

基于 MODIS-NDVI 时序数据的黄土丘陵区 土地覆盖分类研究

陆广勇¹, 杨勤科^{2,1}, 王海江¹

(1. 中国科学院 水利部 水土保持研究所, 陕西 杨凌 712100; 2. 西北大学, 西安 710069)

摘 要: 为了获取年际土地覆盖基础数据以支持区域土壤侵蚀定量评价, 探讨了采用基于时间序列谐波分析法(HA-NTS)对延河流域 2006 年 MODIS-NDVI 时间序列进行滤波; 参考当地年植被变化特征, 根据土地覆盖类型的 NDVI 时序曲线特征, 选择量化的分类特征, 进行决策树分类, 并以此为基础建立各个决策树类别掩膜; 基于 MODIS-NDVI 和 MOD09GA, 在各个掩膜区内进行非监督分类, 结合坡度图调整分类结果, 形成延河流域土地覆盖类型图。结果表明: 与 2006 年 TM 分类结果的总体一致性为 75%, Kappa 系数为 0.620 9, 影像分类结果可靠, 分类方法现实可行。

关键词: 植被指数; 时间序列; 决策树; ISODATA; 土地覆盖分类

中图分类号: P931.1; TP79

文献标识码: A

文章编号: 1005-3409(2011)02-0112-04

Land Cover Classification in Loess Hilly and Gully Region Using Multi-temporal MODIS-NDVI Data

LU Guang-yong¹, YANG Qin-ke^{2,1}, WANG Hai-jiang¹

(1. Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling, Shaanxi 712100, China; 2. Northwest University, Xi'an 710069, China)

Abstract: The temporal features of various land objects were studied based on the MOD13Q1, and built the decision tree model for the land cover classification of Yanhe basin by filtering the MODIS-NDVI time-series data of 2006 in the area via the HNATS (Harmonic Analysis of Time Series) algorithm; analyzing the temporal curve features of NDVI, and selecting the rational quantitative classification features for the decision tree model, building masks for each land cover, and performing the unsupervised classifications in the masks based on the data of MODIS-NDVI or MOD09GA, modifying the results via slope data. The results showed the overall accuracy of the final land cover map was 75.00%, and the kappa coefficient was 0.620 9. It is suggested that applying the temporal features in classification model and integrating decision tree with ISO-DATA for the land cover classification are feasible and promising.

Key words: vegetation index; time-series; decision tree; ISODATA; land cover classification

土地覆盖是全球变化研究的重要内容, 其研究为区域和全球尺度的气候和生态过程模型提供了重要的参数。区域尺度的土地覆盖信息主要基于多光谱遥感影像 Landsat MSS/TM/ETM+ 数据提取, 但所需数据量庞大, 而且关键生长时期的影像会受到云的影响而无法获取, 数据更新受限^[1]。

土地覆盖变化具有季节性, 这种季节变化可以利用植被指数的变化进行定量分析研究^[2]。卫星遥感影像能够监测植被的物候变化^[3-5], 来自遥感影像的区域

植被指数 NDVI 值随时间周期性升高和降低是植被生长周期的典型体现^[6]。反映地表覆盖物季节性变化特征的多时相 NDVI 可以作为土地覆盖的分类特征^[5]。

植被指数时间序列的数据源主要包括 NOAA-AVHRR、SPOT-VGT 和 MODIS 等植被指数数据产品。前两者的合成周期为 10 d, 空间分辨率约为 1 km, 多用于全球或大区域的研究^[3,5,7-8]。MODIS 植被指数产品合成周期为 16 d, 但其空间分辨率较高, 有 500 m 和 250 m 两种分辨率产品^[9-11], 因此应用研

究也越来越多^[1,12-16]。

基于植被指数时间序列影像的土地覆盖分类方法可分成非监督分类、监督分类和分层决策树分类 3 种。盛永伟^[3]、李道峰^[7]、宫攀^[1]等基于气象卫星的 NDVI 时间序列影像、运用非监督分类法进行了植被覆盖类型分类;刘勇洪等基于 MODIS 地表反射率、植被指数和地表温度数据产品,选择 6 种分类特征,通过监督分类,将山东省土地覆盖分成了 8 类^[13];张霞等综合运用 ISODATA 非监督分类法、决策树法,对华北平原进行了土地覆盖分类^[14]。

