

# CONSISTENCY ANALYSIS ON PADDY RICE AREA SURVEY WITH SPOT AND TM IMAGES UNDER THE TOTAL RESTRAINT QUANTITY

Li Le<sup>\*</sup>; Li Chao; Zhu Xiufang; Pan Yaozhong

State Key Laboratory of Earth Surface Processes and Resource Ecology (Beijing Normal University); College of Resources Science & Technology, Beijing Normal University, Beijing, 100875. China - (lile, robbin\_li, zxf, pyz)@ires.cn

Commission VIII, WG VIII/10

**KEY WORDS:** Image Interpretation; Percentage Vegetation Cover; Space Photogrammetry; Comparative Analysis; Crop; Satellite Remote Sensing; Classification

## ABSTRACT:

Many global problems can be estimated and forecasted by remote sensing with its advantage of periodicity, objectivity and large scale scope, besides plant area estimation. Conventional classification methods always lead to different extracting results from different resolution satellite images because of the difference between different remote sensors, different periods, different resolutions as well as other factors. This paper presents an extraction method under the restraint of total area of paddy rice, using a remote sensing sampling method to obtain accurate rice area, and re-extracting information under the restraint of total area of paddy rice, to minimize the inconsistency owing to different data sources. This study will also show that the consistency analysis model proves this method can provide a higher accuracy in extracting plant area from middle and low resolution satellite images, not only in total quantity but also in spatial distribution.

## 1. INTRODUCTION

Statistical work for paddy rice, the head of three main crops in China, is becoming the focus of statistics for crops in China. In a highly developing country such as China, the agricultural planting structure is always different, and the crop acreage is often fluctuating. The traditional methods, signified by report by administration unit and sampling methods for crop yields, used to be interfered by human factors in the course of data acquisition. Therefore, it has been quite difficult for conventional methods to obtain timely and accurate paddy rice area data (Jiao, et al, 2006).

The periodicity, objectivity and large scale scope of remotely sensed data, with its capability to obtain a wide range of land cover information with different temporal resolution and spatial resolution, provide a new bright future for statistics for global resources. Consequently, many of the global problems can be estimated and forecasted by remote sensing. In many foreign countries, this method worked out well in monitoring a wide range of planting structure. In China, some significant characteristics, such as intercropping of crops, land fragmentation, frequent variation of planting types, and vulnerability to natural disasters of rural defense has proposed new challenges for the use of remote sensing in dynamically monitoring and obtaining objective, accurate and timely statistics crop area.

In crop area measurements based on the characteristics of remote sensing images, Su-Qin Han applied mixed supervised classification and non-supervised classification methods in monitoring the area of winter wheat with the EOS / MODIS satellite data (Han, et al, 2003). Following the pixel-oriented

supervision and non-supervisory methods and the object-oriented approach, Dufenglan, combines cognition and object-oriented method to achieve the precise classification of features. In his experiment, segmentation for remote sensing images are carried out, and more representative images of surface entities to extract more supporting classification information are obtained, which leads to the overall classification accuracy of 92.79 percent (Du, et al., 2004). Object-oriented methods used the spatial characteristics and spectral characteristics to process classification, and effectively overcome inherited deficiencies. The information-band of remote sensing images combined with the ecological survey index reflect the spatial information of crop growth, and effectively realize growth monitoring and crop yield estimation. This method, compared to conventional ways, has achieved marked improvement in classification accuracy, and the overall direction is in the distribution of nearly unlimited ideal of the true value.

Crop area measurement technology integrated with statistical sampling techniques is widely used in the global large-scale agricultural survey projects. The EU MARS plan adopted the stratified random sampling method and the United States Large Crop Yield Assessment Plan (LACIE), as well as agriculture and the resources of space remote sensing survey plan (AGRIS2TARS), and the Area Sampling Frame (GALLEGO, 2004). Bing-Fang Wu, in the project of "China agriculture and remote sensing monitoring system," certified the feasibility of two-level stratified sampling to extract crop cultivation area (Wu, et al, 2004). This method made marked achievement in classification accuracy in crop area measurement integrating a statistical sampling technique based on the statistical

---

The work is supported by The National High Technology Research and Development Program of China (2006AA120101).

\* Corresponding author: Li Le Tel:86-13811401540

E\_mail Address:lile@ires.cn

significance, but didn't perform as well in the spatial distribution of extracted information.

This paper presents an extraction method under the restraints of total area of crops, using remote sensing sampling methods to obtain accurate crop area, and re-extracting information under the restraints of total area of crops, to minimize crop area inconsistency owing to different data sources.

