

基于遥感和GIS的中国植被指数变化的驱动因子分析及模型研究^{*}

高志强 刘纪远

(中国科学院遥感应用研究所, 北京 100101)

摘要 在样带和典型区研究的基础上, 采用相关分析和偏相关分析方法, 对影响植被指数变化的因子(水、热和地表植被覆盖类型)进行了分析。结果表明: 中国植被指数的时空变化极其复杂, 虽受水、热和地表植被覆盖类型三个主导因子的影响和控制, 但因时和因地而异, 三者对植被指数影响和控制的主导地位也因时因地而不同; 基于空间上的概念模型 $I_{ndv} = F(x, y, z)$ 只能定性地描述以上三个主导因子时空变化同植被指数的相互关系。

关键词: 土地覆盖; 植被指数; 相关分析; 偏相关分析; 驱动模型

1 引言

土地利用和土地覆盖的研究已成为全球环境变化研究的重要部分和核心内容。“国际地圈与生物圈计划”(International Geosphere-Biosphere Program, IGBP)和“全球环境变化与人类发展计划”(Human Dimensions of Global Environmental Change Programs, HDP)的“全球土地利用与土地覆盖变化”(Land-Use and Land-Cover Change, LUCC)研究报告中, 不仅提出了有关土地利用和土地覆盖变化的五个研究框架问题, 并且提出了土地利用和土地覆盖的三个研究重点: 土地利用变化机制研究、土地覆盖变化机制研究和地区和全球性模型研究, 其中土地利用和土地覆盖的变化机制研究是土地利用和土地覆盖研究的焦点问题, 在全球变化研究中占据重要地位^[1]。

本文在样带和典型区科学选择的基础上, 就植被指数的相关影响因子进行分析, 然后, 研究和设计了植被指数时空变化的驱动模型^[2~4]。

2 中国植被指数时空变化研究的样带和典型区选择

2.1 中国植被指数的样带选择

样带指全球变化或区域环境变化中具有代表性的地带和敏感区。中国地域辽阔, 从南到北跨越热带、亚热带、暖温带和温带, 从东向西包括了湿润、半湿润、半干旱、干旱地区; 自然景观上从东向西分为“三级阶梯”的分异特色。所以, 针对研究目的——中国土地利用和土地覆盖的时空变化及驱动因子, 我们在进行样带选择时, 按纬向选择了

1998-12-30收到, 1999-05-05收到修改稿

^{*} 本文得到中国科学院“九五”重大项目“国家资源与环境信息系统与农情速报”KT95T-01-04和中国科学院知识创新项目“国土资源遥感时空信息分析与数字地球相关理论技术预研究”KZCX1-Y-02-1-02的资助

三条,按经向选择了二条(如图1):按1:100万经纬度网格,纬向选择了处在温带与暖温带过渡的 40°N 附近地带、暖温带与北亚热带过渡的 28°N 附近地带和北热带与南亚热带过渡的北纬 24°N 附近地带;经向选择了中国地形一、二级阶梯(东部沿海平原和中部高原)交接处的 114°E 地带和二、三级阶梯(中部高原和西部荒漠)交接处的 108°E 地带。以上五条样带作为对中国植被指数研究的样带。

2.2 中国植被指数的典型区选择

样带的选择是为了对植被指数的变化在空间分布上进行分析、比较,而典型区的分析着重是对植被指数的变化在时间序列上进行比较分析。

为了研究植被指数在不同区域的时间变化规律,我们选择了3组共13个典型区(如图1)。典型区1、2、4、6、7、8位于季风区,主要研究热量变化对植被指数时间变化的影响,以上6个典型区均代表着一个温度带;典型区3、9、10、11基本沿 40°N 纬线向西延伸, 40°N 纬线是中国干湿变化最明显的地带,故选择了4个典型区来研究干旱变化对土地利用和土地覆盖时间变化的影响;典型区5、12、13位于中国一、二、三级阶梯上,通过对5、12、13三个典型区的选择,研究中国三级阶梯地势对植被指数时间变化的影响。以上13个典型区内有一个国家级气象站,以便将气候因子同其他环境因子进行比较分析。

本文采用的数据为:植被指数NDVI数据采用LCWG/AARS(The information of the Land Cover Working Group of the Asian Association on Remote Sensing)提供的CD-ROM 1994年全球8 km分辨率的AVHRR栅格数据,将中国数据切出再进行

