

doi: 10. 6046/zrzyyg. 2022376

引用格式: 陈昊宇,项磊,高贺,等. 基于分数阶微分的土壤全氮高光谱反演[J]. 自然资源遥感,2023,35(3):170-178. (Chen H Y, Xiang L, Gao H, et al. Hyperspectral inversion of total nitrogen content in soils based on fractional order differential[J]. Remote Sensing for Natural Resources, 2023, 35(3): 170-178.)

基于分数阶微分的土壤全氮高光谱反演

陈昊宇, 项 磊, 高 贺, 牟金燧, 索晓晶, 滑博伟

(中国地质调查局呼和浩特自然资源综合调查中心, 呼和浩特 010010)

摘要: 为快速无损地获取托克托县土壤全氮含量,满足当今精准农业的要求,文章以研究区内 120 个采样点的土壤全氮含量与高光谱数据为数据源,利用分数阶微分(1~2 阶)间隔 0.1 对光谱数据进行处理,筛选敏感波段,利用支持向量机(support vector machine, SVM)与 BP 神经网络模型共建立 24 个土壤全氮反演模型,结果表明:①经过分数阶微分处理后,光谱的波峰、波谷处信息被放大,随分解尺度的增加,其余波段的反射率逐渐趋于 0;②原始光谱与土壤全氮的皮尔森相关系数 $r=0.61$,经分数阶微分处理后,在 1.1 阶处达到最大值 $r=-0.67$,绝对值较之前提升了 0.06;③BP 神经网络预测模型结果优于 SVM 预测模型结果,本研究最佳土壤全氮预测模型为 1.1 阶微分处理后建立的 BP 神经网络模型,建模集 R^2 为 0.75,均方根误差(root mean square error, RMSE)为 0.16,验证集 R^2 为 0.71, RMSE 为 0.16,相对分析误差(relative percent deviation, RPD)为 2.06,可有效反演当地土壤全氮含量,相对于原始光谱建立的 BP 神经网络模型精度有较高提升。因此,利用 1.1 阶微分处理后的高光谱数据建立 BP 神经网络模型可实现对研究区土壤全氮含量的反演预测,可为当地精准农业的发展提供理论参考与技术支持。

关键词: 分数阶微分; BP 神经网络; 支持向量机; 精准农业

中图法分类号: TP 79 **文献标志码:** A **文章编号:** 2097-034X(2023)03-0170-09

0 引言

土壤全氮含量对作物的生长、健康以及品质等有着明显的影响,对叶片、根系、茎的生长起着促进作用,决定着作物的产量^[1],同时土壤全氮的含量也是衡量土壤肥力的重要指标之一^[2],所以精准掌握区域内土壤全氮含量对于发展农业、林业有着重要意义,传统的土壤养分测定方法虽然精度高,但是周期长、成本高,且只能得到瞬时量,无法满足当今精准农业的要求。高光谱遥感具有波段多、波段窄、信息丰富和实时高效等特点,在土壤化学分析领域得到了迅猛发展,相对于传统土壤检测方法而言,利用高光谱遥感手段周期短、成本低、时效性高,为实现土壤养分的快速诊断提供了新思路^[3-4]。

最早于 1986 年 Dalai 等^[5]利用多元回归分析发现 1 700~2 100 nm 是土壤光谱与氮元素关系较为密切的波段, Hummel 等^[6]与 Reeves 等^[7]研究表明土壤全氮在可见光和近红外波段处相关性较高,并利用近红外光谱反射率特征波段对土壤全

氮含量进行有效估测。随着高光谱技术的不断发展,国内基于高光谱土壤全氮反演也逐步趋于成熟,卢艳丽等^[8]利用反射率对数的一阶导数光谱预测黑土全氮含量,决定系数 R^2 均达到 0.8 以上;张娟娟等^[9]对我国中、东部地区 5 种类型土壤全氮含量进行高光谱反演研究,结果表明偏最小二乘法(partial least-square method, PLS)与 BP 神经网络建立高光谱全氮反演模型有较高的精度;王炜超等^[11]基于 CatBoost 算法建立的土壤全氮预测模型 R^2 均高于 0.80,证明 CatBoost 算法在土壤全氮含量预测中有一定的使用价值;赵燕东等^[10]与王世东等^[11]分别运用一阶微分与二阶微分有效预测了土壤全氮含量。

基于上述研究,各学者对于高光谱数据的处理方式均采取一阶微分与二阶微分等传统处理方式,整数阶微分数学变换形式对土壤光谱数据处理时中间空缺较多,可能会导致具体的信息遗漏,与整数阶微分相比,分数阶微分(fractional order derivative, FOD)具有较小的间隔变化,可以确保信噪比转变缓慢,这为检测到某些光谱曲线信号提供更多特

收稿日期: 2022-09-19; 修订日期: 2023-01-05

第一作者: 陈昊宇(1995-),男,硕士,主要从事自然资源调查、高光谱遥感研究。Email: chenhaoyu0807@163.com。

通信作者: 项 磊(1983-),男,学士,高级工程师,主要从事自然资源调查、地质矿产研究。Email: xiang6280@163.com。

征^[12]。此外,因一阶微分与二阶微分波形差异较大,可能会忽略土壤全氮含量造成的光谱差异,因此,为了细化光谱特征,利用 FOD 的方法对光谱数据进行潜在信息的挖掘是非常有必要的^[13],相关研究中竞霞等^[14]通过对原始光谱进行 FOD 处理,提高了光谱反射率与小麦条锈病严重度的相关性;赵慧等^[15]引入 FOD 对土壤高光谱反射率进行 0~2 阶微分处理,发现在 1.6 阶处模型结果较好;田安红等^[16]使用 FOD 的方法对新疆干旱区盐渍土的电导率进行研究时发现在 1.5 阶处估算效果最佳;Wang 等^[17]在对艾比湖的土壤有机质使用 FOD 方法时发现在 1.2 阶处取得了良好效果,综上所述,渐变的 FOD 信息制约模型的精度^[12],基于 FOD 的高光谱估算模型具有较好的稳定性,模型估算的预测能力较好。

