

基于时序特征的玉米轮作制耕地提取方法研究

李百红¹, 彭勃², 董超^{2*}

(1. 山东省泰安市自然资源与规划局, 山东泰安 270000; 2. 山东农业大学信息与工程学院, 山东泰安 271018)

摘要 [目的]该研究对小麦、玉米轮作制下耕地的面积与分布有重要意义。[方法]基于目视判别收集样本点和 GEE 平台, 分析地物的 NDVI、EVI 等指数时序特征, 开发特征识别算法提取 2018 年玉米轮作区面积分布, 同时与 CART 算法监督分类结果进行精度分析。[结果]与基于多光谱的监督分类相比, 基于时序特征的特征识别算法准确率、精准率、F1 Score 和 Kappa 系数分别提高了 0.121、0.110、0.136 和 0.246。[结论]基于时序特征的特征识别算法可有效识别玉米轮作区, 该算法可为轮作区农业生产提供基础数据支持。

关键词 GEE; 时序; 归一化指数; 提取算法

中图分类号 S126 文献标识码 A

文章编号 0517-6611(2021)19-0214-04

doi:10.3969/j.issn.0517-6611.2021.19.056

开放科学(资源服务)标识码(OSID):

**Study on Extraction Method of Wheat-Maize Rotation System Based on Time Series Characteristics**

LI Bai-hong¹, PENG Bo², DONG Chao² (1. Taian Natural Resources and Planning Bureau of Shandong Province, Taian, Shandong 270000; 2. College of Information Science and Engineering, Shandong Agricultural University, Taian, Shandong 271018)

Abstract [Objective] It is important to study the area and distribution of cultivated land under wheat and maize rotation system. [Method] Based on the collection of sample points and GEE platform, we analyzed the NDVI, EVI and other index time series features of ground objects, developed feature recognition algorithm to extract the 2018 wheat-maize rotation area distribution, and conducted precision analysis with the supervised classification results of CART algorithm. [Result] Compared with the multi-spectral supervised classification, the accuracy, F1 Score and Kappa coefficient of the time-series feature recognition algorithm were improved by 0.121, 0.110, 0.136 and 0.246, respectively. [Conclusion] The feature recognition algorithm based on time series features could effectively identify wheat-maize rotation area, and the algorithm could provide basic data support for agricultural production in rotation area.

Key words GEE; Time series; Normalized index; Extraction algorithm

快速准确获取作物种植的分布信息有利于实现对耕地的精确管理、产量估算、制定科学的粮食管理政策^[1]。当前, 作物识别多基于遥感影像光谱差异, 如熊勤学等^[2-3]利用 NDVI 提取小麦的种植面积, 路中等^[4-6]提取玉米的种植面积, 刘佳等^[7-8]完成了中国冬小麦制图和大规模春小麦提取研究。受限于传统数据处理工具的低效, 这些研究只使用了一种指数提取地物信息, 难以发挥高分辨率多光谱遥感数据的优势。多时相的遥感数据进行分类不仅可以提取地物物候特征, 而且可有效避免单一时相存在的异物同谱和同谱异物状况^[9]。Google Earth Engine(GEE)平台可对大尺度区域进行分析, 还支持多源多时相海量数据的协同研究^[10]。鉴于此, 笔者选取华北平原小麦-玉米轮作区, 基于 GEE 平台, 利用多源遥感数据分析地物样本点光谱指数的时序特征, 依据各类地物的时序特征设计有效的识别算法, 通过精度评价评估识别算法的优劣。

1 研究区域与数据

1.1 研究区概况 华北平原小麦-玉米轮作区(32°16'~39°38'N, 112°50'~118°57'E)总面积约 221 838 km², 跨河北、河南、山东 3 省(图 1)。研究区大部分平原区为栽培植被, 其中以冬小麦、夏玉米轮作最为常见, 主要分布在河南北部和山东西北的引黄灌溉区、太行山前平原水源条件较好的区域。研究区内小麦 10 月中上旬播种, 次年 3 月拔节, 6 月上

