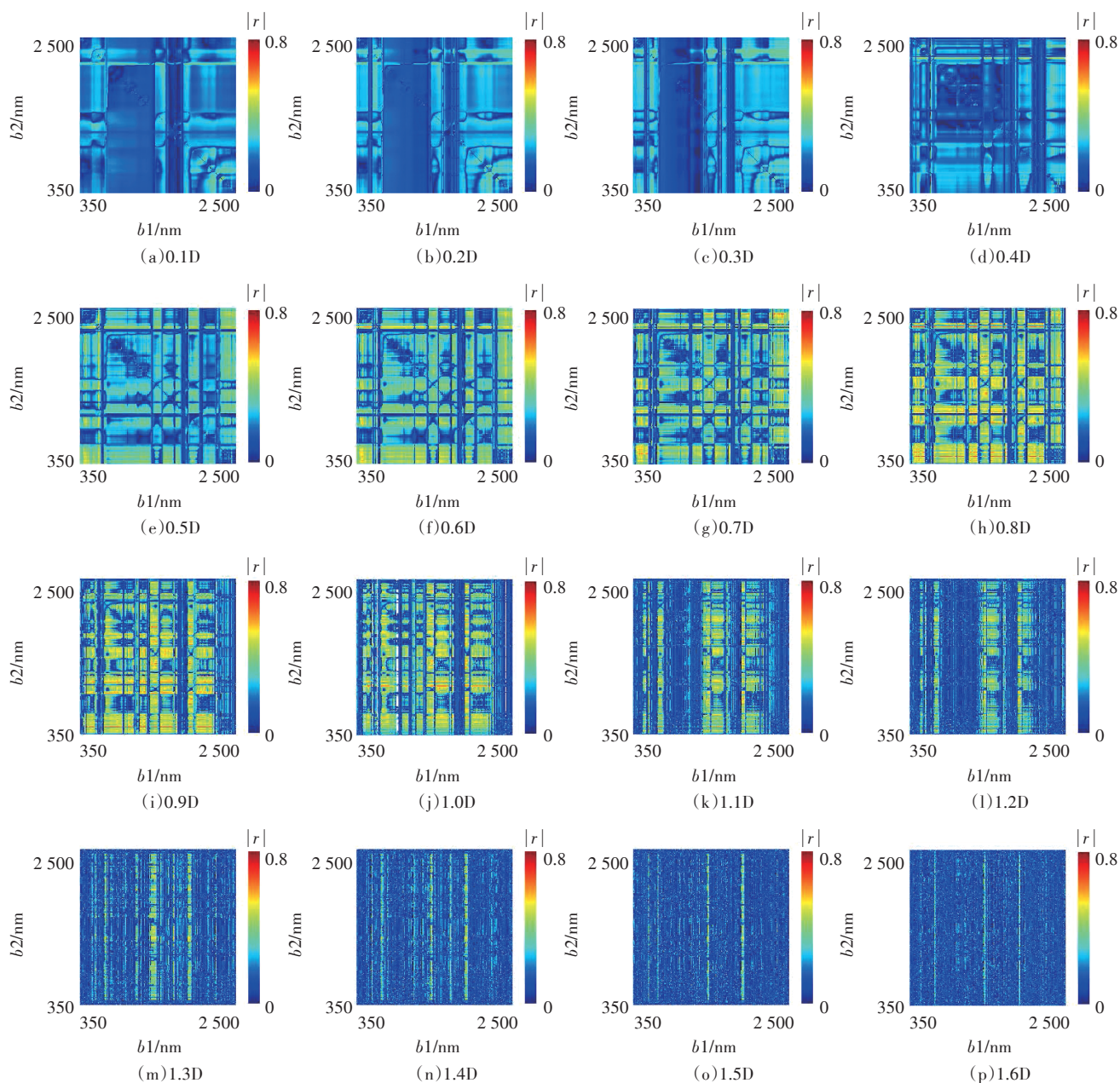
