

基于 MODIS 影像和随机森林算法的 河北坝上林地动态监测

周佳宁¹ 张洁¹ 李天宏^{1,2,†}

1. 北京大学深圳研究生院环境与能源学院, 深圳 518055; 2. 北京大学环境科学与工程学院, 北京 100871;

† 通信作者, E-mail: litianhong@iee.pku.edu.cn

摘要 为了揭示河北省坝上地区林地的动态变化特征, 采用 250 m 空间分辨率的 MODIS 反射率和 NDVI 数据, 利用 TM 遥感影像辅助选择训练样本, 基于随机森林分类算法, 提取 2000—2015 年 8 个时相的林地信息, 并分析其空间变化情况。结果表明, 与常用的最大似然法和神经网络法相比, 随机森林法分类的精度更高, 总体精度和 Kappa 系数分别为 91.89% 和 0.88。通过二进制编码方法, 快捷地揭示了 8 个时相的林地信息在空间上的动态变化, 识别出变化幅度较大的年份和空间分布。结果显示, 林地退化严重的地区集中在丰宁、围场、张北和沽源四县, 时间集中在 2002、2010 和 2013 年。

关键词 坝上林地; MODIS; 随机森林算法; 动态监测

中图分类号 X835

Bashang Forest Change Monitoring with Multi-Temporal MODIS Images and Random Forest Algorithm

ZHOU Jianing¹, ZHANG Jie¹, LI Tianhong^{1,2,†}

1. School of Environment and Energy, Peking University Shenzhen Graduate School, Shenzhen 518055; 2. College of Environmental Sciences and Engineering, Peking University, Beijing 100871; † Corresponding author, E-mail: litianhong@iee.pku.edu.cn

Abstract In order to reveal the dynamic characteristics of the forest in Bashang area of Hebei Province, MODIS reflectivity and NDVI data with a spatial resolution of 250 m were used for forest classification, and a Thematic Mapper (TM) image in 2005 was resorted to aid training sample selection. With Random Forest Algorithm and time series of MODIS images, the forest in Bashang area was monitored from 2000 to 2015 in every two years. Compared with widely used classifiers such as maximum likelihood classifier and BP artificial neural network algorithm, Random Forest classifier showed the best performance with its overall accuracy and Kappa coefficient being 91.89% and 0.88 respectively. Binary coding was applied to the eight phases of forest distribution images, which can easily and rapidly reflect the changing trajectory from phase to phase. It showed that the severe forest changes mainly occurred in counties of Fengning, Weichang, Zhangbei and Guyuan during the years of 2000, 2010, and 2013.

Key words forest in Bashang area; MODIS; Random Forest Algorithm; dynamic monitoring

我国从 20 世纪 70 年代末开始兴建西北—华北—东北地区巨型人工林生态工程(简称三北防护林), 用以减缓日益加速的荒漠化和水土流失进程。工程建设实施 30 多年来, 提高了“三北”地区森林覆被

率, 增加了低中产区粮食产量, 有效地减缓了沙漠化和水土流失进度, 增加了该地区的生物碳储量。河北省北部的坝上林区是三北防护林的重要组成部分和首都生态安全的重要屏障, 在减轻环北京地区

的风沙危害中发挥了重要作用。利用遥感技术,可以快速、准确和大范围地监测林地的生长及恢复情况,为林地的建设、维护和更新提供科学依据。

目前,对林地进行遥感监测的研究主要针对防风固沙林、农田防护林区和沿海防护林区^[1-3]。对于面积较小的区域,主要使用 TM (Thematic Mapper), SPOT 影像^[4-5]和航拍影像^[6]。对于面积较大的区域,由于使用 Landsat TM/ETM+和 SPOT 等资源卫星数据和处理成本太高,因此往往使用空间分辨率较低的免费的中分辨率成像光谱仪(Moderate Resolution Imaging Spectrometer, MODIS)数据^[7-8]。与资源卫星数据相比,MODIS 数据的时间分辨率高,在动态监测方面有明显优势。随着计算机技术和数字图像自动分类技术的日益普及,计算机数字图像分类方法已经越来越多地应用于林地遥感监测中^[2-3,9-10]。从总体上来看,这些研究的监测精度在 75%~96%之间,长时间序列的监测研究较少。

随机森林算法(Random Forest Algorithm, RFA)是近年发展起来的一种分类算法^[11-12],在多光谱、多时相和高维遥感影像分类中表现出较高的分类精度、较快的运算速度和稳定性^[13-14],但目前尚未在林地动态监测中得到应用。本文以河北坝上林区为研究区域,采用 2000—2015 年间多时相 MODIS 数据,运用随机森林算法监测坝上林地生长、消失和恢复状况,以期为该地区林地的维护和更新提供科学依据。