旬收割; 6 月初播种玉米, 9 月下旬收割^[11]。**1.2 数据及预处理**

1.2.1 卫星影像选取。 研究所用数据包括: Landsat7 SR (Surface Reflectance, SR) 数据集; Landsat8 SR 数据集; MOD13Q1.006 数据集的植被指数 (MOD13Q1) 第 6 版数据; Sentinel-1 SAR GRD 数据集。研究影像数据来自 GEE 平台 (<https://earthengine.google.com/>)。

1.2.2 影像预处理。 影像预处理包括: ①使用研究区对影像统一进行掩膜; ②利用 Landsat7 和 Landsat8 影像中由 CF-MASK 算法生成的质量评估波段去云^[12]; ③对 Landsat7 和 Landsat8 计算归一化光谱指数。NDVI 计算公式如下:

$$\text{NDVI} = (\text{NIR} - \text{R}) / (\text{NIR} + \text{R}) \quad (1)$$

根据 NIR、R 与 B 波段计算 EVI 计算公式如下:

$$\text{EVI} = 2.5 \times (\text{NIR} - \text{R}) / (\text{NIR} + 6 \times \text{R} - 7.5 \times \text{B} + 1) \quad (2)$$

根据 G 与 MIR 波段计算 MNDWI, 计算公式如式下:

$$\text{MNDWI} = (\text{G} - \text{MIR}) / (\text{G} + \text{MIR}) \quad (3)$$

将 Landsat7 和 Landsat8 的影像合成指数的数据集, 以弥补 Landsat 系列卫星回访周期较长的缺点。

1.3 样本点与验证点 基于时间序列将研究区地物类别分为小麦-玉米轮作区、林地、城镇建设用地、水域和其他耕地类型。利用在线高分辨影像及历史时间轴, 结合研究区内各区县的相关文献资料和统计报告, 通过目视的方法选取地物样本点和验证点(表 1)。

2 时序数据分类方法

2.1 时序特征的组合 基于指数数据集生成样本点指数时序曲线, 取同类指数的平均值生成各类别指数年积日 (Day of

基金项目 山东省重点研发项目(14032761, 140380198)。**作者简介** 李百红(1985—), 女, 山东青州人, 工程师, 硕士, 从事土地资源利用与信息化工作。* 通信作者, 讲师, 博士, 从事农业信息技术研究。**收稿日期** 2021-02-26