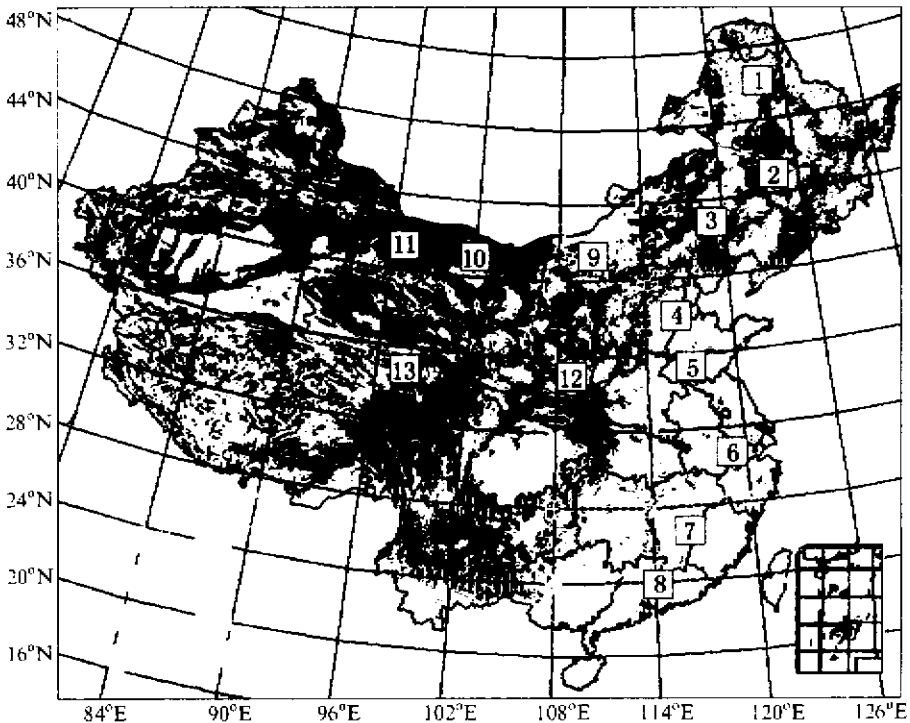


图1 中国植被指数样带及典型区研究示意图

等面积圆锥投影变换, 再将此 $8\text{ km} \times 8\text{ km}$ 的 NDVI 的数据进行重采样变为 $2\text{ km} \times 2\text{ km}$ 的全国 2075×2517 的阵列规格即可; 水热指标是点位数据, 采用《中国地面气候资料 1951~1980》的 671 个站点资料, 将全国 671 个台站经纬度坐标按投影变换公式换算成等面积割圆锥投影体系, 为下面的各种分析所用。

3 中国植被指数时空变化的影响因子分析

土地覆盖是地球陆地表层和近地面层的自然状态, 是自然过程和人类活动共同作用的结果; 由于植被具有明显的年、季节变化特点, 并且是联结土壤、大气和水分的自然“纽带”, 在一定程度上代表土地覆盖的变化; 植被指数是根据植被反射波段的特性计算出来的, 是反映地表植被种类情况的间接指标, 是反映植物生长状态及植被空间分布密度的最佳指示因子, 与植被覆盖分布密度呈线性相关关系。因此, 可利用植被指数作为土地覆盖的一种量化指标, 通过对植被指数同环境因子关系的研究, 寻找土地覆盖变化的主要驱动因子, 并在此基础上建立土地覆盖变化的驱动模型^[5,6]。

3.1 基于典型区研究的中国植被指数变化的相关分析

在典型区分析中, 从 13 个典型区的植被指数和环境因子的变化曲线可以看到, 在水、热因子决定植被指数年际变化曲线主体形状基础上, 下垫面覆盖的不同及变化导致其主体形状产生波动, 风速、日照百分率、蒸发虽对植被指数有影响, 但并不是主导因子, 在此暂不作分析。基于以上工作基础, 将月均温、月日辐射总量、月降水同植被指数进行相关分析 (如表 1)。对相关系数进行显著性检验, 得到 $F_{0.05}(1, 10) = 4.96$, $F_{0.001}(1, 10) = 10.0$ 。

