基于 MODIS 产品的区域土壤遥感分类研究
——以新疆为例

亢 庆[1],张增祥[2],赵晓丽[2]
(1. 珠江水利科学研究院,广东 广州 510611;2. 中国科学院遥感应用研究所,北京 100101)

摘要:以新疆维吾尔自治区全境为研究区域,采用中高分辨率 MODIS 遥感数据和地形数据,在第二次全国土壤普查数据库的支持下,采用自动分类方法,探讨了遥感技术在常规土壤调查工作受限制的干旱地区进行土壤调查的效果和适用性。研究中使用了 MODIS 地表反射率、植被指数、地表昼夜温度等数据产品,提取了多种图像特征,并结合了 DEM 生成的地形参数。研究区土壤分类系统在发生学分类的基础上集合遥感信息特征进行了调整,形成了具有 26 个土壤类型及特殊地表覆被的土壤遥感分类系统。经分类试验,总体精度为 70%左右。

关 键 词:遥感;MODIS;土壤分类;干旱区;土壤地形数据库
中图分类号:TP 79 **文献标识码**:A **文章编号**:1004-0323(2007)06-0690-06

1 引 言

随着 1986 年国际土壤学会 SOTER 项目的启动,采用遥感和 GIS 等新技术进行区域乃至全球数字土壤地形数据库的研究得到了广泛关注并取得了较大进展。尤其针对常规土壤调查方法受到限制或耗资耗时巨大的地区,基于遥感的土壤调查方法具有不可替代的优势。

目前,针对集成卫星数据和地形数据进行土壤资源调查与制图的尝试性研究已有不少,比较典型的有:Vettorazzi[1] 等对 AVHRR 数据土壤模式识别的潜力进行的评估研究;Odeh[2,3] 等对 AVHRR 在区域土壤分类和土壤性质的研究;Dobos[4,5] 等(2000,2001)基于 AVHRR 和 DEM 数据进行的土壤分类研究,其中 Dobos 的应用研究较深入。在其研究中集成了多期的 AVHRR 反射率和 NDVI 数据以及地形参数,分别在匈牙利全国(93 000 km^2)和某地区(9 200 km^2)尺度上进行了分辨率为 500 m 和 1000 m 的土壤分类试验,结果针对 10 种土壤类型的分类精度分别达到了 80%~90%(500m 分辨率)和 67%~93%。

这些研究对 NOAA 卫星 AVHRR 数据和地形数据进行土壤分类进行了探索,证明了在该尺度下粗分辨率卫星数据的适用性。而由美国 1999 年发射 MODIS 在空间和光谱分辨率上相比 AVHRR 更具有优势,其提供的丰富的数据产品在土地资源调查方面具有相当大的应用潜力。但在土壤分类方面的适用性和应用效果尚缺乏研究。本文以新疆地区为研究区,集成了 MODIS 数据产品和地形数据,在分析土壤发生学分类系统基础上的土壤—景观遥感信息特征的基础上,采用自动分类方法对研究区主要土壤类型进行了分类实验,以探讨 MODIS 数据在该区域尺度下进行土壤资源调查的应用潜力和前景。

2 研究区环境背景

将我国新疆维吾尔自治区全境作为研究区,新疆维吾尔自治区位于欧亚大陆中部,地处 73°40′~96°23′E,34°25′~49°10′N 之间,面积 166 万 km^2。四周高山环绕,远离海洋,自然地理环境封闭。境内北部为阿尔泰山,南部为昆仑山,天山横亘中部,与北部的准噶尔盆地和南部的塔里木盆地形成“三山夹两盆”的地貌格局。在辽阔的疆域内,高山占 27.5%沙漠占 22.4%,低山丘陵占 21.8%,平原戈壁占 24.0%,而平原天然绿洲仅占 4.0%,且以串珠

收稿日期:2007-03-02;**修订日期**:2007-09-12
基金项目:中国科学院知识创新工程重要方向项目(KZCX3-SW-334)。
作者简介:亢庆(1973-),男,理学博士,主要从事环境遥感应用研究。