非线性模型因其更加接近客观规律,得到了广泛的应用,而非线性模型的复杂性也促进了变量的筛选^[18]。支持向量机 (support vector machine, SVM) 是建立在统计学习理论中结构风险最小化基础上的监督学习模型^[19],在解决小样本、非线性、高维数等模式识别问题时,SVM 较其他建模方法来说具有更多的优点^[20],相对于线性模型能更好地应用于函数回归问题^[21]。BP 神经网络适用范围广、易操作、扩充性强,所以在实践中得到了广泛应用^[22],一般线性回归无法达到较高精度,神经网络算法以其强大的容错性、自组织自学习和非线性模拟能力在不确定两者之间关系的前提下仍然可以很好地解决非线性问题^[23],在估算建模研究中人们还发现,人工神经网络集成模型的精度要优于单神经网络模型,单神经网络模型则优于多元逐步回归模型^[24]。

所以本文采取 FOD 的处理方法,筛选敏感波段,构建土壤全氮高光谱反演模型,探究 FOD 在土壤全氮高光谱反演研究中深度挖掘光谱信息、有效提升模型反演精度的能力,力图为土壤全氮高光谱数据预处理提供一定的新思路,为土壤全氮高光谱反演提供理论依据。

1 研究区及数据来源

1.1 研究区概况

托克托县(图 1)隶属于内蒙古自治区首府呼和浩特市,位于自治区中部、大青山南麓、黄河上中游分界处北岸的土默川平原上。地理坐标在 E111°02'30"~111°32'21",N40°05'55"~40°35'15"之间,

总面积为 1 416.8 km²,平均海拔为 1 117 m,属于温带大陆性干旱气候,年均气温为 7.3℃,年均降雨量为 362 mm。托克托县耕地总面积达 400 km²,其中古城镇、新营子镇和五申镇的耕地较多,占全县耕地面积的 60% 以上,主要作物包括小麦、玉米和莜麦,工农业及生产生活用水主要来源于大黑河和黄河水资源,整个地形是以大黑河为轴,呈现由丘陵向平原过度的趋势,地势为东南高、西北和西南低;东南向西北土壤类型为栗钙土、砂壤石灰性冲积土、盐渍化石灰性冲积土,土壤类型的不同导致土壤养分存在差异性分布,植被类型从西向东依次为草甸草原、干草原和退化灌丛草原分布^[25]。

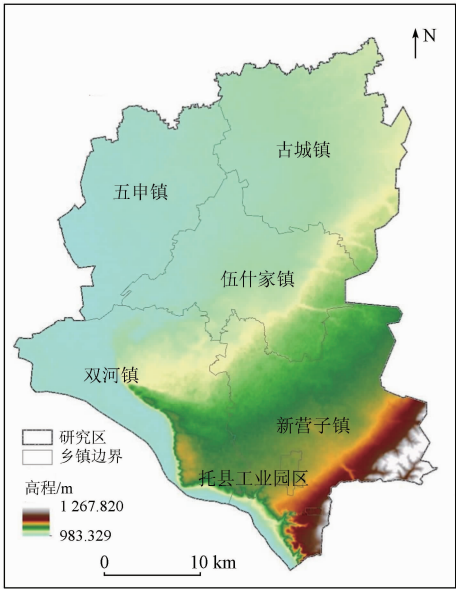


图 1 托克托县研究区概况

Fig. 1 Overview of study area of Togtoh County

1.2 数据来源

1.2.1 土样采集

采样时间为 10—11 月,地面植被较少,方便进入农田,也可减少取土时对地表植被的破坏。采用五点取样法,取样前先剔除地表杂物,取样深度为 0~20 cm,取样约 500 g 放入聚乙烯密封袋中,并记录相应的土地利用类型与地理坐标。采集的土样置于通风干燥的室内进行自然风干、去除杂物、研磨,并分为 2 份,一份过 10 目筛,用于光谱数据测定,一份过 100 目筛,用于土壤全氮含量测定。

1.2.2 土地利用现状

以 Landsat8 OLI 影像为基础数据源,运用人工目视解译与 BP 神经网络分类法得到托克托县 2019 年 7 月份土地利用数据,土壤采集分布示意图如图 2 所示。

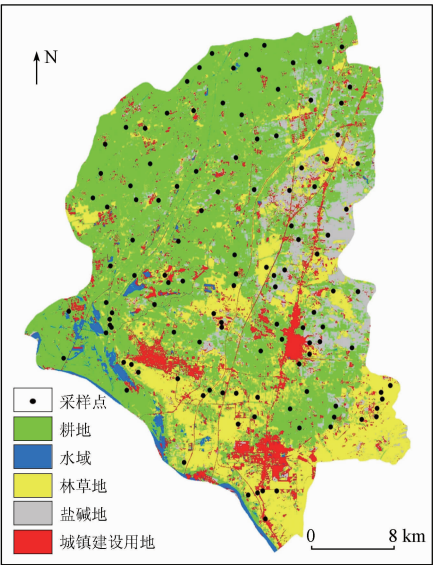


图 2 土壤采集分布示意图

Fig.2 Schematic diagram of soil collection and distribution

1.2.3 土壤全氮测定

土壤全氮采用半微量开氏法^[2],土壤全氮描述性统计见表 1,样品在加速剂的参与下用浓硫酸消煮,各种含氮有机化合物经过复杂的高温分解反应,转化为氨与硫酸结合的硫酸氨,碱化后蒸馏出来的氨用硼酸吸收,用标准酸溶液滴定,求土壤全氮含量(不包括全部硝太氮)。

表 1 土壤养分描述性统计

Tab.1 Descriptive statistics of soil nutrients

类型	最大值/ (g·kg ⁻¹)	最小值/ (g·kg ⁻¹)	均值/ (g·kg ⁻¹)	中位数/ (g·kg ⁻¹)	标准 差	偏度	峰度	变异 系数/ %
全氮	1.74	0.11	0.53	0.41	0.33	1.27	1.16	0.62

1.2.4 土壤光谱测定

土壤光谱于暗室内测量,测量示意图见图 3,采用 SVC HR-1024 便携式光谱仪,光谱范围 350~2 500 nm,在 350~1 000 nm 波段之间光谱分辨率≤3.5 nm,在 1 000~1 850 nm 波段之间光谱分辨率≤

9.5 nm,在 1 850~2 500 nm 波段之间光谱分辨率≤6.5 nm。光源采用与太阳光接近的 50 W 卤素灯,将土壤样品放入深 2 cm,宽 10 cm 的黑色器皿内,用直尺将土壤表面刮平,探头距离土样 10 cm,光源距离土壤表面 30 cm,天顶角为 30°,测量前用白板进行标定,每个土样采集 5 条光谱求平均作为该土样的光谱数据。