表 1 典型区研究中植被指数同水、热因子的相关系数及检验值

典型区	月气温		月太阳辐射总量		月降水	
	相关系数	F值	相关系数	F值	相关系数	F值
典型区 3(内蒙古 23)	0.88	33.21	0.71	10.12	0.91	46.34
典型区 9(内蒙古 32)	0.91	48.3	0.71	10.51	0.94	79
典型区 10(内蒙古 36)	0.68	8.84	0.70	9.8	0.80	17.33
典型区 11(甘肃)	0.71	10.40	0.84	25	0.45	2.64
典型区 5(山东)	0.90	40.41	0.76	13.46	0.82	20.9
典型区 12(陕西)	0.93	63.5	0.85	26.71	0.90	45.24
典型区 13(青海)	0.87	30.8	0.72	10.51	0.92	54
典型区 1(黑龙江)	0.95	91.46	0.86	27.41	0.92	54.56
典型区 2(吉林)	0.92	51.65	0.78	15.7	0.89	37.66
典型区 4(河北)	0.84	24.8	0.64	6.96	0.75	12.94
典型区 6(安徽)	0.98	237	0.91	50.21	0.80	17.88
典型区 7(江西)	0.86	29.5	0.81	19.55	0.65	7.38
典型区 8(广东)	0.82	20.61	0.90	44	0.59	5.51

分析表 1 可知: 植被指数同水、热因子的相关系数同时通过 $\alpha = 0.01$ 水平显著性检验的典型区有: 典型区 6 (安徽)、典型区 4 (河北)、典型区 1 (黑龙江)、典型区 2 (吉林)、典型区 3 (内蒙古 23)、典型区 9 (内蒙古 32)、典型区 13 (青海)、典型区 5

(山东)、典型区 12 (陕西); 而典型区 8 (广东)、典型区 7 (江西)、典型区 11 (甘肃)、典型区 10 (内蒙古 36), 都没有通过 $\alpha=0.01$ 水平的显著性检验。原因是典型区 8 (广东)、典型区 7 (江西) 水热充足, 以自然常绿林覆盖为主, 植被指数年际变化不大, 导致水、热因子同植被指数相关不显著, 而典型区 11 (甘肃)、典型区 10 (内蒙古 36) 是以沙漠植被和无植被地覆盖为主, 植被指数年际变化也不大, 同水、热就无多大关系了。

从以上相关分析可知: 水、热因子同植被指数的相关关系是基于下垫面覆盖状况的, 当下垫面覆盖类型成为主导因子时, 以上相关关系就不明显, 甚至不存在。

3.2 基于样带研究的中国植被指数变化的相关分析

通过样带分析可以看出, 因每条样带经过的空间跨度较大, 水热变化大, 下垫面覆盖类型复杂, 导致植被指数的空间变化非常复杂, 植被指数值在空间上的变化是水分、热量和下垫面覆盖相互作用影响的结果。在 5 条样带上, 植被指数在水、热的总体控制下, 因下垫面的影响导致变化曲线上的许多峰值和低谷。因下垫面覆盖作用很难量化, 现就水、热的作用进行相关分析 (如表 2)。

表 2 基于样带的植被指数同水、热因子的相关系数及检验值

样带	年降水量		$\geq 10^{\circ}\text{C}$ 积温	
	相关系数	F 值	相关系数	F 值
40°N	0.95	839	0.60	49.17
28°N	0.75	56	0.73	48.66
24°N	0.56	17	0.60	20.83
108°E	0.86	103.25	0.76	50.0
114°E	0.88	131.6	0.83	86.1

通过查 F 表可知, $\alpha=0.01$ 水平下, $F(1,30)=7.56$ 。由统计原理知, 当相关系数 >0.70 时, 两者之间有实质性相关关系。结合以上检验值, 表 2 中, 相关系数都通过 $\alpha=0.01$ 水平的显著性检验。比较东西向三条样带, 从北到南, 植被指数同水分的相关系数为 0.95、0.75 和 0.56, 相关关系由显著变为不显著。在 24°N 样带上, 水、热同植被指数的相关关系不明显, 水、热相关系数都 <0.70 , 原因是, 24°N 样带在北回归线附近, 处于南亚热带, 水、热充足, 为常绿阔叶林覆盖区, 植被指数年际变化不大, 其变化是因下垫面覆盖变化导致的, 故同水、热的相关关系不明显。在 28°N 样带上, 植被指数同水、热相关明显; 在 40°N 样带上, 植被指数同降水明显相关, 而同年积温相关系数 <0.70 , 即无实质性相关关系, 因为在 40°N 样带上, 从东向西, 同纬度地区热量变化不大, 水分却从湿润区到干旱区急剧变化, 成为制约植被指数大小的主要因子, 故植被指数同水分指标相关关系显著。在 108°E、114°E 样带上, 植被指数同水、热相关明显, 因为沿 108°E 和 114°E 从南到北水、热变化都是显著的, 对植被指数的影响也是明显的, 故有实质性相关关系。

