

DOI:10.13869/j.cnki.rswc.2025.02.021; CSTR:32311.14.rswc.2025.02.021.

李文戈, 陈楠, 孙阵阵. 地形地貌视角下黄土高原植被 GPP 模拟及空间分异研究[J]. 水土保持研究, 2025, 32(2): 331-339.

Li Wenge, Chen Nan, Sun Zhenzhen. Study on simulation and spatial differentiation of vegetation GPP in the Loess Plateau from the perspective of terrain and geomorphology[J]. Research of Soil and Water Conservation, 2025, 32(2): 331-339.

地形地貌视角下黄土高原植被 GPP 模拟及空间分异研究

李文戈^{1,2}, 陈楠^{1,2}, 孙阵阵^{1,2}

(1.福州大学 空间数据挖掘与信息共享教育部重点实验室,
福州 350108; 2.福州大学 数字中国研究院(福建), 福州 350108)

摘要: [目的]揭示在地势起伏影响下植被 GPP 时空格局特征, 进而深入分析地形地貌与植被 GPP 之间的相互作用机制, 为植被碳通量模拟以及空间分异性研究提供新的视角。 [方法]采用机器学习模型, 基于宏观地形因子构建植被 GPP 模拟模型。通过谱模型提取 6 个典型地貌样区的植被 GPP 空间谱, 并运用定性和定量分析方法研究了其空间异质性。 [结果]XGBoost 模型的模拟精度较好, 且引入宏观地形因子特征组模型的决定系数(R^2)相较于经典特征组提升 11.26%, 与微观地形因子特征组相比提高了 0.94%, 同时均方根误差 (RMSE) 分别降低了 21.27% 和 2.27%。2003—2023 年, 黄土高原植被 GPP 整体上升了 19.12%, 呈现出东南高西北低的空间分布特征。区域内 6 种典型样区的 GPP 在不同地形条件下表现出明显的地形分异性, 且普遍随着地形崎岖度的增加, 呈现先降后升的波动变化趋势。 [结论]地形因子在植被 GPP 的模拟中起到了关键作用, 且宏观地形因子比微观地形因子更能揭示地形起伏对 GPP 的影响。

关键词: 黄土高原; 植被总初级生产力; 机器学习; 数字高程模型; 谱模型

中图分类号: TP181; Q948

文献标识码: A

文章编号: 1005-3409(2025)02-0331-09

Study on simulation and spatial differentiation of vegetation GPP in the Loess Plateau from the perspective of terrain and geomorphology

Li Wenge^{1,2}, Chen Nan^{1,2}, Sun Zhenzhen^{1,2}

(1.Key Lab for Spatial Data Mining and Information Sharing, Ministry of Education, Fuzhou University, Fuzhou 350108, China; 2.The Academy of Digital China (Fujian), Fuzhou University, Fuzhou 350108, China)

Abstract: [Objective] The study aims to reveal the spatiotemporal patterns of gross primary productivity (GPP) of vegetation under the influence of terrain undulation, to further delve into the interaction mechanisms between topography and vegetation GPP, and to provide new perspectives for the simulation of vegetation carbon flux and the study of its heterogeneity. [Methods] Machine learning models were employed to construct a vegetation GPP simulation model based on macro-scale terrain factors. Spectral models were utilized to extract the spatial spectra of GPP in six typical geomorphic sample areas, and qualitative and quantitative analysis methods were applied to investigate their spatial heterogeneity. [Results] The XGBoost model in this study showed the improved accuracy in vegetation GPP simulation. The R^2 of the model with macroscopic topographic features increased by 11.26% over the conventional feature set and by 0.94% over the micro-topographic feature set. Correspondingly, the RMSE was reduced by 21.27% and 2.27%, respectively. From 2003 to 2023, Loess Plateau vegetation GPP rose by

收稿日期: 2024-05-27

修回日期: 2024-06-17

接受日期: 2024-06-27

资助项目: 国家自然科学基金(41771423); 福建省自然科学基金(2023J01421)

第一作者: 李文戈(1999—), 男, 湖南冷水江人, 在读硕士研究生, 研究方向为地理信息系统。E-mail: outlooks@qq.com

通信作者: 陈楠(1975—), 男, 福建厦门人, 博士, 研究员, 主要从事数字地形分析研究。E-mail: chennan@fzu.edu.cn

http://stbcjy.paperonice.org

19.12%, with higher values in the southeast and lower values in the northwest. GPP varied notably across six typical regions, generally peaking after an initial decline with increasing terrain ruggedness. [Conclusion] Topographic factors play a crucial role in simulating vegetation GPP, with macroscopic topographic factors more effectively revealing the impact of terrain undulation on GPP than microscopic ones.

Keywords: Loess Plateau; gross primary production; machine learning; digital elevation model; spectrum

黄土高原作为中国北方的关键生态屏障,其稀疏的植被覆盖和长期的水土流失问题使其成为生态环境研究的焦点^[1]。碳通量作为全球碳循环的关键,涉及生态系统通过光合作用、呼吸作用和分解作用的碳交换;而植被的总初级生产力(Gross Primary Productivity, GPP)是碳通量中一个核心组成部分,作为维持生态系统功能的基础对于评估生态系统健康、预测其对环境变化的响应至关重要,是目前国内外备受瞩目的研究领域^[2-4]。尤其在黄土高原这样的生态敏感地区,深入研究植被 GPP 不仅有助于揭示其碳固定潜力和生产动力动态,而且在指导生态恢复、提高区域生态安全和促进可持续发展等方面扮演着关键角色。

目前用来研究的碳通量数据主要来自通量塔站点的实际观测获得或者对碳通量数据进行模拟生成。由于站点数量有限且分布不均匀,为了弥补数据有限对区域化碳通量研究的缺陷,学者们因此对碳通量进行模拟,以此得到区域性乃至全球性的碳通量数据,更深入研究碳通量的时空分异规律,因此碳通量模拟也成为目前领域内研究的热点。目前碳通量模拟的方法非常丰富,用于研究最多的是基于遥感数据驱动式的模型,典型的模型有 CASA^[5]、EC-LUE^[6] 以及 MODIS 推出的 GPP 产品也是基于 LUE 模型,此类模型简单实用,在如今较为容易获得遥感数据的今天,是大尺度时空范围内最常用的方法;但在全球变化的大背景下,植被与大气之间的碳交换过程变得愈加复杂,生物地球化学循环的复杂性以及微气象条件的多变性导致了模型中各因子之间的非线性相互作用。这种复杂性使得现有模型在精确捕捉这些动态过程方面面临挑战,进而影响了模型的精确性和稳定性,从而存在与实测数据吻合度不高的问题^[7-8]。

