

基于可见-近红外反射光谱的油区农田土壤 石油烃含量估测研究^①

任红艳^{1,2}, 史学正³, 庄大方^{2*}, 徐新良², 江 东², 于信芳²

(1 华东师范大学地理信息科学教育部重点实验室, 上海 200062; 2 资源与环境信息系统国家重点实验室(中国科学院地理科学与资源研究所), 北京 100101; 3 中国科学院南京土壤研究所, 南京 210008)

摘要: 利用可见-近红外光谱技术, 研究油区农田土壤在颗粒为 2 mm、0.25 mm 和 0.15 mm 时的光谱(原始、一阶微分、连续统去除、多次散射校正)反射率与石油烃含量变化的关系, 建立并利用偏最小二乘回归模型估测了石油烃含量。结果表明: 原始、连续统去除、多次散射校正光谱的反射率对石油烃含量变化的响应在 350 ~ 600 nm 最敏感($P < 0.05$), 一阶微分光谱的敏感波段为 2 280 nm ($|r| = 0.81, P < 0.01$); 光谱预处理能明显提高石油烃含量的估测精度, 但土壤研磨对估测精度的影响因光谱预处理方法而异。基于可见-近红外光谱技术, 本研究确立了估测油区农田土壤石油烃含量的颗粒粒径(2 mm)和光谱预处理方法(一阶微分), 实现了石油烃含量的快速、有效估测, 探索了一种可替代传统分析技术的新方法。

关键词: 可见-近红外光谱; 油区; 农田; 土壤; 石油烃含量

中图分类号: S151.9

全球石油总产量 22 亿 t 中的 80% 来自陆地油田^[1], 其中有相当一部分处于农业生产区^[2]。因此, 石油开采对使陆地生态系统尤其农田土壤-植物生态系统面临严峻的污染威胁^[3-6]。中国作为世界上最大的发展中国家及石油生产和消费大国, 由于生产条件、环保技术等方面相对落后, 我国每年有 6×10^5 t 石油经跑、冒、滴、漏等途径进入环境, 造成土壤污染^[1,3,7]。

快速、全面摸清土壤中石油类物质含量是有效、准确评价石油污染农田土壤质量的前提。石油是 C15 ~ C36 的烷烃、多环芳烃(PAHs)、烯烃、苯系物、酚类等数百种物质组成的混合物, 这些物质进入农田土壤后主要在表层土壤形成污染, 使土壤的通透性、水分含量、体积质量、氧含量等物理特性改变^[6,8-10]; 同时使土壤有效氮和速效磷含量降低从而改变了土壤的化学特性^[7,9]; 从而影响农田土壤质量。目前, 测定土壤石油烃含量最通用的方法是通过固相或者液相萃取之后, 利用特殊波长激发其荧光特性从而进行测定。

通过化学分析测定土壤石油烃含量的传统方法费时、费力, 首先需要野外采集大量样品, 进行石油

烃等物质含量的分析, 从而进行土壤石油污染评价; 在油田影响面积较大的区域, 采样、分析的成本高、效率低。随着光谱分析技术和信息科学的发展, 越来越多的研究人员将可见-近红外光谱分析技术应用到土壤有机质含量、土壤湿度、土壤盐分含量、土壤氮素水平等土壤属性的快速估测中^[11-14], 但这些土壤属性的光谱估测精度受到土壤样品准备(颗粒粒径)、样本数量以及光谱预处理方法的影响^[15-17]。

本研究拟以苏北油田真武油区油井周边土壤及其可见-近红外光谱为对象, 比较不同土壤颗粒粒径和光谱预处理下土壤光谱反射率对石油烃含量变化响应的差异, 建立基于全波段光谱反射率的石油烃含量偏最小二乘回归模型, 评价土壤研磨、光谱预处理提高石油烃含量估测精度的适用性, 以期石油烃含量的有效、快捷、准确测定寻找一种可替代传统方法的新技术。

1 材料与方法

1.1 研究区概况

苏北油田是我国南方较大的新生代油田, 真武油

基金项目: 地理信息科学教育部重点实验室开放研究基金项目(KLGIS2011A13)和国家自然科学基金项目(41001279)资助。

* 通讯作者 (dfzhuang@lreis.ac.cn)

