

引用格式: Zhu Shuang, Zhang JinShui. Influence Factors Analysis on Accuracies of Winter Wheat Distribution from Low and Medium Resolution Compositied Remote Sensing Images[J]. Remote Sensing Technology and Application, 2022, 37(3): 608-619. [朱爽, 张锦水. 中低分冬小麦分布提取模型效率的样本特征分析[J]. 遥感技术与应用, 2022, 37(3): 608-619.]

DOI: 10.11873/j.issn.1004-0323.2022.3.0608

## 中低分冬小麦分布提取模型效率的样本特征分析

朱爽<sup>1,2</sup>, 张锦水<sup>2,3</sup>

(1. 北京工业职业技术学院, 北京 100042;

2. 北京市陆表遥感数据产品工程技术研究中心, 北京 100875;

3. 北京师范大学地理科学学部遥感科学与工程研究院, 北京 100875)

**摘要:** 中空间分辨率样本(简称中分样本)的数量、质量是决定中低分辨率复合识别模型效率的关键因素。以冬小麦为研究对象, 中低分辨率影像结合构建支撑向量回归模型(Support Vector Regression, SVR)实现冬小麦的混合像元分解, 提取出冬小麦的空间分布, 定量分析中分样本数量、质量对识别精度的影响。结果表明: 从样本数量上看, 样本量为10%即可保证稳定的冬小麦精度, 在典型冬小麦区的区域精度、像元精度达到98%、92%以上; 从样本质量上看, 识别精度随样本质量(达到60%即可获得较好的识别结果)增加而升高; 对于非中分样本区的冬小麦, 区域精度、像元精度也是随样本数量的增加而提高, 在20%样本量下, 区域精度和像元精度稳定在97%、92%以上, 表明该模型具有较强的空间泛化性能力, 弥补了从低分影像上难以获取有效样本的不足。

**关键词:** 支撑向量回归; 混合像元分解; 样本数量/质量; TM; MODIS

**中图分类号:** S127 **文献标志码:** A **文章编号:** 1004-0323(2022)03-0608-12

### 1 引言

准确的农作物空间分布信息是作物生产管理、估产的基础。遥感是进行大范围农作物识别最有效技术手段, 一般通过传感器接收地表要素反射的太阳电磁波成像, 利用模式识别将影像像元归属为特定的地物类型<sup>[1-2]</sup>。已有研究表明, 通过中高分辨率遥感影像可获得较为纯净的地表要素光谱信息, 确定地物光谱特征, 如能获得关键时期遥感影像就可准确地提取出作物的空间分布<sup>[3]</sup>。但由于受到卫星回访周期长和云雨天气的影响, 无法保证在作物生长关键时期获得有效的较高分辨率影像<sup>[4]</sup>。这样, 不但会造成同期作物识别的干扰, 影响识别精度<sup>[5]</sup>, 也无法保证遥感影像的全覆盖能力, 这限制了利用中高分辨率尺度遥感影像进行作物分布提取的应

用能力。

低空间分辨率遥感影像(简称低分影像), 如MODIS和NOAA, 回访周期短, 能在1d内对地球表面进行观测, 并可在一定时间窗口上进行数据合成, 消除云和噪声的影响, 得到无云的遥感影像<sup>[6-8]</sup>。时间序列影像能够定量地刻画出农作物在整个生长周期的物候特征<sup>[9-10]</sup>, 利用每一种作物在时间尺度上的特有时间序列特征来构建识别模型, 提取出作物的分布<sup>[11-12]</sup>。相对于中高分辨率影像而言, 具有丰富时间序列特征的低分影像能够一定程度上消除“异物同谱、同物异谱”的困扰, 也能够解决作物生长关键期遥感影像缺失的问题<sup>[13]</sup>, 但由于空间分辨率低, 混合像元现象严重<sup>[14]</sup>, 制约了作物分布的识别精度。一般采用混合像元分解来提取出作

收稿日期: 2021-03-08; 修订日期: 2022-05-25

基金项目: 高分辨率对地观测系统重大专项(20-Y30F10-9001-20/22), 北京工业职业技术学院重点课题(BGY2022 KY-14Z)。

作者简介: 朱爽(1981-), 女, 甘肃金昌人, 博士, 副教授, 主要从事资源与环境遥感方面的研究。E-mail: zhushuang@mail.bnu.edu.cn

通讯作者: 张锦水(1978-), 男, 河北沧州人, 博士, 教授, 主要从事资源与环境遥感、人工智能方面的研究。E-mail: zhangjs@bnu.edu.cn

物的组分,可以在一定程度上消除混合像元的影响<sup>[3,11,15]</sup>。线性混合像元分解模型是常用的方法<sup>[12-13]</sup>,其中端元选择的优劣是决定模型性能的关键<sup>[16]</sup>,一般可以采用MNF变换从图像上选择出纯净的端元<sup>[17]</sup>,但由于我国农业种植地区地块破碎,很难在低分影像(如MODIS影像为250 m分辨率)上选择出纯净的端元特征,这影响了利用时间序列的低分影像进行作物提取的能力<sup>[18]</sup>。

综上所述,中低空间分辨率遥感影像(简称中低分影像)在空间、时间分辨率上具备各自的优势,综合各自的优点建立中低分影像的作物识别模型,是解决中等分辨率遥感影像(简称中分影像)获取困难、低分影像提取精度难以保证的有效途径。2012年Pan等<sup>[19]</sup>提出基于MODIS时间序列筛选出特定地物物候特征,并结合中分影像作为目标量,构建冬小麦作物面积指数模型,一定程度上消除了混合像元对作物提取的影响。另外,利用遥感和作物物候模型结合,可减小因作物播种日期差异带来的不确定性,提高作物识别的精度<sup>[20]</sup>。

中低分辨率影像识别方法能够有效地进行作物识别,但中分影像对识别模型的影响仍需深入探讨。本研究以冬小麦为研究对象,利用从中分影像提取的冬小麦分布作为模型的样本目标量,结合时间序列数据为自变量,构建支撑向量回归模型(Support Vector Regression, SVR)实现混合像元分解,提取出冬小麦的空间分布,分析中分影像作为样本,其数量、质量对模型的影响,进而验证中低分作物识别模型的适用能力。

## 2 中低分影像复合的冬小麦提取模型

研究的目的是结合中、低分影像各自的优势,利用中分影像构建样本,训练支撑向量回归模型,构建非线性模型,实现低分时间序列影像的混合像元分解,进行冬小麦分布的识别。

对于一个低分像元,如MODIS,该像元内所包含的冬小麦面积( $S_{\text{MODIS}}$ )为:

$$S_{\text{MODIS}} = S_{\text{MODIS}_0} \cdot P_{\text{MODIS}} \quad (1)$$

其中: $P_{\text{MODIS}}$ 表示某个MODIS像元内冬小麦所占比例; $S_{\text{MODIS}_0}$ 为MODIS像元的面积(即250 m×250 m)。

首先,利用线性混合像元模型来解释像元的光谱特征与组分之间的关系:

$$f = \sum f_m s_m + \varepsilon \quad (2)$$

其中: $f$ 为混合像元的反射率; $f_m$ 为 $m$ 地物类型组分的反射率; $s_m$ 为第 $m$ 种地物类型组分所占的比例; $\varepsilon$ 为残差。这个模型阐释了目标地物所占比例与混合光谱与目标地物光谱特征相关,组分比例越高则混合光谱中呈现出该地物的光谱特征越明显。

但是混合像元光谱特征一般不是简单的线性组合,可以采用非线性模型进行求解。SVR为设定最多有 $2\varepsilon$ 偏差的隔离带,能够保证即使在小样本量的情况下也能得到较好的识别结果<sup>[21]</sup>。研究构建SVR模型进行冬小麦识别,其中因变量为混合像元的丰度,自变量为MODIS的时间序列特征。

为利用SVR混合像元分解进行冬小麦分布的提取,提出两个假设:

假设1:相对于低分影像而言,中分影像(以Landsat5 TM为例)提取出的冬小麦精度高,可以聚合到低分辨率尺度作为SVR回归的目标量。中分冬小麦结果聚合到低分尺度的规则如下:

$$P_m = \frac{n \cdot S_{\text{TM}}}{S_M} \quad (3)$$

其中: $n$ 为MODIS像元所对应TM的冬小麦像元数; $S_{\text{TM}}$ 、 $S_M$ 分别为TM与MODIS的像元分辨率。

假设2:以MODIS像元中的冬小麦面积丰度为目标量( $K$ ),其提取精度与样本量( $n$ )、样本质量( $q$ )密切相关:

$$K = f(n, q) \quad (4)$$

## 3 实验设计

用中低分影像融合的SVR模型来提取冬小麦的流程主要包括:基于TM数据和MODIS时间序列数据分别提取各自冬小麦的分布范围;SVR模型训练样本构建;TM和MODIS共区、非共区区域划分;对冬小麦识别结果进行精度评价。

### 3.1 研究区与数据

#### 3.1.1 研究区概况

研究区位于北京市东南部,跨越北京、河北和天津,覆盖范围为 $116^{\circ}5'13'' \sim 117^{\circ}15'16''$  E,  $39^{\circ}25'25'' \sim 40^{\circ}28'26''$  N,见图2。属于中纬度暖温带,是典型的温带大陆性季风气候。该地区北部为燕山山脉,大部区域为平原,属于黄淮海平原北部地区,北京市辖区位于西部。作为中国典型的农业区,该区域主要以冬小麦和玉米轮作方式为主。

#### 3.1.2 数据与预处理

实验利用两种空间分辨率的遥感影像:Landsat 5 TM和MODIS。用TM数据提取出的冬小麦

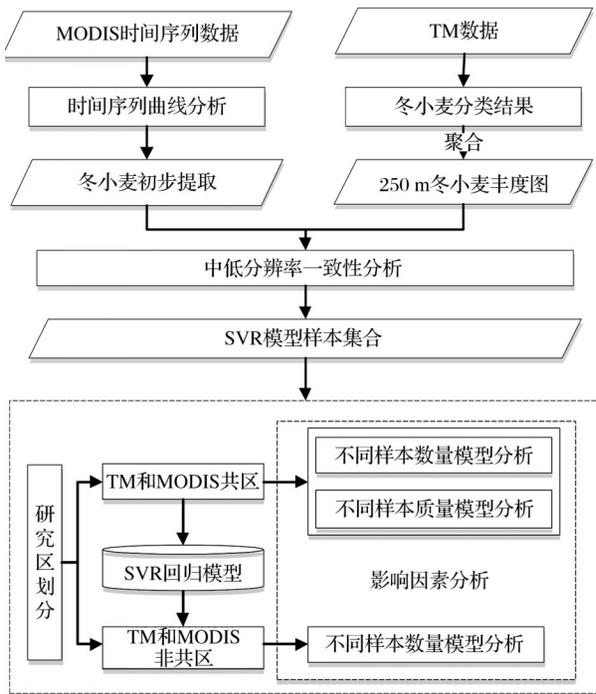


图1 中/低分影像冬小麦遥感测量流程图

Fig.1 Technological flow chart

分布结果作为目标量,然后结合MODIS数据的时间序列特征进行建模。

Landsat 5 TM 数据采集时间为2006年4月7日,为图幅123/32的一个子区范围,图像质量完好,无云,所用波段包括蓝光(0.433~0.515 μm),绿光(0.525~0.600 μm),红光(0.630~0.680 μm)、近红外(0.845~0.885 μm)、短波红外1(1.560~1.660 μm)和短波红外2(2.100~2.300 μm)6个波段。从美国

地质调查局USGS网站上下载的TM影像已经进行了几何精校正,本研究仅使用单期TM数据进行冬小麦提取,不再需要进行大气校正。TM的投影类型设定为Albers等积投影,第一条纬线为北纬25°,第二条纬线为北纬47°,中央经线为东经105°,椭球体为WGS84。通过判读图像可以看出,该地区地物类型主要包括冬小麦、裸地、林地、城镇和人工草地等,农业景观呈现规模连片和破碎分布并存的现象。相对于MDOIS数据(分辨率为250 m),TM影像分辨率(30 m)较高,能够有效地反映出冬小麦的分布,且属于冬小麦拔节期,同期作物少,可作为精确识别冬小麦的数据源。

研究采用时间序列MODIS 16 d合成的归一化植被指数(Normalized Difference Vegetation Index, NDVI)产品,时间跨度为2015年10月上旬至2016年6月下旬,覆盖了冬小麦的整个生长周期,利用SG滤波消除云和噪声的影响。投影与坐标系与TM保持一致。图3看出10月下旬、中旬和上旬合成显示,小麦被清晰显示出来,呈粉红色,这主要是10月下旬冬小麦已经出苗,冠层覆盖地面,而其他植被,如林地,正好处于叶绿素减少,叶子脱落,呈裸地状态,造成二者剧烈的时序遥感特征的反差。

### 3.2 一致样本区的提取

为有效支撑SVR进行冬小麦提取,需要精确的冬小麦分布。在此引用冬小麦“一致性区域”概念,提取本研究中的冬小麦分布作为构建SVR的样本<sup>[22]</sup>。基本流程为,通过阈值设定从MODIS上提

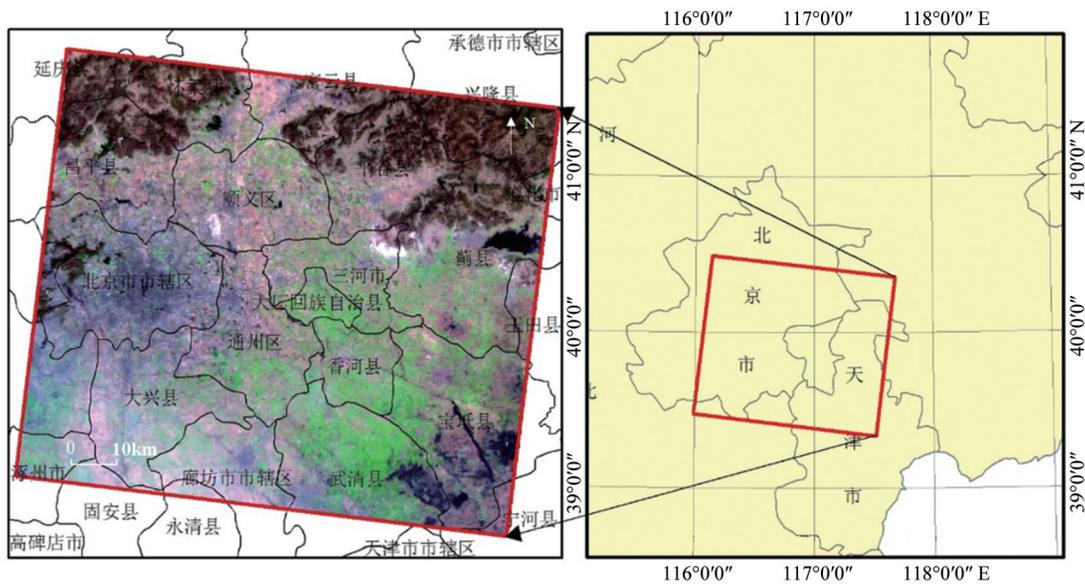


图2 研究区

Fig.2 Study area

取出冬小麦的分布,利用机器学习的方法从TM影像上提取出冬小麦的分布,二者求取交集,得到冬小麦测量结果一致性区域,这样一定程度上消除不同传感器识别冬小麦的误差。定义模式如下:

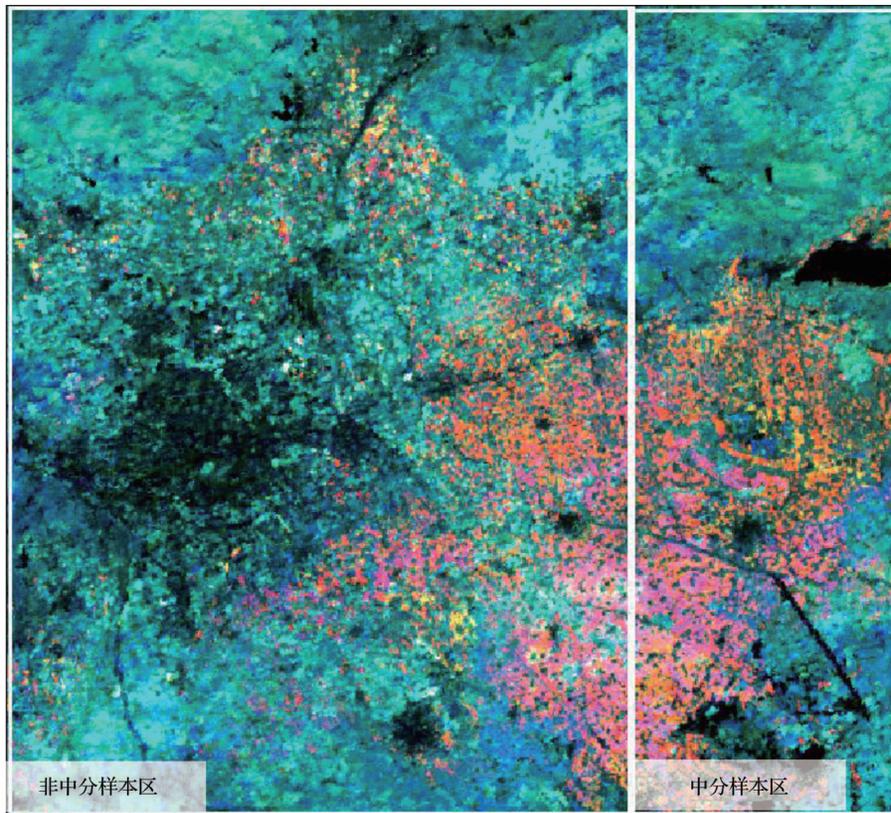
$$C = S_{\text{MODIS}} \cap F_{\text{TM}} \quad (5)$$

其中: $S_{\text{MODIS}}$ 为利用MODIS影像提取的冬小麦丰度值, $F_{\text{TM}}$ 为利用TM影像提取出的冬小麦分布范围。当 $S_{\text{MODIS}}=1$ 且 $F_{\text{TM}} \geq 50\%$ 的像元区为典型冬小麦一致区, $C$ 取值为1;当 $S_{\text{MODIS}}=0$ 且 $F_{\text{TM}} < 50\%$ 的区域为典型非冬小麦一致区域, $C$ 取值为0。具体流程参见文献[22]。

### 3.3 试验区的划分

样本的代表性是SVR构建模型提取冬小麦分

布的关键因素之一。从两个角度设计实验,构建SVR模型:一是从区域自身提取样本来训练SVR模型,该方法是常用的机器学习方法<sup>[23-25]</sup>,保证了样本的代表性,定义该区为I区;二是利用已知区域提取样本构建的模型来直接推断另一个区域,即预测区域与样本区域完全独立,定义II区,目的就是验证SVR提取冬小麦的模型泛化能力,在没有任何样本下提取冬小麦的适用性。根据上述确定的规则,将整个研究区划分为两个部分:中空间分辨率影像样本区(简称中分样本区,I区)和非中空间分辨率影像样本区(简称非中分样本区,II区),分别对应于有样本支持和没有样本支持情形下的冬小麦提取,如图3所示。



(构成波段时相组合为:3:2:1)

图3 MODIS影像示意图

Fig.3 Sketch map of MODIS image

### 3.4 SVR训练样本的选择

采用分层等比例抽样的策略,分层指标为 $250\text{ m} \times 250\text{ m}$ 像元尺度内的冬小麦比例,以从“一致样本库”中有效提取出样本为目的。为验证样本数量对回归提取冬小麦精度的影响,在每一层按照抽样比例为0.1%、0.5%、0.8%、5%、10%、20%、30%、40%、50%、60%、70%、80%和90%的数量进行样

本的选择,分析当样本量为多大规模时,冬小麦的识别精度就能够达到比较高且稳定的程度。

另外,考虑到用于回归的样本精度(即样本质量)会对回归结果产生一定的影响,因此在一致样本区混入了一定比例的误差,以模拟样本混入误差对SVR冬小麦提取结果的影响。在一致区内的像元样本内混入10%、20%、30%、40%、50%和60%

的随机误差(TM的分类误差不能低于40%),也就是在一致区内选择出相应比例的样本量,将冬小麦丰度值 $f$ 取反(即 $1-f$ )。

### 3.5 精度评价指标

将上述方法提取出的样本,输入到LibSVM中采用epsilon-SVR进行回归分析,构建SVR的冬小麦提取模型。本文选用了像元精度、区域精度来定量分析不同尺度下冬小麦的提取精度。

像元精度(Pixel Accuracy, PA),代表区域内像元级别的冬小麦识别精度。

$$PA = \frac{1}{N} \sum \left( 1 - \frac{|\hat{w}_i - w_i|}{w_i} \right) \quad (6)$$

其中: $\hat{w}_i$ 代表第 $i$ 个像元内SVR提取的冬小麦丰度; $w_i$ 代表第 $i$ 个像元对应的冬小麦像元丰度; $N$ 代表用于精度评价的像元个数。

区域精度(Region Accuracy, RA),代表区域内用SVR提取冬小麦的面积精度,反应出区域数量精度。

$$RA = 1 - \frac{|\hat{W} - W|}{W} \quad (7)$$

其中: $\hat{W}$ 为SVR提取出的冬小麦总面积; $W$ 为中分冬小麦总面积。

## 4 结果分析

### 4.1 样本数量对SVR识别精度的影响

按照0.1%、0.5%、0.8%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%,进行5次抽样来训练SVR模型,得到I区内的冬小麦结果(见图4)。从识别结果可以看出,当样本量为0.1%时,非冬小麦区域明显有误差混入的情况;随着训练样本量的增加,误差混入现象逐渐消失,冬小麦被准确地识别出来。

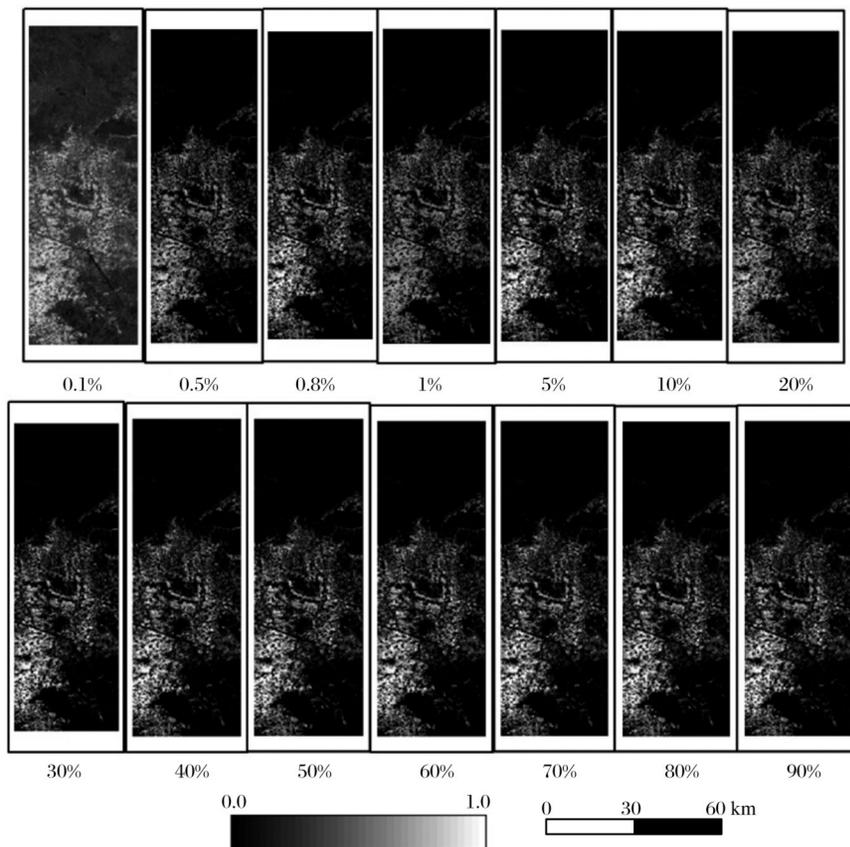


图4 不同样本下MODIS混合像元冬小麦测量结果丰度图

Fig.4 Wheat fractional map generated from MODIS mixed pixel unmixing with various amount of sample data

