

统计数据总量约束下全局优化阈值的冬小麦分布制图

郭文茜^{1,2}, 任建强^{1,2}, 刘杏认³, 陈仲新^{1,2}, 吴尚蓉^{1,2}, 潘海珠^{1,2}

1. 中国农业科学院 农业资源与农业区划研究所, 北京 100081;

2. 农业部农业遥感重点实验室, 北京 100081;

3. 中国农业科学院 农业环境与可持续发展研究所, 北京 100081

摘要: 大范围、长时间和高精度农作物空间分布基础农业科学数据的准确获取对资源、环境、生态、气候变化和国家粮食安全等问题研究具有重要现实意义和科学意义。本文针对传统阈值法农作物识别过程中阈值设置存在灵巧性差和自动化程度低等弱点, 以中国粮食主产区黄淮海平原内河北省衡水市景县为典型实验区, 首次将全局优化算法应用于阈值模型中阈值优化选取, 开展了利用全局优化算法改进基于阈值检测的农作物分布制图方法创新研究。以冬小麦为研究对象, 国产高分一号(GF-1)为主要遥感数据源, 在作物面积统计数据为总量控制参考标准和全局参数优化的复合型混合演化算法SCE-UA (Shuffled Complex Evolution-University of Arizona)支持下, 提出利用时序NDVI数据开展阈值模型阈值参数自动优化的冬小麦空间分布制图方法。最终, 获得实验区冬小麦阈值模型最优参数, 并利用优化后的阈值参数对冬小麦空间分布进行提取。通过地面验证表明, 利用本研究所提方法获取的冬小麦识别结果分类精度均达到较高水平。其中冬小麦识别结果总量精度达到了99.99%, 证明本研究所提阈值模型参数优化方法冬小麦提取分类结果总量控制效果良好; 同时, 与传统的阈值法、最大似然和支持向量机等分类方法相比, 本研究所提阈值模型参数优化法区域冬小麦作物分类总体精度和Kappa系数分别都有所提高, 其中, 总体精度分别提高4.55%、2.43%和0.15%, Kappa系数分别提高0.12、0.06和0.01, 这体现出SCE-UA全局优化算法对提高阈值模型冬小麦空间分布识别精度具有一定优势。以上研究结果证明了利用本研究所提基于作物面积统计数据总量控制以及SCE-UA全局优化算法支持下阈值模型参数优化作物分布制图方法的有效性和可行性, 可获得高精度冬小麦作物空间分布制图结果, 这对提高中国冬小麦空间分布制图精度和自动化水平具有一定意义, 也可为农作物面积农业统计数据降尺度恢复重建和大范围区域作物空间分布制图研究提供一定技术参考。

关键词: 遥感, 作物分布制图, 阈值法, 总量控制, 全局优化算法, 冬小麦

中图分类号: TP79 **文献标志码:** A

引用格式: 郭文茜, 任建强, 刘杏认, 陈仲新, 吴尚蓉, 潘海珠. 2018. 统计数据总量约束下全局优化阈值的冬小麦分布制图. 遥感学报, 22(6): 1023–1041

Guo W Q, Ren J Q, Liu X R, Chen Z X, Wu S R and Pan H Z. 2018. Winter wheat mapping with globally optimized threshold under total quantity constraint of statistical data. *Journal of Remote Sensing*, 22(6): 1023–1041 [DOI:10.11834/jrs.20187468]

1 引言

近些年来, 随着国内外对资源环境、气候变化、生态和灾害等对国家或地区粮食生产系统影响等一系列问题定量研究的逐步深入, 国家农业科学空间基础数据对上述事关国家粮食安全和应对气候变化等问题研究和科学决策具有更加重要

的意义(Qiu 等, 2003; Ye 等, 2013)。其中, 大范围、长时间和高精度农作物分布数据准确获取可为气候变化对国家粮食安全、粮食区域平衡、作物种植制度、生产结构和区域布局等方面影响研究提供重要基础数据, also 具有重要现实意义和科学意义(Barnwal和Kotani, 2013)。

遥感信息由于具有覆盖面积大、探测周期

收稿日期: 2017-11-10; 预印本: 2018-05-16

基金项目: 国家自然科学基金(编号: 41471364, 61661136006); 国家高技术研究发展计划(863计划)(编号: 2012AA12A307)

第一作者简介: 郭文茜, 1992年生, 女, 硕士研究生, 研究方向为农业遥感应用。E-mail: guowenqian1992@163.com

通信作者简介: 任建强, 1974年生, 男, 研究员, 研究方向为农业遥感基础研究与应用。E-mail: renjianqiang@caas.cn

短、资料丰富和费用低等特点,已经在大范围作物空间分布信息获取和作物空间分布制图中发挥了越来越重要作用。冬小麦是中国主要的粮食作物之一,也是重要的商品粮和战略性储备粮。因此,开展准确获取冬小麦分布信息的作物识别和作物分布制图方法研究不仅可为气候变化、资源环境和粮食安全等领域研究提供重要基础科学数据,而且也对农业种植结构调整、农业政策科学制定等具有重要的现实意义。目前,常用的农作物识别遥感提取方法大体分为3类,第1类为基于地物空间光谱信息的差异进行农作物遥感识别的方法,第2类为基于时序遥感数据提取作物物候特征发展出的农作物遥感识别方法,第3类则是基于多源数据融合的农作物遥感识别方法(许文波和田亦陈, 2005)。

基于光谱特征的农作物遥感识别方法较早是以目视解译为主,之后出现了利用计算机基于分类图像统计进行分类的监督分类和非监督分类法,随着多学科交叉领域的发展,以句法结构分类法为主的各种集成分类方法(如神经网络法、模糊数学法、决策树法和基于混合像元分解的方法等)逐渐得到应用(唐华俊等, 2010)。但由于受卫星影像分辨率限制,导致分类过程中“同物异谱”和“同谱异物”现象难以避免,所以单纯基于地物光谱特性进行复杂种植结构作物分类较难获得理想效果(王娜等, 2017)。

对基于时序遥感数据提取作物物候特征的作物遥感识别方法而言,不同作物随着时相的变化具有明显的变化特征,基于时间序列遥感数据提取作物物候时相变化规律可以实现目标作物的识别,该方法通过分析时序数据中作物关键物候期特征进行作物类型识别,可避免利用单一时相影像数据进行作物空间分布提取时因“异物同谱”等原因导致的错分、漏分等现象。其中,基于时序影像单一特征量(如NDVI、EVI等)和作物物候特征的阈值法,不仅利用了作物随时间的生长变化规律,具有操作简单、运算速度快的特点,并且在作物种植类型较为单一的地区取得了较高的识别精度,因此,已经成为农作物特别是冬小麦识别中最常用的方法之一(蔡学良和崔远来, 2009; 张健康等, 2012; 雷光斌等, 2014; 刘吉凯等, 2015; Salahddine等, 2017)。近些年来,国内外学者已经开展较多基于阈值法和光学遥感数据的

多种作物(如冬小麦、玉米、水稻、大豆等)分类研究,分类总体精度一般高于80%(Foerster等, 2012; 黄青等, 2013; 许青云等, 2014; 杨闫君等, 2015; 梁守真等, 2015)。然而,传统阈值法大多根据农作物关键时间点物候历特征及其“光谱—时序”曲线特征确定区分目标作物与其他地物的最佳时间点(胡琼等, 2015; 郝震等, 2017),最佳区分点的植被指数阈值大小主要通过训练样本光谱统计分析方法确定(蔡学良和崔远来, 2009; 刘佳等, 2015; 张森等, 2015),或采用试错法,人工不断调试植被指数阈值的方式,根据经验对各关键物候期阈值不断调整尝试并多次分类,选择分类结果精度较高的一组阈值(张霞等, 2008; 黄青等, 2013)。尽管上述方法能够取得较高精度结果,但传统阈值检测方法阈值设置缺乏灵巧性与智能性,不仅需耗费大量的工作量,主观性强,阈值检测结果存在较大偏差,从而阻碍了阈值方法作物识别精度进一步提高,阈值优化设置问题是阈值法作物识别中重要关键科学问题。近年国内外学者对阈值法阈值合理设置进行了一系列研究(Foerster等, 2012; 张健康等, 2012; 郝鹏宇等, 2012; 王利民等, 2016),如利用地面样点时序植被指数参考曲线相似性阈值确定法、约束条件下的植被指数特征值最小阈值法和基于阈值分割的阈值筛选法等,但将全局优化思想引入阈值优化设置问题中的相关研究未见报道,且阈值设置准确性和自动化程度仍然有待进一步提高。

基于多源数据农作物遥感识别方法一类是遥感信息与非遥感信息融合的作物识别方法,另一类是多种传感器获取的具有不同时间、空间分辨率和不同波段信息的多源遥感数据融合的作物识别方法。其中,由于农作物面积统计数据在气候变化、国家粮食安全等领域是不可替代的重要基础数据,国内外学者已将统计数据作为非遥感信息与遥感信息进行融合,从而实现了大范围作物空间分布信息提取,这为全球变化与粮食安全研究提供了可靠的作物空间分布基础数据(Barnwal和Kotani, 2013; 宋茜等, 2015)。如部分学者基于交叉信息熵原理的作物空间分配模型(SPAM)将遥感信息与农业统计数据融合,已在全球尺度及区域尺度上获得了较高精度的作物空间分布结果(Liu等, 2008; 刘珍环等, 2013)。但是,目前一般遥感信息与非遥感信息融合的作物空间分布制

图技术大多将遥感获取的土地利用、农业灌溉和耕地适宜性等作为辅助信息,而对研究对象作物自身生长机理及其变化规律考虑较少,没有充分地对作物自身遥感信息(特别是时序遥感信息)加以直接应用,从而阻碍了作物空间分布制图精度的进一步提高。近年来,一些学者研究表明,将NDVI等遥感特征参量引入遥感信息与非遥感信息(如统计数据)融合的作物空间分布制图中,降低了分类规则对训练样本的依赖,不仅方法简单,而且容易操作,可有效提高制图精度和效率(Pervez和Brown, 2010; Brown和Pervez, 2014)。

综合以上分析,本研究针对阈值法农作物识别过程中阈值设置存在灵巧性差和自动化程度低等弱点,提出引入全局优化算法实现作物识别模型阈值参数选取的方法,通过选择作物识别总量约束以及利用全局优化算法优化阈值检测方法两个方面进行改进,开展利用全局优化算法改进基于阈值检测的农作物分布制图方法研究,以期实现作物面积提取模型阈值参数快速准确的自动率定,从而进一步提高阈值法农作物识别精度。其中,本文以中国粮食主产区黄淮海地区河北省衡水市景县为典型实验区,以冬小麦为研究对象,以国产高分一号(GF-1)数据为主要遥感数据源,在阈值模型和全局优化算法支持下,开展以下两方面研究:首先,选择合适的总量约束措施,确定全局优化的标准参考;其次,确定合适的全局优

化算法对阈值模型中待优化阈值参数进行设置,以提高参数计算效率和制图精度。通过研究,实现阈值法阈值参数自动优化选取,提高阈值选取客观性和准确性,从而进一步提高阈值法冬小麦空间分布制图的工作效率和分类精度,为获取大范围、长时间和高精度农作物空间分布基础农业科学数据提供一定思路借鉴。

2 实验区域及数据

2.1 实验区概况

本研究实验区域($115^{\circ}54'E-116^{\circ}28'E$, $37^{\circ}28'N-37^{\circ}52'N$)位于中国北方粮食生产基地黄淮海平原内河北省衡水市景县,该县地势较为平坦,海拔在14.1—25 m。县域面积共1183 km²(图1),其中,耕地面积约830 km²,以冬小麦和夏玉米为主要农作物,二者为一年两熟轮作制,种植总面积占总耕地面积的80%以上。实验区所在区域属温带半湿润季风气候,大于或等于0℃年积温4200℃—5500℃,年累计辐射量 $5.0\times 10^6-5.2\times 10^6$ KJ·m⁻²,无霜期170—220 d,年均降水量500—900 mm。实验区冬小麦集中种植时间为上年10月上中旬,11月下旬—12月上旬达到冬前分蘖峰值,12月中旬开始停止生长,第二年下年3月上中旬进入返青期,拔节期为3月下旬至4月上中旬,抽穗期为4月下旬到5月上旬,乳熟期为5月中下旬,成熟期为6月上旬。

