

应用 NOAA / AVHRR 数据测算局地 水稻种植面积方法研究

李郁竹

曾 燕

(中国气象科学研究院 北京 100081) (南京气象学院 南京 210044)

摘 要 本文采用在模糊监督分类中增加迭代过程的方法——模糊监督分类—迭代法,在分解混合像元的基础上,利用 AVHRR 数据估算水稻种植面积。根据稻田与旱地存在温度差异的特点,在分析 AVHRR 数据统计特征的基础上,增加了第 3 和 4 两通道参加模糊监督分类,从而增加了分类像元值矢量维数,增强了对水稻的鉴别能力。所采用的迭代法是收敛且有效的,经本方法输出的水稻种植面积百分含量图与实际水稻分布十分吻合,测算出的面积值与 TM 和实地调查结果相比,有较高的精度。

关键词 NOAA / AVHRR, 水稻, 种植面积, 迭代, 模糊监督分类

1 引 言

水稻产量估测历来是国家农业决策部门的重要经济信息。在产量估测中,种植面积是必不可少的参数。我国的水稻种植具有田块小,农田作物布局复杂,同期生长的绿色植被多等特点,应用卫星遥感测算水稻种植面积存在一定的难度。利用高分辨率陆地卫星(Landsat)的专题制图仪(TM)资料估算水稻种植面积,虽能满足精度要求,但由于存在数据费用昂贵、资料量大、数据源得不到保证等问题而不便推广应用。美国第 3 代业务极轨气象卫星(TIROS—N / NOAA 系列)装载的改进的甚高分辨率辐射仪(AVHRR)数据具有易于获取(双星运行,重复观测周期短,一天同一地区可以获得 4 次资料)、数据量小、处理费用低等优点,比较符合我国国情和实际应用的需要,但由于受空间分辨率低(1.1km)的限制,应用 AVHRR 数据进行水稻遥感估产,尤其是测算水稻种植面积,AVHRR 像元尺度远大于水稻种植田块尺度,同一 AVHRR 像元包含多种地物,混合像元问题突出,增加了应用难度。

本文通过研究混合像元中水稻种植面积信息提取和分析各通道数据的应用潜力,探讨了应用 NOAA—AVHRR 数据测算局地水稻种植面积的方法。

2 资料及处理

本研究区域为天津市津南区,是天津市水稻主产区之一。总辖区面积为 38926.03hm²,全区共分为 10 个乡镇,主要种植作物有水稻、高粱、玉米等。研究中应用的资料:

2.1 遥感资料

遥感资料包括彩红外航片、Landsat / TM 数据和 NOAA / AVHRR 数据 3 种。

彩红外航片拍摄时间为 1995 年 1 月 10 日—13 日,比例尺为 1:3.5 万。在应用 AVHRR 数据测算水稻种植面积时,航片用来提供样本像元内部的各种地物类别信息。

TM 数据接收时间分别为 1995 年 1 月 12 日和 8 月 8 日,均为晴空资料,经过了初步的几何和辐射校正等预处理,空间分辨率为 30m。本研究共选用了 5 个波段,即第 2 至第 6 波段(TM2—TM6),以土地利用现状图为准,选取控制点进行二次多项式拟合校正,处理后的 TM 图像与数字化边界匹配吻合好,误差控制在 2 个像元以内。在研究中,TM 数据主要用于估算研究区水稻种植面积本底准真值。

AVHRR 数据分别取自 NOAA-12(降轨)和 NOAA-14(升轨)的资料,经过系统预处理,空间分

收稿日期:1997-03-21,收到修改稿日期:1997-10-13

辨率为 1.1km。为相对消除大气的影响,仅选取晴空资料。为克服远离星下点 AVHRR 数据分辨率降低、图像产生几何畸变的不足,取研究区位于卫星扫描带中心三分之一以内位置的资料^[1]。经上述条件筛选后,能够应用的时相有:1995 年 6 月 19 日和 27 日,7 月 20 日,9 月 11 日;1996 年 6 月 11 日和 14 日,7 月 4 日,9 月 20 日。经系统定位处理的 AVHRR 数据,在实际应用时,根据地物的地理位置再次进行地理定位处理,定位误差控制在一个像元以内,主要选用了 1 至 4 通道。