鉴于前两类模型存在的复杂性与不足,学者们开始追求更适合的碳通量模拟方案。利用机器学习以及站点数据结合构成的新型数据驱动式模型能摆脱复杂的先验假设和模型参数,在提高植被 GPP 模拟准确性方面取得了显著进展,如 Yang 等^[9]利用支持向量机(SVM)技术,对美国大陆尺度的植被 GPP 进行了精确模拟;Bai 等^[10]则采用随机森林(RF)模型,成功实现了全球尺度的植被 GPP 模拟。在特征选择方面,研究者通常集中于植被的生态特性和气候因

素,构建特征组,包括植被指数、地表温度、叶面积指数(LAI)和植被有效光合辐射吸收比例(fPAR)等^[11-12]。然而,地形因素在 GPP 模拟中的重要性也日益受到关注。李玉洁等^[13]通过引入高程和坡度等宏观地形指标对福建省的植被 GPP 进行了模拟,得出地形特征是机器学习估算 GPP 的重要参数、对于提升植被 GPP 的模拟精度具有显著的积极作用的结论。同样,刘欢欢等^[14]在对黄土高原草地净初级生产力的模拟研究中也顾及了地形因素,证实了在黄土高原这种地形复杂的区域,考虑地形特征对于提高模型的预测性能是有明显益处的。然而,在机器学习模型中集成多源数据时,需要将不同特征重采样至统一分辨率,较低分辨率的 DEM 数据可能导致传统的微观地形因子所包含的地形信息和地形起伏特征的丢失,从而减弱了这些因子在碳通量模拟中的有效性。这构成了当前碳通量模拟中的一个挑战。针对这一问题,本研究尝试引入宏观地形因子,以期在模拟过程中更全面地捕捉地形对植被 GPP 的影响。

黄土高原的植被碳通量时空分布特征已受到广泛研究,但其与该地区复杂地貌的关联尚未得到充分探讨。为了填补这一空白,本研究将引入谱模型提取黄土高原不同地貌类型的植被 GPP 稳定空间谱,综合考虑地形起伏对植被 GPP 空间特征的影响,进而深入分析植被 GPP 的空间分异性,揭示其与黄土高原地貌的相互作用关系。谱模型是一种基于坡谱理论的定量分析工具,通过定量地形因子及其分级下的地学信息构建的直方图模型,如 Chen 等^[15]对黄土高原 6 种典型地貌构建的天文辐射谱和可照时间谱,成功揭示了地形起伏条件下天文辐射和可照时间的稳定空间结构;该理论模型的应用为地貌学领域提供了一种强有力的分析工具,使得地貌特征的系统解读及地貌过程多层次结构的深入研究成为可能^[16]。

1 数据来源与研究方法

1.1 研究区概况

黄土高原位于中国中北部,横跨东经 101°—114°和北纬 34°—41°,是一个以黄土地貌为主导的区域地貌单元,是世界上黄土覆盖面积最广的地区(图 1)。该地区处于温带大陆性季风带,主要气候类型为温带季风气候

和温带大陆性气候。地形上,黄土高原呈现西北高东南低的倾斜特征,显著的地形起伏是其突出的特征。

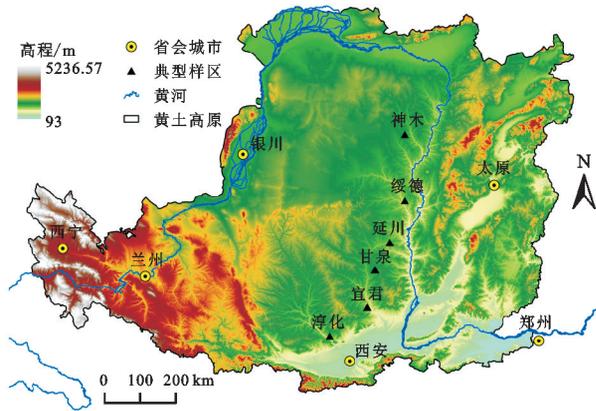


图 1 研究区位置及典型样区分布

Fig. 1 Study area location and distribution of typical sampling areas

黄土高原因其严重的水土流失问题而闻名^[17],同时,其独特的地理特征和地貌价值也引起了全球的广泛关注,成为地貌学研究的重要区域^[18]。为了深入探讨黄土高原典型地貌与植被总初级生产力(GPP)之间的关系,本研究选取黄土高原 6 种不同的地貌类型进行深入分析。由表 1 可知,这些样区所代表的是黄土高原核心区域的典型地貌类型,并且在南北方向上存在逐步过渡的特征^[19],符合黄土高原“黄土梁—黄土塬—黄土峁”的逐层演化规律,对于研究黄土高原的植被 GPP 空间分异具有重要意义。

表 1 黄土高原典型样区基本信息

Table 1 Basic Information Table of Typical Sampling Areas in the Loess Plateau

样区	纬度 (N)	经度 (E)	地貌类型	平均坡度/ 高程范围/ (°) m
神木	38°50'32"	110°29'56"	沙盖黄土低丘	9.45 1035~1323
绥德	37°35'00"	110°18'45"	黄土峁状丘陵沟壑	26.39 820~1180
延川	36°45'00"	109°56'15"	黄土梁状丘陵沟壑	31.30 923~1252
甘泉	36°14'15"	109°36'35"	黄土梁状丘陵沟壑	26.01 1145~1458
宜君	35°27'30"	109°22'30"	黄土长梁残塬沟壑	19.26 767~1158
淳化	34°52'30"	108°26'15"	黄土塬	12.41 768~1188

1.2 数据来源及预处理

遥感数据。基于前人研究经验,本研究选取的以表征生理生境的经典模型特征组遥感指标数据均来自于 MODIS(<https://modis.gsfc.nasa.gov/>)产品,旨在减少不同数据源合并时可能产生的复杂性和困难,具体信息见表 2。为保证模拟一致性,所有数据空间分辨率统一至 1 km。NDWI 数据则将利用 MOD09A1 数据的波段 2 和波段 6 进行计算;且鉴于 EVI 数据的时间分辨率为 16 d,而其他数据为 8 d,

