

DOI:10.11829/j.issn.1001-0629.2018-0091

刘艳, 聂磊, 杨耘. NDVI 相似性分区下天山地区草地总产草量遥感估算. 草业科学, 2018, 35(7): 1754-1764.

Liu Y, Nie L, Yang Y. Estimation of the total production of the herbage in the Tianshan Mountain Area using remote sensing technology with NDVI similarity zoning. Pratacultural Science, 2018, 35(7): 1754-1764.

NDVI 相似性分区下天山地区草地 总产草量遥感估算

刘 艳^{1,2}, 聂 磊³, 杨 耘⁴

(1. 中国气象局乌鲁木齐沙漠气象研究所, 新疆 乌鲁木齐 830002; 2. 中亚大气科学研究中心, 新疆 乌鲁木齐 830002;
3. 武汉大学测绘遥感信息工程国家重点实验室, 湖北 武汉 430079; 4. 长安大学地质工程与测绘学院, 陕西 西安 710054)

摘要:山区草地总产草量遥感估算是定量评价区域牧业生产力的有效手段。常规总产草量地面观测数据准确性较高, 但无法覆盖整个天山山区, 且耗时耗力。针对此问题, 以新疆天山山区为研究区, 选取 MODIS/MOD13Q1 250 m 植被指数(normalized difference vegetation index, NDVI)产品数据, 以县(市)为单元, 基于巴氏距离定量评价研究区植被指数分布区域相似性以得到有效遥感建模分区, 在此分区基础上, 结合草地总产草量实测数据, 建立研究区植被指数—草地总产草量遥感估算模型。结果显示, 1) 基于各县(市)2009—2015 年 7 月底至 8 月初植被生长期多年 NDVI 均值直方图计算巴氏距离, 以巴氏距离 $d > 0.5$ 为阈值, 研究区被划分为 7 个遥感建模区; 2) 各分区内 NDVI—草地总产草量数据拟合方程形式不同, 有线性、指数、幂指数和多项式回归方程几种形式。总体来看, 各分区 NDVI—草地总产草量拟合相关系数在 0.784~0.836。交叉检验除天山北坡西段—伊犁河谷草原畜牧业区 RMSE 值在 2 951 kg·hm⁻² 外, 其他分区 RMSE 值均在 266~928 kg·hm⁻², 原因在于伊犁河谷草原畜牧业区实测草地总产量在 10 000~30 000 kg·hm⁻² 的样点居多, 区域草地总产量较其他区域多。

关键词:天山山区; 草地总产草量; 建模分区; 遥感估算模型; 空间分布

中图分类号: S812; S127

文献标志码: A

文章编号: 1001-0629(2018)07-1754-11*

Estimation of the total production of the herbage in the Tianshan Mountain Area using remote sensing technology with NDVI similarity zoning

Liu Yan^{1,2}, Nie Lei³, Yang Yun⁴

(1. Institute of Desert Meteorology, CMA, Urumqi 830002, Xinjiang, China;

2. Center of central Asia atmospheric science research, Urumqi 830002, Xinjiang, China;

3. State Key Laboratory of Information Engineering in Surveying, Mapping and
Remote Sensing, Wuhan University, Wuhan 430079, Hubei, China;

4. College of Geology Engineering and Geomatics, Chang'an University, Xi'an 710054, Shaanxi, China)

Abstract: An estimation of the total production of herbage in Xinjiang using remote sensing technology is an effective method for quantitative evaluation of regional animal husbandry productivity. It is time-consuming and cannot cover the whole area of the Tianshan Mountains, although the measured accuracy of the total herbage yield using conventional means is very high. To address this problem, we used MODIS/MOD13Q1 vegetation index products, with 250 m GSD as experimental data and city or county as a basic unit for analysis. The Bhattacharyya distance was used to quantitatively evaluate the distribution of the similarity of the vegetation index

* 收稿日期: 2018-02-09 接受日期: 2018-03-21

基金项目: 中央级公益性科研院所基本科研业务费专项资金项目(IDM2016004); 风云三号(02)批气象卫星地面应用系统工程应用示范系统项目(FY-3(02)-UDS-1.5.1); NSFC-新疆联合基金(U1703121)

通讯作者: 刘艳(1978-), 女, 河南扶沟人, 副研究员, 硕士, 主要从事遥感在积雪、植被变化监测等方面的研究。E-mail: liuyan@idm.cn

in the Tianshan Mountain Area as a study case. The purpose was to obtain an effective remote sensing modeling zoning, and then to construct an estimation model of the total production of herbage with respect to the vegetation index using remote sensing technology. Finally, the spatial distribution and feature analysis of the total herbage yield (fresh weight) in the Tianshan Mountains from the year 2009 to 2015 were determined based on the analysis of the spatial distribution and characteristics obtained under the GIS platform. The results showed the following: the seven modeling zones were derived from an analysis of the mean histogram of NDVI data collected during the optimal period (i.e., July and August each year) of vegetation growth for each city or county in the study area via the Bhattacharyya distance with a threshold greater than 0.5. Secondly, the constructed estimation model of the total production of the herbage showed different fitting relationships to the vegetation index, and there were three forms, including the exponential, power index, and a unary regression equation with the second order. On the whole, the fitting correlation coefficient of the constructed estimation model could reach between 0.754 and 0.836 for each zones. The RMSE value of cross-validation in the northern slope of the Tianshan Mountains-Yili Valley was $2\,951\text{ kg}\cdot\text{ha}^{-1}$, and the RMSE value were between 266 and $928\text{ kg}\cdot\text{ha}^{-1}$ in the other zones. This was because more measured samples which total production of the herbage was between 10 000 and $30\,000\text{ kg}\cdot\text{ha}^{-1}$ were collected in the zone. Also there was higher total production of the regional herbage in the zone than others.

Key words: Tianshan Mountains Area; total production of the herbage; modeling zoning; remote sensing estimation model; spatial distribution

Corresponding author: Liu Yan E-mail:liuyan@idm.cn

新疆现有牧草地总面积 $5\,116.07\text{ 万 hm}^2$, 占全国牧草地总面积的 19.52% ^[1]。北疆地区天然草地有效面积约 $2\,377.52\text{ 万 hm}^2$, 占全疆 46.47% , 是新疆草原畜牧业的重要基地^[2]。建立草原产草量监测模型以准确、及时地估算草原产草量对确定载畜量、合理安排畜牧业生产和维护草原生态平衡等具有重要意义。新疆牧草产量估算研究主要为 90 年代初开展的 NOAA/AVHRR 卫星数据和天然草地牧草产量关系模型研究^[2-7]; 地面测产数据和同步 MODIS 遥感数据集合的天山北坡乌鲁木齐草甸、草原和荒漠草原草地牧草估产模型^[8]; 遥感分类结果和产量估测模型集合的阜康市 1990—2008 年近 20 年草地总面积和总产量趋势变化^[9]; 1991—1996 年新疆天山北坡阜康市内不同草地类型实测草地可食产量、环境与遥感资料等综合的多重相关分析和遥感估产技术研究^[10]。这些研究主要集中在 1990—2000 年; 研究区域有限且集中分布在天山北坡的乌鲁木齐(南郊)、阜康和阿勒泰; 模型采样点少且时空分布不均, 致使模型代表性不足; 模型主要针对一个观测年进行, 时间序列性不强。仅仅使用一个观测年内高寒草甸、山地草甸、山地草原等 7 种不同牧草类型逐月产量观测数据, 进行产量变化特征分析和实现单一牧草类型产量计算。新疆草地分布错综复杂, 上述模型代表性和时效性都非常有限。NOAA/NDVI 植被指数产品在空间和光谱分辨率上劣于

