

引用格式: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(5):937-947. [Tan S, Wu B F, Zhang X. Mapping paddy rice in the Hainan Province using both Google Earth Engine and remote sensing images[J]. Journal of Geo-information Science, 2019,21(6):937-947.] DOI:10.12082/dqxxkx.2019.180423.

基于Google Earth Engine与多源遥感数据的海南水稻分类研究

谭 深^{1,2}, 吴炳方^{1,2*}, 张 鑫¹

1. 中国科学院遥感与数字地球研究所 遥感科学重点实验室, 北京 100101; 2. 中国科学院大学, 北京 100049

Mapping Paddy Rice in the Hainan Province Using both Google Earth Engine and Remote Sensing Images

TAN Shen^{1,2}, WU Bingfang^{1,2*}, ZHANG Xin¹

1. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Science, Beijing 100101, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: Rice is one of the main grain crops in China and East Asia, including China. The annual yield of rice has a significant influence on domestic livelihood. Therefore, timely and accurate assessment of rice distribution information is crucial for forecasting rice yields and optimize the allocation of agricultural resources. Remote sensing (RS) images can provide time series surface spectral, and other electro-magnetic, dynamics over a large-scale land surface, which are commonly used for large-scale crop monitoring. However, routine rice classifying strategies provided by the RS images during key growth stages, require spectral patterns at high frequency. This method appears to be impractical within South China, as the number of high quality RS images are difficult to obtain due to cloud contamination caused by the hot and wet weather. A combination of various RS images of rice classification from multi-platforms provide an indirect way of reducing the revisit period in routine rice classification, thus enabling successful crop mapping in cloudy regions. However, this causes difficulty with data manipulation and storage, especially when conducting classification work at province or large area levels. To address these issues, this research utilizes Google Earth Engine (a cloud-based geospatial analysis platform running on the Google server) to collect online optic RS data and micro-wave RS data at diverse resolutions for rice mapping. A distribution map of paddy rice at 10-m spatial resolution in the Hainan Province in 2016 was made by using the combined methods of random forest (RF) classification and a pattern-matching strategy based on conjunct features extracted at monthly level and histogram value distribution. Results showed this method

收稿日期:2018-08-30;修回日期:2018-12-24.

基金项目:中国科学院科技服务网络计划(STS计划)项目(KFJ-STZ-ZDTP-009);国家重点研发计划(2016YFA0600301、2016YFA0600302);中国科学院战略性先导科技专项(A类)(XDA19030201);国家自然科学基金(41561144013、41701496)。 [**Foundation items:** Science and Technology Service Network Initiative(STS), No.KFJ-STZ-ZDTP-009; National Key R&D Program of China, No.2016YFA0600301, 2016YFA0600302; Strategic Priority Research Program of the Chinese Academy of Sciences(A), No.XDA19030201; National Natural Science Foundation of China, No.41561144013, 41701496.]

作者简介:谭 深(1992-),男,辽宁丹东人,博士生,主要从事农业与水资源遥感研究。E-mail: tanshen@radi.ac.cn

*通讯作者:吴炳方(1962-),男,江西玉山人,博士,研究员,主要研究领域包括农业遥感与粮食安全、水资源遥感与耗水管理、生态遥感等。E-mail: wubf@radi.ac.cn

was suitable for rice mapping in Hainan and could show clear feature divergence between the different land surface cover types. Spatial distribution results corresponded well with the actual edges of the field, along with texture information. The rice classification result of the Hainan Province was validated using sample points captured on the ground and achieved overall accuracy of 93.2%, indicating reliability for practical application. Overall, the automatic rice classifying strategy was able to map paddy rice with high efficiency and sufficient accuracy in the Hainan Province, and could be applied to other vast areas.

Key words: Mapping paddy rice; Google Earth Engine; SAR; random forest; remote sensing images; Hainan Province

***Corresponding author:** WU Bingfang, E-mail: wubf@radi.ac.cn

摘要:水稻是中国乃至亚洲的重要粮食作物之一,稻米产量关系到民生福祉。及时、准确的水稻分布信息是监测水稻产量、调控农业资源配给的基础。遥感(Remote Sensing)技术能够提供大范围地表的时间序列光谱变化特征,常用于大尺度范围的作物监测。然而,传统基于水稻生长关键时期光谱特征的分类、提取方法对遥感数据的时间分辨率要求较高。由于我国南方水稻产区湿热,雨季云污染降低了遥感数据的有效时间分辨率,因此上述方法在该地难以推广。融合多源遥感数据的分类方案变相缩短了卫星的重访周期,使多云气候区基于遥感影像的水稻分类成为可能。然而,集成多源数据所需更高的数据处理效率和存储需求也成为限制省级乃至更大范围水稻分类的主要因素。本研究基于谷歌地球引擎(Google Earth Engine)云平台,在线调用中分辨率的光学、微波遥感数据,创新性地采用了按月提取、按直方图大小提取特征的方式,采用随机森林分类器,绘制海南省2016年10 m分辨率水稻种植分布图。实验结果证明,该方法可以用于南方多云地区水稻分类,提取结果能够体现不同地类之间的差异,且与实际地表的地块边界、纹理符合良好。经过地表样本点的验证,总体精度为93.2%,满足实际应用需求。因此,本研究采用的自动分类流程能够准确、高效地提取海南省的水稻种植范围,可以向其他地区大范围推广。

关键词:水稻分类;谷歌地球引擎云平台;微波数据;随机森林;多源遥感数据;海南省

1 引言

近年来,世界人口增加、耕地承载力下降等全球性问题引起全社会对粮食安全的关注。水稻作为主要粮食作物,为世界上近50%的人口提供日常能量来源^[1],全球范围约12%的耕地种植水稻,并且这个数字仍在逐年增加^[2]。中国更是以水稻作为主食的国家:全国约65%的人口以水稻作为主食^[3],水稻产量关系到社会稳定。此外,水稻特殊的种植形式与高耗水的特性,也使其成为水资源管理的重要内容与评价指标^[4-5]。因此,准确、快速地获取水稻种植面积及空间分布信息,对估计粮食产量、优化农业管理具有重要意义^[6]。

传统方法获取水稻种植面积依赖于统计上报与地方实测,缺少空间分布的刻画且耗资巨大。遥感技术能够稳定、准确地提供大范围地表的同步观测结果,已被广泛应用于土地利用、植被监测相关的研究,也为高效获取水稻种植范围提供了可能。总体上,基于遥感数据的水稻提取根据输入数据的不同,可以分为基于光学数据和基于微波数据两种方法。前者通过水稻生长关键时期(如移栽期灌水前后)遥感影像光学波段的反射率或植被指数与其

