DOI：10.11766/trxb202506250309 **CSTR：32215.14.trxb202506250309**

谢平如，洪永胜，徐向华，严国菁，张超，田康，樊亚男，陈剑，胡文友. 分数阶微分和光谱指数结合的典型黑土区有机质光谱预测研究[J]. 土壤学报，2026，63（5）：1490–1503.

XIE Pingru，HONG Yongsheng，XU Xianghua，Yan Guojing，ZHANG Chao，TIAN Kang，FAN Yanan，CHEN Jian，HU Wenyou. Spectral Prediction of Soil Organic Matter in Typical Black Soil Regions by Combining Fractional-Order Derivatives and Spectral Indices[J]. Acta Pedologica Sinica，2026，63（5）：1490–1503.

分数阶微分和光谱指数结合的典型黑土区有机质光谱预测研究*

谢平如[1, 2]，洪永胜[3]，徐向华[1†]，严国菁[2, 4]，张 超[2]，田 康[2]，樊亚男[2]，陈 剑[2, 4]，胡文友[2, 4†]

（1. 南京信息工程大学生态与应用气象学院，南京 210044；2. 土壤与农业可持续发展全国重点实验室（中国科学院南京土壤研究所），南京 211135；3. 武汉大学资源与环境科学学院，武汉 430079；4. 中国科学院大学，北京 100049）

摘 要：土壤有机质（Soil organic matter，SOM）含量的快速、准确预测对土壤肥力评价和农业可持续发展至关重要。本研究基于室内可见-近红外（Visible and near-infrared，Vis-NIR）光谱数据，构建分数阶微分（Fractional-order derivative，FOD）与光谱指数协同建模策略，用于东北典型黑土区 SOM 含量的快速预测。采集 227 份耕地土壤样本，获取室内 Vis-NIR 光谱数据和 SOM 实测值，并对光谱数据进行 0～2 阶（间隔 0.1）的 FOD 处理。计算二维和三维光谱指数，并分析其与 SOM 含量间的相关性，筛选出最优光谱指数。采用随机森林（Random forest，RF）和 Cubist 两种建模方法构建 SOM 光谱反演模型。结果表明，FOD 处理可增强微弱的土壤光谱吸收特征，有效提升模型对 SOM 含量的预测能力。相比原始光谱和整数阶微分（一阶和二阶），基于 0.3 阶微分处理土壤光谱数据建立的 Cubist 模型性能最佳，验证 R^2 为 0.74。RF 模型则适合 1.6～1.9 阶处理的土壤光谱数据，其 R^2 稳定在 0.63～0.65。相比二维光谱指数，三维光谱指数表现出与 SOM 更高的相关性，证实了多波段交互信息在提高特征变量解释力方面的潜力。基于最优光谱指数数据集建立的 RF 和 Cubist 模型在多个分数阶微分处理下展现出良好的模型精度。本研究验证了 FOD 处理和光谱指数在 SOM 预测中的可行性，可为光谱技术在复杂农业场景中进行高精度监测提供重要参考和技术支撑。

关键词：黑土地；土壤监测；土壤有机质；分数阶微分；光谱指数

中图分类号：S153.6 **文献标志码：**A

Spectral Prediction of Soil Organic Matter in Typical Black Soil Regions by Combining Fractional-Order Derivatives and Spectral Indices

XIE Pingru[1, 2], HONG Yongsheng[3], XU Xianghua[1†], Yan Guojing[2, 4], ZHANG Chao[2], TIAN Kang[2], FAN Yanan[2],

* 国家重点研发计划项目课题（2024YFD1501102）、中国科学院南京土壤研究所“十四五”自主部署项目（ISSAS2418）和国家自然科学基金项目（42571068）资助 Supported by the National Key Research and Development Program of China（No.2024YFD1501102），the "14th Five-Year Plan" Self-deployment Project of Institute of Soil Science，Chinese Academy of Sciences（No.ISSAS2418）and the National Natural Science Foundation of China（No.42571068）

† 通讯作者 Corresponding author，E-mail：wyhu@issas.ac.cn，xuxianghua@nuist.edu.cn

作者简介：谢平如（2001—），女，山东济宁人，硕士研究生，主要从事区域土壤遥感监测与数字制图研究。E-mail：prxie@nuist.edu.cn

收稿日期：2025–06–25；收到修改稿日期：2026–01–27；网络首发日期（www.cnki.net）：2026–04–13

CHEN Jian[2, 4], HU Wenyou[2, 4†]

(1. *College of Ecology and Applied Meteorology*, *Nanjing University of Information Science and Technology*, *Nanjing* 210044, *China*; 2. *State Key Laboratory of Soil and Sustainable Agriculture*, *Institute of Soil Science*, *Chinese Academy of Sciences*, *Nanjing* 211135, *China*; 3. *School of Resource and Environmental Sciences*, *Wuhan University*, *Wuhan* 430079, *China*; 4. *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*)

Abstract: 【Objective】Rapid and accurate estimation of soil organic matter (SOM) is crucial for assessing soil fertility, guiding sustainable agricultural management, and supporting carbon accounting at regional and global scales. SOM is a key indicator of soil quality, influencing nutrient cycling, microbial activity, crop productivity, and soil carbon sequestration potential. While reliable, traditional chemical analysis methods are costly, time-consuming, and destructive, making them unsuitable for large-scale or repeated monitoring. Visible and near-infrared (Vis-NIR) spectroscopy provides a rapid, non-destructive, and environmentally friendly alternative for SOM assessment. However, the effectiveness of Vis-NIR spectroscopy is often limited by spectral noise, baseline drift, and low sensitivity to absorption features associated with organic components. Therefore, developing advanced spectral transformation and modeling strategies that can enhance weak spectral signals and extract effective features related to SOM is essential. This study aims to construct a collaborative modeling framework combining fractional-order derivative (FOD) transformation and spectral indices to improve the interpretability and predictive accuracy of Vis-NIR spectral data from typical black soil regions in Northeast China.【Methods】A total of 227 soil samples were collected from representative farmland in the black soil region, an important grain-producing area in Northeast China. The reflectance spectra and SOM content were obtained in the laboratory. The spectral data were processed with FOD ranging from 0 to 2.0 (increment by 0.1 at each step). Two-dimensional (2D) and three-dimensional (3D) spectral indices were calculated to explore the interaction information between different wavelength combinations. The correlation between each spectral index and SOM content was analyzed to determine the most sensitive index. Two machine learning algorithms—random forest (RF) and Cubist—were used to construct prediction models. The input datasets were divided into two categories: (1) FOD-transformed reflectance(FOD dataset); and(2)spectral indices that were most strongly correlated with SOM (index dataset). Four models were thus constructed: FOD-RF, Index-RF, FOD-Cubist, and Index-Cubist. Ten-fold cross-validation was used to evaluate the performance of the model, and the determination coefficient (R^2) and root mean square error (RMSE) were used as evaluation indexes. In addition, this study also analyzes the importance of model characteristics to determine the key wavelength or exponential combination that is helpful for SOM prediction.【Results】FOD transform significantly improved the spectral interpretability and enhanced the detection of weak organic absorption characteristics in the Vis-NIR band. The Cubist model with the 0.3-order derivative spectrum exhibited the best performance and provided a validation R^2 of 0.74. The RF model performs best at higher derivatives (1.6-1.9), and the R^2 value remains between 0.63 and 0.65. In addition, the correlation between 3D spectral index and SOM is stronger than that of 2D index, and 3D spectral index improves the interpretability of features. The characteristic importance analysis showed that the most sensitive spectral regions predicted by SOM were located within 1410-1880 nm and 2200-2350 nm, corresponding to the overtone and combined vibration of C-H, N-H and O-H functional groups.【Conclusion】The combination of FOD preprocessing and spectral index provides a robust, flexible and scalable framework for estimating SOM using Vis-NIR spectroscopy. This study emphasizes a promising direction in the field of intelligent soil remote sensing monitoring and information technology, and provides methodological progress for digital soil mapping, precise nutrient management and sustainable land management. Its application potential is not only limited to the prediction of SOM, but also extended to a wider range of soil property assessment, which provides a theoretical and technical basis for the construction of intelligent soil information system to support the sustainable development of agriculture in major grain producing areas such as northeast China.