本研究将 EVI 应用于两个 8 d 周期,以匹配其他数据的时间合成周期。

地形数据。本研究选取的数字高程模型(DEM)为 ASTER GDEM 的第 2 版,其分辨率为 30 m。本研究将利用高分辨率 DEM 数据提取 1 000 m 分辨率网格尺度下的宏观地形特征,以捕捉地形的起伏变化。此外,本研究还将基于同一数据源,提取 1 000 m 空间分辨率的坡度、高程和坡向作为微观地形指标,用于特征组的对比分析。

表 2 遥感数据信息及来源

Table 2 Remote sensing data information and sources

产品名称	数据来源	时间 分辨率/d	空间 分辨率/m
LAI(叶面积指数)	MOD15A2	8	1000
fPAR(植被有效光合辐射吸收比例)	MOD15A2	8	1000
LST(地表温度)	MOD11A2	8	1000
EVI(增强型植被指数)	MOD13A1	16	500
NDWI(归一化水体指数)	MOD09A1	8	500
land type(土地覆盖类型)	MCD12Q1	1	500

实测数据。本研究主要依据中国通量观测研究联盟(ChinaFLUX, <http://www.chinaflux.org/>)提供的开放数据,针对黄土高原地区土地覆盖类型的匹配局限性,辅以 FLUXNET2015 数据集(<http://fluxnet.fluxdata.org>)以获取日尺度的碳通量数据。为保障数据在区域层面的适用性,本研究采纳 MCD12Q1 产品数据,并依据与 FLUXNET 站点数据一致的国际生物圈计划(IGBP)分类体系,对黄土高原地区的土地覆盖类型进行精确分类,见图 2。其中,黄土高原地区常见的土地覆盖类型及其相应的缩写包括:草地(GRA)、农田(CRO)、落叶阔叶林(DBF)、稀疏草原(SAV)、城市建设区(URB)、混合林(MF)以及灌木稀疏草原(WSA)。在这一分类框架下,本研究选择 17 个土地覆盖类型相匹配的通量监测站点,以增强数据代表性和研究科学性。相关站点的详细信息已系统整理并呈现于表 3 中。

1.3 研究方法

1.3.1 机器学习模型构建 随着机器学习技术的不断进步,其在碳通量模拟领域的应用日益广泛。本研究采用机器学习模型,将通量站点的实测 GPP 数据及特征向量按照 7:3 的比例分为训练集和验证集,其中训练集用于模型的训练,而验证集则用于评估模型的预测性能和泛化能力。基于前人研究的经验及机器学习领域的最新进展,本研究选择了随机森林(Random Forest, RF)、支持向量回归(Support Vector Regression, SVR)和 XGBoost 模型这 3 种方法构建模型进行植被 GPP 模拟,通过对比分析得出在植被 GPP

模拟中更为适用的模型及其关键特征组合。

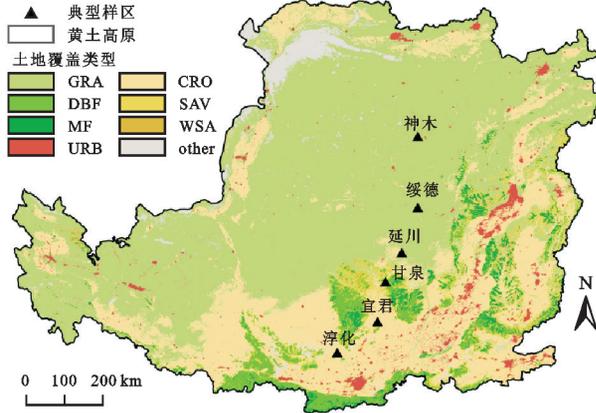


图 2 黄土高原土地覆盖类型

Fig. 2 Land type categories in the Loess Plateau

表 3 实测站点信息

Table 3 Site information table

站点	纬度(N)	经度(E)	时间范围	植被类型
CN-Cha	42.4025	128.0958	2003—2005	MF
CN-Cng	44.5934	123.5092	2007—2010	GRA
CN-Dan	30.4978	91.0664	2004—2005	GRA
CN-Du2	42.0467	116.2836	2006—2008	GRA
CN-Du3	42.0551	116.2809	2009—2010	GRA
CN-HaM	37.37	101.18	2002—2004	GRA
CN-SW2	41.7902	111.8971	2010—2012	GRA
US-AR1	36.4267	-99.42	2009—2012	GRA
CN-YC	36.829	116.5702	2003—2010	CRO
US-ARM	36.6058	-97.4888	2003—2012	CRO
US-CRT	41.6285	-83.3471	2011—2013	CRO
US-Ne1	41.1651	-96.4766	2001—2012	CRO
US-SRM	31.8214	-110.8661	2004—2014	WSA
US-UMB	45.5598	-84.7138	2000—2014	DBF
JP-MBF	44.3869	142.3186	2004—2005	DBF
AU-ASM	-22.283	133.249	2010—2014	SAV
AU-Cpr	-34.0021	140.5891	2010—2014	SAV

随机森林回归是一种集成学习方法,集成多个决策树,通过特征子集的随机抽样增强模型稳健性,并集成各树预测以提升精度。有很好的高维数据处理和抗噪声能力,不过存在过拟合的问题^[10]。支持向量回归(SVR)作为支持向量机(SVM)的回归算法,过引入核函数将数据映射到高维空间以寻找最优超平面,从而解决非线性回归问题,具有较强泛化能力,但容易受到噪声影响^[20-21]。XGBoost 模型基于梯度提升决策树(GBDT)的提升算法,通过逐步构建一系列弱学习器(通常是决策树),并将其组合成一个强学习器,从而提高预测性能,具有高准确性、灵活性的特点,但存在过拟合和欠拟合的问题^[22]。