他地类在空间域或时间域上的变化特征,区分水稻与其他地类。由于提供光学波段的卫星众多,该方法适应性更广,但其原理决定需要至少一景关键时期清晰的遥感影像,以获取所需的特征信息^[7-10]。如Li等^[9]通过不同地类间植被指数的差异研究鄱阳湖周边水稻种植面积的变化;黄青等^[11]通过逐像元检测植被指数在时间域的变化特征提取水稻种植;Mrinal等^[12]将地表分解为特征相近的多边形对象,并检测其特征与水稻特征是否相符。研究表明,在光学数据能有效覆盖整个生长季的前提下,基于上述光谱的差分的分类方法具有较高的精度与可信度^[13]。然而,光学遥感信号受云层遮挡,经常形成无效观测,云污染对于重访周期较长的中、高分辨率数据(如30 m或更高分辨率)影响更为显著,可能造成1-2个月的数据缺失。水稻通常种植在温暖、湿润的地区,产生云污染的几率更高,导致分类精度降低。低分辨率遥感数据虽然通常具有更高的时间频率,但国内农田往往支离破碎,地表异质性强,易形成混合像元,影响分类效果^[14]。因此,选择合适的策略或数据组合,规避云污染对分类结果的影响是水稻分类研究的重点。

基于微波的合成孔径雷达(Synthetic Aperture

Radar, SAR)对云层有着极强的穿透力,可以在无法获取光学数据的天气观测地表。利用极化后的后向散射系数对水体的敏感性可以有效区分水稻及其他旱生作物^[15-16]。然而,受限微波数据获取方式,传统的研究集中在小区域内或空间分辨率较低,不具备大规模推广的条件^[17-19]。欧盟委员会于2016年通过哥白尼计划,完成了 Sentinel-1 卫星的组网,通过两颗卫星连续观测全球地表,提供免费的、近实时的地表雷达观测及处理结果。该数据集空间分辨率为5~20 m,同一地区临近影像获取间隔小于10 d^[20],可以作为大范围水稻分类的数据源,弥补了应用微波数据提取水稻种植范围的数据短板。

受限同一传感器无法同时提供高时-空分辨率的光学遥感数据,融合不同遥感数据的分类方法也成为研究的热点,以结合不同传感器在时间、空间域上的分辨率优势^[11,21-22]。如融合常见的 Sentinel-2 和 Landsat 数据^[23],实现对省级乃至更大范围地表的高频观测。由此带来前所未有的数据存储与处理能力的需求,只能通过云计算实现。谷歌地球引擎(Google Earth Engine, GEE)借助谷歌公司强大的计算资源与海量在线数据,已成为地理、空间相关研究的重要工具^[24]。用户通过调用 GEE 平台收集的大量已发布的地理数据产品,结合其提供的算法与自身的需求,实现在线计算,大大减轻了数据获取、处理的工作量。借由该平台,研究人员可以大大扩展自身原有研究的覆盖范围,提供国家乃至全球尺度的研究成果^[25-26]。

鉴于单一光学遥感数据往往不足以满足湿、热地区水稻分类的需求, Sentinel-1 全球产品的发布为光学-微波数据的结合提供了可能。本研究因此旨在融合 Landsat-8 与 Sentinel-2 光学数据、Sentinel-1 微波 SAR 数据与其他辅助数据,消除南方地区多云天气对水稻提取的限制,探讨一种工程化的分类策略。研究采用随机森林分类器,抓取光学-微波数据在时间域上的特征。整个分类过程借助 GEE,采用云端-本地结合的方式,提取海南省2016年水稻种植范围。

2 实验区概况、数据来源与研究方法

2.1 实验区概况

海南省以海南岛(19°20'N-20°10'N, 108°21'E-111°03'E)为主体,包括三沙群岛及大面积海域,是我国最南端、纬度最低的省份。主岛面积约占该省

陆地总面积的97%,因此下文所有海南省均指代主岛。由于地理位置接近赤道,该省气候湿润多雨,长夏无冬。年平均气温为25℃,≥10℃积温为8200℃,年平均降水量为1639 mm。良好的水热环境为水稻及其他作物提供了理想的生长条件。省内大多数作物一年两到三熟,水稻通常只播种两季,如图1(a)所示为该省水稻的归一化植被指数(NDVI)过程线与物候变化(阴影所标注的植被指数区间为汇总全省水稻样本的标准差):其中,3月至7月上旬为早稻种植季;7月下旬至11月下旬为晚稻种植季^[27]。由于两季水稻种植面积之间存在细微差异,晚稻面积通常较早稻略大。因此,下文水稻提取的面积均以晚稻为准。

海南省湿热的气候有利于作物生长,也令基于光学影像的水稻分类工作难以稳定开展。如图1(b)所示为海南省3-4月(水稻移栽期)可以利用的光学数据像元数量分布图,包括 Landsat-8 和 Sentinel-2 两种数据源。可以看出,受云层影像,超过90%地区在移栽期没有或仅有一景可用的光学影像,全年可用的影像也仅有5~8景,无法基于此开展分类。地面调查范围覆盖海南岛的全部农业种植区,样点通过智能手机安装的 GVG 软件获取^[28-29],分布如图1(c)所示。有效水稻样点共计2166个。分类过程中,50%的样点作为训练样本,其余50%的样点作为精度验证。

2.2 数据来源

2.2.1 遥感数据

研究过程中,所有遥感数据均通过 GEE 平台在线调用、处理,包括:Landsat-8 卫星搭载的陆地成像仪 Operational Land Imager (OLI)、Sentinel-2 A/B 卫星搭载的多光谱成像仪 Multi-Spectral Instrument (MSI);以及 Sentinel-1 卫星的 C-band 提供的雷达回波信息提取地表特征。3种遥感数据源的各波段信息如表1所示。

数据池中的 Landsat-8 影像为经过在轨纠正的反射率产品,几何偏差小于12 m^[30]; Sentinel-2 数据为经过大气和几何校正的 Level-1C 产品。受助于 GEE 平台的运算和数据管理机制,用户在运算过程种无需过多考虑不同数据源之间的分辨率匹配,仅限定分析结果的输出分辨率即可。本研究中应用的 Sentinel 系列数据空间分辨率为10 m,因此, Landsat 数据后台自动采样至10 m与之相配。上述数据通过 GEE 内嵌的算法统一坐标系,保证每个像