1 研究区域

河北坝上地区指北京北端与内蒙古高原南端之间的区域,包括张家口坝上的张北、尚义、康保、沽源四县,承德坝上的丰宁、围场两县,地跨北纬 $40^{\circ}73' - 43^{\circ}34'$,东经 $113^{\circ}50' - 118^{\circ}30'$,总面积为 3 万多 km^2 (图 1)。该地区地形地貌比较复杂,包括山地、丘陵、高原及沙地等几个重要类型,总体上呈典型的波浪起伏状高原景观^[15]。

坝上地区位于中温带亚干旱地区,属东亚温带大陆性季风气候^[16]。降水量年际变化较大,年平均降水量为 340~450 mm。年平均气温在 $-0.3 \sim 3.5^{\circ}\text{C}$ 之间,年蒸发量为 1850 mm。大风日较多,冬季常出现雪暴和沙尘暴。该地区属于农牧交错带,以半干旱草原生态系统为主,主要植被类型有杨树林、榆树林、白桦林和大针茅草原等^[17]。防护林以杨树和榆树为主,兼有少部分柠条、枸杞和沙棘等灌木林。从 20 世纪 50 年代初至今,坝上草场的面积由 86 万 hm^2 减少到 51 万 hm^2 ,牧草的覆盖率从 90%降低到 44%,而耕地面积却扩张近一倍。根据第六次全国人口普查结果,坝上地区总人口超过 1628 万。

2 研究方法

根据区域特征,选择植被生长状况最好的 8 月的图像集进行林地变化监测。主要步骤: 1) 目视解

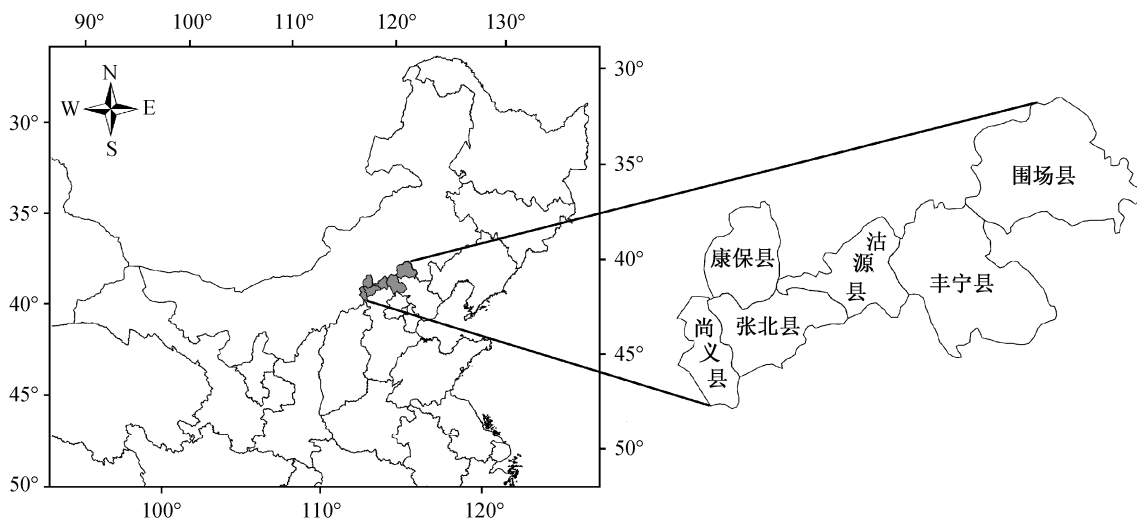


图 1 研究区位置示意图

Fig. 1 Location of the study area

译 2005 年 8 月 Landsat TM 数据, 选择坝上不同土地覆被类型作为训练数据集; 2) 基于随机森林算法, 生成 2005 年 8 月 MODIS 分类模型; 3) 利用 2005 年分类模型, 对其他 7 个年份 8 月的影像进行土地覆被分类, 提取林地分布区域, 对分类结果进行时空分析。

2.1 数据及预处理

本研究使用的 MODIS 数据(<http://modis.gsfc.nasa.gov/>)包括 2000, 2002, 2004, 2005, 2008, 2010, 2013 和 2015 年共 8 个时相。为了减小数据量和预处理的工作量, 选用分辨率为 250 m 的 MOD09Q1 反射率产品数据以及 MOD13Q1 的 8 天合成的归一化植被指数(normalized difference vegetation index, NDVI)数据。