## 2. DATA AND STUDY AREA

### 2.1 Data

In this study, two kinds of satellite images and one kind of auxiliary data are chosen. As the example of high resolution satellite images, the SPOT image is selected. The TM images are selected as examples of middle or low resolution satellite images. Additional, survey data collected in GIS database are opted as auxiliary information.

According to the growth calendar of paddy rice, the data of the SPOT satellite dates from August, 17<sup>th</sup>, 2006, with a resolution of 10 meters. The Landsat TM satellite dates from July, 30<sup>th</sup>, 2006, at nearly the same period as SPOT images.

Phenology	Date (D/M-D/M)
Sow	1/4~20/4
Seed	21/4~30/4
Raise seedling	1/5~10/5
Transplant	10/6~20/6
Re-greening	21/6~30/6
Tillering	1/7~20/7
Booting	21/7~10/8
Heading	11/8~30/8
Milky ripe	1/9~10/9
mature	11/9~20/9

Table 1 The phenology of paddy rice in study area.

### 2.2 Data preparing

The satellite image processing included geometric and radiance correction, merging, and classification. The images are classified in 30 classes to be interpreted. After combining the same classes into the same group, the remaining classes, which can't be defined in the last step, should be reclassified. Repeating the above approaches, the following classification results were obtained: the area of paddy rice extracted from SPOT images is 204497642 square meters, and 192708955 square meters in TM images, with the different rate as 5.8%.

### 2.3 Study area

The PiZhou city in Jiangsu Province was chosen as an experimental area because its large planted area of paddy rice. There are two kinds of paddy rice planted in the study area representing different spectral characters, which provides good samples for interpreting and classification. Unlike in other countries in which paddy rice is planted in large quantities, the plant structure in China is often fragmented, scattered, in small patch, which absolutely represents in PiZhou city.

## 3. THE RESTRAINT TOTAL QUANTITY AND THE CONSISTENCY ANALYSIS MODEL

### 3.1 Steps to calculate the restraining total quantity

This study gives an applied presentation of how to calculate the restraining total quantity by sampling satellite images.

The first step is preparing the sampling frame and the sampling units. The sampling frame is drawn referenced with satellite images. For images sampled by plant units, drawing with fixed path provide a good solution. Polygons of the frame center should not be counted after digital drawing. Spatial analysis results based on the classification of SPOT images in section 2 are chosen as the sampling units. This analysis shows, only the polygons in which wheat is present according to classification results.

A stratification should be defined if pertinent information is available. Considering each stratum independently, the strata are expected more to have a higher stability and a higher sampling rate in each stratum are expected to have a higher variability (Gallego, 2005). While choosing the stratum as 4,5,6...8, the sampling results are determined at a 95% confidence interval. A better balance of the number of sampling units and the sampling accuracy is found with a stratum value of 6. In this case, the number of sampled units is 89.

Based on the sampling products, interpretation depends on the GIS survey database to calculate the total quantity. Taking into the uncertainty factors made by interpreters, the work is repeated 3 times, and then the mean of them is taken.

For the stratification sampling model, the total amount of wheat for all sampling units can be calculated:

$$T \hat{X} = A \bar{x} = A \frac{\sum_{i=1}^n A_i x_i}{\sum_{i=1}^n A_i} \quad (1)$$

A<sub>i</sub> is the number of samples in each stratum, while X<sub>i</sub> is the wheat quantity in each sample, and n is the number of stratum in Eq1.

In this rule, the total quantity of the study area is 240585463 square meters. This value will be taken as the real total quantity in next processing, just named the restraining total quantity.

### 3.2 Extracting wheat area from satellite images under the restrain of the total quantity

Before processing under the total restraining quantity, pre-processing resulted in the classification into three groups: pure paddy rice area, mixed paddy rice area, non-paddy rice planted area. The mixed area was then masked from the original satellite images. The effect of the total restrain quantity in our study will be shown as follows:

1) Obtain the quantity of the paddy rice in mixed areas  
The quantity in all statistic areas satisfies the rule:

$$S_m = S_t - S_p \quad (2)$$

where  $S_m$  = the quantity in the mixing planted paddy rice area  
 $S_t$  = the total quantity  
 $S_p$  = the quantity in the pure planted paddy rice area  
 The extracting results from SPOT and TM images classification before presents in Table 2.