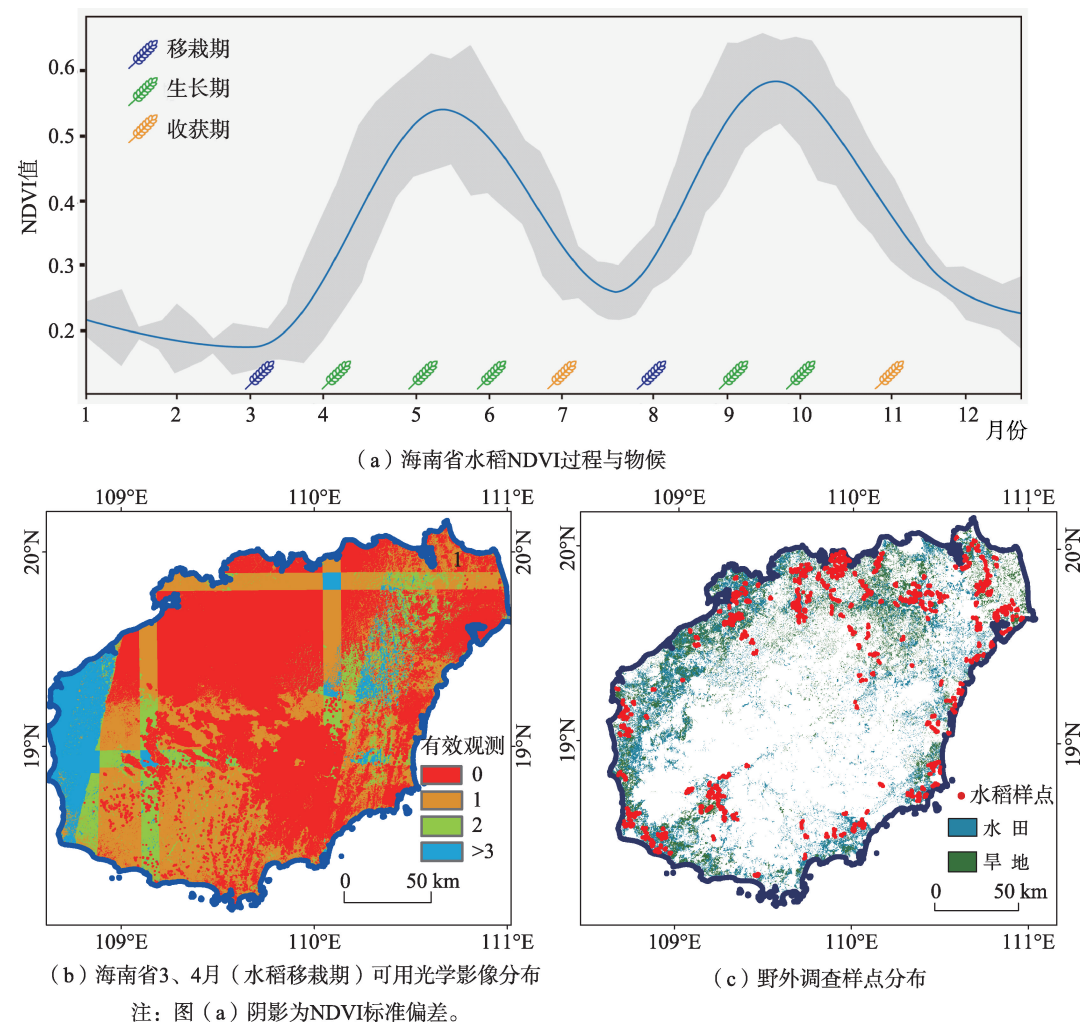


图1 海南省水稻实验区基本信息

Fig. 1 Basic information of research area of Hainan Province

表1 分类所需遥感影像波段参数

Tab. 1 Band information of images employed in the research

传感器		波段	波长/nm	分辨率/m	数据量/景
Sentinel-2 MSI	B2	Blue	490	10	340
	B3	Green	560	10	
	B4	Red	665	10	
	B8	Near-infrared	842	10	
	B11	Short-wave infrared 1	1610	20	
	B12	Short-wave infrared 2	2190	20	
Landsat-8 OLI	B2	Blue	450~510	30	106
	B3	Green	530~590	30	
	B4	Red	640~670	30	
	B5	Near-infrared	850~880	30	
	B6	Short-wave infrared 1	1570~1650	30	
	B7	Short-wave infrared 2	2110~2290	30	
Sentinel-1	VV	Dual-band crosspolarization, vertical transmission		10	月合成
	VH	horizontal receiver		10	

元代表相同的地表范围,运算过程中不会产生几何失配。所有 Landsat-8 数据通过综合考量辐射亮温、表面温度、归一化冰雪指数(NDSI)等特征,标注受云影响的像元^[31]; Sentinel-2 数据则通过自带的 QA60 波段去除云像元^[32]。标注为云的像元均不参与后续计算。而质量合格的像元经过整编后,计算后续分类所需地表参数。

垂直极化后 Sentinel-1 数据在线调用预处理模块完成标定轨道参数、去除热噪声、辐射定标、地表几何纠正等操作。由于 SAR 数据质量很大程度受观测角度影响,即距成像中心距离越大的像元噪声越强烈。因此,预处理后的 SAR 数据以月为单位,逐像元计算其信号在本月内的中位数,以排除信号对成像角度及地表状况过敏感带来的噪声。最终得到空间分辨率 10 m,幅宽 250 km 的月合成产品。处理后的遥感数据量如表 1 所示。

2.2.2 辅助数据

分类结果的精度与训练样本的数量和代表性显著相关。因此,在地面验证样点数量、代表性不足的情况下保证理想的分类精度是衡量分类流程鲁棒性的关键。本研究的目的之一是探讨一种工程化的水稻分类策略,为保证后续推广过程中任何研究区域的训练过程能够输入足够的样本数量,在地面调查样点输入以外,采用基于像元的稻田分类方法(PPPM),选择光学数据时间序列完整的、置信度高的地块作为辅助,共同进入到训练及特征提取中。PPPM 方法具有较高的分类置信程度,但需要完整时间序列的光学数据来检测移栽期后的灌水信号,因此提取出的范围很小,但具有可靠精度,可以作为辅助数据解决地表样本点不足的问题^[13]。

