Brazil Soybean Area Estimation Based on Average Samples Change Rate of Two Years and Official Statistics of a Year Before

Kejian Shen(✉), Weifang Li, Zhiyuan Pei, Fei Wang, Xiaoqian Zhang, Guannan Sun, Jiong You, Quan Wu, and Yuechen Liu

Chinese Academy of Agricultural Engineering, Innovation Team of Crop Monitoring by Remote Sensing, Beijing Chaoyang District Maizidian Street No. 41, Beijing 100125, China
ashenkejian@126.com, guannan_sun@126.com, wfli1988@aliyun.com, {peizhiyuan,wangfei, zhangxiaoqian,liuyuechen}@agri.gov.cn,

Abstract. Comprehensive, reliable and timely information of Brazil's soybean area is necessary for China to make decisions on agricultural related problems. Spatial sampling method which combined remote sensing and sampling survey is widely used. Due to limitations of width and revisit cycle of medium resolution satellite, This study designed a typical investigation method about Brazil soybean area based on average samples change rate of two years and official statistics of a year before, typical samples were selected to survey, sampling frame was constructed on soybean planting state, the sampling unit was designed as 40 km × 40 km, the sampling proportion was 2 %, average samples change rate of two years were 2013 and 2014. Estimated area was compared with Brazil official harvested area in 2014 (published on 2015 April by Brazilian Institute of Geography and Statistics), the relative error is 2.37 %.

Keywords: Brazil soybean area estimation · Change rate of two years · Official statistics of a year before · Sampling

1 Introduction

Chinese soybean planting area has decreased year by year, soybean self-supply ability had been to about 20 % in 2013, mainly imports country are United States, Brazil, and Argentina [1]. Comprehensive, reliable and timely information of Brazil's soybean area is necessary for China to make decisions on agricultural related problems. Compared with the traditional survey method, Remote sensing survey has advantage of large coverage, low cost and less investigation time [2]. Spatial sampling method which combined remote sensing and sampling survey is widely used in the investigation of large scale crop area estimation [3]. The survey accuracy is mainly effect by Population, sampling proportion, sample distribution. Population can be defined by

D. Li and Z. Li (Eds.): CCTA 2015, Part I, IFIP AICT 478, pp. 366–374, 2016.
DOI: 10.1007/978-3-319-48357-3_36

historical cultivated land [4] or administrative divisions with historical statistical data [5]. Sampling proportion is the bigger the better with the premise of meeting the minimum sampling proportion, but the actual survey should consider the accuracy requirements, cost and time. The ideal sample distribution is random distribution, but it is subject to the satellite width limitation and revisit cycle.

With the rapid development of remote sensing technology, medium resolution satellite (10 m–30 m) are gradually meet the sampling survey requirements and even full coverage. But in the soybean growth period, it is difficult to get full coverage image with cloud free. Considering the weather, satellite width limitation and revisit cycle, using landsat7/8 as a data source, this study designed a typical investigation method about Brazil soybean area estimation based on average samples change rate of two years and official statistics of a year before.

2 Study Area and Data Source

2.1 Study Area

Brazil is located in the west by 35 to 74°, 5° north latitude to 35° south latitude. Brazil's total area is about 8514900 square kilometers, which is about 46 % of the South America total area. The terrain of Brazil is divided into two parts, one part is plateau of Brazil with altitude of 500 m above, located in the south of Brazil, the other part is plains with elevation of 200 m below, mainly distributed in the Amazon River Basin in the north and the west. Throughout the terrain is divided into the

