

采用非负矩阵分解方法的秸秆-土壤光谱解混及盐分估测

李 振¹, 曹见飞^{1,2*}, 杨 晗¹, 刘建华¹, 王召海¹, 段欣荣³, 张乐天⁴

(1. 山东师范大学地理与环境学院, 济南 250014; 2. 中科山东东营地理科学研究院, 东营 257000; 3. 山东省国土测绘院, 济南 250013; 4. 中环联合(北京)认证中心有限公司, 济南 250000)

摘 要: 黄河三角洲地区发展节约型农业, 大力推行秸秆还田技术以改善土壤生态环境, 当利用遥感技术对部分秸秆覆盖地区进行土壤盐渍化监测时, 混合像元制约土壤盐分遥感估测的精度。该研究设置 10 组不同秸秆覆盖度的盐渍化地表场景, 多次光谱测量并取平均值作为其反射光谱。探究非负矩阵分解(Non-negative Matrix Factorization, NMF)方法分离农业秸秆和土壤光谱的有效性, 并基于分离出的土壤光谱构建偏最小二乘法土壤盐分估测模型。结果表明: 1) 秸秆覆盖会使土壤光谱反射率增高, 相比于纯土壤光谱在 1 730 nm、2 090 nm 附近出现吸收谷; 2) 通过 NMF 进行混合光谱解混之后, 可以有效地将土壤光谱从秸秆覆盖光谱中分离出来, 且土壤光谱保留了土壤盐分信息; 3) 对于不同秸秆覆盖水平下的土壤盐分估测模型而言, 与原始混合光谱相比, 利用土壤光谱 NMF 分离数据构建的模型精度普遍提高, 建模集决定系数平均提高 0.07, 均方根误差降低 1.21。验证集决定系数、相对分析误差分别平均提高 0.07、0.25, 均方根误差降低 1.22。该研究可为黄河三角洲地区部分秸秆覆盖盐渍土地的近地面遥感估测精度的提高提供方法。

关键词: 土壤; 模型; 盐分; 非负矩阵分解; 近地面遥感

doi: 10.11975/j.issn.1002-6819.2022.08.019

中图分类号: TP79

文献标志码: A

文章编号: 1002-6819(2022)-08-0161-08

李振, 曹见飞, 杨晗, 等. 采用非负矩阵分解方法的秸秆-土壤光谱解混及盐分估测[J]. 农业工程学报, 2022, 38(8): 161-168. doi: 10.11975/j.issn.1002-6819.2022.08.019 http://www.tcsae.org

Li Zhen, Cao Jianfei, Yang Han, et al. Spectral unmixing of straw and soil to estimate the soil salinity using non-negative matrix factorization[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(8): 161-168. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2022.08.019 http://www.tcsae.org

0 引 言

土壤盐渍化是盐分在地表土层中不断累积的结果, 是土地退化的一种常见表现形式^[1-2]。土壤盐渍化的产生受到多种因素的共同影响, 比如气候、降水、地形等自然因素以及不合理的农业灌溉、滥砍乱伐、破坏植被等人为因素^[3-6]。土壤盐渍化是可持续农业的严重威胁, 因为土壤盐渍化会导致土地肥力下降、农作物减产^[7-10]。因此, 加强对土壤盐分的监测和评估在风险管控、土壤修复等方面具有重要作用。

近年来, 遥感技术已经被广泛应用于土壤盐分的监测。尤其是 20 世纪 90 年代, 高光谱数据发展起来, 它比多光谱具有更多的波段, 其精细的光谱分辨率甚至可以反映土壤理化性质的微小差异, 所以被广泛用于土壤理化性质的估测和反演^[11-15]。孙亚楠等^[16]的研究表明高光谱的反演模型精度效果优于多光谱的反演模型, 并通过高-多光谱融合数据构建了反演模型进行土壤盐分的空间反演, 提高土壤盐分反演精度。厉彦玲等^[17]利用超球体色彩空间变换算法, 将 HSI 高光谱影像与 OLI 多光

谱影像进行融合, 通过选择土壤盐分的特征波段, 建立了统计分析模型和机器学习模型, 对黄河三角洲垦利县的土壤盐分进行了遥感反演。

但现实中土壤表面往往覆盖多种物质, 黄河三角洲地区普遍采用秸秆还田措施, 以期达到增强土壤肥力、改良土壤性状的目的, 但秸秆覆盖会引起土壤的整体光谱特征发生改变, 继而对土壤盐分遥感监测结果的精度造成影响, 因此如何解决定量遥感监测中的混合光谱问题对于农田土壤盐分遥感估测具有重要意义。已有一些研究^[18-22]采用光谱分解方法从混合光谱中分离出土壤光谱。其中, Ouergemmi 等^[23]应用盲源分离(Blind Source Separation, BSS)技术从混合光谱中提取土壤光谱, 并使用提取的土壤光谱预测土壤黏粒含量。Bartholomeus 等^[24]首次提出并利用残余光谱分解(Residual Spectra Unmixing, RSU)方法, 分离出玉米对混合成像光谱的影响, 估测土壤中有机碳的含量。赵伟等^[25]利用基于盲源分离的独立分量分析(Independent Components Analysis, ICA)算法, 对混合光谱进行解混, 削减枯枝落叶层对土壤光谱的影响, 提高了土壤有机碳光谱模型的精度。然而上述方法均存在一定的局限性, RSU 法需要事先知道土壤和植被所占的比例及其端元信息。而 ICA 技术则要求源信号相互独立。非负矩阵分解(Non-negative Matrix Factorization, NMF)作为另一种 BSS 技术, 提取的源光谱均为正值, 不需要任何先验知识, 更适合混合高光谱的分解。然而, NMF 是否能够分解

收稿日期: 2021-11-30 修订日期: 2022-04-01

基金项目: 中科山东东营地理研究院开放基金(202102); 山东省重点研发计划项目(2019GSF109034)

作者简介: 李振, 研究方向为土壤遥感监测。Email: lizhen_sdu@163.com

*通信作者: 曹见飞, 讲师, 研究方向为土壤遥感监测与制图。

Email: cjfsdnu@163.com

农作物秸秆-土壤混合光谱, 提取的土壤光谱是否能够成功评价土壤盐分含量, 仍待进一步研究。

因此, 本研究为了验证 NMF 分离农作物秸秆和土壤光谱的有效性, 以黄河三角洲农用地为研究对象, 试验模拟不同秸秆覆盖度盐渍化地表场景, 利用 NMF 方法剔除秸秆覆盖对土壤光谱的影响, 对比解混前后土壤盐分估测模型的精度, 验证非负矩阵分解方法的解混效果, 为盐渍化土壤遥感监测提供方法指导。