另外,训练过程中增加了部分限定条件以提升精度。根据统计,水稻仅生长在 0~8° 的坡面上,因此引入多种基于遥感的辅助数据作为限定条件,如:基于 SRTM 的数字化高程模型(DEM)及其衍生产品(坡度、坡向)^[33];另外,水体、滩涂等地类与移栽期水稻具有类似的光谱特征,通过引入水体掩模去除其对分类结果的影响^[26]。土地利用数据来自于 ChinaCover,该数据集基于面向对象方法制作,全国范围内二级类精度大于 86%,具有很高的可靠性^[34],并且已在其他方向的研究中取得了良好的辅助作用^[35-36]。该数据每五年更新一次,本研究中使用的是 2015 年数据。上述辅助数据的分辨率均为 30 m,且投影与遥感数据相配。

2.3 水稻分类方法

基于遥感数据进行分类的基本理念是通过检测每个像元时间、光谱矢量特征与基于地面样点归纳出的水稻特征之间的差异,提取与水稻接近的区域。其中,水稻生长关键时期(如移栽期、灌浆期)的植被指数或其他指标的突变作为重点检测的特征。本研究融合了 2 种光学遥感数据源,且引入高分辨率 SAR 数据,以补充水稻生长的关键时期缺失的光学数据。此外,考虑到水体、湿地等类别易与水稻产生混淆,引入土地利用、水体、地形等数据作为建立特征库的辅助。所有数据经过整理后,输入随机森林分类器处理。其中,涉及数据预处理及特征检测、识别等需要大量运算、存储资源的操作在云端实现;结果判读、验证等需要人工干涉的工作在本地进行。具体流程如图 2 所示。

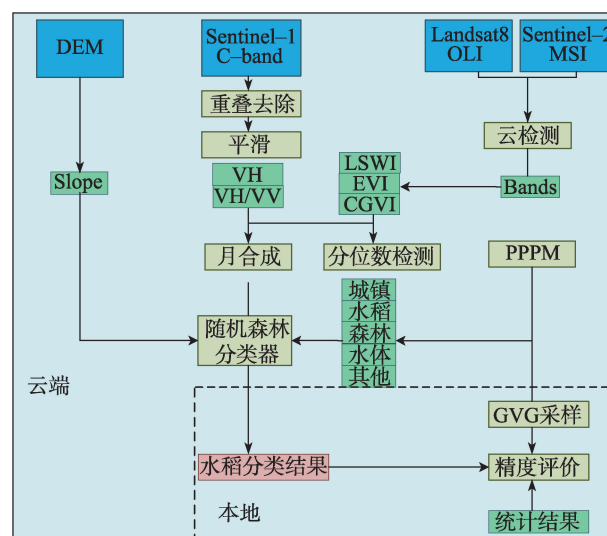


图2 基于GEE的水稻分类流程

Fig. 2 Rice mapping based on GEE

2.3.1 分类指标

预处理后的光学数据用来计算地表参数,其中包括:常用来描述植被长势的NDVI、增强型植被指数(EVI)^[37];描述植被冠层及土壤表层含水量的地表水体指数(LSWI)和归一化水体指数(NDWI)^[38];描述植被光合作用活性的绿色叶绿素指数(GCVI)。具体指标计算公式如表2所示,所得结果均存放在云端的数据池。其他指标包括 Sentinel-2、Landsat-8 的光学波段;经过极化、月合成后的 SAR 波段;以及坡度、坡向等与时间无关,但可以参与辅助判断的信息。所有数据具有相同的空间分辨率与投影,整编成一个时间-空间-光谱的多维数组,进入分类器。

表2 遥感地表参数计算公式
Tab. 2 RS-based surface parameters

地表参数	计算方法
NDVI	$NDVI = (NIR - Red) / (NIR + Red)$
EVI	$EVI = 2.5 \times (NIR - Red) / (NIR + 6 \times Red - 7.5 \times Blue + 1)$
LSWI	$LSWI = (NIR - SWIR) / (NIR + SWIR)$
NDWI	$NDWI = (Green - NIR) / (Green + NIR)$
GCVI	$GCVI = NIR / Green - 1$

注: Red为传感器红光波段; Green为绿光波段; Blue为蓝光波段; NIR为近红外波段; SWIR为短波红外波段。

2.3.2 特征提取

随机森林分类器在训练过程中通过迭代选择有限条件子集,构建相互独立的决策树,可以有效处理大量的输入指标^[39-40]。该方法允许在不显著增加计算量的前提下,提供更为快速、可靠的分类结果,因此在影像分类相关的研究中得到了广泛的应用^[41-42]。与其他遥感分类器相比,该算法更适用于本文中涉及的多源遥感数据所带来的高维指标情境,具有运算、训练速度快且不易过拟合的优点;并且分类器内部隐含对指标分类效果的判别权重,以突出利于分类的特征。另外,该分类器与其他方法相比运算量大的缺陷,在本研究中通过调用云平台海量的计算资源得以弥补^[43]。最终,本文基于GEE平台调用该分类器,构建了300棵具有5个以上节点的决策树用以提取水稻特征。

训练过程中,我们选取部分地面调查获取的水稻样点及通过PPPM提取的特征充分的像元作为样本,提取全部55个指标(包括LSWI、EVI和GCVI等地表参数、全部光学波段反射率、SAR波段,以及地形等参考信息)在整个生长季的变化特征。海南省气候湿润,遥感数据云污染严重。为了解决光学数据缺失的问题,本研究采用了一种更为灵活的,基于百分位数的特征匹配方式^[44]:归纳、统计所有55个波段、参数在完整生长周期内的直方图,检测5%,25%,50%,75%和90%五个百分位数的特征值与出现时间,计算其90%置信区间。逐像元匹配上述所有参数的五个百分位点,至少4个特征点落在置信区间内即视为与样本特征一致。这种方法优点有二:可以有效规避部分地区有效影像无法覆盖完整生长季的问题,即只需每个参数大概率符合水稻的过程特征即可判断;避免出现少部分地区可能出现的播种期提前或滞后现象。