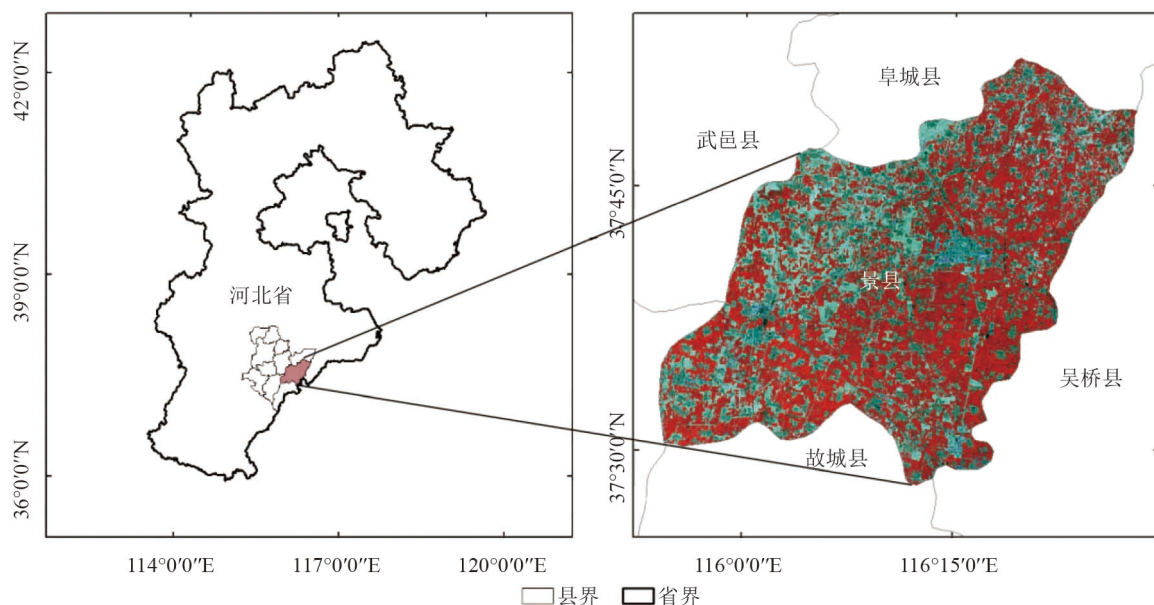


图1 实验区地理位置图

Fig. 1 Location of experimentation area

2.2 数据获取与处理

2.2.1 遥感数据

本研究主要利用国产GF-1高分辨率遥感卫星数据进行冬小麦分布提取。GF-1卫星搭载4台16 m分辨率的多光谱相机(WFV1-WFV4),光谱范围为0.45—0.89 μm ,包含蓝(0.45—0.52 μm)、绿(0.52—0.59 μm)、红(0.63—0.69 μm)和近红外(0.77—0.89 μm)4个波段,多光谱相机组合幅宽达800公里,重访周期为4d。本文数据获取时间为2014年10月至2015年6月,影像来源于中国资源卫星应用中心(<http://www.cresda.com/CN/> [2017-11-10])推送至农业部遥感应用中心的GF-1数据。选取影像时,以影像云量较低质量较好,并且影像时间间隔尽量一致为原则。最终,本研究共获得覆盖2015年冬小麦整个生长期的15景较高质量影像,所获取GF-1遥感数据详细信息见列表1。

表1 实验区高分一号影像信息列表
Table 1 GF-1/WFV image list in experimentation area

时相序号	传感器类型	轨道号	获取时间
01	WFV2	L1A0000393184	2014-10-16
02	WFV4	L1A0000427927	2014-11-02
03	WFV3	L1A0000454185	2014-11-14
04	WFV2	L1A0000580482	2014-12-04
05	WFV1	L1A0000538637	2014-12-24
06	WFV2	L1A0000580482	2015-01-10
07	WFV3	L1A0000639856	2015-02-08
08	WFV4	L1A0000673553	2015-03-01
09	WFV2	L1A0000725498	2015-03-29
10	WFV1	L1A0000750015	2015-04-14
11	WFV1	L1A0000765287	2015-04-22
12	WFV3	L1A0000781757	2015-05-01
13	WFV4	L1A0000812059	2015-05-18
14	WFV1	L1A0000826550	2015-05-25
15	WFV3	L1A0000850828	2015-06-07

2.2.2 遥感影像预处理

研究中,GF-1遥感影像预处理过程主要包括正射校正、辐射定标、大气校正、几何精校正等,数据预处理过程主要在ENVI5.3中完成。其中,借助高分一号影像自带的RPC(Rational Polynomial

Coefficient)文件对原始影像进行无控制点的有理多项式模型区域网平差正射校正。辐射定标利用中国资源卫星应用中心(<http://www.cresda.com/CN/> [2017-11-10])提供的绝对定标系数对经过正射校正后的GF-1影像进行辐射定标获得了影像的辐亮度文件。其中辐射定标公式如下:

$$L(\lambda) = \text{Gain} \cdot \text{DN} + \text{Bias} \quad (1)$$

式中, $L(\lambda)$ 为转换后辐亮度($\text{W}/(\text{m}^2 \cdot \text{sr} \cdot \mu\text{m})$),Gain为定标斜率($\text{W}/(\text{m}^2 \cdot \text{sr} \cdot \mu\text{m})$),DN为观测灰度值,Bias为定标截距。其中,Gain和Bias均来自中国资源卫星应用中心。

采用ENVI5.3软件中自带的FLAASH模型对影像进行大气校正以获得真实地表反射率数据。最后以15 m分辨率的Landsat 8/OLI影像(<https://glovis.usgs.gov/> [2017-11-10])及地面实际采集样点作为参考进行几何精校正,误差控制在一个像元内。

2.2.3 植被指数计算

归一化差值植被指数NDVI(Normalized Difference Vegetation Index)是反映植被生长状态及植被覆盖度的最佳指示因子之一,目前已被广泛应用于农作物识别和面积估算等方面(Zhang等, 2017; Saeed等, 2017)。其中,基于时序植被指数数据提取冬小麦面积与类型的研究方法已经比较成熟(Goodwin和Collett, 2014; Massey等, 2017),因此,本研究也将采用时序高分一号NDVI数据开展冬小麦识别研究。其中,通过大气校正、几何精校正等预处理后,GF-1 NDVI具体计算公式如下:

$$\text{NDVI} = (\rho_{\text{NIR}} - \rho_{\text{R}}) / (\rho_{\text{NIR}} + \rho_{\text{R}}) \quad (2)$$

式中, ρ_{NIR} 为GF-1卫星WFV传感器近红外波段的反射率; ρ_{R} 为GF-1卫星WFV传感器红光波段的反射率。NDVI值域范围为[-1, 1]。本研究中部分GF-1遥感卫星NDVI结果如图2所示。

2.2.4 地面样本点数据

本研究于2015年4月在开展了实验区冬小麦分布提取训练样本、验证样本的信息采集。根据实地调查,由于该地区在冬小麦种植时期主要影响冬小麦识别的地物为林地,因此,实验区共分为冬小麦、林地、其他(裸地、建筑用地等)3个类别。其中,林地及其他在最终分类中划分为非冬小麦类别。实地调查中,地物类型样点分布主要遵循均匀分布原则,并利用差分GPS进行样本点定

位, 最终, 共获得2653个地面实测各种地物类型样点。其中, 冬小麦1860个, 林地385个, 其他类型408个。研究中, 随机选择600个样本点作为训

练样本, 用于提取地物时序特征, 剩下的2053个样本点作为验证样本, 用于后续提取结果的精度验证, 具体样本点数量和分布, 如表2和图3所示。

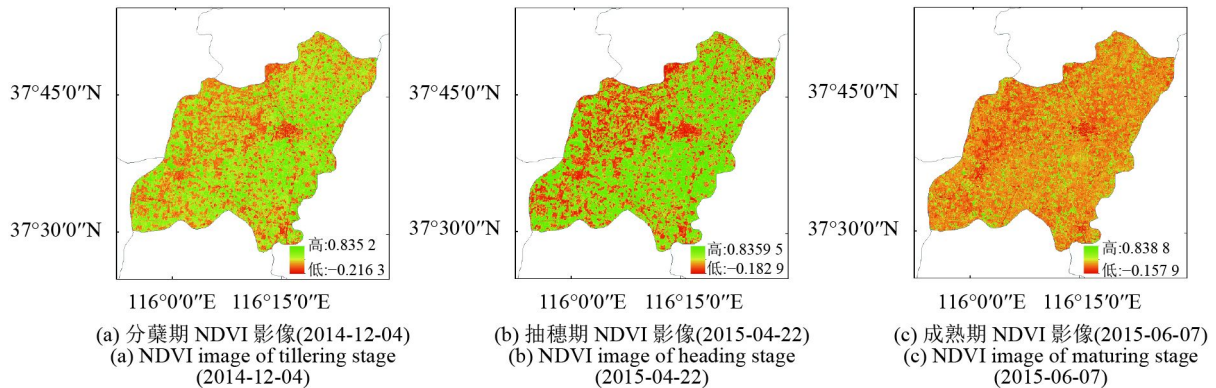


图2 实验区景县GF-1 NDVI计算结果

Fig. 2 GF-1 NDVI data of Jingxian county in 2015

表2 实验区地面样本点采集数量
Table 2 Number of samples in Jingxian county

样本类别	验证点采集数量/个			总计
	冬小麦	非冬小麦		
		林地	其他	
训练样本	400	100	100	600
验证样本	1460	285	308	2053

该数据主要用于实验区冬小麦分布提取方法研究中的行政单元内冬小麦种植面积总量控制参数。

3 研究方法

3.1 方法流程

本研究在冬小麦时序NDVI参考曲线基础上, 利用冬小麦关键物候期特征建立基于阈值法的冬小麦分布提取阈值模型。其中, 模型中的阈值参数作为待优化参数。在此基础上, 以区域农作物种植面积统计数据为作物种植面积总量全局优化的标准参考, 用于和冬小麦面积分布提取模型获取的冬小麦面积建立优化目标函数。然后, 通过SCE-UA全局优化算法寻找最小目标函数值。当遥感提取的区域冬小麦面积与区域冬小麦面积统计数据间差异最小时, 便可得到区域内冬小麦分类模型最优参数, 并在最优参数下输出冬小麦空间分布信息, 从而获得高精度的冬小麦空间分布结果。最后, 完成对冬小麦空间分布结果的精度评价和对比分析, 具体如图4所示。

其中, 利用SCE-UA算法对阈值模型中待优化参数进行全局优化的主要过程是: 首先利用训练样本, 确定各待优化阈值参数的上下界; 然后, SCE-UA全局优化算法按照其自身搜索法则在上下界中寻找一组阈值参数作为初始值, 代入冬小麦分布提取阈值模型中, 获得研究区冬小麦空间分布结果, 并进行像元统计获得区域冬小麦种植面积。将该利用遥感提取的区域冬小麦种植面积数

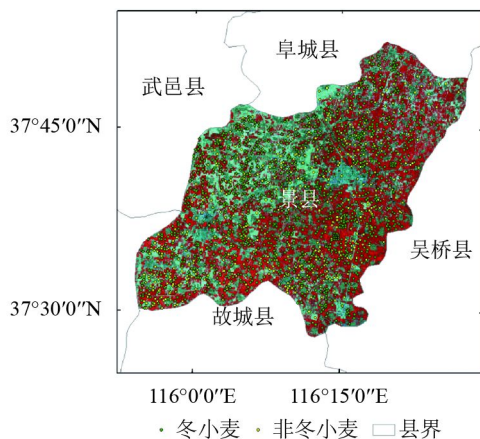


图3 实验区野外采集样本点分布

Fig. 3 Distribution of sample points in experimentation area

2.2.5 其他辅助数据

本研究涉及的辅助数据主要包括研究区行政边界矢量图(县、市、省级)、县级农业统计数据以及冬小麦物候信息等。其中, 景县2015年冬小麦种植面积县级统计数据主要来源于衡水市统计年鉴(衡水市统计局和国家统计局衡水调查队, 2016),

据代入目标函数中, 计算目标函数值(即利用遥感提取的通过像元统计获得的区域冬小麦面积与区域冬小麦面积统计数据间差值), 判断目标函数值是否满足收敛条件: 若否, 则重新初始化阈值, 返回至确定阈值初始值的步骤, 重新赋予一组阈值参数初值, 并进行后续步骤; 若满足收敛条件, 则输出该组阈值, 即为所求的最优化阈值。

3.2 冬小麦NDVI时序构建及特征提取

在野外采集的实验区样本数据库支持下, 利用训练样本在时序遥感影像中分别提取冬小麦及其他各类地物类型的NDVI时间序列数据。最终, 绘制不同地物NDVI均值形成的NDVI时序特征曲线, 如图5所示。由于研究中采用高分辨率遥感影像, 很难获得时间间隔一致的影像, 因此, 图5中显示的特征曲线为非等时间间隔特征曲线。由图5中可以看出, 结合实验区内物候信息, 冬小麦各个生育期在时序影像上的表征为10月上中旬为冬小麦播种期, 此时影像中主要表现为裸地光谱信息, 整体NDVI值偏低; 10月下旬至11月上旬为冬小麦出苗期, NDVI值逐渐升高, 直到冬前分蘖期, NDVI值达到比较高的水平, 此时出现了冬小麦整个生长季的第一个波峰; 之后冬小麦进入越冬期, 叶绿素含量降低, 光合作用减弱, 植株基本停止生长, NDVI值逐渐减小, 并达到波谷; 3月上旬开始, 冬小麦进入返青—拔节期, NDVI值迅速增长, 是NDVI曲线斜率最大的时期, 并于