作者简介: 任红艳(1979—), 男, 湖南岳阳人, 博士, 助理研究员, 主要研究方向为环境遥感应用。E-mail: renhy@lreis.ac.cn

田为其中较大的采油区,主要覆盖江都市北部的真武镇及其周边油田,是江苏省石油和天然气主要产区。江都市平均海拔 5 m 左右,气候属湿润北亚热带气候区,年平均气温 14.9℃,降水量 978.7 mm。真武镇土壤类型主要为江河冲积物上发育而来的潜育型水稻土,大田农作物主要有小麦、油菜、水稻。

1.2 样品采集与分析

在真武镇选择 10 口处于农田中心的油井。以油井为中心,按东、东南、东北、南、西南、西北、西和北面 8 个方向不同距离(5、10、15、20、25、30、35、40 和 50 m)随机布点。每点采集大约 200 g 的农田表层土壤(0~20 cm),共采集 146 个土壤样品。除去样品中的石块、植物根系和凋落物等在室内风干,研磨并分别过 2、0.25、0.15 mm 土壤筛。

在样品风干、过筛后,随机挑选出 40 个样品,委托中国科学院南京土壤研究所分析测试中心进行土壤(2 mm)石油烃(soil petroleum hydrocarbon, SPH)含量、土壤(0.15 mm)全氮(TN)含量、土壤(0.15 mm)有机质(SOM)含量分析,并在获取上述化学分析结果后,计算土壤碳氮比(C:N)

1.3 光谱测量与预处理

在室内太阳光照模拟条件下,使用 ASD_Fieldspec 测定 2、0.25、0.15 mm 土样的反射光谱,仪器波谱特征以及具体测量方法参见文献[14];利用多次散射校正(multiple scattering correction, MSC)、一阶微分(first derivative, FD)和连续统去除(continuum removal, CR)等 3 种方法对土壤光谱数据进行预处理。上述光谱预处理的具体实现过程参见相关文献[14, 18]。

1.4 多元回归模型

线性回归模型在可见-近红外光谱分析技术中,由于进入模型的敏感波段信息有限,在确立最大相关

波段为敏感波段的同时,其他与 SPH 含量变化密切相关的敏感波段信息亦可能被忽略、剔除。偏最小二乘回归(partial least square regression, PLSR)方法在可见-近红外光谱分析技术中应用十分广泛。因此,本研究考虑利用偏最小二乘回归分析方法,建立更稳定、更准确、充分利用全波段(350~2 500 nm)的光谱信息的 PLSR 模型。本研究中,由于样本数量有限(40 个),所以采用逐一法(leave-one-out)建立 SPH 含量的 PLSR 模型,并进行交叉验证(cross-validation)。根据模型的均方根误差(root mean square error of prediction, RMSEP)、决定系数(R^2)、因变量标准误差(standard deviation, SD)与 RMSEP 的比值(the ratio of SD of samples to RMSEP, RPD)的大小对模型的精度进行评价。上述参数的计算和精度分析方法详见文献[18]。

2 结果与分析

2.1 SPH 含量变化

表 1 给出了 40 个土样的 SPH 含量、SOM、TN 含量、C:N 以及它们之间的相关性分析结果。土样的 SPH 含量变化明显(变化范围 = 523.20 mg/kg), SPH 含量变化与样点到油井中心距离之间存在极显著的倒数关系(图 1)。在距离油井 5 m 范围内,SPH 含量较高,随着距离增加其含量急剧降低直至 10 m 范围,而 10~30 m 范围内,SPH 含量下降趋势放缓;30~50 m,其含量变化不再明显。与此同时,SOM 含量、C:N 与 SPH 含量存在极显著正相关。这表明 SPH 含量的增加,可能提高了 SOM 含量和土壤 C:N;而 TN 含量与 SPH 含量存在不显著的负相关,从 TN 含量随样点距离油井的远近变化(图 1)来看,靠近油井的地方(5 m 左右)TN 含量略低,随着距离的增加 TN 含量逐渐增加,大概从 20 m 开始,其含量趋于稳定。

表 1 土壤 SPH 实测含量与 SOM、TN 和 C:N

Table 1 Descriptive statistics of SPH concentration, SOM content, TN content and C:N

项目	变化范围	最小值	最大值	平均值	标准差	相关系数
SPH (mg/kg)	523.20	12.80	536.00	71.07	101.76	
SOM (g/kg)	30.94	10.35	41.29	22.3	6.13	0.47**
TN (g/kg)	1.54	0.53	2.07	1.36	0.35	-0.02
C:N	8.93	7.7	16.63	9.61	1.43	0.85**