上述步骤基于GEE平台实现,全部的遥感影像被汇总成的多维数组,并提取每个像元的特征矢

量,与随机森林建立的判断条件进行比较,提取水稻像元。再通过面积滤波,去除细碎像元造成的椒盐噪声。最终,下载分类完成的结果到本地,通过验证样点评价结果精度。

3 结果及分析

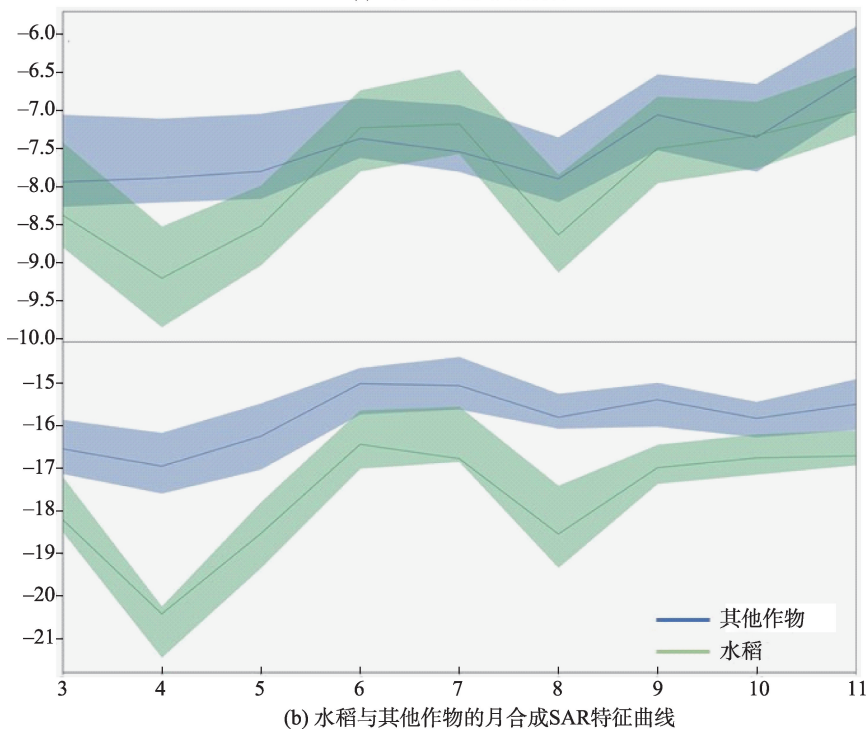
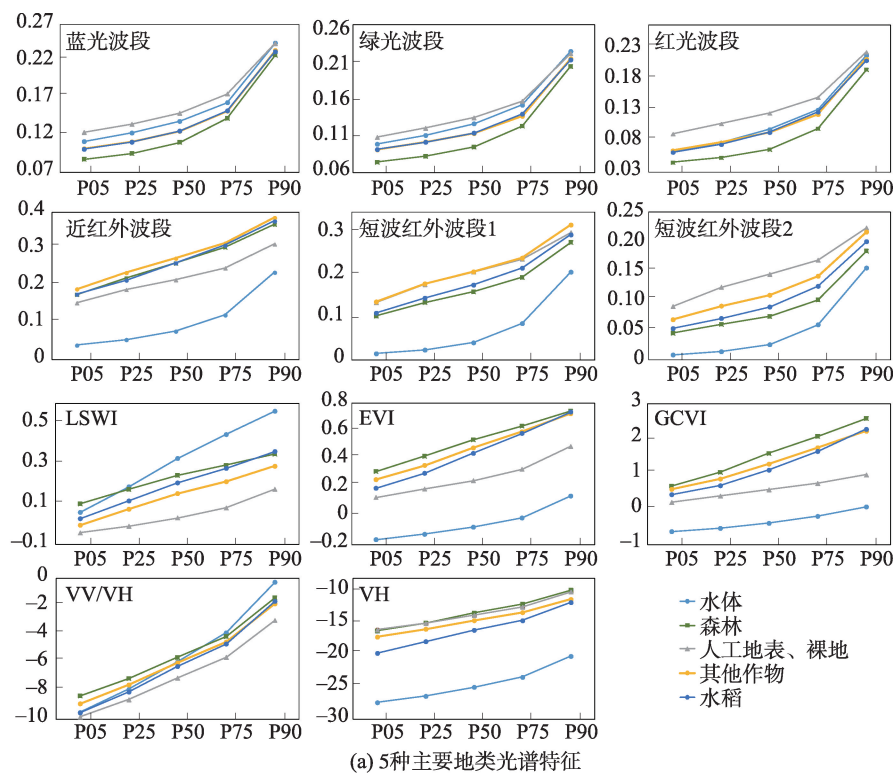
3.1 主要参数特征分析

如上文所述,处理、整编后的遥感数据被输入到随机森林分类器中。考虑到充分体现不同下垫面之间的差异,除水稻外,其他类别被汇总成为水体、森林、城镇(裸地)和其他作物4个类别。提取出的5个类别在11个典型参数的时序特征如图3所示。由于水稻特殊的生长形式,其兼具植被与水体的光谱特征而易造成混淆和错分。叶绿素在蓝光波段具有强烈的吸收峰,因此该波段稻田的特征与其他作物几乎一致;由于海南省地处热带,林木茂盛、叶片密度较大而光合作用强烈,森林在该波段的吸收率较农田更高,可以用来区分农田、森林与水体。各地物类别在绿光波段的光谱特征与蓝光波段接近;而在红光和近红外波段,森林、水稻和其他作物由于植物生长对红光的强烈吸收,三者光谱特征接近。水稻在短波红外波段的光谱特性与其他地类区分度较高。尤其在第二个短波红外波段(2190 μm),3种植被下垫面与水体之间光谱特征差异明显。基于此波段计算的地表水体指数LSWI对表层含水量敏感,可以放大这种差异,作为水体和水稻的区分依据;而常用来描述地表植被状态的EVI和GCVI,在作物生长中、后期无法明显区分出水稻和其他作物。

微波SAR数据不受天气的影响,可以合成完整的逐月数据集。利用水体对微波VH波段数据的敏感性,可以明显区分出水体、水稻和其他地类。图3(a)所示为水稻与其他作物SAR的时间变化曲线。可以看出,常用做水稻分类的VH/VV波段,时间变化

曲线在六月之后与其他作物区分度并不高,二者变化趋势相同且显著置信区间有很大程度的重合,影响分类判断。因此分类过程中我们采用了单一VH

波段,以体现二者的区分度,效果如图3(b)所示:二者重叠区域消失,差异显著;且水稻信号的变化范围更加稳定,基于此的分类效果更好。



注:图(a)中的P代表将全部观测值从小到大排序所获取的百分位数,其中P05、P25、P50、P75、P90分别为第5个、第25个、第50个、第75个、第90个百分位数;图(b)中的实线为平均值,彩色标注为显著置信区间。