4月下旬至5月上旬进入冬小麦抽穗期, NDVI达到最大值, 此时出现冬小麦整个生长季的第2个波峰; 之后进入灌浆乳熟期, NDVI值不断降低, 直至冬小麦成熟。

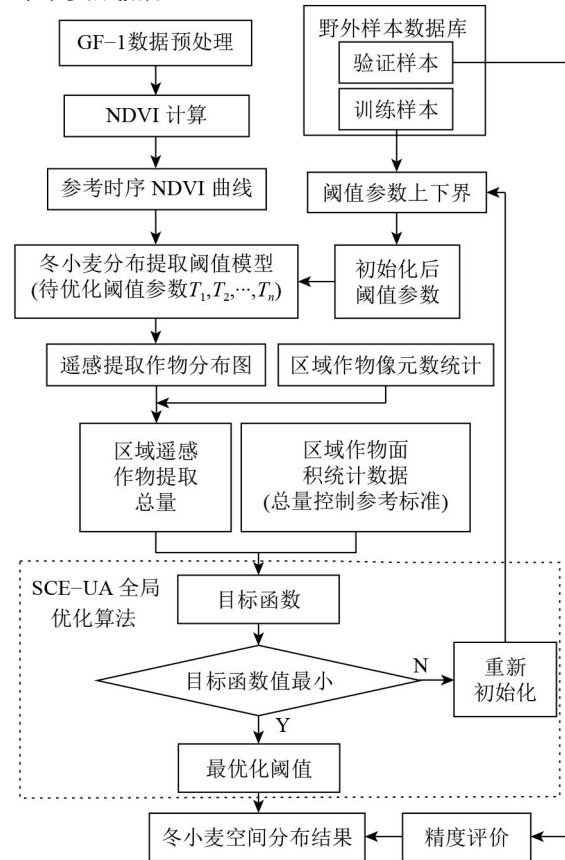


图4 研究方法流程图

Fig. 4 Flowchart of research

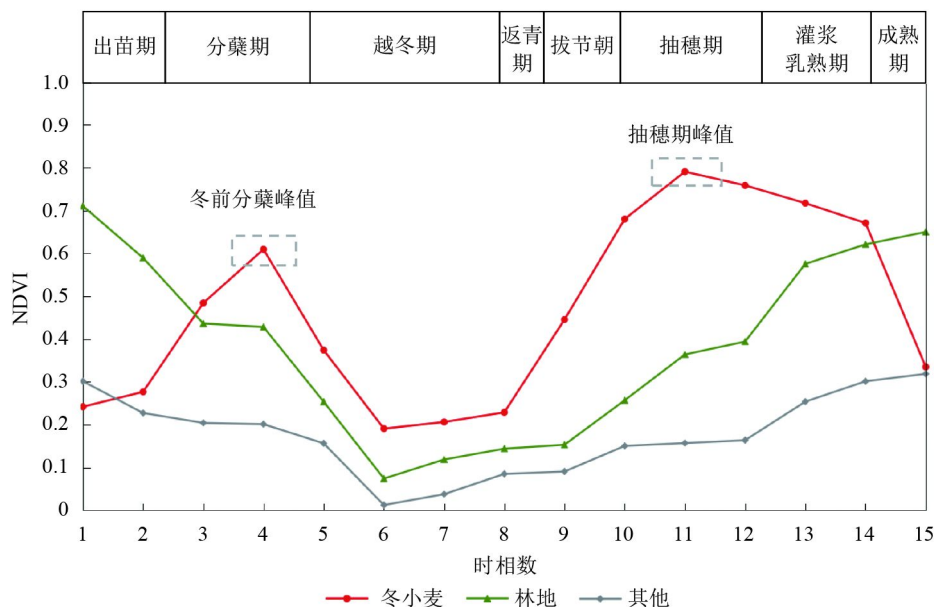


图5 不同地物GF-1/WFV NDVI时序曲线

Fig. 5 GF-1/WFV NDVI time series curve of different features

从图5分析可知,与非冬小麦(如林地和其他地物等)对比,整个生长季内只有冬小麦出现了明显的两次波峰,并且两次波峰的NDVI值均明显大于非冬小麦。在出苗期—分蘖期和抽穗期—成熟期,冬小麦与非冬小麦的时序植被指数特征差异较大。在出苗期到分蘖期,冬小麦NDVI值迅速增大,而林地NDVI值却迅速减小,其他类型NDVI值波动较小。在抽穗期到成熟期,冬小麦NDVI值迅速减小,而林地地物类型NDVI值反而逐渐升高,其他类型NDVI变化幅度较小。因此,利用冬小麦生长期内的关键物候期作为区分冬小麦的识别特征,从而建立冬小麦识别模型。

3.3 阈值法冬小麦空间分布提取模型建立

根据以上对冬小麦及其他各类地物类型的物候特征分析,本研究利用冬小麦出苗期、冬前分蘖期、拔节期、抽穗期和成熟期等关键物候期以及冬小麦双峰波形特征NDVI时序变化规律,确定不同物候期下冬小麦像元需要满足的判别函数和特征参数,从而建立基于时序NDVI和物候特征的冬小麦空间分布提取模型。

(1) 根据冬小麦冬前分蘖期和抽穗期平均NDVI值均达到峰值的特征,建立判别函数:

$$\text{NDVI}_{t_4} > T_1 \text{ 且 } \text{NDVI}_{t_{11}} > T_2 \quad (3)$$

式中, T_1 表示冬小麦第4时相NDVI应满足的阈值, T_2 表示冬小麦第11时相NDVI应满足的阈值。

(2) 根据只有冬小麦出苗—分蘖期和拔节期平均NDVI值呈快速增长的特征,建立判别函数:

$$\text{NDVI}_{t_4} - \text{NDVI}_{t_2} > T_3 \text{ 且 } \text{NDVI}_{t_{11}} - \text{NDVI}_{t_9} > T_4 \quad (4)$$

式中, T_3 表示冬小麦第4时相与第2时相NDVI差值应满足的阈值, T_4 表示冬小麦第11时相与第9时相NDVI差值应满足的阈值。

(3) 根据只有冬小麦进入越冬期和成熟期平均NDVI值呈快速下降的特征,建立判别函数:

$$\text{NDVI}_{t_4} - \text{NDVI}_{t_6} > T_5 \text{ 且 } \text{NDVI}_{t_{14}} - \text{NDVI}_{t_{15}} > T_6 \quad (5)$$

式中, T_5 表示冬小麦第4时相与第6时相NDVI差值应满足的阈值, T_6 表示冬小麦第14时相与第15时相NDVI差值应满足的阈值;

式(3)到式(5)中, $t_i(i=1, 2, \dots, c)$ 代表冬小麦整个生育期开始对应的 c 个时相; NDVI_{t_i} 表示某一像元在 t_i 时相对应的NDVI平均值; $T_i(i=1, 2, \dots, d)$ 表示冬小麦各判别函数对应的阈值,为待优化的参数。本研究中时相数 $c=15$,待优化阈值参数个数 $d=6$ 。以上3个判别函数组成了衡水市景县冬小麦空间分布提取阈值模型,如图6所示。

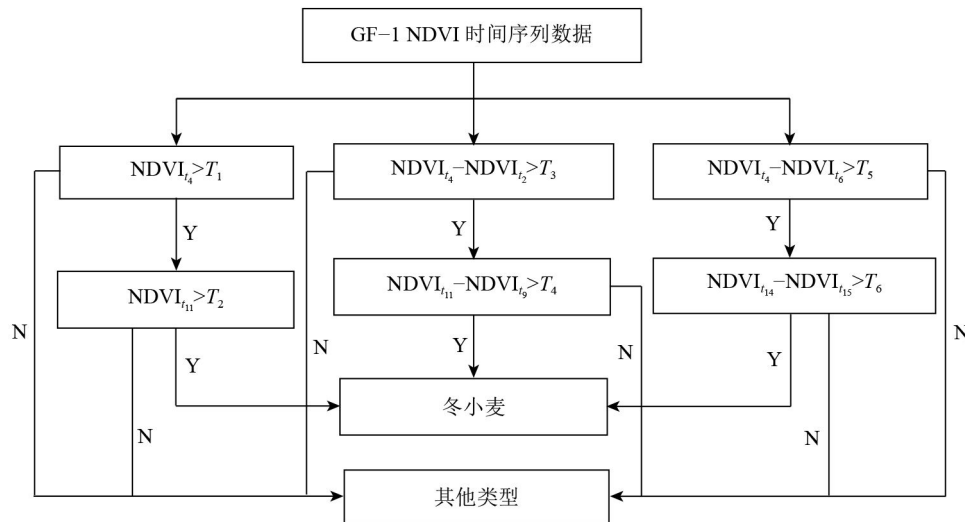


图6 阈值模型冬小麦识别流程

Fig. 6 The flowchart of winter wheat identification by threshold method

3.4 统计数据总量约束下阈值参数全局优化

3.4.1 全局优化算法的选择

针对传统阈值法中阈值设定存在的弱点,提出利用全局优化算法对阈值参数进行优化选取以

提高参数计算效率和制图精度的方法。目前国内常用传统的全局优化算法主要有遗传算法(葛继科等, 2008)、模拟退火算法(林慧君和彭宏, 2006)和蚁群算法等(段海滨等, 2007)。然而, 遗传算法

全局优化效果较好, 但该算法局部搜索能力较弱, 运算量大, 运行时间长, 且对于参数敏感从而优化结果易受影响。模拟退火算法虽然运行效率高, 操作简单, 局部搜索能力强, 但该算法是一种随机搜索算法, 全局搜索能力差, 容易陷入局部最优解因而难以获得全局最优解(王雪梅和王义和, 1997; 卢宇婷等, 2015)。(Duan等, 1994, 1993)针对水文模型率定提出的SCE-UA(Shuffled Complex Evolution-University of Arizona)复合形混合演化算法, 综合了遗传算法和控制随机搜索算法的优点, 是一种对于非线性复杂模型最优化问题寻找全局最优解最为成功的方法之一(董洁平等, 2012)。并且具有可以同时优化模型中的若干个参数并迅速搜索到全局最优解的特点, 收敛速度快, 稳定性高, 且对初始值不敏感(邢会敏等, 2017), 是参数优选问题中非常有效且应用广泛的方法(Liu等, 2005)。

目前SCE-UA算法大多应用于水文模型参数优化中, 由于其综合了多种算法的优点, 全局搜索能力强、效率高, 且针对非线性问题优化效果良好的特点, 在遥感同化作物生长模型的参数优化(闫岩等, 2006; 黄健熙等, 2015; 刘峻明等, 2018)、陆面过程模式参数同化优化(张生雷等, 2015)等领域中也有了广泛的应用。针对本研究中模型没有具体函数表达式, 且具有参数多、非线性强、待优化参数有区间约束的特征, 本研究首次尝试引入SCE-UA全局优化算法作为冬小麦分布提取阈值模型参数自动优化的全局优化算法。本研究利用SCE-UA全局优化算法的目的是以行政单元冬小麦种植面积统计数据作为外部优化数据比较对象, 当提取县级行政单元冬小麦面积与县级行政单元冬小麦面积统计数据间差异最小时, 得到县域内冬小麦分类模型最优参数。最后, 在最优参数下输出冬小麦分类算法提取的冬小麦分布信息。SCE-UA算法进行目标函数最小化问题求解的具体步骤(Duan等, 1993; 董洁平等, 2012; 张景等, 2014)如下:

(1) 初始化。假定待优化问题是 n 维问题, 选取参与进化的复合型个数 $p(p \geq 1)$ 和每个复合型所包含的顶点数目 $m(m \geq n + 1)$, 计算样本点数目 $s = p \times m$ 。

(2) 产生样本点。在可行域内随机产生 s 个样本点 $x_1, \dots, x_s$, 分别计算每一点 x_i 的函数值 $f_i = f(x_i)$, $i = 1, \dots, s$ 。

(3) 样本点排序。把 s 个样本点 (x_i, f_i) 按函数值升序排列, 排序后仍记为 $(x_i, f_i), i = 1, \dots, s$, 其中 $f_1 \leq f_2 \leq \dots \leq f_s$, 记 $D = \{(x_i, f_i), i = 1, \dots, s\}$ 。

(4) 划分复合型群体。将 D 划分为 p 个复合型 $A_1, \dots, A_p$, 每个复合型含有 m 个点, 其中 $A^k = \{(x_j^k, f_j^k) | x_j^k = x_{k+m(k-1)}, f_j^k = f_{j+m(k-1)}, j = 1, \dots, m\}, k = 1, 2, \dots, p$ 。

