

引用格式:焦伟,陈亚宁,李稚,等.基于多种回归分析方法的西北干旱区植被NPP遥感反演研究[J].资源科学,2017,39(3):545-556. [Jiao W, Chen Y N, Li Z, et al. Inversion of net primary productivity in the arid region of Northwest China based on various regressions[J]. *Resources Science*, 2017, 39(3):545-556]. DOI: 10.18402/resci.2017.03.16

基于多种回归分析方法的西北干旱区 植被NPP遥感反演研究

焦伟^{1,2}, 陈亚宁¹, 李稚¹, 李玉朋^{1,2}, 黄晓然^{1,2}, 李海霞^{1,2}

(1. 中国科学院新疆生态与地理研究所 荒漠与绿洲生态国家重点实验室, 乌鲁木齐 830011;
2. 中国科学院大学, 北京 100049)

摘要:植被净初级生产力(Net Primary Productivity, NPP)是评价陆地生态系统的重要参数,植被NPP的反演为全球碳循环的研究提供了重要的参考依据。本文基于2000-2014年的MODIS及气象站点数据,引入逐步线性回归、主成分回归、偏最小二乘回归以及岭回归分析,构建了多元逐步回归模型、主成分回归模型、偏最小二乘回归模型和岭回归模型,并对这4种模型进行了评估。结果表明:①多元逐步回归模型被认为是一种可以简单、高效、准确地反演西北干旱区植被NPP的模型,且模拟结果与MOD17A3产品及实测值具有很强的相关性;②西北干旱区多年平均植被NPP的分布具有明显的区域差异,山区增加,平原区减少,呈现出北部及西北部高,而南部及东南部低的特征;③西北干旱区NPP在年际变化上呈波动增加趋势,线性增长率为 $0.40\text{gC}/(\text{m}\cdot\text{a})$ 。自2000年以来,58.66%区域的NPP呈现增长趋势,13.64%的地区保持不变,27.7%的地区表现为退化趋势。

关键词:净初级生产力;逐步多元回归;主成分回归;偏最小二乘回归;岭回归;西北地区

DOI: 10.18402/resci.2017.03.16

1 引言

植被净初级生产力(Net Primary Productivity, NPP)是生态系统生产能力和碳汇能力的反映^[1,2],是陆地生态系统和大气之间碳循环的重要组成部分^[3-5]。大量研究表明,随着气温升高,植被生产力不断增加,北半球高纬度地区森林线开始向高纬地区移动^[6-8];但也有研究发现美国西南部地区呈现退化的趋势^[9],表明植被生产力的变化具有空间异质性。因此在区域尺度上研究自然和人为因素扰动作用对陆地生态系统碳循环的影响具有重要的意义。西北干旱区作为对气候变化响应最敏感的地区之一,得到国内外学者的广泛关注,过去半个多世纪,气温上升速率高达 $0.39^\circ\text{C}/10\text{a}$,是全球气温上升速率($0.14^\circ\text{C}/10\text{a}$)的2.78倍^[10,11],因此,准确反演

西北干旱区生物量对全球碳循环的影响以及国家气候谈判中关于碳源汇等问题具有重要的意义。

近年来,全球及区域尺度上植被NPP的反演模型不断发展,从传统的估测方法(包括站点实测法^[12,13]、气候相关统计模型^[14]、生态系统过程模型^[15])发展到基于遥感手段的遥感模型估测法^[16,17]。鉴于站点实测方法在数据采集时的局限性,基于遥感手段的模型方法成为研究植被NPP反演的主要途径,能够快速及时获取全球及区域尺度上的陆地生态系统碳储量。早期的植被生物量研究多采用归一化植被指数来检测植被的生长,随着遥感数据的不断完善,更多学者利用植被指数、地形以及气象因子与生物量之间建立回归模型,以寻求更为简洁、快速反演植被生物量的方法。杨存健等对云南省西双

收稿日期:2016-08-22;修订日期:2017-01-18

基金项目:新疆维吾尔自治区自然科学基金(2015211A048)。

作者简介:焦伟,女,山东泰安人,硕士生,主要从事生态水文方面的研究。E-mail: jiaoweisdu@163.com

通讯作者:陈亚宁, E-mail: chenyn@ms.xjb.ac.cn

版纳森林生物量与气象因子建立回归模型并反演生物量,模型的 R^2 为0.589^[18]。范文义等通过构建逐步回归模型、神经网络模型来反演森林生物量并对模型的精度进行评价^[19]。张旭琛等以实测数据为基础结合遥感气象数据进行回归建模,反演并分析了伊犁地区草地植被生物量的空间分布特征^[20]。但是以上研究没有考虑自变量之间的相关性,这在一定程度上会影响模型模拟的可靠性。

在回归分析中,解决自变量之间共线性问题的主要回归分析方法有逐步回归分析^[19,21]、主成分回归分析^[22]、偏最小二乘回归^[23],以及岭回归分析^[24]。Fan W Y等、Liu Y M等采用逐步多元回归模拟了中国太阳总辐射和森林生物量数据,并证明了模型的可靠性^[19,21]。胡馨月等运用主成分回归进行了温度分布的反演^[22];Bin等将偏最小二乘回归运用在多元光谱分析,相对误差的预测精度为1.88%^[23]。高思远等运用岭回归建立ET0的多元反演研究,提高了模型精度和可靠性^[24]。但是关于植被生物量的模拟及模型的综合比较分析方面的研究还比较少。

本研究采用归一化植被指数(Normalized Difference Vegetation Index, NDVI)、MODIS NPP数据、气象站点数据以及样点实测数据,建立生物量与各影响因子之间的逐步回归模型、主成分回归模型、偏最小二乘模型以及岭回归模型,比较分析了这4类模型在生物量反演方面的精度,以寻求高效、准确反演植被NPP的最优模型,并对西北干旱区植被NPP遥感反演进行初步尝试,为探讨西北干旱区植被NPP的分布规律提供参考,实现西北干旱区植被资源的合理利用与生态环境的可持续发展。

2 研究区概况与数据来源

2.1 研究区概况

西北干旱区位于欧亚大陆腹地,介于 $N34^{\circ}54' - N49^{\circ}19'$ 和 $E73^{\circ}44' - E106^{\circ}46'$ 之间,贺兰山、乌鞘岭以西,昆仑山以北包括新疆全部的广大区域,面积约250万 km^2 (图1)。干旱少雨的气候特点、山盆相间的地貌形态、特殊的土壤植被特征,使得这一地区形成了与中国其他地区迥异的植被类型和生态系统^[25]。这一地区山地生态系统垂直分异明显、绿洲景观特色鲜明,荒漠类型复杂多样,形成了独具一格的山地、绿洲、荒漠共存的地理景观;由于干

旱区水资源短缺且蒸发强烈,在水分胁迫作用下,该区域植被覆盖率低,土地荒漠化问题严重,形成了以浅根系植被为主体,存在大量旱生、耐盐碱、抗风沙的干旱植被,区域生态环境极其脆弱。