图3 主要地物光谱特征提取结果

Fig. 3 Spectral pattern of main land cover types

上述结果表明,水稻与其他地类之间的差异可以通过不同参数和波段的组合加以放大。如蓝光波段可用来提取水稻及其他耕地;基于短波红外波段与LSWI指数可以有效区分不同地物类型;单一VH波段对水体和水稻具有较高的区分度。而基于单一时相或生长关键时期可见光遥感数据或植被指数的水稻分类,在海南地区极易受异物同谱的影响产生错分。

3.2 水稻提取结果

提取样本提供的光谱特征后,后台自动建立随机森林并逐一判断输入像元的归属类别,最终

输出海南省2016年10 m空间分辨率水稻分布结果(图4(a))。分类结果空间分布趋势与图1中展示的全省30 m土地利用中的水田分布区域一致:海南岛中部山地由于地形起伏较大,稻田罕见分布;大规模水稻种植区分布在北部和东北部、以及南部沿海地区。全岛水田地块面积破碎、形态不规则,较少出现规整形状的地块。不同地块的物候期可能略有差异,且与周边林地、其他作物等地类产生混淆。由此可见,单一传感器的高分辨率影像数据量不足以满足水稻分类的需求。而采用逐日低分辨率遥感数据提取海南省水稻种植范围,极易受到混合像元的干扰。

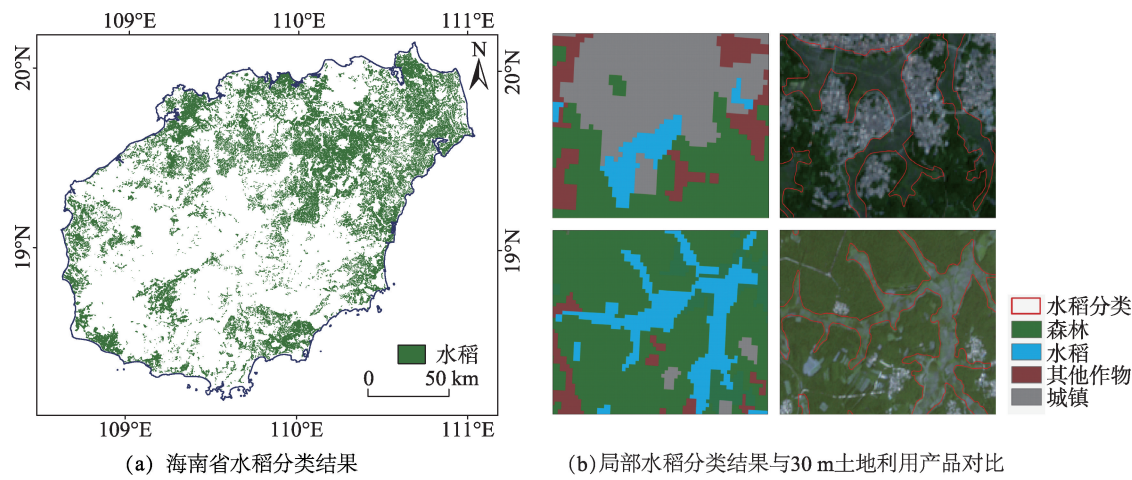


图4 海南省水稻分类结果示意及局部分类结果与30 m土地利用产品对比
Fig. 4 Rice maps of the Hainan Province and comparison of time-series SAR of rice and other crops

图4(a)选取文昌市与乐东县2个典型区域的水稻提取结果与土地利用、原始影像对比。可以看出,土地利用产品基于30 m分辨率Landsat影像制作,边缘与背景10 m分辨率Sentinel-2数据匹配度不高,有明显的锯齿状像元边缘纹理。虽然该产品对森林和城镇的提取保持着良好的准确性,但由于未利用微波数据,对水田和旱地的区分度较低,在图示的2个样例中都产生了小范围混淆与错分。基于本文方法提取的水稻地块边界与10 m分辨率影像契合度很高,在稻田-旱地与稻田-森林等易产生混淆的过度地带没有发生明显的错分与漏分现象,准确性良好。

经过汇总,最终得到海南省水稻种植面积为151.4千hm²,与实际情况接近。常规水稻面积统计更多依赖于结合地籍数据的填图,地块分割的过程中可能包含部分田埂、防护林等,基于遥感数据的分类结果与地表实际状况符合程度更高。检验分类结果精度的混淆矩阵如表3所示。用于验证的水稻样点

共1083个(部分样本通过PPPM方法获取),其中错分78个,生产者精度为92.7%。用于其他地类验证的样点共257个,错分13个,用户精度为98.7%,FScore为95.6%。综上结果表明,基于非监督的随机森林方法对于云污染严重的海南省表现稳定、可靠,能够有效区分水稻与水体、旱作农田等相似地类。分类特征提取过程中不进行人工干预,完全由算法后台分析。因此,本方法满足工程化自动运行的条件。

表3 精度评价混淆矩阵					
Tab. 3 Accuracy assessment of rice in Hainan Province					
海南省水稻		验证样点		总计	用户精度/%
分类验证		其他地类	水稻		
分类结果	其他地类	244	78	322	75.8
	水稻	13	1005	1018	98.7
总计		257	1083	1340	
生产者精度/%		94.9	92.7		
总体精度/%		93.2			
FScore		95.6			

4 讨论

常规的基于光学数据的水稻分类方法在中国南方水稻主产区的分类工作中极易受到多云天气的影响,无法形成有效的光谱时间序列,影响进一步提取水稻生长关键时期的特征。很多方法的基本原理在于检测水稻像元(或分割后对象)在移栽期灌水前后的信号差异。而海南省及相邻水稻大省地处热带,炎热潮湿多云,该时期有效光学数据至多仅有1景,大部分地区甚至没有可用数据,不足以满足分类需求。因此本研究采用光学-微波数据结合的方式,利用不受云层影响的 Sentinel-1 微波数据,形成完整的时间序列 SAR 数据集。并针对南方湿、热地区云污染概率高、水稻播种日期不统一的特点,引入了一种更为灵活的、基于百分位数的光谱信息匹配策略。该方法只需时间窗口内的特征变化趋势大概率符合水稻特征即可判定为水稻像元,尤其适合炎热地区物候模糊的情况,在本研究中取得了良好的分类效果。