(5) 复合型进化。按CCE(复合型进化算法)分别进化各个复合型。

(6) 复合型混合。把进化后的每个复合型的的所有顶点组合成新的点集, 再次按函数值 f_i 升序排列, 排序后不妨仍记为 D , 对 D 按目标函数的升序进行排列。

(7) 收敛性判断。如果满足收敛条件则停止, 否则返回步骤(4)。

3.4.2 阈值法冬小麦分布提取模型参数优化主要过程

模型参数的优化过程就是寻找一组阈值参数, 利用该组阈值参数使利用模型提取的值与实测值之间按照目标函数设定的规则达到最吻合的状态。本文中待优化阈值参数的过程如下:

(1) 冬小麦分布提取模型阈值参数上下界确定。统计采集的冬小麦地面训练样本NDVI值, 计算冬小麦空间分布提取模型中各个判别函数对应的阈值 $(T_1, T_2, \dots, T_6)$ 分布范围(图7)。由于传统阈值法中各关键物候期对应阈值一般采用统计训练样本后得出的最小值或者在最小值附近多次尝试, 寻找精度最高的一组, 因此本研究阈值参数上下界选择设置在最小值附近。又由于SCE-UA优化算法的最大特点是对待优化参数初始值不敏感, 可避免优化过程对先验知识的过分依赖, 因此, 本研究只需保证最优阈值在确定阈值大致范围内即可, 所以最终本研究将统计所得最小值设为下界, 其下四分位数设为上界。将各关键物候期对应的下四分位数NDVI值到NDVI最小值分别作为参数优化的上下界。

(2) 全局优化算法SCE-UA主要参数确定。在SCE-UA算法的具体使用中根据待优化模型的复杂程度和参数个数设置相关参数, 其中大部分参数取值主要采用已有研究的默认值(宋星原等, 2009), 但复合型的个数 p 是唯一需要单独考虑的参数。本研究中, 据已发表文献(张景等, 2014), 取 $p=7$ 。根据算法创始人的推荐(Duan等, 1993), 其

余各参数缺省取值为 $\beta = 2n + 1$, $m = 2n + 1$, $\alpha = 1$, $q = n + 1$ 。本研究利用的SCE-UA算法中各参数具体含义及其取值结果见表3。

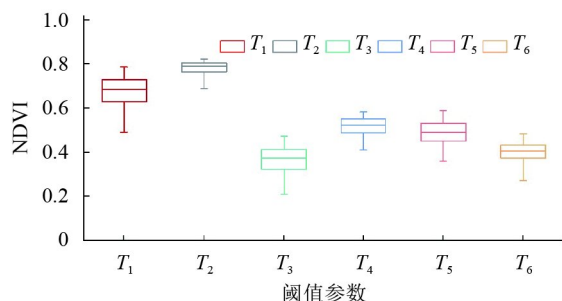


图7 冬小麦训练样本关键物候期阈值范围分布

Fig. 7 Range of threshold value for training samples in critical growth stage of winter wheat

表3 复合型混合演化算法中有关参数的取值
Table 3 Parameters' values of SCE-UA method in this study

待优化参数个数 n	复合型个数 p	复合型中顶点数 m	子复合型顶点数 q	复合型演化次数 β	子复合型演化次数 α
6	7	13	7	13	1

(3) 目标函数的建立。本文为了评价利用模型提取值与实测值之间的吻合程度, 将区域冬小麦种植面积统计数据作为总量控制参考标准, 与根据建立的冬小麦空间分布提取模型, 在SCE-UA全局优化算法支持下, 提取的冬小麦种植面积的差值的绝对值建立优化目标函数:

$$y = \sum_{i=1}^n |X_{\text{simi}} - X_{\text{obsi}}| \quad (6)$$

式中, 目标函数值 y 是利用遥感提取的区域冬小麦空间分布总量与区域冬小麦面积总量控制参考标准间差异绝对值的和, X_{simi} 是区域内基于遥感利用模型提取的冬小麦空间分布进行像素统计后得到的面积总量, X_{obsi} 是区域内冬小麦种植面积的总量控制参考标准(本研究将区域冬小麦种植面积统计数据作为总量控制参考标准), n 为研究区域个数(本研究中, $n=1$)。

(4) 迭代停止规则确定。当优化过程出现以下3种情况之一时停止优化过程: 1)目标函数在临近5个最优目标函数值之差的绝对值小于0.001, 即无法显著改善目标函数值; 2)目标函数的最大迭代次数不超过10000次; 3)待优化参数连续迭代5次后无显著改变。此时认为目标函数达到最优, 迭代停止, 输出最优阈值。

3.5 精度验证与对比

(1) 遥感分类精度验证指标。本研究中, 主要采用野外采集样本点对冬小麦空间分布提取结果进行精度验证, 该方法是遥感分类和冬小麦面积提取精度评价中常用的检验方式(Conese和Maselli, 1992; Janssen和van der Wel, 1994; Knick等, 1997)。研究中, 利用地面点样本数据, 主要采用总量精度、总体精度、Kappa系数、制图精度和用户精度等5个指标对本研究中获取的冬小麦空间分布结果进行精度客观评价。其中, 总量精度计算公式为:

$$K = 1 - \left| \frac{X_{\text{simi}} - X_{\text{obj}}}{X_{\text{obj}}} \right| \times 100\% \quad (7)$$

式中, K 为总量精度(%), X_{simi} 为基于遥感提取冬小麦空间分布进行像素统计获得的面积总量(hm^2), X_{obj} 为研究区内冬小麦种植面积的统计数据(hm^2)。其他指标具体定义及计算方式可参考相关参考文献(Congalton, 1991), 此处不再赘述。

(2) 其他分类方法对比。为进一步验证本研究所提阈值模型参数优化提取冬小麦空间分布的方法效果, 除进行所提方法自身精度验证外, 本研究利用GF-1/WFV遥感数据、地面训练样本数据和地面验证数据, 同时开展与其他方法冬小麦空间分布提取精度对比研究。其他方法主要包含传统阈值方法, 还包括较为常用且精度较高的冬小麦分布提取方法, 如最大似然法(Solberg等, 1994; 杨鹏等, 2000; 黄健熙等, 2017)、支持向量机等方法(Mazzoni等, 2007; Mathur和Foody, 2008; 贾坤等, 2011)。有关两种方法的介绍在此不多作介绍, 详细可见相关参考文献。

4 结果与分析

4.1 冬小麦空间分布提取阈值模型参数优化结果

本研究在冬小麦种植面积统计数据总量参数控制下, 利用章节3.4所述阈值法参数优化过程, 在全局优化算法SCE-UA支持下, 对阈值法冬小麦空间分布提取模型的参数进行优化, 最终得到实验区景县2015年冬小麦空间分布提取模型最优阈值参数 T_i , 具体结果如表4所示。其中, 当像元满足第4时相NDVI值大于0.46且第11时相的NDVI值大于0.73时, 判断该像元为冬小麦; 当像元满足第

4时相比第2时相NDVI值大0.30以上并且第11时相比第9时相NDVI值大0.44以上时,判断该像元为冬小麦;当像元满足第4时相比第6时相NDVI值大0.39以上并且第14时相比第15时相NDVI值大0.37以上时,判断该像元为冬小麦;剩余其他像元归为非冬小麦。利用上述阈值法冬小麦空间分布提取模型最优参数,可实现实验区景县冬小麦空间分布高精度提取。通过对本研究利用上述冬小麦空间分布提取模型优化参数获取的2015年冬小麦空间分布结果分析可知,提取面积与县域冬小麦面积总量控制参考标准间差异 y 仅为 160 m^2 。可见,通过全局优化算法和总量控制获得的阈值法冬小麦空间分布提取模型阈值参数确定的准确性和自动化程度均得到较大程度提高。

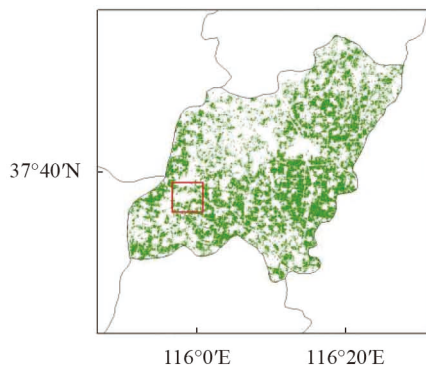
表4 冬小麦识别模型阈值参数优化结果(2015)

Table 4 Parameter optimization results of threshold method for winter wheat identification (2015)

阈值参数	T_1	T_2	T_3	T_4	T_5	T_6
下界	0.46	0.69	0.20	0.39	0.34	0.26
上界	0.55	0.78	0.31	0.51	0.43	0.39
优化阈值	0.46	0.73	0.30	0.44	0.39	0.37

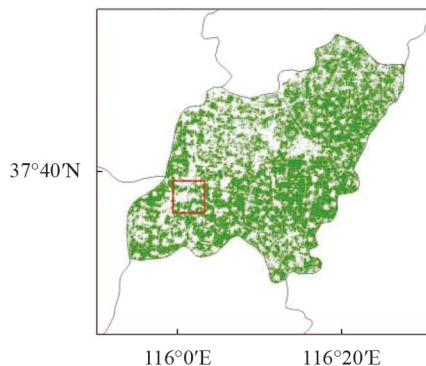
4.2 阈值模型参数优化法冬小麦提取结果精度验证

根据4.1节中冬小麦空间分布提取阈值法最优模型对实验区景县冬小麦空间分布进行提取,具体空间分布结果如图8(a)所示。从图8(a)中的冬小麦空间分布上看,2015年景县冬小麦主要集中分布于东部和东南部,西北部冬小麦分布较为破碎。通过统计可知,利用全局优化算法和总量控制下阈值法提取景县全县冬小麦种植面积为 40917.98 hm^2 ,县级冬小麦种植面积统计数据为 40918.00 hm^2 ,两者间相对误差仅为0.01%,即总量精度达到了99.99%,这说明采用作物面积统计数据总量控制和全局优化算法支持下的冬小麦空间分布提取模型实际提取冬小麦总量精度达到很高水平。最后,本研究将分类后的结果与采集地面验证样本计算混淆矩阵,对方法进行精度验证,具体如表5所示。验证结果表明,本文方法在保证较高总量精度的同时,总体精度为97.03%,Kappa系数为0.94,均达到了较高水平。其中,冬小麦和其他类型的制图精度分别为97.48%和96.00%,用户精度分别为98.25%和94.30%,实现了高精度冬小麦空间分布提取的目标。以上结果证明了阈值参数优化法中利用作物面积统计数据总量控制和全局优化算法的有效性。