2.2 数据来源及预处理

2.2.1 模型变量选择

MOD17A3数据是利用BIOME-BGC模型进行模拟,得到年NPP累积量,时间尺度是一年,空间分辨率是1km。该数据产品目前已在全球和区域NPP与碳循环研究中得到验证和广泛应用^[3,26,27]。但在官网下载的MOD17A3原始数据在沙漠区是空值,因此对于西北干旱区来说,数据缺失严重(图2),进而无法直接利用MOD17A3数据产品对西北干旱区的植被NPP进行相关方面的研究。因此,本研究以NDVI、气温、降水、海拔高度、潜在蒸散(Potential Evapotranspiration, PET)及太阳辐射等作为自变量,MOD17A3产品数据作为因变量进行回归建模,模拟研究区植被NPP的空间分布规律及特征。

对现有气象站点进行筛选,选择有MODIS NPP数据覆盖的38个站点(图1)。因此,针对研究区的自变量多重共线性问题,本研究采用逐步多元回归、主成分回归、偏最小二乘回归以及岭回归方法,建立生物量与各因子之间的回归模型。

2.2.2 数据来源及预处理

本研究采用的基础地理数据包括:

(1)2011年NASA公布全球空间分辨率为30m的DEM数据(ASTER 30m全球高程数据集)(<http://glovis.usgs.gov/index.shtml>)^[28]。

(2)来自NASA网站由Terra卫星MODIS传感器提供的,时间尺度是2000-2014年的MOD13A2 NDVI以及MOD17A3数据(<https://ladsweb.nascom.nasa.gov/data/order.html>)^[29]。MOD13A2 NDVI数据是一种时空连续的植被监测数据,时间和空间分辨率分别为16d和1km。MODIS NPP数据的时间空间分辨率分别为1年和1km。

(3)PET月数据是由东英吉利大学气候研究小组(CRU TS v.3.22)提供,空间分辨率为 0.5° ,时间分辨率为1个月,跨度是2000-2014年^[30]。

(4)气象站点数据来自中国气象数据网(<http://data.cma.cn/>)^[31]的西北干旱区具有完整观测序列的

2017年3月

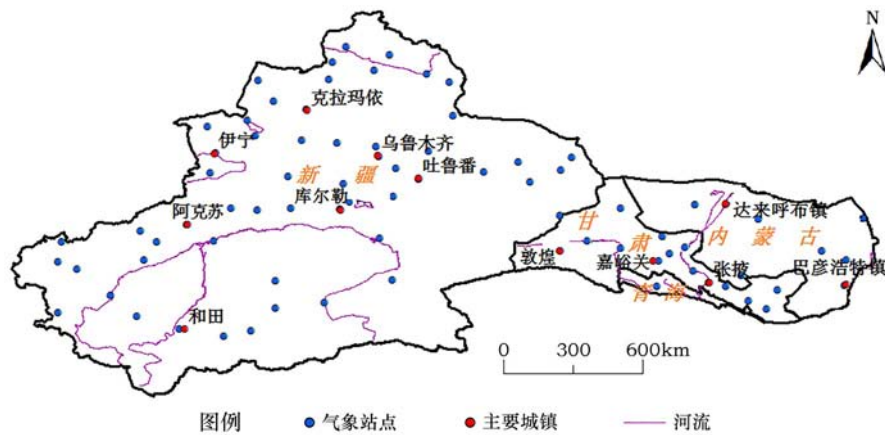


图1 研究区示意

Figure 1 Sketch of study area

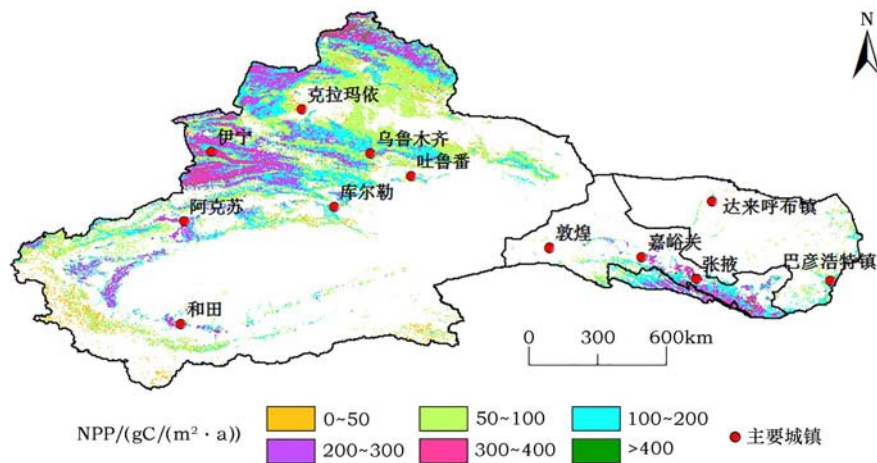


图2 2014年MODIS 17A3数据空间分布

Figure 2 Distribution of MODIS 17A3 data of 2014 in Northwest China

2000-2014年的日气象资料(包括日降水量、日平均气温、日照时数、海拔高程)。

对于MODIS NDVI和NPP遥感数据,首先利用MRT将研究区的影像进行拼接投影处理,进而采用ArcGIS和ENVI软件对NDVI数据进行MVC最大值合成,然后对数据进行格式转换、裁剪等预处理;对于全球CRU PET数据,在MTLAB软件中提取并进行裁剪处理;针对本研究是基于站点来进行回归反演研究,则需要在ArcGIS中把所有的栅格图层转换成站点图层,提取相应站点上的数据来进行回归分析。对于站点气象数据,将站点日数据经过计算得到年均温、年累计降水量以及年累计日照时数等。

2.3 研究方法

研究流程如图3所示。

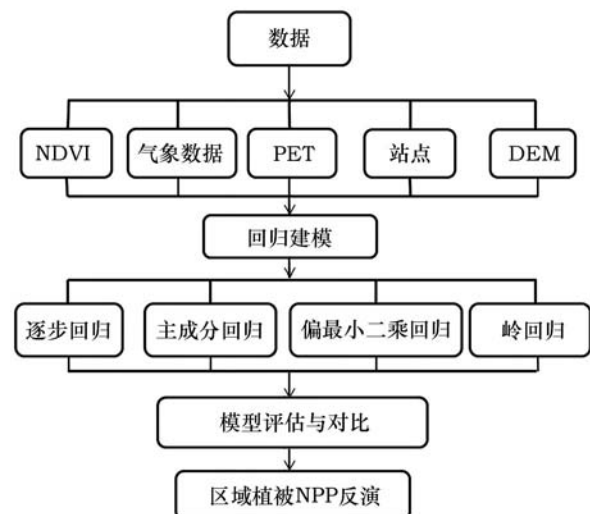


图3 技术路线

Figure 3 Study design

2.3.1 多元逐步回归建模

逐步回归分析的建模思路:根据自变量对因变量的重要性大小以及自变量之间的相关性进行因子筛选,将自变量逐个增加并剔除不重要变量,确定最终模型变量个数,引入方程进行回归分析,进而得到一个稳定可靠的回归模型^[19]。

2.3.2 主成分回归建模

Massy 于 1965 年根据主成分分析的思想提出了主成分回归,并成为回归分析中解决多重共线性比较有效的方法^[32]。建模思路是:从自变量系统中,以选取的主分量尽可能保留原始变量的信息且相互之间不相关为原则,提取主成分,形成新变量,再对新变量采用多元线性回归进行建模,并经过线性转化得到因变量与自变量之间的线性回归模型。在整个提取主成分的过程中未与因变量发生联系,是一种完全独立于因变量 Y 而进行的分析。