状镶嵌在 2 万多 km^2 的交通线上，构成了经济活动空间分散的基本格局。横贯中部的天山使南北疆形成在气候、水文、土壤、生物资源方面分异的两大区域。新疆境内绵连的雪岭，林立的冰峰，形成了独具特色的大冰川，共计 1.86 万余条，总面积 2.4 万多 km^2。沙漠占全国沙漠面积的 2/3，其中塔克拉玛干沙漠的面积为 36.5 万 km^2，是我国最大的沙漠。准噶尔盆地的古尔班通古特沙漠，面积 5.113 万 km^2，为我国第二大沙漠（吴正等，2003）。新疆地区属温带大陆性气候。冬季长、严寒，夏季短、炎热，春秋季节变化剧烈。年平均气温南疆为 10 ℃，气温日均差可达 12～15 ℃，最大可达 20～30 ℃。新疆夏季相对湿度、冬季绝对湿度都不大，形成夏季干热，冬季干冷的特点。全区多年平均降水量为 145 mm，只有全国平均年降水量 630 mm 的 23%，而蒸发量约 200 mm。新疆日照丰富，太阳辐射总量全年为 542.10～646.35 J/cm^2，仅次于青藏高原。新疆风多风大，并呈现北疆大于南疆，戈壁大于山区，盆地边缘大于盆地腹地的特征[6]。

总的来说新疆自然生态环境具有以下特点：地域辽阔但绿洲面积不大；气候干旱气温变化大；水资源总量丰富，但时空分布不均；土地面积大，但林地面积小；土壤质量差且沙化，盐碱化在全国占有重要地位。

3　数据选择

3.1　遥感数据

数据选择在考虑研究区内地物反射差异的同时，考虑了植被类型生育期存在的季节性差异，因此选用了 2004 年获取的 MODIS 反射率 MOD09A1、植被指数 MOD13A1 和地表温度数据 MOD11A2。具体数据特征见表 1。

表 1　选用的 MODIS 数据产品特点

Table 1　Characteristics of MODIS data

产品名称	MOD09A1	MOD13A1	MOD11A2
所选数据	1-7 波段反射率	NDVI 与 EVI	昼、夜地表温度
数据描述	8 d 一次合成，经过大气校正、薄云去除及汽溶胶订正	16 d 一次合成，采用限定视角下的最大值法 CV-MVC 和双向反射率模型 BRDF 进行了去云处理	经过反演的 8 d 一次合成的平均昼夜地表温度
分辨率(m)	500	500	1000
选用日期（2004 年）	7-27 至 8-3	3-5 至 3-20；4-6 至 4-21；5-8 至 5-23；6-9 至 6-24；7-11 至 7-26；8-12 至 8-27；9-13 至 9-28；10-15 至 10-31	3-5 至 3-12；4-6 至 4-13；5-8 至 5-15；6-9 至 6-16；7-11 至 7-19；8-12 至 8-19；9-13 至 9-20；10-15 至 10-22

MOD09A1 产品属 MODIS 的 3 级产品，由每日获取的 2 级反射率产品经合成得到。2 级反射率产品均经过了气体吸收、汽溶胶、卷云效应、临近效应等大气校正处理。

MOD13A1 是 16 d 合成空间分辨率为 500 m 的植被指数产品。合成算法采用的最大值合成方法（CV-MVC）。即逐日计算植被指数，用植被指数合成 MVC（最大值合成）方法选择最佳像元。若只有 1 d 无云，该天资料自动被选中。

MOD11A2 是 8 d 合成空间分辨率为 1000 m 的陆地表面温度产品，包含昼、夜地表温度。地表温度通过建立 31、32 通道亮温线性组合的劈窗算法计算获取，对已知发射率的像素点陆地表面温度的精度为 1 k[7]。

3.2　辅助数据

为了对遥感图像监督分类提供训练数据和验证数据，研究选用 1∶100 万全国土壤数据库和土壤样本数据库及 1∶25 万比例尺的 DEM 数据。

中国 1∶100 万土壤数据库的空间数据源是《1∶100万中华人民共和国土壤图》，它是全国第二次土壤普查的重要成果之一。第二次全国土壤普查依据全国统一的调查技术规程和土壤分类系统，从县和乡的土壤详查做起，在土壤普查过程中分别采集土壤样品和编制大比例尺土壤详查图，然后按地级市、省级和全国逐级汇总成图。该工作始于 1979 年，1994 年结束，历时 16 a。这是我国在扎实的工作基础与丰富的数据资料基础上所编成的第一套 1∶100万土壤图，也是到目前为止最为详细的全国性土壤图[8～10]。

该数字化土壤图基本制图单元为亚类，共有 12 个土纲，61 个土类，227 个亚类，以及 10 个非土壤单元。其中研究区内主要涉及约 31 个土类，89 个亚类，7 个非土壤单元。

