APCAS/14/14.1

February 2014

E

منظمة الأغذية والزراعة للأمم المتحدة | 联合国粮食及农业组织 | Food and Agriculture Organization of the United Nations | Organisation des Nations Unies pour l'alimentation et l'agriculture | Продовольственная и сельскохозяйственная организация Объединенных Наций | Organización de las Naciones Unidas para la Alimentación y la Agricultura

Asia and Pacific Commission on Agricultural Statistics

Twenty-fifth Session
Vientiane, Lao PDR, 18-21 February 2014
Agenda Item 14
Application of spatial information technology in Crop production survey in China

Contributed by: **Minghua Wang, Deputy Director-General,**
Department of Rural Surveys,
National Bureau of Statistics P.R. China
Email: wangmh@gi.stats.cn

Ge Wei
Consultant
Department of Rural Surveys,
National Bureau of Statistics P.R. China
Email: weige@gi.stats.cn

Promoting the Application of Spatial Information Technology in Crop Production Survey (Crop Surveys)

Wang Minghua   Wei Ge

The application of spatial information technology in Chinese agricultural statistics includes three aspects, namely, remote sensing, GIS and GPS, often called the "3S" technology. The core of the space information technology is information integration from the sky, ground and person, is a system of acquisition, management and analysis with spatial characteristics, complete information, accurate positioning.

In 2006, National Bureau of Statistics launched a national “863- project :national statistical sensing operational system key technology research and application", the crop production survey is the main contents of the project, other application aspects include population census , economics census and investment projects monitor survey. Spatial information technology applications in crop production survey, at present mainly in the following two aspects:

1.0 Crop Production Survey based on Area Sampling Frame

The reform of crop acreage and yield survey method aims to further improve the major crops such as wheat, rice, corn, cotton production data quality according to the National Bureau of Statistics on "three advances" (enhance up the quality of statistical data, enhance up the statistical capacity, enhance up the public credibility of official statistics), which is to strengthen the application of modern special information technology as well as in major grain (wheat, rice and corn) provinces to carry out the work required for crop surveys based on the experience of the pilot of crop production area frame sample survey in 2011 and 2012. The National Bureau of Statistics from 2013 onwards decided to officially begin in Jiangsu, Henan, Liaoning, Jilin, Hubei five provinces to implement crop production area frame sample surveys. The survey included sampling design, data collection, quality control and assessment, estimation as follows.

1.1 Sampling design

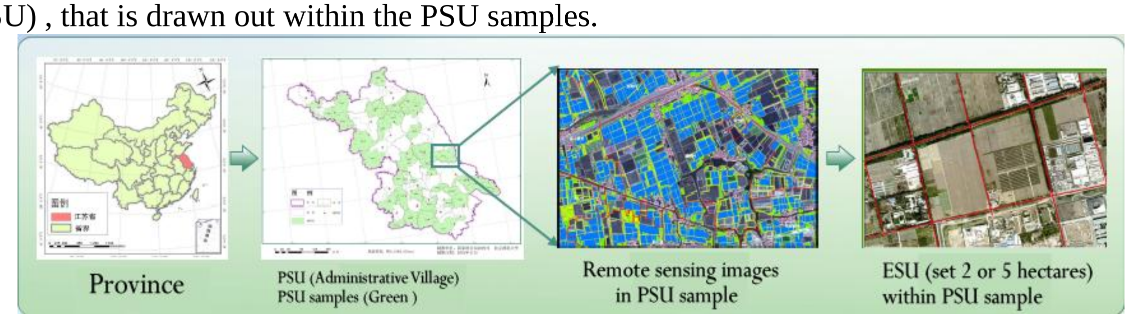
1.1.1 Preparation of the area sampling frame

Source of preparing area sampling frame has the following aspects. First, the second national agricultural census data (2006), which provides a more complete crop information, but there is a problem in the non-spatial and timeliness of the data. Second, the second national land survey results, which provides close to complete spatial information of the cultivated land, but there is a lack of crop information. Third, the high-resolution remote sensing images can produce more update major crop information, but its coverage do not meet the need, while coarse and medial resolution remote sensing images may meet the coverage requirement but lose the accuracy of the crop information identifying. Therefore the remote sensing images are not used directly, but they can be used as auxiliary variables in sample design. So take planting crops according to local circumstances, will combine several sources of information to each other, learn from each other to carry out a comprehensive and prepare the area sampling frame by provinces.