1 研究区概况

研究区位于黄河三角洲, 地理坐标为 $118^{\circ}33' \sim 119^{\circ}20'E$, $37^{\circ}25' \sim 38^{\circ}12'N$, 如图 1 所示。该区属暖温带大陆性季风气候, 受季风的影响, 降水比较集中, 且季节差异性比较大, 土壤易发生季节性积盐和脱盐。海水侵蚀、大气降水和风暴潮等过程带来的大量盐分成为黄河三角洲土壤盐渍化的重要物质基础。另外, 该区太阳辐射强度大, 蒸发量远大于降水量, 地形比较平缓, 地下水高度矿化, 且普遍埋深较浅。地下水中的盐分可以随着毛管水上升, 然后经过蒸发不断在地表累积, 导致土壤盐渍化。黄河三角洲土壤类型主要为滨海潮盐土, 受其影响, 该区植被、物种多样性比较低, 主要为盐生芦苇、柽柳等耐盐碱、耐水湿植物。由于盐荒地大量分布, 该区广种薄收, 产量较低, 严重限制了黄河三角洲地区的可持续发展^[26]。

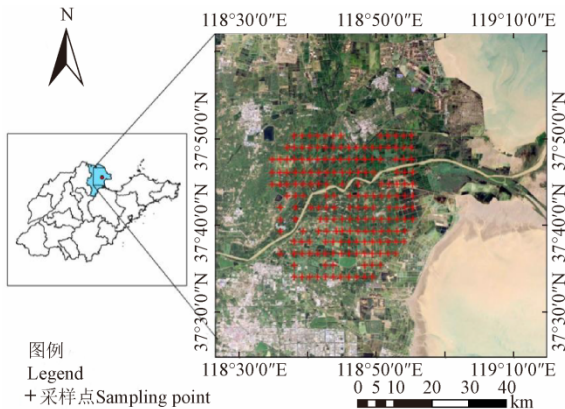


图 1 研究区及采样点的分布

Fig.1 Location of the study area and distribution of sampling points

2 材料与方法

2.1 数据准备

2.1.1 土壤数据采集

本研究于 2018 年 6—7 月进行实地调查和采样。在采样区利用网格法均匀布设采样点, 且采样点分布在各不同土壤盐渍化程度的位置。有效采样点 182 个, 每个采样点采集 0~20 cm 深度的土壤样本约 1 kg, 采样的同时应用手持 GPS 定位仪进行定位。研究区采样点覆盖的土地利用类型包括耕地和未利用地, 其中位于农用地的采样点 154 个, 未利用地的采样点 28 个。采样点的土壤类型为石灰性冲积土。

将采集的土壤样本自然风干, 经过研磨、过筛, 配置土壤溶液, 水土比例为 5:1。利用便携式多参数水质分析仪 (德国 WTW) Multi3420 SET B 测定电导率 ($EC_{1:5}$,

dS/m), 通过经验公式^[27]: $SSC=(0.2882EC_{1:5}+0.0183)$ 计算得到土壤含盐量 (SSC, g/kg)。利用重铬酸钾和硫酸的混合物在 $180^{\circ}C$ 下湿氧化测定土壤有机质 (Soil Organic Matter, SOM) 含量^[28]。

本研究为避免土壤有机质的干扰, 研究区内选择的采样点之间土壤有机质较为均质。对所有土壤样本重新增加有机质含量的测定。结果显示 182 个土壤样本的有机质含量最小值为 0.84 g/kg、最大值为 2.18 g/kg 以及平均值为 1.24 g/kg, 求得土壤有机质变异系数为 23.39%, 说明研究区有机质变异程度较低。土壤盐分和有机质含量统计特征如表 1 所示。

表 1 土壤盐分和有机质含量统计特征

Table 1 Statistical characteristics of soil salt and organic matter content

项目 Item	有机质 Organic matter/(g kg ⁻¹)	含盐量 Salt content/(g kg ⁻¹)
最小值 Min	0.84	0.06
最大值 Max	2.18	35.21
平均值 Mean	1.24	6.98
标准差 Standard deviation	0.29	8.02
变异系数 Variable of coefficient/%	23.39	114.90

2.1.2 试验室模拟秸秆覆盖

在试验室内设置不同土壤盐分含量以及不同秸秆覆盖度的对照试验。首先剔除土壤中的杂物, 干燥至恒定质量, 然后磨成粒径小于 0.149 mm 的细粉。对不同土壤盐分含量样本分别设置 10 组不同秸秆覆盖度模拟, 本研究选择剪碎的小麦秸秆为覆盖材料。

首先将处理过的不同含盐量土壤样本分别等量放置在干净的容器内。

其次把秸秆叶片压平, 切成 2 cm 长 (1 cm 高), 并放置在现场的土壤表面。在不损坏土壤样品的情况下, 每次添加等量秸秆。

最后将数码相机放置于容器正上方拍摄照片, 进行图像二值化处理, 统计秸秆占据照片比例, 得到秸秆覆盖度。根据测量视场面积, 将秸秆覆盖度分别设置为 0%、5%、10%、15%、20%、25%、30%、35%、40%和 50%, 模拟秸秆和土壤各占不同比例的混合像元。

2.2 光谱反射率测量

土壤光谱反射率采用 ASD FieldSpec3 地物光谱仪 (光谱分辨率为 350~2 500 nm) 在试验室内进行测量。

去除采集的土壤样本中的杂物, 然后烘干、研磨, 在透明塑料容器中平铺土壤样本, 分别设置不同秸秆覆盖度模拟试验。选取 25° 视场角, 使测量仪器垂直于土壤表面, 距离土壤样品的高度为 10 cm。在地面 50 cm 高度设置两盏 50 W 的卤素灯作为照射光源, 照明角度为 45° 。每次测量前使用 40 cm×40 cm 硫酸钡 ($BaSO_4$) 参照板进行校准, 为了消除测量的不稳定性, 每个土壤样品进行多次测量, 然后取平均值作为其光谱反射率。

