

文章编号: 1007-4619 (2002) 01-0030-05

# 遥感植被指数对多时相 AVHRR 数据 主成分分析的影响

延 昊, 王长耀, 牛 铮, 姜小光, 王 汶

(中国科学院 遥感应用研究所遥感信息科学开放实验室, 北京 100101)

摘 要: 对中国全年 36 个旬 NOAA-AVHRR 的 1 km 覆盖数据进行两步处理: 分别采用比值植被指数  $RVI$ 、归一化植被指数  $NDVI$ 、土壤调整植被指数  $SAVI$  和修改型土壤调整植被指数  $MSAVI$  最大值合成方法从每 3 旬数据合成每月数据; 对每一种处理后的原始数据计算四种植被指数, 并对这 16 种数据进行主成分变换, 分析不同处理方式对主分量累积方差和各主分量所反映生物学规律的影响。

关键词: 遥感植被指数; 主成分分析; NOAA-AVHRR

中图分类号: TP79 文献标识码: A

## 1 引 言

当前大尺度土地覆盖变化研究主要使用 NOAA-AVHRR 数据, 通过计算植被指数进行土地覆盖分类和变化研究。已提出的各种植被指数<sup>[1,2]</sup> 如比值植被指数  $RVI$ 、归一化植被指数  $NDVI$ 、土壤调整植被指数  $SAVI$  和修改型土壤调整植被指数  $MSAVI$ , 对植被的敏感性、抗土壤和大气的干扰性等不尽相同。由于  $NDVI$  与一些重要的生物物理参数如生物量、叶面积指数和光有效辐射等有密切的联系<sup>[1,3]</sup>, 所以  $NDVI$  被广泛用于植被研究。而主成分分析在多时相 AVHRR 研究中, 作为一种常用的数据处理方法, 用于压缩和提取主要植被信息和变化信息。但对于应用多种遥感植被指数进行主成分分析有何影响的研究较少。

本文对中国全年 NOAA-AVHRR 的 1 km 覆盖数据分别计算上述四种植被指数, 并进行主成分分析, 对比分析不同处理方式对各主分量的方差变化和所反映的生物学特征的影响。

## 2 遥感植被指数

### (1) 比值植被指数 $RVI$

$$RVI = Ch2 / Ch1 \quad (1)$$

其中,  $Ch1$  为可见光通道,  $Ch2$  为近红外通道。为存储和计算方便, 将  $RVI$  乘以 100, 用  $RVI * 100$  表示。 $RVI$  是最早发展的植被指数, 比用单波段信息监测植被更为稳定, 但对大气影响较敏感。

### (2) 归一化植被指数 $NDVI$

$$NDVI = (Ch2 - Ch1) / (Ch2 + Ch1) \quad (2)$$

其中,  $Ch1$  为可见光通道,  $Ch2$  为近红外通道。为存储和计算方便, 将  $NDVI$  乘以 1000, 用  $NDVI * 1000$  表示。 $NDVI$  是目前应用最广的植被指数, 但在植被覆盖稠密的地区,  $NDVI$  容易较早地饱和; 在植被覆盖稀疏的地区,  $NDVI$  受土壤背景影响较大。

### (3) 土壤调整植被指数 $SAVI$

$$SAVI = (Ch2 - Ch1) / (Ch2 + Ch1 + L) * (1 + L) \quad (3)$$

其中,  $Ch1$  为可见光通道,  $Ch2$  为近红外通道,  $L$  取 0.5。为存储和计算方便, 将  $SAVI$  乘以 1000, 用  $SAVI * 1000$  表示。 $SAVI$  能较好地消除土壤背景对植被指数的影响。

### (4) 修改型土壤调整植被指数 $MSAVI$

$$MSAVI = (2Ch2 + 1 - \sqrt{(2Ch2 + 1)^2 - 8(Ch2 - Ch1)}) / 2 \quad (4)$$

其中,  $Ch1$  为可见光通道,  $Ch2$  为近红外通道。为存储和计算方便, 将  $MSAVI$  乘以 1000, 用  $MSAVI * 1000$  表示。

收稿日期: 2000-12-14; 修订日期: 2001-02-20

基金项目: 中国科学院知识创新重大项目 KZCX1-SW-01-02 和中国科学院遥感应用研究创新项目。

作者简介: 延 昊(1973—), 男, 中国科学院遥感应用研究所在读博士, 1994年毕业于西北大学地理系, 1997年毕业于北京师范大学资源所, 获自然地理学硕士学位。现从事资源遥感和全球变化遥感研究工作。

