

基于克隆选择聚类改进的植被变化回归趋势分析*

王佃来^{1,2} 刘文萍¹

(1. 北京林业大学信息学院, 北京 100083; 2. 首钢工学院计算机系, 北京 100144)

摘要: 基于克隆选择聚类算法对植被变化趋势回归分析中线性方程的斜率 k 进行聚类, 并按聚类中心大小分类, 解决了对斜率 k 的划分问题。实验结果表明: 所提算法的 DBI 值 0.65 明显小于与之对比 3 种算法的 DBI 值 0.89、0.91 和 0.89。北京市城区植被变化趋势显著增强的结论与 1998~2011 年北京市城区绿化率和人均绿化面积稳步增长的统计数据一致, 验证了算法的正确性, 因此该方法有一定的实用性, 可应用于长时序植被变化趋势分析。

关键词: 植被变化监测 克隆选择算法 线性回归分析 聚类分析

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1000-1298(2013)12-0215-06

Improved Linear Trend Analysis in Vegetation Cover Change Based on Clonal Selection Clustering Algorithm

Wang Dianlai^{1,2} Liu Wenping¹

(1. College of Information, Beijing Forestry University, Beijing 100083, China

2. Department of Computer Science, Shougang Institute of Technology, Beijing 100144, China)

Abstract: A clonal selection clustering algorithm was proposed to solve classification problem of the slope produced by the linear trend analysis. The experimental results showed that the DBI value of 0.65 was significantly less than the compared three algorithms. The result of vegetation change trend was in accord with the statistic data of greenbelt for everyone and green coverage percentage in Beijing from 1998 to 2011. So the proposed method had some application value and could widely apply in long time series of vegetation change trend analysis.

Key words: Trend monitor in vegetation cover Clonal selection algorithm Linear regression analysis Clustering analysis

引言

植被是土地覆盖的最主要部分, 其变化对全球能量循环及物质的生物化学循环具有重要的影响^[1]。长时间序列植被监测和评价是研究植被生长和受影响特征的核心领域, 也是研究陆地生态系统过程和全球变化的重要环节^[2]。遥感技术的发展为植被长时序监测提供了数据基础。

回归分析法是研究植被长时序变化趋势的重要方法^[2], 能定量的反映多年植被变化趋势。国内外

许多学者使用该方法进行植被变化趋势研究^[3-16]。

文献[10,12~15]和文献[11]使用阈值法将线性回归拟合生成的斜率 k 分别划分为 7 类和 3 类。但是, 上述文献均未明确指出阈值选取的依据和方法, 且划分斜率 k 的阈值各不相同, 无统一规律。文献[1]以逐像元对应的平均 NDVI 值与年份进行回归分析, 得到 Pearson 系数 (R), 如果 R 值通过 0.05 的显著水平, 则认为植被变化趋势为显著增加或显著减少, 否则认为增加或减少不显著。文献[16]基于线性显著性 t -检验将植被变化趋势分为 4 类。

收稿日期: 2013-07-02 修回日期: 2013-08-09

* 国家重点基础研究发展计划(973 计划)资助项目(2009CB421105)和中央高校基本科研业务费专项资金资助项目(YX2011-28)

作者简介: 王佃来, 博士生, 首钢工学院高级工程师, 主要从事数字图像处理和软件设计研究, E-mail: wangdll2345@126.com

通讯作者: 刘文萍, 教授, 博士生导师, 主要从事数字图像处理及模式识别研究, E-mail: wendyl@vip.163.com

由上述分析可见,如何合理对回归分析产生的斜率 k 分类是回归植被变化趋势分析法的重要问题。阈值法对斜率 k 的分类存在人为因素的差异。线性显著性检验有较成熟的统计学理论依据,但仅将 k 划分为3类、4类或5类,在更细致的划分应用中受到限制,如文献[10,12~15]中的划分为7类。为了解决上述问题,本文采用克隆选择算法对回归分析产生的斜率 k 聚类,按聚类中心的大小排序实现分类,并对聚类结果进行评估和验证。

