Remote Sensing Application in China's Crop Acreage Estimation

Wei ZHOU
National Bureau of Statistics, Beijing, China
email:weizhou@263.net

Abstract

To acquire the planted acreage of major grain crops in a timely, accurate and cost -effective way is a goal for the National Bureau of Statistics (NBS) of China. In recent years, NBS has collaborated with some research institutes to apply the remote sensing technology in improving the method for crop surveys, especially in the acreage estimation for major grain crops. In the Chinese context, the remote sensing application in crop survey currently emphasized the following three aspects. First, by using multi-source and multi-temporal remote sensing imagery to extract the land cover types and major planted crops, an area frame is constructed and updated for crop sample design. Second, a strategy of sample selection is developed in order to make reasonable stratification and select samples to be surveyed as ground truth. Finally, combing the ground survey data with the classified grain acreage from remote sensing imagery as an auxiliary data, a linear model is adopted to produce the crop acreage estimation with a satisfied precision. Taking the year 2011's autumn crop survey in Heilongjiang province with the support of remote sensing imagery as an example, this paper illustrates the whole procedures.

Key Words: Area Frame, Support Vector Machine, Sample Selection, Linear Regression Estimation.

1. Introduction

China is a major agricultural production country as well as a large consumption and trade country in terms of farm products. Depending on the various agricultural information especially on the planted acreage statistics of major grain crops, it has been an important gist for national grain and food policy making as well as national economic plan making. To acquire the planted acreage of major grain crops and their spatial distribution in a timely, accurate, quantitative and periodic way is of significance for making the socio-economic development plans as well as plans for import and export of agricultural products at various levels of government, ensuring national food security, steering and adjusting the planting structure of farm crops through macro economic control, improving the operational management for relevant enterprises and farmers (Quarmby and Milnes et al.(1993); Tsiligirides(1998); Blaes and Vanhalle et al.(2005); Tao and Yokozawa et al.(2005); Chauhan and Arora et al.(2008)). Since the year 2003, the National Bureau of Statistics (NBS) of China has started to organize the study and application of the spatial information technology on agricultural statistics which mainly focus on remote sensing application, its potential usage and feasibility on agricultural statistics. Since the beginning of 2010, NBS has collaborated with external research institute to preliminary establish an operational business system—Crop Acreage Estimation by Using Remote Sensing and Sample Survey(Abbreviation :CAERSS).

In parallel with the ongoing crop sample surveys, the CAERSS is developed to produce the planted acreage estimation for major crops at provincial and county levels. By studying and borrowing a series of key programs which aims at the survey for land coverage/land use and crops acreage from the USA (LACIE,1974-1977; AGRISTARS,1980-1986; CDL,1997-2010), European Union (MARS,1998; LUCAS, 2001-2009; Geoland2, 2008(2011)) and ROK (Implement RS Application System, IRSAS, 2008-2012), the CAERSS has provided a feasible solution on acreage estimation for major crops such as wheat, corn, rice, cotton and soybean at provincial level and county level by using remote sensing and survey sampling method. The CAERSS has been taken the considerations of the characteristics and difficulties encountered to Chinese crop production such as plant scattered, complex planting structure, division by the south and the north, multiply planted crops in a year, frequently clouding and raining. Meanwhile the CAERSS has tried to fully take the advantages of the data source of remote sensing satellites for global earth observation. In practice, several provinces from the north to the south which are namely Heilongjiang, Jilin, Henan, Jiangsu, Hubei have been conducted the pilot surveys and achieved satisfied results. Compared with the conventional sampling surveys for crops, the CAERSS has gained significantly in objectiveness, timeliness and cost efficiency for crops acreage estimation, and it could cross validate the results from the ongoing crop sample surveys.

Taking the year 2011's autumn crop survey in Heilongjiang province with the support of remote sensing imagery as an example, this paper illustrates the following procedures: (1) Area frame construction and update based on extracted information of crop land and planted crops from remote sensing imagery. (2) Stratification and sample selection. (3) Field survey facilitated with PDA. (4) Major crops acreage estimation based on linear regression model.

2. Area Frame Construction and Maintenance

The area frame for agricultural statistics targets the full scope of arable land within the jurisdiction of a province, and it is constructed by non-overlap and non-omission grids to be used for sample selection. The area frame construction includes five procedures: preparation of remote sensing imagery, definition of spatial shape, definition of projection, remote sensing classification and crops discrimination, information update.

(1) Data preparation. Historic and real time remote sensing imagery with multi-source and multi temporal is prerequisite for cropland cover classification and planted crop discrimination. All the satellite imagery is pre-processed by radiometric correction, atmospheric correction and geometric correction.