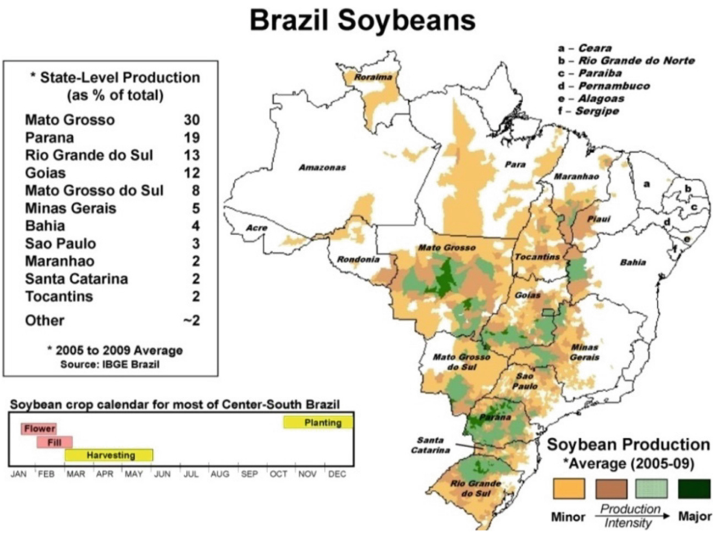

Fig. 1. The sketch map of study area (http://www.usda.gov/oce/weather/pubs/Other/MWCACP/Graphs/Brazil/BrzSoyProd_0509.pdf)

Amazon plain, Paraguay basin, Brazil and the Guyana plateau, the Amazon plain area accounting for about 1/3. Most of the area of Brazil belongs to the tropical climate, parts of the South belongs to the subtropical climate. Annual average temperature of The Amazon plain is 25～28°, the annual average temperature of south is 16 to 19°.

Soybean mainly distribute in central Brazil and southern Brazil (Fig. 1), Due to the tropical climate and long growing season, the crop production cycles are much more complicated. Below is a month-by-month (Table 1) account of what to expect during the growing season.

Table 1. Brazil soybean month-by-month crop cycle

September	• Early soybean begins in Mato Grosso and central Brazil.
October	• Soybean planting in full swing in southern Brazil.
November	• Early November is main planting period.
December	• Finish planting, early-planted soybeans flowering and setting pods. • Begin spraying to control soybean rust.
January	• Soybeans flowering and setting pods. • Some very early soybeans in central Brazil may be harvested this month. • Continue spraying to control soybean rust.
February	• Main pod filling month. • Early soybeans being harvested. • Soybean rust control now focused on later maturing soybeans.
March	• Main soybean harvesting month. • Critical time for soybean rust to affect late maturing soybeans.
April	• Finish soybean harvest.
May	• Rains have ended in central Brazil and dry season has started. • Scattered rains continue to fall in southern Brazil.
June-July-August	• This is the dry season in central Brazil. • Occasional rains can occur in southern Brazil.

From: http://www.soybeansandcorn.com/Brazil-Crop-Cycles

2.2 Data

Landsat Multi-spectral image: Landsat7 and landsat8 Multi-spectral image listed as Table 2 were used. Landsat7 and landsat8 were subset by sampling frame of 40 km × 40 km, only cloud free samples were selected.

Soybean statistical data are downloaded from website of The Brazilian Institute of Geography and Statistics[1], which publishes harvests figures consisting of area, output and average yield for 35 different crops of previous year in the annual 1–4 month.

[1] http://www.ibge.gov.br/english/estatistica/indicadores/agropecuaria/lspa/default_publ_completa.shtm.