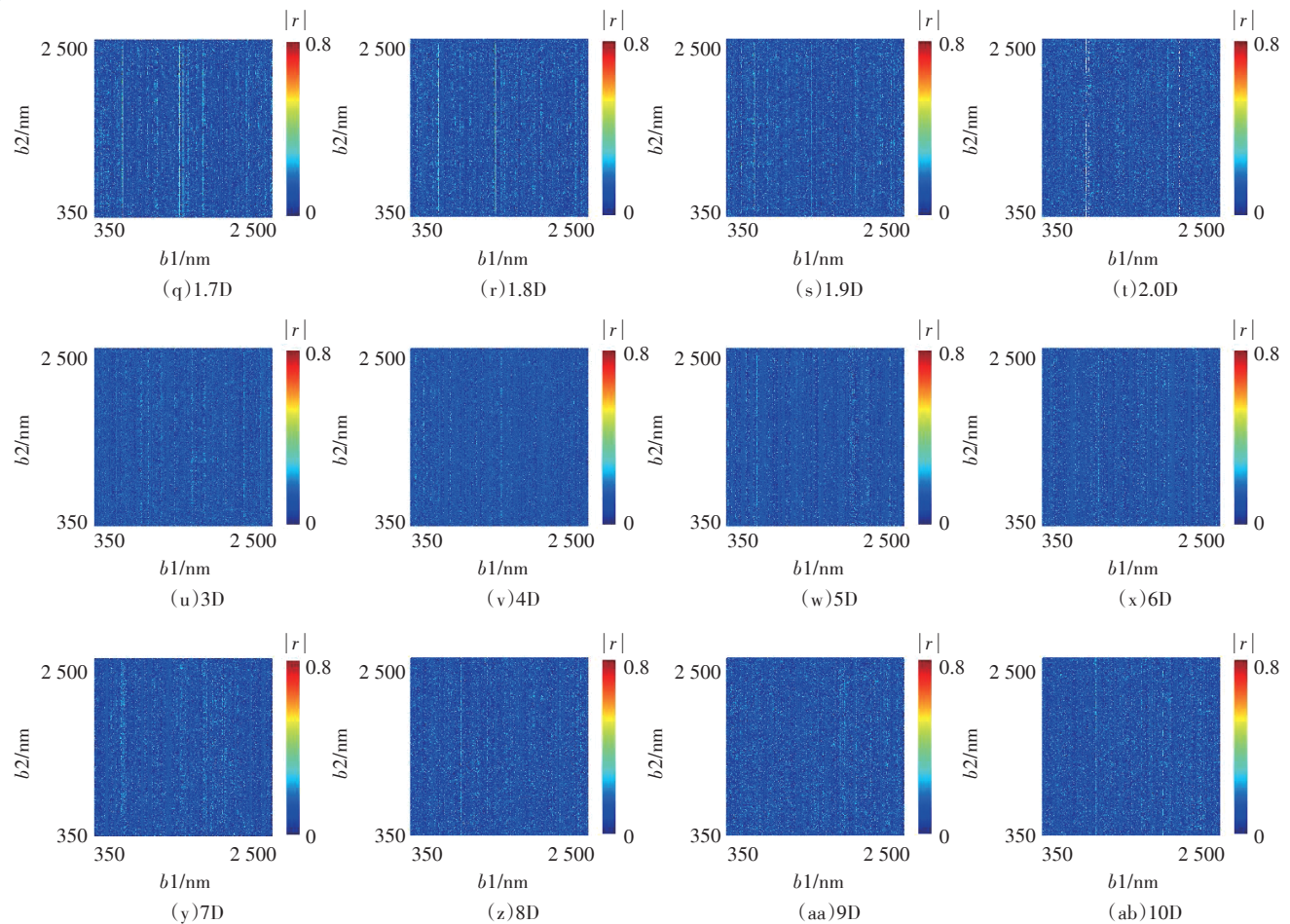