4 工作流程与研究方法

本研究的工作流程见图 1。其中涉及到的数据处理过程和采用的方法将在以下内容中详细叙述。

图 1 工作流程图

Fig. 1 Framework of processing

本研究的自动分类方法采用了最大似然分类法(MLC, Maximum Likelihood Classifier)。最大似然分类法是最常用、最重要的一种监督分类算法,几乎被所有的图像处理系统和软件所采用。最大似然分类的基本思想是:将卫星遥感多波段数据的分布当作多维正态分布来构造判别分类函数。各类已知像元的数据在平面或空间中构成一定的点群;每一类的每一维数据都在自己的数轴上形成一个正态分布,该类的多维数据就构成该类的一个多维正态分布;各类的多维正态分布模型各有其分布特征。根据已知各类的已知数据(通过实地调查或其它方法确定地物类别及其训练区获得),可以构造出各类的多维正态分布模型,实际为各类出现各种数据向量的概率(即概率密度函数或概率分布函数);有了各类的多维分布模型,对于任何一个未知类别的数据向量,都可以反过来求它属于各类的概率;比较这些概率的大小,将该数据向量或像元归于概率大的类别[11]。

5 数据处理与数据集生成

5.1 数据预处理

使用 MODIS 产品处理工具 MRT 将 MODIS 各数据产品进行镶嵌、几何校正和重采样,投影体系为双标准纬线等积圆锥投影(ALBERS),椭球体为 Krasovsky,分辨率统一到 500 m。

5.2 分类特征提取

从 MODIS 数据中选取多波段反射率、植被指数、地表温度数据,并利用多光谱反射率进行湿度指数、亮度指数、纹理特征等分类特征的提取,作为初始分类特征。

5.2.1 地表反射率

MODIS 提供了常见的对地观测的 7 个光学反射率波段(与 TM 波段设置相似,见表 2),而且由于谱段更窄,大气效应减弱,数据质量更高。为提高分类精度,全部 7 个反射率波段均参与分类试验。(以下将 MODIS 的反射率波段表示为 Ref)。

5.2.2 植被指数

植被指数是对地表植被活动的简单、有效的度量,将两个(或多个)光谱通道组合可得到植被指数。

表 2 MODIS 与 TM 波段比较

Table 2 Comparison between MODIS and corresponding TM channels

MODIS 第 1～7 波段(μm)	Ref1 0.62～0.67	Ref2 0.841～0.876	Ref3 0.459～0.479	Ref4 0.545～0.565	Ref5 1.230～1.250	Ref6 1.628～1.652	Ref7 2.105～2.135
TM 波段	TM4 红	TM3 近红外	TM1 蓝	TM2 绿	—	TM5 短红	TM7 短红

植被指数一定程度上反映了不同土壤类型生物特性的差异,例如:生物量、覆盖度等。NDVI 指数是应用最广的植被指数,可用 MODIS 波段表达为:

$$\mathrm{NDVI} = (\mathrm{Ref2} - \mathrm{Ref1})/(\mathrm{Ref2} + \mathrm{Ref1}) \quad (1)$$

为了克服 NDVI 高植被盖度区易饱和、低植被区易受土壤背景影响的特点,MODIS 提供了一种基于本身数据特点的新的植被指数:"增强性植被指数 EVI(Enhanced Vegetation Index)",可用 MODIS 数据表达为:

$$\mathrm{EVI} = 2.5(\mathrm{Ref2} - \mathrm{Ref1})/(\mathrm{Ref2} + C1\mathrm{Ref1} + C2\mathrm{Ref3} + L) \quad (2)$$

其中:$L=1$ 为土壤调节参数,参数 $C1$ 和 $C2$ 分别为 6.0 和 7.5,通过蓝波段(Ref3)来修正大气对红波段(Ref1)的影响[7]。

5.2.3 土壤亮度特征

选用与 TM 相对应的 MODIS 波段反射率数据

(Ref1、Ref2、Ref3、Ref4、Ref6、Ref7 波段)进行 K-T 变换，选择包含土壤信息的亮度分量作为土壤亮度特征。

5.2.4　湿度指数(NDMI)与水体指数(NDWI)

湿度指数 NDMI 可反映植被水分和土壤湿度的信息[12]。该指数基于短波红外波段(相当于 MODIS 的 Ref6 波段)受水吸收带的影响，对湿度、含水量信息非常敏感，且绿波段(相当于 MODIS 的 Ref4 波段)对水体反射较强的特点，选用 MODIS 的这两个波段经标准化处理构建而成，表达式为：

$$NDMI = (Ref4 - Ref6)/(Ref4 + Ref6) \quad (3)$$

另外，为区分湖泊、沼泽等水体，选用了水体指数 NDWI，表达式为：

$$NDMI = (Ref4 - Ref2)/(Ref4 + Ref2) \quad (4)$$

5.2.5　归一化积雪指数(NDSI)