Group	Quantity in SPOT images	Areas in TM images
Pure paddy rice area	206908900	1762523300
Mixing paddy rice area	94667300	197933400
The quantity in mixing paddy rice area	336676562	64062162

Table 2 The extracting classification results from SPOT and TM images

- 2) Eliminate the noise statistics from the mask images of mixing area by ENVI4.4 Minimum Noise Fraction Function.
- 3) Working out the abundance results by selecting the pure ROI from original images and unmixing the masked satellite images by the linear spectral unmixing model. Classify the masked images of mixing area into 10 classes, and then calculate abundant value for each class.
- 4) According to the restrain quantity, the appropriate classes range are selected into the paddy rice area with the pure area.

### 3.3 The Consistency analysis model

Because no earlier model is present, in this experiment, the consistency coefficient in analysis model is defined as:

$$C_i = 1 - | S_{i(S)} / S_{AREA} - S_{i(T)} / S_{AREA} | \quad (3)$$

While  $S_{i(S)}$  and  $S_{i(T)}$  are the quantity area of paddy rice planted in a statistic units(i),  $S_{AREA}$  is the area of the statistic units.

Considering the sampling frame with the detailed patch information has more statistic value than complain standard grid, the patch frame is chosen as the particular statistic unit in the consistency analysis model.

The spatial analysis model can be specified as more efficient while taking the different plant structures and the different plant proportion into accounts. Accordingly, the consistency analysis model is represented as:

$$C_i = \frac{\sum_{i=1}^n [1 - | S_{i(S)} / S_{AREA} - S_{i(T)} / S_{AREA} |]}{n} \quad (4)$$

While  $n$  is the number of patch units in some plant proportion, other parameters are same as Eq(3).

The plant proportion is calculated by the quotient of the plant area and the patch area. The statistic stratum is divided by equal scale of 10%, and then the average of  $C_i$  is calculated in every strata.

### 4. APPLICATION AND ANALYSIS BASED ON THE CONSISTENCY ANALYSIS MODEL

The examples described in this section illustrate the potential of applying the restrain quantity method to extracting plant area from satellite images. The consistency analysis model is the best evaluated method and three comparison situations are considered.

The extracting quantity results by the total restrain quantity method

While the extractions from the SPOT images and TM images all have a foundation of the total quantity by sampling, the differences between the extracting results are comparatively small. The difference is only 0.67%.

Satellite image	The plant area of paddy rice extracted under the total restrain quantity	The restrain total quantity	accuracy
SPOT image	237390000	240585463	98.67%
TM image	242174700	240585463	99.34%

Table 3 The extracting results from satellite images under the control of the restraint total quantity

- 2) The distributing information by the total restrain quantity method

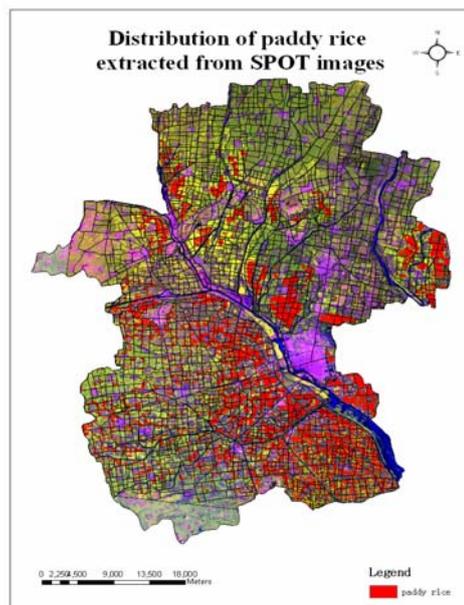


Figure1: The Distribution Of Paddy Rice Extracted From SPOT Images

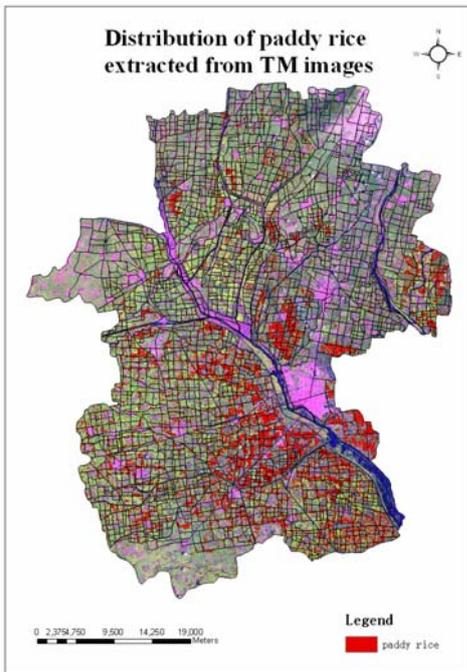


Figure2: The Distribution Of Paddy Rice Extracted From TM Images

The fact that the distribution difference is slight with the method of the restrain total quantity method is easy to be found. The extracting results in southeast and the middle area are exactly alike, even in some area where the plant area is comparatively small, the extracting results are also satisfactory.