(a) 阈值模型参数优化法分类结果

(a) Classification result of the threshold model parameter optimization method



(b) 传统阈值法分类结果

(b) Classification result of the traditional threshold method



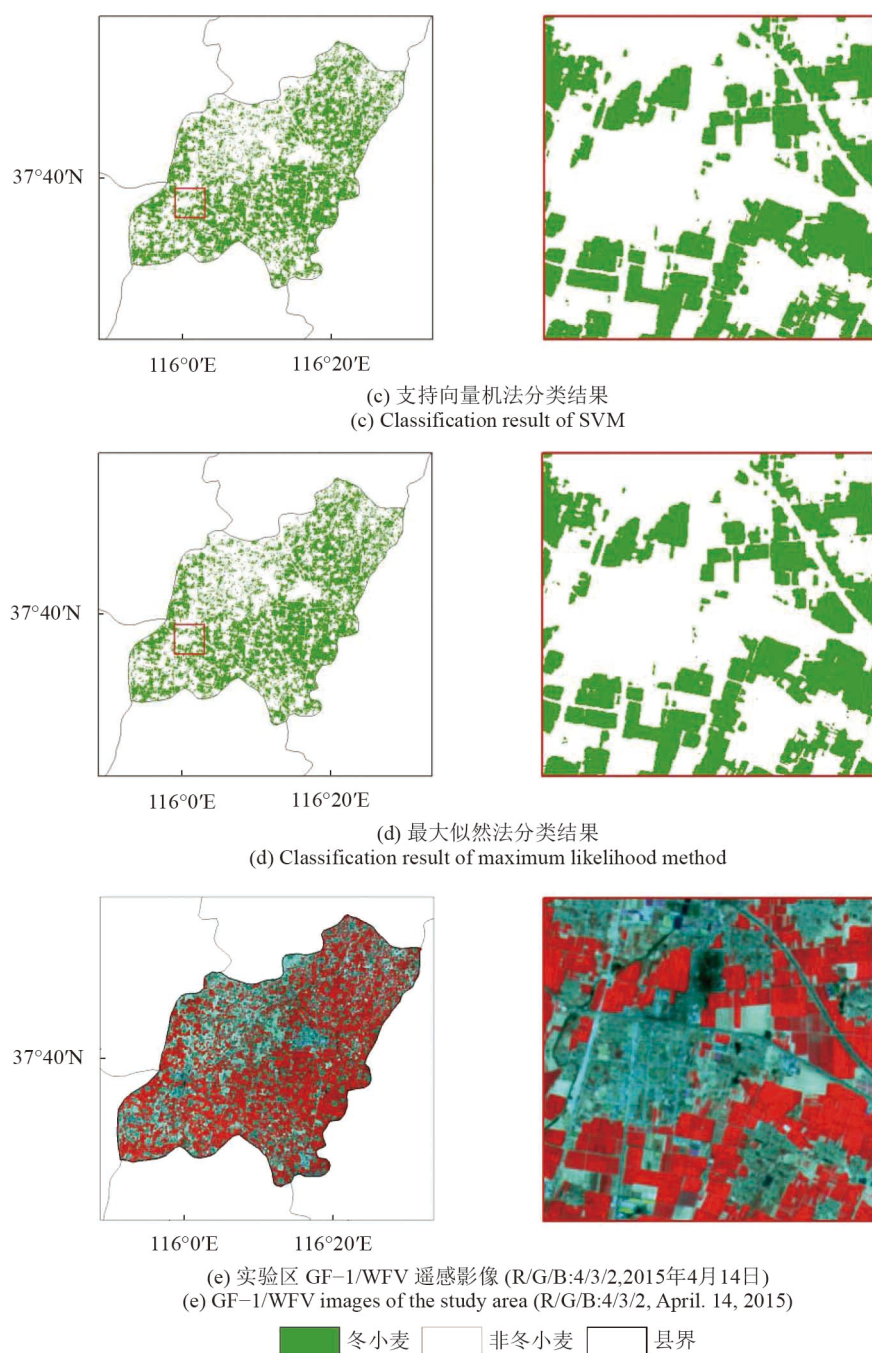


图8 不同冬小麦识别方法结果

Fig. 8 Classification results of different winter wheat identification methods

4.3 阈值模型参数优化法与其他分类方法对比

为进一步验证本研究所提阈值模型参数优化提取冬小麦空间分布的方法效果,并从多个方面进一步证明本研究方法的有效性和可靠性,本研究开展所提阈值模型参数优化方法与传统阈值方法、支持向量机和最大似然法等方法冬小麦提取精度对比研究,各种主要方法获得的冬小麦空间分布结果如图8所示,具体各分类方法分类结果精

度评价如表5所示。其中,在传统阈值法中,本研究阈值确定主要通过对采集的冬小麦所有训练样本NDVI值进行统计,确定冬小麦各关键物候期对应的最小NDVI值作为该时相阈值。支持向量机法以及最大似然法均利用ENVI5.1软件中监督分类模块对冬小麦抽穗期GF-1/WFV单幅影像(2015-04-22)进行分类。其中,最大似然法分类采用ENVI5.1软件系统默认值(王增林和朱大明, 2010);支持向

量机核函数选择最为常用的高斯径向基核函数RBF(Radial Based Function), 内核函数(Gamma in Kernel Function)值为输入影像波段数的倒数, 本研究输入影像为高分一号, 其波段数为4, 因此, 内核函数取值为0.25; 惩罚参数(Penalty Parameter)反映对误差的宽容度, 这个值设置越高, 说明越不能容忍出现误差, 但是将造成负面影响产生严重的“椒盐”现象, 根据经验设置为100(骆剑承等, 2002)。

表 5 不同冬小麦识别方法混淆矩阵

Table 5 Confusion matrix of different winter wheat identification methods

分类方法	地物类型	总体精度/%	Kappa	制图精度/%	用户精度/%
阈值模型参数优化方法	冬小麦	97.03	0.94	97.48	98.25
	非冬小麦			96.00	94.30
传统阈值法	冬小麦	92.48	0.82	95.30	94.00
	非冬小麦			86.00	88.84
支持向量机	冬小麦	96.88	0.93	95.41	100.00
	非冬小麦			100.00	91.12
最大似然法	冬小麦	94.60	0.88	94.00	98.18
	非冬小麦			96.00	87.43

从本研究阈值模型参数优化法与传统阈值法分类结果对比图(图8)中可以看出, 传统阈值法虽操作简单, 运算速度快, 但由于阈值设定仅根据经验并采用简单统计方法确定, 无法准确去除其他背景地物的影响, 导致分类结果冬小麦种植面积总量远高于实际冬小麦种植面积总量, 且存在大量破碎的小斑块, 将非冬小麦错分为冬小麦的现象严重。支持向量机法和最大似然法尽管分类精度也达到一定水平, 但上述两种方法主要基于单一时相影像进行分类, 辨识光谱相近的两种或多种冬小麦时易受“同谱异物”等原因的影响, 一定程度阻碍分类结果精度的进一步提高。此外, 与本研究阈值模型参数优化法结果相比, 支持向量机法和最大似然法获取的冬小麦种植面积总量明显高于实际种植面积总量的问题, 这也体现了本研究提阈值模型参数优化法采用作物面积统计数据总量控制进行优化参数的优势。

由各分类方法结果精度评价结果看(表5), 从总体精度、Kappa、制图精度和用户精度等指标衡量, 传统阈值法精度均低于本研究提阈值模型

参数优化法、支持向量机和最大似然分类方法, 而支持向量机和最大似然法除了总体精度和Kappa系数较高外, 用户精度也达到了较高的精度水平, 证明支持向量机和最大似然法也是分类效果较优的作物分类算法, 这与很多学者的结论是一致的(孙坤和鲁铁定, 2017; 王小明等, 2013)。另外, 从表5也可看出, 本研究方法分类结果比传统阈值法、最大似然法和支持向量机法的总体精度和Kappa系数均有所提高, 且制图精度也达到最高水平。其中, 总体精度分别提高4.55%、2.43%和0.15%, Kappa系数分别提高0.12、0.06和0.01, 制图精度分别提高2.18%、3.48%和2.07%, 尽管本研究所提阈值参数优化法的用户精度比支持向量机略低, 也达到98%以上, 可以满足用户精度需求。相较于传统阈值法而言, 本研究所提阈值参数优化法的用户精度提高了4.25%, 也充分表明了本研究对传统阈值法改进的有效性。尽管支持向量机法在本实验中用户精度高于本研究方法, 但是由于该算法对训练样本依赖性强, 训练样本数目越大, 算法计算复杂程度越高。因此对于大规模训练样本, 利用该算法分类运算量大, 分类耗时长, 因此导致该方法在大尺度区域应用时, 前期工作量大, 计算效率低, 而且若在没有充足训练样本的情况下, 分类精度容易偏低, 使得该方法在大范围、长时间农作物空间分布制图中应用难度增加。最大似然法分类对于地面训练样本同样具有很强的依赖性, 该算法要求采集的训练样本光谱特征值服从正态分布, 但由于实际应用中遥感信息的统计分布很难满足正态分布特点, 因此当区域内选择的样本分布较为离散从而未能满足最大似然法的假设分布条件, 或者样本的选取不具有代表性, 就会导致分类结果精度下降, 这在一定程度上也增加了区域应用中取得较高分类精度的难度。

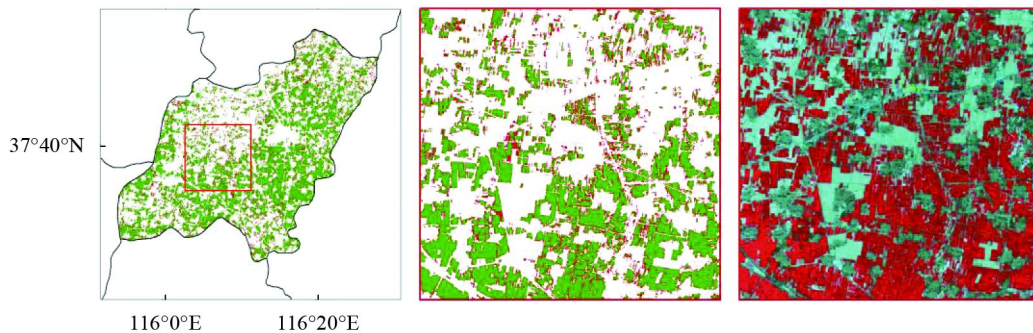
综合看, 所提阈值参数优化法不仅比传统阈值法分类精度有较大提高, 而且与支持向量机和最大似然法等相比, 所提分类方法在保证了冬小麦种植面积总量精度的基础上, 同时分类空间精度也达到较高水平, 可以满足大范围、长时间农业科学基础数据中农作物空间分布制图的精度需求; 并且本方法继承了传统阈值法提取冬小麦空间分布操作简单、计算速度快、能较好地适应当地条件的优点, 对于开展大尺度区域作物空间分

布制图有良好的适用性。其中, 尽管面积提取方法的模型和输入数据时相略有不同, 但是, 通过对比足以证明本研究所提作物面积提取方法常用方法间的精度都达到了较高水平, 且本研究的大部分精度指标均高于上述对比方法, 从而进一步证明本研究作物面积提取方法的有效性和可靠性。

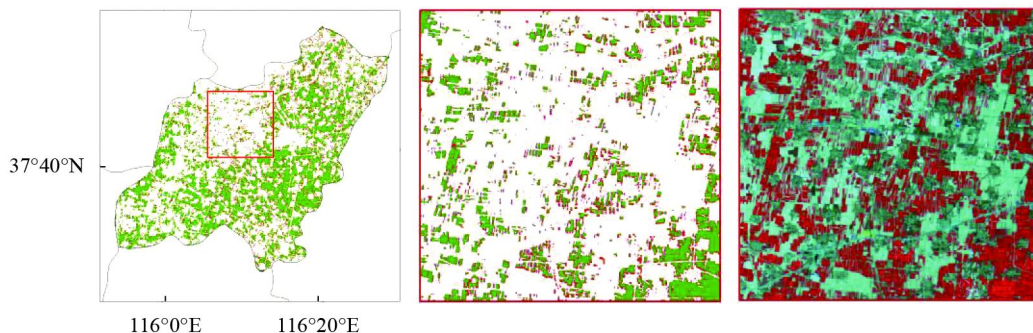
此外, 为进一步对本研究所提阈值模型参数优化方法和传统阈值法、最大似然法、支持向量机法的作物识别结果进行更加细致的空间差异分析, 本研究将所提阈值模型参数优化方法与其他方法的作物空间分布结果之间进行空间叠加对比(图9), 并对结果进行分析。从冬小麦分类结果差异的空间分布看, 本研究所提阈值模型参数优化方法与其他方法间分类结果差异较大的区域主要分布在景县研究区西北部, 如图9中的局部放大区域所示, 主要原因是由于该县西北部地区冬小麦种植集中连片区域相对较少, 地块较为破碎, 导致纯净、混合像元共存现象增加(朱爽和张锦水, 2016), 从而导致西北部地区各分类方法冬小麦识别结果差异较大。由冬小麦分类结果间差异统计结果可知, 传统阈值法分类结果面积总量远高于本研究方法, 分类结果与本研究方法空间分布不一致的比例达到17.59%, 其原因是传统阈值法将

非冬小麦错分为冬小麦的现象较为严重, 导致分类结果中冬小麦面积总量远高于真实总量, 而本研究方法识别结果与常用较优的支持向量机法和最大似然法作物识别结果空间一致比例均达到90%以上, 这说明本研究所提方法不仅能够改进传统阈值法阈值设置主观性强、智能性低等弱点, 而且分类结果精度得到较大水平提高。

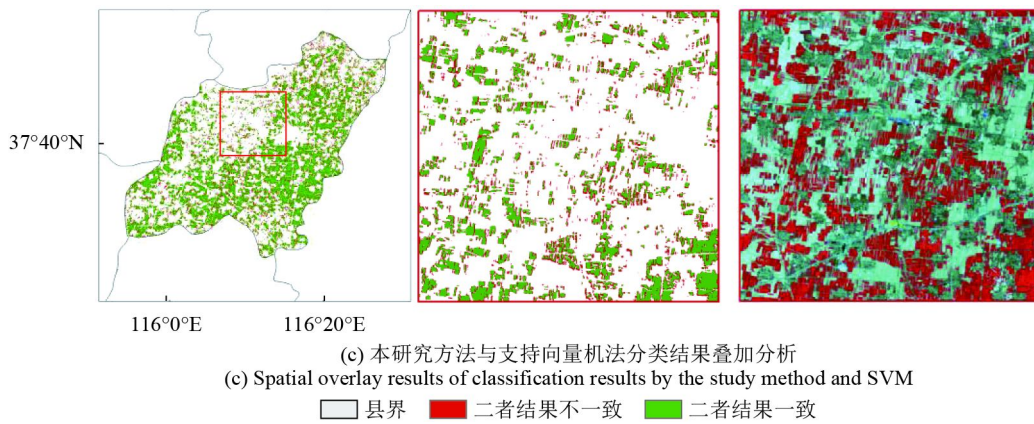
通过以上分析可知, 与传统阈值法相比, 本研究方法利用SCE-UA全局优化算法对阈值参数进行优化, 一定程度上克服了传统阈值法阈值确定主观性和不确定性都较强的缺点, 并且一定程度上避免了以往阈值模型中人工试错确定阈值需耗费大量工作量的缺点, 从而实现了阈值法阈值自动优化的冬小麦空间分布提取, 提高了阈值法分类中阈值选取的智能性与准确性, 较大提高了作物空间分布制图的精度和工作效率。同时, 本研究方法与常用的支持向量机法和最大似然法等分类方法相比, 不仅继承了传统阈值法提取冬小麦空间分布操作简单、计算速度快、能较好地适应当地条件的优点, 作物空间制图精度也达到较高水平, 实现了高精度冬小麦空间分布提取的目标, 为开展大范围、长时间农业科学基础数据中农作物空间分布制图提供了方法借鉴。