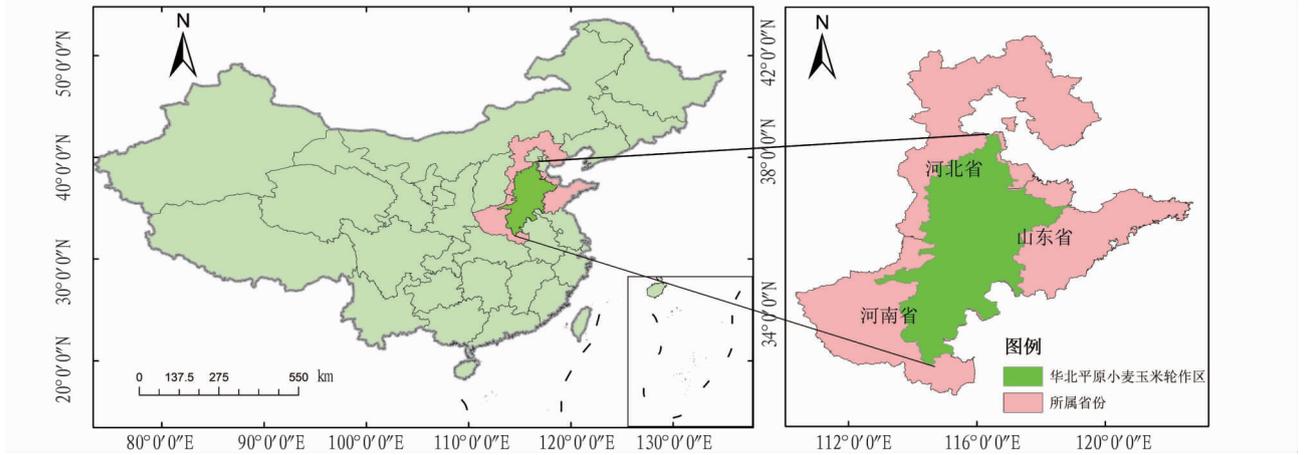


图 1 研究区位置

Fig. 1 Location of the study area

表 1 样本点与验证点个数统计

Table 1 Statistics of sample points and verification points

序号 Code	类别 Group	样本点 Sample site	验证点 Verification site
1	轮作耕地	300	100
2	林地	150	50
3	建设用地	150	30
4	水域	100	20
5	其他耕地	120	30

year, DOY) 时序曲线。指数时序曲线可反映类别特点, 研究基于 NDVI、EVI 和 MNDWI 共 3 种指数的时序特征, 筛选区分地物类别的时序参数。

2.2 轮作区耕地的识别 采用分层分类加决策树的方法, 基于指数的时序特征对研究区中的玉米轮作区外进行分层掩膜^[13], 掩膜顺序为: 水体、建筑、森林、轮作区。由于 Land-

sat 重访周期较长, 受天气影响较大, 在识别算法中使用 MODIS 的 NDVI 产品来对轮作耕地进行最后的掩膜。

2.3 监督分类 利用目视样本点作为训练数据集, 通过 GEE 提供的云评分算法制作研究区最小云量合成影像, 并基于影像的蓝光、绿光、红光、近红外、短波红外等 7 个波段的合成影像, 使用分类回归树 (Classification And Regression Tree, CART) 算法完成轮作区耕地提取^[14]。

2.4 精度评价 利用独立的验证点对结果进行分类评价, 将非轮作区耕地的地类视为一类与轮作区耕地进行二分类评价, 通过准确率、精准率、F1 Score 以及 Kappa 系数对提取精度进行评价。