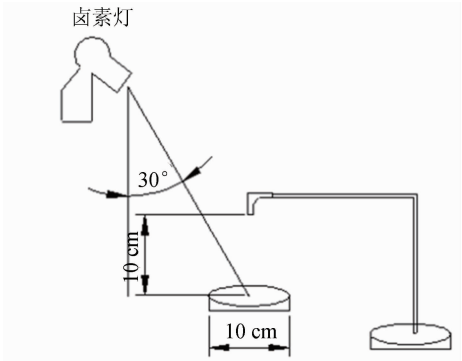


图 3 光谱室内测量示意图

Fig.3 Schematic diagram of spectrum indoor measurement

2 研究方法

2.1 光谱预处理

由于受噪声与仪器暗电流的影响,导致光谱数据混入噪声等信息,因此删除 350~399 nm 和 2 400~2 500 nm 的波段,采用五点平滑法对光谱进行平滑处理,为保留光谱所有信息,将光谱重采样至 1 nm。

2.2 FOD

FOD 由整数阶发展而来,目前比较常用的 FOD 定义为 Grunwald-Letnikov (G-L), Riemann-Liouville (R-L) 和 Caputo 这 3 种形式^[26],其中 G-L 是最为常用的一种形式,当光谱数据重采样至 1 nm 时,其主要形式为:

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

式中:Γ为 Gamma 函数;x 为波长;m 为微分上下限差;v 为任意阶数,当 v 为小数时,则为 G-L 分数阶微分形式,当 v 为整数时则为整数阶微分,即一阶微分、二阶微分,当 v 等于 0 时为原始光谱^[17]。

2.3 模型及精度验证

采用 BP 神经网络与 SVM 建立土壤全氮预测模型,本研究 SVM 的 cmd 参数中“-s”为 SVM 类型,0

代表 C-SVC 模型;1 代表 nu-SVM 模型;2 代表 one-class SVM 模型;3 代表 epsilon-SVR 模型;4 代表 nu-SVR 模型;“-t”为核函数设置类型,0 代表线性核函数,1 代表多项式核函数,2 代表 RBF 核函数,3 代表 sigmoid 核函数;“-p”为 epsilon-SVR 中损失函数 epsilon 的值,默认为 0.01。为了避免造成在高维空间中向量内积的大量运算,对 SVM 方法

进行了各种类型核函数的引用,在实际应用中,根据具体问题选择合适的惩罚参数 c ,灵敏损失参数 f ,核函数类型以及相应的参数;BP 神经网络的迭代次数设置为 1 000,学习率为 0.01,训练的均方根误差小于 0.001,神经元个数则针对不同数据集进行手动调整,有效避免过拟合现象发生。

模型精度采用决定系数 R^2 ,均方根误差 (root mean square error, RMSE) 和相对分析误差 (relative percent deviation, RPD) 共同评价。

R^2 表明模型的稳定性,越接近于 1 模型越稳定,拟合程度越好。公式为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_m - y_p)^2}{\sum_{i=1}^n (y_m - \bar{y})^2}, \tag{2}$$

式中: n 为样本数量; y_m 和 y_p 分别为土壤全氮实测值与预测值; $\bar{y}$ 为土壤全氮实测值的平均值。

RMSE 用来检验模型的预测能力,越小则表明模型的估测能力越好,公式为:

$$RMSE = \frac{\sum_{i=1}^n (y_m - y_p)^2}{n}。 \tag{3}$$

RPD 是样本的标准差与 RMSE 的比值, $RPD < 1.4$ 时,模型无法对样品进行预测; $1.4 \leq RPD < 2$ 时,认为模型预测效果一般,可以用来对样品进行粗略评估; $RPD \geq 2$ 时,模型具有极好的预测能力。计算公式为:

$$RPD = \frac{STDEV}{RMSE}, \tag{4}$$

式中 $STDEV$ 为样本的标准差。

3 结果与分析

3.1 土壤全氮光谱特征分析

采样点室内实测光谱数据曲线经平滑后如图 4 所示,光谱曲线整体变化趋势相同,在 450 ~ 470 nm 波段处反射率最低,1 400 nm 与 1 900 nm 波段处波谷较为明显,主要由土壤水分分子引起的,在 450 ~ 1 000 nm 波段处反射率增加较为明显,1 000 ~ 1 900 nm 波段反射率增长趋势较慢,2 000 ~ 2 300 nm 波段反射率仍增长趋势,350 ~ 450 nm 与 2 400 ~ 2 500 nm 波段光谱曲线较为杂乱且规律性较差,主要受环境以及仪器自身影响,所以本研究不考虑此范围的波段。

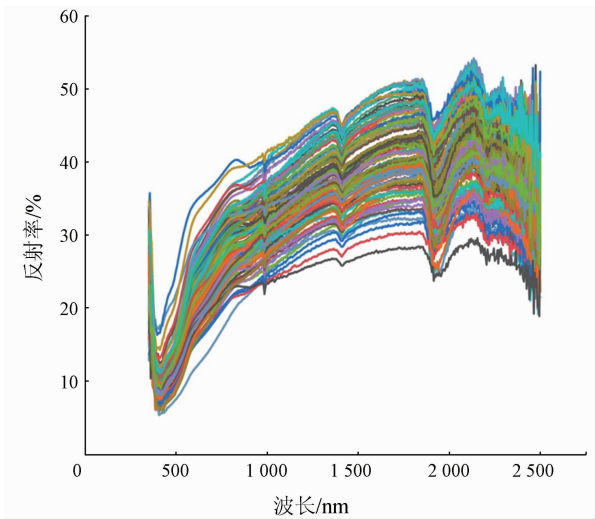


图 4 光谱反射率曲线

Fig.4 Spectral reflectance curve

全氮含量为 0.11 g/kg, 0.63 g/kg, 1.74 g/kg 土样的光谱曲线如图 5 所示。在 350 ~ 470 nm 波段处,不同全氮含量的光谱反射率随波长的增加而逐渐减少,而反射率趋于相同,除 1 000 nm, 1 500 nm, 2 000 nm 和 2 200 ~ 2400 nm 波段外,不同全氮含量的光谱反射率均随波长的增加而逐渐增加,但增长规律并不相同,在 500 ~ 600 nm 波段处,含量为 0.63 g/kg 与 1.74 g/kg 的光谱反射率较为接近,土壤全氮与原始光谱的反射率在 600 ~ 2 200 nm 土壤光谱反射率随全氮含量的增加而减少,在其他波段处全氮含量与光谱反射率没有很好的变换规律。