1 方法及数据

1.1 线性回归趋势分析法

回归分析是研究多个变量之间统计联系的一种重要方法。对一组时间自变量 x 与 NDVI 因变量 y 数据,可描述为

$$y = a + kx + \varepsilon \quad (1)$$

式中 a, k ——未知常数 ε ——随机误差

利用观测值 (x_i, y_i) 可以求出未知参数

$$k = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2)$$

$(i = 1, 2, \dots, n)$

其中 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$

对于 NDVI 长时序数据,采用最小二乘法线性拟合后得到相应的线性方程,方程的斜率 k 说明像元 NDVI 值的多年度变化趋势^[3-5],当 $k > 0$ 时,植被活动增强,当 $k < 0$ 时,植被活动减弱。

1.2 基于克隆选择聚类的植被变化回归趋势分析

1.2.1 克隆选择算法原理

1958年,Burnet等提出了著名的免疫克隆选择学说^[17]。克隆选择原理认为,淋巴细胞除了扩增或者分化成浆细胞以外,也能分化成生命期较长的B记忆细胞。当再次遇到相应的抗原时,记忆细胞将预先被免疫系统选择出来,并迅速活化、增殖、分化为抗体生成细胞,执行高效而持久的免疫功能。

受克隆选择学说启发,de Castro等提出了一种克隆选择算法(Clonal selection algorithm, CSA)^[18],它借助于免疫系统的抗体克隆选择机理,构造适用于人工智能的克隆算子。基于克隆算子的克隆选择算法是一种群体搜索策略,具有并行性和搜索变化的随机性,在搜索中不易陷入局部最优值,能以较大的概率获得问题的全局最优解,且具有较快的收敛速度。

生物免疫系统中的克隆选择原理,描述了免疫

系统对抗原激励做出免疫响应的基本特性^[19]。在基于克隆选择原理的免疫算法中,抗原对应于问题的目标函数,抗体对应于目标函数的优化解,先根据抗体的亲和力对解评价和选择,然后通过记忆细胞保留局部最优解,以保持解的多样性,再次用类似于抗体的亲和力来逐步改善优化过程,最终得到问题的全局最优解。

抗体的亲和力是根据抗原与抗体的亲和力、抗体之间的亲和力进行评价的,若某抗体与抗原之间的亲和力越大,且与其他抗体之间的亲和力越小,则该抗体的亲和力就越大。这种适应值评价方式能保持个体的多样性,提高算法在局部解空间的搜索效率,并能有效摆脱局部最优点,但该适应值的函数表达式难以确定,往往需要通过反复试探。

1.2.2 抗体编码

考虑到算法聚类对象是由回归分析法计算得到的直线的斜率 k ,该值为浮点型数值,所以聚类中心采用实数编码方式,每个抗体 p_i 由 k 个聚类中心组成,它可表示为长度为 $k \times d$ 的实数编码。

$$p_i = \left\{ \frac{p_{i1}, p_{i2}, \dots, p_{id}}{c_1}, \frac{p_{i1}, p_{i2}, \dots, p_{id}}{c_2}, \dots, \frac{p_{i1}, p_{i2}, \dots, p_{id}}{c_k} \right\} \quad (3)$$

式中, c_1, c_2, \dots, c_k 所对应的编码分别为各聚类中心在样本空间中的坐标。

1.2.3 抗体亲和力定义

基于人工免疫系统的聚类算法,通过优化聚类有效性评价函数,搜索与评价函数最优值所对应的聚类中心来获得数据集的最优划分。在聚类研究领域,针对不同类型的数据集,已经提出了多种类型的聚类有效性评价函数。本文采用简单误差平方和准则来定义抗体的亲和力。

抗体解码为对应的 k 个聚类中心,以矩阵 V 的形式表示,则对应于抗体 p_i 的聚类划分的误差平方和准则函数

$$J(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m D^2(c_i, x_j) \quad (4)$$

其中 $U = [u_{ij}]_{k \times n}$ 是模糊指数为 m 隶属度矩阵, $D^2(c_i, x_j)$ 表示 c_i 与 x_j 的欧氏距离,最小值 $J_{\min} = \min\{J(U, V)\}$,所确定的聚类中心对应于最优的聚类划分。基于误差平方和准则的抗体亲和力函数为

$$f(p_i) = \frac{1}{1 + J} \quad (5)$$

1.2.4 个体克隆方式

个体的克隆方式遵循以下原则:如果抗体的浓度低于设定的阈值,其增殖规模较大,否则增殖规模较小,具体定义为

$$N_c = \begin{cases} F(k_1 f(i)) & (C(i) > T_s) \\ F(k_2 f(i)) & (\text{其他}) \end{cases} \quad (6)$$