3 结果与分析

3.1 时序特征分析 由图 2a 可知, 水域 MNDWI 数值全年都为正值, 且均值大于 0.4 明显区别于其他地类, 基于 MNDWI

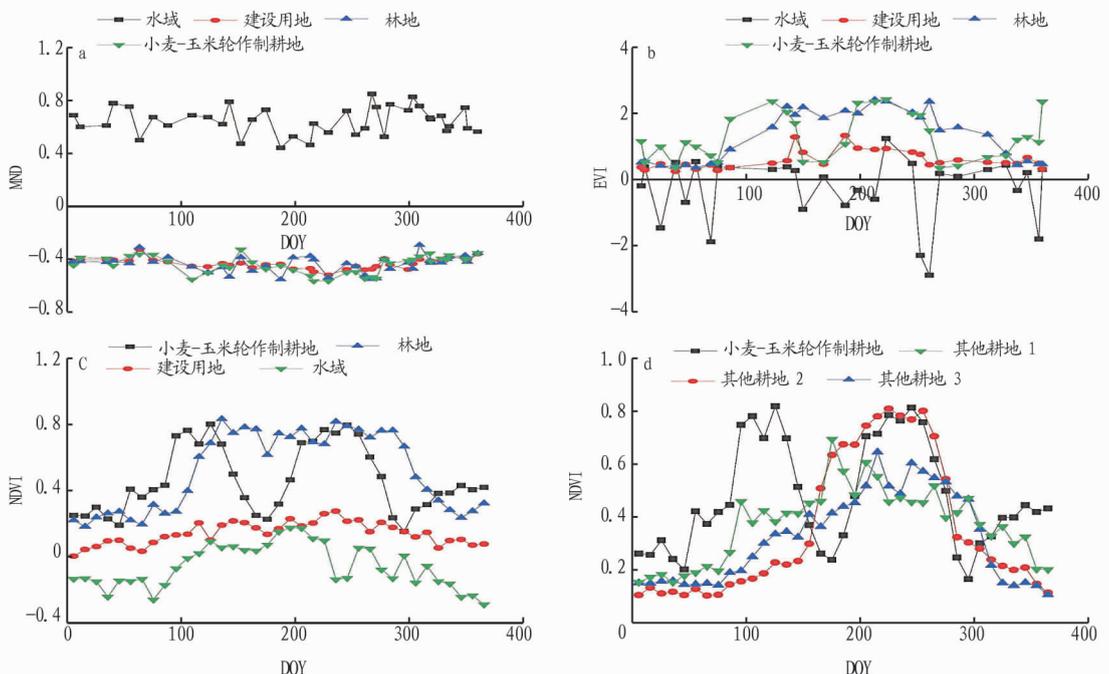


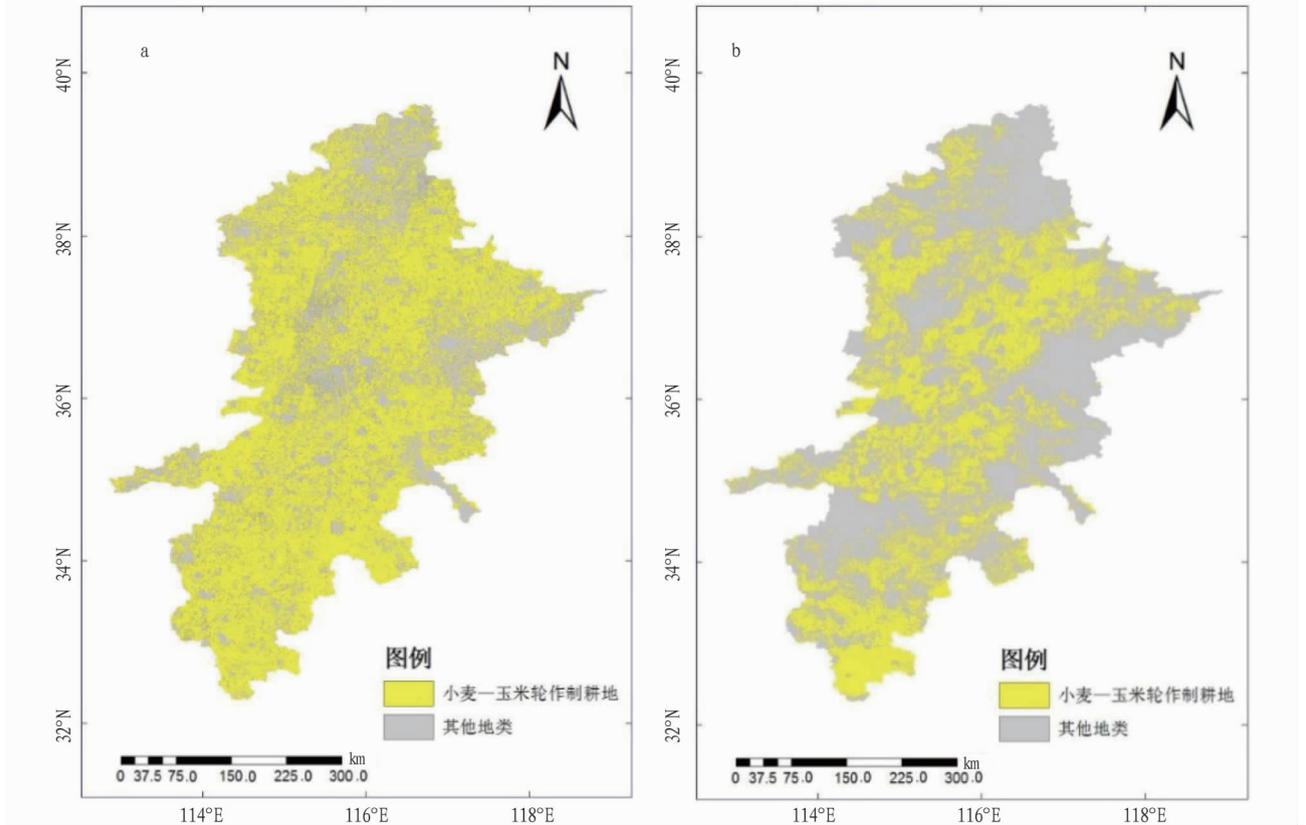
图 2 地类指数时序曲线

Fig. 2 Time series curve of land index

提取水域效果最优。图 2b 中森林和耕地在 EVI 上有较为明显的差异,在 DOY120~273 d,森林的 EVI 均值大于 2,明显区别于小麦收割后裸地。图 2c 中建设用地的 NDVI 曲线较为平缓,全年的数值偏低,且年均值小于 0.3,以此来区分建筑用地。图 2d 中玉米轮作区在 NDVI 时序上具有明显双峰

特征,明显区别于其他地类。儒略日第 79~154 天是第 1 个峰,第 198~271 天是第 2 个峰。

3.2 监督分类结果 为比较算法的精度,采用监督分类与识别算法进行比较。二分类结果如图 3 所示,分类结果中轮作耕地分布过多,明显与调查情况不一致。



注:a. 监督分类结果;b. 时序分类结果

注:a. Multi-spectral classification results;b. Time-series classification results

图 3 轮作区分类结果

Fig. 3 Classification results of rotation area

二分类混淆矩阵如表 2 所示,可见分类结果中,有大量轮作耕地被误分为非轮作耕地,同时也有大量非轮作耕地被误分为轮作耕地,2 种地类相互混淆,提取精度较低。

地和其他耕地被误分为小麦-玉米轮作区。

表 2 基于多光谱的二分类混淆矩阵

Table 2 Confusion matrix based on multispectral with two class

项目 Item	轮作耕地验证点 Verification site for rotation cropland	非轮作耕地验证点 Verification site for non-rotation cropland