1000 表示。MSAVI 是对 SAVI 指数的进一步改进, 较好地消除了土壤背景对植被指数的影响。

### 3 数据处理方法

#### (1) NOAA-AVHRR 数据

研究中使用的美国 NOAA-AVHRR 陆地数据集, 数据的空间分辨率为 1 km, 时间是 1992 年 5 月至 1993 年 4 月共 12 个月的每旬合成数据。数据集所使用的合成方法是最大值合成法(MVC), 即在每个像元取该像元每旬的 *NDVI* 最大值, 同时记录各通道的灰度值。这种数据合成方法有效地去除云, 从而提供了近似于晴空、无云的大气条件<sup>[4]</sup>。进一步, 本文应用阈值法<sup>[5]</sup>进行云检测, 删除受云影响的像元。使用的两个原始通道为 *Ch1* (可见光 0.58 - 0.68  $\mu\text{m}$ ) 和 *Ch2* (近红外 0.725 - 1.10  $\mu\text{m}$ )。

#### (2) 合成每月的原始数据

按照四种植被指数最大值合成法从每三旬数据合成每月数据, 比值植被指数 *RVI* 作为第一种处理, 归一化植被指数 *NDVI* 作为第二种处理, 土壤调整植被指数 *SAVI* 作为第三种处理, 修改型土壤调整植被指数 *MSAVI* 作为第四种处理。

#### (3) 计算每月的四种植被指数

由合成的每月原始数据 *Ch2* 和 *Ch1* 计算四种植被指数, 共形成 16 种数据: *RVI1*、*NDVI1*、*SAVI1*、*MSAVI1*、*RVI2*、*NDVI2*、*SAVI2*、*MSAVI2*、*RVI3*、*NDVI3*、*SAVI3*、*MSAVI3*、*RVI4*、*NDVI4*、*SAVI4* 和 *MSAVI4* (1, 2, 3, 4 分别指前述四种数据合成处理方法)。

#### (4) 主成分分析

主成分分析 (Principal Component Analysis) 在数学上称为 KL (Karhunen-Loeve) 变换, 是一种基于数据统计特征的多维正交线性变换, 也是数字图像处理中常用的一种图像变换方法<sup>[6]</sup>。主成分分析在遥感上主要应用于数据压缩、图像信息的增强, 以及对地表覆盖动态变化的监测<sup>[7, 8]</sup>。进行主成分分析的算法是: 基于每月植被指数图像的均值, 计算均方差矩阵, 并求特征值和特征向量, 接着计算主成分变换矩阵, 最后计算出 12 幅主分量图像。

### 4 结果与分析

遥感植被指数图像均值和主成分变换的各种结果显示: *RVI1*、*RVI2*、*RVI3* 和 *RVI4* 很相似; *NDVI1*、*NDVI2*、*NDVI3* 和 *NDVI4* 很相似; *SAVI1*、*SAVI2*、

*SAVI3* 和 *SAVI4* 很相似; *MSAVI1*、*MSAVI2*、*MSAVI3* 和 *MSAVI4* 很相似, 而 *RVI*、*NDVI*、*SAVI* 与 *MSAVI* 之间的差异较大。应用 4 种植被指数最大值合成法分别从每三旬数据合成每月原始数据, 对随后计算遥感植被指数进行主成分分析影响很小。下面主要分析植被指数 *RVI1*、*NDVI2*、*SAVI3* 和 *MSAVI4* 对主成分分析的影响。

#### 4.1 计算中国每月植被指数图像的均值

计算出每月植被指数图像的均值, 见图 1。

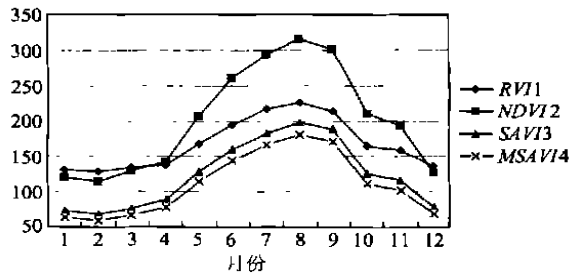


图 1 遥感植被指数各月平均值

Fig. 1 Monthly mean of vegetation indices

显然, 4 种植被指数各月均值曲线形状较相似, 都为单峰形, 2 月份植被指数最低, 8 月份最高, 与实际植被的季节变化特征相符合。

#### 4.2 主成分变换

基于每月植被指数图像的均值, 计算均方差矩阵, 并求特征值和特征向量 (表 1, 图 2), 接着计算主成分变换矩阵, 最后计算出 12 幅主分量图像。

表 1 植被指数各主分量所含信息量 (%)