注: **表示相关性达到 $P < 0.01$ 显著性水平(2-tailed)。

从分析结果来看,该油田开采区 50 m 范围内农田 SPH 含量均值明显低于 500 mg/kg 的临界水平^[7],属于清洁土壤,且在距离油井 5~50 m 范围内 SPH

含量随着距离增加而降低。尽管如此,油井周围 50 m 范围内 SPH 含量的增加显著提高了 SOM 含量和 C:N,且一定程度上降低了土壤氮素水平。

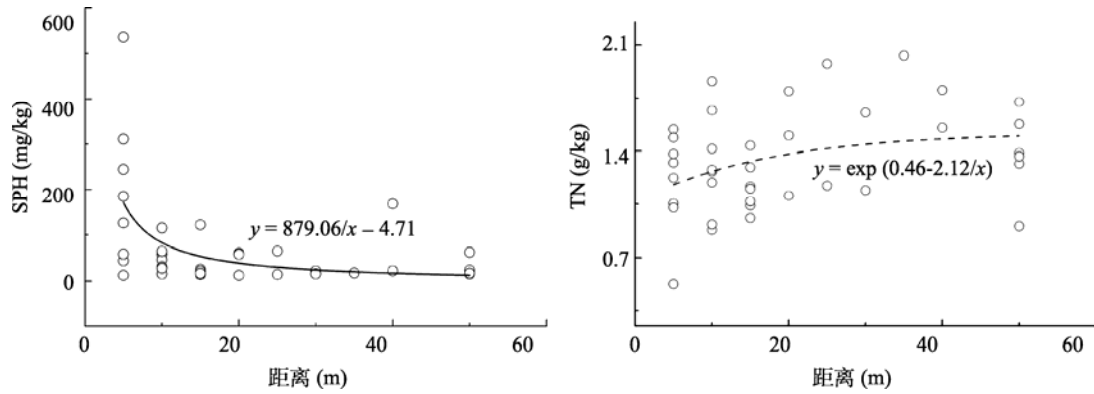


图 1 土壤 SPH 和 TN 含量与采样距离(与油井中心点)变化曲线图
Fig.1 Relation between SPH concentration, TN content and distance of sampled soils from oil well

2.2 土壤光谱反射对 SPH 含量变化的响应

2 mm、0.25 mm、0.15 mm 颗粒粒径土样原始光谱(RAW)反射率-SPH 含量之间的相关性变化趋势一致(图 2) 在 350 ~ 600 nm 区间表现显著(2 mm 土样)、极显著(0.25 mm、0.15 mm 土样)负相关,即土壤光谱反射率随 SPH 含量增加而降低;在 600 ~ 2 500 nm 区间,土壤光谱反射率-SPH 含量相关性随波长增加而逐步减弱;同时,土壤光谱反射率-SPH 含量相关性随着颗粒粒径的增加而降低,即 $|r_{0.15}| > |r_{0.25}| > |r_2|$,这表明土壤研磨可以提高 350 ~ 600 nm 区间土壤光谱反射率对 SPH 含量变化的响应。

与 3 个粒径水平土样的土壤 RAW 光谱反射率-SPH 含量相关型曲线相比,预处理(CR、FD、MSC)后光谱反射率-SPH 含量相关性曲线(图 2)差异变小。其中,在 RAW 光谱反射-SPH 含量显著相关的 350 ~ 600 nm 区间,CR 和 MSC 预处理后 3 个粒径水平土样的光谱反射率-SPH 含量相关性曲线区域重合;此外,与较为平缓的 RAW 光谱反射-SPH 含量相关性曲线相比,经过 CR、FD 以及 MSC 等方法预处理的相关性曲线出现了明显的峰和谷,而且曲线的振幅(最大正负相关系数绝对值之和)均大于原始光谱。由此可知,FD、CR、MSC 等预处理可以显著提