轮作耕地预测 Rotation cropland prediction	86	14
非轮作耕地预测 Non-rotation cropland prediction	19	111

监督分类结果中主要城市、河流湖泊分类准确,但北部只有少量其他耕地被正确识别,且分类结果中林地分布偏少,面积为 33 651.6 km²,轮作耕地分布过多,面积为 121 661.7 km²,与调查情况不符。混淆矩阵如表 3 所示,林地、其他耕地与轮作区耕地之间错分较多,有较大比例的林

表 3 基于多光谱的分类提取多分类混淆矩阵

Table 3 Confusion matrix extracted based on multi-spectral classification

项目 Item	轮作耕地验证点 Verification site for rotation cropland	林地验证点 Verification site for forestland	建筑验证点 Verification site for buildings	水域验证点 Verification site for water area	其他耕地验证点 Verification site for other farmland
轮作耕地预测 Rotation cropland prediction	86	10	0	0	4
林地预测 Forestland prediction	6	39	0	1	4
建筑预测 Buildings prediction	0	0	30	0	0
水域预测 Water area prediction	0	0	0	20	0
其他耕地预测 Other farmland prediction	5	2	1	0	22

3.3 时序分类结果 基于时序数据使用识别算法提取结果精度较高,轮作区耕地面积为 68 032 km²,其中东北部(即黄河流域)、中部靠太行山一侧和偏南部分布较多,分布情况基本与调查结果相符。二分类混淆矩阵如表 4 所示,只有少量轮作耕地和非轮作耕地发生错分,该算法能较为精确地对轮作区进行识别。