本研究的遥感图像处理 and 行政边界数字化等工作是在地理信息系统软件 IDRISI 支持下完成。数字化误差控制在 30m 以内。

2.2 地面资料

地面光谱观测资料 为了了解天津水稻生长期间,水稻及其它有关环境背景地物的光谱特征及其规律性,为遥感影像日视解译和计算机识别提供依据,在津南区,水稻关键发育期内,对水稻及其它作物和背景地物进行了地面光谱测试。

其它资料 地面农学观测和农业气候资料取自农业气象观测站的常规观测资料,土地利用现状资料、地形图等由天津市土地管理局提供,历年作物产量、种植面积资料由天津市统计局提供。

3 研究方法

3.1 模糊监督分类

模糊监督分类方法在 Fangju Wang 1990 年文章^[2]中作了较详细的介绍,该方法将各种地物类别均看成是模糊集合,像元为模糊集合的元素,每一像元均与一组隶属度值相对应,隶属度值代表该像元对某一地物类别模糊集合的归属程度,也即是此像元中所含此种地物类别的面积百分比。模糊监督分类先选择样本像元,根据样本像元计算各种地物类别的模糊均值矢量和模糊协方差矩阵。每一种地物类别的模糊均值矢量为:

$$\mu_c^* = \frac{\sum_{i=1}^n f_c(x_i)x_i}{\sum_{i=1}^n f_c(x_i)} \quad (1)$$

式中 n 为样本像元总数; f_c 为第 c 类别的隶属函数; x_i 为样本像元值矢量 ($1 \leq i \leq n$)。模糊协方差矩阵可用下式表达:

$$\sum_c^* = \frac{\sum_{i=1}^n f_c(x_i)(x_i - \mu_c^*)(x_i - \mu_c^*)^T}{\sum_{i=1}^n f_c(x_i)} \quad (2)$$

μ_c^* 和 \sum_c^* 确定之后,对每一像元确定其对各种类别的隶属度,也即对每一像元进行模糊监督分类,求算各种类别在其内所占面积百分比。采用最大似然分类法,此时,用模糊均值矢量 μ_c^* 和模糊协方差矩阵 \sum_c^* 来代替最大似然分类法中的均值矢量和协方差矩阵(见(4)式),这时属于 c 类别的隶属函数定义为:

$$f_c(x) = \frac{p_c^*(x)}{\sum_{i=1}^m p_i^*(x)} \quad (3)$$

其中

$$p_i^*(x) = \frac{1}{(2\pi)^{N/2} |\sum_i^*|^{1/2}} \cdot \exp\left[-\frac{1}{2}(x - \mu_i^*)^T \sum_i^{*-1}(x - \mu_i^*)\right] \quad (4)$$

式中, N 是像元光谱值矢量的维数, m 是预先定义的类别数, $1 \leq i \leq m$ 。

一个像元矢量 x 的隶属度依赖于其在光谱空间的位置, $f_c(x)$ 随 $(x - \mu_c^*)^T \sum_c^{*-1}(x - \mu_c^*)$ 的减少呈指数增长,即与类别 c 的中心越接近,其值越大。

3.2 模糊监督分类—迭代法

在模糊监督分类法中,要先确定样本像元对各种类别的隶属度,即求出样本像元中各类别的面积百分含量。它可通过地面调查、航片、高分辨率卫星图像分类等方法获得,但无论采用何种方法,求出的样本像元隶属度会存在误差,只能是一靠近真值的相对准确值。同理,由此样本像元求算出的模糊均值矢量和模糊协方差矩阵也存在误差,进而导致求算出的各像元隶属度存在误差,也只能是一靠近真值的相对准确值。为了克服上述初始误差,使算出的模糊均值矢量和模糊协方差矩阵尽量准确,保证像元隶属度的精度,本研究采用了迭代方法,即通过增加迭代过程反复求算模糊监督分类法中的模糊均值矢量和模糊协方差矩阵,使求算出的像元隶属度从靠近真值的相对准确值最终接近误差范围允许之内的真值。终止此过程的条件为:每次求出研究区各像元隶属度之后,以像元面积乘以像元对水稻的隶属度(像元水稻面积百分比)累加作为研究区水稻