2.3.3 偏最小二乘回归建模

偏最小二乘法的建模思路:对于 P 个自变量和 1 个因变量分别构成矩阵系统 $X=[x_1, x_2, x_3, \dots, x_p]$ 和 Y 。在自变量中逐个选取若干彼此之间不相关且对整个系统有解释能力的新的综合变量来代替原有变量,因变量对 m 个主成分进行回归建模,并进行转化得到因变量与自变量之间的回归方程,完成建模,在一定程度上可以有效的消除自变量之间的相关性对模型系数的影响^[33]。

2.3.4 岭回归建模

岭回归是 1962 年由统计界的 A.E.Hoerl 提出,并在 1970 年系统发展的回归分析方法。被认为是求解病态方程的有效方法,是一种为减少均方误差而提出的有偏估计。岭回归公式为:

$$\hat{X}(k) = (X^T P X + k E)^{-1} X^T P Y \quad 0 \leq k < \infty \quad (1)$$

式中 $\hat{X}(k)$ 为参数 \hat{X} 的岭估计; k 为岭参数,一般为正小数; E 为单位矩阵; X 为自变量观测值矩阵; Y 为因变量观测值矩阵;从定义可以看出,岭估计是在最小二乘估计法方程系数阵 $X^T P X$ 的主对角线上加上一个常数 k ,进而改变法方程系数矩阵的态性,得到误差更低、精度更高、更为稳定的模型^[24]。

2.3.5 评价指标

由于研究区自变量之间具有相关性,本研究选择纳什效率系数 (*Nash-Sutcliffe, NS*)、模型判定系数

(R^2)、相对误差 (*PBIAS*)、标准均方根误差 (*SRMSE*)、拟合精度 (*PRECISION*) 来判断。其计算公式为:

$$NS = 1 - \frac{\sum_{i=1}^n (Y_i^{sim} - Y_i^{obs})^2}{\sum_{i=1}^n (Y_i^{obs} - Y_{mean}^{obs})^2} \quad (2)$$

$$R^2 = \left(\frac{\sum_{i=1}^n [(Y_i^{sim} - Y_{mean}^{sim})(Y_i^{obs} - Y_{mean}^{obs})]}{\sqrt{\sum_{i=1}^n (Y_i^{sim} - Y_{mean}^{sim})^2 \sum_{i=1}^n (Y_i^{obs} - Y_{mean}^{obs})^2}} \right)^2 \quad (3)$$

$$PBIAS = \frac{\sum_{i=1}^n (Y_i^{sim} - Y_i^{obs})}{\sum_{i=1}^n (Y_i^{obs})} \quad (4)$$

$$SRMSE = \frac{\sqrt{\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim})^2}{n}}}{\sqrt{\frac{\sum_{i=1}^n (Y_i^{obs} - Y_{mean}^{obs})^2}{n-1}}} \quad (5)$$

$$PRECISION = \frac{1}{n} \sum_{i=1}^n \left(1 - abs \left(\frac{Y_i^{sim} - Y_i^{obs}}{Y_i^{obs}} \right) \right) \times 100\% \quad (6)$$

式中 Y_i^{obs} 、 Y_i^{sim} 分别为第 i 个样本的观测值和模拟值; Y_{mean}^{obs} 、 Y_{mean}^{sim} 分别是样本观测值、模拟值均值。根据文献可知,当 NS 值 > 0.75 时,可认为模拟效果好; $0.36 < NS < 0.75$ 时,模拟效果为令人满意^[34]; R^2 为预测值与实测数据之间的拟合程度,数值越接近于 1 表明其可信度越高; $PBIAS$ 绝对值越小、 $SRMSE$ 越小模拟效果越好; $PRECISION$ 越高,越接近真值。

3 结果与分析

3.1 模型构建

3.1.1 多元逐步回归模型反演植被 NPP

选取 2000-2014 年的研究区站点,总共 570 个,其中 456 个站点数据用来构建模型,剩下的 114 个站点数据作为验证样本对得到的模型进行检验。因变量为 MODIS NPP 数据,年均 NDVI (x_1)、年累积降水量 (x_2)、年均温 (x_3)、PET (x_4)、日照时数 (x_5) 以及海拔高度 (x_6) 为自变量,构建模型。在 SPSS 中通过多元逐步回归将自变量逐个增加并剔除不重要变量,确定最终模型变量个数为 4 个时,模型回归系数的 R^2 (0.82) 和调整的 R^2 (0.819) 达到最大,并通过显著性检验 ($p < 0.05$),模型的系数如表 1 所示。拟合和预测结果分别见图 4a 和图 4b。

2017年3月

表1 回归模型系数、显著性及共线性检验结果

Table 1 Results of regression model coefficients, significance and collinearity

变量	非标准化系数	标准化系数	显著性	共线性
常数	71.073		0.011	
NDVI	701.609	0.878	0.000	1.021
降水	0.088	0.114	0.000	1.085
PET	-0.055	-0.096	0.000	1.028
日照时数	-0.025	-0.070	0.000	1.083

3.1.2 主成分回归模型反演生物量

选取因变量MOD17A3数据和上述6个自变量数据,利用SPSS及MATLAB软件,构建模型。根据徐建华等关于MATLAB中主成分回归程序的相关研究^[35],在MATLAB中对自变量进行主成分回归分析,提取累计贡献率为86%的前4个主成分,并计算得到载荷矩阵。然后对4个主成分进行回归分析,得到回归模型,利用载荷矩阵把标准化回归方程转化为线性回归方程。得到回归模型如下:

$$Y = 551.18x_1 - 0.162x_2 - 0.4047x_3 - 0.074x_4 + 0.131x_5 + 0.01x_6 - 256.86 \quad (7)$$

拟合和预测结果分别见图4a和图4b。

3.1.3 偏最小二乘回归模型反演生物量

将上述6个自变量与因变量(MOD17A3数据)组成矩阵导入MATLAB中,首先在MATLAB中将数据矩阵标准化;其次根据偏最小二乘法的原理,采用交叉有效性原则,在MATLAB中提取3个彼此之间不相关且对整个系统有解释能力的新的综合变量来代替原有变量;最后进行回归分析,并经过线性转换得到自变量与因变量之间的回归方程。参考文献见^[36]。提取的主成分有3个,并得到如下回归模型:

$$Y = 691.372x_1 + 0.042x_2 - 1.763x_3 - 0.056x_4 - 0.036x_5 - 0.0042x_6 + 132.5463 \quad (8)$$

拟合和预测结果分别见图4a和图4b。

3.1.4 岭回归模型反演生物量

根据岭回归的原理,为了改变岭回归中法方程的状态。在最小二乘法的基础上,加上一个 K 值,使得到的回归方程更加稳定,将上述6个自变量与因变量导入SPSS19.0中,调用SPSS软件中自带的岭回归程序文件(ridge regression. sps)中的代码,进行岭回归运算^[37],利用岭迹法判断常数 K 的取值($K=0.05$),并利用验证样本对模型进行分析与评价。SPSS输出模型的拟合结果见表2,得到的模型