图4 各DRSI与样本Pb、Cu污染种类的相关性

Figure 4 Correlation between each DRSI and the Pb and Cu pollution types in the sample



续图4 各DRSI与样本Pb、Cu污染种类的相关性

Continued figure 4 Correlation between each DRSI and the Pb and Cu pollution types in the sample

表3 优势DRSI的性能及构成参数

Table 3 Performance and composition parameters of advantageous DRSI

序号 Number	DRSI[b_1, b_2, b_r]BS	r
1	DRSI [2 412, 1 223, 636]0.9D	0.764 1
2	DRSI [2 398, 984, 636]0.9D	0.751 1
3	DRSI [2 412, 1 205, 636]0.8D	0.747 2
4	DRSI [2 412, 995, 636]0.9D	0.747 1
5	DRSI [2 398, 2 145, 636]0.8D	0.747 0
6	DRSI [2 412, 547, 636]0.9D	0.746 5
7	DRSI [2 398, 2 145, 636]0.9D	0.746 1
8	DRSI [2 412, 984, 636]0.9D	0.742 4
9	DRSI [2 412, 1 205, 636]0.9D	0.741 8
10	DRSI [873, 767, 636]1.0D	0.741 6

(2)在构建的DRSI中,DRSI[2 412, 1 223, 636]0.9D与样本Pb、Cu污染种类的 $|r|$ 最大,为0.764 1;以分数阶导数光谱为基准光谱的指数对Pb、Cu污染种

类的表征能力更强。

(3)基于DFLCPT_{nb}的RFC污染判别模型的效果优于SVC、KNNC、GPC模型,其在训练集与验证集中取得的最高正确率均为100%,精度较好,稳定性较强。

参考文献:

[1] GUO H, YANG K M, CHENG L, et al. Using the characteristic parameters of Hilbert marginal spectrum for indirectly estimating copper content in maize leaves under copper stress[J]. *Remote Sensing Letters*, 2019, 10(11): 1067–1076.

[2] KLADSOMBOON S, JAIYEN C, CHOPRATHUMMA C, et al. Heavy metals contamination in soil, surface water, crops, and resident blood in Uthai District, Phra Nakhon Si Ayutthaya, Thailand[J]. *Environmental Geochemistry and Health*, 2020, 42(2): 545–561.

[3] RAI P K, LEE S S, ZHANG M, et al. Heavy metals in food crops: Health risks, fate, mechanisms, and management[J]. *Environment International*, 2019, 125: 365–385.

[4] WANG F H, GAO J, ZHA Y. Hyperspectral sensing of heavy metals in

表4 各判别模型在训练集和验证集中的正确率

Table 4 Correct rate of each discrimination model in the training set and verification set

模型 Model	<i>n</i>	训练集正确率/% Correct rate of training set	验证集正确率/% Correct rate of validation set	模型 Model	<i>n</i>	训练集正确率/% Correct rate of training set	验证集正确率/% Correct rate of validation set
RFC	1	87	90	SVC	1	84	90
	11	100	100		11	89	90
	21	92	90		21	89	100
	31	100	90		31	87	80
	41	97	90		41	89	100
	51	97	90		51	87	100
	61	95	90		61	84	90
	71	100	100		71	87	90
	81	92	100		81	87	100
	91	95	80		91	89	100
GPC	1	84	90	KNNC	1	84	80
	11	87	100		11	100	90
	21	87	100		21	87	100
	31	89	100		31	87	100
	41	89	100		41	87	100
	51	89	100		51	87	100
	61	89	100		61	89	100
	71	89	100		71	89	100
	81	89	100		81	84	100
	91	92	100		91	87	100

soil and vegetation; Feasibility and challenges[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 136:73–84.

- [5] LIU M L, WANG T J, SKIDMORE A K, et al. Heavy metal-induced stress in rice crops detected using multi-temporal Sentinel-2 satellite images[J]. *Science of the Total Environment*, 2018, 637:18–29.
- [6] 李艳茹, 杨可明, 韩倩倩, 等. 基于SD-SWT的铜胁迫下玉米光谱奇异性甄别与污染监测[J]. *农业环境科学学报*, 2020, 39(9):1869–1877. LI Y R, YANG K M, HAN Q Q, et al. Spectral singularity identification and pollution monitoring of corn under copper stress based on SD-SWT[J]. *Journal of Agro-Environment Science*, 2020, 39(9):1869–1877.
- [7] HARGUINTEGUY C A, COFRE M N, FERNANDEZ-CIRELLI A, et al. The macrophytes *Potamogeton pusillus* L. and *Myriophyllum aquaticum* (Vell.) Verdc. as potential bioindicators of a river contaminated by heavy metals[J]. *Microchemical Journal*, 2016, 124:228–234.
- [8] SHAHID M, DUMAT C, KHALID S, et al. Foliar heavy metal uptake, toxicity and detoxification in plants: A comparison of foliar and root metal uptake[J]. *Journal of Hazardous Materials*, 2017, 325:36–58.
- [9] KANEKO M, KAZATANI T, SHIKATA H. Occupational lead poisoning in a patient with acute abdomen and normocytic anemia[J]. *Internal Medicine*, 2020, 59(12):1565–1570.
- [10] VAHABZADEH M, MEGARBANE B. Abdominal pain related to adulterated opium: An emerging issue in drug addicts[J]. *World Journal of Psychiatry*, 2020, 10(5):95–100.