种植面积值,当连续两次计算的研究区水稻种植面积相对误差小于阈值时,停止迭代。本研究将该阈值取为 0.1%,迭代过程框图见图 1。

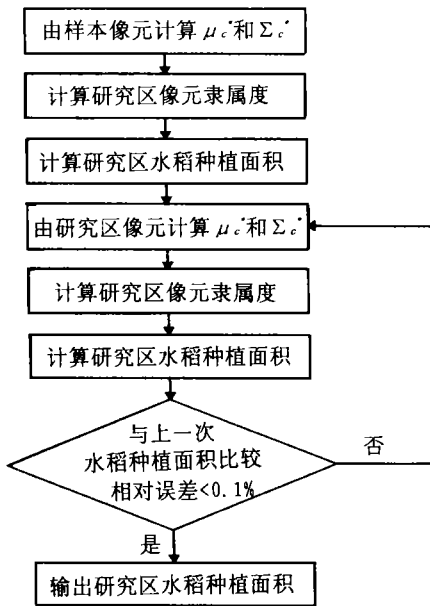


图 1 模糊监督分类—迭代法测算水稻种植面积框图

Fig. 1 Flow chart for evaluating rice planting area using fuzzy supervised classification—iteration method

实际运用结果表明:本迭代过程稳定、收敛快;经迭代之后输出的研究区像元水稻种植面积百分含量图明显优于未经迭代的效果。

3.3 测算水稻种植面积 AVHRR 通道选择

AVHRR 数据共具有 5 个通道,其波长范围分别为:0.58—0.68 μm (CH1), 0.725—1.1 μm (CH2), 3.55—3.93 μm (CH3), 10.3—11.3 μm (CH4), 11.5—12.5 μm (CH5)。长期以来,国内外用于作物估产一般仅用第一通道(相当于可见光波段)和第二通道(相当于近红外波段)^[1,3],对于中红外及热红外通道 CH3—CH5 应用极少。Kerber. A. G^[4]等人在 1986 年曾应用 AVHRR 数据(CH1—CH4)进行森林、植被、城市的分类研究,表明,地物分辨能力以发射辐射波段为最高,其中,CH3 又优于 CH4。为探讨 AVHRR 数据 CH3 和 CH4 对水稻鉴别的可能作用,我们对所具有的 AVHRR 数据做了统计特征分析和单通道图像分析。

对研究区各时相 AVHRR 资料进行的统计特征分析(表略)得出如下结论:AVHRR 数据以 CH3 数据等级最多,方差最大,其次为 CH2 和 CH4,以 CH1 最小。方差越大,表明影像色彩越丰富,包含的信息

量也越大。因此,以 CH3 数据信息量最多,其次为 CH2 和 CH4,最后为 CH1;CH3 和 CH4 数据相关性较高,二者间有一定的信息重复;此外,CH3 和 CH2 两通道对 AVHRR 数据信息量概括最多。

将 6—9 月份研究区各时相 AVHRR CH3 和 CH4 数据分别进行线性拉伸后成图,从所有拉伸后图像上均可看出水稻种植区比非水稻种植区亮温值偏低,呈现出“稻区低温”现象。此区大致呈不规则倒三角形。通过对比分析可看出,由 AVHRR 数据提取的水稻种植区与土地利用现状图上的水稻种植区极为一致,很好地反映了水稻种植区域。

以上分析充分证明,CH3 和 CH4 有助于水稻鉴别。因此,在本研究中增加中红外和热红外通道,采用 1、2、3、4 四个通道的数据进行分类,用 CH1、CH2 体现水稻植被信息特点,CH3、CH4 体现稻田温度特点。

3.4 土地利用分类

在运用模糊监督分类—迭代法测算研究区水稻种植面积以前,需将研究区各种地物进行分类。类别划分的主要原则为,同一种类别具有基本一致或相近的光谱特征,不同类别间光谱特征差异大。结合天津市津南区实际土地利用特点,将地物共划分为 5 种类别,即:旱地,稻田,菜地,居民地,水域。不同类别在卫星图像上光谱特征反映不同,旱地主要包括旱地(高粱、玉米)、果园、林地,在整个生长季的大大部分时间里,都体现为较强的植被信息;在水稻生长季的大大部分时间里,稻田体现为水体和植被的综合信息,在中、热红外图像上,“稻田低温”现象明显;菜地由于其种植时间、种植种类及田间布局特点与旱地差别较大,植被光谱信息一般比旱地弱,因此与旱地分开而单分为一种类别;居民地包括乡镇居民住宅、交通及工矿用地,这些地区本身温度比较高,在中红外和热红外波段特征明显,表现为高亮温值;水域主要包括河流、沟渠、鱼塘、坑塘等,主要体现为水体信息,在可见光和近红外波段反射率比较低,在中红外和热红外波段亮温值低。