3.3 影响植被指数变化的主导因子变化分析

通过以上典型区及样带的相关分析可以看出: 空间上植被指数同水、热和下垫面覆盖状况的关系非常的复杂, 其主导地位因时因地而异。

下面利用全国 671 个站点的月均温、月降水资料, 分析同月植被指数的相关关系的

空间分布规律。

为了分析以上相关关系的空间分布, 首先必须生成相关系数图和偏相关系数图。利用包括 671 个站点在内的空间矢量数据, 将月均温、月降水以公共项联结到属性库中, 然后, 以月植被指数数据为基础, 将前面的 671 个站点的空间矢量数据用 arc/info 中的 pointgrid 命令转为栅格数据, 命名为 pgrid, 再利用 grid 模块中的 combine 命令, 将栅格形式的月植被指数数据同 pgrid 进行 combine 运算, 提出每个气象站点的月植被指数值, 最后, 将得出的 12 个月的、671 个站点的植被指数, 以公共项联结到前面的包括 671 个站点在内的空间矢量数据的属性库中, 这样这个属性库中, 每个站点有一年 12 个月的温度、降水和植被指数值。

以上面的属性库为基础, 利用下面相关系数计算公式计算出每个站点的植被指数同降水和温度的相关系数:

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

利用如下的偏相关系数计算公式计算出基于降水的植被指数与气温的偏相关系数和基于气温的植被指数与降水的偏相关系数:

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}. \quad (2)$$

以包括 671 个站点的矢量数据为基础, 利用 arc/info 中的 arctin 和 tinlattice 命令, 将以上四个相关系数值经插值转换, 变为 2075×2517 的栅格格式, 然后将以上相关系数分级赋色就生成了二张相关系数图和二张偏相关系数图。

相关系数和偏相关系数生成后, 要对相关系数和偏相关系数进行检验。相关系数的检验是在给定的置信水平下, 通过查相关系数的临界表来完成的。当置信水平 $\alpha = 0.05$ 时, 相关系数 $r > 0.576$, 要素之间才有明显相关关系。对偏相关系数的检验, 一般采用 t 检验法:

$$t = \frac{r_{12.3}}{\sqrt{1-r_{12.3}^2}} \sqrt{n-m-1}, \quad (3)$$

其中, $r_{12.3}$ 为基于第三要素的两要素之间的偏相关系数, n 为样本数, m 为自变量数。当置信水平 $\alpha = 0.05$ 时, 偏相关系数 > 0.5493 , 相关关系才明显。

用以上相关系数和偏相关系数临界值, 分析植被指数和温度、降水的相关关系及偏相关系数空间分布状况。

从图 2 可以看出: 植被指数同降水的相关系数 $r > 0.576$ 的地区为深兰和淡兰的地区, 而华北平原、江南丘陵、云贵高原, 因是湿润区, 水分充足, 对植被的生长不起限制作用, 相关性低; 在西北区, 阿拉善高原、塔里木盆地等以沙漠覆盖为主的地区, 因年际降水少, 地表植被覆盖稀少, 植被指数的年际变化不大, 同降水的关系也不大。而在半湿润半干旱地区, 植被指数的年际变化却依赖于降水的多少, 如东北平原、黄土高

原、四川盆地等区。因为对植被指数的影响是多要素的,水分、热量及地表覆盖都会影响植被指数的年际变化,为了研究降水对植被指数的影响,我们将温度视为常数,暂不考虑它对植被指数的影响,求植被指数同降水的偏相关系数(图3)。对照图2和图3,除去温度的影响,降水对植被指数的影响主要分布在半湿润、半干旱区,而湿润区的江南丘陵、黄淮海平原等地区降水充沛,水分对植被指数的年际变化影响不大,西北干旱区,降水稀少,且以荒漠覆盖为主,植被指数年际变化同降水关系不大,只有东北三省、黄土高原、四川盆地、天山南缘、塔里木河谷地、青藏高原区,植被指数的年际变化很大程度依赖水分的多少,这些地区显示出植被指数同降水极大的相关性。