为了提高选取训练样本的准确性, 利用成像时间同为 2005 年 8 月、图像质量好、分辨率为每像素 30 m 的 Landsat TM 影像作为选择训练样本的参考依据。用 TM 近似真彩色合成图像(波段 5, 4, 3 分别赋予红、绿、蓝色显示)来区别城市、农田、植被、水体等地物类型。TM 假彩色合成图像(波段 4, 3, 2 分别赋予红、绿、蓝色显示)主要用来区别草原和林地, 草原因生物量较小而显示为浅红色, 林地显示为深红色/亮红色。通过人工目视解译, 标记坝上地区典型地物类型(农田、水体、林地和草原 4 类)。由于图像中城市的像元数少, 且在 MODIS 250 m 分辨率图像上会与其他地物类别混合, 所以本研究只分 4 种地物类型。对在 TM 图像上标记出的不同地物类型进行随机抽取(30367 个农田像元, 61859 个水体像元, 196111 个林地像元, 13298 个草地像元), 将其中 50% 的兴趣区(region

of interest, ROI)作为训练样本, 剩余的 50% 作为检验样本。通过几何校正和重采样等预处理, 将标记的不同土地覆被类型的感兴趣区加载到 MODIS 影像上进行图像分类和精度检验。

2.2 随机森林分类原理和实现途径

随机森林法是由 Breiman^[18]提出的一种有效的分类器集成策略, 原理是利用 bootstrap (进化树)重抽样方法, 从原始样本中抽取多个样本, 对每个 bootstrap 样本进行决策树建模, 然后组合多棵决策树来预测, 通过投票得出最终预测结果。该方法已应用于商业分析^[19]、生物医学^[20-21]和遥感影像分类^[22-23]等领域。已有理论研究和大量实证研究证明, 随机森林分类具有很高的准确率, 对异常值和噪声具有很好的容忍度, 且不易出现过度拟合^[14]。随机森林是一种自然的非线性建模工具^[24], 其集成思想如图 2 所示。

原始输入的样本 S_1, S_2, \dots, S_n 被随机选取, 构成新的训练集 X 来取代原始样本, 用于降低分类树之间的相关性。再从 X 中随机选取一部分样本作为 bootstrap 训练集, 生成一一对应的回归树。在给定的自变量 X 下, 每个分类树的分类结果都由投票决定。随机森林利用构造不同的训练集来增加分类模型间的差异性, 提高分类模型的外推预测能力。

本研究中, 随机森林模型的具体实施是借助遥感影像处理软件包 ENVI/IDL 和 Weka 软件(<http://www.cs.waikato.ac.nz/ml/weka/>)。首先, 对 2005 年的 MODIS 数据集进行分类, 利用训练数据生成随机森林分类模型。然后, 通过检验数据, 对模型进行验证。模型验证方法采用遥感图像分类精度评价的标准方法, 即基于混淆矩阵^[25]计算得到的总体精

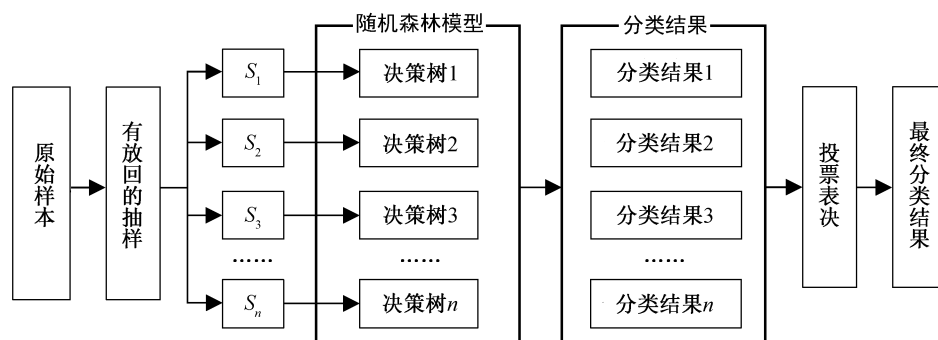


图 2 随机森林算法原理
Fig. 2 Principle of Random Forest Algorithm

度(overall accuracy, OA)、用户精度、生产者精度以及评价分类方法的 Kappa 系数。一般认为 Kappa 系数大于 0.7 即满足最低的判别精度要求^[26], 可用于后续的分析。

将验证后的模型应用于其他 7 个时相的 MODIS 数据集, 进行土地覆被分类, 提取每个时相的林地信息。经过反复试验, 在 Weka 软件中设定分类树的数目, 即 numTrees=100 时效果较佳。其他参数为默认值, 即 maxDepth=0, numFeatures=0, seed=1。

2.3 林地时空变化分析

对遥感图像进行解译或对分类结果进行多相时空分析时, 一般需要将数据转换到地理信息系统, 采用系统中的叠加操作, 分析新生成图斑随时间变化的内涵。这样做不仅需要在图像处理软件与地理信息软件之间进行数据交换, 工作量大, 而且不利于进行细碎的图斑分析处理^[27]。考虑到图像的分辨率较大, 在时空变化图中, 有些年份的变化不是很大, 为了更快捷、直观地表示时空的变化情况, 本研究通过对时间序列的分类图像进行二进制编码来综合林地时空变化信息。与 ArcGIS 空间叠置分析的传统方法相比, 该方法具有不需要进行数据转换、工作量小、空间存储量小和便于表达长时间序列变化等特点^[27]。