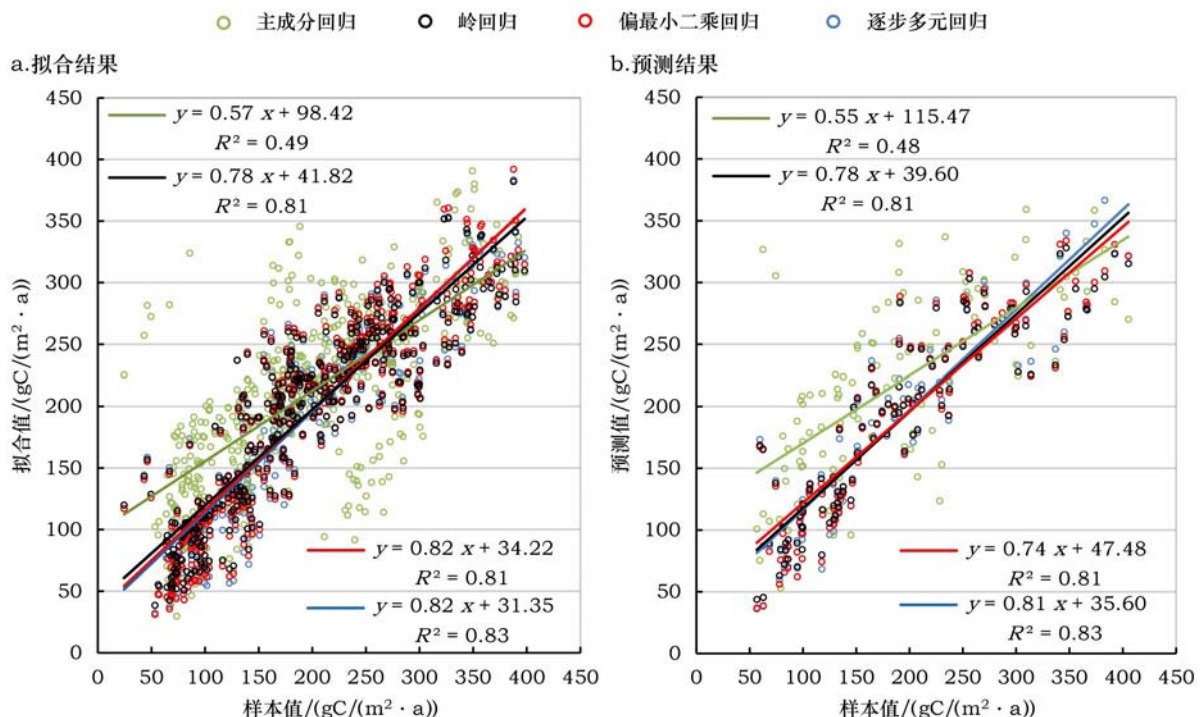


图4 回归模型(a)拟合值和(b)预测值与样本值对比

Figure 4 Comparison of (a) simulation results and (b) predicted results with the sample values

表2 岭回归模型描述

Table 2 Description of Ridge Regression model			
R^2	R^2_{adj}	SRMSE	sig
0.795	0.791	0.44	0.000

如下:

$$Y = 660.06x_1 + 0.053x_2 - 1.386x_3 - 0.051x_4 - 0.033x_5 - 0.035x_6 + 124.30 \quad (9)$$

拟合和预测结果分别见图4a和图4b。

3.2 模型的对比分析与验证

将456个和114个站点自变量数据分别代入上述得到的4种模型中,计算植被NPP,并与因变量(MOD17A3)数据进行对比分析,利用模型指标进行评估。模型拟合结果散点见图4,4种回归模型预测结果(多元线性回归、主成分回归、偏最小二乘回归、岭回归)的变化斜率为0.81、0.55、0.74、0.78, R^2 分别为0.83、0.48、0.81、0.81;针对模型预测结果,从模型评估指标分析(表3),4种模型的NS系数分别为0.85、0.40、0.81、0.81;精度分别为84.43%、60.32%、82.35%、82.63%,模型模拟值与MOD17A3产品值的PBIAS分别为-0.77、14.02、-1.24、-0.87, R^2 分别为0.85、0.50、0.81、0.81,SRMSE分别为0.39、0.77、0.44、0.44。综合模型评估指标以及拟合预测结果散点图发现,4种回归模型中主成分回归模型的模拟精度较低,回归结果不理想。鉴于此,通过阅读大量文献并深入分析模型机理发现,主成分回归方法虽然有效地降低了数据的维数,但是无法避免使新变量受到原始变量中重叠信息的影响,且无法辨识噪音与信息,影响模型的精度;另外在提取主成分时是撇开因变量进行的,造成主成分对因变量的解释能力会变弱,达不到理想的结果。

偏最小二乘回归是一种同时具备主成分分析、典型相关性分析和最小二乘回归的优点,预测精度

较高;岭回归分析是一种有偏估计方法;逐步回归法尽可能地将回归效果显著的自变量引入方程,这3种模型的NS系数均大于0.75,模型的 R^2 在0.8左右,拟合程度较高。但是从模型的简易程度以及模型需要的变量来看,逐步线性回归模型较为简单高效,且容易实现。在此基础上本研究利用逐步线性回归模型模拟2012-2014年西北干旱区38个气象站点的植被NPP值,并用MATLAB提取MOD17A3产品相应像元数据进行对比分析,发现模拟值与产品数据变化趋势基本一致(图5a),拟合程度较好(图5b), $R^2=0.827(p<0.01)$ 。

为了进一步验证模型的适用性,于2016年8月进行了野外采样。由于研究区面积较大,野外工作环境恶劣(尤其是沙漠区和高山积雪区),因此本文选取典型研究区(伊犁河谷)布设19个采样点(图5c),换算成地上、地下生物量,并与相应空间位置上模型模拟结果进行对比,检验模型精度。从(图5d)可知,模型结果与实测值相关性为0.749($p<0.01$),在0.01水平(双侧)上显著相关,模拟结果较好。因此我们可以得出逐步多元线性回归模型是一种可以简单、高效、准确地反演西北干旱区植被NPP的模型。

3.3 研究区植被NPP的反演

3.3.1 植被NPP的空间分布特征

在MATLAB、ArcGIS等软件的基础上,利用逐步回归模型反演西北干旱区2000-2014年的植被NPP(图6a,见第552页)。植被NPP高的地区主要分布在水热条件较好的山区,如阿尔泰山、天山北坡森林区、塔里木河流域、伊犁河谷、昆仑山西段及祁连山等地区,总体上呈现出北部及西北部高,南部及东南部低的特征。伊犁河谷地区的植被NPP值最高,阿尔泰山区、塔河地区、昆仑山西段等高山

表3 不同建模方法的评估指标比较

Table 3 Comparisons of evaluation indexes by different models								
	逐步回归模型		主成分回归		偏最小二乘回归		岭回归	
	拟合	预测	拟合	预测	拟合	预测	拟合	预测
NS	0.83	0.85	0.40	0.40	0.80	0.81	0.79	0.81
Precision	79.94	84.43	55.06	60.32	77.83	82.35	78.11	82.63
PBIAS	-0.68	-0.77	9.42	14.02	0.23	-1.24	-0.35	-0.87
R^2	0.82	0.85	0.45	0.50	0.80	0.81	0.80	0.81
SRMSE	0.42	0.39	0.78	0.77	0.45	0.44	0.45	0.44