但是很多研究应用中,土地覆盖的季节性变化规律没有被作为主要的分类特征,而且非监督分类法与决策树法的应用中都是在动态聚类的基础上进行的,物理意义不够明确。目前在地形复杂区(如黄土丘陵沟壑区)基于 MODIS 植被指数时间序列影像的土地覆盖分类研究较少。本实验综合运用 MODIS 地表反射率与 NDVI 植被指数两种数据产品、DEM,结合当地物候特征,分析了土地覆盖植被指数内变化特征,运用植被指数、湿度指数和纹理特征进行分层决策树分类,利用 ISODATA 法和当地土地覆盖分布特征调整分层决策树结果,为区域环境动态监测、土壤侵蚀评价提供更高时间分辨率的基础数据。

1 研究区域概况

延河流域地处陕北黄土高原中部,地理位置为 36°23′—37°17′N 和 108°45′—110°28′E,面积达 7 684 km²。该流域属暖温带大陆性半干旱季风气候,年平均气温 9℃,年均降雨量为 500 mm,其中 7—9 月降雨量占全年降雨量的 60%以上。流域植被区划上属于森林草原地带。由于土壤侵蚀严重,土壤类型以黄土母质上发育的黄绵土为主,土壤质地均一,土质疏松,抗侵蚀能力差。流域内沟壑纵横,地形复杂,植被覆盖率低,属典型黄土丘陵沟壑区^[17-18]。

2 研究方法

2.1 数据源

- (1)本文采用 NASA USGS 提供的植被指数数据产品 MOD13Q1 的归一化植被指数 MODIS-NDVI,空间分辨率为 250 m,合成周期为 16 d。从 2006 年 1 月 1 日至 2006 年 12 月 30 日,共 23 个时相影像。
- (2)NASA USGS 提供的地表反射率产品数据 MOD09GA,空间分辨为 500 m,影像获取时间为 2006 年 9 月 9 日,共 7 个波段,但第 5 波段存在坏线,故本实验使用了 MODIS 的 1—4,6—7 共 6 个波段的影像。
- (3)2006 年 9 月 10 日成像的 Landsat TM 5 图像

解译生成的 2006 年土地利用图(简称 LU2006)和矢量化的 2000 年延河流域土地利用图(简称 LU2000)。

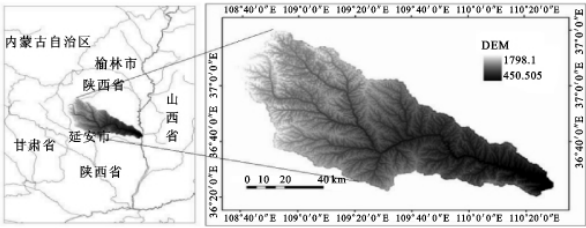


图 1 延河流域地理位置及地势图

(4)DEM:利用 1 : 5 万地形图和 ANUDEM 软件生成的 10 m 分辨率 DEM,用以生成坡度图。

2.2 数据预处理

(1)基于时间序列谐波分析法(HANTS)对 MODIS-NDVI 进行滤波。MODIS-NDVI 影像局部容易受到云等因素的影响,需要进行平滑处理^[15]。本文选择了 HANTS 对 23 景 MODIS-NDVI 进行平滑处理。该方法的核心主要是通过傅里叶变换和最小二乘法,利用谐波重构时间序列数据^[19]。通过 HANTS 滤波后,植被指数时间序列更加光滑(见图 2),更有利于有效揭示物候周期性变化规律,便于地表覆盖信息提取^[19-20]。

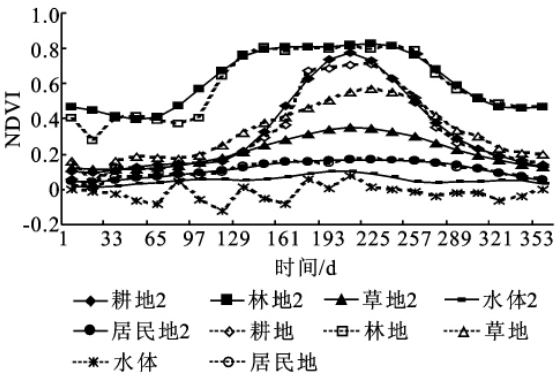


图 2 平滑处理前后典型土地覆盖类型 NDVI 时间序列曲线

- (2)延河流域植被年内变化可划分为休眠阶段、恢复阶段、生长季节、退化阶段 4 个时段^[18]。基于 IDL 语言编程,计算平滑后全年和生长期(GD, growth duration)MODIS-NDVI 时间序列的最大值(max, GDmax)、最小值(min, GDmin)、差值(最大值—最小值, D, GDd)。(3)对平滑后的全年 MODIS-NDVI 影像进行主成分分析(PCA)。基于反映年内植被覆盖差异较大耕地信息 PCA 第二组分,在 3 × 3 窗口下,进行纹理特征局部平滑(Homogeneity, Homo),以便提取耕地信息。
- (4)利用 ENVI 系统的 GLT(Geometry Lookup Table)对 MOD09GA 和 MODIS-NDVI 进行几何校正^[21]。将滤波后的 MODIS-NDVI 及其生成数