3.5 样区选取和样本像元隶属度确定

在用模糊监督分类求像元隶属度时,需先选取样本像元之后再由已知的样本像元隶属度来确定模糊均值向量和模糊协方差矩阵。由于 AVHRR 数据空间分辨率低,在 1 个像元误差的定位精度下,采用单像元抽样将存在较大误差。因此,在选取样本像元时改为

选取样区,即以 一个样区状况代表一个样本像元状况。每一样区由 4 个、6 个或 9 个 AVHRR 像元构成一正方形或长方形区域。样区选取以具有代表性和均匀分布于整个研究区为原则。AVHRR 的每一像元具有 4 个通道的数据,相当于一个 4 维矢量,将样区所包含像元的 AVHRR 数据值矢量平均值和各种地物类别在样区中所占面积百分比作为样本像元值矢量和样本像元隶属度。本研究共选取了 10 个样区,共包括 56 个 AVHRR 像元。样区中各种地物类别所占面积百分比是利用航片求取的。根据土地利用现状图,在航片上勾绘出每一样区,通过对航片的目视解译,最后利用方格法求算出每个样区中各种地物类别所占面积百分比。由此,确定样本像元隶属度。

4 计算结果与分析

4.1 水稻种植面积测算最佳时期的确定

面积提取精度在空间上取决于作物品种布局和田块分布特征,在时间上取决于选择合适的测算时期。因此,恰当地选取水稻种植面积测算时期,对于保证最终面积测算精度是必不可少的条件。综合分析研究区水稻和其它作物生育期,及 AVHRR 单通道图像,确定水稻种植面积测算的最佳时期为 7 月中旬至 8 月上旬。在此时段内,植被、水体和城市在 AVHRR 的 CH1 和 CH2 图像上反差大,水稻与旱地作物长势旺,但二者植被信息强度又略有差别;在 CH3 和 CH4 图像上有“稻区低温”现象。这样,综合 4 个通道信息利于鉴别水稻。在水稻生长季内,由于受天气状况等因素的影响,最佳时期内的资料有时不一定能够获取,此时可考虑应用 8 月中旬至 9 月中旬的资料,此时段,植被、水体和城市在 AVHRR 的 CH1 和 CH2 图像上的信息层次虽不及 7 月中旬至 8 月上旬大,但亦有一定的反差,通过 CH3 和 CH4 图像的“稻区低温”现象,也可体现水稻信息。

4.2 水稻种植面积测算及效果检验

4.2.1 水稻种植面积准真值的获取

本研究区统计上报的水稻种植面积因受种种因素的影响,可信度较差。为得到一相对准确的研究区水稻种植面积值,作为检验标准,采用陆地卫星 TM 数据获取的水稻种植面积为准真值。

我们获取的 TM 数据时相为 1995 年 1 月 12 日和 8 月 8 日,选用的 5 个波段分别为: 0.52—0.60 μm (TM2), 0.63—0.69 μm (TM3), 0.76—0.90 μm

(TM4), 1.55—1.75 μm (TM5), 10.40—12.50 μm (TM6)。在对 5 个波段数据进行统计特征分析的基础上,确定波段组合最佳方案,选用 TM2, 3, 4 和 TM3, 4, 5 两种组合。经边界配准和研究区提取并以 2.5% 进行饱和和线性拉伸后生成假彩色合成图像。在 1 月 12 日生成的假彩色图像,因农田中基本没有作物,整幅图像色调较暗,经计算机亮度级识别,稻田与旱地差异不明显,不好区分。在 8 月 8 日生成的假彩色图像地物信息多,以 TM3, 4, 5 分别赋予绿, 红, 蓝合成图像效果最好,稻田与旱地亮度级差异明显,易于区分。最后选用此幅图像作为测算 TM 水稻种植面积的本底图。

对 TM 本底图进行中值滤波和均值滤波处理,结合航片和土地利用现状图,目视解译判识其中稻田,并生成研究区稻田影像(图略)。经实地考查证明,所生成的研究区稻田影像与实际水稻分布情况吻合极好。最后用 IDRISI 求算面积模块,求算研究区水稻种植面积,其值为 7595.2 hm^2 。

