AREA ACCURACY ASSESSMENT FOR
CHINESE LAND-USE DYNAMIC DETECTION OF REMOTE SENSING PROGRAM

SUN Xiao-xia[a] ZHANG Ji-xian[a,b] LI Hai-tao[a]

[a] Chinese Academy of Surveying and Mapping, Haidian District Beijing 100039, sun.xiaoxia@163.com,lhtao@163.net
[b] Lab. Of Land Use, The Ministry of Land Resources, Haidian District Beijing, 100029, stecsm@public.bta.net.cn

KEY WORDS: Remote sensing, Land Use, Change Detection, Sampling, Accuracy , Monitoring

ABSTRACT:

A suitable area accuracy assessment scheme for change detection is presented.The technique consists of two parts which are the necessary stages of accuracy assessment:(1)sampling method and (2)accuracy assessment. A area stratified random sampling is applied based on the fact that area errors is related to a certain area range (called stratum in following sampling). The approach involves subdividing the area range into strata,and within each stratum a spectific number of sample polygons are randomly chosen. Three experiments about polygons of remote sensing monitoring have been done which aim at finding out the distribution law of area error,the conclusions that the error of polygon area follows normal distribution and the mean is zero are drawed.This give accuracy assessment of remote sensing-derived change detection a science foundation. Based on relative error theory and error spread law ,three area accuracy indices including relative root mean square,relative error,and average relative error are built.Using these indices,the formulas of single polygon error and the whole monitoring region error are produced. Experimental results and practical work show that the method in this paper is simple, effective and applicable.

1. INTRODUCTION

1.1 General Instructions

The Ministry of Land and Resources P.R.China is conducting a continuous Land use Dynamic Detection project using Thematic Mapper(TM) 30-meter and SPOT(10-meter or 5 -meter)data.The objective of this large project is providing the government a generalized, consistent, and reliable land use change data every year. The accuracy of product is very important not only for users but also for makers, and how to assess the accuracy of remote sensing-derived change detection is the focus problem currently.

The rest of the paper is orgnized as follows.We carry out the experiments about area error of polygons of remote sensing monitoring in section 2.In section 3,We present area stratified random sampling technique and an approach to calculate sample size. In section 4, the area accuracy indices are built and an example of area accuracy assessment is given in section 5.In section 6, We conclude with discussions.

2. EXPERIMENTS ON ERROR DISTRIBUTION

Three experiments are made as follows, aiming at finding out the distribution law of area error.

2.1 Histogram Experiment

196 polygons from the monitoring results using radarsat image of BeiJing region are sampled. We divided the sample polygons into four stratums according to their areas ,there are 75 ones under 10mu[*], 43 ones within 10~20 mu, 50 ones within 20~50 mu, 28 ones above 50 mu. In each stratum , relative error is calculated.The histogram of relative error is shown in Figure 1. Clearly, Fig1 displays normal distribution law of area errors in every area range.
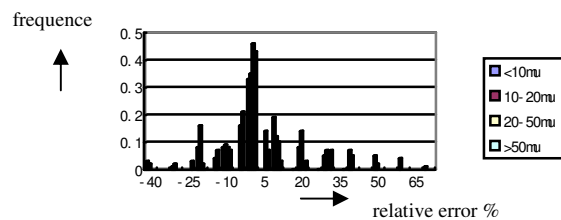
---

* 1 acre = 6.075 mu.

Figure1. The histogram of relative error.

2.2 Mean Test

In order to test whether the mean of relative error is zero, or whether the error is systematic, we make a hypothesis testing on monitoring polygons from a variety of images. We use normal population test technique and construct statistic:

$$T = \frac{\overline{x} - \mu_0}{S/\sqrt{n}} \qquad (1)$$

original hypothesis $H_0 : \mu = \mu_0 = 0$, alternative hypothesis $H_1 : \mu \neq 0$, rejective range: $|T| \geq t_{\alpha/2}(n-1)$, α 0.01. The result is shown in Table 2.

image	n	μ	S	T	$t_{\alpha/2}(n-1)$
SPOT(10m)	25	0.03683	15.9%	1.1553	2.7969
IRS(5.8m)	26	0.03508	11.5%	1.5552	2.7874
IKONOS(1m)	40	0.01206	4.6%	1.6444	2.7079
SAR(6.25m)	31	0.04300	16.3%	1.4688	2.7500

Table2. The result of mean test.