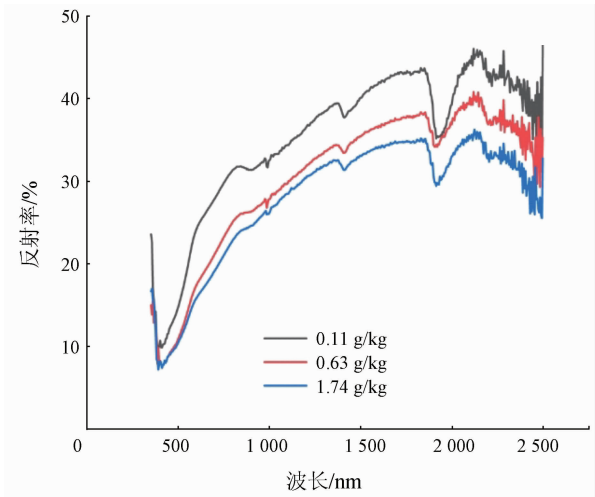


图 5 不同全氮含量光谱反射率

Fig.5 Spectral reflectance of different total nitrogen content

3.2 光谱数据 FOD

经 1 ~ 2 阶 (间隔 0.1) FOD 处理后,如图 6 所示,600 nm, 1 000 nm, 1 500 nm, 2 000 nm 波段处出现明显波动,对比原始光谱,在 500 ~ 1 000 nm 波段处,反射率随波长的增加而快速增加,在 600 nm 波

段处斜率达到最大值,在 1 000 nm,1 500 nm,2 000 nm 波段处存在 3 个显著的吸收峰,在 2 000 ~ 2 400 nm 波段处主要为水汽波段与环境噪声影响波段,光谱曲线变换频率较快,表明 FOD 处理光谱数据能够将光谱的波峰、波谷、斜率变化较快的波段信息放大,有效捕捉光谱的特征值,有利于挖掘光谱的潜在信

息,其余光谱反射率随分解尺度的增加逐渐趋于 0,进一步剔除光谱无用信息,保留有效信息。综上,FOD 处理对于光谱的波峰、波谷、变化较为明显的波段有较高的敏感性,对于放大信号,挖掘数据有显著作用,更有利于获取更精细的光谱数据。

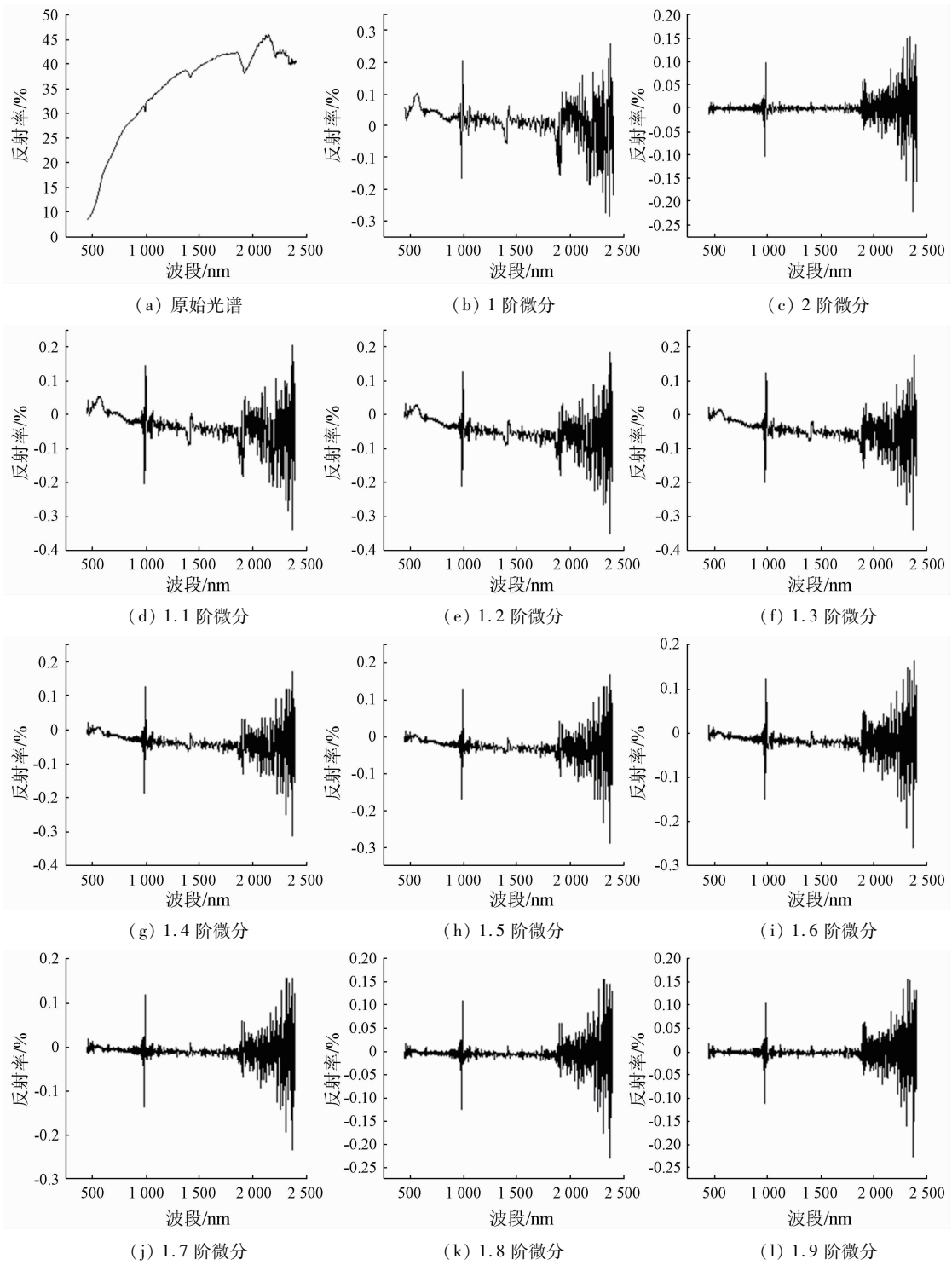


图 6 光谱对照图
Fig. 6 Spectrum comparison chart

3.3 相关性分析

光谱数据对应全氮含量皮尔森相关系数($p <$

0.01)计算结果如图 7 所示,原始光谱敏感波段主要集中于 500 ~ 600 nm 呈负相关、800 ~ 900 nm 呈正

相关,进行微分处理后,呈正相关的敏感波段范围明显增加,集中于 600 ~ 1 000 nm 波段处,在 1 ~ 1.4 阶微分范围内, r 均大于 0.60,随分解尺度的增加,范围逐渐减少;呈负相关的波段随分级尺度的增加,范围逐渐减少,但相关系数在 1.1 阶微分处达到