MODIS/NDVI 植被指数产品, 相对获取不便, 特别是近年来的植被指数产品^[11]。MODIS/NDVI 植被指数产品空间分辨率有 250 m、500 m 和 1 km, 共 3 种, 有单日、16 日和月合成 3 类数据, 产品时空分辨率能很好地反映草地植被时空变化和估算草地生物量^[12-18]。因此, 根据前人已有研究成果不足, 考虑到如何有效合理地利用有限样点数据, 本研究选取 MOD13Q1 归一化植被指数(NDVI)16 d 合成产品为试验数据, 从遥感监测植被指数分布特征入手, 基于计算机图像直方图识别手段定量评价遥感监测植被指数的区域相似性, 以县(市)为研究单元, 计算多年植被指数均值直方图巴氏距离, 定量评价各单元及其临近单元 NDVI 频数直方图相似性以得到有效遥感分区, 以此分区为基础, 建立 NDVI—草地产草量统计模型, 将其推广至临近具有相似性 NDVI 的区域, 从而获取整个天山山区牧草(鲜重)总产草量遥感估算模型。

1 数据来源与研究方法

1.1 研究区概况

按草地植被型组归纳新疆草地有草原、荒漠、草甸、沼泽 4 个类组。草原草地除阿尔泰山南麓冲积扇上部和天山北麓山前倾斜平原存在平原荒漠草原片段外都发育在山地。荒漠草地在新疆分布广、面积大, 包括温性草原化荒漠、温性荒漠和高寒荒漠类。草甸草

地分布面积仅次于荒漠和草原草地,包括低地草甸、山地草甸和高寒草甸 3 种类型。沼泽草地分布面积不大,分别占新疆草地面积和可利用土地面积的 0.47% 和 0.50%,主要分布在南北疆各大河流下游或河口、河漫滩、湖泊周围、山前潜水溢出带山间低位盆地的湖滨和山地高位盆地等,植被是由湿中生多年生根茎禾草和莎草类组成。由表 1 和图 1 可见,本研究区域草地在类型上具有一定的典型性和完整性,在空间分布上具有一定的连续性,可作为草地总产草量遥感估算的典型试验区域。

表 1 研究区 4 类典型草地统计面积及比例

Table 1 Area and proportion of four kinds of typical grasslands in the study area

草地类型	面积	比例
Grassland type	Area/km ²	Proportion/%
草甸草地 Meadow	50 515	18.07
草原草地 Grassland	48 038	17.19
荒漠草地 Desert	60 939	21.80
沼泽草地 Swamp	1 263	0.45
研究区域 Research area	279 505	100.00

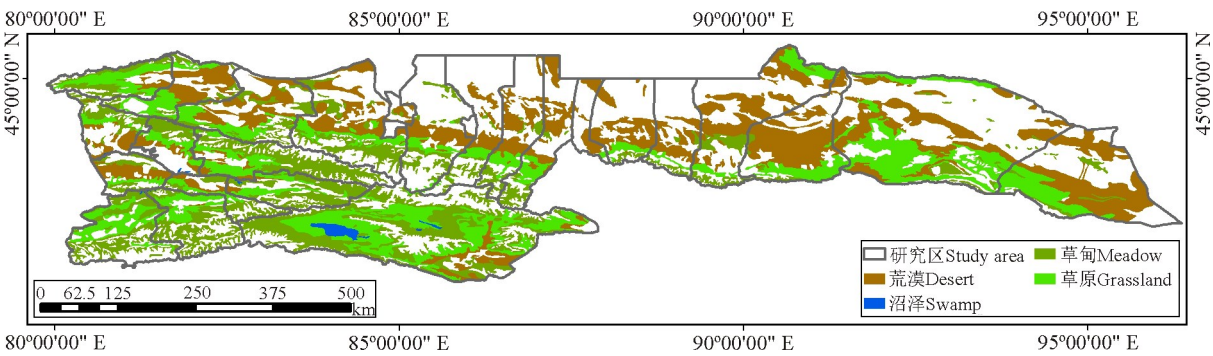


图 1 研究区 4 类典型草地空间分布图

Fig. 1 Spatial distribution of four kinds of typical grasslands in the study area

1.2 数据来源

数据来源包含地面监测和卫星遥感监测数据。地面监测主要包括草地总产草量观测,观测集中在 2009—2015 年 7 月底至 8 月初(图 2)。样地选择在相应群落典型地段且生境条件、植物群落种类组成、群落结构、利用方式和利用强度等具有相对一致性;样地间具有异质性,每个样地能控制的最大范围内地貌、植被等具有同质性,即地貌及植被生长状况相似。草原植被样地面积不小于 100 hm²。矮小草本和小半灌木草地样方面积为草原 1 m²,荒漠草原 2 m²,草甸 0.5 m²;株丛小的灌

木和高大草本测产面积为 4 m×4 m,灌丛或半灌木、高大草本分布均匀的地段设 20 m×5 m 或 10 m×10 m 样方。设置一个灌木样方,灌木样方内设置一个草本样方。每个样方要求测定植物名称、植物盖度和草地总产草量。遥感监测植被指数数据选用美国国家宇航局 250 m 的 MOD13Q1 归一化植被指数(NDVI) 16 d 合成产品,时间是 2009 年 1 月至 2015 年 12 月,轨道号 h23v04、h23v05、h24v04 和 h24v05,共计 644 景影像。

1.3 研究方法

1.3.1 植被指数处理 利用MODIS数据重投影工具

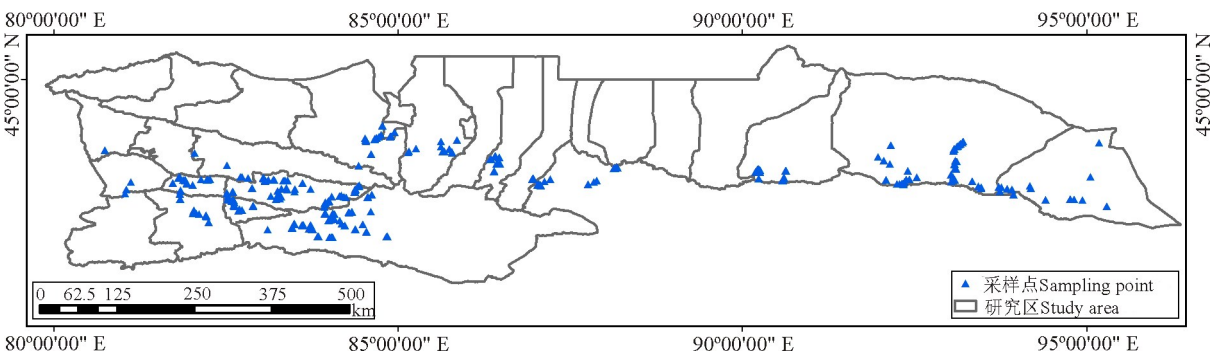


图 2 2009—2015 年 7 月底至 8 月初牧草总产量实测位置分布图

Fig. 2 Spatial distribution of the measured locations of the total production of the herbage from the end of July to the beginning of August during the years 2009—2015