4.2.2 AVHRR 水稻种植面积测算

AVHRR 数据采用 4 个通道,将求算出的样本像元隶属度和样本像元值矢量代入(1)—(4)式,即可求出研究区各像元中水稻种植面积百分比。由于本研究区范围比较小(38.845°—39.075°N, 117.230°—117.550°E),将 AVHRR 图像提取的研究区国土面积除以研究区像元总数作为单像元面积值。在迭代算法中,以单像元面积值乘以单像元水稻种植面积百分比累加,即可求得研究区水稻种植面积值,由

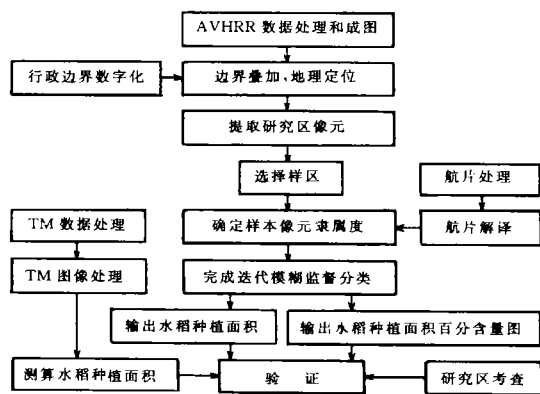


图2 NOAA/ AVHRR 数据测算水稻种植面积流程图
Fig. 2 Flow of Evaluating Rice Planting Area from NOAA/AVHRR data

AVHRR 数据测算水稻种植面积流程图见图 2。以水稻种植面积百分含量作为像元亮度值,经迭代模糊监督分类之后输出研究区像元水稻种植面积百分含量图(图版 1 图 3),图像中蓝, 红, 黄, 绿依次对应水稻

种植面积百分比逐渐升高。将 1995 年 AVHRR 4 个时相(6 月 19 日,6 月 27 日,7 月 20 日,9 月 11 日)资料分别代入模糊监督分类-迭代法,输出未经迭代的和经过迭代后的研究区像元水稻种植面积百分含量图。得到如下结论:迭代运算稳定、收敛快;各时相经迭代之后输出图均比未经迭代的效果好(图版 I 图 3,图 4);在所有输出图中,以 7 月 20 日迭代输出图最能反映实际情况,与各乡实际水稻种植情况吻合好,从而证明了最佳时期内的资料最有利于水稻种植面积的提取。为了比较不同的 AVHRR 通道选择对水稻种植面积测算的作用,我们将 7 月 20 日 AVHRR 资料以单通道,两个通道和三个通道组合分别代入算法,输出研究区像元水稻种植面积百分含量图,并与四个通道输出图进行比较,结果表明:所有情况中,增加了 CH3 和 CH4 的四个通道输出图效果最好(图版 I 图 3),明显优于仅用 CH1 和 CH2 的输出图(图版 I 图 5),进一步证实了 CH3 和 CH4 对水稻鉴别的作用。从通道选择分析可知,CH2 和 CH3 概括了 AVHRR 资料(CH1, 2, 3, 4)的大部分信息,用这两通道的输出图虽不及四个通道的输出图效果好,但也能反映研究区实际情况,当四个通道数据收集不全时,可考虑用此二通道输出图代替。

为检验 AVHRR 种植面积推算精度,我们对比分析了几种不同资料所确定的研究区水稻种植面积值(表 1)。其中 1990 年天津市土地利用现状调查的水稻种植面积值,由航空遥感 1:1 万彩色红外航片图结合《土地利用现状调查技术规程》(全国农业区

表1 天津市津南区水稻种植面积一览表
Table 1 Tabulation of Rice Planting area in South of Tianjin

应用资料	水稻种植面积
1995年7月20日AVHRR资料	7552.8(hm ²)
1995年8月8日TM资料	7595.2(hm ²)
1990年天津市土地利用现状调查	8035.7(hm ²)
1990年统计上报	4159.8(hm ²)
1995年统计上报	4092.0(hm ²)