由以上分析可知:降水对植被指数的年际影响,因区域不同,地表覆盖类型不一样,其相关关系在空间分布上因地制宜,非常复杂。

图4和5给出了植被指数和温度的相关关系和偏相关关系。在相关系数图(图4)上,云贵高原和横断山脉、台湾、海南岛等南方亚热带区,因一年热量丰富,对植被的生长没限制作用,植被指数同温度相关性不好;在西北干旱区,虽年际温度变化大,但下垫面以荒漠覆盖为主,植被指数年际变化很少,导致植被指数同温度的相关性也不好;其余中国绝大部分地区,温度和植被指数的相关性是较好的。在偏相关系数图(图5)上,除去降水对植被指数的影响,可以看到相关系数小的地区分三种类型:(1)热量充足的地区,如亚热带及热带地区;(2)热量不足的地区,如青藏高原;(3)以农耕地覆盖为主的地区,如黄淮海及长江中下游平原、黄土高原、四川盆地、三江平原等地区。这些地区为植被指数同温度相关性较小的地区。其原因为热量充足的地区热量丰富,对植被的生长没限制作用,而热量不足的地区,因热量不足,导致植被年际变化不大,导致这些地区植被指数的年际变化同温度的关系不大;而农垦地区,因植被覆盖的变化受人为的作用,年际变化同温度不同步,也造成植被指数同温度不相关。而其余地区都显示了植被指数同温度不同的相关性。

由以上分析看出:植被指数同温度的相关性在空间上的分布也是因地制宜,非常复杂。

4 中国植被指数的时空变化驱动模型研究

由前面的典型区、样带分析及基于空间的相关分析,可以看出,水、热及下垫面覆盖对植被指数的年际变化的影响因地制宜,三者在对植被指数的影响中,其主导地位因地因时而变化,使得建立一个基于空间的模拟水、热、下垫面覆盖对植被指数影响的定量化模型非常困难。由典型区分析可知,建立基于一个点上的除去下垫面覆盖变化的影响在时间序列上的水、热同植被指数的定量相关关系是可能的。在前面13个典型区分析中,已看到了水、热同植被指数的年际变化基本是同步的,水、热同植被指数具有线性相关关系(表1),除去典型区8(广东)、典型区7(江西)、典型区11(甘肃)、典型区10(内蒙古36)4个典型区中的相关关系没有通过 $\alpha=0.01$ 水平的显著性检验外,别的全通过 $\alpha=0.01$ 水平的显著性检验。