为全面分析不同情况下SVR冬小麦识别精度,将一致性区域分为两个部分:典型冬小麦区和典型非冬小麦区。

下面分别对全区、典型冬小麦区和典型非冬小麦区进行分析。

#### 4.1.1 全区精度评价

从图5可以看出,随着样本量的增加,MODIS提取出的冬小麦区域精度、像元精度的最大值、最小值和平均值都不断提高。RA和PA的最大值和最小值之间的差值逐渐缩小,呈收缩喇叭口状。在

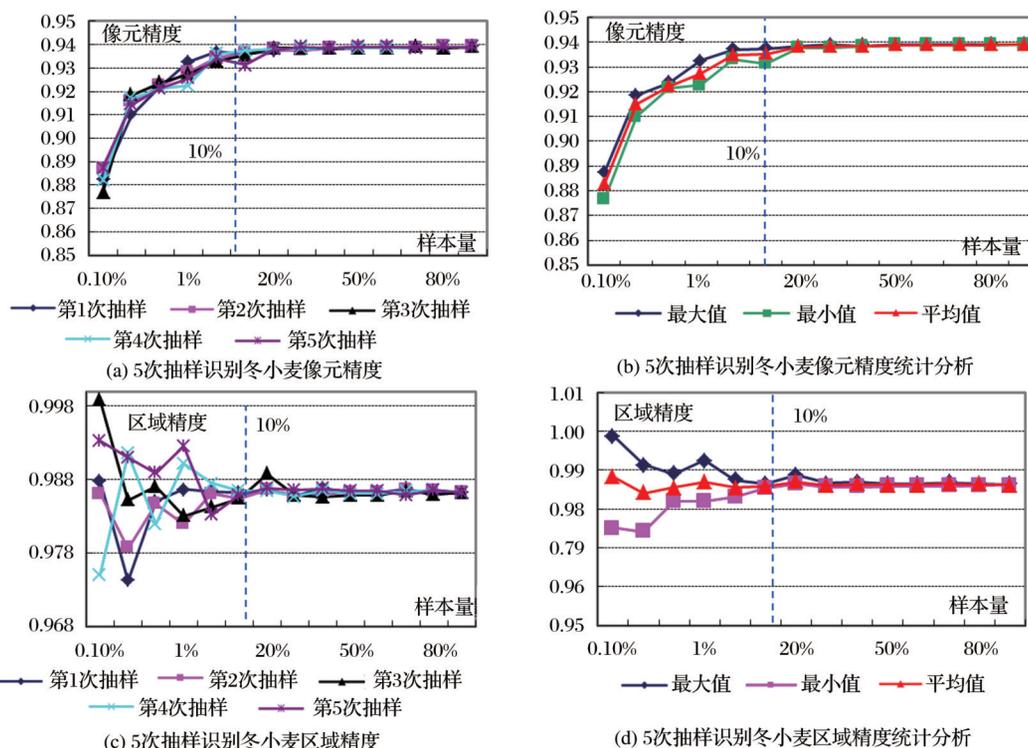


图 5 全区不同样本量下 SVR 提取出的冬小麦精度

Fig.5 Wheat accuracy based on SVR with various amount of sample data

样本量达到 10% 时,像元精度和区域精度达到最高,趋于平稳状态。可知在一致样本区内提取 10% 的样本量就可以满足 SVR 混合像元分解的要求,能够保证区域精度和像元精度的较高一致性(PA 达到 92% 以上,RA 达到 98% 以上)。

4.1.2 典型冬小麦区域精度评价

典型冬小麦区的区域精度、像元精度的结果如图 6 所示。与全区区域精度、像元精度的规律相似,区域精度、像元精度随着样本量的增加而提高。当样本量达到 10% 时,两个精度指标已经比较稳定,像元精度达到 92% 以上,区域精度达到 99% 以上。可见在典型冬小麦区 SVR 能够准确提取出冬小麦分布。

4.1.3 典型非冬小麦精度评价

典型非冬小麦区域的冬小麦面积为 0,因此该区域只需要用像元精度一个指标即可,计算方式是将 SVR 所提取的非冬小麦丰度代入式(6)。图 7 表明,随着抽样数量的增加,非冬小麦区像元精度逐渐提高。当样本量超过 10% 时,像元精度趋于稳定,不再发生变化,只有少许的冬小麦混入到非冬小麦像元。

上述研究表明,当样本量达到 10% 以后,PA 和 OA 均比较高且稳定,因此选择某一次 10% 样本量下的 SVR 冬小麦提取结果进一步分析。从 MODIS 和 TM 冬小麦、非冬小麦的一致区分层随机选择