2.3 光谱分析方法

2.3.1 非负矩阵分解

非负矩阵分解由 Lee 等^[29]提出, 简单来说对于任意一个给定的非负矩阵 X , NMF 算法能够寻找到更小的一

个非负矩阵 \mathbf{A} 和一个非负矩阵 \mathbf{S} , 使得满足:

$$\mathbf{X} \approx \mathbf{A}\mathbf{S} \quad (1)$$

由于 \mathbf{X} 和 \mathbf{A} 、 \mathbf{S} 乘积往往不能相等, 所以这里采用近似值尽量逼近。其中, $\mathbf{X} \in \mathbb{R}_+^{n \times m}$; $\mathbf{A} \in \mathbb{R}_+^{n \times r}$; $\mathbf{S} \in \mathbb{R}_+^{r \times m}$ 。通常在一般情况下, r 小于矩阵行列数, 即满足 $(n \pm m)r < nm$ 。本研究中 r 为端元数目。

另外, NMF 是一个优化求解的过程, 需要构造合理的目标函数^[30]

$$\min f(\mathbf{A}, \mathbf{S}) = \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 \quad (2)$$

$$\text{s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0, \mathbf{1}_r^T \mathbf{S} = \mathbf{1}_m^T$$

通过迭代的方式交替求解 \mathbf{A} 和 \mathbf{S} , 从而获得局部最优解。在不考虑噪声等误差影响时, 目标函数取得最小值 0, 并且高光谱像元解混中除了非负约束外, 同时要满足 $\mathbf{1}_r^T \mathbf{S} = \mathbf{1}_m^T$ 表达的丰度和为一 (Abundance Sum-to-one Constraint, ASC) 约束。为了考虑丰度 ASC 约束, 全约束最小二乘法方法 (fully constrained least squares, FCLS) 被应用。在这个方法中采用了增广矩阵^[30]

$$\hat{\mathbf{X}} = \begin{bmatrix} \mathbf{X} \\ \delta \mathbf{1}_m^T \end{bmatrix}, \hat{\mathbf{A}} = \begin{bmatrix} \mathbf{A} \\ \delta \mathbf{1}_r^T \end{bmatrix} \quad (3)$$

式中 $\delta > 0$ 用于调节丰度满足 ASC 约束的程度。

本研究将试验室模拟秸秆覆盖光谱作为输入矩阵 \mathbf{X} , 对其进行非负矩阵分解。通过 NMF 模型输出端元矩阵 \mathbf{A} 和丰度矩阵 \mathbf{S} , 比较提取的两种光谱 (秸秆光谱与土壤光谱) 与裸土平均光谱的光谱夹角分析, 其中光谱夹角值较小的光谱为土壤光谱, 从而对土壤光谱进行判别与选择。

2.3.2 模型构建和验证

本研究以土壤盐分作为因变量, 以试验室测量的全波长的光谱反射率作为自变量, 建立偏最小二乘回归模型。偏最小二乘回归 (Partial Least Squares Regression, PLSR)^[31] 是一种的多元统计数据分析方法, 被广泛应用于高光谱定量反演中。它综合了主成分分析和逐步回归分析法, 不仅能解决多重共线性问题, 且相对于主成分回归分析, PLSR 新生成的主成分对因变量拥有最强的解释力。

本研究中混合地物主要是土壤和秸秆, 端元数目 r 等于 2。为满足 NMF 模型 $(n \pm m)r < nm$ 中的要求, 输入矩阵 \mathbf{X} 中光谱数目需要大于端元数目。为了探究非负矩阵分解 (NMF) 方法分离农业秸秆和土壤光谱的有效性, 设置 2 种模型验证策略:

策略 1: 为了从总体上验证非负矩阵分解 (NMF) 方法分离农业秸秆和土壤光谱的有效性, 随机选取 127 个土壤样本作为建模集, 剩余的 55 个土壤样本作为验证集。并将土壤样本分别进行秸秆覆盖模拟。同一样本对应的不同秸秆的覆盖光谱作为 NMF 光谱分解方法的输入。NMF 分解后的纯土壤光谱用于模型构建与验证。

策略 2: 为了探究不同秸秆覆盖条件下非负矩阵分解 (NMF) 方法分离农业秸秆和土壤光谱的有效性, 将 182 个采集样本按照含盐量进行排序, 挑选 68 个不同含盐量的土壤作为模拟试验的基础数据, 挑选的土壤样本的含盐浓度范围为 0.56~34.66 g/kg, 每组 4 个样本, 组内土

壤盐分相差不超过 0.15 g/kg, 共分为 17 组。组与组之间土壤盐分浓度差在 0.5 g/kg 左右, 但由于所采样本盐分浓度具有较高的变异性, 后 4 组浓度间隔较大, 依次间隔为 2.04、4.28、9.42 和 11.94 g/kg。每组的平均土壤盐分含量如表 2 所示。其中每组的土样盐分含量相近的 4 个样本对应的混合光谱作为 NMF 光谱分解方法的输入。由此本试验在每种秸秆覆盖水平下获得 17 个土壤光谱, 共计 170 条土壤光谱用于偏最小二乘回归建模。并将剩余的土壤样本重复进行模拟秸秆覆盖试验, 将其作为独立的验证集对所建模型的精度进行验证, 其中同一样本对应的不同秸秆的覆盖光谱作为 NMF 光谱分解方法的输入。NMF 分解后的纯土壤光谱用于模型验证。

表 2 不同含盐量土壤样本分组

Table 2 Groups of soil samples with different salt content

组别 Group	含盐量 Salt content (g kg ⁻¹)	组别 Group	含盐量 Salt content (g kg ⁻¹)	组别 Group	含盐量 Salt content (g kg ⁻¹)
1	0.56	7	3.5	13	6.98
2	1	8	4	14	9.02
3	1.5	9	4.5	15	13.3
4	2.02	10	5	16	22.72
5	2.56	11	5.54	17	34.66
6	3.08	12	6.04		

2.3.3 模型精度评估

采用建模集的决定系数 R_c^2 、均方根误差 $RMSE_c$ 和验证集的决定系数 R_p^2 、均方根误差 $RMSE_p$ 、相对分析误差 RPD_p 来评价土壤盐分估测模型的精度。