- [11] BREWER G J. Copper-2 ingestion, plus increased meat eating leading to increased copper absorption, are major factors behind the current epidemic of alzheimer's disease[J]. *Nutrients*, 2015, 7(12):10053–10064.
- [12] KARPENKO M N, ILYICHEVA E Y, MURUZHEVA Z M, et al. Role of copper dyshomeostasis in the pathogenesis of parkinson's disease[J]. *Bulletin of Experimental Biology and Medicine*, 2018, 164(5):596–600.
- [13] BULCKE F, DRINGEN R, SCHEIBER I F. Neurotoxicity of copper[J]. *Advances in Neurobiology*, 2017, 18:313–343.
- [14] ZHU Q B, XING Y C, LU R F, et al. Visible/shortwave near infrared spectroscopy and hyperspectral scattering for determining bulk density and particle size of wheat flour[J]. *Journal of Near Infrared Spectroscopy*, 2017, 25(2):116–126.
- [15] RATHOD P H, BRACKHAGE C, VAN DER MEER F D, et al. Spectral changes in the leaves of barley plant due to phytoremediation of metals – results from a pot study[J]. *European Journal of Remote Sensing*, 2015, 48(1):283–302.
- [16] HEDE A N H, KASHIWAYA K, KOIKE K, et al. A new vegetation index for detecting vegetation anomalies due to mineral deposits with application to a tropical forest area[J]. *Remote Sensing of Environment*, 2015, 171:83–97.
- [17] HAN Q Q, YANG K M, LI Y R, et al. SVD-ANFIS model for predicting the content of heavy metal lead in corn leaves using hyperspectral

- data[J]. *Spectroscopy and Spectral Analysis*, 2021, 41(6):1930–1935.
- [18] BANERJEE B P, RAVAL S, ZHAI H, et al. Health condition assessment for vegetation exposed to heavy metal pollution through airborne hyperspectral data[J]. *Environmental Monitoring & Assessment*, 2017, 189(12):604.
- [19] MIRZAEI M, VERRELST J, SAFAR M, et al. Eco-friendly estimation of heavy metal contents in grapevine foliage using in-field hyperspectral data and multivariate analysis[J]. *Remote Sensing*, 2019, 11(23):2731.
- [20] LI Y R, YANG K M, GAO W, et al. A spectral characteristic analysis method for distinguishing heavy metal pollution in crops: VMD-PCA-SVM[J]. *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy*, 2021, 255:119649.
- [21] FU P J, ZHANG W, YANG K M, et al. A novel spectral analysis method for distinguishing heavy metal stress of maize due to copper and lead: RDA and EMD-PSD[J]. *Ecotoxicology and Environmental Safety*, 2020, 206:111211.
- [22] LI Y R, YANG K M, HAN Q Q, et al. A new method that combines spectral indexes and Naive Bayes to distinguish heavy metal pollution in crops[J]. *Remote Sensing Letters*, 2021, 12(7):666–673.
- [23] HONG Y S, LIU Y, CHEN Y Y, et al. Application of fractional-order derivative in the quantitative estimation of soil organic matter content through visible and near-infrared spectroscopy[J]. *Geoderma*, 2019, 337:758–769.
- [24] TONG P J, DU Y P, ZHENG K Y, et al. Improvement of NIR model by fractional order Savitzky-Golay derivation (FOSGD) coupled with wavelength selection[J]. *Chemometrics and Intelligent Laboratory Systems*, 2015, 143:40–48.
- [25] WANG J Z, TIYIP T, DING J L, et al. Desert soil clay content estimation using reflectance spectroscopy preprocessed by fractional derivative[J]. *PLoS One*, 2017, 12(9):19.
- (责任编辑:叶飞)