其中 $C(i) = \frac{N_n}{N} \quad k_1 = 0.25k_2$

式中 $f(i)$ ——个体的亲和度函数

$C(i)$ ——个体浓度 $F(x)$ ——取整函数

N_n ——个体 x_i 的邻域内的个体数目, 邻域

$\delta(i) = \{s(x_i, x_j) \leq T_d\}$

s ——个体之间的相似性度量函数

T_d ——相似度阈值 N ——种群规模

k_1, k_2 ——个体浓度大于、小于指定阈值时的比例系数阈值

N_c ——个体克隆个数

T_s ——个体浓度阈值

1.2.5 克隆个体的变异

每个克隆增殖产生的新个体, 其每一位以较大的变异概率产生随机变异。采用变异范围内均匀分布的随机数代替原值, 则变异范围与个体的亲和度值呈反比。

1.2.6 算法主要步骤

(1) 计算北京市地区遥感影像中每一个像素点的 NDVI 值 y 与年份 x 进行线性拟合, 得到直线方程的斜率 k , 对斜率 k 进行步骤(2)~(8)的操作。

(2) 随机初始化抗体群 P , 设定模糊隶属度的加权指数 m , 指定聚类数目 c , 定义样本之间的距离度量函数 D_{ik} , 设定算法的终止条件。

(3) 对种群 P 中的每一个体实施一步 FCM 操作, 形成新种群 P_1 。

(4) 从种群 P_1 中按一定比例选择亲和度高的个体, 再基于式(6)复制产生一定数目的新个体, 形成新种群并记为 P_2 。

(5) 对种群 P_2 中的每一个体采用高变异率的位变异操作, 形成种群 P_3 。

(6) 计算 P_3 中每一个体的亲和度和浓度, 用所提免疫选择策略获取更新的记忆细胞种群 M 。

(7) 用种群 M 和一定数目、随机产生的新个体, 按一定比例替换种群 P 中亲和度低的个体, 形成新一代种群。

(8) 若满足结束条件, 算法终止, 否则转到步骤(3)继续循环。

1.3 实验数据与实验参数

1.3.1 实验数据

实验数据包括两部分: ①1998年4月~2011年12月空间分辨率为 1 km 的 SPOT-4 VGT S10 Normalized Difference Vegetation Index (NDVI) 数据, 该数据下载于 <http://free.vgt.vito.be/>, 数据区域为

东南亚(SE-Asia)。为保证数据的精确性, 进行以下预处理: 首先使用 VGT Extract 软件从下载的东南亚数据根据北京市地区经纬度范围大致裁剪出北京市地区数据, 然后结合北京市地区行政区划矢量图使用 ENVI 软件的掩模功能裁剪出北京市地区的精确遥感数据。②北京市行政区划矢量图。以上两部分数据投影格式为 Geographic Lon/Lat。

裁剪后的北京市地区遥感数据经过归一化处理, 将 NDVI 值和年份归一化到 [0, 1] 之间, NDVI 值取每年对应像素点的最大值, 也就是每年植被最茂盛时期的 NDVI 值, NDVI 值作为因变量 y , 年份作为自变量 x , 两者组成时间序列, 使用最小二乘法进行线性拟合, 得到相应直线的斜率 k , 然后对 k 进行聚类。

1.3.2 实验参数

克隆选择算法参数设置: $k_1 = 1, k_2 = 4, T_s = 0.1, T_d = 0.5$, 变异率 $P_m = 0.3$, 选择概率 $P_s = 0.3$, 记忆细胞在种群中所占比例为 45%, 种群的淘汰率为 40%, 最大迭代次数为 50 次; 模糊聚类算法中 $m = 2.5$ 。

2 实验结果与分析

文中算法均由 C++ 实现, 测试机的 CPU 为 P4 2.66 GHz, 内存为 2 GB, 操作系统为 Windows XP, 使用开发工具为 VC++ 6.0, 遥感数据处理软件 ENVI 4.5。

为了使结果有较好的可比性, 回归分析中的斜率 k 统一分为 4 类, 按照聚类中心从小到大的顺序植被变化趋势依次为严重退化、轻微退化、轻微改善和明显改善。

为了验证文中提出算法的有效性, 本文采用 DBI^[20] (Davies-Bouldin index) 参数及运算速度 2 种测度分别与 Otsu、K-means 和 FCM 聚类算法进行了定量比较。