(2) Spatial shape definition. From within the boundaries of the province as a whole population space, arable lands are delimitated by grids which are in size of approximately 3 km * 3 km. For those grids overlapped with the boundaries, the areas which is located out of the boundaries are eliminated, and the rest parts are merged

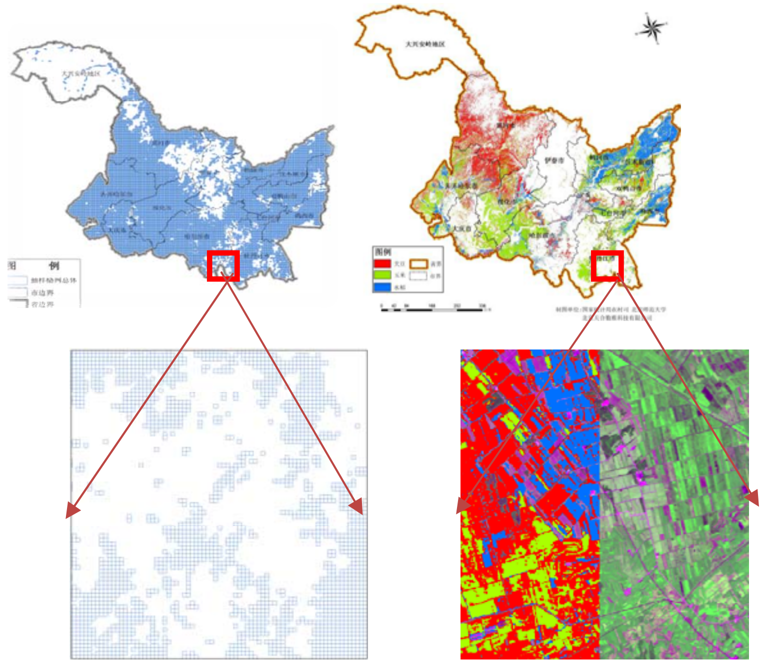
with adjacent grids within the scope of province by certain rules. In addition, the grids intersected with the county boundaries would be automatically delineated to each affiliated counties by merging and adjustment.

(3)Projection definition. The geo-spatial coordinates is defined, in our operational system of CAERSS we adopt the Albers Conic Equal Area projection and WGS-84 coordinate system.

(4) Remote sensing classification and crops discrimination. The classification of remote sensing imagery is a key technique and procedure to develop a sampling frame. The crops distribution is processed by using the Support Vector Machine (SVM) classification based on key phenological period of crops, supplemented by multi-temporal dynamic monitoring by fully use of historic and real-time remote sensing imagery. The extracted information includes the arable land, planted acreage of all crops, planted acreage by grain crops, spatial characteristic distribution from inverse model.

(5) Information update. After the complete construction of sampling frame, the maintenance is needed to annually update the information of land cover and planted acreage of major grain crops.

Within the boundary of Heilongjiang province, there were altogether more than 60 thousand grids in size of 3km*3km. In support of historic high resolution imagery (2.5m), those grids incapable of planting crops were eliminated so that approximately 45 thousand grids were finally left. The complete coverage of sampling frame was constructed by: 1)taking 144 scenes of Landsat TM imagery (time range from year 2010 to year 2011); 2)extracting the scope of arable lands, overall planted acreage and individual crop acreage by each county; 3) assessing classification precision by error matrix, and a threshold of overall precision above 80% was set; 4) constructing the sampling frame from extracted information associated with each grid.



(a) Definition of spatial shape of grid (b) Discrimination of planted crops

Figure 1 Area Frame Construction

3. Stratification and Sample Selection

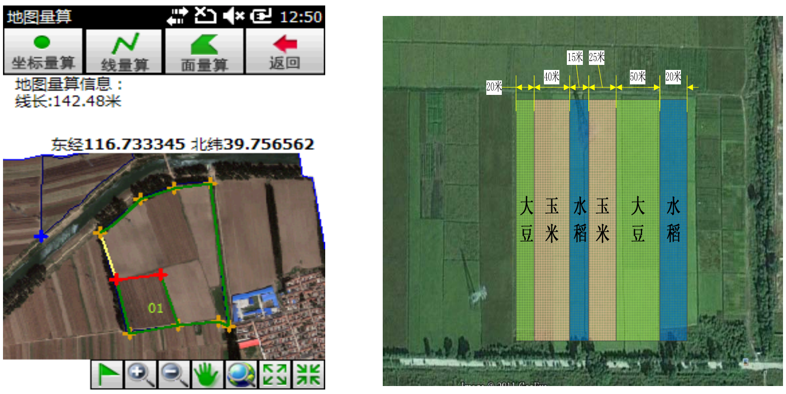
To select a good representative sample is a key procedure to gain the correction of crops classification from remote sensing imagery based on ground truth. Taking the advantages of information from the sampling frame for the purpose of stratification, we adopt the K-means cluster method by choosing multi-variables such as arable land and planted acreage of corn, rice, soybean. The result of stratification has gained the homogeneity within the stratum, and then a two stage sampling is applied.

In the first stage, a probability proportional to size of arable land is used to draw grids. There are 150 sampled grids in the first stage. In the second stage, 5 sample segments in size of 150m*150m each are selected from each sampled grid for ground survey. Altogether there are 750 segments were selected for ground survey. During the sample selection, we adopt the replicates technique to calculate the Mean Squared Error (MSE) as precision assessment.