一个字节(Byte)有 8 位(Bit), 分别代表 2000—2015 年 8 个时相中林地“有”(赋值为 1)或“无”(赋值为 0)的状态(表 1)。在生成的编码结果图上, 通过解读某像元的二进制值, 就可以简便直观地获得该像元上林地的年际变化情况。例如, 某像元的编码值为 72, 对应的二进制位为 01001000, 表示该像元在 2002 年和 2008 年为林地, 而在其他年份林地消失。

表 1 林地多时相变化信息的编码方法
Table 1 Coding method for multiple temporal change of forestry

字节的位置	影像年份
Bit7	2000
Bit6	2002
Bit5	2004
Bit4	2005
Bit3	2008
Bit2	2010
Bit1	2013
Bit0	2015

3 结果和分析

3.1 分类精度评价

遥感图像分类中, 常用的传统算法和智能算法分别是最大似然分类(maximum likelihood classification, MLC)和反向传播(backpropagation algorithm, BP)神经网络分类(artificial neural network, ANN)两种。我们将这两种分类方法的结果与本文方法的结果进行比较(见表 2~4), 3 种算法运行中所取的训练样本和检验样本完全相同。

从表 2~4 可以看出, 3 种分类方法的总体精度都在 80%以上, 随机森林法得到更高的总体精度(分别为 91.65%, 91.89% 和 91.80%)和 Kappa 系数(分别为 0.85, 0.88 和 0.87)。在林地监测中, 一个比较大的困难是有效地分离草原与林地两种地物。草原与林地都是绿色植物, 光谱相近且相互掺杂, 传统的分类方法对草原和林地的识别结果精度大多不高。表 2~4 显示, 最大似然法和人工神经网络法草原识别结果的用户精度都不高, 均不超过 60%; 而随机森林分类法的用户精度超过 70%。林地的分类结果中, 随机森林分类的用户精度和生产者精度均超过 90%, 比最大似然法和人工神经网络法的精度更高、更稳定。总的来说, 随机森林法对本研究关注的地物(林地)有很好的识别和区分能力。

在利用遥感技术监测林地空间变化的研究中, Liknes 等^[8]利用 1 m 分辨率的航拍影像, 结合随机森林算法, 得到 84.8%的总体精度; 张春桂等^[2]利用 MODIS 数据监测林地的总体精度为 79.53%, Kappa 系数为 0.76; 朱晓玲等^[3]利用 TM 数据监测林地的精度为 84.56%, Kappa 系数为 0.81。相比之下, 本研究采用的方法有以下优势: 1) 将粗分辨率的 MODIS 数据与随机森林算法相结合, 取得更高的总体精度(91.89%)和 Kappa 系数(0.88); 2) 能够更好地区分林地与草原这两类地物; 3) 模型有很好的扩展性, 根据某年(如本研究选用的 2005 年)建立的随机森林识别模型可以应用于其他年份的监测, 避免了传统分类方法中每年重复选择感兴趣区的繁重工作。

3.2 林地面积年际变化

本研究利用 2005 年生成的随机森林模型, 对其他 7 个时相的 MODIS 数据集进行地物分类, 统计各个县域的林地面积变化情况, 结果如表 5 所示。

表 2 3 种分类方法对 2000 年 MODIS 影像的分类精度
Table 2 Accuracies of MLC, ANN and RF on MODIS images in 2000

类型	最大似然法		人工神经网络法		随机森林法	
	生产者精度/%	用户精度/%	生产者精度/%	用户精度/%	生产者精度/%	用户精度/%
林地	81.65	92.65	80.44	95.94	91.68	96.28
草原	73.36	56.23	81.45	59.69	79.12	78.58
河流	89.61	91.74	90.68	93.59	91.87	98.78
农田	89.56	92.93	92.63	91.86	96.92	92.96
总体精度/%	83.39		85.27		91.65	
Kappa	0.76		0.79		0.85	

表 3 3 种分类方法对 2005 年 MODIS 影像的分类精度
Table 3 Accuracies of MLC, ANN and RF on MODIS images in 2005

类型	最大似然法		人工神经网络法		随机森林法	
	生产者精度/%	用户精度/%	生产者精度/%	用户精度/%	生产者精度/%	用户精度/%
林地	83.76	91.65	82.44	96.74	90.70	94.30
草原	73.36	56.23	80.45	59.16	78.02	78.16
河流	89.59	90.83	91.86	94.42	91.85	98.54
农田	88.45	92.86	92.48	91.04	96.83	93.16
总体精度/%	84.95		87.20		91.89	
Kappa	0.78		0.81		0.88	

表 4 3 种分类方法对 2010 年 MODIS 影像的分类精度
Table 4 Accuracies of MLC, ANN and RF on MODIS images in 2010