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

$$RPD = \frac{SD}{RMSE} \quad (6)$$

式中 n 是验证样本的数量, y 是实际测量值, \bar{y} 是实际测量值的平均值, \hat{y}_i 是回归估计值。SD 为实际测量值的标准差。

R^2 越大, $RMSE$ 越小, 说明模型的精度越高。根据 RPD 值确定六类模型类别^[32]: 当 $RPD < 1.0$ 时, 模型效果非常差; 当 $1.0 \leq RPD < 1.4$ 时, 表示模型效果较差, 仅可以分辨高值和低值; 当 $1.4 \leq RPD < 1.8$ 时, 表示模型可用于评估; 当 $1.8 \leq RPD < 2.0$ 时, 表示模型能够进行良好的预测; 当 $2.0 \leq RPD < 2.5$ 时, 表示模型效果很好; 当 $RPD > 2.5$ 时, 表明模型非常优秀。

3 结果与分析

3.1 秸秆覆盖对土壤光谱反射率的影响

采样区土壤中的主要可溶性盐是 NaCl 和 $MgCl_2$, 不同含盐量的土壤光谱如图 2a 所示。在整个波长范围内, 光谱反射率随着盐分的增加逐渐降低, 但曲线变化趋势保持相似。从光谱波段观察, 光谱突出特征主要分布在

肩部 800 和 2 350 nm 波段,波谷 1 410, 1 940 和 2 210 nm 附近。在到达第一个吸收特征 1 410 nm 前,光谱间的差异相对较小。在之后的光谱区域,随着盐分增加,反射率落差较为明显,但光谱曲线波峰波谷位置相对稳定。

不同秸秆覆盖度含盐土壤表面的反射光谱曲线如图 2b 所示。可以看出,不同秸秆覆盖度的光谱曲线虽然在局部区域存在差异,但是总体趋势一致,光谱反射率显著高于土壤反射率。并且秸秆覆盖土壤光谱的形状总体上与秸秆光谱保持一致,随着秸秆覆盖度的增加,光谱越来越类似于秸秆光谱。在大于 500 nm 的波段范围内,当秸秆覆盖度越高,光谱反射率越高。在小于 2 250 nm 的波段范围内,光谱反射率由低到高依次为土壤、秸秆覆盖土壤、秸秆。在大于 2 250 nm 的波段范围内,秸秆、秸秆覆盖土壤与土壤的反射光谱曲线开始出现交叉现象。与土壤光谱不同,秸秆覆盖土壤光谱在 1 450 nm 附近吸收谷的起始点之间的波段宽度明显大于裸土土壤光谱,并且在 1 730 nm、2 090 nm 附近有吸收谷。

3.2 基于 NMF 的混合光谱解混

以盐分含量为 5.54 g/kg 的土壤制成的秸秆覆盖样本为例,基于 NMF 对混合光谱进行解混获得的纯端元光谱结果如图 3 所示。从图中可以看出,NMF 能够将混合光谱分离为两个可区分的光谱,并且光谱曲线与秸秆覆盖度相关。当秸秆覆盖度低于 25%时,提取的两条光谱曲线分别与秸秆光谱和裸土光谱相似。这表明 NMF 成功分离了秸秆与土壤光谱。当秸秆覆盖度增加到 25%时,提取的土壤

光谱曲线依然能保持土壤反射特征,但秸秆光谱在 1 730 nm 附近反射率陡然下降。当秸秆覆盖度增加到 30%以上,提取的土壤和秸秆光谱曲线形状发生明显变化,土壤反射特征出现不稳定现象。这一结果说明,当秸秆覆盖度较高时,随着秸秆覆盖度的增加,NMF 的性能变差。

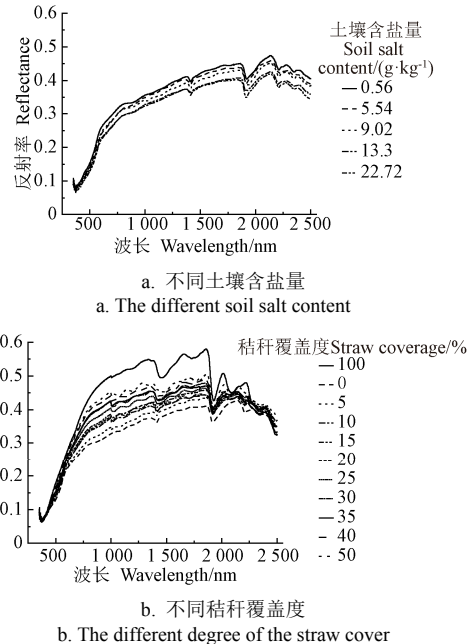


图 2 不同含盐量以及不同秸秆覆盖度的土壤光谱反射率
Fig.2 Spectral reflectance of soil with different salt content and different degree of the straw cover