为了提取积雪信息，使用了基于反射特性的归一化差分积雪指数(NDSI)，其算法具有普遍的实际操作性意义，精度高，分类合理，是提取积雪信息的有效技术手段[7,13,14]。归一化差分积雪指数(NDSI)是基于雪对可见光与短波红外波段的反射特性和反射差的相对大小的一种测量方法。NDSI 类似于归一化植被指数(NDVI)，对大范围的光照条件不敏感，对大气作用可使其局地归一化并且不依赖于单通道的反射，其基本运算如下：

$$NDSI = (Ref1 - Ref4)/(Ref1 + Ref4) \quad (5)$$

5.2.6　昼夜地表温度

昼夜地表温度作为一个重要的环境变量，是区分土壤类型的重要特征因子，也是 MODIS 的一个重要产品。

5.2.7　纹理特征

纹理特征使用了表达可视纹理的局部平稳(Homogeneity)特征和基于信息理论的特征——熵(Entropy)参与分类。通过对 7 个波段进行可视化对比与分析后，选用 Ref2 波段进行纹理特征提取。

5.2.8　地形参数

鉴于地形与土壤类型存在着密切的关系，试验中采用 DEM 数据进行了多个地形特征参数的提取，包括：高程、坡度、坡向、表面曲率等。

5.3　分类数据集生成

将以上根据 MODIS 数据和地形数据提取的各种分类特征进行配准和叠合、裁切，最终形成具有多层数据的分类数据集。

此外，将分类数据集与整理后的土壤矢量数据进行叠合，便于下一步分析。

表 3　训练数据采集方案

Table 3　Scheme of collecting training data

样本编号	土　类	训练样本数	验证样本数
1	草甸土＋潮土	350	135
2	草毡土＋冷钙土	700	232
3	风沙土	400	140
4	栗钙土	200	74
5	灌漠土	100	28
6	灌淤土	150	45
7	龟裂土	100	34
8	寒冻土	300	87
9	寒钙土	300	96
10	寒漠土	200	68
11	黑毡土＋黑钙土	600	186
12	灰钙土	150	46
13	灰褐土	250	85
14	灰漠土	250	85
15	灰色森林土	200	66
16	灰棕漠土	500	154
17	石质土	450	138
18	盐土	400	128
19	沼泽土	80	26
20	棕漠土	620	212
21	棕钙土	488	158
22	棕色针叶林土	70	15
23	湖泊、水库	150	48
24	冰川、雪被	350	108
25	漠境盐土	160	55
26	西北盐壳	150	48

6　分类系统分析与训练数据采集

鉴于研究区涉及的土壤亚类较多(89 个)，可能造成分类过于复杂。因此，本研究的土壤分类系统以土类为分类基本单元(31 个)。对研究区内各土类的图斑面积进行统计，舍弃分布面积极小的土壤类型，剩余土类及非土壤单元共 29 个，分别为：棕色针叶林土、灰褐土、灰色森林土、黑钙土、栗钙土、棕钙土、灰钙土、灰漠土、灰棕漠土、棕漠土、龟裂土、风沙土、石质土、草甸土、潮土、沼泽土、盐土、漠境盐土、灌淤土、灌漠土、草毡土、黑毡土、寒钙土、冷钙土、寒漠土、寒冻土等土类，以及湖泊水库、冰川雪被、西北盐壳等非土壤土地覆被类型。

6.1　训练数据样本采集

将 1∶100 万土壤数据库矢量数据以统一的地理坐标叠合到待分类数据集上，在各土类图斑上分别对 29 个分类对象选取样点，按照 3∶1 的比例分别采集训练样本和验证样本。

6.2　土壤类型可分度分析

为了判定分类特征对各土壤类型的可分性大小的影响，本研究采用了 Jeffries-Matusita 距离(J-M 距离)进行了可分度分析。J-M 距离是类对间统计可分性的一种度量，是两个类别的密度函数之间的平均差异的一种度量[15]。J-M 距离取值在 0～2.0