据、MOD09GA 进行 Albers 等面积投影。并以 DEM 为基准,首先将 TM 5 影像和 2006 年土地利用图等进行空间配准,然后基于配准后的 TM 5 影像对 MODIS-NDVI、MOD09GA 进行配准。空间配准的误差都控制在一个像元以内。

(5) 基于 MOD09GA 提取了影像的湿度指数 (NDMI, Normalized Difference Water Index)^[13], 其表达式如式(1)。

$$\text{NDMI} = (\rho_4 - \rho_6) / (\rho_4 + \rho_6) \quad (1)$$

式中: ρ_4, ρ_6 ——经过标准化处理的 MODIS 可见光绿色波段(第 4 波段)和短红外波段(第 6 波段)。波段 ρ_6 处于水的吸收带,对湿度、含水量信息尤为敏感,而 ρ_4 对水体的反射率较高,因此利用 NDMI 提取水体的外边界。

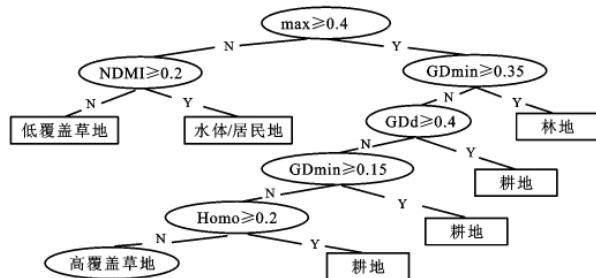
2.3 影像分类方法

不同的土地覆盖类型具有不同的植被时间序列曲线。根据当地的年内植被变化特征,基于植被指数响应曲线特征,进行影像分类。本文综合运用 MODIS 数据的时间、光谱、空间信息,选取了植被指数 NDVI、湿度指数、纹理特征作为分层决策树的分类特征,基于阈值进行分层决策树分类,然后运用 ISO- DATA 法分类别进行动态聚类,调整决策树分类结果,最后利用 DEM 调整耕地类型,最终形成土地覆盖类型图。

(1) 分析 NDVI 时间序列曲线,参考 LU2000、LU2006 和植被年内变化特点^[18],选择各个指标阈值,进行分层决策树分类,并生成各土地覆盖类型掩膜。如图 3 所示,耕地的 NDVI 曲线在生长期(第 113~289 d)上升和下降梯度最大,波峰明显,差值最大;典型林地的 NDVI 值全年都较大,在生长期最小值略小于 0.4;草地也存在较明显的波峰,但其变化梯度小于耕地,且峰值一般小于林地和耕地;而居民地和水体 NDVI 年内变化较小,因为居民地内往往与一些植被(如城市绿地)混合,所以居民地的 NDVI 值略大于水体。根据各种土地覆盖的时间响应曲线特征,选择合适的阈值,进行决策树分类。

(2) 基于决策树分类类型掩膜,利用迭代自组织数据分析法(ISODATA 法),对 MODIS-NDVI 图像进行动态聚类(非监督分类)^[14],参考土地利用图和 MOD09GA 多光谱影像,调整决策树分类结果。MODIS-NDVI 主要反应植被信息,因此,基于 MODIS-NDVI 时间序列影像对林地、高覆盖草地进行非监督分类。在植被盖度较低区域,NDVI 指标反应植被盖度不够灵敏,单纯依靠 NDVI 时间序列难以正确识别土地覆盖类型^[6],因此基于地表反射率影像

MOD09GA,对低覆盖草地、水体与居民地进行非监督分类。



注: max、min 分别代表全年 NDVI 的最大值和最小值; GDmin 代表生长周期内 NDVI 的最小值; GDd 是指生长周期内 NDVI 最大值与最小值之差; HOMO 是指局部平滑纹理特征值; NDMI 是指影像湿度指数。Y 表示椭圆内的条件成立, N 则表示不成立