(MODIS reprojection tools, MRT)对 MOD13Q1 数据集进行批量 HDF-TIF 格式和定义 WGS84 投影,应用 GIS 平台计算 2009—2015 年 7 月底至 8 月初多年 NDVI 均值和最大值合成法(maximum value composition, MVC)合成 2009—2015 年 7 月底至 8 月初最大 NDVI,以表征对应年草地总产草量实测期内牧草平均和最好长势情况,基于 GPS 测点信息提取地面采样点对应 NDVI。

1.3.2 巴氏距离计算 以县(市)为研究单元,统计 2009—2015 年 7 月底至 8 月初时段多年 NDVI 均值频数直方图,利用巴氏距离(Bhattacharyya distance)测量各县(市)及其临近县(市)NDVI 直方图分布的相似性。巴氏距离定义为^[19-21]:

$$d(H_1, H_2) = 1 - \sqrt{1 - \frac{\sum_I \sqrt{H_1(I) \cdot H_2(I)}}{\sqrt{H_1 \cdot H_2} N^2}} \quad (1)$$

式中: H_1 与 H_2 为县(市)尺度多年 NDVI 均值频数直方图, N 为频数直方图组数, I 为频数直方图组编号, $H_1(I)$ 为县(市)频数直方图中组数 I 对应的频数。 $d(H_1, H_2)$ 越接近 1 说明两个直方图越相似,即所在县(市)被划分为一个遥感分区。本研究中,定义 $d > 0.5$,即为相似区域。

1.3.3 模型精度验证 应用留一交叉验证(leave-one-out cross validation, LOOCV)对草地总产草量遥感模型模拟结果进行验证^[22]。假设有 n 个样本,从中选择一个观测值作为验证数据,其他 $n-1$ 个样本作为训练样本来建立回归模型,用回归模型模拟选出的单个验证数据的预测值来检验模型精度,如此重复 n 次,用 n 个结果平均值来衡量模型模拟精度。每个模型预测能力由均方根误差(RMSE)及拟合系数 R^2 决定。RMSE 用来量化模型精度, R^2 用来评估模型的准确性。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n [Fw(y_i) - y_i]^2}{n}} \quad (2)$$

式中: $Fw(y_i)$ 表示第 i 个样方草地总产草量(鲜重)实测值, y_i 为模型模拟第 i 个预测值, n 是观测样本总数。RMSE 越小表明回归模型越精确, R^2 越接近于 1 表示模型精度越高。

2 基于巴氏距离的 NDVI 均值相似性分析和建模分区

各县(市)间 NDVI 均值巴氏距离计算结果如表 2 所列,将巴氏距离大于 0.5 的县市定义为相似县(市),

研究区域被分为 7 个建模区域(图 3)。

各建模区域所含县(市)多年 NDVI 均值—频数直方图如图 4 所示,可见在地形、气候条件、土壤环境等方面具有很大相似性的区域应该具有相似的牧草生物量,从而使得该区域 NDVI 值也相近。而对于地形、气候条件、土壤环境等方面空间差异较大的地区,其牧草生物量也呈现较大的差异性,对应区域 NDVI 值也具有大的离散度。因此,本研究依据 NDVI 值的局部相似性准则进行牧草生物量建模分区是合理的。

对各个分区分别采用指数、线性、多项式以及幂函数 4 类函数建立植被指数与草地总产草量回归模型(表 4)。结果显示:I、Ⅲ、Ⅵ分区选择幂指数函数、Ⅱ、Ⅳ、Ⅶ分区选择指数函数、V 分区采用多项式进行建模具有最小的拟合误差。这可能是因为不同分区,气候、地形、牧草生长环境等因子存在差异,使得不同分区 NDVI 空间分布差异较大,从而使得产草量随 NDVI 的变化速率及趋势也不同。因此,利用 NDVI 进行建模分区,不同分区采用不同的拟合模型才能更精确地估算整个研究区的牧草产草量。

如图 5a 所示,分区 I 采用幂指数模型拟合草地总产草量与 NDVI 之间的函数关系,效果最佳,能体现牧草鲜重随 NDVI 呈明显的快速递增现象。该分区内 NDVI 在 0.25~0.85 变化,表明该分区内大部分区域草地长势较好、产草量较高。图 5b 所示,分区 II 采用指数模型拟合草地总产草量与 NDVI 之间的函数关系效果较好,能够很好地反映牧草鲜重随 NDVI 值增长的变化趋势。该分区内 NDVI 主要在 0.04~0.65 间变化。当少数子区域 NDVI 增加至 0.7 左右时,对应产草量骤然变大。当 NDVI 在 0.04~0.50 内时牧草鲜重随 NDVI 呈近似线性变化。图 5c 所示,分区 III 采用幂指数模型拟合草地总产草量与 NDVI 之间的函数关系效果较好,能够反映牧草鲜重随 NDVI 的变化趋势,该分区内 NDVI 主要在 0.10~0.47 变化,但是,少数子区域产草量超出分区产量均值,异常大。如图 5d 所示,分区 IV 采用指数模型拟合草地总产草量与 NDVI 之间的函数关系,误差最小,该分区内 NDVI 值大都大于 0.25。但是,存在一些子区域的产草量异常大。图 5e 所示,分区 V 采用多项式拟合模型拟合草地总产草量与 NDVI 之间的函数关系,误差最小。该分区内 NDVI 值都大于 0.05,牧草鲜重随 NDVI 值变化速度较慢。图 5f 所示,分区 VI 采用幂指数模型拟

表 2 归一化 NDVI 均值—频数直方图巴氏距离和所属建模分区

Table 2 The Bhattacharyya distance of the normalized mean-frequency histogram of NDVI data and the corresponding modeling zone in the study area