表 1 是对 Otsu 法、K-means、FCM 和克隆选择聚类算法各项指标比较的汇总表。表 1 可以看出: 以 DBI 参数为测度, 本文提出算法的 DBI 值最小, 并且与比较算法的 DBI 值差异明显, 因此本文提出的算法优于比较算法。3 种比较算法的 DBI 值很接近, 最大差值仅为 0.02。其中 K-means 的 DBI 最大。产生该结果的主要原因是克隆选择聚类算法具有全局搜索能力, 通过初始种群的多样化防止聚类过程收敛于局部最优解, 并且有较好的抗噪声能力。

表 2 是 Otsu 法、K-means、FCM 和克隆选择聚类算法运行速度的比较。其中 K-means、FCM 和克隆选择聚类算法的最大迭代次数均为 50 次。从表 2

表 1 克隆选择聚类与 Otsu、K-means 和 FCM 结果比较

Tab.1 Comparison results of Otsu, K-means, FCM and clonal selection clustering algorithm

算法		严重退化	轻微退化	轻微改善	明显改善
Otsu	聚类中心	-0.176 8	-0.042 2	0.049 7	0.125 6
	类内方差	0.058 9	0.030 4	0.023 9	0.060 1
	DBI		0.89		
K-means	聚类中心	-0.154 5	-0.023 9	0.053	0.122 3
	类内方差	0.061 8	0.027 8	0.019 4	0.058 6
	DBI		0.91		
FCM	聚类中心	-0.160 62	-0.030 93	0.047 6	0.116 7
	类内方差	0.061 0	0.027 8	0.019 7	0.056 2
	DBI		0.89		
克隆选择聚类	聚类中心	-0.135 6	0.007 7	0.090 4	0.384 2
	类内方差	0.064 0	0.030 7	0.030 0	0.114 3
	DBI		0.65		

表 2 各算法运行时间比较

Tab.2 Comparison of running time for Otsu, K-means, FCM and clonal selection algorithm ms

算法	运行时间
Otsu	3 670.0
K-means	4 969.0
FCM	7 281.0
克隆选择聚类	13 625.0

可以看出克隆选择聚类算法运行时间最长。引起这种现象的主要原因是该算法需要在解空间中反复探测最优解,而增加了算法运行的时间开销。其他 3 种算法没有该步骤所以运行速度较快。

图 1 ~ 4 分别为使用自动阈值法 (Otsu)、K-means、FCM 和克隆选择聚类算法对线性回归法产生的斜率 k 进行分类得到的植被变化趋势图与北京市行政区划图叠加的结果。分析图 1 ~ 4 可以看出: Otsu、FCM 和克隆聚类算法产生的植被变化趋势图差别不大;肉眼较难看出明显差异。但是 K-means 生成的植被变化趋势图与上述三者差异明显,特别是在植被明显改善区域有很大的差别。从图 4 中可以看出北京城区呈深绿色,也就是说北京城区的植被变化趋势是明显改善,而植被严重退化(褐