本研究采用网格搜索技术来优化机器学习模型

的超参数,并使用 R^2 (决定系数)和 RMSE(均方根误差)作为模型评估指标。 R^2 衡量模型对观测数据的拟合程度,其值越接近 1,说明模型的解释能力越强。RMSE 则衡量预测值与实际值之间的偏差,其值越低,表示模型的预测精度越高,从而反映出模型对数据的拟合更为紧密。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{FLUX}_{\text{observe}} - \text{FLUX}_{\text{predict}})^2} \quad (1)$$

$$R^2 = \frac{\left[\sum_{i=1}^n (\text{FLUX}_{\text{observe}} - \overline{\text{FLUX}_{\text{observe}}}) (\text{FLUX}_{\text{predict}} - \overline{\text{FLUX}_{\text{predict}}}) \right]^2}{\sum_{i=1}^n (\text{FLUX}_{\text{observe}} - \overline{\text{FLUX}_{\text{observe}}})^2 \sum_{i=1}^n (\text{FLUX}_{\text{predict}} - \overline{\text{FLUX}_{\text{predict}}})^2} \quad (2)$$

式中: $\text{FLUX}_{\text{observe}}$ 为观察的碳通量值(即植被 GPP 值); $\text{FLUX}_{\text{predict}}$ 则为模拟结果值; $\overline{\text{FLUX}_{\text{observe}}}$ 与 $\overline{\text{FLUX}_{\text{predict}}}$ 分别为所有观测值与模拟值的均值。

1.3.2 特征数据集构建 本研究在现有研究框架内,集成生理生态与地形指标,旨在探索最优的特征组合。在生理生态指标选择方面,本研究基于前人研究选定 EVI,NDWI,LST(昼夜 LST),LAI,fPAR 及时间特征作为解释变量,构建一组经典的代表性特征向量组。为了更精确地捕捉地形的复杂性,本研究尝试引入宏观地形因子。地形起伏作为地貌学的核心研究内容,已有众多学者建立了完善的分析指标体系,并发展了成熟的数字地形分析理论。通过高分辨率 DEM 数据,本研究提取宏观地形因子,以期获得更详尽的地形起伏信息,从而全面反映地形特征,所采用的宏观地形因子及计算公式如下所示:

地形起伏度(ER):

$$ER = H_{\max} - H_{\min} \quad (3)$$

地表切割深度(SCD):

$$SCD = H_{\text{mean}} - H_{\min} \quad (4)$$

高程变异系数(EVC):

$$EVC = \frac{H_{\text{sqr}}}{H_{\text{mean}}} \quad (5)$$

地表粗糙度(SR):

$$SR = \frac{1}{\cos\left(\frac{\text{slope} \cdot 3.14}{180}\right)} \quad (6)$$

平均高程(H_{mean}):

$$H_{\text{mean}} = \frac{h(x_1) + h(x_2) + \dots + h(x_n)}{n} \quad (7)$$

式中: H_{\max} 为尺度范围内的高程最高值(m); H_{\min} 为高程最低值(m); H_{sqr} 为所有栅格高程值的平方根(m);slope 为坡度($^{\circ}$); $h(x_1) + h(x_2) + \dots + h(x_n)$ 为所有栅格的高程值的总和(m); n 为区域内栅格总数。

此外,为了与宏观地形因子进行对比,本研究还

根据前人引入地形因子进行植被 GPP 模拟的研究,设计两组引入微观地形因子的特征对比组。一组采用传统的微观地形因子,即坡度和高程;另一组则综合考虑包括坡向在内的所有微观地形因子。因此,本研究共设计 4 个特征组合对比组,具体见表 4。通过这种细致的特征数据集构建,本研究期望能够为植被 GPP 模拟提供一个全面、深入的特征分析框架,进而寻找对 GPP 模拟具有显著预测效能的特征组合。

表 4 模型输入特征向量组合

Table 4 Model input feature vector groups

特征组	经典特征 向量组	微观地形因子 (坡度、高程)	微观地形因子 (坡度、高程、坡向)	宏观 地形因子
A	√			
B	√	√		
C	√		√	
D	√			√

注:√表示特征组包含此特征向量。

1.3.3 谱模型构建 本研究借鉴坡谱、天文辐射谱的思想构建地形起伏配置下的植被 GPP 空间谱,该方法的核心在于,通过在预设区域内对栅格数据应用矩形窗口的扩张,并结合地形因子与植被 GPP 值的统计分析,通过判断是否满足给定的误差植被以确定不同地貌的稳定空间谱结构。鉴于模拟所得植被 GPP 数据空间分辨率进行谱提取稳定性存在不足,本研究通过重采样技术将其精度提升至 500 m,且由于数字高程模型在低分辨率时在表达坡度信息时的局限性^[23],本研究选择均方根高程作为描述地形起伏的指标,图 3 展示的是黄土高原地区均方根高程信息。获取黄土高原不同地貌类型的均方根高程-年植被总初级生产力空间谱(本研究将其简称为 GPP 空间谱)。

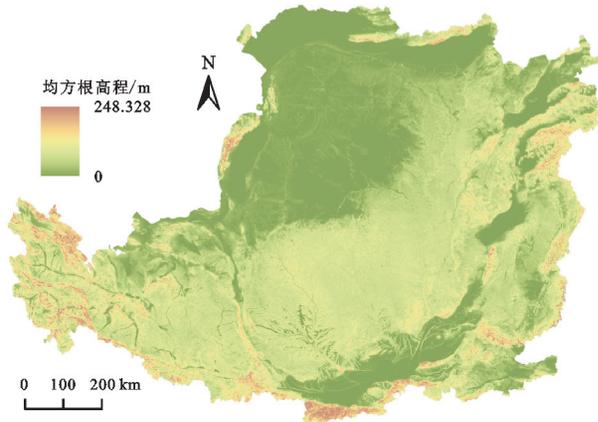


图 3 黄土高原均方根高程分布

Fig. 3 Loess Plateau root mean square elevation distribution

GPP 空间谱提取流程如下:在各样区中任意选取一个栅格单元 A 作为中心点,运用 GIS 移动窗口分析法,逐步构建尺寸递增的分析窗口,窗口为 3×3,

5×5,依此类推直至(2n+1)×(2n+1)。在每个分析窗口内,对栅格单元的配置地形因子进行等差量化,划分为 11 个等级区间,具体如下:(1) 0~4 m;(2) 4~8 m;(3) 8~12 m;(4) 12~16 m;(5) 16~20 m;(6) 20~24 m;(7) 24~28 m;(8) 28~32 m;(9) 32~36 m;(10) 36~40 m;(11)大于 40 m(均方根高程大于 40 m 区间面积在样区占比不达 3%,可以不再继续进行级别划分)。以初始的 3×3 分析窗口为例,计算该窗口内各均方根高程等级区间的栅格单元对应的植被 GPP 平均值,并以此构建初始 GPP 谱表示为 $GPP^1 = [GPP_1^1, GPP_2^1, \dots, GPP_9^1]$,其中 GPP_j^i 为第 i 个扩大分析窗口内第 j 个均方根高程等级区间的平均植被 GPP 值。