Table 2. Landsat multi-spectral image

Path/row	Number of samples	Data	Acquired time	Purpose
221/78	16	Lansat7	10-jan-13	2013 soybean extract
		Lansat8	21-jan-14	2014 soybean extract
221/79	6	Lansat7	10-jan-13	2013 soybean extract
		Lansat8	21-jan-14	2014 soybean extract
222/79	14	Lansat7	17-jan-13	2013 soybean extract
		Lansat8	28-jan-14	2014 soybean extract
222/81	9	Lansat7	06-mar-13	2013 soybean extract
		Lansat8	28-jan-13	2014 soybean extract
223/77	11	Lansat7	24-jan-13	2013 soybean extract
		Lansat8	19-jan-14	2014 soybean extract
224/68	8	Lansat7	30-dec-12	2013 soybean extract
		Lansat7	16-feb-13	
		Lansat7	13-apr-13	
		Lansat8	02-jan-14	2014 soybean extract
		Lansat8	27-feb-14	
		Lansat8	08-apr-14	
224/69	5	Lansat7	30-dec-12	2013 soybean extract
		Lansat7	16-feb-13	
		Lansat7	05-apr-13	
		Lansat8	02-jan-14	2014 soybean extract
		Lansat8	27-feb-14	
		Lansat8	08-apr-14	
225/75	14	Lansat7	21-dec-12	2013 soybean extract
		Lansat7	07-feb-13	
		Lansat7	20-apr-13	
		Lansat8	30-nov-13	2014 soybean extract
		Lansat8	02-feb-14	
		Lansat8	07-apr-14	

Soybean official harvested area in 2013 is 27736 thousand hectares, which is used for are estimation. Soybean official harvested area in 2014 is 30241 thousand hectares, which is used for accuracy assessment.

3 Methodology: Sampling Design

The flow of this experiment (Fig. 2) includes: (1) Construction of sampling frame; (2) Determine the Sampling proportion and distribution of samples, (3) Soybean extraction by unsupervised classification and visual interpretation; (4) The average change rate of samples between 2013 and 2014; (5) Area estimation; (6) Accuracy assessment.

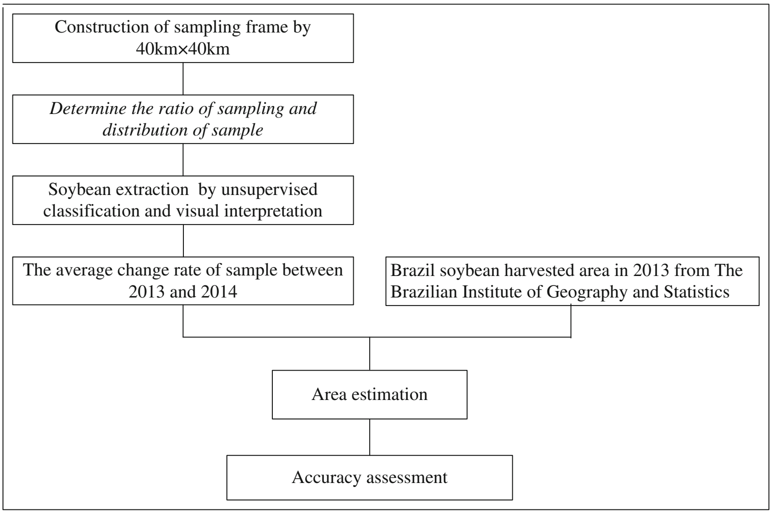

Fig. 2. Flow chart of experiment

3.1 Construction of Sampling Frame

The sampling frame covers 17 soybean planting states (from 2013 Brazil official statistics), which can be seen from Fig. 3, the sampling unit was designed as 40 km × 40 km. The population is 4215, which is shown on Fig. 3.

Fig. 3. The sketch map of image samples

3.2 Determine the Sampling Proportion and Distribution of Samples

We considered relevant research to determine the sampling proportion. The rate of sampling of Monitoring Agriculture with Remote Sensing (MARS) of European Union is about 1 % (60sites × 40 km × 40 km/10160000 km^2) [6]; The rate of paper about paddy rice area estimation using a stratified sampling method with remote sensing in China is 1.3 % [7]. In order to improve the estimation accuracy, the rate of sampling is increased to 2 %. The number of samples is 83. Selected samples could be seen from Fig. 3.

Distribution of samples should consider several factors. Firstly, the samples should cover the major and minor soybean planting areas; Secondly, in the soybean growth season, the images can effectively extract the spatial distribution of soybean. Thirdly, sample is cloud free. The samples are shown in Fig. 3.