Pic1. Resolution remote sensing images

1.1.2 Formation of the sampling unit

The area unit for crop production survey should be directly investigated, namely elementary area sampling unit (ESU). Currently more direct investigation is the use of on site investigation, so the area unit should not be too large. The optimal unit size is taking into account sampling variance of different survey variables when comparing different size units, proportion of non-zero value, the survey costs, investigator daily workload, the physical boundaries of availability pertinent to area unit on the map. After considering these factors, we set Jiangsu, Hubei area unit by two hectares (30 “mu”) ; Liaoning, Henan, Jilin area unit by five hectares (75 “mu”). Under the current conditions of the sampling frame data, not directly to the area unit as the preparing unit to the sampling frame, instead of using a multi-stage approach, beginning from the larger unit by stage preparation. Preparation of primary cell sampling frame, within the overall sample space area to the administrative boundary of space is divided according to the administrative villages, the use of the second national agricultural census villages crop data, combined with the current or recent situation of high resolution remote sensing images and the second national land survey results for calibration, convergence to the preparation of the administrative villages as a primary sampling unit(PSU). By the way, instead to the administrative villages we set the spatial information grid while enough of the current or recent (2 or 3 years before) situation of high resolution remote sensing images. The Elementary area sampling unit(ESU) , that is drawn out within the PSU samples.



Pic2. Setting PSU and ESU

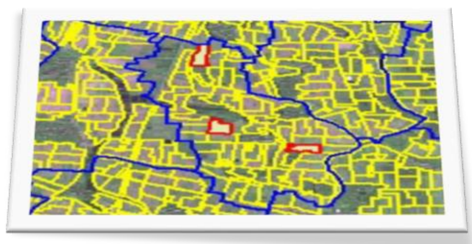
1.1.3 Random sample selection by probability

1.1.3.1 Selection of sown area survey sample

A multi-stage sampling method is adopted to select acreage survey sample, specific requirements are as follows :

1) The first stage, with the method of PPS (the probability proportional to cultivated land area) by farmland stratified selected sample of PSU (administrative villages or spatial information grid), requires investigation crops (total sown area, the total sown area of grain, the main crops sown) with relative error coefficient less than 5% , the number of samples is calculated on the basis of workload and cost affordability.

2) The second stage, within PSU sample, by simple random sampling (SRS) method selected elementary area sampling uni t(ESU) as the sown area survey sample. Detailed steps are: within the PSU (administrative villages or spatial information grid) drawn out the elementary area sampling unit on the spatial vector diagram, then sequentially encoded by using table of random numbers, random select three or five elementary sampling unit as the sown area survey sample.



Pic3 ESU (yellow) and ESU samples (red)

1.1.3.2 Selection of the yield per sown area survey sample

Take two-phase sampling method selected sample for yield per sown area survey base on the sown survey sample. Steps are as follows :

(1) Data collection for yield estimation . In the crop production harvest season, collect data for yield estimation from all crop sown area survey samples.

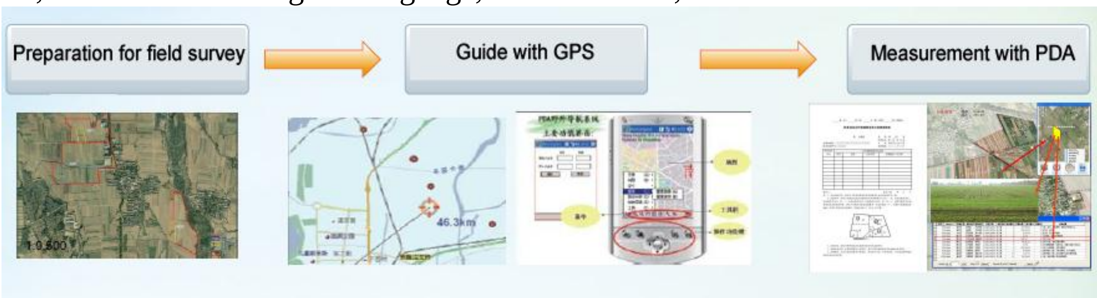
(2) The yield survey sample was drawn. For target survey crop (wheat, rice, corn, etc.), sorted by the target crop estimation result, with 1 / 3 sampling ratio, a random starting point, the final number of sown area survey samples of area expansion weight is used to determine the cumulative expansion of acreage, equidistant selected yield per area survey sample. Relative error coefficient investigation crops controlled within 2 %, if the sample 1/3 ratio not meet the need, should be expanded the appropriate sample size.