随着分析窗口的逐级扩大至 5×5,7×7 等,重复上述计算过程,得到一系列 GPP 谱,分别为 $GPP^2 = [GPP_1^2, GPP_2^2, \dots, GPP_9^2]$, GPP^3 等。在分析窗口逐步扩大的过程中,为了评估所得 GPP 空间谱是否达到了稳定,本研究设定了两个误差指标 E_1 和 E_2 ,公式如下所示:

$$E_1 = \max \left(\left| \frac{GPP_i^{n+1} - GPP_i^n}{GPP_i^n} \right| \right) \quad (8)$$

$$E_2 = \sum_{i=1}^n \left| \frac{GPP_i^{n+1} - GPP_i^n}{GPP_i^n} \right| \quad (9)$$

式中: GPP_i^{n+1} 为 GPP_i^n 扩大一次分析窗口后的第 i 个均方根高程等级区间的 GPP 空间谱。

本研究分别以 $E_1 < 0.015$ 和 $E_2 < 0.010$ 为准则作为误差评估标注,这些误差指标是参考坡谱和天文辐射谱理论,并结合本研究的试验情况制定的,旨在确保空间谱的稳定性和可靠性。若连续 5 次扩张所得的空间谱均满足上述误差标准,则认为获得了该地貌类型下的稳定 GPP 空间谱。

2 结果与分析

2.1 顾及地形起伏的植被 GPP 模拟比较

本研究使用了 XGBoost、随机森林(RF)和支持向量回归(SVR)3 种机器学习算法,对各特征组的模拟精度进行深入对比分析,其结果展示在图 4 中。从图中可以得出,在顾及地形起伏下构建的宏观地形指标特征组 D,在 3 种机器学习模型中均优于其他特征组。

在 XGBoost 模型中,考虑地形起伏的宏观地形指标特征组 D 的决定系数(R^2)比仅考虑植被生境的特征组 A 提高了 11.26%,均方根误差(RMSE)降低了 21.27%。与采用微观地形指标的特征组 B 和 C 相比,特征组 D 的 R^2 分别提升了 1.18%和 0.94%,而 RMSE 则分别降低了 2.93%和 2.27%。在随机森林模型中,相较于特征组 A,B,C,特征组 D 的 R^2 分

别提升了 5.55%, 1.18%, 0.82%, 而 RMSE 分别下降了 16.36%, 7.33%, 5.92%。在支持向量回归模型中, 特征组 D 相对于特征组 A, B, C 的 R^2 分别提高了 7.54%, 0.48%, 0.12%, 而 RMSE 分别减少了 10.54%, 1.12%, 0.22%。因此, 本研究选择特征集 D 来构建植被 GPP 的模拟模型。在顾及宏观地形的特征集 D 中, 于不同机器学习模型的模拟性能比较下, XGBoost 模型表现出最高的模拟精度, 其 R^2 值达到了 0.859, 相较于 RF 和 SVR 模型分别提高了 0.35% 和 2.14%; 同时, 其 RMSE 值降至 1.591 $\text{g C}/(\text{m}^2 \cdot \text{d})$, 相较于 RF 和 SVR 模型分别降低了 0.87% 和 13.63%。

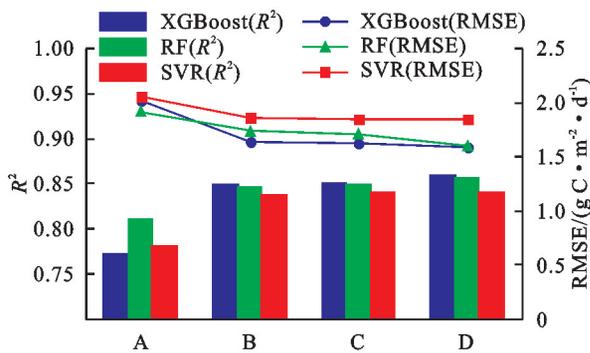


图 4 各特征组在不同模型下性能对比

Fig. 4 Performance comparison chart of different models for each feature group

综上所述, 相较于随机森林与支持向量回归模型, XGBoost 模型在融合地形特征后性能得到的提升更为明显, 精度也更加高。XGBoost 算法的这一优势可能源于其基于梯度提升决策树 (GBDT) 的框架, 该框架通过逐步添加新的树来纠正前序树的错误, 从而实现复杂非线性关系的建模^[24]。此外, XGBoost 算法的正

则化项有助于防止模型过拟合, 这对于处理具有大量特征的高维数据集尤为重要。由此可见, 在顾及地形起伏的植被 GPP 模拟模型中, XGBoost 是一种相对较为高效、可靠且拥有较高精度的模型。

2.2 黄土高原的模拟结果及时空格局分析

基于研究结果, 本文将采用 XGBoost 模型并构建以顾及地形起伏的宏观地形因子作为地形指标的特征组对地形复杂且观测数据匮乏的黄土高原地区进行 2003—2023 年的模拟, 其模拟结果如图 5A 所示。

近 20 年黄土高原植被 GPP 的模拟结果显示, 该地区植被生产力在空间上呈现明显的异质性, 并总体呈东南高、西北低的分布趋势。在黄土高原南部的陕西和甘肃地区, 特别是渭河流域, 由于秦岭的森林覆盖和关中平原的农业条件, 具有较高的 GPP 值。相比之下, 黄土高原西北部的内蒙古和宁夏地区, 由于沙漠和沙地的影响, 植被覆盖度较低, GPP 值相对较低。尽管如此, 在黄土高原西北部的黄河沿岸的河套平原和宁夏平原作为重要的农业区, 其 GPP 值仍然较高。