2017年3月

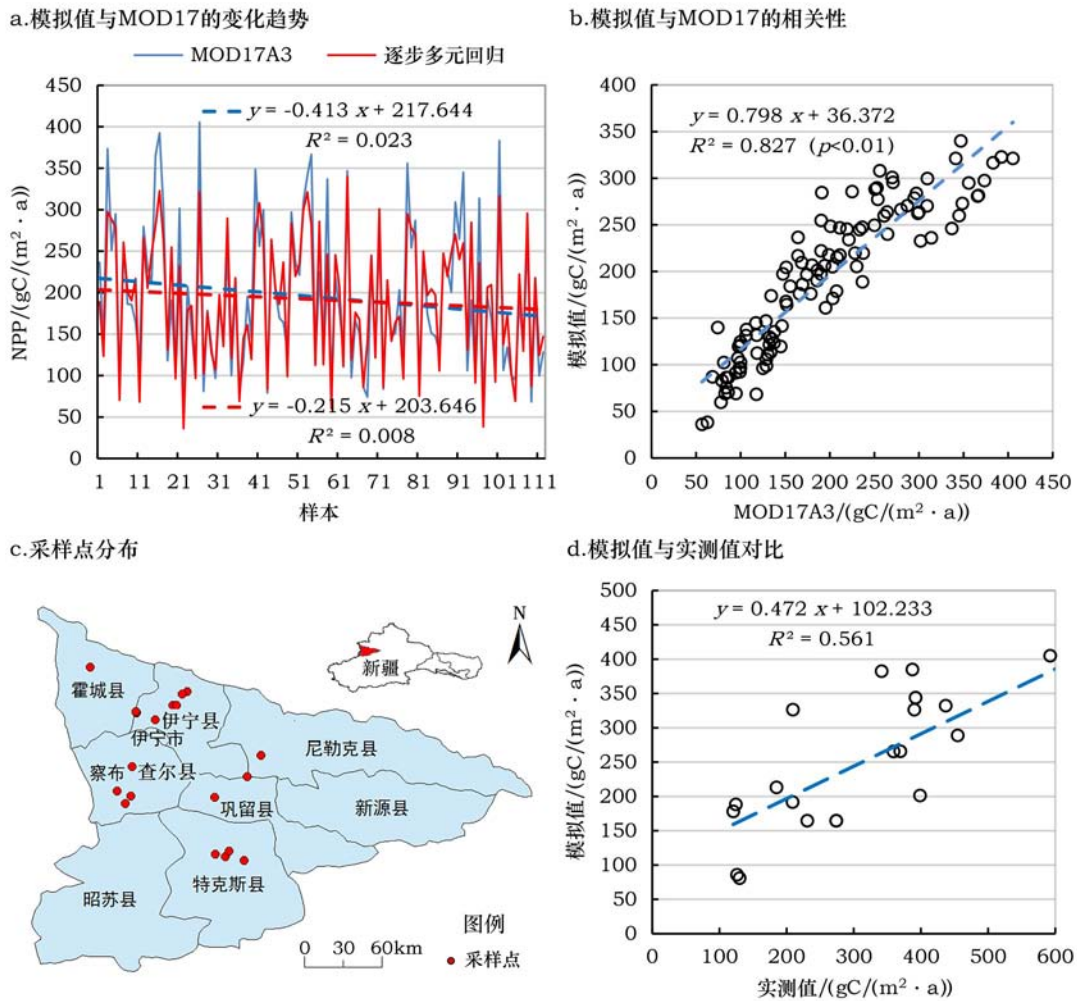


图5 逐步回归模型模拟结果的对比与验证分析

Figure 5 The comparison and verification of the simulated results by Multi-stepwise regression model

地区植被NPP的量次之,南疆的塔克拉玛干沙漠地区及河西走廊,植被覆盖率极低,NPP低,甚至低于 $10.00\text{gC}/(\text{m}^2 \cdot \text{a})$ 。

3.3.2 植被NPP的变化特征

从研究区2000-2014年植被NPP的年际变化看(图6c),西北干旱区NPP在波动中呈增加趋势,线性增长率为 $0.40\text{gC}/(\text{m}^2 \cdot \text{a})$ 。植被NPP从2000年的 $46.04\text{gC}/\text{m}^2$ 增加至2014年的 $51.587\text{gC}/\text{m}^2$,其中,最大值为2012年的 $56.85\text{gC}/\text{m}^2$ 。从西北干旱区植被NPP的年际波动来看,植被NPP 2000-2009年变化趋势较为平稳,线性增长斜率为 $0.22\text{gC}/(\text{m}^2 \cdot \text{a})$;2009年以后呈增加的趋势,线性增长斜率为 $0.45\text{gC}/(\text{m}^2 \cdot \text{a})$ 。

从研究区植被NPP时间序列的空间变化趋势(图6b)看,西北干旱区2000-2014年间植被NPP变

化除了伊犁河谷地区植被出现明显的退化现象外,其余地区均表现出“好的愈好,劣的愈劣”的状态。在水热条件较好的地区,如阿尔泰山北坡、塔城地区、昆仑山北坡、甘肃东部等地区植被NPP增幅较高,多数NPP碳值变化大于 $1.50\text{gC}/(\text{m}^2 \cdot \text{a})$;天山北坡、塔里木河流域以及甘肃河西走廊地区植被NPP碳值变化甚至超过 $3.00\text{gC}/(\text{m}^2 \cdot \text{a})$;伊犁谷地、天山山脉以南部分地区、阿尔泰山南部、塔克拉玛干沙漠东部、昆仑山西段、甘肃西部等地区植被NPP出现减少的趋势,部分地区植被NPP减少的趋势显著,植被NPP碳值变化甚至超过 $3.00\text{gC}/(\text{m}^2 \cdot \text{a})$ 。

从整体上来看,西北干旱区2000-2014年植被NPP呈现改善的趋势,58.66%的地区呈现出不同程度的增长趋势,13.64%的地区保持不变,27.7%的地区出现不同程度的植被退化现象。