(3) Place the measured samples. Within yield per area survey sample (the ESU or if the ESU too large and complex then random sampling sub ESU, called “segment”), evenly placed 3-5 real measured cutting small samples(usually one square meter), by the straight line, equilateral triangle, required number of samples within each survey must be consistent, the area must be consistent with the measured sample.

1.2 Data collection

Include two aspects content: one is “one time survey”, at the beginning survey, include the ESU samples finding, position, and measurement with the GPS and PDA, named ESU samples registration. Another is “seasonal survey”, for each target crop sown and harvest season, include target crop sown area and yield per sown area with PDA and measurement tools.

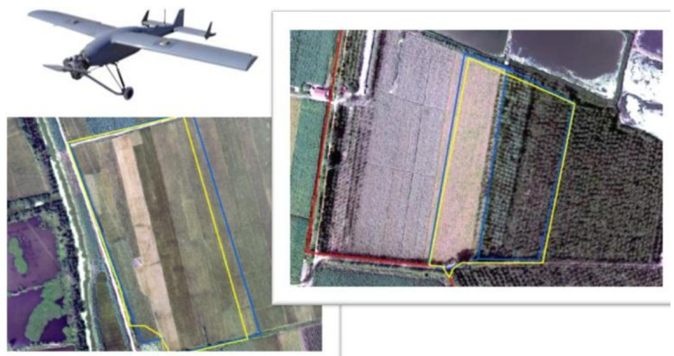
For the yield per sown area of field survey, still using the traditional method, in the target crop harvest the investigators place the real measured cutting small samples and cut real measurement by hand, the main measuring tool is gauge, mini threshers, scales and so on.



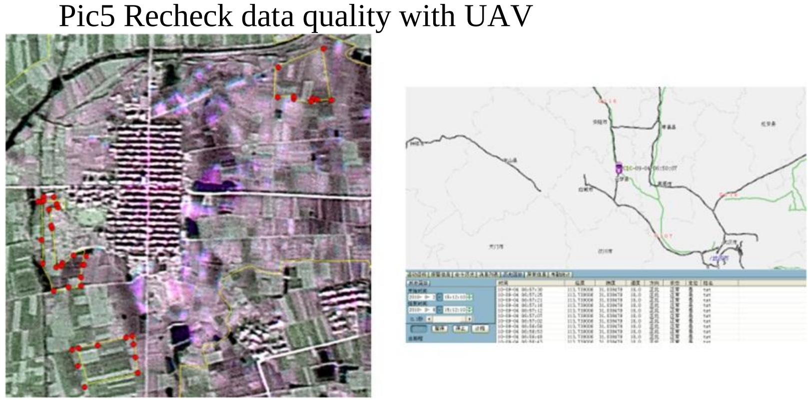
Pic4 Field survey with GPS+PDA

1.3 Data quality control and assessment

In the process of survey, the UAV(Unmanned Aerial Vehicle) is used for recheck sown area of ESU samples; positioning monitoring system is used for timely survey direction, harmonization, control and supervision; also the digital images of field give a new resources and method for data quality control and assessment.



Pic5 Recheck data quality with UAV



Pic6 Timely survey control and supervision with GPS and GSM

Sampling error assessment to all results of target crop sown area calculated using the jackknife method. Specific steps are as follows:

1) The sample stratified by village groups to sort through the equidistance method the sample is then assigned to R non-overlapping groups.

2) To each group (identified by r) the complement of a knife called a first group of sub-samples (Replicate).