建模分区	县(市)	察布查尔	伊宁县	霍城县	伊宁市	尼勒克县	巩留县	特克斯县	新源县	昭苏县
Modeling zone	County/City	Chabhar	Yining	Huocheng	Yining	Nilka	Gongliu	Tekesi	Xinyuan	Zhaosu
		County	County	County	City	County	County	County	County	County
I	察布查尔 Chabhar County	1.00	0.66	0.82	0.66	0.70	0.72	0.76	0.51	0.63
	伊宁县 Yining County	0.66	1.00	0.77	0.66	0.78	0.78	0.71	0.72	0.69
	霍城县 Huocheng County	0.82	0.77	1.00	0.64	0.82	0.86	0.86	0.64	0.75
	伊宁市 Yining City	0.66	0.66	0.64	1.00	0.60	0.58	0.56	0.47	0.51
	尼勒克县 Nilka County	0.70	0.78	0.82	0.60	1.00	0.80	0.84	0.75	0.86
	巩留县 Gongliu County	0.72	0.78	0.86	0.58	0.80	1.00	0.85	0.72	0.75
	特克斯县 TekesiCounty	0.76	0.71	0.86	0.56	0.84	0.85	1.00	0.69	0.82
	新源县 Xinyuan County	0.51	0.72	0.64	0.47	0.75	0.72	0.69	1.00	0.74
	昭苏县 Zhaosu County	0.63	0.69	0.75	0.51	0.86	0.75	0.82	0.74	1.00
建模分区	县(市)	乌苏市			博乐市		温泉县		精河县	
Modeling zone	County/City	Wusu City			Bole City		Wenquan County		Jinghe County	
II	乌苏市 Wusu City	1.00			0.79		0.75		0.80	
	博乐市 Bole City	0.79			1.00		0.76		0.66	
	温泉县 Wenquan County	0.75			0.76		1.00		0.67	
	精河县 Jinghe County	0.80			0.66		0.67		1.00	
建模分区	县(市)	石河子市		沙湾县		玛纳斯县		呼图壁县		奎屯市
Modeling zone	County/City	Shihezi City		Shawan County		Manas County		Hutubi County		Kuitun City
III	石河子市 Shihezi City	1.00		0.71		0.57		0.58		0.71
	沙湾县 Shawan County	0.71		1.00		0.78		0.73		0.64
	玛纳斯县 Manas County	0.57		0.78		1.00		0.73		0.54
	呼图壁县 Hutubi County	0.58		0.73		0.73		1.00		0.59
	奎屯市 Kuitun City	0.71		0.64		0.54		0.59		1.00
建模分区	县(市)	昌吉市			米泉市			阜康市		
Modeling zone	County/City	Changji City			Miquan City			Fukang City		
IV	昌吉市 Changji City	1.00			0.81			0.77		
	米泉市 Miquan City	0.81			1.00			0.83		
	阜康市 Fukang City	0.77			0.83			1.00		
建模分区	县(市)	奇台县			木垒自治县			吉木萨尔县		
Modeling zone	County/City	Qitai County			Mulei Autonomous County			Jimsar County		
V	奇台县 Qitai County	1.00			0.75			0.73		
	木垒自治县	0.75			1.00			0.74		
	Mulei Autonomous County	0.73			0.74			1.00		
	吉木萨尔县 Jimsar County	0.73			0.74			1.00		
建模分区	县(市)	巴里坤县			伊吾县					
Modeling zone	County/City	Barlikun County			Yiwu County					
VI	巴里坤县 Barlikun County	1.00			0.76					
	伊吾县 Yiwu County	0.76			1.00					
建模分区	县(市)	和静县								
Modeling zone	County/City	Hejing County								
VII	和静县 Hejing County	1.0								

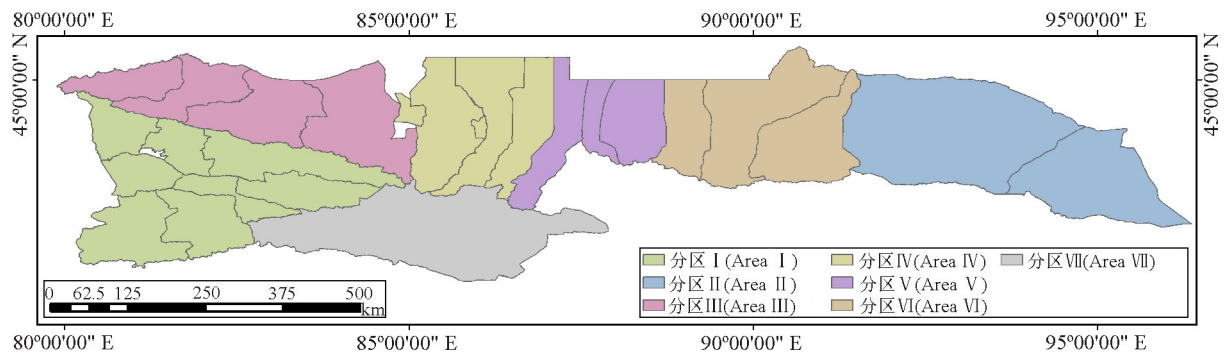


图 3 基于巴氏距离的 NDVI 时间序列相似性分区示意图

Fig. 3 Diagram of similarity zoning of NDVI time serial data based on the Bhattacharyya distance