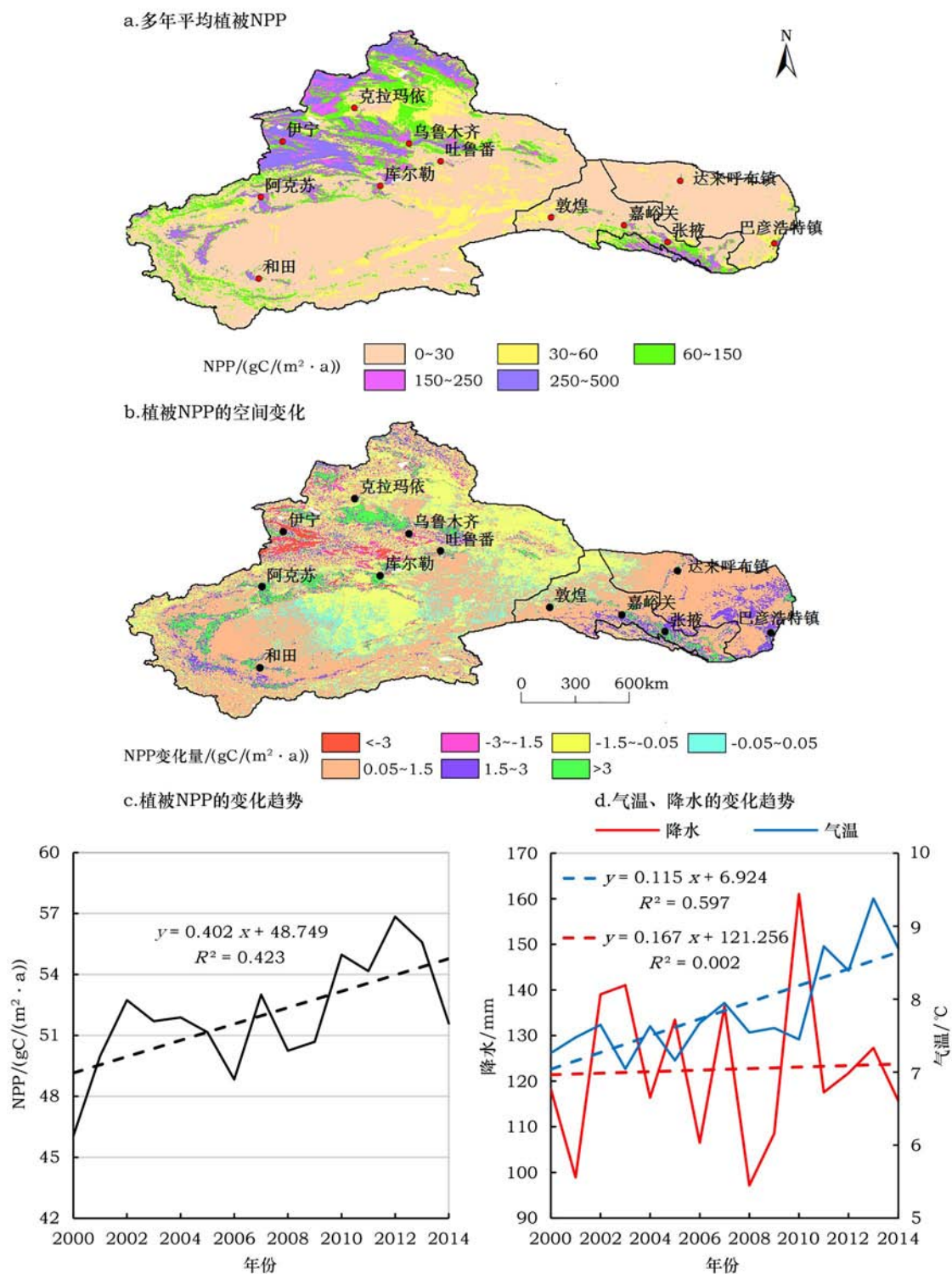


图6 2000–2014年西北干旱区植被 NPP 时空变化格局及气温、降水的变化趋势分析

Figure 6 The spatial-temporal change pattern of vegetation NPP and the changing trend of precipitation and temperature in Northwest China from 2000 to 2014

4 讨论

通过对比分析研究区气温、降水的变化趋势发现(图6d),植被 NPP 与降水和气温均呈显著地正相

关,相关系数分别为 0.475 ($p<0.05$) 和 0.504 ($p<0.05$),偏相关系数分别为 0.660 ($p<0.01$) 和 0.720 ($p<0.01$)。近年来,西北干旱区多年平均气温为

2017年3月

7.84℃,年降水量在80~180mm之间,年降水量波动较大,且降水量的区域差异显著,山区降水量较多,平原区降水稀少。低海拔平原区,干旱少雨,降水量的增加会导致植被NPP增多,气温升高或太阳辐射加强会加速土壤中水分的流失^[10,11,38],使植被生长受到胁迫,进而限制植被的生长^[39,40];而海拔较高的山区,在地形作用的影响下,降水多、气温相对较低,随着气温升高,高山冰川消融加剧,一定程度补充了山区的水资源供给^[41],促进植被生长,使植被NPP增加。总之,植被NPP的影响因素很多,水热组合状况、植被类型的地带性差异、不同类型植被对水分的要求以及植被的生理生态特征等均会影响植被NPP的变化,进而导致研究区植被NPP的分布存在空间差异。

另外,本研究模型建模过程中存在一定的不足。首先,选用有MOD17A3数据覆盖的气象站点,南疆沙漠区站点较少,影响模型的建模精度;其次,气象数据的插值精度和质量有待提高。本文虽然采用具有较强的统计学意义的克里金插值方法,将站点气象数据插值到面上,但是气象数据往往与所处的地理位置及下垫面性质关系密切,插值过程往往不可避免的出现误差,因此发展更为可靠的插值方法,对植被NPP反演精度的提高具有重要意义。

目前对于干旱半干旱区碳循环的相关性研究还不够深入,在过去的50年,研究区冬季增温明显,极端最低气温升高了约3.11℃^[10]。蒸发潜力增大、冰雪融水加剧,给干旱区生态系统稳定性带来很大的挑战^[42]。气候及土地利用变化会导致物种发生变化,进而在一定程度上影响生态系统生产力,且气候变化对植被的影响具有时滞效应^[4,43,44]。因此植被对不同季节降水量、气候的时滞效应以及土地覆被变化对生物量影响的定量研究有待进行深入的分析。

5 结论

本文采用逐步线性回归、主成分回归、偏最小二乘回归以及岭回归等四种方法构建模型,探讨了西北干旱区NPP时空变化,并对其进行了评估,得到如下结论:

(1)综合比较4种回归模型的模型指标及拟合散点图可知,逐步回归模型被认为是一种可以简

单、高效、准确地反演西北干旱区植被NPP的模型,且模拟结果与MOD17A3产品及实测值具有很强的相关性。

(2)西北干旱区NPP自2000年以来呈波动增加趋势,线性增长率为0.40gC/(m²·a)。其中,在2000-2009年间,植被NPP变化趋势较为平稳,线性增长斜率为0.22gC/(m²·a);2009年以后呈明显增加趋势,线性增长斜率为0.45gC/(m²·a)。

(3)西北干旱区NPP变化的空间分布表现为周边山区(阿尔泰山、伊犁河谷、天山北坡、昆仑山西段)的生物量较高,而平原盆地及塔克拉玛干沙漠地区植被覆盖率较低,尤其塔克拉玛干沙漠低于10gC/(m²·a)。在过去的10余年间,西北干旱区58.66%区域的植被NPP呈现增长趋势,13.64%的地区保持不变,27.7%的地区表现为退化趋势。退化的区域主要集中在沙漠边缘及戈壁地带。

参考文献(References):