从图 5B 中我们可以观察到, 黄土高原植被 GPP 在近 20 年间呈现明显的增长趋势, 其中, 最低值出现在 2003 年, 随后整体呈现稳定上升趋势。在这近 20 年的时间跨度内, 植被 GPP 的年均值从 241.89 $\text{g C}/(\text{m}^2 \cdot \text{a})$ 升至 299.09 $\text{g C}/(\text{m}^2 \cdot \text{a})$, 整体增长了 19.12%, 这一结论与前人对黄土高原碳通量时序变化的研究结果相一致^[25-26]。黄土高原自实施包括退耕还林还草在内的一系列生态恢复措施以来, 该地区的植被覆盖得到了显著恢复, 植被生产力在长期范围内显著提高, 这反映了黄土高原生态环境的积极变化。

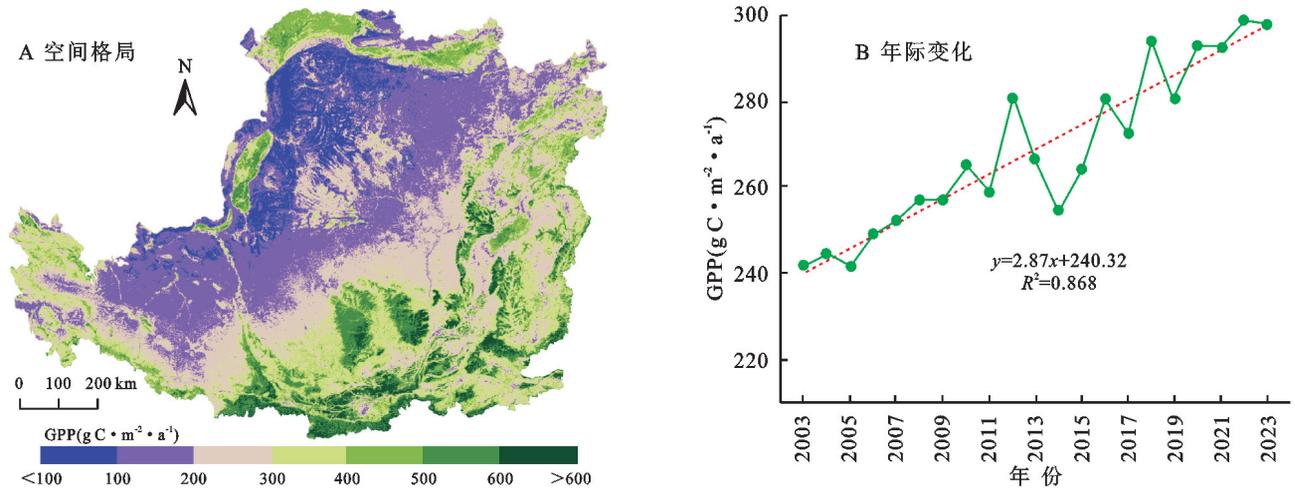


图 5 黄土高原 2003—2023 年年均植被 GPP 时空格局

Fig. 5 spatial and temporal patterns of annual average vegetation GPP on the Loess Plateau from 2003 to 2023

综上所述, 本研究对于黄土高原近 20 年的模拟结果在空间格局和趋势变化上与前人研究保持一致,

且与 MODIS 的 MOD17A2H 产品数据估算的黄土高原植被 GPP 的空间分布相吻合, 年度 GPP 的波动

趋势与年内变化特征也呈现出良好的对应关系。

2.3 植被 GPP 谱及空间分异分析

如图 6 所示,对于黄土高原的 6 种典型地貌类型样区,我们运用谱模型成功构建了各样区的均方根高程-年植被总初级生产力空间谱(GPP 空间谱),结果如图呈现;我们观察到,随着均方根高程分级的变化,植被 GPP 表现出一定的变化规律。通过对 GPP 谱进行定性与量化分析,我们深入探讨了地形起伏与植被生产力之间的复杂相互作用,并得到了以下发现:

(1) 从地形信息图谱的视角分析,黄土高原的 6 个典型地貌样区在反映地形起伏的均方根高程的地形配置下,其植被 GPP 表现出一定的地形分异性。如绥德、延川和淳化样区的 GPP 变化较为平缓,可以反映出其 GPP 对地形变化的响应较为稳定,而淳化样区的 GPP 谱尤其稳定,未随地形变化出现显著波动,这归因于该区域黄土塬及残塬的平缓地形,降低了 GPP 对地形的敏感性。而甘泉和宜君样区的变化较大,表明这些区域的 GPP 对地形变化更为敏感,说明以林地为主的样区,在地形起伏下植被生长条件的差异性较大,导致 GPP 随地形复杂度的增加而有更

明显变化。尽管不同地貌存在差异,但它们均展现出一致的动态变化趋势,即随着地形崎岖复杂程度的增加,植被 GPP 普遍呈现出先下降后上升的波动变化。

(2) 分析黄土高原各样区植被 GPP 平均值揭示了显著的纬度分异规律。沿纬度从北至南,GPP 均值呈现出与纬度梯度正相关的增长趋势,表现为淳化、延川、绥德至神木的递增序列。甘泉和宜君样区的 GPP 值尤为突出,这一现象与它们的地理位置和植被类型紧密相关。这两个样区坐落于洛河流域南部,其植被以落叶阔叶林和混交林为主,这些森林生态系统不仅结构复杂、生物多样性丰富,而且在光合作用和碳固定效率上具有显著优势。林地在地形起伏的区域能够提供多样化的水热条件,满足多样的生态位和微气候需求,进而积极影响植被 GPP。特别是甘泉样区,得益于其位于洛河流域中游的地理位置,拥有湿润的气候和强大的植被恢复力,加之处于子午岭和黄龙山之间,使得该样区拥有广阔的林地面积。这些因素综合作用,使得甘泉样区在所研究样区中 GPP 平均值最高,体现了植被 GPP 的空间分布与纬度、地形、气候条件以及植被类型的复杂相互作用。