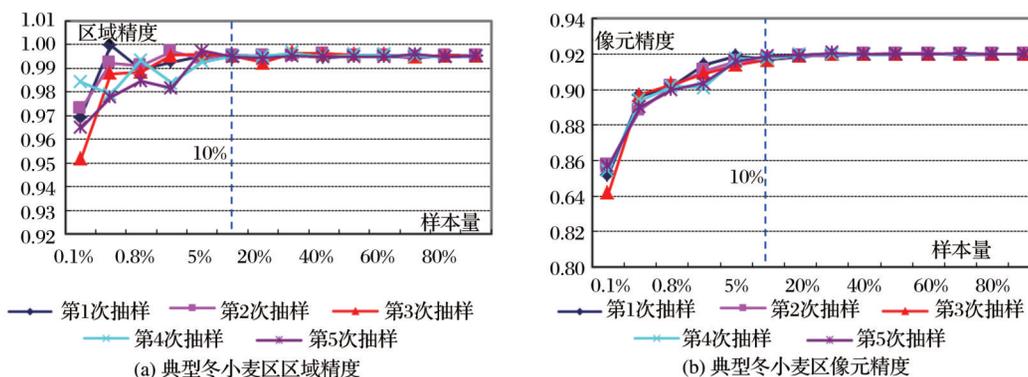


图 6 典型冬小麦区区域精度和像元精度

Fig.6 Regional and pixel accuracy within representative wheat areas

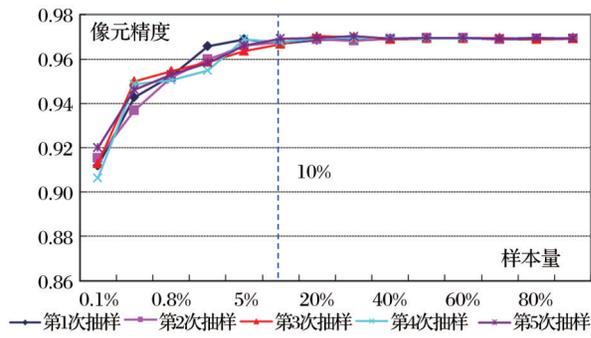


图7 非冬小麦区的像元精度

Fig.7 Pixel accuracy with non-wheat area

10%的样本,制作散点图、误差直方图和累积误差直方图,分析估计结果和检验样本之间的关系。

图8展示的是5次10%样本量抽样下识别结果

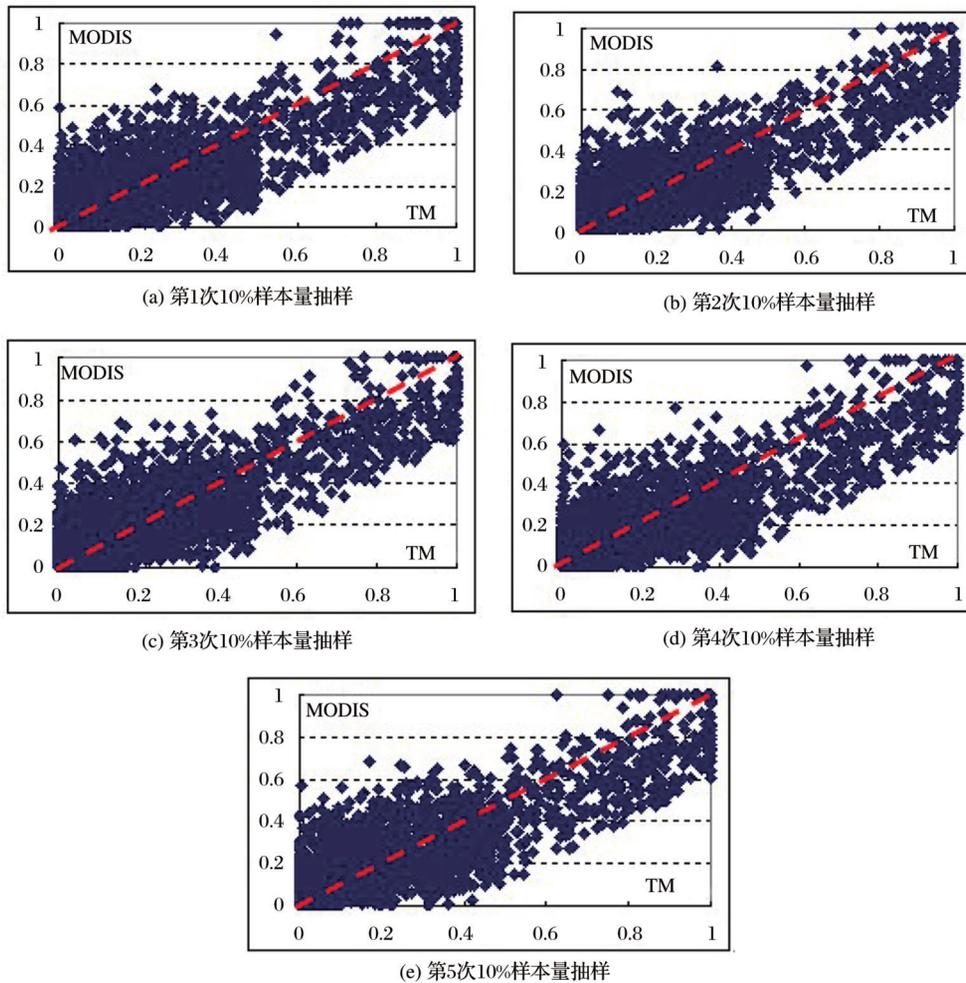


图8 5次10%样本量抽样下TM与MODIS冬小麦测量结果的散点图

Fig.8 The scatter dot diagram calculated between TM and MODIS with 10% sample

#### 4.2 样本质量对SVR提取冬小麦的影响

一致区内的样本精度直接影响SVR模型提取冬小麦的能力。下面将一致区内混入一定的误差,分析样本质量对结果的影响。对一致区内的像元

和检验样本之间的关系。从这5次结果可以看出,SVR识别冬小麦结果与检验样本的结果非常接近,集中分布在1:1线两侧,确保了区域面积精度的准确性。

将MODIS冬小麦的识别结果与TM结果作差,并将差值结果进行分段 $[-1, -0.9], [-0.9, -0.8], \dots, [0.9, 1.0]$ 统计差值个数,制作差值直方图,如图9所示。差值结果多集中在 $[-0.1, 0.1]$ 范围之内,说明SVR能够准确地提取出冬小麦结果。图10为差值绝对值的累积直方图,差值绝对值在 $[0, 0.1]$ 之间达到80%左右,说明SVR混合像元分解得到的冬小麦结果与检验值接近,验证了SVR提取冬小麦的能力。

混入一定的误差后,分层随机5次抽选出10%样本量构建SVR冬小麦提取模型。

图11是某一次冬小麦的提取结果。当混入10%~40%误差量时,对SVR分解提取的冬小麦分

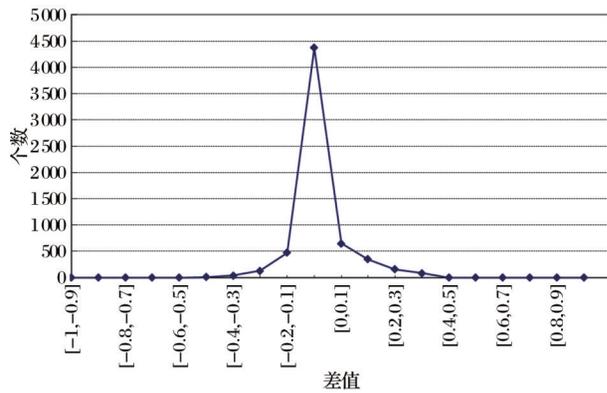


图9 10%样本量下5次结果的误差直方图

Fig.9 Error histograms of five experiments with 10% sample

布影响不大,冬小麦的分布比较一致且稳定;当误差量达到50%时,山区的绿色植被、水体等被误分

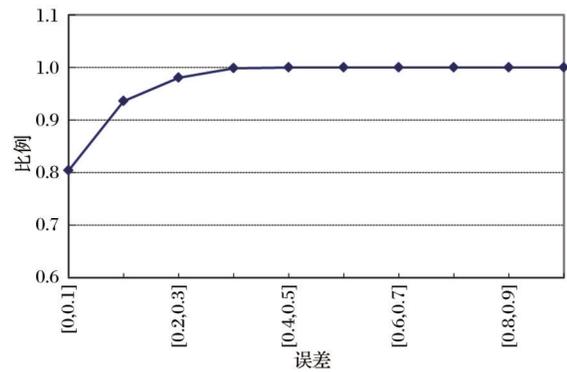


图10 10%样本量下5次结果的误差累积直方图

Fig.10 Accumulated error histograms of five experiments with 10% sample

解成冬小麦。可见,输入样本的质量对SVR分解模型有一定影响。

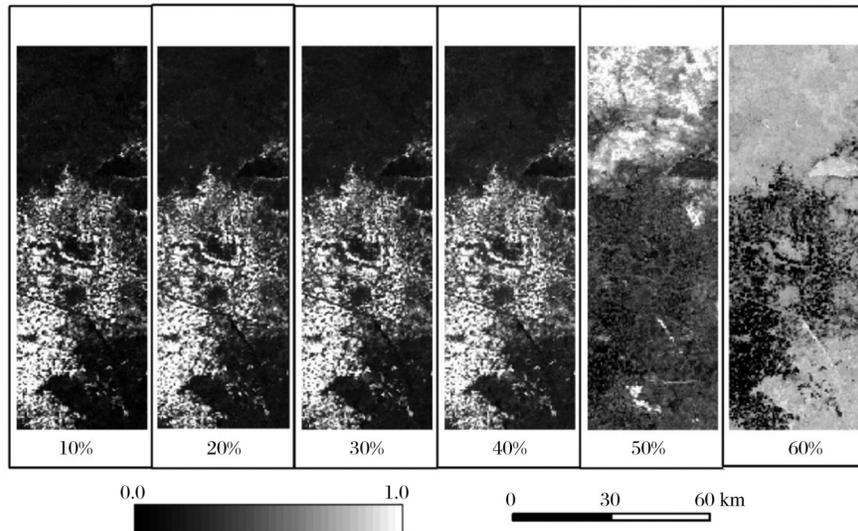


图11 不同误差量下MODIS冬小麦混合结果

Fig.11 MODIS based wheat results with different error fraction

从图12可以看出,RA、PA随着误差量的增加而降低。当误差控制在10%~40%范围内,精度降低较慢,区域精度和像元精度分别在89%、86%以上。当误差量超过50%,区域精度、像元精度发生陡变,精度明显下降,原因是混入误差过多,识别过程将非冬小麦像元分解成冬小麦像元,而冬小麦像元被分解成非冬小麦像元。TM提取冬小麦精度相比MODIS而言能够保证高于50%,且TM提取样本精度越高则SVR提取冬小麦的精度越高。