Key words: Black soil; Soil monitoring; Soil organic matter; Fractional-order derivative; Spectral indices

土壤有机质（Soil organic matter，SOM）是影响土壤肥力、土壤结构稳定性和碳循环的重要指标，对维持农业生产力和生态系统功能具有关键作用[1]。黑土地因肥力高、性状好、适宜农耕，被誉为“耕地中的大熊猫”[2]。东北黑土区作为我国重要的粮食生产基地，对保障国家粮食安全和农业可持续发展具有战略意义[3]。然而，受长期重用轻养和气候变化等因素影响，东北黑土地的 SOM 含量持续下降，严重制约了黑土地的可持续利用和农业的可持续发展[4]。传统的化学分析方法虽然精度较高，但土壤样品前处理复杂、耗时费力，不适合区域大范围快速监测与评价[5]。近年来，随着检测技术的发展，在 SOM、重金属含量与 pH 等属性的快速高效检测方面均取得了显著进展[6-7]。可见-近红外（Visible and near-infrared，Vis-NIR）光谱技术因其速度快、成本低和无损的优势，已成为 SOM 等理化性质快速预测的重要手段[8]。但原始高光谱数据存在噪声干扰严重、波段冗余度高等问题，难以充分挖掘 SOM 细微的光谱吸收特征。微分处理是提升模型预测精度的有效手段之一，其中，整数阶光谱微分（如一阶和二阶）在一定程度上突出吸收特征，但易放大噪声，使光谱解释更加复杂[9]。分数阶微分（Fractional-order derivative，FOD）可在整数阶之间进行任意插值，既能增强光谱细节，又能抑制噪声累积，在光谱建模研究中展现出较好的应用潜力[10-11]。

在变量选择方面，光谱指数通过两波段或多波段组合，可以挖掘波段间协同变异信息，减少光谱数据冗余问题，常用的二维（Two-dimensional，2D）光谱指数在土壤属性预测中已取得一定成效[12]。三维（Three-dimensional，3D）光谱指数引入第三个波段，能够较好地挖掘波段间的交互作用，进而提升模型精度[13]。已有研究将 FOD 与 2D 光谱指数结合用于土壤属性反演[14]，但关于 FOD 与 3D 光谱指数结合预测 SOM 的研究仍较少。此外，建模方法也对预测精度具有重要影响。早期研究以线性偏最小二乘回归模型（Partial least squares regression，PLSR）为主，不能有效处理土壤属性与光谱数据之间的非线性关系[15]。近年来，随着机器学习方法的发展，随机森林（Random forest，RF）、支持向量机（Support vector machine，SVM）、Cubist 等方法被广泛应用[16-18]。一般而言，在样本量充足时，机器学习方法构建的模型预测精度优于线性模型[19]。本研究选择 RF 与 Cubist 模型进行对比研究，原因在于二者在处理高维、共线性光谱数据方面均表现优越，但其机理各异。RF 具有高精度和抗过拟合等优点，而 Cubist 模型则结合了回归树和线性模型的双重优势，能生成易于解释的规则集，并且能够进一步提升预测能力。这种对比不仅旨在筛选出用于 SOM 监测的最佳模型，也期望从模型可解释性角度深化对光谱特征与 SOM 之间关系的理解。

本研究以东北典型黑土区为研究对象，基于室内 Vis–NIR 高光谱数据与实测 SOM 含量，评估了 FOD 与光谱指数在特征提取中的协同作用。通过构建不同 FOD 处理下的光谱指数特征集，分别采用 RF 与 Cubist 模型对 SOM 含量进行建模与预测，明确了不同 FOD 阶数与特征组合的适配关系，筛选出了最优模型组合。研究结果可为东北黑土区 SOM 含量的快速高精度预测提供方法和技术支撑。

1 材料与方法

1.1 研究区概况

研究区位于吉林省西部及黑龙江省西南部（图 1a），地处我国东北典型黑土区的核心地带[20]，海拔主要介于 91～200 m 之间，属温带大陆性季风气候区，是我国最重要的商品粮生产基地之一。该区域土壤发育良好，典型的黑土、黑钙土和草甸土广泛分布，具有代表性强、土壤类型典型的特点[21]。其中，黑土主要分布在研究区东部和北部，是中国黑土分布最集中、有机质累积最丰富的区域之一；黑钙土集中分布在中部；草甸土则分布于河漫滩及地下水位较浅的洼地。此外，研究区是我国“黄金玉米带”的核心区域和重要的大豆主产区之一，农业生产对土壤质量高度依赖[22]。该区域集中了典型黑土、主要粮食作物产区和高强度农业利用三大特征[3]，在东北黑土区土壤质量监测及黑土地保护与可持续利用研究中具有很好的代表性和重要的现实意义。

1.2 土壤样品采集与测试

研究采用网格布点法在大面积耕地上均匀布设样点，避开道路、堆肥区及排灌口等干扰源，采用五点混合采样法，共采集 227 个 0～20 cm 表层土壤样品（图 1b）。同时，使用便携式 GPS 设备记录每个样本中心点的坐标。将采集的土壤样品送至实验

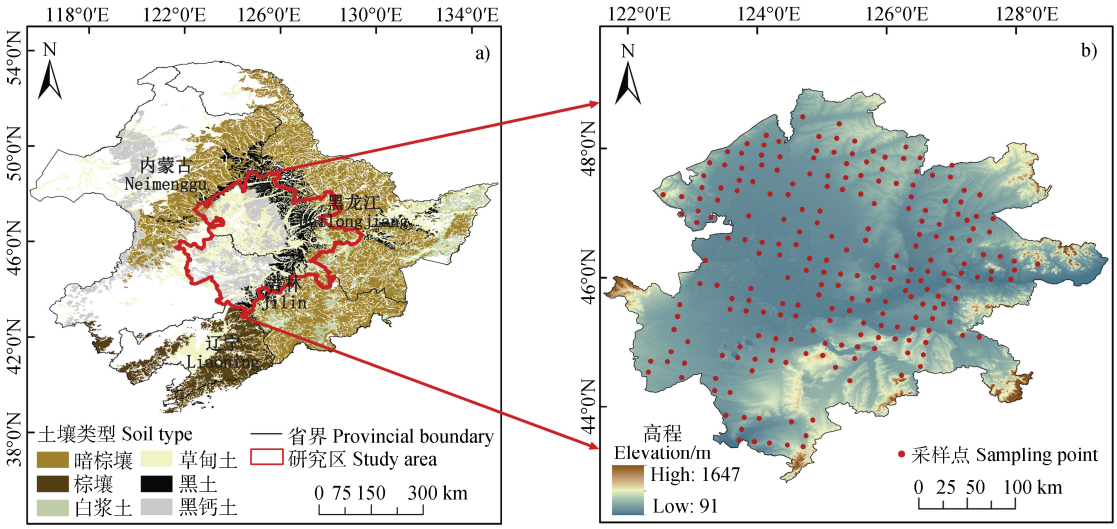


图 1　研究区位置（a）与采样点分布（b）

Fig. 1　Study area location（a）and sampling point distribution（b）

室分析，风干，去除石砾和植物碎片，研磨过 2 mm 孔径筛。将制备好的土样分为两份，一份用于光谱数据的测定，一份采用重铬酸钾容量法测定 SOM 含量[23]。

使用 ASD FieldSpec 3 光谱仪测定土壤样本的 Vis-NIR 光谱数据。该光谱仪波长范围为 350～2500 nm，在 350～1000 和 1000～2500 nm 区域的光谱分辨率分别为 3 和 10 nm，重采样间隔为 1 nm，输出 2151 个波段。所有光谱测量均在暗室中进行，以尽量减少杂散光的干扰。光源入射角设置为 45 °，传感器垂直放置于土壤样本上方，每测定 10 个样本进行 1 次白板校正。每个样本采集 10 次反射光谱，取平均值作为最终的反射光谱数据。去除噪声过大的边缘波段（350～399 和 2401～2500 nm）。对 400～2400 nm 范围内的光谱数据进行 Savitzky-Golay 平滑处理，窗口大小为 11，多项式阶数为 2。对每个土壤样品的反射光谱进行间隔 10 nm 采样，得到 201 个波段[9]。

1.3　分数阶微分处理

FOD 可将传统的整数阶微分推广至任意阶，可实现微分之间的连续插值。本研究采用 Grünwald–Letnikov（G–L）定义[9]，FOD 的数学公式如下：

在[a，b]（G-L）区间上的 v 阶微分公式：

$$d^{v}f(x)=\lim_{h\to 0}\frac{1}{h^{v}}\sum_{m=0}^{\frac{b-a}{h}}(-1)^{m}\frac{\Gamma(v+1)}{m!\Gamma(v-m+1)}f(x-mh) \quad (1)$$