在指标的选择上,我们选用了两个卫星平台的全部光学、红外波段、微波后向散射系数以及反应地表植被、水分状况的遥感参数等,共计55个参数作为特征提取的基础,以充分区分不同下垫面类型之间的差异。随机森林分类器在无需额外人工干预的情况下,自动选取特征,建立决策树,实现逐像元的特征匹配。该方式提取特征较常规人工建立决策树的方法更为客观。大体量的随机森林能够遍历全部特征,挖掘更具有普适性的判定组合,更适用于研究区域较大,而且地表状况复杂地区的水稻分类^[45-46]。基于该方法进行的全省范围的特征提取、匹配操作所需要大量的存储与运算资源通过 GEE 平台实现;而结果判读、精度验证等需要人工干预的操作在本地实现。这种云端-本地结合的方式能够利用云端多线程运算资源,允许建立更为复杂、完善的分类体系,从而更高效的完成省级乃至全国尺度的分类工作。

本研究所采用的分类策略在操作过程中无需人工干预,只需输入地表样本点的位置与属性,由服务器端自动提取、归纳对应位置的特征信息。只有在结果判读、质量控制的阶段需要人工干预。考虑到海南等地区样本点所在位置影像质量可能不足的情况,引入 PPPM 方法提取置信程度较高的水稻像元,作为辅助样本提取其对应的特征。这种工程化的分类方案具有更好的稳定性,在各种气候条

件、样本采集状况下都能够输出客观、稳定的分类结果,具备向其他区域推广的能力。

目前,很多研究人员在分类过程中融合水稻地块的纹理信息,采用面向对象的方法,取得了良好的分类精度与空间效果^[12,47-48]。本研究采用的分类方法基于随机森林实现。但该方法的分类流程目前仍无法有效结合面向对象的思维,在特征提取和分类过程中引入影像中隐含的纹理和形状信息。只能对基于像元的分类结果进行面积滤波去除噪声,而该策略无法去除地块内的噪声像元。因此,该方法在海南等地块形态不规则、分布破碎的地区仍具有改进的空间。未来研究的重点将放在影像分割与特征匹配过程的融合,即以对象为单位提取光谱以及空间纹理特征。

中国幅员辽阔,水稻种植区自北端黑龙江省向南端延伸,水稻物候也由单季稻、单双混合向双季稻过度。本研究选取了典型种植双季水稻的海南省。未来的研究计划基于当前分类框架,探索单-双混合种植区和单季稻种植区水稻分类的方法。实现覆盖全国的、工程化的10 m分辨率水稻自动分类。

5 结论

本研究基于谷歌地球引擎,融合 Sentinel-2、Landsat-8 可见光数据和 Sentinel-1 微波 SAR 数据等多个传感器的不同波段和参数组合,绘制海南省10 m空间分辨率2016年水稻种植区域分布图。根据1083个地面样本点的验证精度为92.7%,表现良好。空间精细度、准确性较基于30 m分辨率的分类产品更好,能够契合海南地区不规则的地块、地类边界,且对于常规光学影像易产生混淆的地类区分度较高。该分类策略能够有效规避我国南方湿、热气候条件下,云层对光学遥感影像的污染,且无需额外的人工干预,在样本不足的情况下也能够得到稳定的数据结果,云端-本地结合的方式也使其具备大规模推广的条件。总体上看,基于该方法得到的分类结果满足进一步研究的需求。

参考文献(References):

- [1] Elert E. Rice by the numbers: A good grain[J]. Nature, 2014,514(7524):50-51.
- [2] Khush G S. What it will take to feed 5.0 billion rice consumers in 2030[J]. Plant Molecular Biology, 2005,59(1):1-6.
- [3] Zhang Q, Zhang W, Li T, et al. Projective analysis of staple food crop productivity in adaptation to future climate

- change in china[J]. *International Journal of Biometeorology*, 2017,61(8):1-16.
- [4] Bouman, B. How much water does rice use[J]? *Rice Today*, 2009,69:115-133.
- [5] Tao F, Hayashi Y, Zhang, Z, Sakamoto T, et al. Global warming, rice production, and water use in china: Developing a probabilistic assessment[J]. *Agricultural & Forest Meteorology*, 2008,148(1):94-110.
- [6] Knauer K, Knauer K. Remote sensing of rice crop areas [J]. *International Journal for Remote Sensing*, 2013,34(6): 2101-2139.
- [7] Ramankutty N, Foley J. Characterizing patterns of global land use: An analysis of global croplands data[J]. *Global Biogeochemical Cycles*, 1998,12(4):667-685.
- [8] Ramankutty N, Foley J A. Estimating historical changes in global land cover: Croplands from 1700 to 1992[J]. *Global Biogeochemical Cycles*, 1999,13(4):997-1027.
- [9] Li P, Feng Z M, Jiang, et al. Changes in rice cropping systems in the poyang lake region, china during 2004-2010[J]. *Journal of Geographical Sciences*, 2012,22(4):653-668.
- [10] Thenkabail P S. Mapping rice areas of south asia using modis multitemporal data[J]. *Journal of Applied Remote Sensing*, 2011,5(1):863-871.
- [11] 黄青,吴文斌,邓辉,等.2009年江苏省冬小麦和水稻种植面积信息遥感提取及长势监测[J].*江苏农业科学*,2016(6):508-511. [Huang Q, Wu W B, Deng H, et al. Mapping area and condition monitoring of winter wheat and rice in Jiangsu province in 2009[J]. *Jiangsu Agricultural Sciences*, 2016(6):508-511.]
- [12] Singha M, Wu B, Zhang M. An object-based paddy rice classification using multi-spectral data and crop phenology in assam, northeast india[J]. *Remote Sensing*, 2016,8(6):479.
- [13] Xiao X, Boles S, Liu J, et al. Mapping paddy rice agriculture in southern china using multi-temporal modis images [J]. *Remote Sensing of Environment*, 2005,95(4):480-492.
- [14] Zhang X, Zhang M, Zheng Y, et al. Crop mapping using proba- v time series data at the yucheng and hongxing farm in china[J]. *Remote Sensing*, 2016,8(11):915.
- [15] Choudhury I, Chakraborty M. Sar signature investigation of rice crop using radarsat data[J]. *International Journal of Remote Sensing*, 2006,27(3):519-534.
- [16] Kurosu T, Fujita M, Chiba K. Monitoring of rice crop growth from space using the ers-1 c-band sar[J]. *Geoscience & Remote Sensing IEEE Transactions on Geoscience and Remote Sensing*, 1995,33(4):1092-1096.
- [17] Chen J, Lin H, Pei Z. Application of envisat asar data in mapping rice crop growth in southern china[J]. *IEEE Geoscience & Remote Sensing Letters*, 2007,4(3):431-435.
- [18] Panigrahy S, Jain V, Patnaik C. Identification of rice crop in bangladesh using temporal c-band sar - a feasibility study[J]. *Journal of the Indian Society of Remote Sensing*, 2012,40(4):599-606.
- [19] Prucker S, Meier W, Stricker W. Evaluation of radarsat standard beam data for identification of potato and rice crops in india[J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 1999,54(4):254-262.
- [20] Torres R, Snoeij P, Geudtner D, et al. Gmes sentinel-1 mission[J]. *Remote Sensing of Environment*, 2012,120:9-24.
- [21] 吴炳方,许文波,孙明,等.高精度作物分布图制作[J].*遥感学报*,2004,8(6):688-695. [Wu B F, Xu W B, Sun M, et al. Mapping Crop Distribution with High Accuracy[J]. *Journal of Remote Sensing*, 2004,8(6):688-695.]
- [22] Thenkabail P S, Biradar C M, Tural H, et al. An irrigated area map of the world (1999) derived from remote sensing [M]. *Iwmi Books Reports*, 2006,36:600-605.
- [23] Mandanici E, Bitelli G. Preliminary comparison of sentinel-2 and landsat 8 imagery for a combined use[J]. *Remote Sensing*, 2016,8(12):1014.
- [24] Gorelick N, Hancher M, Dixon M, et al. Google earth engine: Planetary-scale geospatial analysis for everyone[J]. *Remote Sensing of Environment* 2017,202:18-27.
- [25] Hansen M C, Potapov P V, Moore R, et al. High-resolution global maps of 21st-century forest cover change[J]. *Science*, 2014,342(6160):850-853.
- [26] Pekel J F, Cottam A, Gorelick N, et al. High-resolution mapping of global surface water and its long-term changes[J]. *Nature*, 2016,540(7633):418-422.
- [27] 王斌,陈小敏,钟曼茜,等.海南水稻生育期的时空变化特征及对气候变暖的响应[J].*热带作物学报*,2017,38(3): 415-420. [Wang B, Chen X M, Zhong M X, et al. Spatial-temporal variation and response to climate warmer upon phenology of Hainan rice[J]. *Chinese Journal of Tropical Crops*, 2017,38(3):415-420.]
- [28] Li Q, Wu B, Xu W. Accuracy assessment of crop type proportion using GVG instrument on transect line[J]. *Journal of Remote Sensing*, 2004,6:253.
- [29] 吴炳方,田亦陈,李强子. GVG农情采样系统及其应用 [J].*遥感学报*,2004,8(6):570-580. [Wu B F, Tian Y C, Li Q Z. GVG a crop type proportion sampling instrument[J]. *Journal of Remote Sensing*, 2004,8(6):570-580.]
- [30] Chander G, Markham B L, Helder D L. Summary of current radiometric calibration coefficients for landsat mss, tm, etm+, and eo-1 ali sensors[J]. *Remote Sensing of Environment*, 2009,113(5):893-903.
- [31] Oreopoulos L, Wilson M J, Várnai T. Implementation on landsat data of a simple cloud-mask algorithm developed