最大值(−0.67),正负相关均在 1.5 阶微分后敏感波段范围与相关系数逐渐减少。如图 7 所示,经过 FOD 处理后,相关系数较原始光谱($r=0.61$) 在 1.4 阶微分尺度内均有一定的提高,进一步证明 FOD 对于放大光谱信号,挖掘光谱信息有一定的意义。

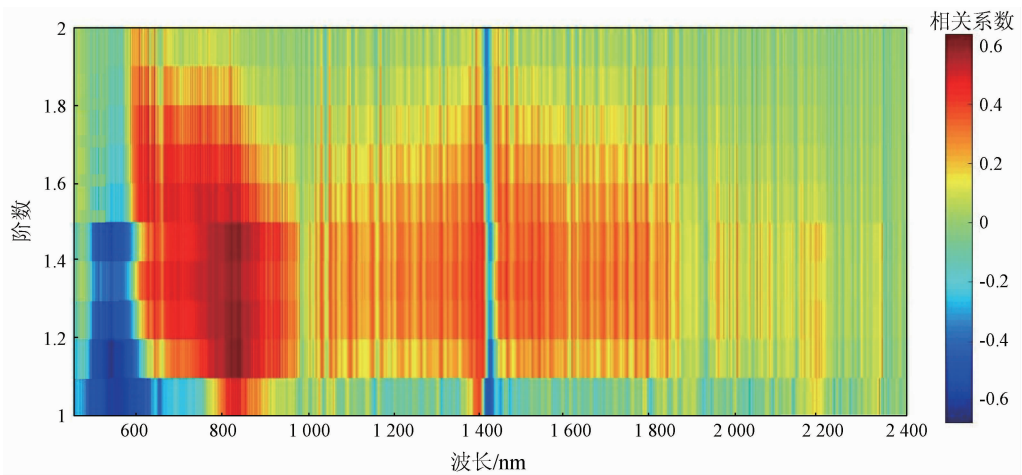


图 7 相关性分布
Fig. 7 Correlation distribution

3.4 模型构建

利用皮尔森相关系数筛选敏感波段,结果如表 2 所示,利用 90 个土壤全氮数据建立模型,30 个土壤全氮数据验证模型,建模集、验证集描述性统计如表 3 所示。

表 2 敏感波段筛选
Tab.2 Sensitive band screening

阶数 ^①	波长范围/nm	波段数
0	815 ~ 819,830 ~ 835	11
1	828 ~ 829,534 ~ 535,548 ~ 554	11
1.1	535 ~ 539,541 ~ 542,817 ~ 818,834 ~ 835,839	12
1.2	535 ~ 536,816 ~ 819,830 ~ 835,837,839 ~ 840	15
1.3	785,810,815 ~ 819,822,830 ~ 835,837,839 ~ 840	17
1.4	535 ~ 536,816 ~ 819,830 ~ 835	12
1.5	624,639,640,746 ~ 747,769,816 ~ 818	9
1.6	621,624 ~ 625,638 ~ 640,680,727,746 ~ 747,816 ~ 818	13
1.7	604,621,624 ~ 625,638 ~ 639,666,727,816 ~ 817	10
1.8	604,621,624,638 ~ 639	5
1.9	604,1 411 ~ 1 412,1 415	4
2	1 404,1 407 ~ 1 408	3

①阶数为 0 时为原始光谱。

表 3 全氮建模集和验证集描述性统计
Tab.3 Descriptive statistics of modeling set and prediction set of total nitrogen

数据集	最大值/ (g · kg ⁻¹)	最小值/ (g · kg ⁻¹)	均值/ (g · kg ⁻¹)	标准差/ (g · kg ⁻¹)	土样 数/个
建模集	1.74	0.12	0.52	0.29	90
验证集	1.44	0.11	0.52	0.36	30

3.4.1 SVM 模型

筛选敏感波段建立 SVM 预测模型结果如表 4 所示,原始光谱建立的模型 R^2 分别为 0.28 与 0.27, $RMSE$ 分别为 0.22 与 0.28, RPD 为 1.18,模型效果较差,无法对土壤全氮进行粗略估测,经过 FOD 处理后, R^2 均得到了一定提升,在 1.8 阶之后,模型预测效果逐渐下降,12 个预测模型整体来看, RPD 达到 1.4 以上的模型为 1.1 阶和 1.2 阶,建模集和验证集的 R^2 分别为 0.55 和 0.42,0.54 和 0.41, $RMSE$ 分别为 0.16 和 0.22,0.17 和 0.23, RPD 分别为 1.50 与 1.43,较原始光谱建模集 R^2 分别提升了 0.27 和 0.26, $RMSE$ 分别减少了 0.06 和 0.05,验证集 R^2 分别提升了 0.15 和 0.14,可实现对土壤全氮的粗略估算,1.6 阶与 1.4 阶处,模型建模集 R^2 虽达到 0.5 以上,但验证集 R^2 效果较差, RPD 也未达到模型评

表 4 SVM 预测模型
Tab.4 SVM prediction model

阶数 ^①	建模集		验证集		RPD
	R^2	$RMSE$	R^2	$RMSE$	
0	0.28	0.22	0.27	0.28	1.18
1	0.44	0.19	0.33	0.26	1.27
1.1	0.55	0.16	0.42	0.22	1.50
1.2	0.54	0.17	0.41	0.23	1.43
1.3	0.45	0.19	0.35	0.26	1.27
1.4	0.52	0.17	0.38	0.24	1.38
1.5	0.43	0.19	0.38	0.25	1.32
1.6	0.50	0.17	0.34	0.26	1.27
1.7	0.47	0.18	0.32	0.26	1.27
1.8	0.39	0.28	0.23	0.28	1.18
1.9	0.28	0.22	0.13	0.31	1.06
2	0.28	0.21	0.15	0.30	1.10