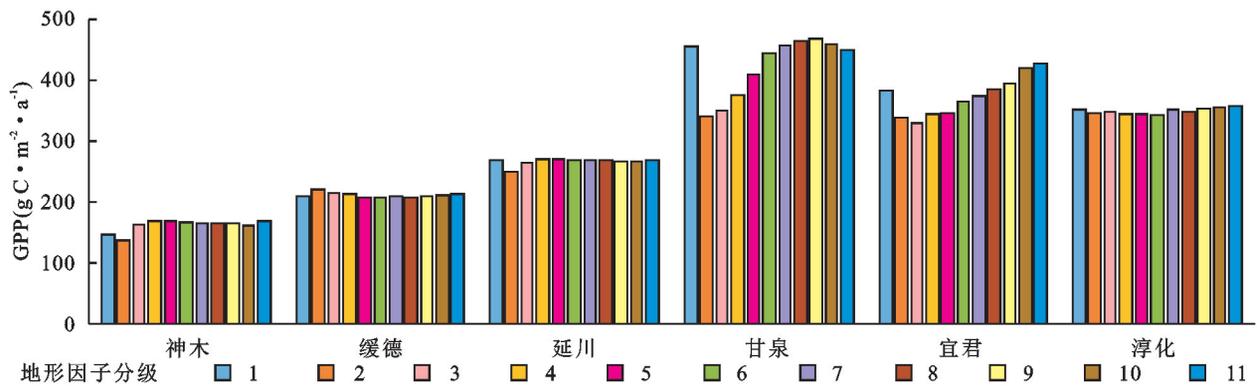


图 6 黄土高原典型样区 GPP 空间谱

Fig. 6 Loess Plateau typical sampling area GPP spatial spectrum

3 讨论

本文构建基于机器学习算法顾及地形起伏的植被 GPP 模拟模型,由结果我们可以观察到顾及地形的特征数据集在提升模型模拟精度方面具有显著优势,即融合地形特征指标后,相较于未融合地形特征的经典植被 GPP 模拟模型,展现出了更高的预测能力。这一现象表明,地形特征的纳入对于提高模型的泛化能力和准确性具有重要作用,这与李玉洁等^[13]在其研究中所指出的地形因子对于增强植被 GPP 模拟精度的积极影响相一致。同时本研究进一步发现,宏观地形指标特征组在模拟精度上优于微观地形指标特征组,这是因为宏观地形指标是由高分辨率

DEM 计算得来,捕捉格网区域范围内地形的连续性和复杂性,表达了区域的地形起伏变化情况以及地表侵蚀程度。而相比之下,微观地形因子通常通过重采样高分辨率 DEM 得到,这一过程可能会导致原始数据的空间信息损失,降低模型对小尺度地形变化的敏感性。由此可得出宏观地形指标能够更全面地反映地形起伏对植被生长的影响,从而为植被 GPP 的模拟提供了更为精确的数据支持。

本研究在模拟植被 GPP 时,与 MODIS 数据产品的估算值及其他研究进行了定量比较。结果表明,本研究的模拟值相对较低。此差异主要源于所采用模拟方法之间的根本差异。本研究所采用的基于机器学习的数据驱动模型在模拟植被碳通量方面,其结

果在很大程度上受限于训练数据集的质量和选择性。尽管本研究综合利用了现有的开放数据资源和区域特征,但数据资源的限制可能导致模拟结果的不稳定性。未来研究在获取更高质量、更广覆盖范围的数据集支持下,有望显著提升模型的训练效果,进而提高碳通量模拟的精度。并且随着遥感技术和通量观测网络的不断发展,我们期望将本研究所建立的植被 GPP 模拟模型扩展到更大尺度的研究中,以实现更为精确和全面的植被碳通量评估。

本研究创新性地在地形信息图谱方法应用于生态领域,通过构建植被总初级生产力空间谱,深入探究了黄土高原 6 种典型地貌类型下起伏地形对碳通量稳定空间结构的影响。研究揭示了这些地貌类型 GPP 谱的共性与个性,为理解植被 GPP 与地形复杂性之间的关系提供了新的视角。与以往利用谱模型理论对黄土高原地区太阳辐射空间谱的研究相比,我们发现植被 GPP 并非随地形复杂度的增加而线性变化,这表明植被 GPP 与地形之间的关系并非单一线性,而是存在更为复杂的相互作用,这也与前人基于黄土高原草地净初级生产力与地形特征关系的研究相吻合,均发现植被碳通量与地形特征之间的关系并非为简单线性^[26-27]。本研究的发现为先前的认识提供了新的补充,并突出了植被 GPP 与地形关系研究的复杂性。

4 结论

(1) 将地形特征纳入机器学习模型显著增强了植被 GPP 的模拟精度,而宏观地形指标特征组在模拟精度上的表现优于微观地形指标特征组,表明宏观地形指标更能有效捕捉和反映地形起伏的综合信息。而在机器学习算法中,XGBoost 模型在顾及地形起伏的植被 GPP 模拟上展现了较好的性能。

(2) 2003—2023 年,黄土高原的植被 GPP 整体呈现上升趋势,这一现象验证了退耕还林还草政策的有效性,并反映了该地区生态环境的积极改善。植被 GPP 在黄土高原的空间分布呈现出东南高、西北低的格局,其中林地的 GPP 尤为显著。

(3) 地形信息图谱分析显示,黄土高原 6 种典型地貌的植被 GPP 在地形起伏的影响下表现出明显的分异性。但整体上各地貌类型的植被 GPP 随地形崎岖复杂度的增加普遍呈现先降后升的非线性动态变化,反映了植被碳通量与地形崎岖程度的复杂关联。

参考文献(References):

[1] 张琨,吕一河,傅伯杰,等.黄土高原植被覆盖变化对生

态系统服务影响及其阈值[J].地理学报,2020,75(5):949-960.

Zhang K, Lyu Y H, Fu B J, et al. The effects of vegetation coverage changes on ecosystem service and their threshold in the Loess Plateau[J]. Acta Geographica Sinica, 2020,75(5):949-960.

[2] Erb K H, Kastner T, Plutzer C, et al. Unexpectedly large impact of forest management and grazing on global vegetation biomass[J]. Nature, 2018,553:73-76.

[3] Wang S H, Zhang Y G, Ju W M, et al. Tracking the seasonal and inter-annual variations of global gross primary production during last four decades using satellite near-infrared reflectance data[J]. Science of the Total Environment, 2021,755:142569.

[4] 朴世龙,岳超,丁金枝,等.试论陆地生态系统碳汇在“碳中和”目标中的作用[J].中国科学:地球科学,2022,52(7):1419-1426.

Piao S L, Yue C, Ding J Z, et al. Perspectives on the role of terrestrial ecosystems in the ‘carbon neutrality’ strategy[J]. Science China Earth Sciences, 2022,52(7):1419-1426.

[5] Potter C S, Randerson J T, Field C B, et al. Terrestrial ecosystem production: a process model based on global satellite and surface data[J]. Global Biogeochemical Cycles, 1993,7(4):811-841.