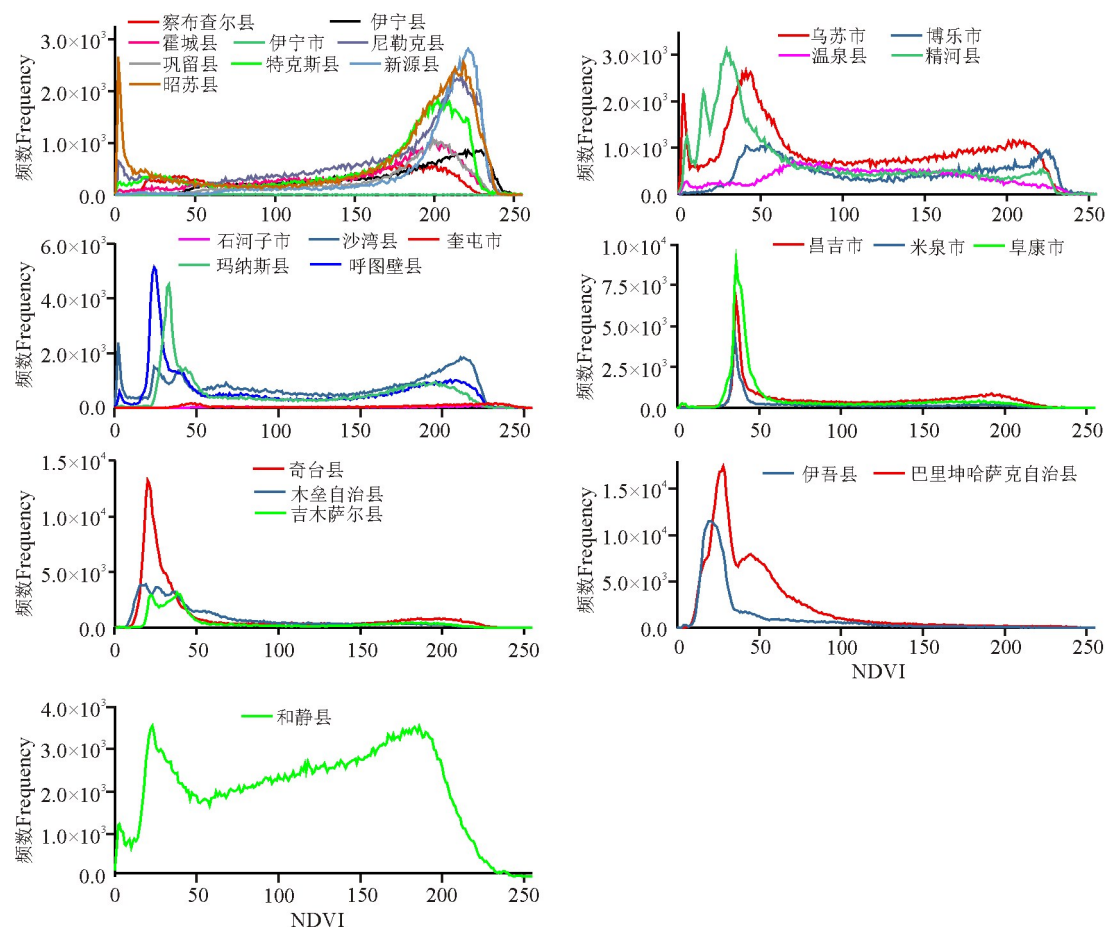


图 4 归一化 NDVI 均值-频数直方图

Fig. 4 The normalized mean-frequency histogram of NDVI data in the study area

察布查尔 Chabhar County,伊宁县 Yining County,霍城县 Huocheng County,伊宁市 Yining City,尼勒克县 Nilka County,巩留县 Gongliu County,特克斯县 Tekesi County,新源县 Xinyuan County,昭苏县 Zhaosu County,乌苏市 Wusu City,博乐市 Bole City,温泉县 Wenquan County,精河县 Jinghe County,石河子市 Shihezi City,沙湾县 Shawan County,玛纳斯县 Manas County,呼图壁县 Hutubi County,奎屯市 Kuitun City,昌吉市 Changji City,米泉市 Miquan City,阜康市 Fukang City,奇台县 Qitai County,木垒自治县 Mulei Autonomous County,吉木萨尔县 Jimsar County,巴里坤县 Barlikun County,伊吾县 Yiwu County,和静 Hejing County.

合草地总产草量与 NDVI 之间的函数关系,误差最小。该分区内 NDVI 值都大于 0.1,当 NDVI 在 0.1~0.3 牧草鲜重随 NDVI 呈近似的线性增加。图 5g 所示,分区 VII 采用指数模型拟合草地总产草量与 NDVI 之间的函数关系是合适的,该分区内 NDVI 都大于 0.3,主要集中在 0.3~0.8,牧草鲜重随 NDVI 上升的速度较快。

3 讨论

3.1 与不分区建模对比

忽略 NDVI 的空间异质性,采用研究区内所有采样点建立 NDVI—草地总产草量回归模型(表 5),与考虑 NDVI 空间异质性采用相似性分区的回归模型对比,以探究相似性分区回归模型在各个区域建模的提升效果。可以看出,除线性回归模型决定系数

R^2 为 0.389,其余回归模型的决定系数 R^2 均在 0.570~0.600,最高的幂指数模型为 0.600。同时,除线性回归模型的估计误差 $RMSE$ 大于 2 000 $\text{kg} \cdot \text{hm}^{-2}$,为 2 205.219 $\text{kg} \cdot \text{hm}^{-2}$ 外,其余回归模型的估计误差 $RMSE$ 均在 2 000 $\text{kg} \cdot \text{hm}^{-2}$ 以下,最低为幂指数模型,为 1 478.640 $\text{kg} \cdot \text{hm}^{-2}$ 。结合 R^2 与 $RMSE$ 两个指标,可知适合全区域产草量估算的回归模型为幂指数模型。

表 4 NDVI-草地总产草量回归模型及交叉检验结果列表
Table 4 List of regression models and the results of cross validation of total production of the herbage-NDVI

分区 Zone	回归模型 Modeling type	回归系数 Regression coefficients	R^2	调整 R^2 Adjusted correlation coefficient	$RMSE/$ ($\text{kg} \cdot \text{hm}^{-2}$)	样点数 Sample number
I	$y = ae^{bx} + c$	$a = 748.278; b = 0.416; c = -1\,480.075$	0.692 3	0.686 8	2 956.008	115
	$y = ax + b$	$a = 2\,919.790; b = -7\,871.904$	0.637 4	0.634 2	3 447.071	
	$y = a + bx + cx^2$	$a = 2\,749.592; b = -2\,228.115; c = 520.651$	0.688 4	0.682 8	2 946.565	
	$y = ax^b + c$	$a = 27.225; b = 3.143; c = 164.135$	0.690 4	0.684 9	2 951.245	
II	$y = ae^{bx} + c$	$a = 0.099\,3; b = 1.656; c = 970.249$	0.699 1	0.693 1	878.249	104
	$y = ax + b$	$a = 754.154; b = -430.20$	0.358 9	0.352 6	1 319.649	
	$y = a + bx + cx^2$	$a = 2\,619.598; b = -1\,865.506; c = 363.745$	0.545 6	0.536 6	1 190.839	
	$y = ax^b + c$	$a = 6.368\text{E}-06; b = 10.915; c = 996.389$	0.698 6	0.692 6	879.623	
III	$y = ae^{bx} + c$	$a = 166.206; b = 0.393; c = 3.438$	0.693 3	0.676 3	327.564	39
	$y = ax + b$	$a = 355.045; b = -50.744$	0.624 3	0.614 1	358.947	
	$y = a + bx + cx^2$	$a = 887.640; b = -245.462; c = 76.127$	0.699 8	0.683 2	328.892	
	$y = ax^b + c$	$a = 10.929; b = 2.746; c = 635.403$	0.698 8	0.682 1	326.929	
IV	$y = ae^{bx} + c$	$a = 0.060\,1; b = 1.686; c = 875.501$	0.640 2	0.629 2	574.803	68
	$y = ax + b$	$a = 598.760\,8; b = -1\,289.092$	0.404 1	0.395 1	686.268	
	$y = a + bx + cx^2$	$a = 4\,228.837; b = -2\,134.950; c = 310.400$	0.527 8	0.513 3	650.753	
	$y = ax^b + c$	$a = 5.315\text{E}-06; b = 10.837; c = 922.885$	0.633 6	0.622 3	584.147	
V	$y = ae^{bx} + c$	$a = 165.239; b = 0.383; c = 578.310$	0.570 0	0.547 7	392.536	42
	$y = ax + b$	$a = 273.982; b = 401.964$	0.500 4	0.487 9	413.103	
	$y = a + bx + cx^2$	$a = 1\,103.633; b = -249.927; c = 74.293$	0.591 0	0.570 0	375.322	
	$y = ax^b + c$	$a = 12.938; b = 2.618; c = 839.286$	0.581 6	0.560 1	381.863	
VI	$y = ae^{bx} + c$	$a = 1.812; b = 0.748; c = 5.871$	0.673 1	0.666 1	268.850	96
	$y = ax + b$	$a = 276.191; b = 130.997$	0.511 4	0.506 2	328.578	
	$y = a + bx + cx^2$	$a = 1\,069.387; b = -435.972; c = 103.711$	0.648 4	0.640 8	274.912	
	$y = ax^b + c$	$a = 0.600; b = 4.394; c = 653.542$	0.672 6	0.665 6	266.568	
VII	$y = ae^{bx} + c$	$a = 253.621; b = 0.417; c = 258.955$	0.616 2	0.610 5	928.710	136
	$y = ax + b$	$a = 1\,182.533; b = -3\,351.688$	0.565 5	0.562 3	1 000.695	
	$y = a + bx + cx^2$	$a = 3\,412.629; b = -1\,399.207; c = 233.912$	0.612 8	0.607 0	937.632	
	$y = ax^b + c$	$a = 5.068; b = 3.429; c = 1\,015.085$	0.615 3	0.609 5	932.101	