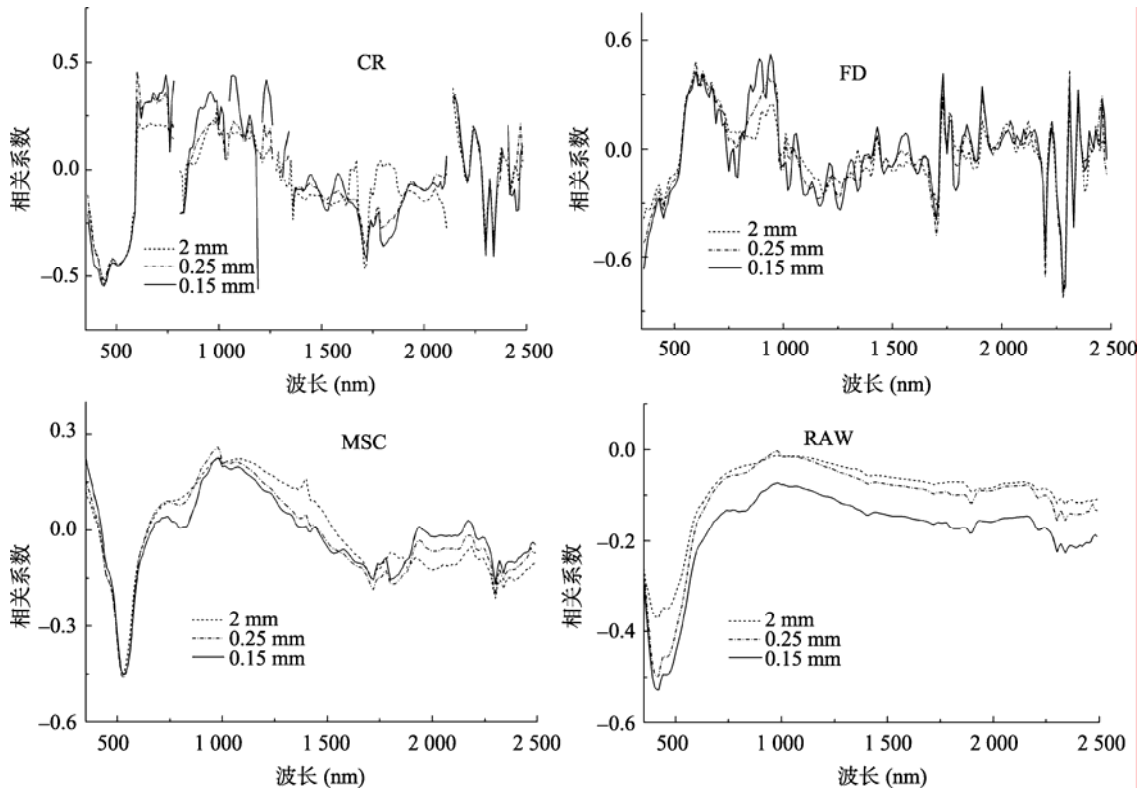


图 2 SPH 含量与土壤光谱反射率的相关性
Fig. 2 Correlation between SPH concentrations and reflectance spectra of soils at three particle sizes

高光谱反射对 SPH 含量变化的响应,同时可以降低土样颗粒粒径变化导致的光谱反射-SPH 含量相关性差异。

从 2 mm、0.25 mm、0.15 mm 土样的光谱反射率-SPH 含量最大相关系数及其对应的波段分析结果(表 2)可知,对于 RAW 和 CR 光谱而言,反射率-SPH 含量的最大相关对应的土壤颗粒粒径均为 0.15 mm,对应的波段分别为 420 nm 和 1 190 nm;对于 FD 和 MSC 光谱而言,最大相关对应的土壤颗粒粒径均为 0.25 mm,对应的波段分别为 2 280 nm 和 530 nm。0.25 mm 土样的 FD 光谱 2 280 nm 处反射率-SPH 含量相关系数最大($|r| = 0.82, P < 0.01$),与 2 mm 土样 FD 光谱 2 280 nm 处的相关性($|r| = 0.81, P < 0.01$)相比,相关系数并未显著提高(仅为 1.23%),而 2 mm 土样比 0.25 mm 土样要便于获取。

表 2 不同颗粒大小土壤反射光谱与 SPH 含量(mg/kg)最大相关波段

Table 2 Wavelengths with maximal correlation between SPH concentration and reflectance of soils at various particle sizes

颗粒粒径 (mm)	RAW		CR		FD		MSC	
	$ r $	波长(nm)	$ r $	波长(nm)	$ r $	波长(nm)	$ r $	波长(nm)
2	0.37	420	0.54	440	0.81	2 280	0.45	530
0.25	0.50	410	0.52	440	0.82	2 280	0.46	530
0.15	0.53	420	0.56	1 190	0.78	2 280	0.45	530