①阶数为 0 时为原始光谱。

价标准。将 R^2 , $RMSE$ 和 RPD 综合分析后,1.1 阶处理后筛选敏感波段土壤全氮可以进行粗略观测,也初步证明 FOD 处理后,有效提升了模型反演能力。

3.4.2 BP 神经网络模型

筛选敏感波段建立 BP 神经网络预测模型结果如表 5 所示,原始光谱建立模型建模集和验证集的 R^2 均为 0.58, $RMSE$ 分别为 0.25 和 0.31, RPD 为 1.06,无法对土壤全氮含量进行粗略估测,经过 FOD 处理后,1,1.1,1.2 和 1.3 阶模型 RPD 均大于 1.40,1.4 阶后预测模型 $RMSE$ 过高,导致 RPD 过低,无法对土壤全氮含量进行预测,综上最佳预测模型为 1.1 阶处理后建立的土壤全氮反演模型,建模集和验证集的 R^2 分别为 0.75 和 0.71, $RMSE$ 都为 0.16, RPD 为 2.06,可有效预测土壤全氮含量。

表 5 BP 神经网络预测模型
Tab.5 BP neural network prediction model

阶数 ^①	建模集		验证集		RPD
	R^2	$RMSE$	R^2	$RMSE$	
0	0.58	0.25	0.58	0.31	1.06
1	0.66	0.23	0.64	0.22	1.50
1.1	0.75	0.16	0.71	0.16	2.06
1.2	0.74	0.19	0.65	0.23	1.43
1.3	0.73	0.19	0.56	0.21	1.57
1.4	0.75	0.45	0.51	0.31	1.06
1.5	0.69	0.22	0.58	0.3	1.10
1.6	0.75	0.20	0.53	0.31	1.06
1.7	0.69	0.47	0.63	0.28	1.18
1.8	0.60	0.25	0.63	0.29	1.14
1.9	0.54	0.26	0.37	0.33	1.00
2	0.53	0.26	0.46	0.32	1.03

①阶数为 0 时为原始光谱。

4 讨论与结论

4.1 结论

本文以托克托县为研究区,进行野外采样,实验室土壤养分测定以及光谱测量,利用 FOD(0.1 间隔)将光谱数据进行 1~2 阶处理,分别建立 SVM 模型和 BP 神经网络模型,采用综合评价指标建立托克托县土壤养分最佳反演模型,得到如下结论:

- 1)不同含量土壤全氮高光谱反射率受众多因素干扰。
- 2)FOD 处理后光谱与全氮之间的最大相关系数绝对值由 0.61 提升至 0.67。
- 3)本研究区内 BP 神经网络预测模型优于 SVM 模型,最佳预测模型为 1.1 阶处理后建立的 BP 神经网络模型,建模集和验证集的 R^2 分别为 0.75 和 0.71, $RMSE$ 均为 0.16, RPD 为 2.06。

4.2 讨论

本研究区内,600~2 200 nm 波段处的土壤光谱

反射率随全氮含量的增加而减少,在其他波段处全氮含量与原始光谱反射率没有很好的变换规律,由于不同类型的土壤、理化性质不同,光谱特征也存在差异^[27],当全氮含量较低时,光谱反射特征对土壤类型、粒径和盐分等信息的响应^[28]。

FOD 对于提升土壤全氮与光谱数据之间的相关性有显著效果。对光谱数据进行 FOD(间隔 0.1)处理后,与土壤全氮建立相关性分析得出,未处理前,原始光谱最大相关系数 $r = 0.61$,FOD 处理后,在 1~1.4 阶内,相关系数均有一定提升,在 1.1 阶处相关系数绝对值达到最大值($r = -0.67$),较未处理前提高了 0.06,与田安红等^[16]和邵丽冰等^[29]研究结果一致,进一步证明分数阶微分对于挖掘光谱信息具有一定意义。

BP 神经网络模型对于研究区内土壤全氮含量预测效果最佳。相对 2 种模型来看,通过原始光谱建立土壤全氮高光谱估测模型效果较差,在 SVM 模型当中,以 1.1 和 1.2 阶处理后建立模型对土壤全氮可进行粗略估测,最佳预测模型建模集和验证集的 R^2 分别为 0.55 和 0.42, $RMSE$ 分别为 0.16 和 0.22, RPD 为 1.50。在 BP 神经网络当中,1~1.3 阶处理后模型均可实现土壤全氮估测,最佳预测模型建模集和验证集的 R^2 分别为 0.75 和 0.71, $RMSE$ 均为 0.16, RPD 为 2.06,可以实现土壤全氮含量有效预测,与竞霞等^[14]和李长春等^[30]研究结果一致。

在预测模型中,SVM 预测效果整体低于 BP 神经网络预测。由于针对不同类型土壤以及光谱数据,模型并不能良好运算,虽然 SVM 在处理非线性回归问题中有较强的能力,但本身模型中存在不稳定性,对模型的环境设置同样要求较高,所以未能对所有数据集进行良好的预测。由于光谱数据受土壤类型影响较大,不同区域光谱差异性较大,对验证集的敏感性不同,本文研究表明研究区土壤全氮含量适合 1.1 阶 FOD 处理后,建立 BP 神经网络模型进行含量估测。

参考文献(References):

[1] 王炜超,杨 玮,崔玉露,等.基于 CatBoost 算法与图谱特征融合的土壤全氮含量预测[J].农业机械学报,2021,52(s1):316-322.

Wang W C, Yang W, Cui Y L, et al. Prediction of soil total nitrogen based on CatBoost algorithm and fusion of image spectral features [J]. Transactions of the Chinese Society of Agricultural Machinery, 52(s1):316-322.

[2] 殷彩云,白子金,罗德芳,等.基于高光谱数据的土壤全氮含量估测模型对比研究[J].中国土壤与肥料,2022(1):9-15.