化委员会编)求得,由天津市土地局提供。表中的统计上报值,因受有关农业税等问题的影响,比其它方法得到的水稻种植面积值明显偏小,但年际间变化趋势基本一致。比较统计上报值可看出,研究区的水稻种植面积 1995 年比 1990 年略有减少。比较 1995 年由 AVHRR 和 TM 资料推算的水稻种植面积值和 1990 年天津市土地利用现状调查的水稻种植面积值,也呈减少趋势。由此可看出,由 AVHRR

资料推算出的水稻种植面积值是客观、可信的。若以 1995 年 TM 目视解译推算的研究区水稻种植面积为准真值,则 1995 年用 AVHRR 数据推算的水稻种植面积相对误差为 -0.6%。

采用模糊监督分类-迭代法,利用 1996 年能够应用的 4 个时相的 AVHRR 资料,试测算了 1996 年天津津南区水稻种植面积。由于这些时相的资料均不在“最佳时期”内,故实际输出的研究区像元水稻种植面积百分含量图均不甚理想,以 9 月 20 日 CH2 和 CH3 两通道图效果相对最好(CH1 和 CH4 由于存在不明噪音而未被采用)(图略),用此时相资料测算出的研究区 1996 年水稻种植面积为 7626.5hm²。此值与 1995 年水稻种植面积值相比,水稻种植面积略有增加。将 1996 年 9 月 20 日输出图与 1995 年 7 月 20 日输出图叠加分析,可看出 1996 年研究区水稻种植面积较 1995 年增加的地区为葛沽镇、八里台、北闸口三乡。上述结论经实地考察,完全得到了证实。

5 结束语

通过本项研究,初步结论如下:

1. 模糊监督分类-迭代法用于 AVHRR 数据的局地水稻种植面积测算,可得到较高的测算精度,且计算方法简便,实际运用的迭代过程效果明显,收敛快,完全符合实际推广应用的需要。

2. 通过混合像元分解,提取混合像元内各种类别面积信息,可在一定程度上克服 AVHRR 数据空间分辨率低的局限。

3. 增加 CH3 和 CH4 两中、热红外通道,可增强水稻识别能力,提高水稻种植面积提取精度。

由于有迭代过程的存在,本研究方法对样区各种地物类别所占面积百分比的推算精度要求比较宽松。因此,该百分比值除可由小比例尺航片获取外,也可采用地面调查或由高分辨率 TM 数据分类获得。

本方法亦可用于大范围作物种植面积的提取,可在划分作物生态类型区——区域分层的基础上,选择样区,完成测算。

参 考 文 献

- 1 Li Yuzhu. Estimating Production of Winter Wheat by Remote Sensing and Unified Ground Network, II Nationwide Estimating of Wheat Yields. Applications of Remote Sensing in Agriculture, Butterworths UK, 1990, 149--157.

- 2 Fangju Wang. Fuzzy supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 1990, 28(2): 194—201.
- 3 李郁竹等. 冬小麦气象卫星遥感综合测产技术与试验. 中国遥感进展. 北京: 万国学术出版社, 1992, 130—136.
- 4 Kerber, A. G. et al. Land—cover mapping. *Photogrammetric Engineering and Remote Sensing*, 1986, 1877—1883.

作 者 简 介

李郁竹, 女, 1935年5月生, 正研级高级工程师。毕业于苏联敖德萨水文气象学院农业气象专业, 现从事农业遥感(遥感估产), 在国内外发表论文38篇, 主编论文集1册, 合著2册。

Study on Methods of Rice Planting Area Estimation at Regional Scale Using NOAA/ AVHRR Data

Li Yuzhu

Zeng Yan

(Chinese Academy of Meteorological Sciences, Beijing, 100081)

(Nanjing Institute of Meteorology, Nanjing, 210044)

Abstract Fuzzy supervised classification—iteration method was used in evaluating the rice planting area on a mixed pixel signal unmaxing basis. Channels 3 and 4 of AVHRR were used in this study according to the fact that the temperature of rice fields different from that of rainfalled—crop fields and based on the analysis of statistical characteristic. The additional channels in middle infrared and thermal infrared allowed rice to be identified easier than using the visible and near—infrared channels alone, and added the dimensional number of the pixel value vector in the classification. The iterative process is effective and convergent. The pixel proportion images of rice planting area produced in this study have a good consistence with the ground truth Compared with rice area obtained from TM image and field survey, this method has a high accuracy.

Key words NOAA / AVHRR, Rice, Planting area, Iteration, Fuzzy supervised classification