表 3 SPH 含量的 PLSR 估测模型精度(n = 40)

Table 3 Statistical parameters of PLS regression models for predicting SPH concentration

项目	RAW			CR			FD			MSC		
	2	0.25	0.15	2	0.25	0.15	2	0.25	0.15	2	0.25	0.15
Factors	8	6	6	2	2	2	5	3	3	7	5	6
R^2	0.58	0.59	0.51	0.56	0.61	0.6	0.65	0.61	0.59	0.53	0.54	0.6
RMSEP	66.64	65.79	72.07	68.48	64.76	65.11	60.58	64.74	65.71	70.8	69.76	65.2
RPD	1.53	1.55	1.41	1.49	1.57	1.56	1.68	1.57	1.55	1.44	1.46	1.56
Slope	0.62	0.75	0.48	0.43	0.53	0.49	0.70	0.59	0.54	0.55	0.81	0.66

此外,就表 3 中 2 mm、0.25 mm、0.15 mm 土样 SPH 含量的估测精度而言, MSC-PLSR 模型的表现(R^2 均值 = 0.56, RPD 均值 = 1.49)与 RAW-PLSR 模型(R^2 均值 = 0.56, RPD 均值 = 1.48)相当,而 CR-PLSR 模型的表现(R^2 均值 = 0.59, RPD 均值 = 1.54)要优于 RAW-PLSR 模型,而 FD-PLSR 模型的表现更佳(R^2 均值 = 0.62, RPD 均值 = 1.60)。这表明对土壤光谱进行 FD 预处理能显著提高全波段光谱 PLSR 模型估测 SPH 含量的精度。

因此,对于 SPH 含量的全波段光谱 PLSR 估测,土壤颗粒大小对估测精度的影响因光谱预处理方法而异;若采取 FD 方法对土样光谱进行预处理,在获得 2 mm 土样后,继续研磨并不能提高估测精度。图 3

2.3 SPH 含量的 PLSR 估测模型

为了充分利用全波段(350 ~ 2500 nm)的光谱信息,分别构建基于 2 mm、0.25 mm、0.15 mm 土样光谱(经过多种光谱预处理)反射率的 SPH 含量 PLSR 估测模型,所有模型的因子数(Factor)、决定系数(R^2)、估测均方根误(RMSEP)、斜率(Slope)以及标准误差-均方根误比值(RPD)等参数见表 3。从该表可知,随着土样研磨程度增加(颗粒粒径越小),CR 光谱和 MSC 光谱的 PLSR 模型估测 SPH 含量的精度逐渐增加(RPD 和 R^2 增大、RMSEP 减小),而 RAW 光谱和 FD 光谱 PLSR 模型估测 SPH 含量的精度随之降低(RPD 和 R^2 减小、RMSEP 增加)。这说明研磨可以提高 CR-PLSR 和 MSC-PLSR 模型估测 SPH 含量精度,但不能改善 RAW-PLSR 和 FD-PLSR 模型的估测精度。

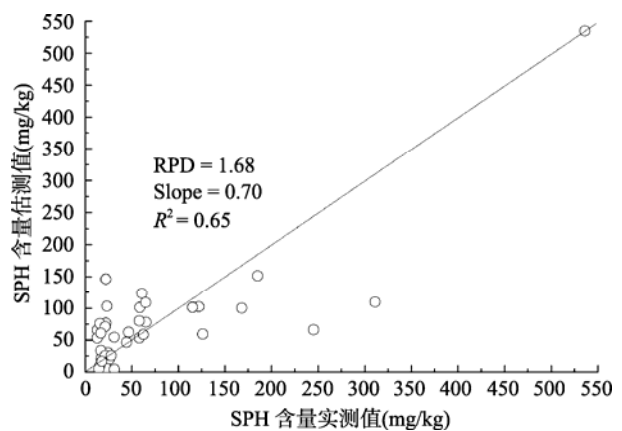


图 3 SPH 含量估测模型的估测-实测值对比图

Fig. 3 Predicted v.s measured values of SPH concentration yielded by predicting models

是 FD-PLSR 模型估测 SPH 含量的估测值-实测值对比图, 其中落在 1:1 线上及其附近区域的点占有样本数量的 1/2 左右——尽管这种估测精度不尽如人意, 但该模型为本研究中的最佳估测模型。