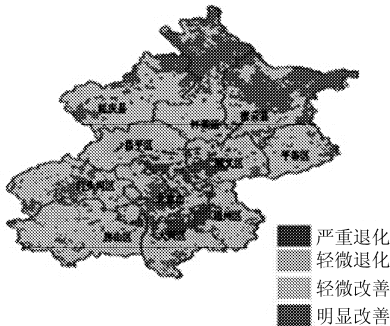


图 1 自动阈值法 (Otsu) 植被变化趋势图

Fig.1 Diagram of vegetation cover change trend made by Otsu

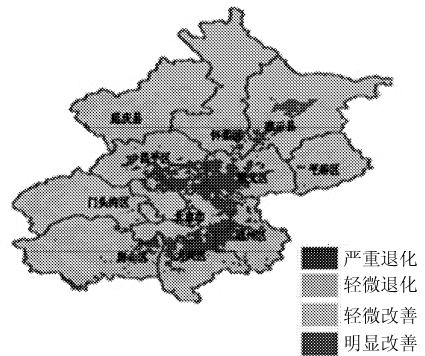


图 2 K-means 植被变化趋势图

Fig.2 Diagram of vegetation cover change trend made by K-means

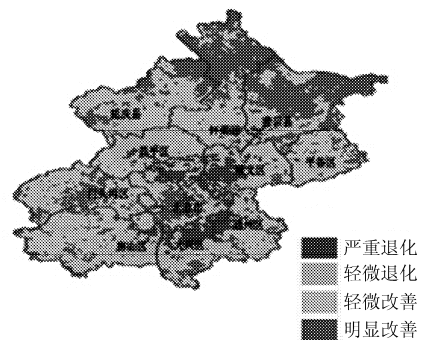


图 3 FCM 植被变化趋势图

Fig.3 Diagram of vegetation cover change trend made by FCM

色)或轻微退化(浅褐色)的区域以马蹄形环绕北京市城区。

由于时间跨度长和验证数据的不完整等原因,长时间序列植被覆盖变化趋势验证是个难题。为了客观地评价本文所提算法与实际北京市植被覆盖变化趋势结果的符合程度,本文采取定性分析的方法,对本文算法生成的植被覆盖变化趋势分析结果进行验证。表 3 是北京市 1998 ~ 2011 年城区绿化覆盖率 and 人均绿地面积。从表 3 中可以看出,城区的覆

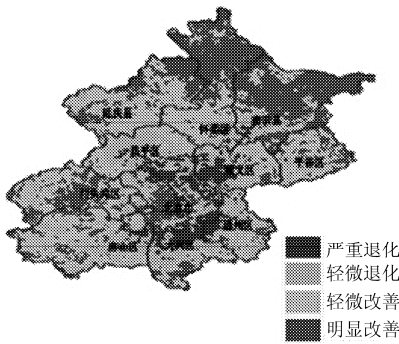


图 4 克隆选择植被变化趋势图

Fig. 4 Diagram of vegetation cover change trend made by clonal selection algorithm

盖率和人均绿地面积是逐年增加的。图 5 是 1998 ~ 2011 年北京市城区绿化率的变化图。从图 5 中可以看出,近 14 年北京市城区绿化率接近直线增长,

表 3 1998 ~ 2011 年北京市城区植被覆盖率和人均绿地面积

Tab. 3 Statistic data of greenbelt for everyone and green coverage percentage in Beijing from 1998 to 2011

年份	人均绿地面积/ m^2	植被覆盖率/%
1998	9.00	35.60
1999	9.10	36.30
2000	9.66	36.50
2001	10.07	38.78
2002	10.66	40.57
2003	11.43	40.87
2004	11.45	41.91
2005	12.00	42.00
2006	12.00	42.50
2007	12.60	43.00
2008	13.60	43.50
2009	14.50	44.40
2010	15.00	45.00
2011	15.30	45.60

决定系数 $R^2 = 0.95$, 超过显著水平 $\alpha = 0.05$ 时的值 0.532 4, 因此北京市城区植被覆盖变化趋势是显著增强的。这与本文算法生成的植被覆盖变化趋势图(图 4)中北京市城区植被覆盖变化趋势是明显改善的结论一致,从而间接地证实了本文算法的正确性。

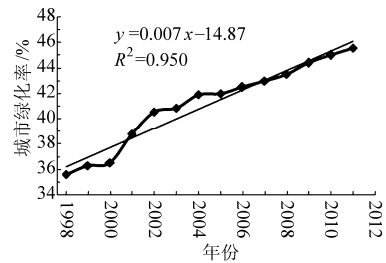


图 5 1998 ~ 2011 年北京市城区绿化率变化

Fig. 5 Change map of greening rate in urban area of Beijing from 1998 to 2011

3 结束语

将克隆选择聚类算法应用于解决回归分析植被趋势方法生成斜率 k 的分类问题上,在不需人工干预的情况下实现了自动化植被变化趋势分析。实验结果表明,该算法通过初始化种群的多样性使算法有较强的全局搜索能力,能很好避免聚类过程中陷入局部最优解的问题,提高了聚类的稳定性及可靠性。其 DBI 明显小于与之对比的 3 个算法(Otsu、K-means和 FCM),聚类效果优于比较算法。北京市城区植被变化趋势显著增强的实验结果与北京市 1998 ~ 2011 年城区植被覆盖率和人均绿地面积线性增长的统计数据一致,进而验证了算法的正确性。综上所述,该算法有一定的实用性,可以应用于长时序植被变化趋势分析。值得注意的是该算法存在算法的时间复杂度高、在大数据量计算时耗时较长的不足。