Table 1 Image segment of the change vector magnitude image for *NDVI*

主分量	<i>RVI1</i>	<i>NDVI2</i>	<i>SAVI3</i>	<i>MSAVI4</i>
1	62.47	69.42	68.61	67.26
2	14.33	12.40	11.09	10.98
3	6.67	4.05	4.85	5.39
4	3.97	3.41	3.85	4.01
5	3.31	2.28	2.58	2.914
6	3.12	2.09	2.08	2.37
7	2.43	1.60	1.92	2.03
8	1.27	1.34	1.52	1.54
9	0.89	1.28	1.19	1.18
10	0.73	0.80	0.93	0.99
11	0.50	0.73	0.84	0.83
12	0.32	0.59	0.55	0.52

主分量按方差的大小排序, 第一主分量的方差最大, 包含了信息量的绝大部分, 最大的 *NDVI2*, 达 69.42%, 其次是 *SAVI3*, 为 68.61%, 最小的是 *RV11*, 仅为 62.47%, 第二主分量、第三主分量和第四主分

量的方差依次减少, 所含信息量也依次减少, 到最后几个主分量几乎全是噪声信息。不同植被指数前 4 个主分量的方差累积值都超过了 85%, 几乎包含了所有信息量, 最大的 *NDVI2*, 达 89.28%, 其次是

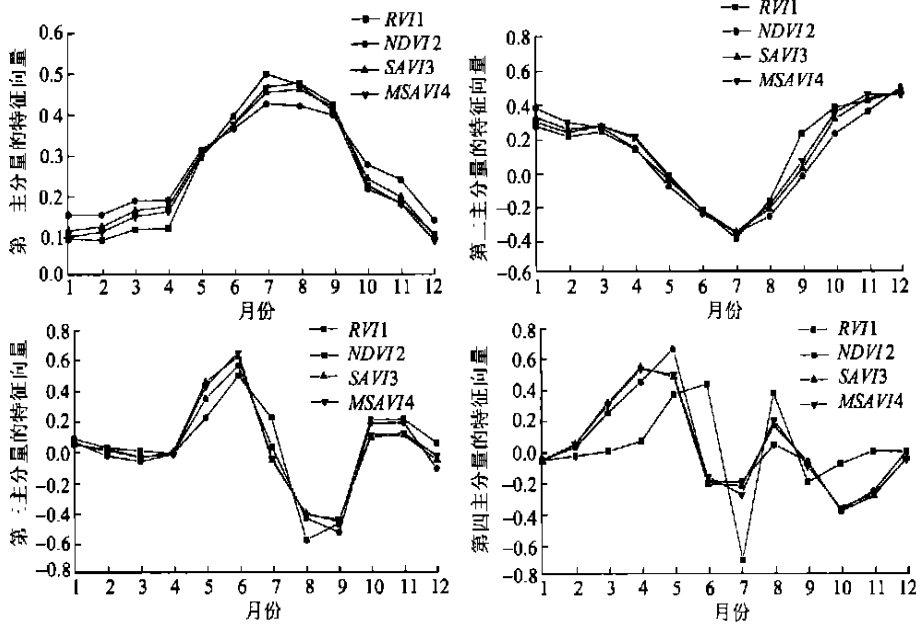


图 2 植被指数的前四个主分量的特征向量曲线  
Fig. 2 Eigenvector curve of the first four PCA components of vegetation indices

*SAVI3*, 为 88.40%, 最小的是 *RV11*, 为 87.44%, 顺序与第一主分量完全相同, 所以确定采用前 4 个主分量进行分析比较和植被分类是可行的。

从图 2 知, 遥感植被指数的前四个主分量的特征向量曲线基本相似, *SAVI3* 和 *MSAVI4* 最相似, *NDVI2* 与前两个较相似, *RV11* 与前三个的相似度最差。NOAA-AVHRR 多时相图像对于第一主分量的贡献都为正值, 最大值为 7、8 月份, 最小值为 1、2 和 12 月份。对于第二主分量, 多时相图像的贡献主要来自冬季的 1、11 和 12 月份和夏季的 6、7 和 8 月份。对于第三主分量, 多时相图像的贡献主要来自春季的 5 和 6 月份和夏季的 8 和 9 月份。对于第四主分量, 多时相图像的贡献主要来自春季的 4 和 5 月份和秋季的 10 和 11 月份。