式中，v 为微分阶数；x 为波段光谱值；h 为步长；[（b-a）/h]为（b-a）/h 的整数部分。Γ（x）表示伽马函数，特征为：

$$\Gamma(z)=\int_{0}^{\infty}exp(-u)u^{z-1}\,du=(z-1)! \quad (2)$$

本研究中的反射光谱已被重采样至 10 nm 的分辨率，h 设置为 1。式（1）可以转换为：

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

在 Matlab R2023b 中实现从 0 到 2 阶（间隔 0.1）的 FOD 变换过程。

1.4　土壤光谱指数构建

本研究选取差值指数（Difference Index，DI）、比值指数（Ratio Index，RI）和归一化指数（Normalized Difference Index，NDI）三个 2D 光谱指数[24]。DI、RI 和 NDI 的计算公式如下：

$$DI = R_i - R_j \quad (4)$$

$$RI = R_i / R_j \quad (5)$$

$$NDI = (R_i - R_j) / (R_i + R_j) \quad (6)$$

式中，R_i 和 R_j 分别为波段选取的反射率值，且 $i \neq j$。

通过在 2D 光谱指数中引入一个新的波段，生成 3D 光谱指数。参考先前研究，计算了 TBI1、TBI2、TBI3、TBI4 四个三维光谱指数[25]。

$$TBI1 = (R_i - R_j) / (R_i + R_k) \quad (7)$$

$$TBI2 = R_i / (R_j + R_k) \quad (8)$$

$$TBI3 = R_i / (R_j \times R_k) \quad (9)$$

$$TBI4 = (R_i - R_j) / (R_j - R_k) \quad (10)$$

式中，R_i、R_j、R_k 分别为波段选取的反射率值，且 $i \neq j \neq k$。

使用 Matlab R2023b 计算 FOD 处理后的光谱数据的 2D 和 3D 光谱指数。将获得的光谱指数与 SOM 进行相关性分析，确定最大相关系数对应的波长组合。

1.5 建模与验证

本研究将土壤有机质含量进行排序，采用间隔为 3∶1 的分层系统抽样，将 227 个样本划分为训练集（170 个）与验证集（57 个）。

RF 是一种集成学习算法，通过采样构建多棵决策树，在节点分裂时随机选择特征子集，最终通过投票或平均获得预测结果。RF 模型是在 R 语言中使用 *randomForest* 和 *caret* 包构建的。Cubist 是一种结合决策树与多元线性回归的模型，其核心是通过生成一系列规则将数据划分为子集，并在每个子集内建立独立的线性回归模型。Cubist 模型在 R 语言中通过 *Cubist* 包和 *caret* 包构建。

为评估不同 FOD 阶数下光谱指数在土壤高光谱数据中的应用潜力，本研究比较了两种光谱数据集的建模性能：FOD 变换后的光谱反射率（FOD 光谱数据集）和相关性最高的光谱指数数据（Index 数据集）。采用 RF 和 Cubist 两种机器学习模型进行建模，共构建四种方法：FOD-RF（RF 模型应用于 FOD 光谱数据集）、Index-RF（RF 模型应用于 Index 数据集）、FOD-Cubist（Cubist 模型应用于 FOD 光谱数据集）以及 Index-Cubist（Cubist 模型应用于 Index 数据集）。

计算 SOM 含量的一些基本描述性统计，包括最小值、最大值、平均值、标准偏差（Standard Deviation，SD）和变异系数（Coefficient of Variation，CV）。应用两个统计指标来评估和比较模型的性能：决定系数（Coefficient of Determination，R^2）、均方根误差（Root Mean Square Error，RMSE）。较高的 R^2 和较低的 RMSE 表明模型更有效。

2 结 果

2.1 研究区土壤有机质含量的统计特征

研究区 227 个样点的描述统计分析（表 1）结果表明，总体数据集的 SOM 含量范围为 4.31～66.59 g·kg^{-1}，均值为 29.55 g·kg^{-1}，标准差为 12.54 g·kg^{-1}，CV 为 42.43%。研究区 SOM 含量存在较大的变异性，有助于提高校准模型的预测能力。建模集的 SD 和 CV 分别为 12.47 g·kg^{-1} 和 42.24%，验证数据集的 SD 和 CV 分别为 12.84 g·kg^{-1} 和 43.36%。结果表明，本研究采用的训练集与验证集划分方法有效，能够充分代表数据集的总体特征。

表 1 土壤有机质含量统计特征

Table 1 Statistical characteristics of soil organic matter content

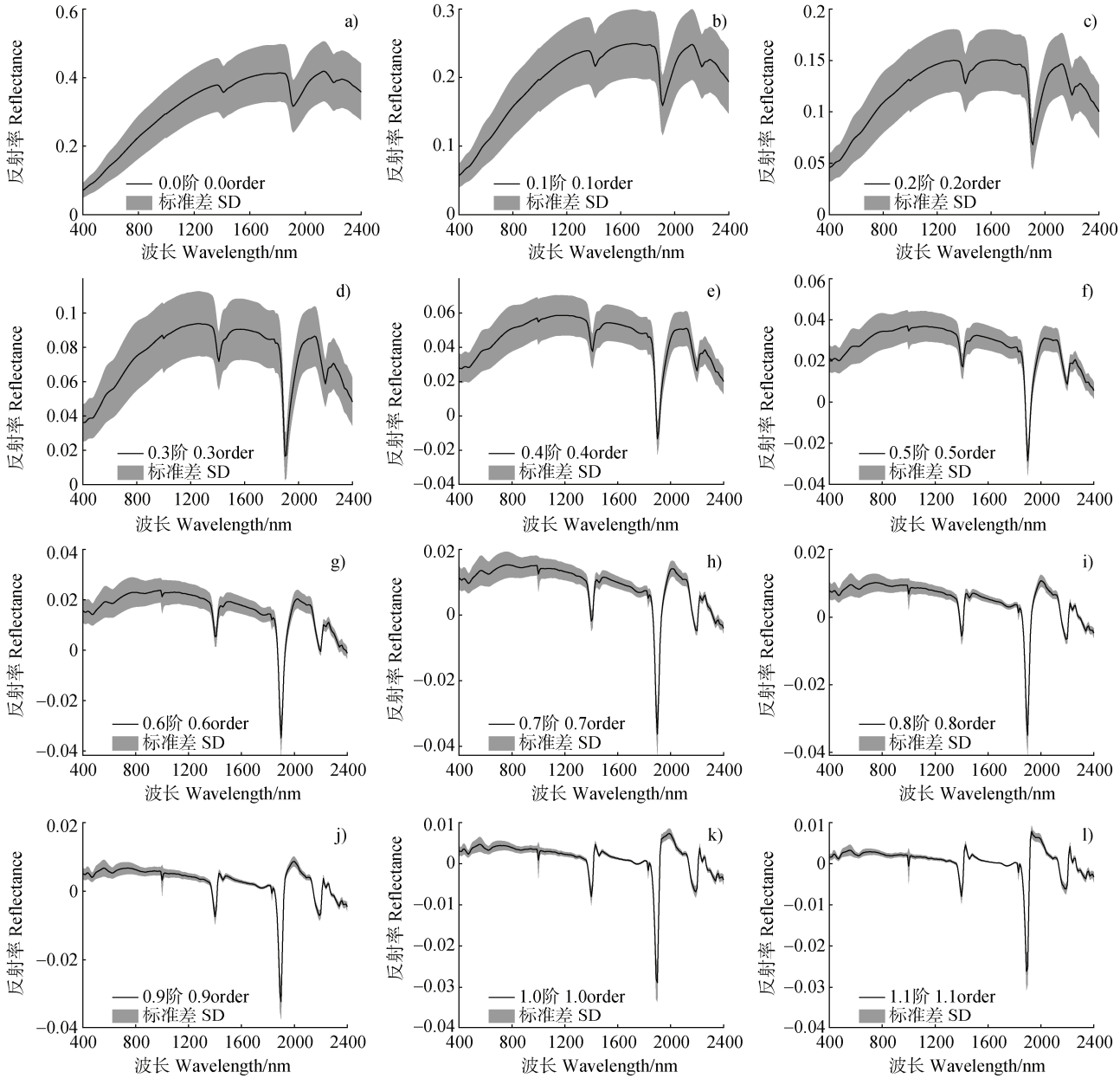
数据集 Data set	样本数 Sample size	最小值 Minimum/（g·kg^{-1}）	最大值 Maximum/（g·kg^{-1}）	均值 Mean/（g·kg^{-1}）	标准差 SD/（g·kg^{-1}）	变异系数 CV/%
全集 Whole set	227	4.31	66.59	29.55	12.54	42.43
建模集 Calibration set	170	4.31	66.59	29.52	12.47	42.24
验证集 Validation set	57	4.81	64.61	29.62	12.84	43.36