之间，值越大，类对间可分性越好。一般认为大于1.9的类之间具有较好的可分性，本研究对小于1.9的类进行合并。

采用训练样本点对分类数据集的采样分析，J-M距离值小于1.9的类别见表4。

表4 几种土类的J-M距离值

Table 4 J-M distances of several soil unit pairs

土 类		J-M距离值
黑毡土	黑钙土	1.81
草甸土	潮土	1.84
草毡土	冷钙土	1.87

将可分度不佳的土壤类型训练样本合并，对遥感土壤分类的系统方案进行调整，调整后的土壤和地表基本类型为26类。同时相应地调整数据采样方案(见表3)。

7 分类

采用最大似然监督分类方法，结果如图2(见图版Ⅱ)所示。

8 精度分析

本研究采用3种方法对分类结果进行了验证和评价，即混淆矩阵、Kappa指数以及其它方法获取的典型土壤分布面积。

8.1 混淆矩阵和Kappa指数

使用验证样本，以混淆矩阵分析方法计算总分类精度R和一致性指数Kappa。结果总体分类精度达到69.95%，整体Kappa统计值为0.69。

用户精度表示从分类结果图中任取随机样本所具有的类型与地面实际类型相同的条件概率，它反映分类图中各类别的可信度，即这幅图的可靠性。制图精度表示参考图的任意一个随机样本与分类图上同一地点的分类结果相一致的条件概率，它反映用于生产这张分类图的方法的好坏。制图精度和用户精度对应的是漏分误差和错分误差。漏分误差显示的是实际的某一类地物有多少被错误地分到其它类别；而错分误差显示的则是图像中被划为某一类地物实际上有多少应该是别的类别。表5列出了分类结果各类的制图精度和用户精度。从中可看出，漏分误差较大的类有：草毡土＋冷钙土、灰漠土、棕钙土等，错分现象较严重的主要有：栗钙土等。这几种土壤类型在景观特征上具有相似性，这是造成分类精度较差的主要原因。此外，不同土类的土壤中可能存在景观较相似的土壤亚类，这也对分类的精度造成了限制。

表5 分类精度

Table 5 Accuracy of classification

土 类	制图精度%	用户精度%
冰川雪被	97.43	96.88
草甸土＋潮土	53.14	87.32
草毡土＋冷钙土	28.86	86.32
风沙土	98.00	85.78
栗钙土	70.00	38.57
灌漠土	100.00	60.98
灌淤土	96.67	70.05
龟裂土	76.00	58.46
寒冻土	46.00	76.63
寒钙土	67.67	60.06
寒漠土	94.50	55.26
黑毡土＋黑钙土	58.83	62.81
湖泊、水库	99.33	100.00
灰钙土	99.33	95.51
灰褐土	94.00	70.57
灰漠土	41.20	59.54
灰色森林土	97.00	55.59
灰棕漠土	88.60	61.27
石质土	68.22	58.59
盐土	65.50	74.86
沼泽土	100.00	90.91
棕漠土	68.37	83.27
棕钙土	31.06	58.61
棕色针叶林土	94.03	94.03
漠境盐土	92.50	60.66
西北盐壳	100.00	93.75

8.2 不同方法典型土壤分布面积比较

表6中比较了不同调查方法获得的几类土壤分布面积。其中，遥感分类与常规调查和基于全国第二次土壤普查的1∶100万土壤数据库数据有一定的差异，其原因大致有以下几点：

(1)几种数据的指标定义不一致。例如：常规调查涉及的风沙土实际上是沙漠面积，其面积较大于土壤分类结果。

(2)数据的时态不同。例如：1∶100万土壤数据库的数据源于第二次全国土壤普查，该工作始于1979年，1994年结束，历时16 a。而遥感分类基础为2004年现状，两种方法的基础景观特征已有一定的差异。

(3)因采用的方法不同，各调查结果均具有各自的误差。

表6 新疆几类土壤的分布面积比较(单位：10^4 km^2)

Table 6 Comparing areas of several soil units in Xinjiang

土壤类型	常规调查	1∶100万土壤数据库	遥感分类
风沙土(新疆沙漠)	43.81[16]	37.5429	30.3702
风沙土(塔克拉玛干沙漠)	36.5[17]	27.793	24.968
风沙土(古尔班通古特沙漠)	5.113[17]	3.7293	3.7899
盐土	13.36[6]	6.3783	8.6021
漠境盐土	1.6833	1.7987	2.8141
西北盐壳		1.8456	2.5702

9　结　语

本研究选择采用 MODIS 遥感数据和地形数据,参照土壤发生学分类系统,对新疆地区进行了遥感分类研究,并对分类结果进行了精度验证和分析。研究结果表明,MODIS 数据系列产品与地形数据的结合使用对 500～1 000 m 分辨率较大区域尺度上的土壤分类研究是适用的,并且具有相当大的应用潜力。而结合地面调查和遥感信息特征,对土壤分类系统进行分析和调整,并建立适合遥感分类的土壤分类系统是本研究方法需进一步研究和改进的地方。

参考文献:

[1] Vettorazzi C A , Bayramin I, Baumgardner M F. Evaluation of AVHRR Data for Delineating Regional Soil Patterns[R]. Post-doctoral Research Report, Agronomy Department, Purdue University, 1995.