参 考 文 献

- 1 朴世龙,方精云.最近 18 年来中国植被覆盖的动态变化[J].第四纪研究,2001,21(4):294~302.
Piao Shilong, Fang Jingyun. Dynamic vegetation cover change over the last 18 years in China[J]. Quaternary Sciences, 2001, 21(4):294~302. (in Chinese)
- 2 蔡博峰,于嵘.基于遥感的植被长时序趋势特征研究进展及评价[J].遥感学报,2009,13(6):1177~1186.
Cai Bofeng, Yu Rong. Advance and evaluation in the long time series vegetation trends research based on remote sensing[J]. Journal of Remote Sensing, 2009, 13(6): 1177~1186. (in Chinese)
- 3 Weiss E, Marsh S E, Priman E S. Application of NOAA-AVHRR NDVI time-series data to assess changes in Saudi Arabia's rangelands[J]. International Journal of Remote Sensing, 2001,22(6): 1005~1027.
- 4 Fuller D O. Trends in NDVI time series and their relation to rangeland and crop production in Senegal, 1987 ~ 1993 [J]. International Journal of Remote Sensing, 1998, 19(10): 2013~2018.
- 5 Camberlin P, Martiny N, Philippon N, et al. Determinants of the interannual relationships between remote sensed photosynthetic activity and rainfall in tropical Africa[J]. Remote Sensing of Environment, 2007,106(2): 199~216.
- 6 Rigina O, Rasmussen M S. Using trend line and principal component analysis to study vegetation changes in Senegal 1986 ~ 1999 from AVHRR NDVI 8km data[J]. Geografisk Tidsskrift, 1996,103(1): 31.
- 7 Tottrup C, Rasmussen M S. Mapping long-term changes in savannah crop productivity in Senegal through trend analysis of time series of remote sensing data[J]. Agriculture, Ecosystems & Environment, 2004, 103(3): 545~560.
- 8 Herrmann S M, Anyamba A, Tucker C J. Recent trends in vegetation dynamics in the African Sahel and their relationship to

- climate[J]. *Global Environmental Change Part A*, 2005, 15(4): 394~404.
- 9 Evans J P, Geerken R. Discrimination between climate and human-induced dryland degradation[J]. *Journal of Arid*, 2004, 57(4): 535~554.
- 10 宋怡,马明国. 基于 SPOT VEGETATION 数据的中国西北植被覆盖变化分析[J]. *中国沙漠*, 2007, 27(1): 89~93.
Song Yi, Ma Mingguo. Study on vegetation cover change in Northwest China based on SPOT VEGETATION data[J]. *Journal of Desert Research*, 2007, 27(1): 89~93. (in Chinese)
- 11 韩秀珍,李三妹,罗敬宁,等. 近 20 年中国植被时空变化研究[J]. *干旱区研究*, 2008, 25(6): 753~759.
Han Xiuzhen, Li Sanmei, Luo Jingning, et al. Study on spatiotemporal change of vegetation in China since 20 years [J]. *ARD Zone Research*, 2008, 25(6): 753~759. (in Chinese)
- 12 邱海军,曹明明. 基于 SPOT VEGETATION 数据的中国植被覆盖时空变化分析[J]. *资源科学*, 2011, 33(2): 335~340.
Qiu Haijun, Cao Mingming. Spatial and temporal variations in vegetation cover in China based on SPOT VEGETATION data [J]. *Resources Science*, 2011, 33(2): 335~340. (in Chinese)
- 13 陈燕丽,龙步菊,潘学标,等. MODIS NDVI 和 AVHRR NDVI 对草原植被变化监测差异[J]. *遥感学报*, 2011, 15(4): 831~845.
Chen Yanli, Long Bujie, Pan Xuebiao, et al. Differences between MODIS NDVI and AVHRR NDVI in monitoring grasslands change[J]. *Journal of Remote Sensing*, 2011, 15(4): 831~845. (in Chinese)
- 14 贺振,贺俊平. 基于 SPOT-VGT 的黄河流域植被覆盖时空演变[J]. *生态环境学报*, 2012, 21(10): 1655~1659.
He Zhen, He Junping. Spatio-temporal variation of vegetation cover based on SPOT-VGT in Yellow River Basin [J]. *Ecology and Environmental Sciences*, 2012, 21(10): 1655~1659. (in Chinese)
- 15 李双双,延军平,万佳. 近 10 年陕甘宁黄土高原区植被覆盖时空变化特征[J]. *地理学报*, 2012, 67(7): 960~970.
Li Shuangshuang, Yan Junping, Wan Jia. The spatial-temporal changes of vegetation restoration on Loess Plateau in Shaanxi-Gansu-Ningxia region[J]. *Acta Geographica Sinica*, 2012, 67(7): 960~970. (in Chinese)
- 16 孙晓鹏,王天明,葛剑平. 基于 MODIS 的北京绿化隔离地区植被格局与趋势分析[J]. *地理与地理信息科学*, 2012, 28(6): 20~23.
Sun Xiaopeng, Wang Tianming, Ge Jianping. Pattern and trend analysis of the isolated afforestation regions in Beijing using MODIS [J]. *Geography and Geo-Information Science*, 2012, 28(6): 20~23. (in Chinese)
- 17 Burnet M F. A modification of Jerne's theory of antibody production using the concept of clonal selection[J]. *Austrian Journal of Science*, 1957, 20(1): 67~76.
- 18 de Castro L N, von Zuben F J. Artificial immune systems: part I—basic theory and applications [R]. RT-DCA, 1999.
- 19 de Castro L N, von Zuben F J. Learning and optimization using the clonal selection principle [J]. *IEEE Transactions on Evolutionary Computation*, 2002, 6(3): 239~251.
- 20 Davies D I, Bouldin D W. A cluster separation measure [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1979, 1(2): 224~227.