### 4.3 计算遥感植被指数前四个主分量图像的相关系数

计算出四种植被指数的前四个主分量图像间的相关系数, 列于表 2。

不同植被指数的第一主分量图像间的相关系数

表 2 四种植被指数间主分量图像的相关系数阵  
Table 2 Correlation matrix of PCA components of vegetation indices

第 1 主分量	<i>RV11</i>	<i>NDVI2</i>	<i>SAVI3</i>	<i>MSAVI4</i>
<i>RV11</i>	1.00	0.96	0.95	0.96
<i>NDVI2</i>		1.00	0.98	0.97
<i>SAVI3</i>			1.00	1.00
<i>MSAVI4</i>				1.00
第 2 主分量	<i>RV11</i>	<i>NDVI2</i>	<i>SAVI3</i>	<i>MSAVI4</i>
<i>RV11</i>	1.00	0.92	0.89	0.89
<i>NDVI2</i>		1.00	0.96	0.94
<i>SAVI3</i>			1.00	1.00
<i>MSAVI4</i>				1.00
第 3 主分量	<i>RV11</i>	<i>NDVI2</i>	<i>SAVI3</i>	<i>MSAVI4</i>
<i>RV11</i>	1.00	0.89	0.75	0.73
<i>NDVI2</i>		1.00	0.88	0.84
<i>SAVI3</i>			1.00	0.99
<i>MSAVI4</i>				1.00
第 4 主分量	<i>RV11</i>	<i>NDVI2</i>	<i>SAVI3</i>	<i>MSAVI4</i>
<i>RV11</i>	1.00	0.35	0.30	0.34
<i>NDVI2</i>		1.00	0.90	0.87
<i>SAVI3</i>			1.00	0.99
<i>MSAVI4</i>				1.00

都大于 0.95, 第二主分量图像间的相关系数都大于 0.85, 显示由 4 种植被指数获取的前 2 个主分量较相似。 $RVI1$  的第三分量图像与其他植被指数图像的相关系数较小, 最小值为 0.73, 而另外三个植被指数第三分量图像间的相关系数都大于 0.84。 $RVI1$  的第四分量图像与其他植被指数图像相关系数小于 0.35, 而另外 3 个植被指数第四分量图像间的相关系数都大于 0.87。

$SAVI3$  与  $MASVI4$  的主分量图像相关系数始终较高, 大于 0.99。

遥感植被指数前 4 个主分量图像的相关系数阵显示,  $SAVI3$  和  $MASVI4$  最相似,  $NDVI2$  与前两个较相似,  $RVI1$  与前 3 个植被指数的相似度最差。

## 5 结论与讨论

通过对中国全年 NOAA-AVHRR 的 1 km 遥感植被指数图像进行主成分分析, 结果表明:

(1) 应用比值植被指数  $RVI$ 、归一化植被指数  $NDVI$ 、土壤调整植被指数  $SAVI$  和修改型土壤调整植被指数  $MASVI$  进行主成分分析, 结果有一定差异。就前 4 个主分量的方差累积值来讲, 最大的是  $NDVI$ , 达 89.28%, 其次是  $SAVI$ , 为 88.40%, 最小是  $RVI$ , 仅为 87.44%。

(2) 前 4 个主分量图像的相关系数阵从统计分析角度对遥感植被指数作出了一定评价:  $SAVI$  和  $MASVI$  最相似,  $NDVI$  与前两个较相似,  $RVI$  与前三个植被指数的相似度最差。

(3) 主成分变换的前 4 个主分量包含了丰富的生物学信息: 第一主分量反映了植被指数的全年累积值, 第二分量提取了冬夏两季的植被变化信息, 第三主分量提取了春夏两季的植被变化信息, 第四主分量提取了春秋两季的植被变化信息。

(4) 对单一变量的多时相图像进行主成分分析, 其第一主分量反映该变量的一般特征信息, 而第二和其余主分量反映的是一种权重逐渐减小的变化特征信息<sup>[9]</sup>。植物在红光波段的强烈吸收及近红外波段的强烈反射, 所形成的光谱信息与生物量密切相关<sup>[1]</sup>。遥感植被指数最初就是为了定量表达植物

的生物量, 但不同植被指数具有不同的特点: 与叶面积指数的相关性、对植被的敏感性、抗土壤和大气的干扰性等不尽相同。正是遥感植被指数在定量表达植物生物量上存在差异, 当使用相同的主成分分析方法时, 产生有差异的结果。