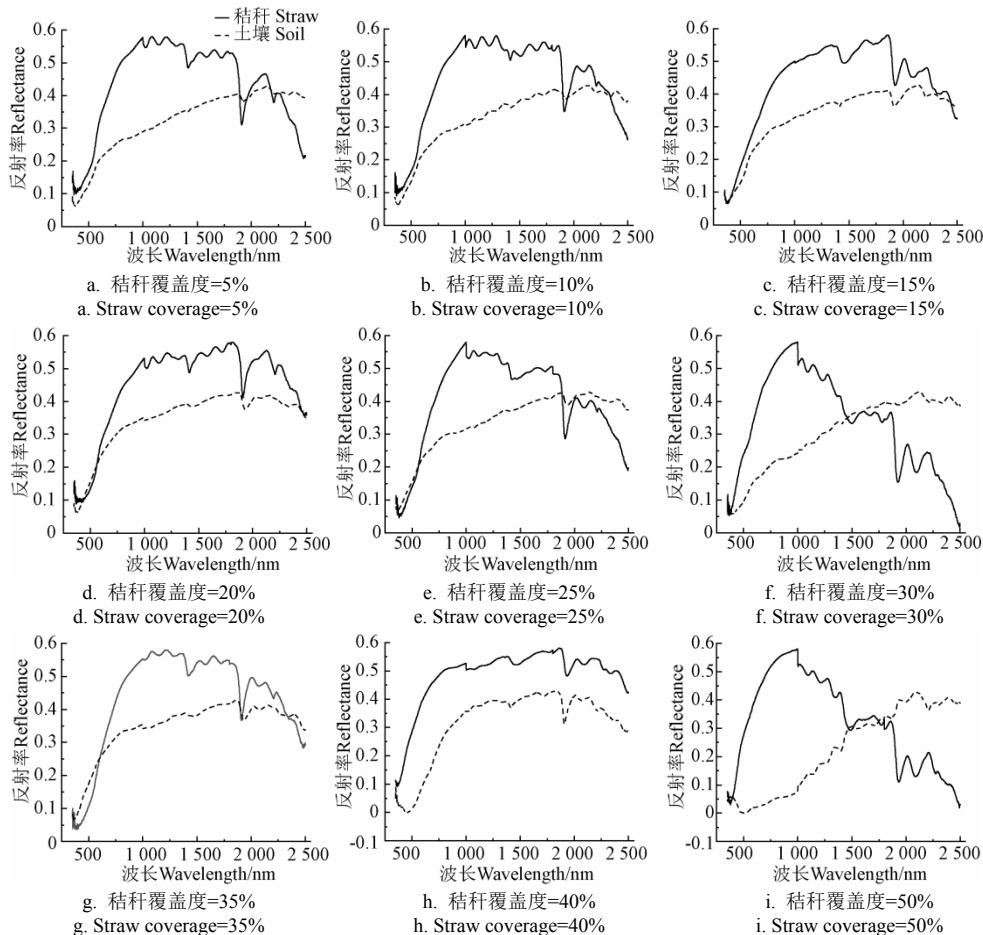


图 3 利用 NMF 提取不同秸秆覆盖度下的端元光谱

Fig.3 Extracted endmember spectra with different degree of the straw cover by Non-negative Matrix Factorization (NMF)

3.3 光谱解混前后的土壤盐分估测对比

依据验证策略 1, 利用建模集和验证集分别获得 NMF 光谱解混后的 PLSR 模型的预测值和实测值的关系如图 4 所示, 从总体上验证 NMF 方法分离农业秸秆和土壤光谱的有效性。经 NMF 方法解混获得土壤光谱构建的模型精度有所提升, 其中, 建模集决定系数 R_c^2 由 0.65 提高至 0.77, 均方根误差 RMSEc 由 7.17 降低至 5.62; 验证集决定系数 R_p^2 由 0.61 提高至 0.70, 均方根误差 RMSEp 由 8.57 降低至 6.80, 相对分析误差 RPDp 由 1.16 提高至 1.48。从散点图上可以看出经过 NMF 分解后的土壤光谱构建的模型精度有所提升。

3.4 不同秸秆覆盖水平的土壤盐分估测比较

依据验证策略 2, 在同一秸秆覆盖下, 每组 4 个土壤样本的混合光谱组成 NMF 模型的输入矩阵 X , 计算出 1 个土壤光谱, 共计 17 个土壤光谱。利用 17 个土壤光谱与每组平均土壤盐分含量构建 PLSR 模型, 然后依次对 10 个秸秆覆盖水平进行混合光谱解混, 对比解混前后的模型精度分析 NMF 分离农业秸秆和土壤光谱的有效性,

如表 3 所示。相比无秸秆覆盖的裸土光谱, 混合光谱造成模型精度下降。随着秸秆覆盖度的增加, 模型精度呈现下降趋势, 表明模型的精度会随着秸秆覆盖度的增加而不断下降。另一方面相比原始混合光谱, NMF 分解后的土壤光谱在一定程度提高了土壤估测模型精度, 缓解了混合光谱干扰。

经 NMF 解混后, 提取的土壤光谱数据构建的土壤盐分估测模型的建模集精度决定系数平均值从 0.65 提高至 0.72, 提高了 0.07; 均方根误差平均值由 7.42 降低至 6.21, 降低了 1.21; 验证集精度平均值为决定系数由 0.61 提高至 0.68, 提高了 0.07; 相对分析误差由 1.22 提高到 1.47, 提高了 0.25; 均方根误差平均值由 7.56 降低到了 6.34, 降低了 1.22。并且结果也表明当秸秆覆盖度增加, 模型的精度逐渐降低, NMF 提取的土壤光谱有效性也逐渐降低。当秸秆覆盖度大于 15% 时, 解混前的混合光谱模型相对分析误差 RPD 值小于 1.4, 表明此时估测模型估测结果已经失去定量能力。NMF 分解后的土壤光谱将估测模型定量能力提高至秸秆覆盖度 25%。

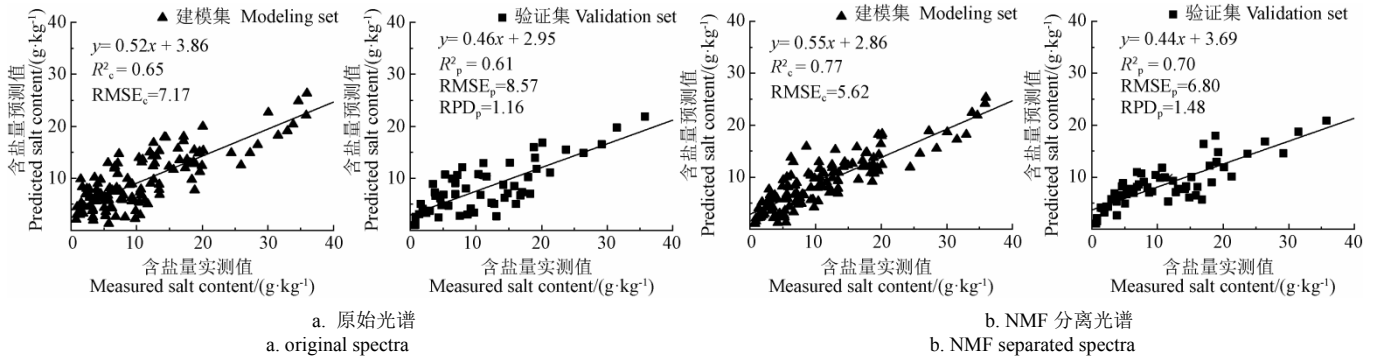


图 4 基于原始光谱和 NMF 分离光谱的含盐量实测值与预测值散点图

表 3 NMF 分解前后土壤盐分估测模型精度对比
Table 3 Accuracy comparison of soil salt estimation models before and after NMF