表 4 基于时序特征的二分类混淆矩阵

Table 4 Confusion matrix based on time series feature with two class

项目 Item	轮作耕地验证点 Verification site for rotation cropland	非轮作耕地验证点 Verification site for non-rotation cropland
轮作耕地预测 Rotation cropland prediction	97	3
非轮作耕地预测 Non-rotation cropland prediction	2	128

3.4 精度对比 基于时序特征的识别算法准确率、精准率、F1 Score 和 Kappa 系数分别为 0.987、0.970、0.975 和 0.956,分别比监督分类高出 0.121、0.110、0.136 和 0.246。准确率、精准率、召回率和 F1 Score 越高,分类精度越高,这 4 个指标都是时序特征识别算法占优。Kappa 系数表示混淆矩阵的不平衡度,监督分类结果的 Kappa 系数远低于时序识别算法。

4 结论

该研究基于 GEE 平台,分别利用识别算法与监督分类,对研究区小麦、玉米轮作区耕地进行提取,并进行精度对比分析,得到以下结论:

(1) 该研究实现的提取算法可以快速有效地实现大区域小麦-玉米轮作耕地的提取,其二分类准确率、精准率、F1 Score 和 Kappa 系数分别为 0.987、0.970、0.975 和 0.956,分别比基于多光谱的分类高出 0.121、0.110、0.136 和 0.246。

(2) 某些特殊时期,林地、轮作耕地及其他耕地具有相似的光谱特征,传统基于多光谱数据的监督分类难以区分,导致其难以准确地提取轮作耕地。

(3) 研究区 2018 年约有小麦-玉米轮作耕地 68 032 km²,约占研究区总面积的 30.7%,主要分布在河北东南部、河南东北部和山东西北部区域。

利用遥感数据的各类归一化指数的时序特征,能够比利用多光谱数据的监督分类更准确地地区分具有类似光谱特征的地物,如轮作耕地与其他耕地,而在 GEE 平台的支持下,能快速获取并处理大区域、长时间的遥感数据,从而快速通过时序特征来区分不同地物。