3) For the sample layer h i, is defined

$$w_{i(r)} = \begin{cases} w_i n_h /(n_h - n_{hr}) & \textit{when ”i” in the “r”th jackknife subsample} \\ 0 & \textit{other (when ”i” in the “r”th group)} \end{cases}$$

4) When a continuous statistic variable $\hat{\theta}$ is calculated by the w_i estimator, the estimated amount of each sub-sample $\hat{\theta}_{(r)}$ by a set of its own $w_{i(r)}$ calculated the packet Jackknife (DAG) MSE estimator is:

$$\hat{V}_{DAGJK}(\hat{\theta}) = [(R-1)/R]\sum^{R}(\hat{\theta} - \hat{\theta}_{(r)})^2$$

Like the above formula as MSE estimators usually called "Packet Jackknife estimator", usually set R = 15, the confidence interval of 14 degrees of freedom using the Student's t distribution.

Accordingly, the relative standard error for the sample as follow:

$$CV(\hat{\theta}) = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}}.$$

After the actual test, the CV of target crop sown are less than 5% .

1.4 Method of estimation

For target crop sown area, with each ESU sample ‘s extension weight which is reciprocal of the sampling probability , calculate the target crop sown area. Formula is as follows:

$$\hat{X} = \sum_{i=1}^{m}\sum_{j=1}^{n_i} x_{ij} w_{ij} w_i$$

In the formula, x_{ij} is survey target crop sown area in ESU samples, and w_{ij} is the extension weight from ESU sample to PSU sample, and w_i is the extension weight from PSU sample to target population (province).

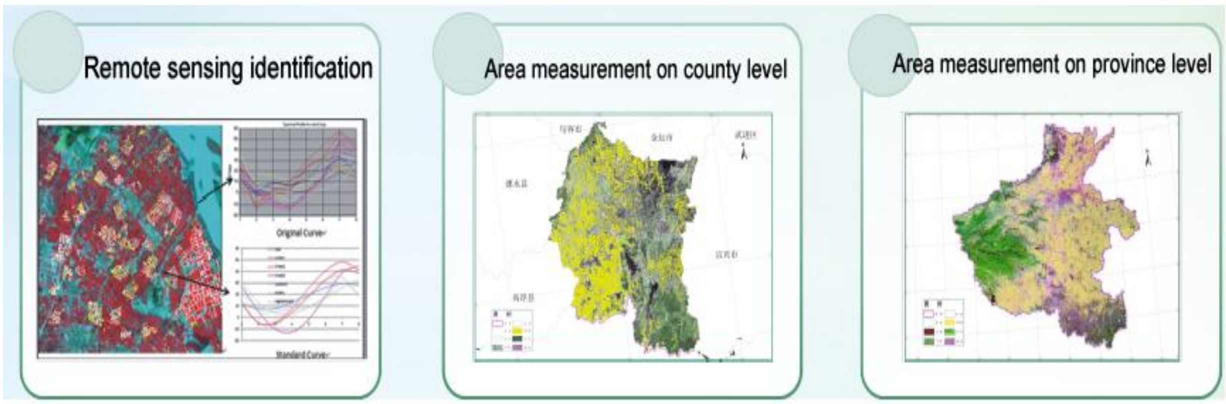
For target crop yield per sown area ,we use simple average of ESU real cutting and measurement. Formula is as follows:

$$\hat{\bar{y}} = \frac{1}{n}\sum_{i=1}^{n} \bar{y}_i$$

2.0 Remote sensing of crop area measurement and yield trend estimation

Facing the problems of crop data interference links exist in statistical investigation, data acquisition cycle is long, the investigation could not cover all, through the application of modern remote sensing technology, exploration of wheat, rice, cotton and other crops of the whole process, full coverage of the monitoring.