类型	最大似然法		人工神经网络法		随机森林法	
	生产者精度/%	用户精度/%	生产者精度/%	用户精度/%	生产者精度/%	用户精度/%
林地	82.58	93.21	81.02	95.76	90.99	96.87
草原	74.01	56.87	81.82	59.95	79.95	78.63
河流	89.54	91.31	91.82	92.97	90.27	98.82
农田	88.96	91.98	92.89	92.01	95.46	92.88
总体精度/%	83.34		85.17		91.80	
Kappa	0.75		0.78		0.87	

康保、尚义、张北三县的林地比丰宁、围场两县的林地面积小很多。2013 年,尚义、张北两县林地面积分别减少 35.81%和 51.94%,康保县林地面积 2013 年变化较小,减少 3.71%。林地面积减少的主要原因可以归结为树龄超过生理期、连年干旱、地下水超采等^[28]。2015 年,各县的林地面积均增加,康宝、尚义两县林地面积增长幅度较大,特别是尚义县,增加 63.9%。丰宁、围场两县 2002 年林地面积减幅较大,分别为 17.23% (7.50×10^4

hm^2)和 45.24% ($2.31 \times 10^5 \text{ hm}^2$)。虽然丰宁县 2013 年林地面积只减少 10.63%,但是减少的面积却有 $5.50 \times 10^4 \text{ hm}^2$,相当于张北县 2013 年林地面积的 4.53 倍,康保县 2013 年林地面积的 9.16 倍。2015 年各县林地面积都有增加,主要原因是 2013 年林地大面积退化后,根据国务院批复的《河北张家口坝上地区退化林分改造试点实施方案(2014—2016 年)》,当地政府采取了相应措施,将退化林进行更新改造,使林地得以恢复,有的县林地面积甚至超

表 5 8 个时相的林地面积变化
Table 5 Areal changes of forestry over 8 phases

年份	康保县		尚义县		张北县	
	面积/hm ²	增长率/%	面积/hm ²	增长率/%	面积/hm ²	增长率/%
2000	1568.75		3712.50		5568.75	
2002	2200.00	40.24	3950.00	6.40	9843.75	76.77
2004	1275.00	-42.05	4825.00	22.15	10262.50	4.25
2005	2387.50	87.25	8512.50	76.42	12193.75	18.82
2008	3393.75	42.15	5862.50	-31.13	15737.50	29.06
2010	6237.50	83.79	9093.75	55.12	25256.25	60.48
2013	6006.25	-3.71	5837.50	-35.81	12137.50	-51.94
2015	8435.40	40.44	9569.26	63.90	18381.00	51.44

年份	沽源县		丰宁县		围场县	
	面积/hm ²	增长率/%	面积/hm ²	增长率/%	面积/hm ²	增长率/%
2000	20956.25		4.36×10 ⁵		5.11×10 ⁵	
2002	21643.75	3.28	3.61×10 ⁵	-17.23	2.8×10 ⁵	-45.24
2004	31725.00	46.58	4.95×10 ⁵	36.98	4.04×10 ⁵	44.53
2005	31862.50	0.43	5.01×10 ⁵	1.30	5.21×10 ⁵	28.83
2008	38250.00	20.04	5.16×10 ⁵	2.86	5.19×10 ⁵	-0.30
2010	41031.25	7.27	5.11×10 ⁵	-0.86	4.39×10 ⁵	-15.43
2013	36575.00	-10.96	4.56×10 ⁵	-10.63	4.78×10 ⁵	8.90
2015	39435.40	7.80	4.86×10 ⁵	6.50	5.13×10 ⁵	7.30

过 2010 年。

3.3 林地空间分布年际变化

坝上林区林地 2000—2015 年 8 个时相的空间变化如图 3 所示, 其中 2002 年的变化最显著, 沽源、丰宁、围场三县大面积林地的消失, 康保、张北、尚义三县草地和林地零星状地消失。随后几年, 坝上的林地得以恢复。

图 3 显示, 2008—2015 年坝上林地面积的空间分布变化不大。我们利用 2.3 节的方法, 得到 8 个时相的林地面积变化信息图像, 对该图像中的灰度值(编码值)进行统计后得到图 4, 直观地反映了 8 个时相的林地面积变化轨迹。图 4 中编码值 255 对应的像元数最多, 为 82424 个, 255 的二进制数为 11111111, 表示 8 个时相都为林地。像元数最多的编码值有 2, 4, 8 和 191, 分别占林地总像元数的 0.34%, 0.39%, 0.41%和 0.67%。其中, 2 表示 2013 年扩展的林地面积为 8639 个像元(53993.75 hm²), 4 表示 2010 年扩展的林地面积为 9968 个像元(62300 hm²), 8 表示 2008 年扩展的林地面积为 10504 个像元(65650 hm²), 191 表示 2002 年消失的林地面积有