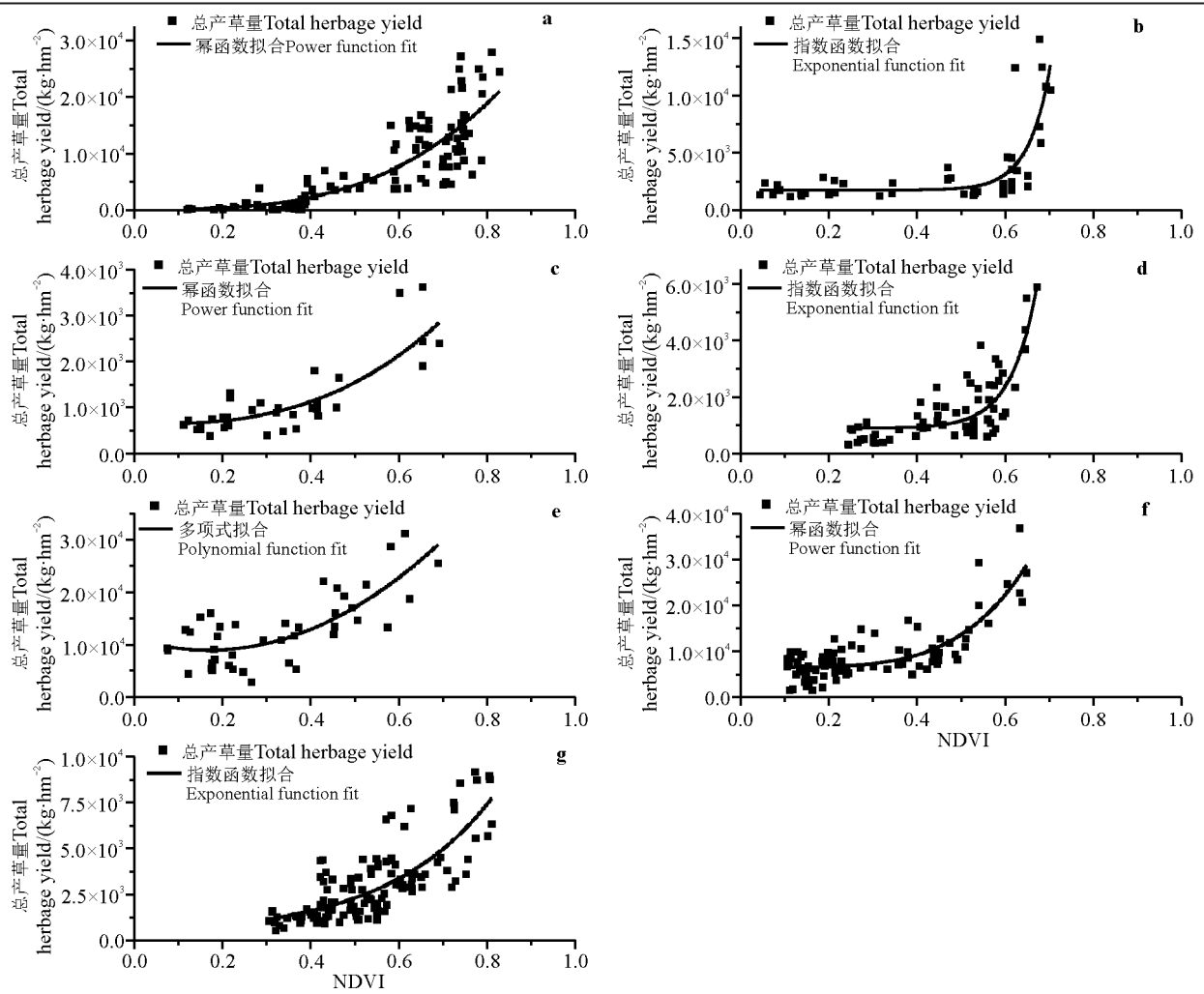


图 5 各个建模分区草地总产草量—NDVI 拟合关系图

Fig. 5 The fitting relation curves of the total production of the herbage—NDVI for each modeling zone

表 5 不分区情况下 NDVI—草地总产草量回归模型及交叉检验结果

Table 5 The regression models and results of cross validation of the total production of the herbage—NDVI in the unzoned case					
回归模型 Modeling type	回归系数 Regression coefficients	R^2	调整 R^2 Adjusted correlation coefficient	RMSE/ ($\text{kg} \cdot \text{hm}^{-2}$)	样点数 Sample number
$y = ae^{bx} + c$	$a = 57,185; b = 0.703; c = 425.156$	0.597	0.596	1 492.482	600
$y = ax + b$	$a = 1\,330,417; b = -2\,619,260$	0.389	0.389	2 205.219	
$y = a + bx + cx^2$	$a = 3\,820,722; b = -2\,682,155; c = 482.320$	0.573	0.571	1 680.967	
$y = ax^b + c$	$a = 0.430; b = 5.035; c = 750.253$	0.600	0.599	1 478.640	

对比表 4 中各分区的最优模型以及表 5 中不分区情况下的最优模型,不分区回归模型的决定系数 R^2 低于各个分区回归模型。说明采用 NDVI 相似性分区后,回归模型自变量能解释更多因变量的变异,更适合用于牧草产量的估算。同时,除分区 I 外,其余分区回归模型的估计误差 RMSE 均在 $1\,000\text{ kg} \cdot \text{hm}^{-2}$ 以

内,最低可到 $266\text{ kg} \cdot \text{hm}^{-2}$,均优于不分区回归模型。但分区 I 回归模型的估计误差达到了 $2\,951.910\text{ kg} \cdot \text{hm}^{-2}$,这主要是因为分区 I 内实测草地总产量在 $10\,000 \sim 30\,000\text{ kg} \cdot \text{hm}^{-2}$ 的样点多;而不分区回归模型估计误差低于分区 I 的原因是研究区内其余区域的实测草地总产草量较低,低于 $10\,000\text{ kg} \cdot \text{hm}^{-2}$,由交

叉验证原理可知,较低的实测值可将较高的模型估计误差平均。因此,分区回归模型能在各个区域内提升牧草产量的估算。

3.2 与随机分区建模对比

分区建模能够在各个区域内很好提升牧草产量的估算精度。根据采样点的聚集程度进行随机分区(图 6),并在各个随机分区内建立 NDVI—草地总产草量回归模型(表 5),以探究相似性分区回归建模对牧草产量估算的提升。可以看出,随机分区 I 的最佳回归模型为幂指数模型,其决定系数 R^2 为 0.549,估计误差 $RMSE$ 为 $2\,508.562\text{ kg}\cdot\text{hm}^{-2}$ (表 6);随机分区 II 的最佳回归模型为抛物线模型,其 R^2 为 0.524,

$RMSE$ 为 $479.584\text{ kg}\cdot\text{hm}^{-2}$;随机分区 III 最佳回归模型为指数模型,其模型决定系数 R^2 达到了 0.692,模型的估计误差为 $604.101\text{ kg}\cdot\text{hm}^{-2}$ 。