- 3) The consistent coefficient by the restrain total quantity method

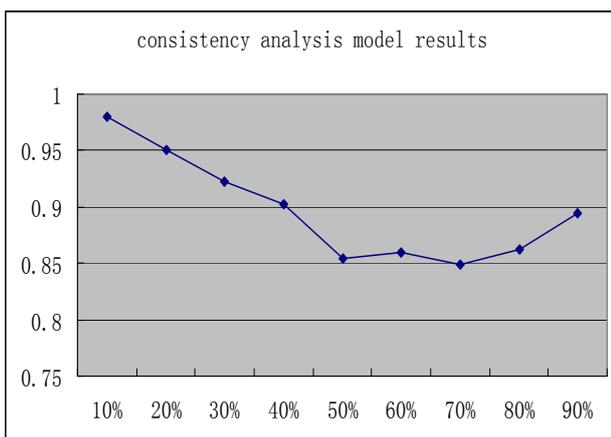


Figure2: The Diagram Of The Evaluation Result By The Consistency Analysis model

All the consistency coefficients can reach above 0.85 in any plant proportions. While plant proportions vary from 1% to 45%, the consistency coefficients in the model descend. Plant proportions vary from nearly 45% to 90%, and the stability and accuracy of the consistency coefficients behave well.

The diagram also shows that the performance is better when the plant proportions are low, which can make this method useful in more applications where the patches are small and cracked.

### 5. CONCLUSION AND DISCUSSION

The total restrain quantity method presented in this paper provided us a way to get a higher accuracy in extracting plant area from middle or low resolution satellite images, not only in total quantity but also in spatial distribution. This method makes use of the sampling technique in presenting a more accurate distribution map for plant area.,

Because of the veracity of total quantity gained from sampling, the uncertainty in extracting paddy rice from middle or low resolution satellite images is overcome. Different results between different resolution images are separated by results from re-extracting under total quantity. All the results from above allow us to get a better outcome for middle or low resolution satellite images, which can inspire a more widespread application.

In particular, the consistency analysis model indicated great accuracy in comparison between different plant proportions. However, this method has a number of limitations and may need more improvements.

The accuracy of classification before sampling hopes to be clearer, although the quantity accuracy is not so important. The classes which are grouped into non-plant area will no longer be ranked in the re-class group.

Evaluation of the experimentation outcome in the new-created consistency model is proven to be efficient, but the analyst can decide to use more effective consistency models. Further studies need to be taken into account, along with bias estimations.

### REFERENCE

DU Feng-lan, TIAN Qing-jiu, XIA Xue-qi, HUI Feng-ming,2004a. Object-oriented Image Classification Analysis and Evaluation. *Remote Sensing Technology And Applicatio*, 19(1),pp.20-24.

F. J. GALLEGO,2004a.Remote sensing and land cover area estimation. *International Journal of Remote Sensing*. 25(15),pp.3019–3047.

F. Javier Gallego,2005a. Stratified sampling of satellite images with a systematicgrid of points. *ISPRS Journal of Photogrammetry & Remote Sensing*, 59,pp. 369–376.

Han Suqing, Liu Shumei, 2003. Monitoring wheat area using EOS/MODIS satellite data.*Agriculture and entironment*, China, pp.298-300.

Jiao, Xianfeng, Yang Bangjie, & Pei Zhiyuan, 2006a. Paddy rice area estimation using a stratified sampling method with remote sensing in China. *Transaction of the CSAE*.,22(5),pp.105-111.

Lu D.S., Weng, Q.H., 2006a. Spectral mixture analysis of ASTER images for examining the relationship between urban thermal features and biophysical descriptors in Indianapolis,

Indiana, USA. *Environment of Remote Sensin*, 104(2),pp.157-167.

Volker, W., 2004a. Object-based classification of remote sensing data for change detection. *IRPRS Journal of Photogrammetry & Remote Sensing* 58,pp.225-238.

WuBing-fang, Li Qiang-zi,2004a. Crop Acreage Estimation Using Two Individual Frameworks with Stratification. *Journal of remote sensing* ,8(6),pp.551-569.

#### **ACKNOWLEDGEMENTS**

The work is supported by The National High Technology Research and Development Program of China (2006AA120101).