2.2　原始光谱与分数阶微分处理后的光谱特征

FOD 处理后的土壤光谱曲线发生显著变化（图 2）。图 2a 为原始光谱反射率曲线，在 1400、1900 和 2200 nm 附近存在明显的吸收峰。原始高光谱信号往往受波段冗余、基线漂移和噪声干扰等因素影响，直接建模易导致信噪比较低、模型泛化性能不足。图 2b～图 2u 展示了不同 FOD 处理下的光谱变化。经过 FOD 处理后，光谱曲线发生显著变化。当阶数从 0 增加至 1 时，3 个吸收峰变得越来越明显。阶数从 1 增加至 2 时，光谱曲线形状变化不明显。随着阶数的增加，大部分反射率逐渐趋近于 0。

2.3　光谱指数与土壤有机质的相关性

2.3.1　二维光谱指数与土壤有机质相关性　不同 2D 光谱指数对 FOD 处理的响应不同（图 3）。在所有 2D 光谱指数中，基于 1.5 阶 FOD 处理计算的 NDI 与 SOM 最大相关系数达到 0.82。基于 0.5 阶 FOD 反射率的 DI 与 SOM 的最大相关系数为-0.80，基于 1.4 阶 FOD 反射率的 RI 与 SOM 的最大相关系数为 0.81。相比之下，原始反射光谱（0 阶）的 DI、RI、NDI 最大的相关系数分别为–0.70、0.73 和 0.72。除 1.8 阶、1.9 阶和 2 阶的 RI 以及 2 阶的 NDI 外，其余所有 FOD 处理下的最大相关系数均高于原始光



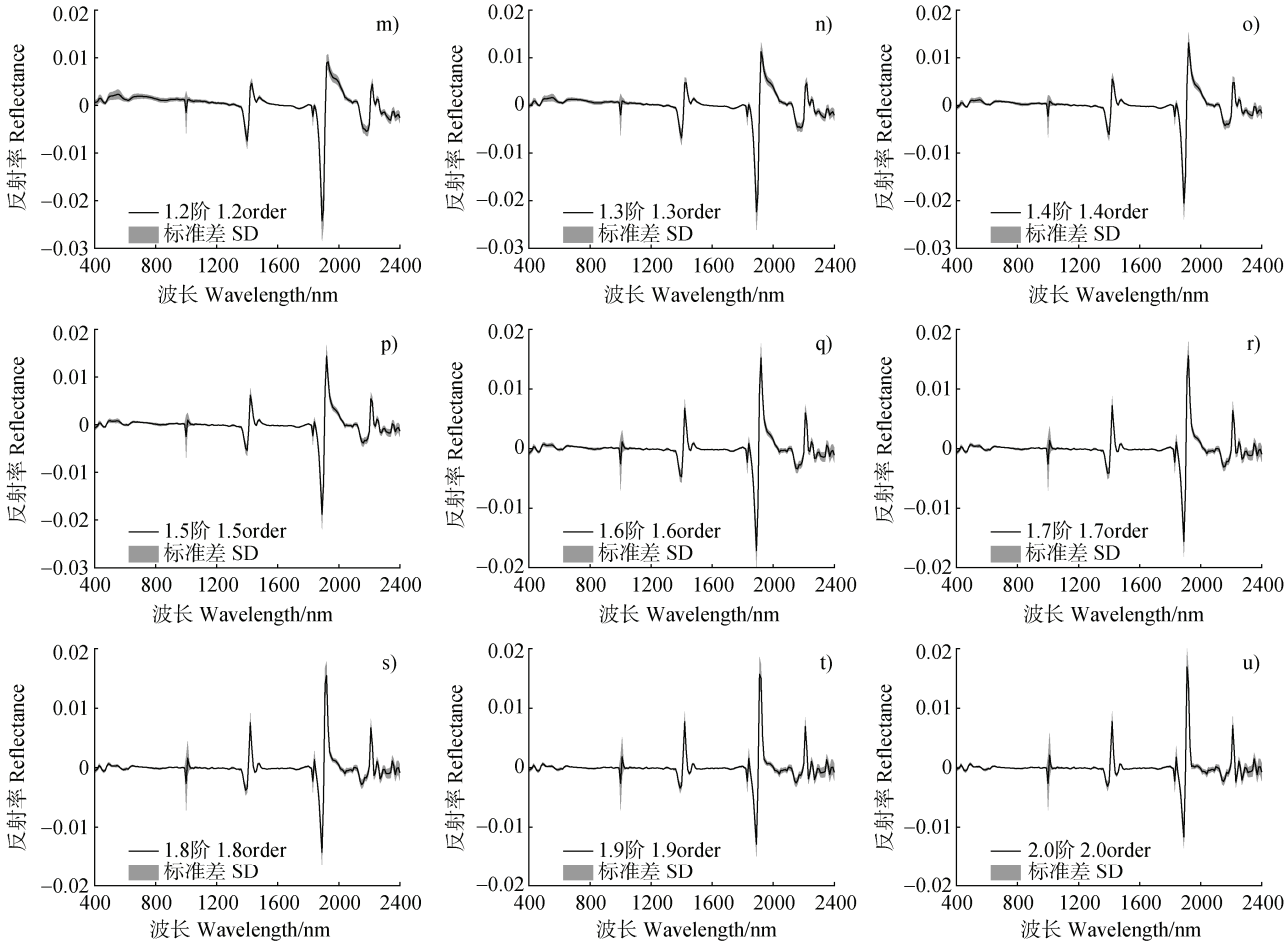


图 2 不同分数阶微分下的土壤有机质平均光谱曲线

Fig. 2 Average spectral curves of SOM at different fractional-order derivatives

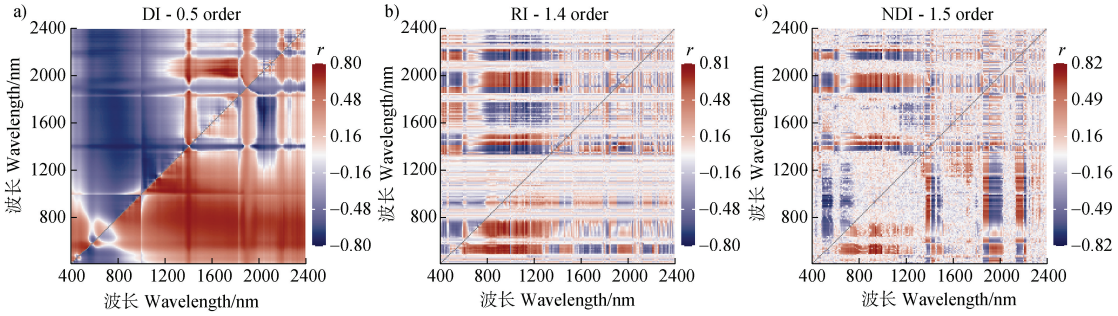


图 3 土壤有机质与最优二维光谱指数之间的相关性

Fig. 3 Correlations between SOM and the optimal 2D spectral indices

谱。结果进一步表明 FOD 可以提供与 SOM 相关的更多详细光谱信息。

DI 指数对 SOM 的敏感波段主要集中在 430～560、1390～1450、1860～2050 以及 2360 nm 等区域。RI 指数对 SOM 的敏感波段主要分布在 490～590、940～990 和 1390～1420 nm 区域。NDI 指数对 SOM 的敏感波段主要分布在 500～550、940～990、1400～1440 以及 2380 nm 等区域。在多个波段组合中出现 990 nm，表明该波段在 SOM 光谱响应中具有较强的辨识能力。

2.3.2 三维光谱指数与土壤有机质相关性 与 2D 光谱指数相比，3D 光谱指数增加了一个新的波段，增强了识别与 SOM 相关的更复杂的光谱响应模式的潜力。四个 3D 光谱指数均呈现出与 SOM 的高

相关性（图 4）。TBI1 在 1.4 阶的相关系数 r 值达到 0.85，对应波段为 990、1670 和 1420 nm；TBI2 在 0.4 阶表现最佳，r 值为 0.85，波段组合为 1400、2380 和 1430 nm；TBI3 的最高相关性出现在 1.2 阶，r 值为 0.82，波段组合为 1050、1920 和 750nm；而 TBI4 的最优相关性出现在 0.2 阶，r 值为 0.86，波段为 1360、2390 和 1390 nm。与初始光谱反射率（0 阶）相比，相关性分别提高了 16.44%、10.39%、15.49%和 4.88%。同时，3D 光谱指数普遍优于 2D 指数。