[2] Odeh I O A, McBratney A B. Using NOAA Advanced Very High Resolution Radiometric Imageries for Regional Soil Inventory[A]. Proceedings of the 16th World Congress of Soil Science[C]. Montpellier, France. 1998.

[3] Odeh I O A, McBratney A B, Chittleborough D J. Further Results on Prediction of Soil Properties from Terrain Attributes: Heterotropic Cokriging and Regression Kriging[J]. Geoderma, 1995,67:215-226.

[4] Dobos E, Micheli E, Baumgardner M F, *et al*. Use of Combined Digital Elevation Model and Satellite Radiometric Data for Regional Soil Mapping[J]. Geoderma, 2000, 97:367-391.

[5] Dobos E, Montanarella L, Negre T, *et al*. A Regional Scale Soil Mapping Approach Using Integrated AVHRR and DEM Data[J]. International Journal of Applied Earth Observation and Geoinformation, 2001, 3(1):30-42.

[6] 新疆维吾尔自治区农业厅,新疆维吾尔自治区土壤普查办公室. 新疆土壤[M]. 北京:科学出版社,1996.

[7] 刘玉洁,杨忠东. MODIS 遥感信息处理原理与算法[M]. 北京:科学出版社,2001:1-3.

[8] 张定祥,史学正,于东升,等. 中国 1∶500 万土壤数据库建设的基础[J]. 地理学报,2002,57:82-86.

[9] 张定祥,潘贤章,史学正,等. 中国 1∶500 万土壤数据库建设中的几个问题[J]. 土壤通报,2003,34(2):81-84.

[10] 史学正,于东升,潘贤章,等. 我国 1∶100 万土壤数据库及其应用,面向农业与环境的土壤科学[M]. 北京:科学出版社,2004.

[11] 曾志远. 卫星遥感图像计算机分类与地学应用研究[M]. 北京:科学出版社,2004.

[12] 赵英时. 美国中西部沙山地区环境变化的遥感研究[J]. 地理研究,2001,20(2):213-219.

[13] 王建. 卫星遥感雪盖制图方法对比与分析[J]. 遥感技术与应用,1999,14(4):29-36.

[14] 惠凤鸣,田庆久,李英成,等. 基于 MODIS 数据的雪情分析研究[J]. 遥感信息,2004,4:35-38.

[15] 赵英时. 遥感应用分析原理与方法[M]. 北京:科学出版社,2003.

[16] 钟德才. 中国沙海动态演化[M]. 兰州:甘肃文化出版社,1998.

[17] 吴正. 风沙地貌与治沙工程学[M]. 北京:科学出版社,2003.

A Regional Scale Soil Classification Approach Based on MODIS Data

KANG Qing[1], ZHANG Zeng-xiang[2], ZHAO Xiao-li[2]

(1. *Pearl River Hydraulic Research Institute*, *Guangzhou* 510611, *China*;
2. *Institute of Remote Sensing Applications*, *CAS*, *Beijing* 100101, *China*)

Abstract: The aim of this study was to test a method for small-scale soil classification using MODIS and digital elevation data in Xinjiang province, northwest of China, where the soil survey is restrained commonly. A series of MODIS data products including surface reflectance, vegetation indices, land surface temperature of daytime and nighttime, etc., were processed to extract features of images, while several terrain descriptors were achieved with DEM, such as elevation, slope, curvature, aspect, and so on. These layers of data were integrated to a dataset for classification. Considering the complexity of classification, we adopt the soil class of soil genesis taxonomy as unit of classification. With the data of soil genesis taxonomy achieved in the second soil survey in China, training points were collected and used to analyze separability of soil units. According to the J-M distances, some pairs of soil classes with same landscapes have been merged. At last, there are 26 units confirmed. Using the maximum likelihood classifier, the dataset was classified. A performance of about 70% was achieved, assessed with confusion matrixes.

Key words: Remote sensing, MODIS, Soil classification, Arid area, Soil-terrain database