3.3 Soybean Extraction by Unsupervised Classification and Visual Interpretation

Each sample has soybean classification results of 2013 and 2014, samples were classified by unsupervised classification. The classification results were corrected by visual interpretation of ArcGIS software. Statistics of samples classification results are shown as Table 3.

3.4 The Average Change Rate of Samples Between 2013 and 2014

$$Change\_rate = (Area2014_{sample\_i} - Area2013_{sample\_i})/Area2013_{sample\_i} \quad (1)$$

where change_rate represents change rate of samples between 2013 and 2014, sample_i represents sample number, i from 1 to 83. The average change rate of 83 samples between 2013 and 2014 is 6.45 %, which is shown on Table 3.

3.5 Area Estimation

$$\hat{P} = Area2013_{official} \times (1 + Average\, change\_rate) \quad (2)$$

where $\hat{P}$ represents estimated area; $Area2013_{official}$ represents Brazil soybean statistical data of 2013, which are downloaded from website of The Brazilian Institute of Geography and Statistics.

3.6 Accuracy Assessment

Sampling results was appraised by relative error r, which is defined as follow:

Table 3. Statistics of samples classification result

Sample	Path/row	Area13	Area14	Rate of change	Sample	Path/row	Area13	Area14	Rate of change
1	221/78	226	195	−14.00 %	43	222/81	375	408	8.75 %
2	221/78	66	60	−9.19 %	44	222/81	321	350	8.84 %
3	221/78	138	130	−5.59 %	45	222/81	15	55	273.35 %
4	221/78	162	159	−2.21 %	46	223/77	422	424	0.46 %
5	221/78	516	515	−0.33 %	47	223/77	621	646	3.95 %
6	221/78	281	289	2.96 %	48	223/77	92	96	4.48 %
7	221/78	67	69	3.56 %	49	223/77	210	226	7.64 %
8	221/78	382	398	4.09 %	50	223/77	574	626	9.08 %
9	221/78	190	201	6.09 %	51	223/77	437	479	9.41 %
10	221/78	325	349	7.21 %	52	223/77	517	644	24.65 %
11	221/78	223	239	7.38 %	53	223/77	488	624	27.99 %
12	221/78	404	435	7.48 %	54	223/77	687	942	37.15 %
13	221/78	343	371	8.04 %	55	223/77	82	157	91.33 %
14	221/78	356	386	8.45 %	56	223/77	15	34	120.32 %
15	221/78	241	265	9.98 %	57	224/68	264	188	−28.87 %
16	221/78	231	262	13.40 %	58	224/68	34	26	−24.84 %
17	221/79	23	9	−61.04 %	59	224/68	420	379	−9.69 %
18	221/79	94	53	−43.82 %	60	224/68	566	538	−5.05 %
19	221/79	263	162	−38.50 %	61	224/68	515	497	−3.59 %
20	221/79	454	328	−27.79 %	62	224/68	353	366	3.89 %
21	221/79	80	59	−26.40 %	63	224/68	58	61	5.55 %
22	221/79	303	332	9.46 %	64	224/68	55	66	18.95 %
23	221/79	78	47	−39.41 %	65	224/69	795	736	−7.39 %
24	221/79	146	120	−18.30 %	66	224/69	485	470	−3.07 %
25	221/79	35	35	0.32 %	67	224/69	1295	1300	0.32 %
26	221/79	725	734	1.24 %	68	224/69	288	309	7.32 %
27	221/79	326	332	2.11 %	69	224/69	144	164	14.49 %
28	221/79	12	13	8.66 %	70	225/75	513	372	−27.43 %
29	221/79	293	318	8.69 %	71	225/75	513	429	−16.41 %
30	221/79	526	573	8.92 %	72	225/75	675	637	−5.61 %
31	221/79	801	878	9.70 %	73	225/75	8	8	−3.38 %
32	221/79	253	281	10.96 %	74	225/75	798	779	−2.41 %
33	221/79	289	335	15.90 %	75	225/75	282	278	−1.28 %
34	221/79	594	695	17.15 %	76	225/75	401	405	1.07 %
35	221/79	78	93	19.55 %	77	225/75	23	23	2.61 %
36	221/79	227	314	38.14 %	78	225/75	388	399	2.81 %
37	222/81	352	348	−1.39 %	79	225/75	376	388	3.23 %
38	222/81	101	103	1.61 %	80	225/75	117	123	4.94 %
39	222/81	186	193	3.78 %	81	225/75	585	614	5.05 %
40	222/81	290	304	4.89 %	82	225/75	201	212	5.29 %
41	222/81	260	275	5.46 %	83	225/75	195	208	6.76 %
42	222/81	193	207	7.56 %	Average change rate of samples				6.45 %