秸秆覆盖度 Degree of the straw coverage/%	原始光谱 Original spectra					NMF 分解光谱 NMF separated spectra				
	建模集 Modeling set		验证集 Validation set			建模集 Modeling set		验证集 Validation set		
	R_c^2	RMSE _c	R_p^2	RMSE _p	RPD _p	R_c^2	RMSE _c	R_p^2	RMSE _p	RPD _p
0	0.93	3.96	0.91	4.31	2.06	-	-	-	-	-
5	0.81	5.45	0.77	5.49	1.61	0.90	4.59	0.88	4.64	1.91
10	0.76	5.74	0.74	5.81	1.52	0.89	4.66	0.86	4.71	1.88
15	0.72	6.25	0.68	6.21	1.43	0.84	4.89	0.80	4.89	1.81
20	0.70	6.55	0.64	6.56	1.35	0.81	5.27	0.75	5.30	1.67
25	0.65	7.12	0.61	7.74	1.14	0.76	6.11	0.70	6.24	1.42
30	0.63	8.33	0.58	8.54	1.04	0.67	6.87	0.62	7.03	1.26
35	0.59	8.52	0.54	8.73	1.01	0.62	7.32	0.58	7.44	1.19
40	0.52	9.11	0.51	9.28	0.95	0.54	7.70	0.54	8.13	1.09
50	0.44	9.69	0.42	9.67	0.92	0.46	8.52	0.43	8.68	1.02

4 讨论

本研究结果表明, 秸秆覆盖土壤的光谱反射率明显高于纯土壤光谱反射率, 且在大于 2 250 nm 的波段范围内, 秸秆、秸秆覆盖土壤与土壤的反射光谱曲线开始出现交叉现象, 这是秸秆纤维的吸收作用导致的^[33]。当秸秆覆盖度较高时, 光谱曲线越来越类似于秸秆光谱, 这

也导致了 NMF 方法从混合光谱中获取土壤信息更加困难。模型结果显示, 随着秸秆覆盖度增加盐分估测模型精度呈现先缓后急的下降趋势。

已有研究^[23-24]利用 RSU 法和 BSS 法对混合光谱信号进行分解, 但这些方法的应用依赖端元和丰度等先验知识。NMF 对混合光谱和源光谱信号都有限制, 只需要确定最大迭代次数这一个参数^[34]。因此, NMF 在方法应

用方面更具有优势。NMF 混合分解方法是建立在不同端元线性组合的基础上。虽然在本研究试验中尽可能的避免光谱不受其他因素的干扰,但是这种绝对条件几乎无法实现,土壤中的组份差异、秸秆的褶皱仍然影响 NMF 分解过程。现实场景中的土壤盐渍化地区下垫面是由复杂的地物组成,大部分地物存在空间变异性与 NMF 的纯端元线性假设冲突,这将会导致 NMF 在卫星影像应用中存在较大局限性^[30]。另一方面,为限制 NMF 解空间,需要设计相应的约束条件,引导土壤光谱提取的最优化^[35]。本文仅是将 NMF 粗略的应用于秸秆-土壤混合光谱分解,为进一步提高 NMF 方法的有效性,作者将继续深入研究空谱约束解混条件下的 NMF 土壤盐分模型构建方法。

尽管 NMF 并不能完全消除秸秆的影响,但与秸秆覆盖光谱相比,提取的土壤光谱为估测模型提供了更多有用的信息。这为秸秆覆盖条件下的土壤盐分估测提供了新的建模思路,同时可将该方法推广其他混合光谱应用中。对于卫星遥感数据,NMF 计算过程的不确定性、现实场景的地物复杂性和变异性以及背景环境干扰都会对估测模型造成误差。因此,如何将该方法可靠稳定的推广至卫星影像仍是需要进一步研究的内容。

5 结 论

1) 秸秆覆盖土壤的光谱反射率显著高于土壤光谱反射率,且随着秸秆覆盖度的增加,秸秆覆盖土壤的光谱反射率不断升高。对于试验室模拟秸秆覆盖数据,土壤盐分估测模型的精度会随着秸秆覆盖度的增加而降低。

2) NMF 可以有效地剔除秸秆的影响,将土壤光谱从混合光谱中分离出来。随着秸秆覆盖度的增大,NMF 的性能变差,提取的土壤光谱曲线形状与反射特征出现不稳定现象。

3) 经过 NMF 分解后,提取的土壤光谱可以用来进行土壤盐分估测,模型精度相比于解混前有所提升。其中,建模集决定系数 R_c^2 平均值由 0.65 提高至 0.72;均方根误差 $RMSE_c$ 平均值由 7.42 降低至 6.21。当秸秆覆盖度小于 25% 时,经 NMF 解混提取的土壤光谱构建的土壤盐分估测模型的相对分析误差 RPD 均大于 1.4,且精度相比于原始混合光谱均有较大提升。

[参 考 文 献]

- [1] Wang Z, Zhang F, Zhang X L, et al. Regional suitability prediction of soil salinization based on remote-sensing derivatives and optimal spectral index[J]. *Science of the Total Environment*, 2021, 775(12): 145807.
- [2] Boudibi S, Sakaa B, Zineddine B, et al. Spatial prediction and modeling of soil salinity using simple cokriging, artificial neural networks, and support vector machines in El Outaya plain, Biskra, southeastern Algeria[J]. *Acta Geochimica*, 2021, 40(3): 390-408.
- [3] 边玲玲, 王卷乐, 郭兵, 等. 基于特征空间的黄河三角洲垦利县土壤盐分遥感提取[J]. *遥感技术与应用*, 2020, 35(1): 211-218.
Bian Lingling, Wang Juanle, Guo Bing, et al. Remote