基于以上分析,可以用 $z = ax + by + c$ 二元函数来模拟点上的水、热同植被指数的

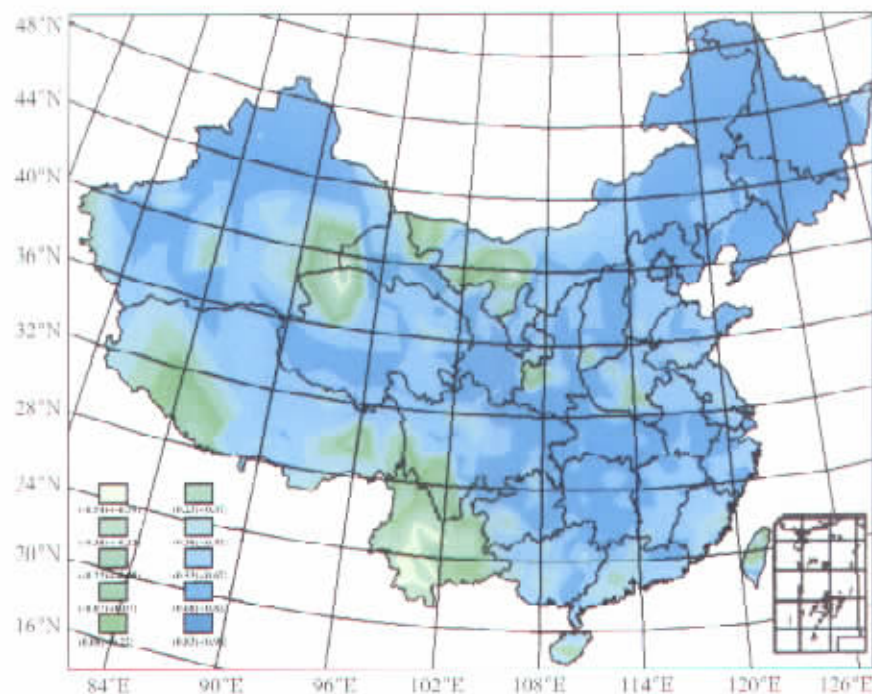


图4 植被指数和气温的相关系数图

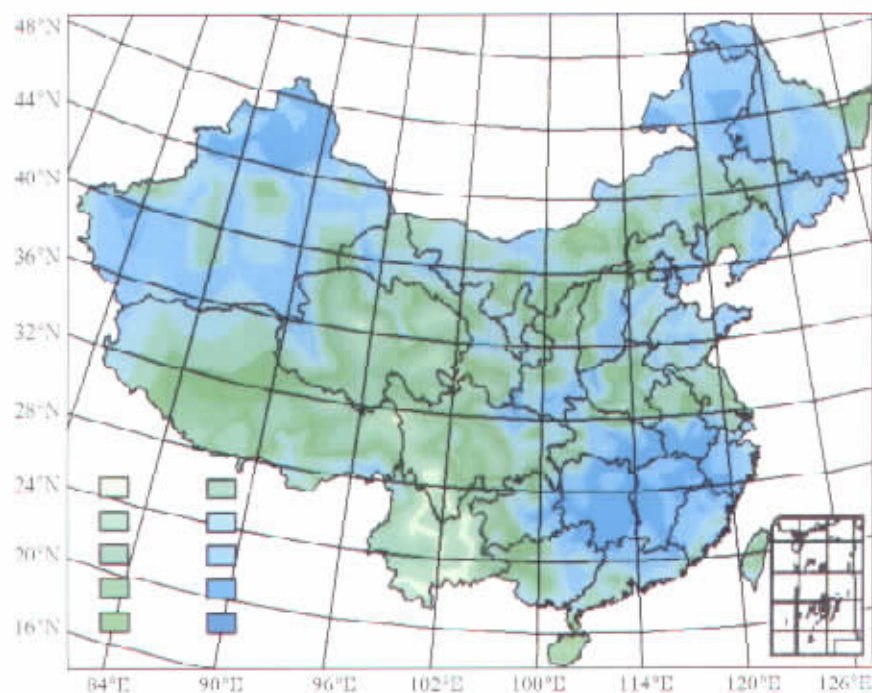


图5 基于降水的植被指数和气温的偏相关系数图

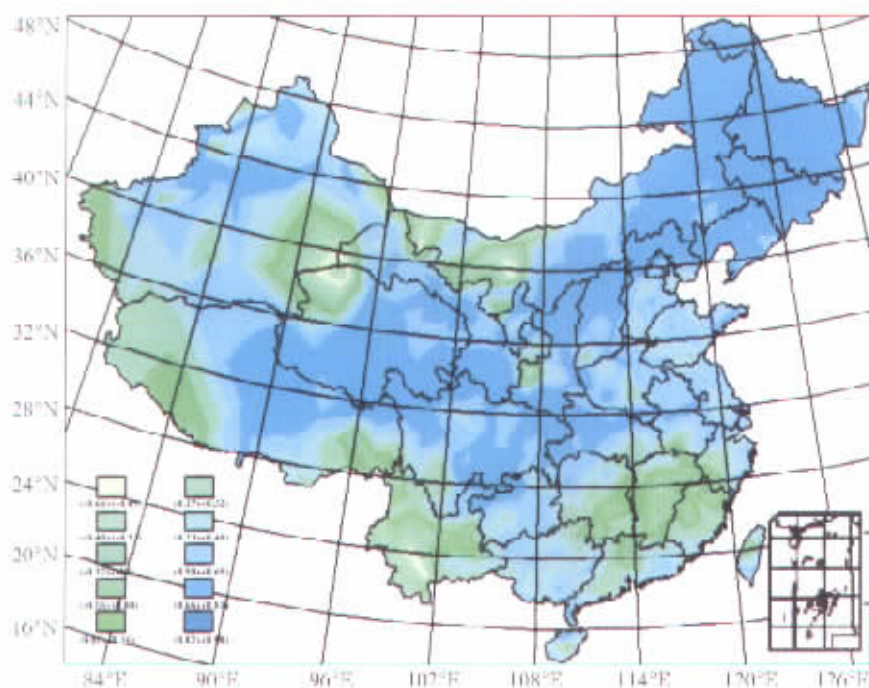


图2 植被指数和降水的相关系数图

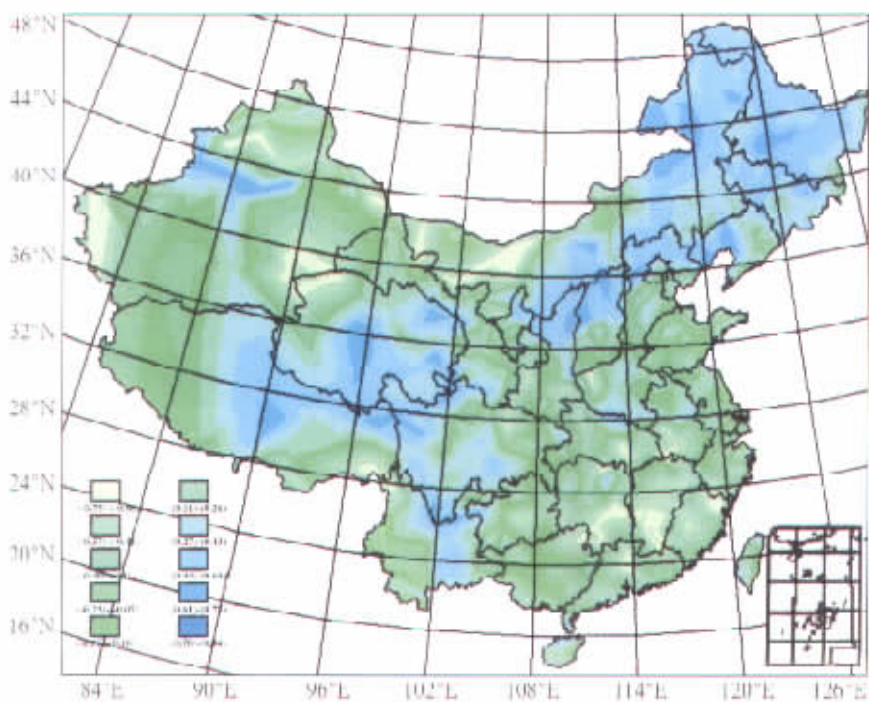


图3 基于气温的植被指数和降水量的偏相关系数图

定量关系。表 3 列出了经计算后 9 个典型区的二元回归方程。对表 3 的回归方程进行 F 检验，查 F 表，置信水平 $\alpha=0.001$ 时， $F(2, 9)=8.02$ ，置信水平 $\alpha=0.05$ 时， $F(2, 9)=4.26$ 。

由表 3 中所列 F 值可知：所有方程均通过 $\alpha=0.05$ 水平的显著性检验，即以上回归方程反映了这些点的水、热同植被指数的定量函数关系。

表 3 基于典型区的植被指数同气温和降水的回归方程

典型区	回归方程	F 值
典型区 6(安徽)	$I_{ndv} = 1.55066T - 0.10137P + 144.247$	115.7
典型区 4(河北)	$I_{ndv} = 0.8582T + 0.7945P + 141.779$	7.58
典型区 1(黑龙江)	$I_{ndv} = 1.0777T + 0.31P + 149.7414$	64.94
典型区 2(吉林)	$I_{ndv} = 0.733T + 0.231P + 137.65$	33.32
典型区 3(内蒙古 23)	$I_{ndv} = 0.2056T + 0.2236P + 137.73$	22.53
典型区 9(内蒙古 32)	$I_{ndv} = 0.3086T + 0.2811P + 136.26$	67.03
典型区 13(青海)	$I_{ndv} = -0.12068T + 0.526P + 132.26$	23.95
典型区 5(山东)	$I_{ndv} = 0.994T + 0.0596P + 142.8$	12.33
典型区 12(陕西)	$I_{ndv} = 1.042T + 0.0919P + 144.76$	29.69