- [1] Pan S F, Tian H Q, Dangal S, *et al.* Impacts of climate variability and extremes on global net primary production in the first decade of the 21st century[J]. *Journal of Geographical Sciences*, 2015, 25(9):1027-1044.
- [2] Li X, Zhu Z, Zeng H, *et al.* Estimation of gross primary production in China(1982-2010)with multiple ecosystem models[J]. *Ecological Modelling*, 2016, 324:33-44.
- [3] Li Z, Chen Y N, Wang Y, *et al.* Dynamic changes in terrestrial net primary production and their effects on evapotranspiration[J]. *Hydrology and Earth System Sciences*, 2016, 20(6):2169-2178.
- [4] Guo Q, Hu Z, Li S, *et al.* Contrasting responses of gross primary productivity to precipitation events in a water-limited and a temperature-limited grassland ecosystem[J]. *Agricultural and Forest Meteorology*, 2015, 214-215(3):169-177.
- [5] Flanagan L B, Sharp E J, Gamon J A. Application of the photosynthetic light-use efficiency model in a Northern Great Plains grassland[J]. *Remote Sensing of Environment*, 2015, 168(6):239-251.
- [6] Beck P S A, Judayg P, Alix C, *et al.* Changes in forest productivity across Alaska consistent with biome shift[J]. *Ecology Letters*, 2011, 14(4):373-379.
- [7] Van B R, Haneca K, Hoogesteger J, *et al.* A century of tree line changes in sub-Arctic Sweden shows local and regional varia-

- bility and only a minor influence of 20th century climate warming[J]. *Journal of Biogeography*, 2011, 38(5): 907-921.
- [8] Mekonnen Z A, Grant R F, Schwalm C. Contrasting changes in gross primary productivity of different regions of North America as affected by warming in recent decades[J]. *Agricultural and Forest Meteorology*, 2016, 218-219(1-2): 50-64.
- [9] Williams A P, Allen C D, Millar C I, et al. Forest responses to increasing aridity and warmth in the southwestern United States [J]. *Proceedings of the National Academy of Sciences*, 2010, 107(50): 21289-21294.
- [10] 陈亚宁, 李稚, 范煜婷, 等. 西北干旱区气候变化对水文水资源影响研究进展[J]. *地理学报*, 2014, 69(9): 1295-1304. [Chen Y N, Li Z, Fan Y T, et al. Research progress on the impact of climate change on water resources in the arid region of North-west China[J]. *Journal of Geographical Sciences*, 2014, 69(9): 1295-1304.]
- [11] 陈亚宁, 杨青, 罗毅, 等. 西北干旱区水资源问题研究思考[J]. *干旱区地理*, 2012, 35(1): 1-9. [Chen Y N, Yang Q, Luo Y, et al. Ponder on the issues of water resources in arid region of northwest China[J]. *Arid Land Geography*, 2012, 35(1): 1-9.]
- [12] 罗天祥, 李文华, 冷允法, 等. 青藏高原自然植被总生物量的反演与净初级生产量的潜在分布[J]. *地理研究*, 1998, 17(4): 337-344. [Luo T X, Li W H, Leng Y F, et al. Estimation of total biomass and potential distribution of net primary productivity in the Tibetan plateau[J]. *Geographical Research*, 1998, 17(4): 337-344.]
- [13] 方精云. 全球生态学[M]. 北京: 高等教育出版社, 2000. [Fang J Y. *Global Ecology*[M]. Beijing: Higher Education Press, 2000.]
- [14] Lieth H. Modeling the primary productivity of the world[J]. *Indian Forester*, 1972, 98(6): 327-331.
- [15] 高志强, 刘纪远. 中国植被净生产力的比较研究[J]. *科学通报*, 2008, 53(3): 317-326. [Gao Z Q, Liu J Y. A comparative study of the net primary productivity in China[J]. *Chinese Science Bulletin*, 2008, 53(3): 317-326.]
- [16] 朱文泉, 潘耀忠, 龙中华, 等. 基于GIS和RS的区域陆地植被NPP反演-以中国内蒙古为例[J]. *遥感学报*, 2005, 9(3): 300-307. [Zhu W Q, Pan Y Z, Long Z H, et al. Estimating net primary productivity of terrestrial vegetation based on GIS and RS: A case study in Inner Mongolia, China[J]. *Journal of Remote Sensing*, 2005, 9(3): 300-307.]
- [17] Liang W, Yang Y T, Fan D M, et al. Analysis of spatial and temporal patterns of net primary production and their climate controls in China from 1982 to 2010[J]. *Agricultural and Forest Meteorology*, 2015, 204: 22-36.
- [18] 杨存建, 刘纪远, 张增祥. 热带森林植被生物量遥感反演探讨[J]. *地理与地理信息科学*, 2004, 20(6): 22-25. [Yang C J, Liu J Y, Zhang Z X. Study on the estimation of the tropical forest vegetation biomass by using remote sensing data, meteorological data and topographical data[J]. *Geography and Geo-Information Science*, 2004, 20(6): 22-25.]
- [19] Fan W Y, Zhang H Y, Yu Y, et al. Comparison of three models of forest biomass estimation[J]. *Chinese Journal of Plant Ecology*, 2011, 35(4): 402-410.
- [20] 张旭琛, 牛华忠, 钟华平, 等. 新疆伊犁地区草地植被地上生物量遥感反演[J]. *草业学报*, 2015, 24(6): 25-34. [Zhang X S, Niu H Z, Zhong H P, et al. Assessment of above-ground Biomass of grassland using remote sensing, Ili, Xinjiang[J]. *Acta Prataculturae Sinica*, 2015, 24(6): 25-34.]
- [21] Lu Y M, Yue T X, Chen C F, et al. Solar radiation modeling based on stepwise regression analysis in China[J]. *Journal of Remote Sensing*, 2010, 14(5): 852-864.
- [22] 胡馨月, 高明希, 任玉, 等. 基于主成分回归的温度分布反演研究[J]. *光谱学与光谱分析*, 2012, 32(10): 2789-2793. [Hu X Y, Gao M X, Ren Y, et al. Study on inversion of temperature distribution based on principal components analysis[J]. *Spectroscopy and Spectral Analysis*, 2012, 32(10): 2789-2793.]
- [23] Bin J, Ai F F, Liu N, et al. Supervised principal components: A new method for multivariate spectral analysis[J]. *Journal of Chemometrics*, 2013, 27(12): 457-465.
- [24] 高思远, 崔晨风, 范玉平. 基于岭估计的青海省东部农业区ET0遥感反演研究[J]. *自然资源学报*, 2016, 31(4): 693-702. [Gao S Y, Cui C F, Fan Y P. Remote sensing inversion of ET0 in eastern agricultural area of Qinghai Province based on ridge estimation[J]. *Journal of Natural Resource*, 2016, 31(4): 693-702.]
- [25] 张杰, 潘晓玲, 高志强, 等. 干旱生态系统净初级生产力反演及变化探测[J]. *地理学报*, 2006, 61(1): 15-25. [Zhan J, Pan X L, Gao Z Q, et al. Satellite estimates and change detection of net primary productivity of oasis-desert based on ecosystem process with remotely sensed forcing in Arid Western China[J]. *Journal of Geographical Sciences*, 2006, 61(1): 15-25.]
- [26] 崔林丽, 杜华强, 史军, 等. 中国东南部植被NPP的时空格局变化及其与气候的关系研究[J]. *地理科学*, 2016, 36(5): 787-793. [Cui L L, Du H Q, Shi J, et al. Spatial and temporal pattern of vegetation NPP and its relationship with climate in the Southeastern China[J]. *Journal of Geographical Sciences*, 2016, 36(5): 787-793.]
- [27] 吴珊珊, 姚治君, 姜丽光, 等. 基于MODIS的长江源植被NPP时空变化特征及其水文效应[J]. *自然资源学报*, 2016, 31(1): 39-50. [Wu S S, Yao Z J, Jiang L G, et al. The spatial-temporal variations and hydrological effects of vegetation NPP based on MODIS in the source region of the Yangtze River[J]. *Journal of Natural Resources*, 2016, 31(1): 39-50.]
- [28] San B T, Suzen M L. Digital elevation model (DEM) generation