Yin C Y, Bai Z J, Luo D F, et al. Comparative study on estimation models of soil total nitrogen content based on hyperspectral data

- [J]. Soil and Fertilizer Sciences in China, 2022(1): 9–15.
- [3] 钟浩, 李西灿, 翟浩然, 等. 基于表层土壤光谱的耕层土壤有机质间接估测[J]. 安徽农业大学学报, 2020, 47(3): 421–426.
- Zhong H, Li X C, Zhai H R, et al. Indirect estimation of organic matter content in plough layer based on topsoil spectrum[J]. Journal of Anhui Agricultural University, 2020, 47(3): 421–426.
- [4] 王延仓, 杨秀峰, 赵起超, 等. 二进制小波技术定量反演北方潮土土壤有机质含量[J]. 光谱学与光谱分析, 2019, 39(9): 2855–2861.
- Wang Y C, Yang X F, Zhao Q C, et al. Quantitative inversion of soil organic matter content in northern tidal soil by binary wavelet technique[J]. Spectroscopy and Spectral Analysis, 2019, 39(9): 2855–2861.
- [5] Dalai R C, Henry R J. Simultaneous determination of moisture, organic carbon, and total nitrogen by in France spectrometry[J]. Soil Science Society of America Journal, 1986, 50: 120–123.
- [6] Hummel J W, Sudduth K A, Hollinger S E. Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor[J]. Computers and Electronics in Agriculture, 2001, 32(2): 149–165.
- [7] Reeves J B, Mccarty G W, Meisinger J J. Near infrared reflectance spectroscopy for the analysis of agricultural soil[J]. Journal of Near Infrared Spectroscopy, 1997, 7(3): 179–193.
- [8] 卢艳丽, 白由路, 王磊, 等. 黑土土壤中全氮含量的高光谱预测分析[J]. 农业工程学报, 2010, 26(1): 256–261.
- Lu Y L, Bai Y L, Wang L, et al. Hyperspectral prediction analysis of total nitrogen content in black soil[J]. Transactions of the Chinese Society of Agricultural Engineering, 2010, 26(1): 256–261.
- [9] 张娟娟, 田永超, 姚霞, 等. 基于高光谱的土壤全氮含量估测[J]. 自然资源学报, 2011, 26(5): 881–890.
- Zhang J J, Tian Y C, Yao X, et al. Estimating soil total nitrogen content based on hyperspectral analysis technology[J]. Journal of Natural Resources, 2011, 26(5): 881–890.
- [10] 赵燕东, 皮婷婷. 北京地区粘壤土全氮含量的光谱预测模型[J]. 农业机械学报, 2016, 47(3): 144–149.
- Zhao Y D, Pi T T. Spectral prediction model of total nitrogen content of clay loam in Beijing[J]. Transactions of the Chinese Society of Agricultural Machinery, 2016, 47(3): 144–149.
- [11] 王世东, 石朴杰, 张合兵, 等. 基于高光谱的矿区复垦农田土壤全氮含量反演[J]. 生态学杂志, 2019, 38(1): 294–301.
- Wang S D, Shi P J, Zhang H B, et al. Inversion of soil total nitrogen content in reclaimed farmland of mining area based on hyperspectral imaging[J]. Chinese Journal of Ecology, 2019, 38(1): 294–301.
- [12] 赵启东, 葛翔宇, 丁建丽, 等. 结合分数阶微分技术与机器学习算法的土壤有机碳含量光谱估测[J]. 激光与光电子学进展, 2020, 57(15): 253–261.
- Zhao Q D, Ge X Y, Ding J L, et al. Combination fractional order differential and machine learning algorithm for spectral estimation of soil organic carbon content[J]. Laser and Optoelectronics Progress, 2020, 57(15): 253–261.
- [13] 李武耀, 买买提·沙吾提, 买合木提·巴拉提. 基于分数阶微分的土壤有机质含量高光谱反演研究[J]. 激光与光电子学进展, 2023, 60(7): 404–411.
- Li W Y, Mamat S, Maihemuti B. Hyperspectral inversion of soil organic matter content based on fractional differential[J]. Laser and Optoelectronics Progress, 2023, 60(7): 404–411.
- [14] 竞霞, 张腾, 邹琴, 等. 基于分数阶微分光谱指数的小麦条锈病遥感监测模型构建[J]. 农业工程学报, 2021, 37(17): 142–151.
- Jing X, Zhang T, Zou Q, et al. Construction of remote sensing monitoring model of wheat stripe rust based on fractional differential spectral index[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(17): 142–151.
- [15] 赵慧, 李新国, 靳万贵, 等. 基于分数阶微分的博斯腾湖湖滨绿洲土壤电导率高光谱估算[J]. 甘肃农业大学学报, 2021, 56(1): 118–125.
- Zhao H, Li X G, Jin W G, et al. Hyperspectral estimation of soil conductivity in oasis along Bosten Lake based on fractional differential[J]. Journal of Gansu Agricultural University, 2021, 56(1): 118–125.
- [16] 田安红, 赵俊三, 张顺吉, 等. 基于分数阶微分的盐渍土电导率高光谱估算研究[J]. 中国生态农业学报(中英文), 2020, 28(4): 599–607.
- Tian A H, Zhao J S, Zhang S J, et al. Hyperspectral estimation of electrical conductivity of saline soil based on fractional differentiation[J]. Chinese Journal of Eco-Agriculture, 2020, 28(4): 599–607.
- [17] Wang X P, Zhang F, Kung H T, et al. New methods for improving the remote sensing estimation of soil organic matter content (SOMC) in the Ebinur Lake Wetland National Nature Reserve (ELWNNR) in northwest China[J]. Remote Sensing of Environment, 2018, 218: 104–118.
- [18] 卢志宏, 刘辛瑶, 常书娟, 等. 基于BP神经网络的草原矿区表层土壤N/P高光谱反演模型[J]. 草业科学, 2018, 35(9): 2127–2136.
- Lu Z H, Liu X Y, Chang S J, et al. Hyperspectral inversion of the surface soil N/P ratio in a grassland mining area based on the BP network[J]. Pratacultural Science, 2018, 35(9): 2127–2136.
- [19] 谭念, 孙一丹, 王学顺, 等. 基于主成分分析和支持向量机的木材近红外光谱树种识别研究[J]. 光谱学与光谱分析, 2017, 37(11): 3370–3374.
- Tan N, Sun Y D, Wang X S, et al. Research on near infrared spectrum with principal component analysis and support vector machine for timber identification[J]. Spectroscopy and Spectral Analysis, 2017, 37(11): 3370–3374.
- [20] 朱文静, 毛罕平, 周莹, 等. 基于高光谱图像技术的番茄叶片氮素营养诊断[J]. 江苏大学学报(自然科学版), 2014, 35(3): 290–294.
- Zhu W J, Mao H P, Zhou Y, et al. Nitrogen nutrition diagnosis of tomato leaves based on hyperspectral image technology[J]. Journal of Jiangsu University (Natural Science Edition), 2014, 35(3): 290–294.
- [21] 董毅, 程伟, 张燕平, 等. 基于SVM的先分类再回归方法及其在产量预测中的应用[J]. 计算机应用, 2010, 30(9): 2310–2313.
- Dong Y, Cheng W, Zhang Y P, et al. The method of classification and regression based on SVM and its application in yield prediction[J]. Journal of Computer Applications, 2010, 30(9): 2310–2313.
- [22] 郭云开, 刘宁, 刘磊, 等. 土壤Cu含量高光谱反演的BP神经网络模型[J]. 测绘科学, 2018, 43(1): 135–139, 152.
- Guo Y K, Liu N, Liu L, et al. Hyper-spectral inversion of soil Cu