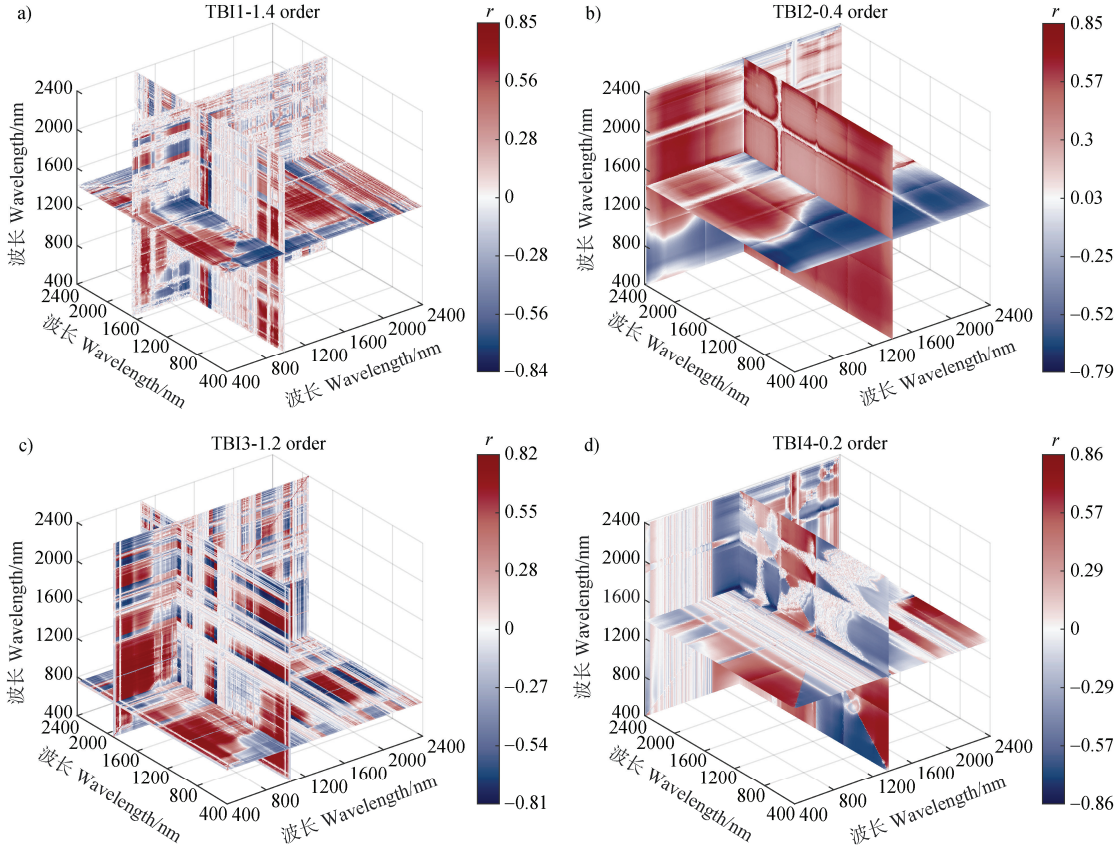


图 4　土壤有机质与最优三维光谱指数之间的相关性

Fig. 4　Correlations between SOM and the optimal 3D spectral indices

相关性最高的波段组合根据 FOD 变换而变化。SOM 对 3D 光谱指数响应的敏感波段集中于多个关键区域，可见光区主要集中在 430～570 nm，近红外区主要集中在 940～1050、1390～1440、1580～1670 以及 1880～2390 nm。不同阶数下，SOM 含量与 3D 光谱指数的最大相关系数分别为 0.82、–0.85、0.86、–0.85、0.85、0.84、0.83、0.84、0.83、0.82、0.82、0.81、0.82、0.83、0.85、0.84、0.82、–0.82、–0.81、–0.79、–0.79，其中最大相关系数出现在 0.2 阶 TBI4，r = 0.86。

2.4　模型性能比较

为探究 FOD 处理与光谱指数对模型预测土壤有机质性能的影响，本文基于不同 FOD 处理（0～2 阶，间隔 0.1）和光谱指数构建了 FOD-RF、Index-RF、FOD-Cubist 以及 Index-Cubist 四种模型，并采用 R^2 和 RMSE 作为评价指标（图 5）。结果表明，大多数 FOD 预处理提升了模型性能。对于 FOD-RF 模型，其 R^2在 1.7 阶最高为 0.65，RMSE 为 7.53 g·kg^{-1}。光谱指数构建的 Index-RF 模型在 1.6 阶微分下，R^2达 0.69，优于 FOD-RF 模型（R^2为 0.63）。FOD-Cubist 模型在所有处理方法中表现最优，在 0.3 阶达到的最高 R^2为 0.74，随后随阶数增加而逐步下降。Index-Cubist 模型在 1.6 阶达到的最高 R^2 为 0.73。

通过 SOM 含量实测值与 SOM 含量预测值的散点图来展现建模结果（图 6）。基于 1.7 阶 FOD 光谱的 RF 模型结果、基于 1.6 阶光谱指数的 RF 模型结