- Sensing extraction of soil salinity in Yellow River Delta Kenli County based on feature space[J]. *Remote Sensing Technology and Application*, 2020, 35(1): 211-218. (in Chinese with English abstract)
- [4] Gorji T, Sertel E, Tanik A. Monitoring soil salinity via remote sensing technology under data scarce conditions: A case study from Turkey[J]. *Ecological Indicators*, 2017, 74: 384-391.
- [5] Singh A. Managing the salinization and drainage problems of irrigated areas through remote sensing and GIS techniques[J]. *Ecological Indicators*, 2018, 89: 584-589.
- [6] Liu L, Ji M, Buchroithner M. A case study of the forced invariance approach for soil salinity estimation in vegetation-covered terrain using airborne hyperspectral imagery[J]. *ISPRS International Journal of Geo-Information*, 2018, 7(2):48-59.
- [7] 王丹阳, 陈红艳, 王桂峰, 等. 无人机多光谱反演黄河口重度盐渍土盐分的研究[J]. *中国农业科学*, 2019, 52(10): 1698-1709.
Wang Danyang, Chen Hongyan, Wang Guifeng, et al. Salinity inversion of severe saline soil in the Yellow River Estuary based on UAV multi-Spectra[J]. *Scientia Agricultura Sinica*, 2019, 52(10): 1698-1709. (in Chinese with English abstract)
- [8] Wei G F, Li Y, Zhang Z T, et al. Estimation of soil salt content by combining UAV-borne multispectral sensor and machine learning algorithms[J]. *PeerJ*, 2020, 8(2): e9087.
- [9] Wang J Z, Ding J L, Yu D L, et al. Machine learning-based detection of soil salinity in an arid desert region, Northwest China: A comparison between Landsat-8 OLI and Sentinel-2 MSI[J]. *The Science of the Total Environment*, 2020, 707: 136092.
- [10] Peng J, Biswas A, Jiang Q, et al. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China[J]. *Geoderma*, 2019, 337: 1309-1319.
- [11] 韩霖昌, 李晓明. 盐碱地利用障碍因子高光谱遥感反演研究[J]. *光谱学与光谱分析*, 2013, 33(7): 1932-1935.
Han Jichang, Li Xiaoming. Research on hyperspectral remote sensing inversion of barrier factors in saline-alkaline land use[J]. *Spectroscopy and Spectral Analysis*, 2013, 33(7): 1932-1935. (in Chinese with English abstract)
- [12] 李娟, 陈超, 王昭. 基于不同变换形式干旱区土壤盐分高光谱特征反演[J]. *水土保持研究*, 2018, 25(1): 197-201.
Li Juan, Chen Chao, Wang Zhao. Inversion of high spectral characteristics of soil salt in arid area based on different transform forms[J]. *Research of Soil and Water Conservation*, 2018, 25(1): 197-201. (in Chinese with English abstract)
- [13] 吴裕, 申广荣, 刘璐, 等. 土壤有机质高光谱特征及其反演研究[J]. *上海交通大学学报(农业科学版)*, 2019, 37(4): 37-44.
Wu Yu, Shen Guangrong, Liu Lu, et al. Hyperspectral characteristics of soil organic matter and inversion methods[J]. *Journal of Shanghai Jiaotong University(Agricultural Science)*, 2019, 37(4): 37-44. (in Chinese with English abstract)
- [14] 彭金英, 蒋金豹, 崔希民, 等. 宁夏盐碱地土壤盐分高光谱反演[J]. *矿山测量*, 2021, 49(1): 69-73.

- Peng Jinying, Jiang Jinbao, Cui Ximin, et al. Hyperspectral inversion of the soil salt content in saline-alkali land of Ningxia[J]. *Mine Surveying*, 2021, 49(1): 69-73. (in Chinese with English abstract)
- [15] 朱赞, 申广荣, 项巧巧, 等. 基于不同光谱变换的土壤盐含量光谱特征分析[J]. *土壤通报*, 2017, 48(3): 560-568. Zhu Yun, Shen Guangrong, Xiang Qiaoqiao, et al. Spectral characteristics of soil salinity based on different pre-processing methods[J]. *Chinese Journal of Soil Science*, 2017, 48(3): 560-568. (in Chinese with English abstract)
- [16] 孙亚楠, 李仙岳, 史海滨, 等. 基于多源数据融合的盐分遥感反演与季节差异性研究[J]. *农业机械学报*, 2020, 51(6): 169-180. Sun Yanan, Li Xianyue, Shi Haibin, et al. Remote sensing inversion of soil salinity and seasonal difference analysis based on multi-source data fusion[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(6): 169-180. (in Chinese with English abstract)
- [17] 厉彦玲, 赵庚星, 常春艳, 等. OLI与HSI影像融合的土壤盐分反演模型[J]. *农业工程学报*, 2017, 33(21): 173-180. Li Yanling, Zhao Gengxing, Chang Chunyan, et al. Soil salinity retrieval model based on OLI and HSI image fusion[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2017, 33(21): 173-180. (in Chinese with English abstract)
- [18] Ouerghemmi W, Gomez C, Naceur S, et al. Semi-blind source separation for the estimation of the clay content over semi-vegetated areas using VNIR/SWIR hyperspectral airborne data[J]. *Remote Sensing of Environment*, 2016, 181: 251-263.
- [19] 罗红霞, 龚健雅, 蹇代君. 亚热带低山丘陵区土壤遥感监测图像处理[J]. *长江流域资源与环境*, 2006(1): 41-47. Luo Hongxia, Gong Jianya, Jian Daijun. Image processing of soil survey by remote sensing in subtropical hilly lands[J]. *Resources and Environment in the Yangtze Basin*, 2006(1): 41-47. (in Chinese with English abstract)
- [20] 李燕丽, 吴士文, 刘娅, 等. 基于FastICA盲源分离法去除土壤干扰的小麦生物量高光谱估算[J]. *生态学杂志*, 2017, 36(4): 1158-1164. Li Yanli, Wu Shiwen, Liu Ya, et al. Applying fast independent component analysis algorithm of blind source separation method to remove soil effects on hyperspectral data for wheat biomass estimation[J]. *Chinese Journal of Ecology*, 2017, 36(4): 1158-1164. (in Chinese with English abstract)
- [21] L Y, P X Z, Shi R J. et al. Predicting soil salt content over partially vegetated surfaces using non-negative matrix factorization[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(11): 5305-5316.
- [22] 刘娅, 潘贤章, 石荣杰, 等. 基于盲源分离的稀疏植被区土壤含盐量反演[J]. *土壤学报*, 2016, 53(2): 322-331. Liu Ya, Pan Xianzhang, Shi Rongjie, et al. Estimation of soil salt content over partially vegetated areas based on blind source separation[J]. *Acta Pedologica Sinica*, 2016, 53(2): 322-331. (in Chinese with English abstract)
- [23] Ouerghemmi W, Gomez C, Naceur S, et al. Applying blind source separation on hyperspectral data for clay content estimation over partially vegetated surfaces[J]. *Geoderma*, 2011, 163(3): 227-237.
- [24] Bartholomeus H, Kooistra L, Stevens A, et al. Soil organic carbon mapping of partially vegetated agricultural fields with imaging spectroscopy[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2011, 13(1): 81-88.
- [25] 赵伟, 包妮沙, 刘善军, 等. 地表枯枝落叶层影响下的土壤混合光谱特征及解混方法研究[J]. *光谱学与光谱分析*, 2020, 40(7): 2188-2193. Zhao Wei, Bao Nisha, Liu Shanjun, et al. Plant litter effect of the soil organic carbon estimation and unmixing method based on the visible-near infrared spectra[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(7): 2188-2193. (in Chinese with English abstract)
- [26] 黄静, 赵庚星, 奚雪, 等. 光谱与纹理信息结合的黄河三角洲土壤盐渍化信息提取: 以垦利区为例[J]. *农业资源与环境学报*, 2022, 39(3): 594-601. Huang Jing, Zhao Gengxing, Xi Xue, et al. Extraction of soil salinization information by combining spectral and textured data in the Yellow River Delta: A case study in Kenli District, Shandong Province[J]. *Journal of Agricultural Resources and Environment*, 2022, 39(3): 594-601. (in Chinese with English abstract)
- [27] 王伦平, 陈亚新, 曾国芳. 内蒙古河套灌区灌溉排水与盐碱化防治[M]. 北京: 水利电力出版社, 1993.
- [28] Liu Y L, Jiang Q H, Fei T, et al. Transferability of a visible and near-infrared model for soil organic matter estimation in riparian landscapes[J]. *Remote Sensing*, 2014, 6(5): 4305-4322.
- [29] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization[J]. *Nature*, 1999, 401(6755): 788-791.
- [30] 曹见飞. 基于空谱约束解混的土壤盐分高光谱遥感反演[D]. 济南: 山东师范大学地理与环境学院, 2021. Cao Jianfei. Hyperspectral Remote Sensing Inversion of Soil Salinity Based on Spatial-spectral Constrained Unmixing[D]. Jinan: College of Geograph and Environment, Shandong Normal University, 2021. (in Chinese with English abstract)
- [31] Wold S, Martens H, Wold H. The multivariate calibration problem in chemistry solved by the PLS method[J]. *Lecture Notes in Mathematics*, 1983, 973: 286-293.
- [32] Viscarra Rossel R A, McGlynn R N, McBratney A B. Determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy[J]. *Geoderma*, 2006, 137: 70-82.
- [33] Yue J B, Tian Q J, Tang S F, et al. A dynamic soil endmember spectrum selection approach for soil and crop residue linear spectral unmixing analysis[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2019, 78: 306-317.
- [34] Bioucas-Dias J M, Plaza A, Dobigeon N, et al. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*,