3 结语

通过对位于真武油田 10 口油井周边 50 m 范围内土壤采样、分析, 发现距离油井 5 ~ 50 m 范围内土壤石油烃含量随着距离增加而降低。该油田区农田土壤的石油污染浓度低于 500 mg/kg 的临界值, 处于清洁水平, 这可能与该油田处于重要农业生产区, 对石油生产采取了较好的环保措施有关; 尽管如此, 进入土壤的石油类污染物明显提高了土壤有机质含量、碳氮比, 并在一定程度上降低了土壤氮素水平。

利用可见-近红外光谱技术, 本研究分析了 2 mm、0.25 mm、0.15 mm 土样不同预处理后 350 ~ 2 500 nm 区间光谱反射率与土壤石油烃含量变化的关系, 结果表明: 对于土样 RAW、CR 及 MSC 光谱, 其反射率对 SPH 含量变化的响应在 350 ~ 600 nm 最为灵敏 ($P < 0.05$), 而 FD 光谱的敏感波段则在 2 280 nm ($|r| = 0.81, P < 0.01$); FD、CR、MSC 等光谱预处理方法和土壤研磨均可以明显提高土壤光谱反射对土壤石油烃含量变化的响应, 而 CR、MSC 预处理能有效减弱土壤颗粒粒径变化导致的光谱反射率-石油烃含量相关性差异。此外, 基于全波段(350 ~ 2 500 nm)光谱反射率, 利用偏最小二乘回归方法建立了土壤石油烃含量的估测模型。CR 和 FD 预处理可以明显提高土壤石油烃含量的估测精度, 而土壤研磨对土壤石油烃含量估测精度的影响因光谱预处理方法而异; 若采取 FD 方法对土样光谱进行预处理, 对 2 mm 土样继续研磨并不能提高全波段光谱 PLSR 模型的估测精度, 即基于 2 mm 土样 FD 光谱构建的偏最小二乘回归模型可以实现土壤石油烃含量的有效估测, 这与土壤属性光谱估测相关研究中最适土壤颗粒粒径的结论一致^[19-20]。

土壤石油烃含量的测定是评价采油区农田土壤质量的前提。基于可见-近红外光谱分析技术, 本研究确立了估测土壤石油烃含量的土壤颗粒粒径(2 mm)、土壤光谱预处理方法(FD)及偏最小二乘回归模型, 实现了土壤石油烃含量的快速、有效估测。尽管其估测精度尚需通过提高采样数量、增加模型因变量变化范围等措施进一步提高, 但本研究探索的可见-近红外光谱分析技术不失为一种可替代传统分析方法的有效选择。

参考文献:

- [1] 任磊, 黄廷林. 土壤的石油污染[J]. 农业环境保护, 2000, 19(6): 360-363
- [2] 刘五星, 骆永明, 滕应, 李振高, 吴龙华. 我国部分油田土壤及油泥的石油污染初步研究[J]. 土壤, 2007, 39(2): 247-251
- [3] 焦杏春, 陶澍, 卢晓霞, 刘煜, 沈伟然. 水稻根系中多环芳烃的动态变化[J]. 环境科学学报, 2007, 27(7): 1 203-1 208
- [4] Edwards NT. Polycyclic aromatic hydrocarbons (PAH's) in the terrestrial environment: A review[J]. Journal of Environmental Quality, 1983, 12(4): 427-441
- [5] Meudec A, Dussauze J, Deslandes E, Poupart N. Evidence for bioaccumulation of PAHs within internal shoot tissues by a halophytic plant artificially exposed to petroleum-polluted sediments[J]. Chemosphere, 2006, 65: 474-481
- [6] 张亚非. 油田开发对生态环境的影响[J]. 油气田环境保护, 2006, 12: 28-32
- [7] 刘五星, 骆永明, 滕应, 李振高, 吴龙华. 石油污染土壤的生态风险评价和生物修复 II. 石油污染土壤的理化性质和微生物生态变化研究[J]. 土壤学报, 2007, 44(5): 848-853
- [8] Koa JY, Daya JW. A review of ecological impacts of oil and gas development on coastal ecosystems in the Mississippi Delta[J]. Ocean & Coastal Management, 2004, 47: 597-623
- [9] Osuji LC, Onojake CM. Field reconnaissance and estimation of petroleum hydrocarbon and heavy metal contents of soils affected by the Ebocha-8 oil spillage in Niger Delta, Nigeria[J]. Journal of Environmental Management, 2006, 79: 133-139
- [10] 李明. 钻井废水对土壤及主要粮食作物的影响研究[J]. 天然气工业, 1996, (9): 72-77
- [11] Ben-Dor E, Banin A. Near - infrared analysis as a rapid method to simultaneously evaluate several soil properties[J]. Soil Science Society of America Journal, 1995, 59: 364-372
- [12] Ben-Dor E. Quantitative remote sensing of soil properties[J]. Advances in Agronomy, 2000, 75: 173-243
- [13] 王淼, 解宪丽, 周睿, 王宝良, 王昌昆, 刘娅, 潘剑君, 沈润平, 潘贤章. 基于可见光-近红外漫反射光谱的红壤有机质预测及其最优波段选择[J]. 土壤学报, 2011, 48(5): 1 083-1 089
- [14] 任红艳, 庄大方, 邱冬生, 潘剑君. 矿区农田土壤砷污染的可见-近红外光谱分析研究[J]. 光谱学与光谱分析, 2009, 29(1): 114-118
- [15] 何挺. 土地质量高光谱遥感监测方法研究(博士学位论文)[D]. 北京: 中国科学院遥感应用研究所, 2003