### 4.3 SVR识别冬小麦的泛化能力分析

大区域范围内采用冬小麦信息提取时,经常会出现自身区域缺乏样本来支撑训练SVR模型的情况,因此利用其他区域已形成的SVR模型来推断该区域冬小麦分布是一种常用的策略。分别利用中

分样本区0.1%、0.5%、0.8%、1%、5%、10%、20%、30%、40%、50%、60%、70%、80%以及90%抽样样本量下形成的SVR模型对非中分样本区进行冬小麦提取。

图13是非TM样本区的冬小麦提取结果。目视来看,0.1%的样本量下,山区和城区误差混入情况严重;当样本量为0.5%和0.8%时,冬小麦的分布比较明显,但山区的误差混入情况仍比较严重;当样本量高于1%时,冬小麦的分布较为准确且稳定。

图14表明,非样本区(II区)混合像元分解精度随着输入样本量的增加而上升。在样本量为[0.1%,20%]范围内,RA和PA一直呈上升趋势。在II区,当样本量达到20%后,RA和PA趋于稳定,

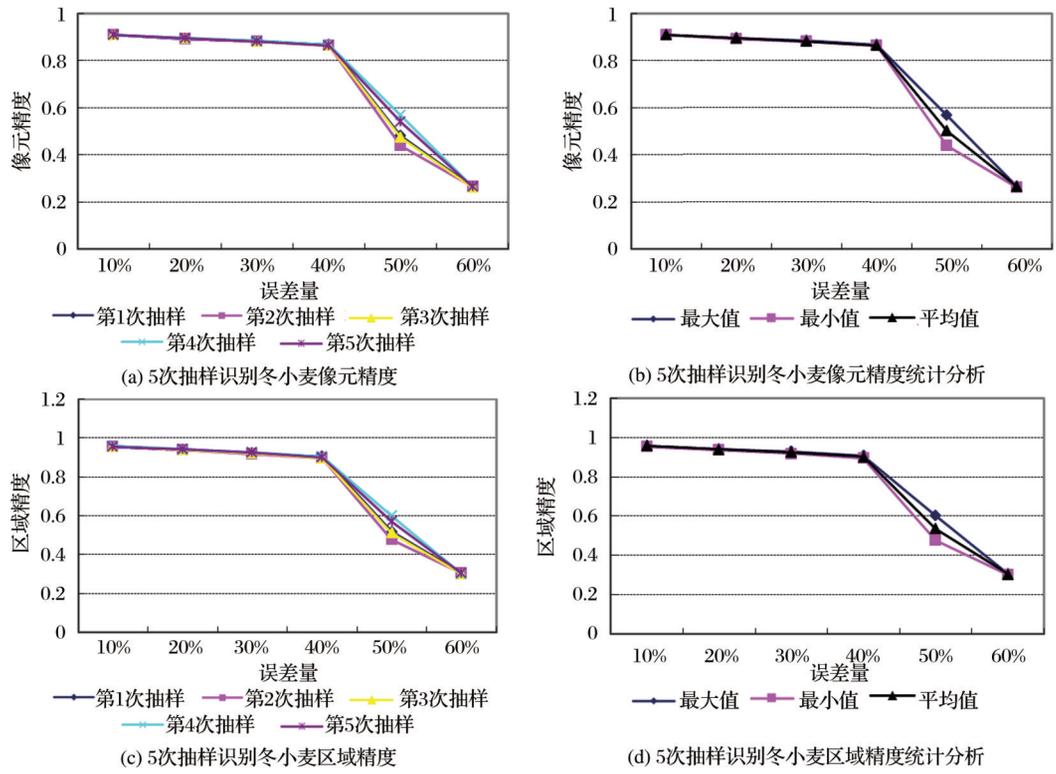


图 12 样本混入不同误差量下的 PA 和 RA

Fig.12 PA and RA with different error fraction mixture

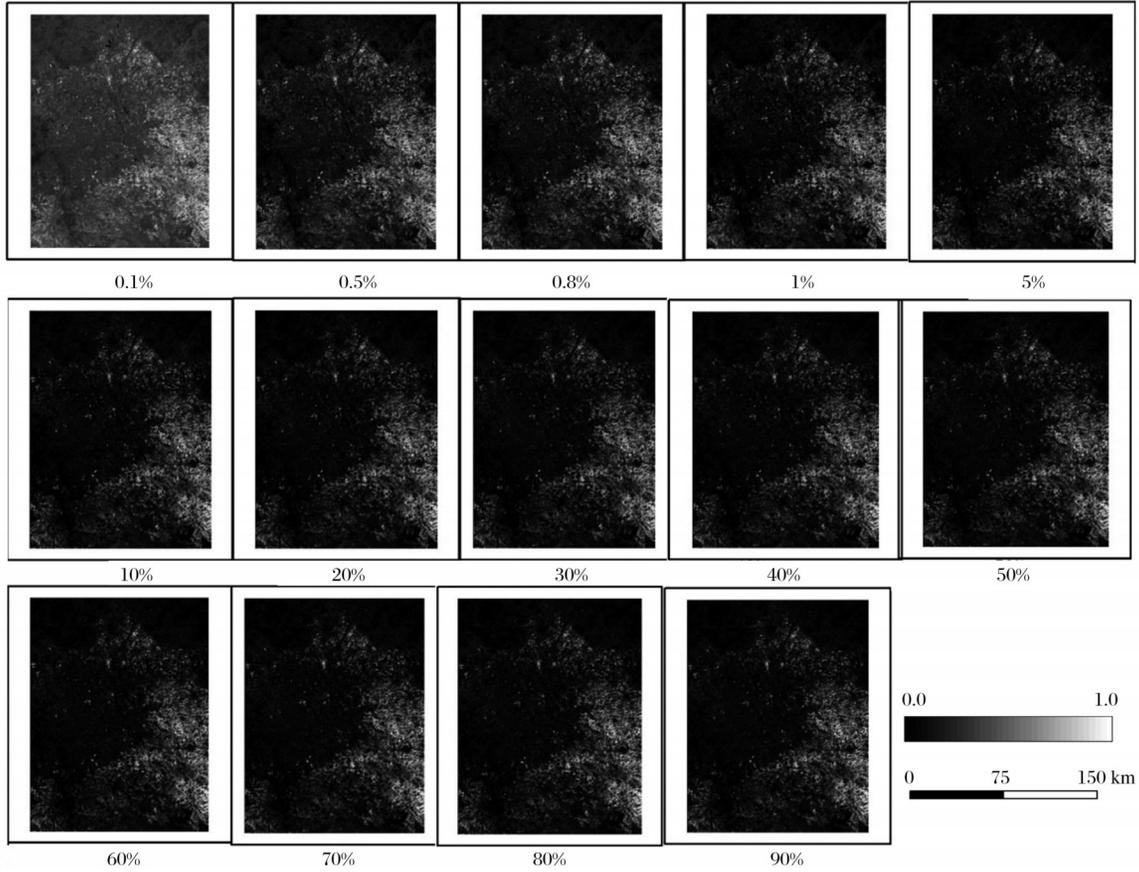


图 13 非 TM 样本区 MODIS 冬小麦提取结果

Fig.13 MODIS based wheat extent within TM sample area

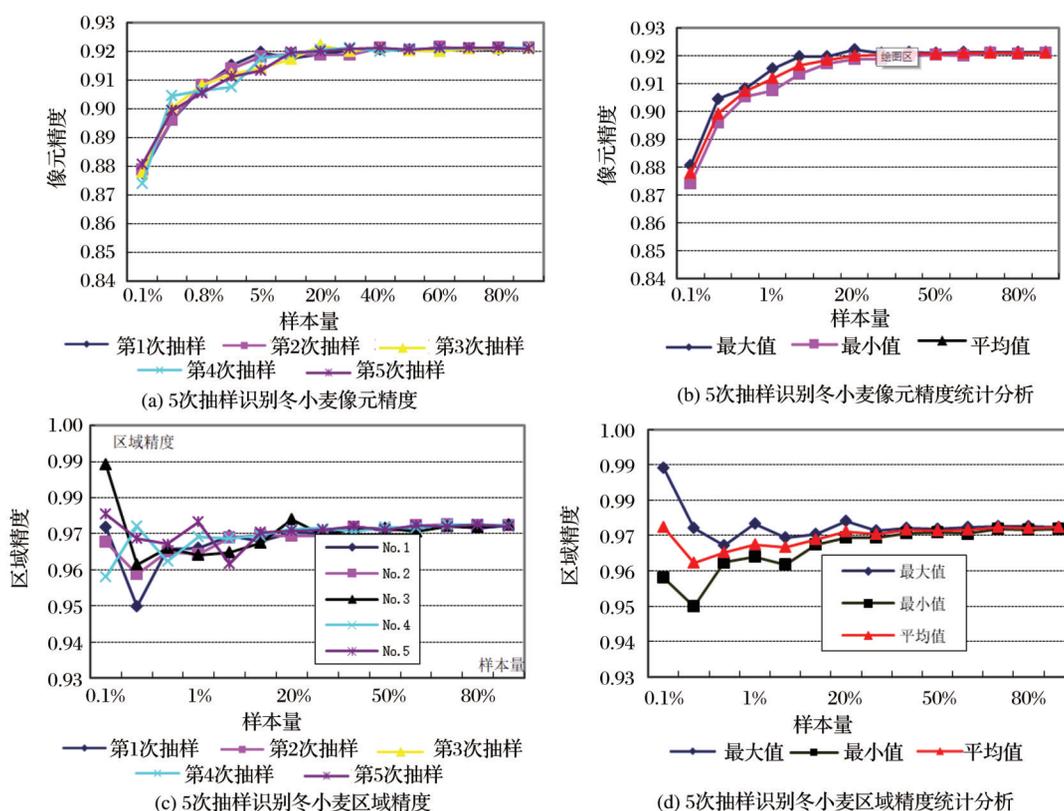


图 14 非 TM 样本区 SVR 冬小麦识别精度随样本量的变化

Fig.14 The change of SVR based wheat accuracy accompanying with sample increment

分别在 97%、92% 以上,样本量要高于 I 区 SVR 提取冬小麦达到比较高精度所需样本量,究其原因从自身区域选取样本的代表性会更强。