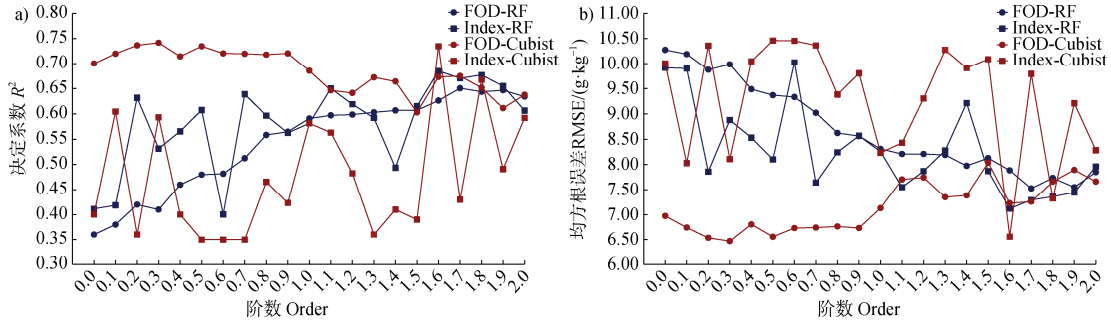


图 5 不同 FOD 变换下四种模型的精度

Fig. 5 Accuracy metrics for four models under different FOD transformations

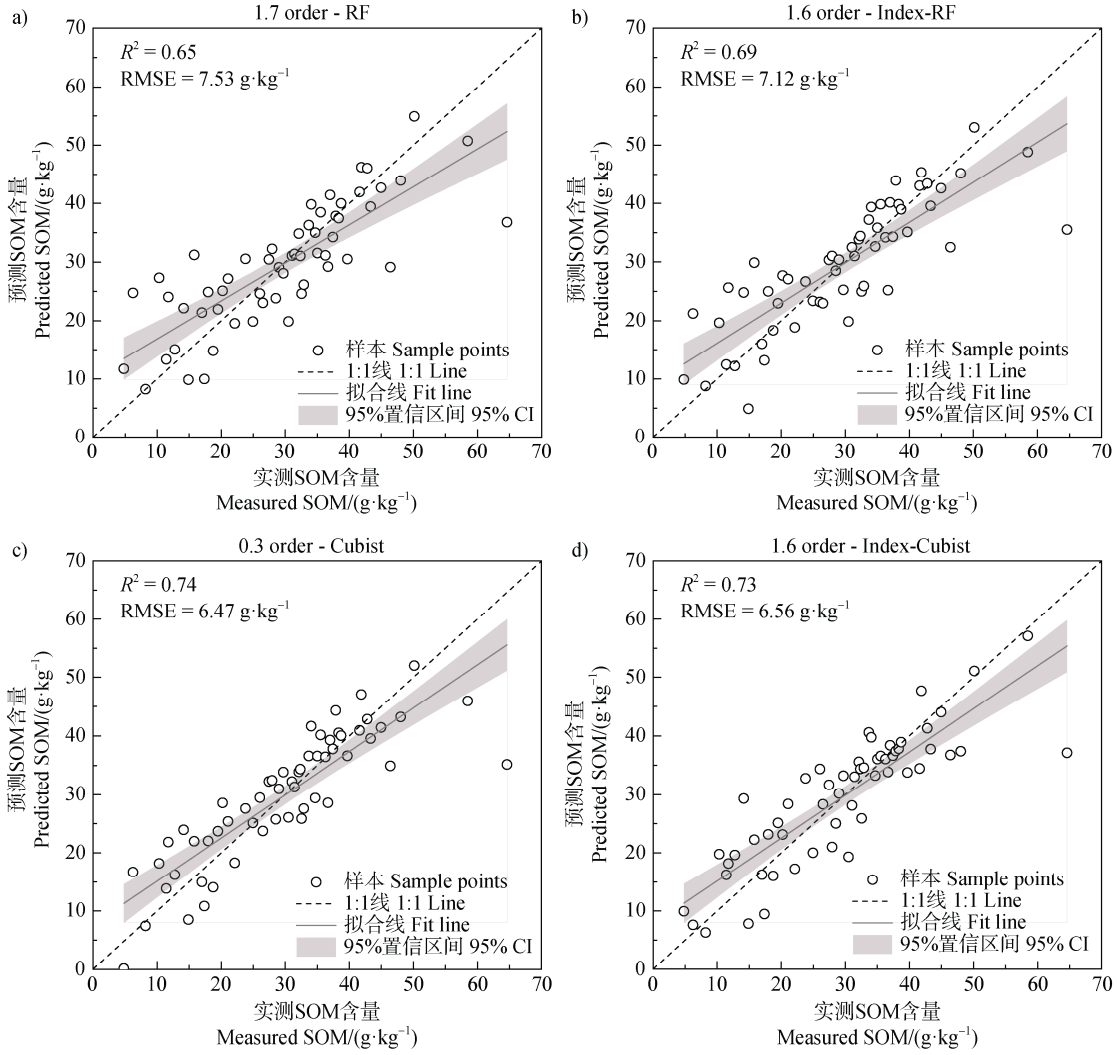


图 6 土壤有机质含量实测值与预测值比较

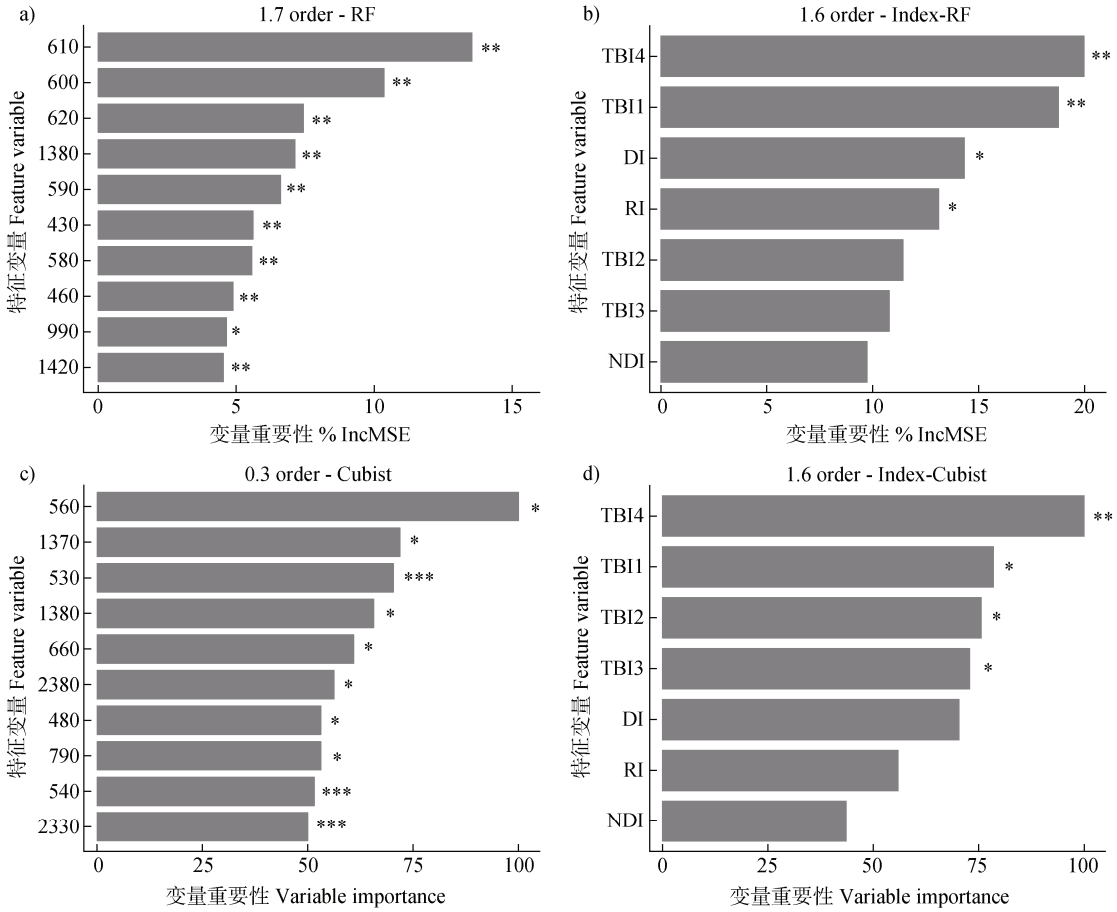
Fig. 6 Scatter plots of the measured SOM content vs. predicted SOM content

果、基于 0.3 阶 FOD 光谱的 Cubist 模型结果以及基于 1.6 阶光谱指数的 Cubist 模型结果的回归线均沿着 1∶1 线很好地分布。该分布表明，本研究构建的模型可以很好地预测 SOM 含量。

2.5 光谱变量重要性

不同建模方法的变量重要性不同（图 7）。对于 RF 模型，变量重要性以均方误差增加百分比（%IncMSE）度量，其值越大表明该变量对预测模型

准确性的贡献越大；对于 Cubist 模型，则由变量的使用频率决定。对于 FOD-RF 模型，关键波长变量主要集中于三个区间：580～620、990 和 1380～1420 nm。对于 Index-RF 模型，模型对光谱指数特征的敏感性较高。FOD-Cubist 模型结果显示 560 nm 波段的重要性极高，此外还在 530～660、1370～1380 和 2330～2380 nm 区间呈现多个重要波长。Index-Cubist 模型结果显示 3D 光谱指数重要性极高。Index-RF 模型和 Index-Cubist 模型光谱指数对应的波段为 500～540、990、1370～1420、1880～2180 nm。综上，四种模型的重要波段主要集中在 460～640 和 1390～1420 nm。



注 Note：* P<0.05；** P<0.01；*** P<0.001.

图 7　不同模型变量的重要性分析

Fig. 7　Variable importance analysis of different models

3　讨　论

土壤光谱经过微分处理，基线偏移效应和重叠峰逐渐消除。传统整数阶导数（一阶和二阶）可在一定程度上强化谱线特征、抑制背景干扰，但也容易放大高频噪声，造成信息损失与过拟合风险[26]。FOD 作为整数阶微分的广义扩展，能在保留谱线整体结构的同时增强局部细节。但随着阶数的增加，光谱噪声也有所增加。因此，应选择适当的阶数处理以达到增加细节和减少噪声之间的平衡。

2D 光谱指数（如 RI、DI、NDI）通过构造波段差值、比值等形式压缩数据维度、突出谱间差异，已被广泛用于土壤属性预测。分析 2D 光谱指数与 SOM 的相关性，可研究两个不同波段的协同效应[24]。然而，2D 光谱指数的表达能力受限于两波段的线性组合，难以捕捉更复杂的光谱交互行为。本研究在 FOD 处理基础上进一步引入 3D 光谱指数，将第三波段纳入指数计算，形成更高维度的组合特征，以更全面地反映光谱结构变化[27]。结果表明，在 0.2 阶 FOD 处理下构建的 TBI4 与 SOM 的相关系数较

同阶 2D 指数提升 2.38%～19.44%，证实了张子鹏等[28]在艾比湖流域提出的三维指数在增强弱吸收峰表达、缓解谱间共线性方面的理论优势，也显示 FOD 有助于提升光谱指数的构建质量。

模型对比结果进一步揭示了 RF 与 Cubist 算法对预处理策略和特征组合的响应差异。Cubist 本质上是基于规则集的分段线性回归模型，依赖于特征变量之间存在明确的分界与线性关系，适合处理特征维度较低的数据。在本研究中，低阶 FOD（如 0.3 阶）能有效增强谱线细节而不引入显著噪声，使 Cubist 模型更易拟合高频特征并获得最佳性能[29]。相比之下，RF 模型更适合处理维度较高、信息复杂的数据[30]。在中高阶 FOD 处理下（如 1.7 阶），RF-FOD 模型 R^2为 0.65，较原始光谱（0 阶）模型提升 80.56%。Index-RF 模型预测精度进一步提升，基于 1.6 阶光谱指数数据集构建的 RF 模型 R^2为 0.69，较 0 阶光谱指数数据集构建的 RF 模型提升 68.29%。在 21 个分数阶处理中，共有 15 个处理的 Index-RF 模型 R^2 优于同阶 FOD-RF 模型，表明基于光谱指数建模方法对噪声的抵抗力与对关键信息的提取能力进一步增强[31]。然而，Index-RF 模型不同分数阶微分处理的 R^2 差异较大，如在 0.6 阶（R^2 为 0.40）和 1.4 阶（R^2 为 0.49）处显著降低。Index-Cubist 模型亦存在类似问题。尽管如此，光谱指数建模所需变量较少，模型运行效率显著优于光谱数据建模方法。因此，选择合适的 FOD 处理阶数进行光谱指数计算并建模，可在保证精度的同时，大幅提升模型运算速度。模型算法在处理不同特征维度、噪声结构、非线性强度等方面表现出的差异，揭示了建模策略必须结合具体数据特性与目标变量结构进行优化选择。这一结论亦与马瑜蔓等[32]在 2023 年基于 FOD 和无人机高光谱数据预测油菜产量的研究相印证，表明 FOD 处理技术具备良好的通用性和适应能力，可拓展至其他土壤理化参数（如 pH、电导率、含水率等）及作物生长指标的建模预测[33]。