content based on BP neural network model[J]. Science of Surveying and Mapping, 2018, 43(1): 135 – 139, 152.

[23] Pan H, Wang X Y, Chen Q, et al. Application of BP neural network based on genetic algorithm [J]. Computer Applications, 2005, 25(12): 2777 – 2779.

[24] 刘 润, 张绍良, 侯湖平, 等. 基于思维进化优化 BP 神经网络的大豆叶片叶绿素含量高光谱反演[J]. 江苏农业科学, 2018, 46(13): 212 – 216.

Liu R, Zhang S L, Hou H P, et al. Hyperspectral retrieval of chlorophyll content in soybean leaves based on mind evolutionary optimization BP neural network[J]. Jiangsu Agricultural Sciences, 2018, 46(13): 212 – 216.

[25] 陈昊宇, 杨 光, 韩雪莹, 等. 基于连续小波变换的土壤有机质含量高光谱反演[J]. 中国农业科技导报, 2021, 23(5): 132 – 142.

Chen H Y, Yang G, Han X Y et al. Hyperspectral inversion of soil organic matter content based on continuous wavelet transform[J]. China Agricultural Science and Technology Guide, 2021, 23(5): 132 – 142.

[26] 刘 杨, 冯海宽, 孙 乾, 等. 基于无人机高光谱分数阶微分的马铃薯地上生物量估算[J]. 农业机械学报, 2020, 51(12): 202 – 211.

Liu Y, Feng H K, Sun Q, et al. Estimation of potato above – ground biomass based on fractional differential of UAV hyperspectral[J]. Transactions of the Chinese Society of Agricultural Machinery, 2020, 51(12): 202 – 211.

[27] 徐彬彬. 土壤剖面的反射光谱研究[J]. 土壤, 2000(6): 281 – 287.

Xu B B. Study on the reflection spectrum of soil profile[J]. Soil, 2000(6): 281 – 287.

[28] 刘 凡, 马 玲, 杨 光, 等. 灰漠土土壤全氮含量的高光谱特征分析及估测[J]. 新疆农业科学, 2017, 54(1): 140 – 147.

Liu F, Ma L, Yang G, et al. Hyperspectral characteristic analysis and estimation of total nitrogen content in grey desert soil[J]. Xinjiang Agricultural Sciences, 2017, 54(1): 140 – 147.

[29] 邵丽冰, 陈奕云, 徐 璐, 等. 基于分数阶微分的土壤含水量高光谱响应特征与估测模型构建[J]. 测绘地理信息, 2022, 47(s1): 131 – 136.

Shao L B, Chen Y Y, Xu L, et al. Analysis on soil moisture content hyperspectral response and construction of estimation model based on fractional – order derivative[J]. Surveying and Mapping Geographic Information, 2022, 47(s1): 131 – 136.

[30] 李长春, 施锦锦, 马春艳, 等. 基于小波变换和分数阶微分的冬小麦叶绿素含量估算[J]. 农业机械学报, 2021, 52(8): 172 – 182.

Li C C, Shi J J, Ma C Y, et al. Estimation of chlorophyll content in winter wheat based on wavelet transform and fractional differential [J]. Transactions of the Chinese Society of Agricultural Machinery, 2021, 52(8): 172 – 182.

Hyperspectral inversion of total nitrogen content in soils based on fractional order differential

CHEN Haoyu, XIANG Lei, GAO He, MU Jinyi, SUO Xiaojing, HUA Bowei
(Hohhot General Survey of Natural Resources Center, China Geological Survey, Hohhot 010010, China)

Abstract: This study aims to determine the total nitrogen content (TNC) in soils in Tuoketuo County quickly and nondestructively, thus meeting the requirements of precision agriculture. With the soil TNC and hyperspectral data of 120 sampling sites in the study area as the data source, this study processed the hyperspectral data using the 1 ~ 2 orders fractional order differential (FOD) interval of 0.1 to screen the sensitive wavebands. Then, this study built 24 inversion models for soil TNC using the support vector machine (SVM) and the back propagation neural network (BPNN). The results are as follows: ① After FOD processing, the information at the wave crests and troughs of the spectra was amplified, and the reflectance of the remaining wavebands approached zero gradually with an increase in the decomposition scale. ② The Pearson correlation coefficient between original spectra and soil TNC was $r = 0.61$. This correlation coefficient was up to a maximum of 0.67 at 1.1 – order after FOD processing, increasing by 0.06. ③ The BPNN prediction models outperformed the SVM prediction models. The optimal soil TNC prediction model was the BPNN model built after 1.1 – order differential processing. This model yielded an R^2 of 0.75 and a root mean square error (RMSE) of 0.16 for the modeling set and an R^2 of 0.71 and an RMSE of 0.16 for the verification set, with a relative percent deviation (RPD) of 2.06. This model produced effective inversion results of the soil TNC in the study area, with a much higher accuracy than the BPNN model built using original spectra. Therefore, the BPNN model built using hyperspectral data through 1.1 – order differential processing allows for the inversion – based prediction of soil TNC in the study area, providing a theoretical reference and technical support for local precision agriculture.

Keywords: fractional order differential; BPNN; SVM; precision agriculture

(责任编辑: 张 仙)