- for modis land bands[J]. IEEE Geoscience & Remote Sensing Letters, 2011,8(4):597-601.
- [32] The European Space Agency. Sentinel-2 user handbook [R]. European Space Agency (ESA), 2015:64.
- [33] Farr T G, Kobrick M. The shuttle radar topography mission: A global dem[J]. American Geophysical Union, 1999,45:37-55.
- [34] Ouyang Z, Zheng H, Xiao Y, et al. Improvements in ecosystem services from investments in natural capital[J]. Science, 2016,352(6292):1455-1459.
- [35] Tan S, Wu B, Yan N, et al. An ndvi-based statistical et downscaling method[J]. Water, 2017,9(12):995.
- [36] Tan S, Wu B, Yan, N, et al. Satellite-based water consumption dynamics monitoring in an extremely arid area [J]. Remote Sensing, 2018,10(9):1399.
- [37] Huete A, Liu H, Batchily K, et al. A comparison of vegetation indices over a global set of tm images for eos-modis [J]. Remote Sensing of Environment, 1997,59(3):440-451.
- [38] Gao B C. In NdwI-a normalized difference water index for remote sensing of vegetation liquid water from space [J]. Imaging Spectrometry, 1995,58(3):257-266.
- [39] Breiman L. Bagging predictors machine learning[J]. Machine Learning 1996,24(2):123-140.
- [40] Ho T. The random subspace method for constructing decision forests[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1998,20(8):832-844.
- [41] Pelletier C, Valero S, Inglada J, et al. Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas[J]. Remote Sensing of Environment, 2016,187:156-168.
- [42] Sharma R, Tateishi R, Hara K, et al. Production of the japan 30-m land cover map of 2013-2015 using a random forests-based feature optimization approach[J]. Remote Sensing, 2016,8(5):429.
- [43] Pal, M. Random forest classifier for remote sensing classification[J]. International Journal of Remote Sensing, 2005,26(1):217-222.
- [44] Azzari G, Lobell D. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring[J]. Remote Sensing of Environment, 2017,202:64-74.
- [45] 朱良,平博,苏奋振,等.多时相tm影像决策树模型的水稻识别提取[J].地球信息科学学报,2013,15(3):446-451. [Zhu L, Ping B, Su F Z, et al. Mapping paddy rice based on multi-temporal TM images and decision tree model[J]. Journal of Geo-Information Science, 2013,15(3):446-451.]
- [46] 魏新彩,王新生,刘海,等. HJ 卫星图像水稻种植面积的识别分析[J].地球信息科学学报,2012,14(3):382-388. [Wei X C, Wang X S, Liu H, et al. Mapping paddy rice based on HJ images[J]. Journal of Geo-Information Science, 2012,14(3):382-388.]
- [47] 陈燕丽,莫伟华,莫建飞,等.基于面向对象分类的南方水稻种植面积提取方法[J].遥感技术与应用,2011,26(2): 163-168. [Chen Y L, Mo W H, Mo J F, et al. Mapping paddy rice in Southern China based on object-oriented method[J]. Remote Sensing Technology and Application, 2011,26(2):163-168.]
- [48] 吴金胜,刘红利,张锦水.无人机遥感影像面向对象分类方法估算市域水稻面积[J].农业工程学报,2018,34(1): 70-77. [Wu J S, Liu H L, Zhang J S. Mapping paddy rice in city based on UAV images and object-oriented method [J]. Transactions of the Chinese Society of Agricultural Engyneering, 2018,34(1):70-77.]