2012, 5(2): 354-379.

[35] Xu X, Shi Z W. Multi-objective based spectral unmixing for

hyperspectral images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 124: 54-69.

Spectral unmixing of straw and soil to estimate the soil salinity using non-negative matrix factorization

Li Zhen¹, Cao Jianfei^{1,2*}, Yang Han¹, Liu Jianhua¹, Wang Zhaohai¹, Duan Xinrong³, Zhang Letian⁴

(1. College of Geography and Environment, Shandong Normal University, Jinan 250014, China; 2. Shandong Dongying Institute of Geographic Sciences, Dongying 257000, China; 3. Shandong Provincial Institute of Land Surveying and Mapping, Jinan 250013, China; 4. China Environmental United (Beijing) Certification Center Co., LTD., Jinan 250000, China)

Abstract: Straw returning to the field has been used to promote the soil ecological environment in the Yellow River Delta. A remote sensing estimation of soil salt can greatly contribute to monitor the soil salinization in the straw covered areas. However, the mixed data has posed a great challenge on the pixel accuracy, due to both straw and soil spectral signals in the same pixel. In this study, ten groups of salinized surface scenes were firstly collected with different straw coverage. Multiple spectral measurements were then carried out on each sample, where the average value was taken as the reflection spectrum. The Non-negative Matrix Factorization (NMF) was selected to treat the mixed spectrum under the non-negative restriction in the process of spectral unmixing. The endmember matrix and abundance matrix were updated simultaneously by an iterative method, and finally two endmember spectra were extracted. The spectra of soil and straw were separated to evaluate the performance of the straw removal from the mixed spectrum, according to the retaining salinity information. Finally, the soil salt model was constructed using the Partial Least Square Regression (PLSR), further to verify the effectiveness of NMF for the spectral unmixing and extracting soil spectrum. The results show that: 1) The soil spectral reflectance decreased gradually with the increase of the salinity in the whole wavelength range, but the trend of curves remained similar. More importantly, the straw coverage increased the spectral reflectance of soil. The spectrum of straw coverage soil was much more similar to the straw spectrum, as the degree of the straw cover increased. In addition, there were the absorption characteristics near 1 730 and 2 090 nm in the straw covered soil, compared with the pure soil. 2) The soil spectrum was effectively separated from the straw mixed spectrum after NMF unmixing. There were more outstanding characteristics of straw in the separated straw spectrum on the whole with the increase of straw coverage, when the degree of the straw cover reached 15%. The typical characteristics of straw were achieved in the separated straw spectrum, when the straw coverage was less than 25%. There was a significant fluctuation in the spectral curve of the extracted soil and straw, indicating the unstable soil reflection, when the straw coverage increased to more than 30%. It infers that the performance of NMF was confined to the low straw coverage in this case. 3) The soil spectra that extracted from the straw mixed spectra was successfully utilized to estimate the soil salinity. The high accuracy of the model was achieved using the soil spectral data after NMF spectral unmixing, compared with the original mixed ones. Specifically, the average coefficient of determination, R^2 , of the soil salt model was 0.68, the average Root-Mean-Square Error (RMSE) was 6.34, and the average ratio of Prediction Deviation (RPD) was 1.47 after NMF unmixing. Among them, the average R^2 and RPD were 0.07 and 0.25 higher than the original mixing, respectively, whereas, the average RMSE was 1.22 lower. The RPD of the improved model was also greater than 1.4, when the degree of the straw cover was less than 25%. The findings can provide a practical basis to efficiently improve the near-ground remote-sensing estimation accuracy of straw covered saline lands in the Yellow River Delta.

Keywords: soils; model; salts; Non-negative Matrix Factorization (NMF); ground-based remote sensing