图 3 决策分类树示意图

(3) 退耕还林(草)实施后,延河流域内大于 25° 的坡耕地基本都进行了退耕,因此基于 DEM 形成的坡度图,将以上遥感分类结果中大于 25° 的耕地调整为草地。考虑到所用数据源分辨率比较低,因而本次研究将要划分的土地利用类型为耕地、林地、草地、居民地和水体 5 类。

3 结果分析

3.1 空间格局与数量统计分析

基于上述分类方法得到延河流域 2006 年土地覆盖分类结果,结果见图 4。植被分布由东南到西北逐渐减少,与流域年降雨量空间分布一致,林地主要分布在流域的南部,与李正国的研究结果一致^[18],占流域总面积的 15.9%;草地遍布整个流域,是主要土地覆盖类,占流域总面 67.3%,其中约 15% 的草地是根据坡度土从遥感影像分类结果中的耕地调整而来;耕地主要沿沟道分布,在流域上游也呈现集中分布,为流域总面积的 16.5%;居民地与水体两者所占比重不足 1%,主要原因是两种流域面积比重较少,分布零散,在 250 m 的影像上,像元混合现象严重^[14]。

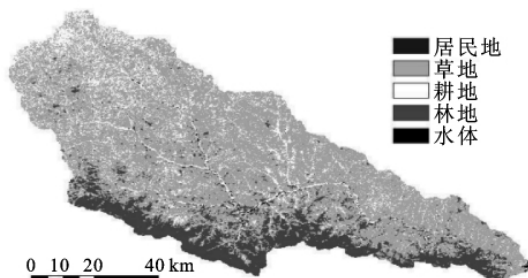


图 4 2006 年土地分类结果

3.2 精度分析

选择 2006 年基于 TM 的土地利用/覆盖分类图 LU 2006 进行比较,同时参考 LU 2000,进行分类精度评价。精度分析在 ENVI 4.7 软件下完成,随机选择耕地、草地和林地各覆盖类型的 0.1% 作为验证像

元;由于居民地、水体的像元较少,故每一类各选择 10 个像元作为验证,5 类地表覆盖类型共 152 个像元,占总流域面积近 0.12%^[22]。混淆矩阵(见表 1)表明总体分类精度为 75.00%,Kappa 系数为 0.620 9,除了居民地,其它 4 类的用户精度到大于 70%,这就说明分类效果良好。用户精度是分类结果可靠性指标^[6],是最常用的精度指标^[1,13-14,23]。表 1 表明,实际地表的林地和耕地被错分为草地的情况较

表 1 误差矩阵

项目	居民地	草地	耕地	林地	水体	总和	制图精度	用户精度
居民地	5	2	3	0	0	10	62.50	50.00
草地	0	64	14	10	1	89	90.14	71.91
耕地	1	5	16	0	0	22	48.48	72.73
林地	0	0	0	21	0	21	67.74	100.00
水体	2	0	0	0	8	10	88.89	80.00
总和	8	71	33	31	9	152		

4 结论与讨论

(1)以季节性变化特征作为分类基础,根据植被指数时序曲线、影像纹理和湿度指数等分类特征,进行分层决策树分类,并结合其它辅助数据(如地表反射率、DEM 等),按类别进行非监督分类,最后调整土地覆盖分类结果获得土地覆盖分类图,具有一定的物理意义。分类结果与 2006 年土地利用总体一致性为 75.00%,Kappa 系数为 0.620 9,图像分类结果良好,能为区域环境监测、土壤侵蚀评价提供可靠的年际土地覆盖基础数据。

(2)根据退耕还林后大于 25°陡坡上耕地稀少的特点,利用 DEM 生成的坡度图,调整遥感分类结果中的耕地类型,能在一定程度上减小土地覆盖类型分布破碎对分类结果带来的影响,改善分类效果。

(3)MODIS-NDVI 虽然在空间分辨率上较 NOAA-AVHRR 和 SPOT-VGT 有了提高,但是在地形复杂、斑块破碎的黄土丘陵沟壑区,在 250 m 的影像上像元混合现象严重,小斑块地类较难识别,导致居民地、水体在分类结果中面积比重小于实际水平。

(4)分类结果中相当一部分耕地和林地被错分为草地,主要原因是当地土地覆盖类型分布破碎,影像分辨率较粗,像元混合现象严重。下一步工作可以结合当地植被的物候特征、分布特点,通过 NDVI 时间序列曲线的变化特征进行更加细致地定量分析,参考混合像元分解方法,对区域土地覆盖类型进行分类,提高分类精度。