(5) 应用不同的植被指数最大值合成方法从每 3 旬数据合成每月原始数据, 再计算每月植被指数, 对进行主成分分析的结果影响很小。由于本文所使用的原始数据是经过  $NDVI$  最大值合成的每旬数据, 因此对使用不同最大值合成法合成每旬数据对主成分分析的影响将在下一步工作中展开。

## 参 考 文 献 (References)

- [1] Zhang Renhua. Experimental Model of Remote Sensing and Field Work[M]. Beijing: Chinese Science Press, 1996. [张仁华. 实验遥感模型及地面基础[M]. 北京: 科学出版社, 1996.]
- [2] Tian Qingjiu, Min Xiangjun. Advances in Study on Vegetation Indices [J]. *Advanc in Earth Sciences*, 1998, 13(4): 219-225. [田庆久, 闵祥军. 植被指数研究进展[J]. 地球科学进展, 1998, 13(4): 219-225.]
- [3] Ruimy, A, Saugier, B. Methodology for the estimation of terrestrial net primary production from remotely sensed data[J]. *Journal of Geophysical Research*, 1994, 99(3): 5263-5283.
- [4] Hollen B. N. Characteristics of maximum-value composite images for temporal AVHRR data [J]. *International Journal of Remote Sensing*, 1986, 7(11): 1417-1434.
- [5] Saunders, R. W., and Kriebel, K. T. An improved method for detecting clear sky and cloudy radiances from AVHRR data [J]. *International Journal of Remote Sensing*, 1988, 9(1): 123-150.
- [6] Rong Guanao. Processing of Computer Image [M]. Beijing: Tsinghua University Press, 1996. [容观澳. 计算机图像处理[M]. 北京: 清华大学出版社, 2000.]
- [7] Fung, T., Ledrew, E. Application of principal components analysis to change detection [J]. *Photogrammetry Engineering & Remote Sensing*, 1987, 53(12), 1649-1658.
- [8] Hiroswa, Y., Marsh, S. E., Kliman, D. H. Application of standardized principal components analysis to land-cover characterization using multitemporal AVHRR data [J]. *Remote Sensing of Environment*, 1996, 58(3): 267-281.
- [9] Eastman, J. Ronald, and Fulk Michele. Long Sequence Time Series Evaluation Using Standardized Principal Components [J]. *Photogrammetry Engineering & Remote Sensing*, 1993, 59(8), 1307-1312.

## Application of Principal Component Analysis by Using Different Vegetation Index Derived from Multitemporal AVHRR Data

YAN Hao, WANG Chang-yao, NIU Zheng, JIANG Xiao-guang, WANG Wen

(LARSIS, Institute of Remote Sensing Applications, CAS, Beijing 100101, China)

**Abstract:** Based on the 12 months' 1km AVHRR data in China, this paper computes four kinds of vegetation index (VI), that are ratio vegetation index (*RVI*), normalized vegetation index (*NDVI*), soil adjusted vegetation index (*SAVI*) and modified soil adjusted vegetation index (*MSAVI*). Then, we make the same principal components analysis (*PCA*) to them, and find that the *PCA*-transformed first four principal components (*PCA1*, *PCA2*, *PCA3*, *PCA4*) contribute about 88% cumulative variance, and *PCA1* represents VI cumulation of whole year, *PCA2* represents VI difference of winter and summer, *PCA3* represents VI difference of spring and summer, *PCA4* represents VI difference of spring and autumn. In other words, for multitemporal vegetation index of one year, *PCA* not only compresses the information to the first four principal components, but also extracts the key change information. The *PCA1* expresses the basic land cover information, the others extract the seasonal change information of vegetation.

However, the outcome of different vegetation index has some differences. As to the cumulative variance of the first four eigenvectors, the biggest is *NDVI*, 89.28%, the second is *SAVI*, 88.40%, and the smallest is *RVI*, only 87.44%. As to the correlation matrix of four vegetation index, *SAVI* and *MSAVI* are the most similar, *NDVI* is much similar with the first two vegetation indices, and *RVI* is the least similar. Although the primary purpose of VI is to indicate the biomass of vegetation, due to the different features of VI, such as different correlation with leaf area index, different sensitivity to vegetation and different anti-disturbance of soil and atmosphere, different VI indicates different biomass for the same vegetation, that is, when we use the same *PCA* to different VI, the result is not uniform.

**Key words:** vegetation index; *PCA*; NOAA-AVHRR