2017年3月

- and accuracy assessment from ASTER stereo data[J]. *International Journal of Remote Sensing*, 2005, 26(22): 5013-5027.
- [29] Gu Y, Hunt E, Wardlow B, *et al.* Evaluation of MODIS NDVI and NDWI for vegetation drought monitoring using Oklahoma Mesonet soil moisture data[J]. *Geophysical Research Letters*, 2008, 35(22): 1092-1104.
- [30] Harris I, Jones P D, Osborn T J, *et al.* Updated high-resolution grids of monthly climatic observations-the CRU TS3.10 Dataset [J]. *International Journal of Climatology*, 2014, 34(3): 623-642.
- [31] 中国气象数据网[EB/OL]. (2015-05)[2016-08-22]. http://data.cma.cn/data/detail/dataCode/SURF_CLI_CHN_MUL_MON_650.html. [China Meteorological Data Service Centre[EB/OL]. (2015- 05) [2016- 08- 22]. http://data.cma.cn/data/detail/dataCode/SURF_CLI_CHN_MUL_MON_650.html.]
- [32] Sun J G. A correlation principal component regression analysis of NIR data[J]. *Journal of Chemometrics*, 1995, 9(1): 21-29.
- [33] Gusnanto A, Pawitan Y, Huang J, *et al.* Variable selection in random calibration of near-infrared instruments: Ridge regression and partial least squares regression settings[J]. *Journal of Chemometrics*, 2003, 17(3): 174-185.
- [34] Pan S F, Tian H Q, Dangal S, *et al.* Modeling and monitoring terrestrial primary production in a changing global environment: Toward a multiscale synthesis of observation and simulation[J]. *Advances in Meteorology*, 2014, doi: 10.1155/2014/965936.
- [35] 徐建华. 计量地理学[M]. 北京: 高等教育出版社, 2005. [Xu J H. Quantity Geography [M]. Beijing: Higher Education Press, 2005.]
- [36] 曾九孙, 刘祥官, 罗世华. 主成分回归和偏最小二乘法在高炉冶炼中的应用[J]. 浙江大学学报(理学), 2009, 36(1): 33-36. [Zeng J S, Liu X G, Luo S H. Application of principal component regression and partial least square in blast furnace iron-making[J]. *Journal of Zhejiang University (Science Edition)*, 2009, 36(1): 33-36.]
- [37] 尹康. 常用统计软件关于岭回归计算原理的比较分析[J]. 统计研究, 2013, 30(2): 109-112. [Yi K. Comparison and analysis of the principle of ridge regression in common statistical software [J]. *Statistical Research*, 2013, 30(2): 109-112.]
- [38] Jobbagy E G, Sala O E, Paruelo J M. Patterns and controls of primary production in the Patagonian steppe: A remote sensing approach[J]. *Ecology*, 2002, 83(2): 307-319.
- [39] Yang Y T, Long D, Guan H D, *et al.* GRACE satellite observed hydrological controls on interannual and seasonal variability in surface greenness over mainland Australia[J]. *Journal of Geophysical Research: Biogeosciences*, 2014, 119(12): 2245-2260.
- [40] 李鹏飞, 孙小明, 赵昕奕. 近50年中国干旱半干旱地区降水量与潜在蒸散量分析[J]. 干旱区资源与环境, 2012, 26(7): 57-63. [Li P F, Sun X M, Zhao X Y. Analysis of precipitation and potential evapotranspiration in arid and semiarid area of China in recent 50 years[J]. *Journal of Arid Land Resources and Environment*, 2012, 26(7): 57-63.]
- [41] 姚俊强, 杨青, 陈亚宁, 等. 西北干旱区气候变化及其对生态环境影响[J]. 生态学杂志, 2013, 32(5): 1283-1291. [Yao J Q, Yang Q, Chen Y N, *et al.* Climate change in arid areas of Northwest China in past 50 years and its effects on the local ecological environment[J]. *Chinese Journal of Ecology*, 2013, 32(5): 1283-1291.]
- [42] Chen Y N, Li Z, Fan Y T, *et al.* Progress and prospects of climate change impacts on hydrology in the arid region of northwest China[J]. *Environ Res*, 2015, 139: 11-19.
- [43] Arredondo T, garcia-Moya E, Huber-Sannwald E, *et al.* Drought manipulation and its direct and legacy effects on productivity of a monodominant and mixed-species semi-arid grassland[J]. *Agricultural and Forest Meteorology*, 2016, 223: 132-140.
- [44] 李霞, 李晓兵, 陈云浩, 等. 中国北方草原植被对气象因子的时滞响应[J]. 植物生态学报, 2007, 31(6): 1054-1062. [Li X, Li X B, Chen Y H, *et al.* Temporal responses of vegetation to climate variables in temperate steppe of northern China[J]. *Journal of Plant Ecology*, 2007, 31(6): 1054-1062.]

Inversion of net primary productivity in the arid region of Northwest China based on various regressions

JIAO Wei^{1,2}, CHEN Yaning¹, LI Zhi¹, LI Yupeng^{1,2}, HUANG Xiaoran^{1,2}, LI Haixia^{1,2}

(1. State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography,

Chinese Academy of Sciences, Urumqi 830011, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Vegetation Net Primary Productivity (NPP) is an important parameter when evaluating terrestrial ecosystems and provides a significant reference for research into global carbon cycles. Based on MODIS data and meteorological station data from 2000 to 2014, we used multi-stepwise regression, principal components regression, partial least-squares regression and ridge regression and estimated vegetation NPP and temporal-spatial distribution patterns in the arid region of Northwest China. We found that the multi-stepwise regression model was superior and the model simulation results were coordinated with measured values in the spatial distribution. A multi-stepwise regression model can be used to retrieve NPP in arid and semi-arid areas and reflects vegetation growth and distribution in the study area. There were obvious regional differences in NPP distribution in the arid region of Northwest China; vegetation NPP in the mountains was increased and decreased in the plains. High values were found in the north, northwest and southeast of the arid region and low values in the south and southeast. At an annual scale, NPP in Northwest China slightly increased from 2000 to 2014 at a rate of $0.40 \text{ gC}/(\text{m}^2 \cdot \text{a})$. Trends in annual vegetation NPP are different among areas, vegetation NPP in 58.66% areas increased, 13.64% of areas remained relatively stable, and 27.7% of areas slightly decreased from 2000.

Key words: net primary productivity; multi-stepwise regression; principal components regression; partial least-squares regression; ridge regression; Northwest China