$$r = 100 \times (P - \hat{P})/P \quad (3)$$

where r represents relative error, $\hat{P}$ represents estimated area, P represents true area, Brazil soybean harvested area in 2014 are used.

4 Results

Soybean official harvested area in 2013 is 27736 thousand hectares, average change rate between 2013 and 2014 is 6.45 %, Estimated soybean harvested area in 2014 is:

$$\hat{P} = Area2013_{official} \times (1 + Average\,change\_rate) = 27736 \times (1 + 6.45\ \%)$$
$$= 29525 \text{ thousand hectares}$$

Soybean official harvested area in 2014 is 30241 thousand hectares, relative error is:

$$r = 100 \times (P - \hat{P})/P = 100\ \% \times (30241 - 29525)/30241 = 2.37\ \%.$$

5 Discussion and Conclusion

Discussion: in the previous study of stratified sampling, stratified variable often from Modis data or land use/cover data or statistical data, the location of sample is determined by stratified variable. However, the number of sample with determined location is hard to be satisfied with cloud free image. A question then worth asking is How to maximize the use of available images?

This study designed a typical investigation method about Brazil soybean area based on average samples change rate of two years and official statistics of a year before, typical samples were selected to survey, sampling frame was constructed on soybean planting state, the sampling unit was designed as 40 km × 40 km, the sampling proportion was 2 %, average samples change rate of two years were 2013 and 2014. Estimated area was compared with Brazil official harvested area in 2014 (published on 2015 April by Brazilian Institute of Geography and Statistics), the relative error is 2.37 %.

Acknowledgment. Funds for this research was provided by National Natural Science Foundation of China (No. 41301506) and Civil Space Project in the 12th Five-Year Plan of China (2011–2015).

References

1. Liu, Z.: Some thoughts concerning development strategy for soybean industry in China. Soybean Sci. **32**(3), 283–285 (2013)
2. Ahmed, R., Sajjad, H.: Crop acreage estimation of Boro Paddy using remote sensing and GIS techniques: a case from Nagaon district, Assam, India. Adv. Appl. Agric. Sci. **3**(3), 16–25 (2015)
3. Huajun, T., Wu, W., Yang, P., et al.: Recent progresses in monitoring crop spatial patterns by using remote sensing technologies. Sci. Agric. Sin. **43**(14), 2879–2888 (2010)
4. Wu, B.-F., Li, Q.-Z.: Crop acreage estimation using two individual sampling frameworks with stratification. J. Remote Sens. **8**(6), 551–569 (2005)
5. Chen, Z., Liu, H., Zhou, Q., et al.: Sampling and scaling scheme for monitoring the change of winter wheat acreage in China. Trans. Chin. Soc. Agric. Eng. **16**(5), 126–129 (2000)
6. Gallego, F.J., Delincé, J., Carfagna, E.: Two-stage area frame sampling on square segments for farm surveys. Surv. Methodol. **20**(2), 107–115 (1994)
7. Jiao, X., Yang, B., Pei, Z., et al.: Paddy rice area estimation using a stratified sampling method with remote sensing in China. Trans. Chin. Soc. Agric. Eng. **22**(5), 105–110 (2006)