结合表 5 和表 6,随机分区后模型对因变量变异的解释程度提升不大,只有随机分区 III 回归模型决定系数大于不分区回归模型,这表明分区建模并不总是带来好的估算结果。同时,模型估计精度 $RMSE$ 也只在随机分区 II 与随机分区 III 低于不分区回归模型,表明分区建模对产草量估算的提升具有区域性。结合表 4 和表 6,采用相似性分区后,回归模型在模型决定系数 R^2 、模型估计精度 $RMSE$ 两个指标上都有显著提升,除了分区相似性分区 I 回归模型的 $RMSE$ 。

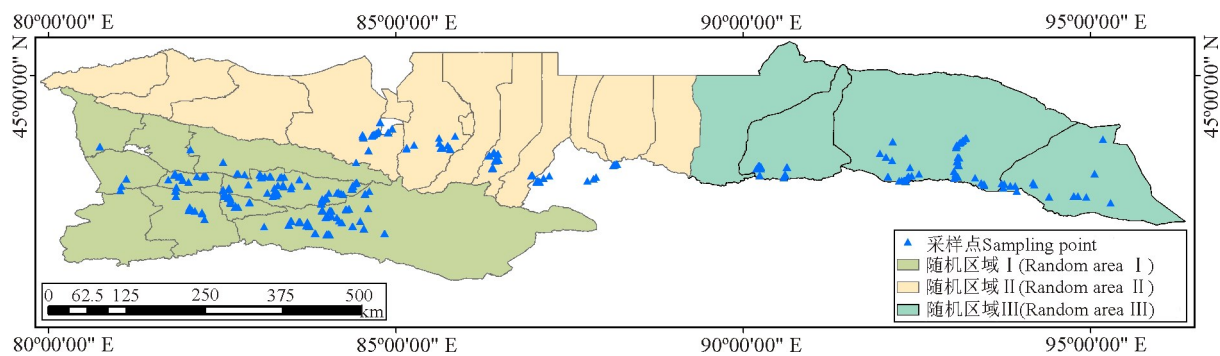


图 6 NDVI 时间序列随机分区示意图

Fig. 6 Diagram of random zoning of NDVI time serial data

表 6 各个随机分区 NDVI—草地总产草量回归模型及交叉检验结果列表

Table 6 List of regression models and results of cross validation of the total production of the herbage—NDVI in each random zone

随机分区 Random zone	回归模型 Modeling type	回归系数 Regression coefficients	R^2	调整 R^2 Adjusted correlation coefficient	$RMSE$ / ($\text{kg}\cdot\text{hm}^{-2}$)	样点数 Sample number
I	$y = ae^{bx} + c$	$a = 333.549; b = 0.485; c = -690.408$	0.547	0.544	2 516.493	251
	$y = ax + b$	$a = 2\,383.587; b = -7.565\text{E}+03$	0.470	0.468	2 951.910	
	$y = a + bx + cx^2$	$a = 4\,929.915; b = -3\,038.495; c = 529.284$	0.545	0.541	2 552.954	
	$y = ax^b + c$	$a = 5.675; b = 3.782; c = 459.552$	0.549	0.545	2 508.562	
II	$y = ae^{bx} + c$	$a = 21.182; b = 0.720\,6; c = 686.125$	0.529	0.522	501.783	149
	$y = ax + b$	$a = 366.143; b = -82.785$	0.404	0.400	540.456	
	$y = a + bx + cx^2$	$a = 1\,530.595; b = -662.416; c = 133.370$	0.524	0.517	497.584	
	$y = ax^b + c$	$a = 0.533; b = 4.449; c = 776.787$	0.530	0.524	500.581	
III	$y = ae^{bx} + c$	$a = 0.036\,0; b = 1.802; c = 860.036$	0.692	0.689	604.101	200
	$y = ax + b$	$a = 615.016; b = -389.069$	0.322	0.318	915.760	
	$y = a + bx + cx^2$	$a = 2\,441.925; b = -1\,643.751; c = 316.854$	0.520	0.515	821.791	
	$y = ax^b + c$	$a = 8.301\text{E}-07; b = 11.965; c = 882.226$	0.689	0.686	610.138	

4 结论

不同类型草地因其种类构成、长势、盖度、产量差异引起卫星植被指数的变化,使得充分利用卫星资料进行大面积周而复始地监测草地生产力动态变化成为可能^[5]。遥感估算天山地区草地总产量是一项具有重要实用价值的工作。天山北坡草地垂直分异明显,类型复杂,给遥感识别分类和监测带来了很大的困难,本研究提出的 NDVI 相似性分区下天山山区草地总产草量遥感估算技术实用可行,为后期高级估算方法开发打下了良好基础,将研究区域分为若干个分区可以减少估产模型的异质性,使得回归模型中的误差项完全符合理想假设(误差独立同分布期望为零,从而提高模型对数据的整体拟合度)。根据 NDVI 直方图显著的空间差异性特征和邻近区域特征相似性,分区域对 2009—2015 年内 7—8 月牧草总产量(鲜重)和 NDVI 进行单一响应因

子建模,发现不同建模区域拟合方程形式也不同,总产草量与植被指数具有线性、幂指数、指数关系、一元二次回归方程多种拟合关系。总体来看,7 个建模分区 NDVI—草地总产草量拟合相关系数在 0.836~0.784。交叉检验除天山北坡西段—伊犁河谷草原畜牧业区 RMSE 值为 2 951 kg·hm⁻² 外,其他分区 RMSE 值在 266~928 kg·hm⁻²,原因在于伊犁河谷草原畜牧业区实测草地总产量在 10 000~30 000 kg·hm⁻² 的样点多。

后续研究中,会在 NDVI 分区建模的基础上,综合考虑 NDVI 分布的巴式距离分区和 NDVI 空间分布(NDVI 空间结构信息),增加确定性自变量的一般回归模型和带有观测误差自变量的扩展回归模型(遥感反演误差到回归统计误差的传播)的比较分析,加入草地长势其他有关解释变量进行建模,设计区域差异、区域效应在回归系数中,如空间滤波和空间面板回归模型里的空间或个体效应变量。

参考文献 References:

- [1] 刘新平,吕晓.新疆牧草地资源利用动态变化及其绩效分析.干旱区地理,2009,32(1):81-85.
Liu X P, Lyu X. Dynamics and performance evaluation of grassland resources utilization in Xinjiang. Arid Land Geography, 2009, 32(1): 81-85. (in Chinese)
- [2] 黄敬峰,胡新博.新疆北部天然草地产草量遥感监测预测模型研究.浙江农业大学学报,1999,25(2):125-129.
Huang J F, Hu X B. Studies on grass yield monitoring and predicting models of natural grassland using remote sensing data in Northern Xinjiang. Journal of Zhejiang Agricultural University, 1999, 25(2): 125-129. (in Chinese)
- [3] 黄敬峰,桑长青,冯振武.天山北坡中段天然草场牧草产量遥感动态监测模式.自然资源学报,1993,8(1):10-17.
Huang J F, Sang C Q, Feng Z W. The remote sensing dynamic monitoring model of the grass yield of natural grassland in the middle section of the Tianshan North slope. Journal of Natural Resources, 1993, 8(2): 10-17. (in Chinese)
- [4] 黄敬峰,王秀珍.天山北坡中东段天然草地光谱植被指数特征.山地学报,1999,17(2):119-124.
Huang J F, Wang X Z. The characteristics of natural grassland spectral vegetation indices in eastern and middle part of Northern Tianshan Mountain. Journal of Mountain Science, 1999, 17(2): 119-124. (in Chinese)
- [5] 黄敬峰,王秀珍,王人潮,蔡承侠,胡新博.天然草地牧草产量与气象卫星植被指数的相关分析.农业现代化研究,2000,21(1):33-35.
Huang J F, Wang X Z, Wang R C, Cai C X, Hu X B. Relation analysis between the production of natural grassland and satellite vegetation indices. Research of Agricultural Modernization, 2000, 21(1): 33-35. (in Chinese)
- [6] 黄敬峰,王秀珍,王人潮,胡新博.天然草地牧草产量遥感综合监测预测模型研究.遥感学报,2001,5(1):69-73.
Huang J F, Wang X Z, Wang R C, Hu X B. A study on monitoring and predicting models of grass yield in natural grassland using remote sensing data and meteorological data. Journal of Remote Sensing, 2001, 5(1): 69-73. (in Chinese)
- [7] 黄敬峰,王秀珍,蔡承侠,胡新博.利用 NOAA/AVHRR 资料监测北疆天然草地生产力.草业科学,1999,16(15):62-65.
Huang J F, Wang X Z, Cai C X, Hu X B. Using NOAA/AVHRR data monitoring natural grassland productivity in the northern Xinjiang uygur autonomous region. Pratacultural Science, 1999, 16(15): 62-65. (in Chinese)
- [8] 王新欣,朱进忠,范燕敏,武鹏飞.利用 EOS/MODIS 植被指数建立草地估产模型的研究.新疆农业科学,2008,45(5):843-846.
Wang X X, Zhu J Z, Fan Y M, Wu P F. Estimation model of establishing grassland with EOS/MODIS vegetation Indexes. Xin-

- jiang Agricultural Sciences, 2008, 45(5): 843-846. (in Chinese)
- [9] 钱育蓉, 杨峰, 李建龙, 于炯, 贾振红. 基于 3S 的新疆阜康典型草地产草量及草畜平衡分析. 草业科学, 2013, 30(9): 1330-1337.
Qian Y R, Yang F, Li J L, Yu J, Jia Z H. Yield and animal-feed balance of typical grassland in Xinjiang Fukang using 3S techniques. Pratacultural Science, 2013, 30(9): 1330-1337. (in Chinese)
- [10] 李建龙, 蒋平, 戴若兰. RS, GPS 和 GIS 集成系统在新疆北部天然草地估产技术中的应用进展. 生态学报, 1998, 18(5): 504-510.
Li J L, Jiang P, Dai R L. Advances in study on the remote sensing technology and gps and gis integration systems in estimating grassland yield applications in the north of Xinjiang, China. Acta Ecologica Sinica, 1998, 18(5): 504-510. (in Chinese)
- [11] 严建武, 李春娥, 袁雷, 陈全功. EOS-MODIS 数据在草地资源监测中的应用进展综述. 草业科学, 2008, 25(4): 1-9.
Yan J W, Li C E, Yuan L, Chen Q G. Application summary of EOS-MODIS data in the monitoring of grassland resources. Pratacultural Science, 2008, 25(4): 1-9. (in Chinese)
- [12] Jin Y X, Xu B, Yang X C, Li J Y, Wang D L, Ma H L. Remote sensing dynamic estimation of grass production in Xilinguole, Inner Mongolia. Scientia Sinica, 2011, 41(12): 1185-1195.
- [13] Leisher C, Hess S, Boucher T M, Beukering P V, Sanjayan M. Measuring the impacts of community-based grasslands management in Mongolia's Gobi. PLoS One, 2012, 7(2): e30991.
- [14] Ouyang W, Hao F H, Skidmore A K, Groen T A, Toxopeus A G, Wang T. Integration of multi-sensor data to assess grassland dynamics in a Yellow River sub-watershed. Ecological Indicators, 2012, 18: 163-170.
- [15] Zhao F, Xu B, Yang X, Jin Y, Li J, Xia L, Chen S, Ma H L. Remote Sensing estimates of grassland aboveground biomass based on MODIS Net Primary Productivity (NPP): A case study in the Xilingol grassland of northern China. Remote Sensing, 2014, 6(6): 5368-5386.
- [16] Jin Y X, Yang X C, Qiu J J, Li J Y, Gao T, Wu Q, Zhao F, Ma H L, Yu H D, Xu B. Remote sensing-based biomass estimation and its spatio-temporal variations in temperate grassland, Northern China. Remote Sensing, 2014, 6(2): 1496-1513.
- [17] 贺俊杰, 锡林郭勒草地 NDVI 和牧草估产产量的变化特征. 中国农学通报, 2015, 31(17): 1-5.
He J J. Variation characteristics of the NDVI and grass estimation yield in Xilingol grassland. Chinese Agricultural Science Bulletin, 2015, 31(17): 1-5. (in Chinese)
- [18] 姚兴成, 曲恬甜, 常文静, 尹俊, 李永进, 孙振中, 曾辉. 基于 MODIS 数据和植被特征估算草地生物量. 中国生态农业学报, 2017, 25(4): 530-541.
Yao X C, Qu T T, Chang W J, Yin J, Li Y J, Sun Z Z, Zeng H. Estimation of grassland biomass using MODIS data and plant community characteristics. Chinese Journal of Eco-Agriculture, 2017, 25(4): 530-541. (in Chinese)
- [19] 徐佳, 陈媛媛, 黄其欢, 何秀凤. 综合灰度与纹理特征的高分辨率星载 SAR 图像建筑区提取方法研究. 遥感技术与应用, 2012, 27(5): 692-698.
Xu J, Chen Y Y, Huang Q H, He X F. Built-up areas extraction in high resolution space borne SAR image based on the integration of grey and texture features. Remote Sensing Technology and Application, 2012, 27(5): 692-698. (in Chinese)
- [20] 李侃, 平西建. 基于图像内容和特征融合的隐写盲检测. 应用科学学报, 2013, 31(1): 97-103.
Li K, Ping X J. Blind steg analysis based on image content and feature fusion. Journal of Applied Sciences, 2013, 31(1): 97-103. (in Chinese)
- [21] 赵小敏, 孙志刚, 夏明. 基于局部学习的车辆图像识别方法. 浙江工业大学学报, 2017, 45(4): 439-444.
Zhao X M, Sun Z G, Xia M. A vehicle image recognition method based on local learning. Journal of Zhejiang University of Zhejiang University of Technology, 2017, 45(4): 439-444. (in Chinese)
- [22] Olden J D, Jackson D A. Torturing data for the sake of generality: How valid are our regression models. Ecoscience, 2000, 7(4): 501-510.

(责任编辑 王芳)