四种模型对应的相关波段均与土壤中的水分、氧化铁、烷基、羧酸、氨基、羟基、高岭石等成分的光谱吸收特性密切相关[34-36]。580～620 nm 区段涵盖了可见光谱中与氧化铁的特征吸收有关的波段，氧化铁作为土壤颜色的重要影响因素，其含量高低影响 SOM 反射光谱[37-38]。这与彭杰等[39]得出的土壤有机质反射光谱特征波段 570～620 nm 基本重叠。而 1380～1420 nm 则对应 O-H 的伸缩振动，反映土壤水分与部分黏土矿物成分的变化，水分与有机质含量密切相关，有机质含量高的土壤，吸附水含量高[40]。Index-RF 和 Index-Cubist 模型所对应的光谱指数重要波段与 FOD 处理数据集建模的重要波段高度一致，表明所构建的指数能够有效捕捉对土壤有机质变化敏感的关键光谱区域。光谱指数变量的重要性较高，说明光谱指数变量能够有效地增强土壤光谱的特异性。这一结果进一步表明，光谱指数可有效提取波段间的交互信息，削弱特征间的冗余性，并提高光谱变量的解释力与代表性。分数阶微分处理与光谱指数构建的协同作用，为提升模型对土壤有机质相关光谱特征的响应能力提供了可靠机制，从而提高了基于 Vis-NIR 光谱的土壤属性估算的精度与稳定性。

尽管本研究在多阶 FOD 与光谱指数框架下，系统比较了 RF 与 Cubist 两类模型，并取得了较好的预测精度，但仍存在可改进之处。首先，本研究采用的建模方法以机器学习回归算法为主，尚未引入基于记忆学习或深度学习（如卷积神经网络）等新兴框架。此类方法在捕获样本间复杂相似性与非线性交互方面具备潜力[41-42]，值得在后续研究中进一步探索与比较。其次，当前研究主要基于光谱数据，未来研究可将空间辅助因子（如土壤类型、地形坡度、植被覆盖等）纳入建模，以提高模型在时空上的可迁移性与泛化能力。

4 结 论

FOD 处理可有效增强土壤光谱的吸收特征，有效提升模型对 SOM 含量的预测能力。其中，Cubist 模型在以 0.3 阶 FOD 处理光谱构建的数据集上表现最优，$R^2 = 0.74$，RMSE = 6.47 g·kg^{-1}；RF 模型更适合 1.6～1.9 阶 FOD 处理的光谱数据集，其 R^2稳定在 0.63～0.65，RMSE 最低达 7.53 g·kg^{-1}；RF 模型采用最优光谱指数建模的整体性能优于 FOD 全光谱建模方法，而 Cubist 模型采用最优光谱指数的建模精度则低于 FOD 全光谱建模方法。Cubist 模型在低阶 FOD（0.3 阶）处理下对 560 nm 与 1370 nm 等波段的微弱非线性变化更为敏感，表现出更优的拟合效果；而 RF 模型则在中高阶 FOD（1.6 阶）处理下对 610 nm 波段的信息保持较强的稳定性。本研究

提出并验证了 FOD 结合光谱指数的协同建模策略，系统评估了 RF 与 Cubist 模型在不同阶数与特征构建方式下的响应机制，明确了各自的建模优势，可为黑土地土壤有机质快速预测与评价提供方法和技术参考。

参考文献（References）

[1] Zhao Y C，Wang M Y，Hu S J，et al. Economics- and policy-driven organic carbon input enhancement dominates soil organic carbon accumulation in Chinese croplands[J]. Proceedings of the National Academy of Sciences of the United States of America，2018，115（16）：4045—4050.

[2] Zhang J B，Sun B，Zhu J J，et al. Black soil protection and utilization based on harmonization of mountain-river-forest-farmland-lake-grassland-sandy land ecosystems and strategic construction of ecological barrier[J]. Bulletin of Chinese Academy of Sciences，2021，36（10）：1155—1164. [张佳宝，孙波，朱教君，等. 黑土地保护利用与山水林田湖草沙系统的协调及生态屏障建设战略[J]. 中国科学院院刊，2021，36（10）：1155—1164.]

[3] Wang J K，Xu X R，Pei J B，et al. Current situations of black soil quality and facing opportunities and challenges in Northeast China[J]. Chinese Journal of Soil Science，2021，52（3）：695—701. [汪景宽，徐香茹，裴久渤，等. 东北黑土地区耕地质量现状与面临的机遇和挑战[J]. 土壤通报，2021，52（3）：695—701.]

[4] Li R，Hu W Y，Jia Z J，et al. Soil degradation：A global threat to sustainable use of black soils[J]. Pedosphere，2025，35（1）：264—279.

[5] Jiang Y，Wang J，Yang J W. Research progress analysis of black soil region cultivated land quality evaluation index by remote sensing[J]. Engineering of Surveying and Mapping，2023，32（1）：1—7，13. [姜芸，王军，杨继文. 基于遥感技术的黑土区耕地质量评价指标体系研究进展分析[J]. 测绘工程，2023，32（1）：1—7，13.]

[6] Deng Y，Fan Y N，Wu Q M，et al. Rapid diagnosis and assessment of heavy metal pollution in greenhouse vegetable production soils based on portable X-ray fluorescence spectroscopy[J]. Acta Pedologica Sinica，2024，61（1）：174—186. [邓源，樊亚男，吴秋梅，等. 基于便携式 X 射线荧光光谱速测的设施菜地土壤重金属污染诊断与评价[J]. 土壤学报，2024，61（1）：174—186.]

[7] Zhang F，Hu W Y，Wang L L，et al. Quick acquisition of black soil pH values from Northeast China by portable soil multi-parameter rapid detection instrument[J]. Acta Pedologica Sinica，2024，61（6）：1568—1576. [张凤，胡文友，王亮亮，等. 便携式土壤多参数快速检测仪原位测量东北黑土 pH[J]. 土壤学报，2024，61（6）：1568—1576.]

[8] Soriano-Disla J M，Janik L，McLaughlin M J，et al. The use of diffuse reflectance mid-infrared spectroscopy for the prediction of the concentration of chemical elements estimated by X-ray fluorescence in agricultural and grazing European soils[J]. Applied Geochemistry，2013，29：135—143.

[9] Hong Y S，Liu Y L，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.

[10] Hong Y S，Chen S C，Liu Y L，et al. Combination of fractional order derivative and memory-based learning algorithm to improve the estimation accuracy of soil organic matter by visible and near-infrared spectroscopy[J]. Catena，2019，174：104—116.

[11] Li W Y，Mamat Sawut，Maihemuti Balati. Fractional differential-based hyperspectral inversion of soil organic matter content[J]. Laser & Optoelectronics Progress，2023，60（7）：404—411. [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.]

[12] Zhang X H，Meng X T，Tang H T，et al. Random forest prediction model for the soil organic matter with optimized spectral inputs[J]. Transactions of the Chinese Society of Agricultural Engineering，2023，39（2）：90—99. [张笑寒，孟祥添，唐海涛，等. 优化光谱输入量的土壤有机质随机森林预测模型[J]. 农业工程学报，2023，39（2）：90—99.]

[13] Zhu C M，Zhang Z P，Wang H W，et al. Assessing soil organic matter content in a coal mining area through spectral variables of different numbers of dimensions[J]. Sensors，2020，20（6）：1795.

[14] Zhang Z P，Ding J L，Wang J Z，et al. Prediction of soil organic matter in northwestern China using fractional-order derivative spectroscopy and modified normalized difference indices[J]. Catena，2020，185：104257.

[15] Zhang H L，Xie C Y，Tian P，et al. Measurement of soil organic matter and total nitrogen based on visible/near infrared spectroscopy and data-driven machine learning method[J]. Spectroscopy and Spectral Analysis，2023，43（7）：2226—2231. [章海亮，谢潮勇，田彭，等. 基于可见/近红外光谱和数据驱动的机器学习方法测量土壤有机质和总氮[J]. 光谱学与光谱分析，2023，43（7）：2226—2231.]