(上接第 179 页)

- 23 González A, Llorens A, Cervera M L, et al. Elemental fingerprint of wines from the protected designation of origin Valencia [J]. *Food Chemistry*, 2009, 112(1): 26~34.
- 24 Makris D P, Kallithraka S, Mamalos A. Differentiation of young red wines based on cultivar and geographical origin with application of chemometrics of principal polyphenolic constituents [J]. *Talanta*, 2006, 70(5): 1143~1152.
- 25 Almeida C M, Vasconcelos M T S D. ICP-MS determination of strontium isotope ratio in wine in order to be used as a fingerprint of its regional origin [J]. *Journal of Analytical Atomic Spectrometry*, 2001, 16(6): 607~611.
- 26 Ferreira V, Fernández P, Cacho J F. A study of factors affecting wine volatile composition and its application in discriminant analysis [J]. *LWT-Food Science and Technology*, 1996, 29(3): 251~259.
- 27 López R, Ferreira V, Hernández P, et al. Identification of impact odorants of young red wines made with Merlot, Cabernet Sauvignon and Grenache grape varieties: a comparative study [J]. *Journal of the Science of Food and Agriculture*, 1999, 79(11): 1461~1467.
- 28 Rosillo L, Salinas M R, Garijo J, et al. Study of volatiles in grapes by dynamic headspace analysis: application to the differentiation of some Vitis Vinifera varieties [J]. *Journal of Chromatography A*, 1999, 847(1): 155~159.
- 29 Isabel Spranger M, Cristina Clímaco M, Sun B, et al. Differentiation of red winemaking technologies by phenolic and volatile composition [J]. *Analytica Chimica Acta*, 2004, 513(1): 151~161.
- 30 Aznar M, López R, Cacho J F, et al. Identification and quantification of impact odorants of aged red wines from Rioja. GC-Olfactometry, quantitative GC-MS, and odor evaluation of HPLC fractions [J]. *Journal of Agricultural and Food Chemistry*, 2001, 49(6): 2924~2929.
- 31 Culleré L, Cacho J, Ferreira V. Analysis for wine C₅-C₈ aldehydes through the determination of their O-(2,3,4,5,6-pentafluorobenzyl)oximes formed directly in the solid phase extraction cartridge [J]. *Analytica Chimica Acta*, 2004, 524(1~2): 201~206.
- 32 Nasi A, Ferranti P, Amato S, et al. Identification of free and bound volatile compounds as typicalness and authenticity markers of non-aromatic grapes and wines through a combined use of mass spectrometric techniques [J]. *Food Chemistry*, 2008, 110(3): 762~768.
- 33 Lee S J, Noble A C. Characterization of odor-active compounds in Californian Chardonnay wines using GC-Olfactometry and GC-MS spectrometry [J]. *Journal of Agricultural and Food Chemistry*, 2003, 51(27): 8036~8044.