(a) 本研究方法与传统阈值法分类结果叠加
(a) Spatial overlay results of classification results by the study method and the traditional threshold method



(b) 本研究方法与最大似然法分类结果叠加分析
(b) Spatial overlay results of classification results by the study method and maximum likelihood method



(左图: 叠加结果 中图: 局部放大图 右图: 2015年4月14日GF-1/WFV遥感影像(R/G/B:4/3/2))

(Left: Spatial overlay results Middle: Partial enlargement Right: GF-1/WFV images on April 14, 2015(R/G/B:4/3/2))

图9 本研究方法与其他分类结果空间叠加对比

Fig. 9 Spatial overlay results of classification results by the study method with other methods

5 讨论

(1) 研究方法应用前景。与一般农作物面积遥感早期识别和动态监测不同, 本文的主要目的为获取大范围、长时间和高精度农作物空间分布基础农业科学数据进行方法探索。由于农作物空间分布基础农业科学数据要求的时效性低且一般针对历史时段进行作物空间分布制图, 所以尽管统计数据具有一定的滞后性, 本方法在农作物空间分布基础农业科学数据制图中仍具有广泛的应用前景。与一般基于遥感信息和非遥感信息进行多源数据融合农作物空间分布制图方法相比, 本方法在保证数量精度较高的情况下, 也能获得较高的空间位置精度。因此, 本研究方法也可作为农作物面积农业统计数据降尺度恢复重建和大范围区域冬小麦空间分布制图研究提供一定技术参考。

(2) 研究方法的不足与改进。目前基于全局优化算法支持下的阈值模型参数优化法提取冬小麦分布信息取得了较好结果, 但是研究中仍然存在一些问题或不足。第一, 由于受云雨天气的影响, 研究中较难获得等时间间隔且云量低的影像, 因此, 目前研究仅采用了非等时间间隔影像; 第二, 本研究目前只在县域范围进行研究, 且研究对象为种植结构简单的冬小麦, 利用的数据为较高空间分辨率的光学遥感数据, 且仅采用了单一遥感特征参数NDVI进行阈值法冬小麦面积提取; 第三, 本研究采用统计数据作为全局优化的总量控制标准参考, 由于统计数据具有一定的

滞后性, 导致本文目前仅用过去年份的数据开展农作物空间分布制图方法应用验证研究, 而对本方法适用的时间范围扩展性有待进一步开展研究; 此外, 本文目前只侧重于全局优化算法支持阈值检测的冬小麦空间分布制图方法研究, 而对方法精度的可能影响因素与控制措施研究存在一定不足, 如不同物候特征时间组合、统计数据误差等因素对作物空间分布制图精度的影响。同时, 本研究中SCE-UA全局优化算法的参数设置目前采用了该算法推荐的默认参数, 并未开展优化算法参数变化对本文研究结果的影响研究。

针对方法存在的可能问题或不足, 在基于总量控制和全局优化的前提下, 今后在以下几个方面进行深入研究: (1)利用NDVI最大值合成方式获取等时间间隔的遥感数据, 开展对冬小麦分布提取精度影响研究; (2)利用其他中低、高空间分辨率遥感数据在更大区域进行冬小麦分布提取的有效性、结果精度验证和优化参数结果应用及评价等研究。另外, 在现有计算框架下进一步讨论应用其他遥感特征参量(如增强型植被指数EVI、宽范围动态植被指数WDRVI等)进行冬小麦空间分布制图研究; (3)进一步探究利用本研究方法获得研究区历史长时序阈值模型参数规律, 通过对算法阈值标准化, 开展将作物识别模型历史阈值用于现在和未来作物空间分布监测及预测的方法研究, 明确阈值应用的时间扩展性和有效性, 从而使得本方法具有更广泛的适用性; (4)开展统计数据自身误差、物候特征组合以及其他总量控制(如

基于抽样技术确定区域总量)的全局优化标准参考等对本研究方法的精度影响研究,从而提高本研究方法的精度,降低可能影响因素对方法精度的不确定性影响,从而进一步增强本方法的适用性;(5)进一步开展SCE-UA算法不同参数设置对本研究方法的精度影响研究。

6 结 论

为获取大范围、长时间和高精度农作物空间分布基础农业科学数据,本研究针对阈值法农作物识别过程中阈值设置存在灵巧性差和自动化程度低等弱点,在中国粮食生产基地黄淮海地区典型实验区内,开展了利用全局优化算法改进基于阈值检测的冬小麦分布制图方法探索研究。主要结论如下:

(1)在改进传统阈值方法基础上,本文引入冬小麦种植面积统计数据作为全局优化的总量控制参考标准,提出了基于时序NDVI遥感数据的阈值模型参数优化冬小麦空间分布制图方法,成功提取了区域冬小麦作物空间分布信息。通过地面验证表明,本研究所提阈值模型参数优化方法冬小麦提取分类结果精度均达到较高水平,可以满足大范围、长时间农业科学基础数据中农作物空间分布制图的精度需求。与传统的阈值法、最大似然和支持向量机等分类方法相比,本研究所提阈值模型参数优化法区域冬小麦作物分类总体精度和Kappa系数分别都有所提高,其中,总体精度分别提高4.55%、2.43%和0.15%,Kappa系数分别提高0.12、0.06和0.01,这证明了本研究所提基于统计数据总量控制的全局优化算法阈值模型参数优化作物分布信息提取方法的有效性和可行性,可获得高精度冬小麦作物空间分布制图结果。

(2)SCE-UA全局优化算法可实现阈值法冬小麦分布提取模型中阈值自动寻优的目的。本文首次尝试将SCE-UA全局优化算法引入冬小麦分布提取阈值模型进行参数自动优化,将遥感提取的实验区冬小麦空间分布总量与实验区冬小麦面积总量控制参考标准间差异绝对值的和作为目标函数,对待优化的阈值参数进行自动优化。结果表明,利用SCE-UA全局优化算法对阈值模型阈值参数进行优化选取,在一定程度上避免了传统阈值模型阈值确定通过人工试错方法特有的主观性和

不确定性,不仅大大提高了工作效率,并且提高了冬小麦提取阈值模型参数确定的精确性和自动化程度,对提高中国冬小麦空间分布制图精度和自动化水平具有重要意义。

(3)从本文开展的作物面积统计数据总量约束下全局优化阈值的冬小麦分布制图方法和精度验证研究可知,本方法由于采用统计数据作为总量控制参考标准,在保证总量精度的基础上,位置精度也达到较高水平,可以获得总量精度和位置精度双优的作物空间分布制图结果。另外,本文所提阈值模型参数优化冬小麦分布制图方法继承了传统阈值法提取作物空间分布操作简单、计算速度快、区域适用性强等优点,不仅作物空间制图精度可达到较高水平,实现了高精度冬小麦空间分布制图的目标,而且在大尺度区域作物空间分布制图中具有良好的适用性,为开展大范围、长时间农业科学基础数据中农作物空间分布制图提供了方法借鉴。

参考文献(References)

- Barnwal P and Kotani K. 2013. Climatic impacts across agricultural crop yield distributions: an application of quantile regression on rice crops in Andhra Pradesh, India. *Ecological Economics*, 87: 95–109 [DOI: [10.1016/j.ecolecon.2012.11.024](https://doi.org/10.1016/j.ecolecon.2012.11.024)]
- Brown J F and Pervez M S. 2014. Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture. *Agricultural Systems*, 127: 28–40 [DOI: [10.1016/j.agsy.2014.01.004](https://doi.org/10.1016/j.agsy.2014.01.004)]
- Cai X L and Cui Y L. 2009. Crop planting structure extraction in irrigated areas from multi-sensor and multi-temporal remote sensing data. *Transactions of the CSAE*, 25(8): 124–130 (蔡学良, 崔远来. 2009. 基于异源多时相遥感数据提取灌区作物种植结构. *农业工程学报*, 25(8): 124–130) [DOI: [10.3969/j.issn.1002-6819.2009.08.023](https://doi.org/10.3969/j.issn.1002-6819.2009.08.023)]
- Conese C and Maselli F. 1992. Use of error matrices to improve area estimates with maximum likelihood classification procedures. *Remote Sensing of Environment*, 40(2): 113–124 [DOI: [10.1016/0034-4257\(92\)90009-9](https://doi.org/10.1016/0034-4257(92)90009-9)]
- Congalton R G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1): 35–46 [DOI: [10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B)]
- Dong J P, Li Z J and Dai J N. 2012. Application of SCE-UA algorithm to optimization of Xin'anjiang model parameters. *Journal of Hohai University (Natural Sciences)*, 40(5): 485–490 (董洁平, 李致家, 戴健男. 2012. 基于SCE-UA算法的新安江模型参数优化及应用. *河海大学学报(自然科学版)*, 40(5): 485–490) [DOI: [10.11911/hh.2012.05009](https://doi.org/10.11911/hh.2012.05009)]

10.3876/j.issn.1000-1980.2012.05.002]