参考文献

- [1] 史舟,梁宗正,杨媛媛,等. 农业遥感研究现状与展望[J]. 农业机械学报,2015,46(2):247-260.
- [2] 熊勤学,黄敬峰. 利用 NDVI 指数时序特征监测秋收作物种植面积[J]. 农业工程学报,2009,25(1):144-148.
- [3] 王菲,李红华,苏浩. 基于 MODIS-NDVI 的邯郸市冬小麦和夏玉米面积提取[J]. 现代农业科技,2018(24):3-5.
- [4] 路中,雷国平,马泉来,等. 基于重构的 Landsat 8 时间序列数据和温度植被指数的区域旱情监测[J]. 水土保持研究,2018,25(5):371-377,384.
- [5] 苏伟,姜方方,朱德海,等. 基于决策树和混合像元分解的玉米种植面积提取方法[J]. 农业机械学报,2015,46(9):289-295,301.
- [6] 翟世常,王学强,颜耀文. 黑河流域中游盆地玉米作物遥感估产研究[J]. 安徽农业科学,2015,43(31):350-353.
- [7] 刘佳,王利民,杨福刚,等. 基于 GF-1 卫星影像的中国冬小麦制图研究[J]. 中国农学通报,2019,35(9):155-164.
- [8] 刘沼辉,柳林,郭慧,等. 基于 GF1-NDVI 时序影像对春小麦进行提取研究[J]. 北京测绘,2018,32(6):643-646.
- [9] 赵丽花,李卫国,杜培军. 基于多时相 HJ 卫星的冬小麦面积提取[J]. 遥感信息,2011,26(2):41-45,50.
- [10] 谭深,吴炳方,张鑫. 基于 Google Earth Engine 与多源遥感数据的海南水稻分类研究[J]. 地球信息科学学报,2019,21(6):937-947.
- [11] 侯学会,隋学艳,姚慧敏,等. 中国北方麦区冬小麦物候期对气候变化的响应[J]. 麦类作物学报,2019,39(2):202-209.
- [12] FOGA S, SCARAMUZZA P L, GUO S, et al. Cloud detection algorithm comparison and validation for operational Landsat data products[J]. Remote sensing of environment, 2017,194:379-390.
- [13] 向海燕,罗红霞,刘光鹏,等. 基于 Sentinel-1A 极化 SAR 数据与面向对象方法的山区地表覆被分类[J]. 自然资源学报,2017,32(12):2136-2148.
- [14] DIDAN K. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [EB/OL]. [2020-11-05]. <https://doi.org/10.5067/MODIS/MOD13Q1.006>.
- [15] 龙江水利科技,2020,48(4):158-160.
- [16] 常晖,曾艾林,贾志伟,等. 新疆红花种子质量分级标准研究[J]. 中药材,2019,42(3):501-504.
- [17] 张龔,任海龙,王豪杰,等. 新疆红花地方种种子含油率与主要农艺性状间的灰色关联分析[J]. 分子植物育种,2020,18(5):1637-1648.
- [18] 王秀珍,周相军,王丽. 不同种植密度对红花产量及农艺性状的影响[J]. 新疆农业科学,2010,47(8):1557-1560.
- [19] 彭云承,艾合买提江,加力肯,等. 伊犁河谷红花 3414 肥效试验[J]. 安徽农业科学,2017,45(13):119-120.
- [20] 宋魁,谭勇,龚昌禄,等. 种植密度和灌溉次数对红花生长发育和产量的影响[J]. 中国农村小康科技,2008(7):60-62.
- [21] 温学萍,金波,赵玮,等. 宁夏红花新品种种植密度试验研究[J]. 宁夏农林科技,2017,58(10):6-8.
- [22] 牛新霞. 新疆维吾尔自治区红花栽培技术[J]. 中国农技推广,2020,36(8):50-51.
- [23] 贾宏涛,谭勇,孙霞,等. 施肥对红花生长和产量的影响[J]. 新疆农业大学学报,2010,33(5):394-397.
- [24] 杨会峰. 新疆红花丝机械化采收研究现状[J]. 新疆农业机械化,2020(5):34-37.
- [25] 李荣. 新疆红色产业发展状况分析与对策[J]. 新疆农业大学学报,2003,26(3):92-96.

(上接第 201 页)

- [8] 王兆木. 发挥新疆资源优势 发展特种油料作物[J]. 新疆农业科学,2003,40(1):10-15.
- [9] 王果平,帕丽达,李晓瑾,等. 药用植物红花新疆产地适应性数值分析[J]. 中国民族民间医药,2010,19(23):49-50.
- [10] 李春艳,顾秋丽,高明清. 伊犁河谷红花产业现状及发展对策[J]. 中国种业,2020(2):27-28.
- [11] 任红松,陈彤,黄润,等. 基于 SWOT 分析察布查尔县红花产业现状分析与对策研究[J]. 新疆农业科学,2014,51(8):1556-1563.
- [12] 宋玉龙. 名贵中药材红花的资源调查及质量评价研究[D]. 乌鲁木齐:新疆医科大学,2015.
- [13] 任超翔,吴沂芸,唐小慧,等. 红花的起源与产地变迁[J]. 中国中药杂志,2017,42(11):2219-2222.
- [14] 王引权,王艳,陈红刚,等. 海拔梯度对药用植物品质形成影响的研究进展[J]. 中国现代中药,2012,14(5):41-44.
- [15] 康东健,谭勇,阚萌萌,等. 不同海拔梯度新疆红花品质分析研究[J]. 中药材,2017,40(1):127-130.
- [16] 王兆木,陈跃华,买买提明,等. 红花新品种选育简报[J]. 新疆农业科学,2000,37(4):165-167.
- [17] 孙兆强. 膜下滴灌技术在中西部地区农田灌溉中的运用研究[J]. 黑