17051 个像元(106568.75 hm²)。像元数次多的编码值有 16, 128, 159 和 253, 分别占林地总像元数的 0.23%, 0.19%, 0.18%和 0.15%。其中, 16 表示 2005 年扩展的林地面积为 6808 个像元(11300 hm²), 128 表示 2002 年消失的林地面积为 4702 个像元(29387.5 hm²), 159 表示 2005 年扩展的林地面积为 4700 个像元(29375 hm²), 253 表示 2013 年消失的林地面积为 4055 个像元(2534.34 hm²)。图 4 还显示, 编码值在 1~14 之间时, 曲线起伏较多, 表示 2008—2015 年林地面积变化频繁。随后, 除在个别编码值(如 16, 128, 159)处曲线变化较大外, 曲线较平坦。编码值为 191 处的峰值主要是 2002 年围场大面积林地消失导致的。

本研究对 8 个时相坝上林地的变化图进行假彩色密度分割, 得到图 5。8 个时相均为林地的像元设为绿色, 为无变化区; 1~3 个时相林地像元消失的设为黄色, 为消失中度区; 4~7 个时相林地像元消失的标记为红色, 为严重消失区; 8 个时相都与林地无关的区域为不相关区。

图 5 显示, 康保县和尚义县林地的消失呈散点

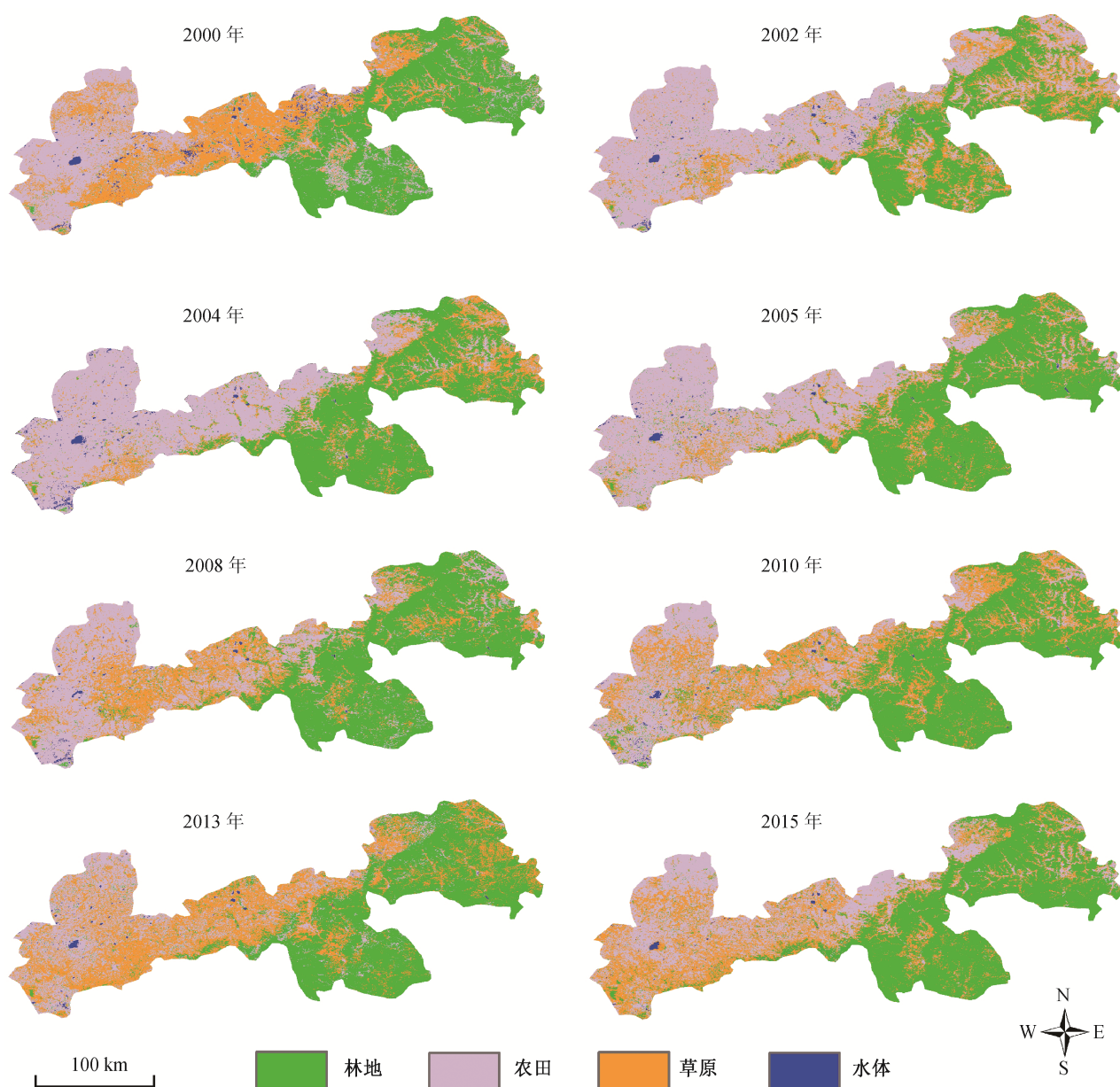


图 3 8 个时相的林地空间变化
Fig. 3 Forest area spatial changes over 8 phases

状, 主要因为这两个县以农田防护林为主, 夹杂灌木等, 由于遥感影像的分辨率和混合像元问题, 使得监测效果有一定的局限性^[28]。张北、沽源、丰宁和围场四县出现条带状的林地消失图像, 主要分布在城市和农田周围。这种林地消失较严重的现象与这些年份的干旱和地下水位下降有关, 也与造林质量不高(当时多采用扦插造林的方法, 树体生命力较弱)和后续抚育经营不到位有关。