### 5 结 语

本文从样本数量和质量两个角度,分析了对中低分辨率影像复合提取冬小麦识别效率的影响,分别在中分样本区和非中分样本区两个区域进行验证,得到如下结论:

(1)在中分样本区内,冬小麦提取的区域精度和像元精度随着用于 SVR 模型训练的样本量逐步增多而提高。当样本量超过 10% 时,全区区域精度和像元精度稳定在 98%、93% 以上,典型冬小麦区的区域精度和像元精度达到 98%、92% 以上,非典型冬小麦区的区域精度和像元精度接近且都达到 96% 以上。说明 SVR 能够准确地提取出冬小麦的分布,错入、错出现象较少。因此,10% 的样本量就可以进行 MODIS 冬小麦的混合像元分解,准确地提取出冬小麦面积。

(2)SVR 冬小麦提取结果的精度与样本质量之间具有相关性。样本质量越高,冬小麦提取结果的精度就会越高。在 5 次 10% 抽样样本的情况下,当

混入误差在 10%~40% 时,提取的冬小麦区域精度、像元精度平均值分别在 89%、86% 以上;当样本混入误差为 50%~60% 时,区域精度和像元精度陡然降低,都低于 50%。可见,SVR 分解对于样本的准确性有一定要求,当输入样本精度高于 60% 时可以获得较好的识别结果。

(3)基于中分样本区训练的 SVR 模型可以成功迁移到非中分样本区进行冬小麦的提取,具有较强的空间泛化性能力。与中分样本区的实验相同,非中分样本区提取的冬小麦区域精度、像元精度随着输入样本数量的增加而提高,在 20% 样本量下,区域精度和像元精度达到稳定,且精度分别在 97% 和 92% 以上。

在一定样本数量、质量的前提下,利用 SVR 方法能够有效地提取出冬小麦的分布。这种方法弥补了从低分影像上难以获取有效样本的不足,验证了将模型迁移进行冬小麦提取的有效性。在后续的研究中,需要深入探讨 SVR 模型参数对提取结果的影响并验证该方法在更大范围和对其他作物提取的适用性。另外,模型迁移学习需要设置迁移网络层、学习效率和方式,还需要注意不同农业景观对迁移学习的影响,这些问题都有待深入研究。