- Duan H B, Wang D B and Yu X F. 2007. Ant colony algorithm: survey and prospect. *Engineering Science*, 9(2): 98–102 (段海滨, 王道波, 于秀芬. 2007. 蚁群算法的研究现状及其展望. *中国工程科学*, 9(2): 98–102) [DOI: 10.3969/j.issn.1009-1742.2007.02.017]
- Duan Q Y, Gupta V K and Sorooshian S. 1993. Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications*, 76(3): 501–521 [DOI: 10.1007/BF00939380]
- Duan Q Y, Sorooshian S and Gupta V K. 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, 158(3/4): 265–284 [DOI: 10.1016/0022-1694(94)90057-4]
- Foerster S, Kaden K, Foerster M and Itzerott S. 2012. Crop type mapping using spectral-temporal profiles and phenological information. *Computers and Electronics in Agriculture*, 89: 30–40 [DOI: 10.1016/j.compag.2012.07.015]
- Ge J K, Qiu Y H, Wu C M and Pu G L. 2008. Summary of genetic algorithms research. *Application Research of Computers*, 25(10): 2911–2916 (葛继科, 邱玉辉, 吴春明, 蒲国林. 2008. 遗传算法研究综述. *计算机应用研究*, 25(10): 2911–2916) [DOI: 10.3969/j.issn.1001-3695.2008.10.008]
- Goodwin N R and Collett L J. 2014. Development of an automated method for mapping fire history captured in Landsat TM and ETM+time series across Queensland, Australia. *Remote Sensing of Environment*, 148: 206–221 [DOI: 10.1016/j.rse.2014.03.021]
- Hao P Y, Niu Z, Wang L, Wang X L and Wang C Y. 2012. Multi-source automatic crop pattern mapping based on historical vegetation index profiles. *Transactions of the Chinese Society of Agricultural Engineering*, 28(23): 123–131 (郝鹏宇, 牛铮, 王力, 王秀兰, 王长耀. 2012. 基于历史时序植被指数库的多源数据作物面积自动提取方法. *农业工程学报*, 28(23): 123–131) [DOI: 10.3969/j.issn.1002-6819.2012.23.017]
- Hao Z, Zhao H L and Jiang Y Z. 2017. Extraction of winter wheat area information based on the improved NDVI density slicing method. *South-to-North Water Transfers and Water Science and Technology*, 15(3): 67–72, 93 (郝震, 赵红莉, 蒋云钟. 2017. 基于改进的NDVI密度分割方法的冬小麦面积信息提取. *南水北调与水利科技*, 15(3): 67–72, 93) [DOI: 10.13476/j.cnki.nsbdkq.2017.03.012]
- Hengshui Municipal Bureau of Statistics, Hengshui Survey Team of the National Bureau of Statistics. 2016. *Hengshui statistical yearbook 2016*. Beijing: China Statistics Press (衡水市统计局, 国家统计局衡水调查队. 2016. *衡水统计年鉴2016*. 北京: 中国统计出版社)
- Hu Q, Wu W B, Song Q, Yu Q Y, Yang P and Tang H J. 2015. Recent progresses in research of crop patterns mapping by using remote sensing. *Scientia Agricultura Sinica*, 48(10): 1900–1914 (胡琼, 吴文斌, 宋茜, 余强毅, 杨鹏, 唐华俊. 2015. 农作物种植结构遥感提取研究进展. *中国农业科学*, 48(10): 1900–1914) [DOI: 10.3864/j.issn.0578-1752.2015.10.004]
- Huang J X, Hou Y Z, Wu H F, Liu J M and Zhu D H. 2017. Crop type mapping method based on time-series MODIS data in Heilongjiang province. *Transactions of the Chinese Society for Agricultural Machinery*, 48(10): 142–147, 285 (黄健熙, 侯裔焯, 武洪峰, 刘峻明, 朱德海. 2017. 基于时间序列MODIS的农作物类型空间制图方法. *农业机械学报*, 48(10): 142–147, 285) [DOI: 10.6041/j.issn.1000-1298.2017.10.017]
- Huang J X, Ma H Y, Tian L Y, Wang P X and Liu J M. 2015. Comparison of remote sensing yield estimation methods for winter wheat based on assimilating time-sequence LAI and ET. *Transactions of the Chinese Society of Agricultural Engineering*, 31(4): 197–203 (黄健熙, 马鸿元, 田丽燕, 王鹏新, 刘峻明. 2015. 基于时间序列LAI和ET同化的冬小麦遥感估产方法比较. *农业工程学报*, 31(4): 197–203) [DOI: 10.3969/j.issn.1002-6819.2015.04.028]
- Huang Q, Tang H J, Wu W B, Li D D and Liu J. 2013. Remote sensing based dynamic changes analysis of crop distribution pattern—taking northeast China as an example. *Scientia Agricultura Sinica*, 46(13): 2668–2676 (黄青, 唐华俊, 吴文斌, 李丹丹, 刘佳. 2013. 农作物分布格局动态变化的遥感监测——以东北三省为例. *中国农业科学*, 46(13): 2668–2676) [DOI: 10.3864/j.issn.0578-1752.2013.13.005]
- Janssen L L F and van der Wel F J M. 1994. Accuracy assessment of satellite derived land-cover data: a review. *Photogrammetric Engineering and Remote Sensing*, 60(4): 419–426
- Jia K, Li Q Z, Tian Y C and Wu B F. 2011. A review of classification methods of remote sensing imagery. *Spectroscopy and Spectral Analysis*, 31(10): 2618–2623 (贾坤, 李强子, 田亦陈, 吴炳方. 2011. 遥感影像分类方法研究进展. *光谱学与光谱分析*, 31(10): 2618–2623) [DOI: 10.3964/j.issn.1000-0593(2011)10-2618-06]
- Knick S T, Rotenberry J T and Zarriello T J. 1997. Supervised classification of Landsat Thematic Mapper imagery in a semi-arid rangeland by nonparametric discriminant analysis. *Photogrammetric Engineering and Remote Sensing*, 63(1): 79–86
- Lei G B, Li A N, Bian J H, Zhang Z J, Zhang W and Wu B F. 2014. A practical method for automatically identifying the evergreen and deciduous characteristic of forests at mountainous areas: a case study in Mt. Gongga region. *Acta Ecologica Sinica*, 34(24): 7210–7221 (雷光斌, 李爱农, 边金虎, 张正健, 张伟, 吴炳方. 2014. 基于阈值法的山区森林常绿、落叶特征遥感自动识别方法——以贡嘎山地区为例. *生态学报*, 34(24): 7210–7221) [DOI: 10.5846/stxb201310112440]
- Liang S Z, Chen J S, Wu B F and Chen G. 2015. Extracting rubber plantation with decision tree model based on object-oriented method. *Journal of Remote Sensing*, 19(3): 485–494 (梁守真, 陈劲松, 吴炳方, 陈工. 2015. 应用面向对象的决策树模型提取橡胶林信息. *遥感学报*, 19(3): 485–494) [DOI: 10.11834/jrs.20153337]
- Lin H J and Peng H. 2006. Application of simulated annealing in global query optimization. *Computer Technology and Development*, 16(4): 155–157 (林慧君, 彭宏. 2006. 模拟退火算法在全局查询优化中的应用. *计算机技术与发展*, 16(4): 155–157) [DOI: 10.3969/j.issn.1673-629X.2006.04.055]
- Liu J, Wang L M, Yang F G, Yang L B and Wang X L. 2015. Remote

- sensing estimation of crop planting area based on HJ time-series images. *Transactions of the Chinese Society of Agricultural Engineering*, 31(3): 199–206 (刘佳, 王利民, 杨福刚, 杨玲波, 王小龙. 2015. 基于HJ时间序列数据的农作物种植面积估算. *农业工程学报*, 31(3): 199–206) [DOI: [10.3969/j.issn.1002-6819.2015.03.026](https://doi.org/10.3969/j.issn.1002-6819.2015.03.026)]
- Liu J G, Fritz S, van Wesenbeeck C F A, Fuchs M, You L Z, Obersteiner M and Yang H. 2008. A spatially explicit assessment of current and future hotspots of hunger in Sub-Saharan Africa in the context of global change. *Global and Planetary Change*, 64(3/4): 222–235 [DOI: [10.1016/j.gloplacha.2008.09.007](https://doi.org/10.1016/j.gloplacha.2008.09.007)]
- Liu J K, Zhong S Q and Liang W H. 2015. Extraction on crops planting structure based on multi-temporal Landsat8 OLI images. *Remote Sensing Technology and Application*, 30(4): 775–783 (刘吉凯, 钟仕全, 梁文海. 2015. 基于多时相Landsat8 OLI影像的作物种植结构提取. *遥感技术与应用*, 30(4): 775–783) [DOI: [10.11873/j.issn.1004-0323.2015.4.0775](https://doi.org/10.11873/j.issn.1004-0323.2015.4.0775)]
- Liu J M, Pan P Z, Wang P X, Cui Z Z and Hu X. 2018. Parameters optimization of wheat spike differentiation stages model based on SCE-UA algorithm. *Transactions of the Chinese Society for Agricultural Machinery*, 49(4): 232–240 (刘峻明, 潘佩珠, 王鹏新, 崔珍珍, 胡新. 2018. 基于SCE-UA算法的小麦穗分化期模拟模型参数优化. *农业机械学报*, 49(4): 232–240) [DOI: [10.6041/j.issn.1000-1298.2018.04.026](https://doi.org/10.6041/j.issn.1000-1298.2018.04.026)]
- Liu Y Q, Gupta H V, Sorooshian S, Bastidas L A and Shuttleworth W J. 2005. Constraining land surface and atmospheric parameters of a locally coupled model using observational data. *Journal of Hydrometeorology*, 6(2): 156–172 [DOI: [10.1175/JHM407.1](https://doi.org/10.1175/JHM407.1)]
- Liu Z H, Li Z G, Tang P Q, Li Z P, Wu W B, Yang P, You L Z and Tang H J. 2013. Spatial-temporal changes of rice area and production in China during 1980–2010. *Acta Geographica Sinica*, 68(5): 680–693 (刘珍环, 李正国, 唐鹏钦, 李志鹏, 吴文斌, 杨鹏, 游良志, 唐华俊. 2013. 近30年中国水稻种植区域与产量时空变化分析. *地理学报*, 68(5): 680–693) [DOI: [10.11821/xb201305009](https://doi.org/10.11821/xb201305009)]
- Lu Y T, Lin Y Y, Peng Q Z and Wang Y Z. 2015. A review of improvement and research on parameters of simulated annealing algorithm. *College Mathematics*, 31(6): 96–103 (卢宇婷, 林禹枚, 彭乔姿, 王颖喆. 2015. 模拟退火算法改进综述及参数探究. *大学数学*, 31(6): 96–103) [DOI: [10.3969/j.issn.1672-1454.2015.06.020](https://doi.org/10.3969/j.issn.1672-1454.2015.06.020)]
- Luo J C, Zhou C H, Liang Y and Ma J H. 2002. Support vector machine for spatial feature extraction and classification of remotely sensed imagery. *Journal of Remote Sensing*, 6(1): 50–55 (骆剑承, 周成虎, 梁怡, 马江洪. 2002. 支撑向量机及其遥感影像空间特征提取和分类的应用研究. *遥感学报*, 6(1): 50–55) [DOI: [10.11834/jrs.20020110](https://doi.org/10.11834/jrs.20020110)]
- Massey R, Sankey T T, Congalton R G, Yadav K, Thenkabail P S, Ozdogan M and Meador A J S. 2017. MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types. *Remote Sensing of Environment*, 198: 490–503 [DOI: [10.1016/j.rse.2017.06.033](https://doi.org/10.1016/j.rse.2017.06.033)]
- Mathur A and Foody G M. 2008. Crop classification by support vector machine with intelligently selected training data for an operational application. *International Journal of Remote Sensing*, 29(8): 2227–2240 [DOI: [10.1080/01431160701395203](https://doi.org/10.1080/01431160701395203)]
- Mazzoni D, Garay M J, Davies R and Nelson D. 2007. An operational MISR pixel classifier using support vector machines. *Remote Sensing of Environment*, 107(1/2): 149–158 [DOI: [10.1016/j.rse.2006.06.021](https://doi.org/10.1016/j.rse.2006.06.021)]
- Pervez M S and Brown J F. 2010. Mapping irrigated lands at 250-m scale by merging MODIS data and national agricultural statistics. *Remote Sensing*, 2(10): 2388–2412 [DOI: [10.3390/rs2102388](https://doi.org/10.3390/rs2102388)]
- Qiu J J, Tang H J, Frolking S, Boles S, Li C S, Xiao X M, Liu J Y, Zhuang Y H and Qin X Q. 2003. Mapping single-, double-, and triple-crop agriculture in china at 0.5°×0.5° by combining county-scale census data with a remote sensing-derived land cover map. *Geocarto International*, 18(2): 3–13 [DOI: [10.1080/10106040308542268](https://doi.org/10.1080/10106040308542268)]
- Saeed U, Dempewolf J, Becker-Reshef I, Khan A, Ahmad A and Wajid S A. 2017. Forecasting wheat yield from weather data and MODIS NDVI using Random Forests for Punjab province, Pakistan. *International Journal of Remote Sensing*, 38(17): 4831–4854 [DOI: [10.1080/01431161.2017.1323282](https://doi.org/10.1080/01431161.2017.1323282)]
- Salahddine D, Housni F E, Najine A, Wafik A, Aadraoui M, Hafiane F Z and Toro H B. 2017. Mapping and characterization of agricultural systems from time series of Normalized Difference Vegetation Index (NDVI) in the northeast area of Tadla, Morocco. *Natural Resources*, 8(1): 73805 [DOI: [10.4236/nr.2017.81002](https://doi.org/10.4236/nr.2017.81002)]
- Solberg A H S, Jain A K and Taxt T. 1994. Multisource classification of remotely sensed data: fusion of Landsat TM and SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 32(4): 768–778 [DOI: [10.1109/36.298006](https://doi.org/10.1109/36.298006)]
- Song Q, Zhou Q B, Wu W B, Hu Q, Yu Q Y and Tang H J. 2015. Recent progresses in research of integrating multi-source remote sensing data for crop mapping. *Scientia Agricultura Sinica*, 48(6): 1122–1135 (宋茜, 周清波, 吴文斌, 胡琼, 余强毅, 唐华俊. 2015. 农作物遥感识别中的多源数据融合研究进展. *中国农业科学*, 48(6): 1122–1135) [DOI: [10.3864/j.issn.0578-1752.2015.06.09](https://doi.org/10.3864/j.issn.0578-1752.2015.06.09)]
- Song X Y, Shu Q Y, Wang H B and Liao W M. 2009. Comparison and application of SCE-UA, genetic algorithm and simplex method. *Engineering Journal of Wuhan University*, 42(1): 6–9, 15 (宋星原, 舒全英, 王海波, 廖为民. 2009. SCE-UA、遗传算法和单纯形优化算法的应用. *武汉大学学报(工学版)*, 42(1): 6–9, 15)
- Sun K and Lu T D. 2017. Comparison of supervised classification methods in remote sensing image classification. *Jiangxi Science*, 35(3): 367–371, 468 (孙坤, 鲁铁定. 2017. 监督分类方法在遥感影像分类处理中的比较. *江西科学*, 35(3): 367–371, 468) [DOI: [10.13990/j.issn1001-3679.2017.03.009](https://doi.org/10.13990/j.issn1001-3679.2017.03.009)]
- Tang H J, Wu W B, Yang P, Zhou Q B and Chen Z X. 2010. Recent progresses in monitoring crop spatial patterns by using remote sensing technologies. *Scientia Agricultura Sinica*, 43(14): 2879–2888 (唐华俊, 吴文斌, 杨鹏, 周清波, 陈仲新. 2010. 农作