4 结论

本文利用 250 m 空间分辨率的 MODIS 影像, 用 TM 影像辅助进行训练样本选择, 采用随机森林算法监测河北坝上林区 2000—2015 年 8 个时相的林地时空变化, 得到以下结论: 1) 随机森林分类方法比传统的分类方法具有更高的精度, 可作为林地动态监测的一种有效手段; 2) 利用 2005 年数据建

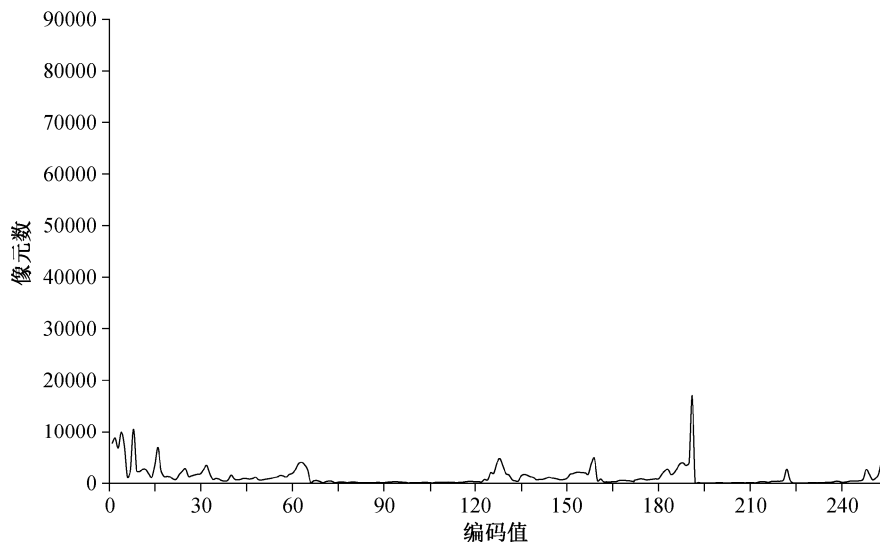


图 4 8 个时相林地面积随时间的变化

Fig. 4 Histogram of forestry area changes over 8 phases

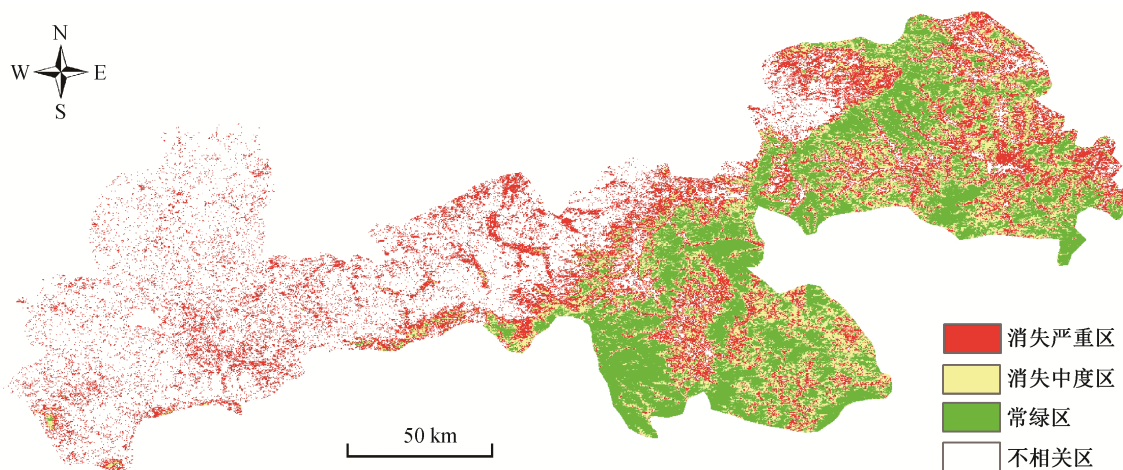


图 5 林地消失程度的空间分布

Fig. 5 Extent of forest spatial changes

立的随机森林识别模型可用于其他年份林地的动态变化监测，能有效地减少工作量；3) 坝上林地退化严重的地区集中在丰宁、围场、张北和沽源四县，退化的时间集中在 2002, 2010 和 2013 年。

250 m 空间分辨率的 MODIS 数据对小面积林地的监测有局限性，其影像上林地与草原的混合像元易造成分类错误，使得利用 MODIS 数据的林地监测结果不能提供精确的林地面积变化数据，但其揭示的林地面积变化趋势对林地管理有参考价值。

参考文献

[1] Peng Dailiang, Zhang Bing, Liu Liangyuan, et al.