4. Field Survey

The ground truth from the field sample survey plays an important role in crop acreage estimation. In order to improve the survey efficiency and accuracy, the CAERSS is developed two approaches to conduct survey, which includes: High resolution imagery assisted and Tablet PC assisted. There are pros and cons for each of these two survey approaches. The high resolution imagery make it easy to discriminate the various grain crops of samples, but the acquisition of the imagery could not be guarantee. The tablet PC assisted survey which integrated with GPS technology could meet the functions of navigation and real-time questionnaire fill-in. Due to its adaptiveness to survey operation under China’s official agricultural statistics, the tablet PC assisted approach has been widely used in regular business operation under the system of CAERSS .

On-site survey is conducted to collect land cover and planted acreage of crops at each of 750 sampled segments. The survey was facilitated with tablet PC (PDA) to navigate direction, and meanwhile high resolution satellite imagery is used as a base cartograph to be filled in the survey information. Please see Figure 2 as a result of field survey.



（a）Navigation by PDA （b）Results on base cartograph

Figure 2 Field Survey and Results

5. Acreage Estimation

By utilizing the survey data and remote sensing information from the sampling frame, we produced the estimates for corn, rice and soybean in three approaches for Heilongjiang province, which are direct expansion from sample, calibration by error matrix, linear regression respectively. Among the three estimation for crops acreage of corn, rice and soybean, the estimates from the linear model is relatively better in terms of coefficient of variation (CV).

The linear regression model is set as follows:

$$\hat{Y}_{reg} = \sum_{h=1}^{H} N_h [\bar{y}_h + \hat{b}_h (\bar{X}_h - \bar{x}_h)]$$

Where, $\hat{Y}_{\mathrm{reg}}$ is the planted acreage estimate for a certain crop,

N_h is the total number of segements in stratum *h*,

$\bar{y}_h$ is the mean acres per segment from the field survey in stratum h,

$\hat{b}_h$ is the coefficient from regression of acres on pixel counts in stratum h,

$\bar{X}_h$ is the mean categorized pixel count for all segments (scenes) in stratum h,

$\bar{x}_h$ is the mean categorized pixel count for sampled segments in stratum h.

The coefficient of variation (CV) of estimates by linear regression is given in Table 1.

Table 1 The CV of Acreage Estimates for Three Major Crops

Corn	**Rice**	**Soybean**
4.2%	5.1%	7.0%

6. Conclusions

In the context of Chinese characteristics of agricultural statistics and the capacity of remote sensing application, the system of CAERSS has broken through the difficulties of key techniques in crops discrimination and acreage estimation in face of either complete or incomplete remote sensing data. Take the year 2011’s autumn crop acreage survey in Heilongjiang province by using CAERSS as an example, the CAERSS has been tested its feasibility for crop acreage estimation.

With the development of remote sensing technology in China, the acquisition capability of multi-source remote sensing data especially high resolution satellite imagery is becoming more strengthening. National Bureau of Statistics(NBS) of China would take the full advantage of remote sensing application in agricultural statistics to yield more reliable acreage estimation for major crops. In the coming few years, NBS is going to gradually roll out the CAERSS to 13 major grain provinces in terms of their production and outputs throughout the nation.

References

[1] Delince, J. (2001) "A European approach to area frame survey", *Proceedings of the Conference on Agricultural and Environmental Statistical Applications in Rome*, Vol.2(XXV), 1-10.

[2]Gallego, F.J. (2004) "Remote sensing and land cover area estimation", *International Journal of Remote Sensing*, 25(15), 3019-3047.

[3] Stehman, S.V. (2009) "Model-assisted estimation as a unifying framework for estimating the area of land cover and land-cover change from remote sensing", *Remote Sensing of Environment*, 113, 2455-2462.

[4] Stehman, S.V., Sohl, T.L. and Loveland, T.R. (2003) "Statistical sampling to characterize recent United States land-cover change", *Remote Sensing of Environment*, 86, 517-529.

[5] Battese, G.E., Harter, R.M. and Fuller, W.A. (1988) "An error-components model for prediction of county crop areas using survey and satellite data", *Journal of the American Statistical Association*, 83, 28-36.

[6] Blaes, X. and Vanhalle, L. et al. (2005) "Efficiency of crop identification based on optical and SAR image time series", *Remote sensing of environment*, 96 (3-4): 352-365.

[7]Craig, M. (2009) *A Brief History of the Cropland Data Layer at NASS*[EB/OL], http://www.nass.usda.gov/research/Cropland/SARS1a.htm.

[8]Chauhan, H. J. and Arora, M. K. et al. (2008) "Estimating land cover class area from remote sensing classification", *Journal of Applied Remote Sensing*, 2 (023514), 1-12.

[9]Gallego, J. and Bamps C. (2008) "Using CORINE land cover and the point survey LUCAS for area estimation", *International Journal of Applied Earth Observation and Geoinformation*, 10 (4), 467-475.

[10] Benedetti R., Bee M., Espa G. and. Piersimoni F.(2010), Agricultural Survey Method, Wiley.