## 参考文献(References):

- [1] Lobell D B, Asner G P. Cropland distributions from temporal unmixing of MODIS data [J]. *Remote Sensing of Environment*, 2004, 93(3): 412-422. DOI: 10.3390/rs4102890.
- [2] Pan Y, Hu T, Zhu X, *et al.* Mapping cropland distributions using a hard and soft classification model [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(11): 4301-4312. DOI:10.1109/TGRS.2012.2193403.
- [3] Zhu Shuang, Zhang Jinshui, Li Changqing, *et al.* Review and prospect of land cover mapping by remote sensing [J]. *Bulletin of Surveying and Mapping*, 2019(11): 8-11. [朱爽, 张锦水, 李长青, 等. 土地覆盖遥感制图方法研究现状与展望 [J]. *测绘通报*, 2019(11): 8-11.]
- [4] Li Lingling, Pan Yaozhong, Zhang Jinshui, *et al.* Method of winter wheat planting area estimation based on support vector machine and post-classification changed vector analysis [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2010, 26(9): 210-217. [李玲玲, 潘耀忠, 张锦水, 等. 支持向量机与分类后验概率空间变化向量分析法相结合的冬小麦种植面积测量方法 [J]. *农业工程学报*, 2010, 26(9): 210-217.]
- [5] Zhao Yingshi. Principle and method of remote sensing application analysis [M]. Beijing: Science Press, 2003. [赵英时. 遥感应用分析原理与方法 [M]. 北京: 科学出版社, 2003.]
- [6] Gu Xiaohe, Pan Yaozhong, Zhu Xiufang, *et al.* Consistency study between MODIS and TM on winter wheat plant area monitoring a case in small area [J]. *Journal of Remote Sensing*, 2007, 11(3): 350-358. [顾晓鹤, 潘耀忠, 朱秀芳, 等. MODIS与TM冬小麦种植面积遥感测量一致性研究——小区域实验研究 [J]. *遥感学报*, 2007, 11(3): 350-358.]
- [7] Wang P X, Xun L, Li L, Wang L, *et al.* Extraction of planting areas of main crops based on sparse representation of time-series leaf area index [J]. *Journal of Remote Sensing*, 2019, 23(5): 959-970. [王鹏新, 荀兰, 李俐, 等. 基于时间序列叶面积指数稀疏表示的作物种植区域提取 [J]. *遥感学报*, 2019, 23(5): 959-970.]
- [8] Han Lidong, Lin Xiangguo, Liang Yong. An automatic cloud detection method for high resolution satellite images [J]. *Science of Surveying and Mapping*, 2021, 46(9): 115-121. [韩利冬, 林祥国, 梁勇. 一种高分辨率卫星影像自动云检测方法 [J]. *测绘科学*, 2021, 46(9): 115-121.]
- [9] Yang X, Lo C P. Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area [J]. *International Journal of Remote Sensing*, 2002, 23(9): 1775-1798. DOI:10.1080/01431160110075802.
- [10] Chen Shiyang, Liu Jia. Evaluation of deep learning algorithm for crop identification based on GF-6 time series images [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2021, 37(15): 161-168. [陈诗扬, 刘佳. 基于GF-6时序数据的农作物识别深度学习算法评估 [J]. *农业工程学报*, 2021, 37(15): 161-168.]
- [11] Zhu Shuang, Zhang Jinshui, Shuai Guanyuan, *et al.* Winter wheat mapping by soft and hard land use/cover change detection [J]. *Journal of Remote Sensing*, 2014, 18(2): 476-496. [朱爽, 张锦水, 帅冠元, 等. 通过软硬变化检测识别冬小麦 [J]. *遥感学报*, 2014, 18(2): 476-496.]
- [12] Ji Z, Pan Y, Zhu X, *et al.* Prediction of crop yield using phenological information extracted from remote sensing vegetation index [J]. *Sensors*, 2021, 21: 14064. DOI:10.3390/s21041406.
- [13] Wan Changjun, Wu Xiaodan, Lin Xingwen. Impact of spatial and temporal scales of remote sensing data on the spatiotemporal change in geographic elements [J]. *Journal of Remote Sensing*, 2019, 23(6): 1064-1077. [万昌君, 吴小丹, 林兴稳. 遥感数据时空尺度对地理要素时空变化分析的影响 [J]. *遥感学报*, 2019, 23(6): 1064-1077.]
- [14] Yuan Zhoumiqi, Zhang Jinshui. Fusion of Spatiotemporal remote sensing data for changing surface characteristics [J]. *Journal of Beijing Normal University (Natural Science Edition)*, 2017, 53(6): 727-734. [袁周米琪, 张锦水. 面向地表特征变化区域的时空遥感数据融合方法研究 [J]. *北京师范大学学报(自然科学版)*, 2017, 53(6): 727-734.]
- [15] Cho M A, Debba P, Mathieu R, *et al.* Improving discrimination of savanna tree species through a multiple-endmember spectral angle mapper approach: Canopy-level analysis [J]. *IEEE Transactions on, Geoscience and Remote Sensing*, 2010, 48(11): 4133-4142. DOI: 10.1109/TGRS. 2010. 2058579.
- [16] Li Henghai, Ou Bin, Liu Yuting, *et al.* Citrus recognition methods of hyperspectral remote sensing image based on spectral unmixing model [J]. *Remote Sensing Technology and Application*, 2017, 32(4): 743-750. [李恒凯, 欧彬, 刘雨婷, 等. 基于混合像元分解的高光谱影像柑橘识别方法 [J]. *遥感技术与应用*, 2017, 32(4): 743-750.]
- [17] Su Yang, Qi Yuan, Wang Jianhua, *et al.* Land cover extraction in the Ejina Oasis by hyperspectral remote sensing [J]. *Remote Sensing Technology and Application*, 2018, 33(2): 202-211. [苏阳, 祁元, 王建华, 等. 基于航空高光谱影像的额济纳绿洲土地覆被提取 [J]. *遥感技术与应用*, 2018, 33(2): 202-211.]
- [18] Ping Yuepeng, Zang Shuying. Crop identification based on MODIS NDVI time-series data and phenological characteristics [J]. *Journal of Natural Resources*, 2016, 31(3): 503-513. [平跃鹏, 臧淑英. 基于MODIS时间序列及物候特征的农作物分类 [J]. *自然资源学报*, 2016, 31(3): 503-513.]
- [19] Pan Y, Li L, Zhang J, *et al.* Winter wheat area estimation from MODIS-EVI time series data using the crop proportion phenology index [J]. *Remote Sensing of Environment*, 2012, 119: 232-242. DOI:10.1016/j.rse.2011.10.011.

- [20] Rezaei E E, Ghazaryan G, Gonzalez J, *et al.* The use of remote sensing to derive maize sowing dates for large-scale crop yield simulations[J]. *International Journal of Biometeorology*, 2020,4(65): 565-576. DOI:10.1007/s00484-020-02050-4.
- [21] Zhang Jinshui, Shen Kejian, Pan Yaozhong, *et al.* HJ-1 remotely sensed data and sampling method for wheat area estimation[J]. *Scientia Agricultura Sinica*, 2010, 43(16): 3306-3315.[张锦水, 申克建, 潘耀忠, 等. HJ-1号卫星数据与统计抽样相结合的冬小麦区域面积估算[J]. *中国农业科学*, 2010, 43(16): 3306-3315.]
- [22] Zhu Shuang, Zhang Jinshui. Medium resolution remote sensing-based winter wheat mapping corrected by low-resolution time series remote sensing images [J]. *Remote Sensing for Land and Resources*, 2020, 32(1):19-26.[朱爽, 张锦水. 时间序列低分影像修正中分遥感冬小麦分布[J]. *国土资源遥感*, 2020,32(1): 19-26.]
- [23] Xie Qiaoyun, Huang Wenjiang, Cai Shuhong, *et al.* Comparative Study on remote sensing inversion methods for estimating winter wheat leaf area index [J]. *Spectroscopy and Spectral Analysis*, 2014, 34(5): 1352-1356.[谢巧云, 黄文江, 蔡淑红, 等. 冬小麦叶面积指数遥感反演方法比较研究[J]. *光谱学与光谱分析*, 2014, 34(5): 1352-1356.]
- [24] Liang Dong, Guan Qingsong, Huang Wenqing, *et al.* Remote sensing inversion of leaf area index based on support vector machine regression in winter wheat [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2013, 29(7): 117-123. [梁栋, 管青松, 黄文江, 等. 基于支持向量机回归的冬小麦叶面积指数遥感反演[J]. *农业工程学报*, 2013, 29(7): 117-123.]
- [25] Sheng Hui, Chi Haiming, Xu Mingming, *et al.* Inland water chemical oxygen demand estimation based on improved SVR for hyperspectral data [J]. *Spectroscopy and Spectral Analysis*, 2021, 41(11): 3565-3571.[盛辉, 池海旭, 许明明, 等. 改进SVR的内陆水体COD高光谱遥感反演[J]. *光谱学与光谱分析*, 2021, 41(11): 3565-3571.]

## Influence Factors Analysis on Accuracies of Winter Wheat Distribution from Low and Medium Resolution Compositd Remote Sensing Images

Zhu Shuang<sup>1,2</sup>, Zhang JinShui<sup>2,3</sup>

(1. *Beijing Polytechnic College, Beijing 100042, China;*

2. *Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China;*

3. *Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China)*

**Abstract:** The quality and quantity of sample dataset from medium resolution remote sensing images is the key factor to contribute to the efficiency of low and medium resolution identification model. For winter wheat in this paper, we constructed a support vector regression model coupled with low and medium resolution images, to decomposed of mixed pixels, and exact winter wheat extent. Then analyzed the influences of sample quantity and quality of medium resolution remote sensing images respectively. The results states that only 10% quantity of samples are enough to achieve stable accuracy. Under this quantity, regional accuracy and pixel accuracy could reach higher than 98% and 92% respectively in typical winter wheat area. In terms of sample quality, the accuracy of result improved accompanying with the sample quality increment. We found that high accuracy could achieved when the sample quality is better than 60%. While in the area where medium resolution sample did not exist in area with medium samples, regional accuracy and pixel accuracy also increased accompanying with the sample amount and quality increment. In this area, 20% quantity of medium resolution sample was needed enough to achieve 97% of regional accuracy and 92% of pixel accuracy respectively. The above demonstrate the successful generalization of winter wheat identification by medium resolution sample to non-medium resolution area.

**Key words:** Support Vector Regression(SVR); Pixel unmixing; Sample quantity/quality; TM; MODIS