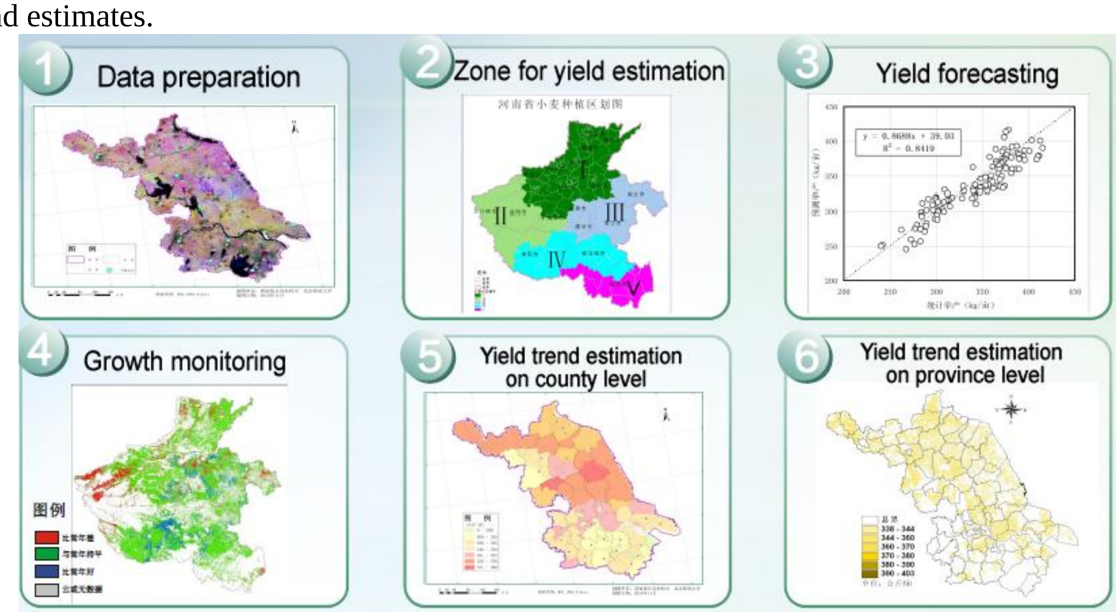
In support of framework with spatial technology and field survey system, the establishment of major crop plant area remote sensing systems business, including: standardization of data acquisition and processing, remote sensing identification, sampling verification, field survey, measurement correction and formed remote sensing measurement products.



Pic6 Remote sensing crop area measurement

From 2010 to now, in Beijing, Jilin, Henan, Jiangsu, Hubei and Shandong provinces to carry out remote sensing of crops. Currently, Beijing has grain acreage into agricultural remote sensing survey system, forming a regular operation. Starting in 2011, carried out in Xinjiang cotton remote sensing measurements.

Establishment of crop yield monitoring and remote sensing systems, including: data preparation, growth monitoring and yield forecasting, production estimates at the county level, provincial yield trend estimates.



Pic7 Crop yield monitoring and remote sensing

3.0 Effect and prospects of spatial information technology application

Through the application of research and pilot projects, "3S" technology to improve the sampling frame, improved sampling design, improve the quality of data acquisition technology and etc, presents the application effect of spatial information technology in agricultural statistics in business process effect and prospect.

1) Reference to improve the sampling frame with the low-middle resolution spatial Image. The wide coverage, time acquisition is mainly derived from satellite remote sensing images and census data, survey data fast combination, so that the a great quantity of census, the survey data play a

greater role in basic research, as the structure unit (administrative village, grid etc.) spatial sampling frame provides the possibility, also greatly reduces the cost.

2) Reference to improve the sampling design with high resolution spatial image. With satellite, aircraft, UAV(Unmanned Aerial Vehicle) as reference for sampling unit division, in order to improve the sampling design, provide auxiliary information beneficial to improve the sampling efficiency.

3) Based on GPS to improve data collection technology. Investigation of equipment based on the GPS technology application, combining the space image and electronic map data, not only can complete survey route design and goal seeking on field, but also locate the sample unit, area boundary outline, area measurement work etc. to enhance the comprehensive ability of data collection.

4) Real-time monitor for field investigation in combination with spatial and network communication technology. The combination of spatial information and network, communication technology, carry out investigation and field data collection, real-time monitoring of the field investigation, to improve the quality management level and efficiency of field survey.

The spatial information technology in stereo the global earth observation system and the technical means for the characteristics of the development, application field of the scale, industry continues to expand, already became to agricultural statistics reform motivation in the statistical methods, techniques and management system " change with the times ", catch up with the advanced international level, to achieve the main driving force of development by leaps and bounds.