[16] Luo D F，Peng J，Feng C H，et al. Inversion of soil organic matter fraction in southern Xinjiang by visible-near-infrared and mid-infrared spectra[J]. Spectroscopy and Spectral Analysis，2021，41（10）：3069—3076. [罗德芳，彭杰，冯春晖，等. 可见光-近

红外、中红外光谱的土壤有机质组分反演[J]. 光谱学与光谱分析，2021，41（10）：3069—3076.]

[17] Deiss L，Margenot A J，Culman S W，et al. Tuning support vector machines regression models improves prediction accuracy of soil properties in MIR spectroscopy[J]. Geoderma，2020，365：114227.

[18] Kong X Q，Gao W C，Pan W J，et al. Prediction of soil carbon-to-nitrogen ratio based on visible-near infrared and mid-infrared spectroscopy[J]. Acta Pedologica Sinica，2025，62（3）：653—664. [孔祥麒，高维常，潘文杰，等. 土壤碳氮比的可见-近红外与中红外光谱预测[J]. 土壤学报，2025，62（3）：653—664.]

[19] Tan Y. Analysis on influence facters and model optimization of soil organic matter prediction accuracy based on Vis-NIR spectra[D]. Changchun：Jilin University，2022. [谭洋. 高光谱土壤有机质含量反演精度影响因素分析及模型优化[D]. 长春：吉林大学，2022.]

[20] Liu B Y，Zhang G L，Xie Y，et al. Delineating the black soil region and typical black soil region of northeastern China[J]. Chinese Science Bulletin，2021，66（1）：96—106. [刘宝元，张甘霖，谢云，等. 东北黑土区和东北典型黑土区的范围与划界[J]. 科学通报，2021，66（1）：96—106.]

[21] Dou S，Guo D. Soil type distribution and black soil land protection in Jilin Province[J]. Journal of Jilin Agricultural University，2018，40（4）：449—456. [窦森，郭聃. 吉林省土壤类型分布与黑土地保护[J]. 吉林农业大学学报，2018，40（4）：449—456.]

[22] Tan T H. Evaluation of soil degradation and productivity of agricultural land in the typical black soil region of Songnen[D]. Wuhan：Huazhong Agricultural University，2023. [澹腾辉. 松嫩典型黑土区农田土壤退化和生产力评价[D]. 武汉：华中农业大学，2023.]

[23] Lu R K. Analytical methods for soil and agro-chemistry（In Chinese）. Beijing：China Agricultural Science and Technology Press，2000. [鲁如坤. 土壤农业化学分析方法. 北京：中国农业科技出版社，2000.]

[24] Shang T H，Chen R H，Zhang J H，et al. Estimation of soil organic matter content in Yinchuan Plain based on fractional derivative combined with spectral indices[J]. Chinese Journal of Applied Ecology，2023，34（3）：717—725. [尚天浩，陈睿华，张俊华，等. 基于分数阶微分联合光谱指数估算银川平原土壤有机质含量[J]. 应用生态学报，2023，34（3）：717—725.]

[25] Geng J，Lü J W，Pei J，et al. Prediction of soil organic carbon in black soil based on a synergistic scheme from hyperspectral data：Combining fractional-order derivatives and three-dimensional spectral indices[J]. Computers and Electronics in Agriculture，2024，220：108905.

[26] Hong Y S，Guo L，Chen S C，et al. Exploring the potential of airborne hyperspectral image for estimating topsoil organic carbon：Effects of fractional-order derivative and optimal band combination algorithm[J]. Geoderma，2020，365：114228.

[27] Chen R H. Inversion study of soil salinization in Yinchuan Plain based on spectral index and elimination of moisture effects[D]. Yinchuan：Ningxia University，2023. [陈睿华.基于光谱指数和消除水分影响的银川平原土壤盐渍化反演研究[D]. 银川：宁夏大学，2023.]

[28] Zhang Z P，Ding J L，Wang J Z，et al. Quantitative estimation of soil organic matter content using three-dimensional spectral index：A case study of the ebinur lake basin in Xinjiang[J]. Spectroscopy and Spectral Analysis，2020，40（5）：1514—1522. [张子鹏，丁建丽，王敬哲，等. 利用三维光谱指数定量估算土壤有机质含量:以新疆艾比湖流域为例[J]. 光谱学与光谱分析，2020，40（5）：1514—1522.]

[29] Chen S C，Xue J，Shi Z. Spectral-guided ensemble modelling for soil spectroscopic prediction[J]. Geoderma，2023，437：116594.

[30] Wang S，Guan K Y，Zhang C H，et al. Using soil library hyperspectral reflectance and machine learning to predict soil organic carbon：Assessing potential of airborne and spaceborne optical soil sensing[J]. Remote Sensing of Environment，2022，271：112914.

[31] Zhao M S，Xie Y，Lu L M，et al. Modeling for soil organic matter content based on hyperspectral feature indices[J]. Acta Pedologica Sinica，2021，58(1)：42—54. [赵明松，谢毅，陆龙妹，等. 基于高光谱特征指数的土壤有机质含量建模[J]. 土壤学报，2021，58（1）：42—54.]

[32] Ma Y M，Duan B，Xu B C，et al. Rapeseed yield prediction based on fractional-order differentiation and UAV hyperspectral index optimization[J]. Transactions of the Chinese Society of Agricultural Engineering，2025，41（10）：166—175. [马瑜蔓，段博，徐宾灿，等. 基于分数阶微分和无人机高光谱指数优选的油菜产量预测[J]. 农业工程学报，2025，41（10）：166—175.]

[33] Hong Y S，Chen S C，Chen Y Y，et al. Comparing laboratory and airborne hyperspectral data for the estimation and mapping of topsoil organic carbon：Feature selection coupled with random forest[J]. Soil and Tillage Research，2020，199：104589.

[34] Liu H J，Zhang B，Zhao J，et al. Spectral models for prediction of organic matter in black soil[J]. Acta Pedologica Sinica，2007，44（1）：27—32. [刘焕军，张柏，赵军，等. 黑土有机质含量高光谱模型研究[J]. 土壤学报，2007，44（1）：27—32.]

[35] Yang L J，Yang S，Zhang S Y，et al. Hyperspectral characteristics and quantitative monitoring of soil organic carbon in farmland[J]. Acta Laser Biology Sinica，2024，33（4）：316—325. [杨林婧，杨莎，张圣杨，等. 农田

土壤有机碳高光谱特征及定量监测研究[J]. 激光生物学报，2024，33（4）：316—325.]

[36] Ji W J，Shi Z，Zhou Q，et al. VIS-NIR reflectance spectroscopy of the organic matter in several types of soils[J]. Journal of Infrared and Millimeter Waves，2012，31（3）：277—282. [纪文君，史舟，周清，等. 几种不同类型土壤的 VIS-NIR 光谱特性及有机质响应波段[J]. 红外与毫米波学报，2012，31（3）：277—282.]

[37] Zhao H L，Gan S，Wang J J，et al. Characteristic wavelength selection and hyperspectral inversion of soil iron oxide[J]. Chinese Journal of Eco-Agriculture，2022，30（8）：1336—1345. [赵海龙，甘淑，王俊杰，等. 土壤氧化铁的特征波长选择和高光谱反演[J]. 中国生态农业学报，2022，30（8）：1336—1345.]

[38] Lu Q，Wang S J，Bai X Y，et al. Rapid inversion of heavy metal concentration in karst grain producing areas based on hyperspectral bands associated with soil components[J]. Microchemical Journal，2019，148：404—411.

[39] Peng J，Zhou Q，Zhang Y Z，et al. Effect of soil organic matter on spectral characteristics of soil[J]. Acta Pedologica Sinica，2013，50（3）：517—524. [彭杰，周清，张杨珠，等. 有机质对土壤光谱特性的影响研究[J]. 土壤学报，2013，50（3）：517—524.]

[40] Shi Z，Wang Q L，Peng J，et al. Development of a national VNIR soil-spectral library for soil classification and prediction of organic matter concentrations[J]. Scientia Sinica：Terrae，2014，44（5）：978—988. [史舟，王乾龙，彭杰，等. 中国主要土壤高光谱反射特性分类与有机质光谱预测模型[J]. 中国科学：地球科学，2014，44（5）：978—988.]

[41] Wang Z，Chen S C，Lu R，et al. Non-linear memory-based learning for predicting soil properties using a regional vis-NIR spectral library[J]. Geoderma，2024，441：116752.

[42] Yang J C，Guo B. Estimating salinization in coastal areas using FOD-CNN spectral indices[J]. Transactions of the Chinese Society of Agricultural Engineering，2025，41（15）：76—85. [杨吉存，郭兵. 基于 FOD-CNN 光谱指数的滨海地区盐渍化估算模型[J]. 农业工程学报，2025，41（15）：76—85.]

（**责任编辑**：檀满枝）