Characteristics and drivers of global NDVI-based FPAR from 1982 to 2006. *Global Biogeochemical Cycles*, 2012, 26(3): 1–15

[2] 张春桂, 朱晓铃, 陈惠, 等. 福建沿海防护林资源的卫星遥感监测. *中国农业气象*, 2007, 28(4): 446–449

[3] 朱晓玲, 汪小钦, 陈芸芝. 漳州市沿海防护林变化的遥感动态监测. *遥感技术与应用*, 2005, 20(2): 243–246

[4] Deng Rongxin, Wang Wenjuan, Li Ying, et al. A retrieval and validation method for shelterbelt vegetation fraction. *Journal of Forestry Research*, 2013, 24 (2): 357–360

- [5] Wiseman G, Kort J, Walker D. Quantification of shelterbelt characteristics using high-resolution imagery. *Agriculture, Ecosystem and Environment*, 2009, 131: 111–117
- [6] Jin S, Sader S A. MODIS time-series imagery for forest disturbance detection and quantification of patch size effects. *Remote Sensing of Environment*, 2005, 99(4): 462–470
- [7] Ranson K J, Kovacs K, Sun G, et al. Disturbance recognition in the boreal forest using radar and Landsat-7. *Canadian Journal of Remote Sensing*, 2003, 29(2): 271–285
- [8] Liknes G C, Perry C H, Meneguzzo D M. Assessing tree cover in agricultural landscapes using high-resolution aerial imagery. *The Journal of Terrestrial Observation*, 2010, 2(1): 38–55
- [9] Nelmes S, Belcher R E, Wood C J. A method for routine characterization of shelterbelts. *Agriculture, Ecosystem and Environment*, 2001, 106: 303–315
- [10] Pal M. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 2005, 26(1): 217–222
- [11] Rodriguez-Galiano V F, Ghimire B, Rogan J, et al. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2012, 67: 93–104
- [12] Gislason P O, Benediktsson J A, Sveinsson J R. Random Forests for land cover classification. *Pattern Recognition Letters*, 2006, 27: 294–300
- [13] 雷震. 随机森林及其在遥感影像处理中的应用研究[D]. 上海: 上海交通大学, 2012
- [14] 王强. 基于 GIS 的张家口坝上地区农业分区研究[D]. 河北: 河北师范大学, 2008
- [15] 李玄珠, 常春平, 王仁德. 河北坝上土地利用方式对农田土壤风蚀的影响. *中国沙漠*, 2014, 34(1): 23–28
- [16] 赵雪, 李进, 赵文智, 等. 河北坝上地区植被及其资源利用研究——以丰宁试验区为例. *中国沙漠*, 1994, 14(4): 29–36
- [17] 孙红艳. 河北省坝上土地荒漠化机制及生态环境评价[D]. 北京: 中国地质大学, 2005
- [18] Breiman L. Random forests. *Machine Learning*, 2001, 45(1): 5–32
- [19] Schwende H, Zucknick M, Ickstadt K, et al. A pilot study on the application of statistical classification procedures to molecular epidemiological data. *Toxicology Letters*, 2004, 151: 291–299
- [20] Díaz-Uriarte R, Alvarez de Andrés S. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, 2006, 7: 1–13
- [21] Timm B C, McGarigal K. Fine-scale remotely-sensed cover mapping of coastal dune and salt marsh ecosystems at Cape Cod National Seashore using Random Forests. *Remote Sensing of Environment*, 2012, 127: 106–117
- [22] Chan J C W, Paelinckx D. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment*, 2008, 112: 2999–3011
- [23] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述. *统计与信息论坛*, 2011, 26(3): 32–38
- [24] Janssen L L F, Van Der Wel F J M. Accuracy assessment of satellite derived land-cover data: a review. *Photogrammetric Engineering & Remote Sensing*, 1994, 60(4): 419–426
- [25] Congalton R, Green K. Assessing the accuracy of remotely sensed data: principles and practices. Boca Raton: CRC/Lewis Press, 1999
- [26] 陈佳祥. 高分辨率遥感影像草地和树木分类方法研究[D]. 厦门: 集美大学, 2011
- [27] 李天宏, 赵智杰, 韩鹏. 深圳河河口红树林变化的多时相遥感分析. *遥感学报*, 2002, 6(5): 364–370
- [28] 孙雷刚, 刘剑锋, 徐全洪. 河北省坝上地区植被覆盖变化遥感时空分析. *国土资源遥感*, 2014, 26(1): 167–172