- 物空间格局遥感监测研究进展. 中国农业科学, 43(14): 2879–2888 [DOI: [10.3864/j.issn.0578-1752.2010.14.006](https://doi.org/10.3864/j.issn.0578-1752.2010.14.006)]
- Wang L M, Liu J, Yang L B, Yang F G, Teng F and Wang X L. 2016. Remote sensing monitoring winter wheat area base on weighted NDVI index. Transactions of the Chinese Society of Agricultural Engineering, 32(17): 127–135 (王利民, 刘佳, 杨玲波, 杨福刚, 滕飞, 王小龙. 2016. 基于NDVI加权指数的冬小麦种植面积遥感监测. 农业工程学报, 32(17): 127–135) [DOI: [10.11975/j.issn.1002-6819.2016.17.018](https://doi.org/10.11975/j.issn.1002-6819.2016.17.018)]
- Wang N, Li Q Z, Du X, Zhang Y, Zhao L C and Wang H Y. 2017. Identification of main crops based on the univariate feature selection in Subei. Journal of Remote Sensing, 21(4): 519–530 (王娜, 李强子, 杜鑫, 张源, 赵龙才, 王红岩. 2017. 单变量特征选择的苏北地区主要农作物遥感识别. 遥感学报, 21(4): 519–530) [DOI: [10.11834/jrs.20176373](https://doi.org/10.11834/jrs.20176373)]
- Wang X M, Mao M Q, Zhang C J and Xu Y. 2013. Comparative study on classification of remote sensing image by support vector machine. Geomatics and Spatial Information Technology, 36(4): 17–20, 23 (王小明, 毛梦祺, 张昌景, 许勇. 2013. 基于支持向量机的遥感影像分类比较研究. 测绘与空间地理信息, 36(4): 17–20, 23) [DOI: [10.3969/j.issn.1672-5867.2013.04.005](https://doi.org/10.3969/j.issn.1672-5867.2013.04.005)]
- Wang X M and Wang Y H. 1997. The combination of simulated annealing and genetic algorithms. Chinese Journal of Computers, 20(4): 381–384 (王雪梅, 王义和. 1997. 模拟退火算法与遗传算法的结合. 计算机学报, 20(4): 381–384)
- Wang Z L and Zhu D M. 2010. A study of maximum likelihood classification algorithm based on remote sensing image. Henan Science, 28(11): 1458–1461 (王增林, 朱大明. 2010. 基于遥感影像的最大似然分类算法的探讨. 河南科学, 28(11): 1458–1461) [DOI: [10.13537/j.issn.1004-3918.2010.11.030](https://doi.org/10.13537/j.issn.1004-3918.2010.11.030)]
- Xing H M, Li Z H, Xu X G, Feng H K, Yang G J and Chen Z X. 2017. Multi-assimilation methods based on AquaCrop model and remote sensing data. Transactions of the Chinese Society of Agricultural Engineering, 33(13): 183–192 (邢会敏, 李振海, 徐新刚, 冯海宽, 杨贵军, 陈召霞. 2017. 基于遥感和AquaCrop作物模型的多同化算法比较. 农业工程学报, 33(13): 183–192) [DOI: [10.11975/j.issn.1002-6819.2017.13.024](https://doi.org/10.11975/j.issn.1002-6819.2017.13.024)]
- Xu Q Y, Yang G J, Long H L, Wang C C, Li X C and Huang D C. 2014. Crop information identification based on MODIS NDVI time-series data. Transactions of the Chinese Society of Agricultural Engineering, 30(11): 134–144 (许青云, 杨贵军, 龙慧灵, 王崇倡, 李鑫川, 黄登成. 2014. 基于MODIS NDVI多年时序数据的农作物种植识别. 农业工程学报, 30(11): 134–144) [DOI: [10.3969/j.issn.1002-6819.2014.11.017](https://doi.org/10.3969/j.issn.1002-6819.2014.11.017)]
- Xu W B and Tian Y C. 2005. Overview of extraction of crop area from remote sensing. Journal of Yunnan Agricultural University, 20(1): 94–98 (许文波, 田亦陈. 2005. 作物种植面积遥感提取方法的研究进展. 云南农业大学学报, 20(1): 94–98) [DOI: [10.16211/j.issn.1004-390x\(n\).2005.01.019](https://doi.org/10.16211/j.issn.1004-390x(n).2005.01.019)]
- Yan Y, Liu Q H, Liu Q, Li J and Chen L F. 2006. Methodology of winter wheat yield prediction based on assimilation of remote sensing data with crop growth model. Journal of Remote Sensing, 10(5): 804–811 (闫岩, 柳钦火, 刘强, 李静, 陈良富. 2006. 基于遥感数据与作物生长模型同化的冬小麦长势监测与估产方法研究. 遥感学报, 10(5): 804–811) [DOI: [10.11834/jrs.200605119](https://doi.org/10.11834/jrs.200605119)]
- Yang P, Tang H J and Liu J. 2000. Progress and outlook in automatic classification methods for crop yield estimation using remote sensing data. Journal of China Agricultural Resources and Regional Planning, 21(3): 57–60 (杨鹏, 唐华俊, 刘佳. 2000. 作物遥感估产中自动分类方法研究进展与展望. 中国农业资源与区划, 21(3): 57–60) [DOI: [10.3969/j.issn.1005-9121.2000.02.016](https://doi.org/10.3969/j.issn.1005-9121.2000.02.016)]
- Yang Y J, Zhan Y L, Tian Q J, Gu X F, Yu T and Wang L. 2015. Crop classification based on GF-1/WFV NDVI time series. Transactions of the Chinese Society of Agricultural Engineering, 31(24): 155–161 (杨闫君, 占玉林, 田庆久, 顾行发, 余涛, 王磊. 2015. 基于GF-1/WFV NDVI时间序列数据的作物分类. 农业工程学报, 31(24): 155–161) [DOI: [10.11975/j.issn.1002-6819.2015.24.024](https://doi.org/10.11975/j.issn.1002-6819.2015.24.024)]
- Ye L M, Xiong W, Li Z G, Yang P, Wu W B, Yang G X, Fu Y J, Zou J Q, Chen Z X, van Ranst E and Tang H J. 2013. Climate change impact on China food security in 2050. Agronomy for Sustainable Development, 33(2): 363–374 [DOI: [10.1007/s13593-012-0102-0](https://doi.org/10.1007/s13593-012-0102-0)]
- Zhang J, Zhang W T and Xing B L. 2014. Application of SCE-UA algorithm to water environmental system optimization. Water Resources Protection, 30(3): 61–64 (张景, 张文婷, 邢宝龙. 2014. SCE-UA算法在水环境系统优化问题中的应用. 水资源保护, 30(3): 61–64) [DOI: [10.3969/j.issn.1004-6933.2014.03.012](https://doi.org/10.3969/j.issn.1004-6933.2014.03.012)]
- Zhang J K, Cheng Y P, Zhang F W, Yue D P, Guo X X, Dong H, Wang J P and Tang H C. 2012. Crops planting information extraction based on multi-temporal remote sensing images. Transactions of the CSAE, 28(2): 134–141 (张健康, 程彦培, 张发旺, 岳德鹏, 郭晓晓, 董华, 王计平, 唐宏才. 2012. 基于多时相遥感影像的作物种植信息提取. 农业工程学报, 28(2): 134–141) [DOI: [10.3969/j.issn.1002-6819.2012.02.024](https://doi.org/10.3969/j.issn.1002-6819.2012.02.024)]
- Zhang M, Wu B F, Yu M Z, Zou W T and Zheng Y. 2015. Concepts and implementation of monthly monitoring of uncropped arable land: a case study in Argentina. Journal of Remote Sensing, 19(4): 550–559 (张森, 吴炳方, 于名召, 邹文涛, 郑阳. 2015. 未种植耕地动态变化遥感识别——以阿根廷为例. 遥感学报, 19(4): 550–559) [DOI: [10.11834/jrs.20154142](https://doi.org/10.11834/jrs.20154142)]
- Zhang S L, Chen L F, Su L and Jia L. 2015. A data assimilation-based method for optimizing parameterization schemes in a land surface process model. Science China Earth Sciences, 45(12): 1907–1922 (张生雷, 陈良富, 苏林, 贾立. 2015. 一种基于数据同化优化陆面过程模式参数化方案的方法. 中国科学: 地球科学, 45(12): 1907–1922) [DOI: [10.1007/s11430-015-5180-9](https://doi.org/10.1007/s11430-015-5180-9)]
- Zhang X, Jiao Q J, Zhang B and Chen Z C. 2008. Preliminary study on cropping pattern mapping using MODIS_EVI image time series. Transactions of the CSAE, 24(5): 161–165 (张霞, 焦全军, 张兵, 陈正超. 2008. 利用MODIS_EVI图像时间序列提取作物种植模式初探. 农业工程学报, 24(5): 161–165) [DOI: [10.3321/j.issn:1002-6819.2008.05.035](https://doi.org/10.3321/j.issn:1002-6819.2008.05.035)]
- Zhang X L, Wu S, Yan X D and Chen Z J. 2017. A global classifica-

tion of vegetation based on NDVI, rainfall and temperature. International Journal of Climatology, 37(5): 2318–2324 [DOI: 10.1002/joc.4847]

Zhu S and Zhang J S. 2016. Impact of land fragmentation on identification of winter wheat distribution accuracy by soft and hard change

detection method. Transactions of the Chinese Society of Agricultural Engineering, 32(10): 164–171 (朱爽, 张锦水. 2016. 地块破碎度对软硬变化检测法识别冬小麦分布精度的影响. 农业工程学报, 32(10): 164–171) [DOI: 10.11975/j.issn.1002-6819.2016.10.023]

Winter wheat mapping with globally optimized threshold under total quantity constraint of statistical data

GUO Wenqian^{1,2}, REN Jianqiang^{1,2}, LIU Xingren³, CHEN Zhongxin^{1,2}, WU Shangrong^{1,2}, PAN Haizhu^{1,2}

1. Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China;

2. Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture, Beijing 100081, China;

3. Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract: Using remote sensing technology to accurately obtain the spatial distribution information of crops in a large region is critical in improving the research level of resources and environment, strengthening the reaction capability of climate change, and ensuring national food security. Threshold method is a common method used to extract crop spatial distribution information based on time-series NDVI remote sensing data. However, several subjectivities and uncertainties were observed in traditional threshold methods when the threshold values were determined by humans, which hindered the improvement of threshold method identification accuracy. Considerable studies on determining the significant threshold value have been conducted in recent years. However, the accuracy and automation should be improved when the threshold values are determined.

Considering the above problems, Jingxian County of Hengshui City in Hebei Province, which is the main grain producing region in China, was selected as the typical experimental area. A global optimization algorithm was applied to select the threshold value in the threshold model, and a crop distribution mapping method using global optimization algorithms based on threshold detection was developed. In this study, winter wheat was selected as the research crop, and domestic Gaofen-1 satellite images that cover the entire winter wheat growing season were used as major data source. A new spatial distribution mapping method for winter wheat based on time-series NDVI data for the automatic optimization of threshold model parameters was proposed in this study by using the global optimization algorithm Shuffled Complex Evolution-University of Arizona (SCE-UA) and crop area statistics as the total quantity constraint. First, the NDVI curves of winter wheat growth season were established, and the key phenological characteristics of winter wheat were extracted based on crop phenological and training sample data. Second, the general threshold model of winter wheat spatial distribution extraction was established, and the threshold parameters to be optimized were determined based on the seasonal variation and characteristics of time-series winter wheat NDVI at key phenological stages (such as seeding, tillering, jointing, heading, and maturation stages). Third, the optimal parameters of winter wheat spatial distribution model were obtained by the global optimization algorithm SCE-UA by using the crop planting area statistics at county level as total quantity constraint data and objective function. Finally, the optimal parameters of winter wheat threshold model were obtained in the study region, the spatial distribution of winter wheat was extracted by using the optimized threshold parameters, and the winter wheat distribution extraction results were validated by the ground verification samples.

The final validation results showed that the accuracy of winter wheat identification results reached 99.99%, which proved that proposed threshold parameter optimization method had a good total quantity constraint effect and other classification accuracies had reached a high level. The overall accuracy and kappa coefficient were 97.03% and 0.94, respectively. Compared with the traditional threshold, support vector machine, and maximum likelihood methods, the overall classification accuracy of the proposed method increased by 4.55%, 2.43%, and 0.15%, respectively. The kappa coefficient of the proposed method increased by 0.12, 0.06, and 0.01, respectively. The above performances indicated that the crop distribution mapping optimization method based on the threshold model parameters of statistical data total quantity constraint and global optimization algorithms was effective and feasible to accurately obtain the spatial distribution mapping results of crops in a large region. This study can serve as basis to improve the accuracy and automation level of crop spatial distribution identification in China and provide several technical support and recommendations for other studies on crop spatial distribution extraction and crop mapping under complex planting pattern at large scale.

Key words: remote sensing, crop mapping, threshold method, total quantity constraint, global optimization algorithm, winter wheat

Supported by National Natural Science Foundation of China (No. 41471364, 61661136006); National High Technology Research and Development Program of China (863 Program) (No. 2012AA12A307)