表 4 比较了典型区 5（山东）植被指数的模拟值和实际值。由表 4 可知：差值大的月份是 6、8、10 月。6 月是山东的夏季小麦收获季节，小麦收获后，导致地表裸露，使植被指数急剧下降，比理论预测值要小；8 月是山东进入雨季，秋季作物加速生长导致植被指数比理论预测值要大；10 月是山东的秋收季节，地表裸露，导致植被指数比预测值要小。由此，进一步验证了植被指数的大小受水、热和地表覆盖三方面因子影响和控制的结论。以上拟合方程是基于点的，仅考虑了水、热两个因子的影响，而没有考虑地表覆盖变化的影响，以上检验值的变化都可以从地表覆盖的变化中加以解释。

表 4 典型区 5（山东）的植被指数模拟值和实际值的比较

月 份	气 温	降 水	植被指数	拟合值	差值	差值百分比
1 月	-1.90	9.00	136.15	141.45	-5.30	-3.89
2 月	0.90	10.50	141.76	144.32	-2.56	-1.81
3 月	7.10	19.10	152.14	151.00	1.14	0.75
4 月	14.30	43.20	169.54	159.59	9.95	5.87
5 月	20.30	38.50	168.00	165.27	2.73	1.62
6 月	25.40	86.70	156.46	173.21	-16.75	-10.71
7 月	26.90	234.40	178.42	183.51	-5.09	-2.85
8 月	26.00	149.00	191.43	177.52	13.91	7.26
9 月	20.80	64.30	172.60	167.31	5.29	3.07
10 月	14.70	35.50	150.49	159.53	-9.04	-6.01
11 月	7.20	21.10	152.65	151.21	1.44	0.94
12 月	0.50	12.00	148.44	144.01	4.43	2.98

基于以上分析，可以将影响植被指数的三个主导因子：水分、热量和地表覆盖，抽象出一个空间上的概念模型。设 z 为下垫面覆盖类型变化对植被指数的影响因子，它因地因时而异，代表了因覆盖类型的变化对植被指数的影响， x 代表温度，是热量因子， y 代表降水量，是水分因子，函数 $F(x, y, z)$ 代表一种基于空间上 x, y, z 对植被指数影

响的量化函数关系, I_{ndv} 代表某时刻空间上某一点的植被指数值, 基于时间和空间上的植被指数同其三个主导因子: 水、热和下垫面覆盖类型的概念模型关系为

$$I_{ndv} = F(x, y, z).$$

5 结论

中国植被指数的时空变化极其复杂, 虽受水、热和地表植被覆盖类型三个主导因子的影响和控制, 但因时和因地而异, 三者对植被指数影响和控制的主导地位也因时因地而不同, 所以, 很难设计出一个可在时空变化上模拟水、热和地表植被覆盖类型变化对植被指数影响的定量函数模型, 只能从概念上, 用 $I_{ndv} = F(x, y, z)$ 函数形式, 定性地描述植被指数和三个主导因子间的关系。

参 考 文 献

- 1 刘纪远主编: 中国资源环境遥感宏观调查与动态研究, 北京: 中国科学技术出版社, 1996, 262~275.
- 2 庄大方, 刘纪远, 中国土地利用程度的区域分异模型研究, 自然资源学报, 1997, 12(2), 10~14.
- 3 陈百明, 试论中国土地利用和土地覆被变化及其人类驱动力研究, 自然资源, 1997, No.2, 12~15.
- 4 李秀彬, 全球环境变化研究中的核心领域, 地理学报, 1996, 51(6), 23~28.
- 5 吴传钧、郭焕成, 中国土地利用, 北京: 科学出版社, 1994, 1~17.
- 6 叶笃正等, 中国的全球变化预研究(总论), 北京: 气象出版社, 1992, 23~54.

The Study on Driving Factors and Models of NDVI Change Based on Remote Sensing and GIS in China

Gao Zhiqiang and Liu Jiyuan

(Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101)

Abstract On the basis of research on sample strips and typical area, the factors of precipitation, temperature and types of land vegetation cover which effect the changes of the normalized difference vegetation index (NDVI) were analyzed by correlation analysis method. The conclusions are that the changes of NDVI are complicated in China, although the changes are controlled by the factors of precipitation, temperature and types of land vegetation cover, but the status of three factors which can control the changes of NDVI is different in different place and at different time, the concept model of three factors (precipitation, temperature and types of land vegetation cover) effecting the changes of NDVI is $I_{ndv} = F(x, y, z)$, the model describes the relations between the changes of NDVI and three factors from the concept point of view.

Key words: land-cover; normalized difference vegetation index; correlation analysis; partial correlation analysis; driving model