参考文献:

[1] 宫攀,陈仲新,唐华俊,等. 基于 MODIS 温度/植被指数的东北地区土地覆盖分类[J]. 农业工程学报,2006,22

为严重,分类结果中草地类型被高估,耕地和林地被低估。这主要原因在于延河流域地表破碎,林地、耕地分布不集中,在影像分辨率较粗的情况下,像元内信息趋于平滑。同时,居民地与水体两者地表反射率都较低,反射特征相近,所以两者之间也存在混分现象。居民地(包括交通用地)和水体的斑块一般较小,在粗分辨的遥感影像上很难识别,因此分类结果中两者容易被低估。

(9):94-99.

[2] 布和敖斯尔,刘纪远,吴祖南. 基于季相及经度特征的中国土地覆盖变化遥感研究[J]. 地理学报,1998,53(12): 52-60.

[3] 盛永伟,陈维英,肖乾广,等. 利用气象卫星植被指数进行我国植被的宏观分类[J]. 科学通报,1995,40(1):68-71.

[4] Jonsson P,Eklundh L. Seasonality extraction by function fitting to time-series of satellite sensor data[J]. Ieee Transactions on Geoscience and Remote Sensing,2002, 40(8):1824-1832.

[5] 汪权方,李家永. 基于时序 NDVI 数据的中国红壤丘陵区土地覆被分类研究[J]. 农业工程学报,2005,21(2): 72-77.

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

[7] 李道峰,田英,郝芳华. 基于 NDVI 数据的黄河流域地表植被覆盖综合分类研究[J]. 水土保持研究,2003,10 (4):88-91.

[8] 卢玲,李新,董庆罕,等. SPOT4-VEGETATION 中国西北地区土地覆盖制图与验证[J]. 遥感学报,2003,7(3): 214-221.

[9] Vegetation Indices 16-Day L3 Global 500m[EB/OL] . https://lpdaac.usgs.gov/lpdaac/products/modis_products_table/vegetation_indices/16_day_l3_global_500m/mod13a1.

[10] Vegetation Indices 16-Day L3 Global 250m[EB/OL] . https://lpdaac.usgs.gov/lpdaac/products/modis_products_table/vegetation_indices/16_day_l3_global_250m/myd13q1.

[11] 王正兴,刘闯,Huetealfredo. 植被指数研究进展:从 AVHRR-NDVI 到 MODIS-EVI[J]. 生态学报,2003,23 (5):979-987.

参考文献:

- [1] 刘正茂,夏广亮,赵艳波. 三江平原地下水开采对生态过程的影响[J]. 水利发展研究,2006(12):38-41.
 - [2] 王韶华,苏轶醒,刘昆鹏. 三江平原水资源的合理开发利用[J]. 中国农村水利水电,2005(7):26-28.
 - [3] 王韶华,刘文朝,刘群昌. 三江平原农业需水量及适宜水稻种植面积的研究[J]. 农业工程学报,2004,20(4):50-53.
 - [4] 田光忠. 昌平区地下水位动态研究及效果[J]. 北京水务,2007(5):44-47.
 - [5] 管新建,逯洪波,徐清山. 基于 BP 神经网络的区域地下水位动态预测[J]. 人民黄河,2006,28(8):40-41,79.
 - [6] 王文圣,丁晶,李跃清. 水文小波分析[M]. 北京:化学工业出版社,2005.
 - [7] 李贤彬,丁晶,李后强. 水文时间序列的子波分析法[J]. 水科学进展,1999,10(2):144-149.
 - [8] 王文圣,向红莲,黄伟军,等. 基于连续小波变换的径流分维研究[J]. 水利学报,2005,36(5):598-601.
 - [9] Bohyun Chon, Yong-Suk Choi. Modeling of three-dimensional groundwater flow using the method to calculate fractal dimension[J]. Korean Journal of Chemical Engineering, 2001, 18(3):382-386.
 - [10] 武晋芳,冯玉广. 山西省城镇体系结构的分形分析[J]. 系统科学学报,2008,16(4):59-63,94.
 - [11] 钟亮,许光祥. 分形理论在紊流与泥沙研究中的应用现状[J]. 重庆交通大学学报:自然科学版,2008,27(5):806-811.
 - [12] 孙金花,冯英俊,胡健. 基于分形理论的股票时序数据
- ~~~~~
- (上接第 115 页)
- [12] Guerschman J P, Hill M J, Renzullo L J, et al. Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 Hyperion and MODIS sensors[J]. Remote Sensing of Environment, 2009, 113(5):928-945.
 - [13] 刘勇洪,牛铮. 基于 MODIS 遥感数据的宏观土地覆盖特征分类方法与精度分析研究[J]. 遥感技术与应用, 2004, 19(4):217-224.
 - [14] 张霞,孙睿,张兵,等. 基于 MODIS 植被指数时间谱的华北平原土地覆盖分类[J]. 农业工程学报,2006,22(12):128-132.
 - [15] 左丽君,张增祥,董婷婷,等. MODIS/NDVI 和 MODIS/EVI 在耕地信息提取中的应用及对比分析[J]. 农业工程学报,2008,24(3):167-172.
 - [16] 王红说,黄敬峰. 基于 MODIS NDVI 时间序列的植被覆盖变化特征研究[J]. 浙江大学学报:农业与生命科学版,2009,35(1):105-110.
 - [17] 谢红霞. 延河流域土壤侵蚀时空变化及水土保持环境
- 离群模式挖掘研究[J]. 运筹与管理,2008,17(5):135-140.
 - [13] Paramanathan P R. Uthayakumar. Application of fractal theory in analysis of human electroencephalographic signals[J]. Computers in Biology and Medicine, 2008, 38(3):372-378.
 - [14] 刘起方,马光文,刘群英,等. 基于分形插值理论的径流预测探讨[J]. 水力发电学报,2008,27(4):20-25.
 - [15] 田丰,王晋国,王明祥. 基于分形理论的地震数据分析[J]. 长安大学学报:自然科学版,2008,28(2):108-110.
 - [16] Mandelbrot B B. The Fractal Geometry of Nature[M]. New York:W. H. Freeman and Company,1982.
 - [17] 刘东,付强. 小波随机耦合模型在三江平原井灌区地下水位预测中的应用[J]. 数学的实践与认识,2008,38(16):54-62.
 - [18] 刘东,付强. 基于小波变换的三江平原低湿地井灌区年降水序列变化趋势分析[J]. 地理科学,2008,28(3):380-384.
 - [19] 丁晶,刘国东. 日流量过程分维估计[J]. 四川水力发电,1999,18(4):74-76.
 - [20] 王文圣,向红莲,赵东. 水文序列分形维数估计的小波方法[J]. 四川大学学报:工程科学版,2005,37(1):1-4.
 - [21] Wornell G. Signal Processing with Fractal: a wavelet based approach[M]. Znc N J:Prentice Hall PTR,1995.
 - [22] 张洁瑜,马克明. 建三江农垦分局湿地植物群落分类和排序[J]. 湿地科学,2008,6(3):359-365.
 - [23] 李相莉,韩金超,王士兰,等. 建三江分局水利工程现状及主要问题分析[J]. 现代化农业,2008(1):17-18.
- ~~~~~
- 效应评价研究[D]. 西安:陕西师范大学,2008.
 - [18] 李正国,王仰麟,吴健生. 等. 不同土地利用方式对黄土高原植被覆盖季节变化的影响:以陕北延河流域为例[J]. 第四纪研究,2005(06):762-769.
 - [19] Roerink G M, Menent W. Verhoef. Reconstructing cloudfree NDVI composites using Fourier analysis of time series[J]. International Journal of Remote Sensing, 2000, 21(9):1911-1917.
 - [20] 王丹,姜小光,唐伶俐,等. 利用时间序列傅立叶分析重构无云 NDVI 图像[J]. 国土资源遥感,2005(2):29-32.
 - [21] 张树誉,王钊,李星敏. 提高卫星遥感资料利用率的方法[J]. 测绘学院学报,2004,21(1):30-33.
 - [22] 刘咏梅,杨勤科,温仲明. 地形复杂地区遥感图像分类方法应用研究:以黄土丘陵沟壑地区坡耕地遥感调查为例[J]. 水土保持通报,2003,23(4):30-32.
 - [23] Kumar U N, Kerle T V. Ramachandra. Constrained Linear Spectral Unmixing Technique for Regional Land Cover Mapping Using MODIS Data[J]. Innovations and Advanced Techniques in Systems, Computing Sciences and Software Engineering, 2008: 416-423, DOI: 10.1007/978-1-4020-8735-6_78.