- [16] Jacquemoud S, Baret F, Hanocq JF. Modelling spectral and bidirectional soil reflectance[J]. *Remote Sensing of Environment*, 1992, 41: 123–132
- [17] Cierneiewski J, Karnieli A. Virtual surfaces simulating the bidirectional reflectance of semiarid soils[J]. *International Journal of Remote Sensing*, 2003, 24(7): 1 469–1 486
- [18] Ren HY, Zhuang DF, Singh AN, Pan JJ, Qiu DS, Shi RH. Estimation of As and Cu contamination in agricultural soils around a mining area by reflectance spectroscopy: A case study[J]. *Pedosphere*, 2009, 19(6): 719–726
- [19] Cañasveras JC, Barrón V, Campillo MC, Torrent J, Gómez JA. Estimation of aggregate stability indices in Mediterranean soils by diffuse reflectance spectroscopy[J]. *Geoderma*, doi:10.1016/j.geoderma. 2009.09.004
- [20] Nduwamungu C, Ziadi N, Tremblay G, Parent L. Near-infrared reflectance spectroscopy prediction of soil properties: Effects of sample cups and preparation[J]. *Soil Science Society American Journal*, 2009, 73(6): 1 896–1 903

Visible-near-infrared Spectroscopy in Estimation of Petroleum Hydrocarbon Concentration in Soil

REN Hong-yan^{1,2}, SHI Xue-zheng³, ZHUANG Da-fang^{2*}, XU Xin-liang², JIANG Dong², YU Xin-fang²

(1 *Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai 200062, China*; 2 *State Key Laboratory of Resource and Environmental Information System (Institute of Geographic Science and Resource Research, Chinese Academy of Sciences), Beijing 100101, China*; 3 *Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China*)

Abstract: Visible-near-infrared spectroscopic characteristics of soils sampled in cropland in oil well region were analyzed and these soil samples were meshed by 2 mm, 0.25 mm, 0.15 mm. Reflectance spectra was pretreated by first derivative, continuum removal, and multiple scattering correction before study on relationship between spectral reflectance and soil petroleum hydrocarbon concentration. Partial least square regression models for predicting soil petroleum hydrocarbon concentration were built on the basis of the full wavebands from 350 to 2 500 nm. The results showed that the most sensitive wavebands of RAW, continuum-removal, and multiple-scattering-correction spectra were located in the region from 350 to 600 nm with 0.05-level significance, and that those of first-derivative spectra were located at 2 280 nm ($|r| = 0.81$, $P < 0.01$). Estimation accuracy of partial least square regression models for predicting soil petroleum hydrocarbon concentration was increased by spectral pretreatment; nevertheless, effects of soil milling on estimation accuracy depended on spectral pretreatment. In this study, 2 mm and first derivative were respectively considered as suitable soil particle size and spectral pretreatment in building models for predicting soil petroleum hydrocarbon concentration on the basis of visible-near-infrared spectroscopy and spectral analysis, which provides a novel alternative of traditional analysis methods.

Key words: Visible-near-infrared spectra, Petroleum exploiting region, Cropland, Soil, Petroleum hydrocarbon concentration