[6] Yuan W P, Liu S G, Zhou G S, et al. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes[J]. Agricultural and Forest Meteorology, 2007,143(3/4):189-207.

[7] Frank D, Reichstein M, Bahn M, et al. Effects of climate extremes on the terrestrial carbon cycle: concepts, processes and potential future impacts[J]. Global Change Biology, 2015,21(8):2861-2880.

[8] Anav A, Friedlingstein P, Beer C, et al. Spatiotemporal patterns of terrestrial gross primary production: a review[J]. Reviews of Geophysics, 2015,53(3):785-818.

[9] Yang F H, Ichii K, White M A, et al. Developing a continental-scale measure of gross primary production by combining MODIS and AmeriFlux data through Support Vector Machine approach[J]. Remote Sensing of Environment, 2007,110(1):109-122.

[10] Bai J, Zhang H L, Sun R, et al. Estimation of global GPP from GOME-2 and OCO-2 SIF by considering the dynamic variations of GPP-SIF relationship[J]. Agricultural and Forest Meteorology, 2022,326:109180.

[11] Xiao J F, Zhuang Q L, Law B E, et al. A continuous

- measure of gross primary production for the conterminous United States derived from MODIS and AmeriFlux data[J]. *Remote Sensing of Environment*, 2010, 114(3):576-591.
- [12] 张坤,刘乃文,高帅,等.数据驱动的植被总初级生产力估算方法研究[J].*遥感技术与应用*,2020,35(4):943-949.
Zhang K, Liu N W, Gao S, et al. Data-driven estimation of gross primary production[J]. *Remote Sensing Technology and Application*, 2020,35(4):943-949.
- [13] 李玉洁,江洪,刘宣广.基于 CatBoost 的植被总初级生产力遥感模拟方法及在福建省的应用[J].*地球信息科学学报*,2023,25(9):1908-1922.
Li Y J, Jiang H, Liu X G. Simulation of vegetation gross primary productivity and its application in Fujian Province of China using remote sensing and CatBoost algorithm [J]. *Journal of Geo-Information Science*, 2023,25(9):1908-1922.
- [14] 刘欢欢,陈印,刘悦,等.基于随机森林模型的黄土高原草地净初级生产力时空格局及未来演变趋势模拟[J].*干旱区研究*,2023,40(1):123-131.
Liu H H, Chen Y, Liu Y, et al. Simulation of spatial pattern and future trends of grassland net primary productivity in the Loess Plateau based on random forest model[J]. *Arid Zone Research*, 2023,40(1):123-131.
- [15] Chen N. Deriving the slope-mean shielded astronomical solar radiation spectrum and slope-mean possible sunshine duration spectrum over the Loess Plateau[J]. *Journal of Mountain Science*, 2020,17(1):133-146.
- [16] Tang G A, Song X D, Li F Y, et al. Slope spectrum critical area and its spatial variation in the Loess Plateau of China[J]. *Journal of Geographical Sciences*, 2015, 25(12):1452-1466.
- [17] 李慧,张少伟,赵阳.黄河流域黄土高原生态质量时空演变格局及其驱动因素[J].*水土保持研究*,2024,31(4):233-240.
Li H, Zhang S W, Zhao Y. Spatiotemporal evolution pattern of ecological quality of the Loess Plateau and its driving factors[J]. *Research of Soil and Water Conservation*, 2024,31(4):233-240.
- [18] Lin S W, Chen N. Employment of continuous slope cumulative frequency spectrum in geomorphology quantitative analysis: a case study on Loess Plateau[J]. *Geomorphology*, 2022,415:108405.
- [19] Zhu H C, Tang G A, Qian K J, et al. Extraction and analysis of gully head of Loess Plateau in China based on digital elevation model[J]. *Chinese Geographical Science*, 2014,24(3):328-338.
- [20] Ueyama M, Ichii K, Iwata H, et al. Upscaling terrestrial carbon dioxide fluxes in Alaska with satellite remote sensing and support vector regression[J]. *Journal of Geophysical Research: Biogeosciences*, 2013,118(3):1266-1281.
- [21] Cho S, Kang M, Ichii K, et al. Evaluation of forest carbon uptake in South Korea using the national flux tower network, remote sensing, and data-driven technology [J]. *Agricultural and Forest Meteorology*, 2021,311:108653.
- [22] 谢军飞,张海清,李代伟,等.基于 Lightgbm 和 XGBoost 的优化深度森林算法[J].*南京大学学报:自然科学*,2023,59(5):833-840.
Xie J F, Zhang H Q, Li D W, et al. Optimized deep forest algorithm based on Lightgbm and XGBoost[J]. *Journal of Nanjing University: Natural Science*, 2023, 59(5):833-840.
- [23] Chang K T, Tsai B W. The effect of DEM resolution on slope and aspect mapping[J]. *Cartography and Geographic Information Systems*, 1991,18(1):69-77.
- [24] Dhaliwal S S, Nahid A A, Abbas R. Effective intrusion detection system using XGBoost[J]. *Information*, 2018,9(7):149.
- [25] 裴宏泽,赵亚超,张廷龙.2000—2020 年黄土高原 NEP 时空格局与驱动力[J].*干旱区研究*,2023,40(11):1833-1844.
Pei H Z, Zhao Y C, Zhang T L. Analysis of spatial and temporal patterns and drivers of local regional NEP in the Loess Plateau from 2000 to 2020[J]. *Arid Zone Research*, 2023,40(11):1833-1844.
- [26] 刘洋洋,王倩,杨悦,等.黄土高原草地净初级生产力时空动态及其影响因素[J].*应用生态学报*,2019,30(7):2309-2319.
Liu Y Y, Wang Q, Yang Y, et al. Spatial-temporal dynamics of grassland NPP and its driving factors in the Loess Plateau, China[J]. *Chinese Journal of Applied Ecology*, 2019,30(7):2309-2319.
- [27] 张佑铭,郎梦凡,刘梦云,等.土地利用转变与海拔高度协同作用黄土高原植被固碳变化特征[J].*生态学报*,2022,42(10):3897-3908.
Zhang Y M, Lang M F, Liu M Y, et al. Vegetation carbon sequestration in the Loess Plateau under the synergistic effects of land cover change and elevations [J]. *Acta Ecologica Sinica*, 2022,42(10):3897-3908.