We can see that all the statistics in table1 are less than the threshold, the conclusion is drawed that the error mean is zero. That is to say , the error is casual, not systematic.

2.3 Error Distribution Test

We make a χ^2 test for proving the normal distribution of the error of polygon area.

First, we divide them into k groups by the range of these relative error, and count the number of samples in each group. Secondly, Sample mean and mean square root error are calculated. Thirdly, supposing H_0 is true, we calculate groups probability p_i and theroy frequence np_i, and at the last, we calculate χ^2 using the following formula

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i} \qquad (2)$$

	f_i	p_i	np_i	$f_i - np_i$	$(f_i - np_i)^2 / np_i$
x≤-15	1	0.0099	0.8118		
-15<x≤-10	2	0.0386	3.1652	-1.407	0.15
-10<x≤-5	9	0.1150	9.43		
-5<x≤0	18	0.2186	17.9252	0.0748	0.00
0<x≤5	29	0.2622	21.5004	7.4996	2.62
5<x≤10	16	0.2088	17.1216	-1.1216	0.07
10<x≤15	1	0.1055	8.651		
15<x≤20	4	0.0336	2.7552	-5.0786	2.14
x≥20	2	0.0082	0.6724		
	82				4.98

Table3. The result of error distribution test.

The result is shown in table 3. This table displays $\chi^2_{0.05}(k-r-1) = \chi^2_{0.05}(5-2-1) = 5.991 > 4.98$ so we accept H_0, that is to say, the errors of polygonal area follow normal distribution and their mean is zero.

3. SAMPLE TECHNIQUE

3.1 Sample Design

The area accuracy of monitoring polygons depends on change detection techniques and image resolutions.It's found that area errors have a regular distribution in a certain area range (called stratum in following sampling)., so a area stratified random sampling is applied. The approach involes subdividing the area range into strata,and within each stratum a spectific number of sample polygons are randomly chosen.When using LANDSAT TM(30meter) and SPOT(10meter) images,we usually divide the whole polygons into four stratums: under 10mu, 10~20 mu, 20~50 mu and above 50 mu,which have been proved reliable and reasonable.

3.2 Sample Size

Because of the normal distribution, for a large sampling population in which the sampling fraction n/N is negligible,in another word, n/N does not exceed 5 or 10 percent (Cochran,1977,p.25), we use sampling size equation based on OC function. The equation has the form

$$\sqrt{n} \geq \frac{(z_\alpha + z_\beta) \cdot \sigma}{\delta} \qquad (3)$$

Where n = the number of sample units,
   N = the total number of units in the population,
   α = the propobility of omission,
   β = the propobility of commission,
   z_α = the critical normal distribution value of α,
   z_β = the critical normal distribution value of β,
   σ = a prior estimate of the population standard deviation.
   δ = a prior estimate of the limit of error.

It should be noted that using this formula parameter σ and δ should be known beforehand.

When the population is not enough large, we can give sample size by our experience. If the number of population is less than 50, we usually sample all the polygons in order to ensure the reliability, If the number is not less than 50 and that the area of the whole monitoring region is below 10 thousand mu, the minimum sample size are 50 , but we should sample more than 75 or 100 sample units when the area of the whole monitoring region is over 10 thousand mu.

4. ACCURACY ASSESSMENT

The accuracy assessment consists of the estimate of a single polygon error and the whole monitoring region error. Based on the standard normal distribution, we establish three area accuracy indices including relative error, average relative error , and relative mean square root error using relative error theory , and using the error spread law, the formulas of single polygon error and the whole monitoring region error are build .The relative error expression is

$$dv = \frac{\sum (A_i - B_i)}{\sum B_i} \qquad (4)$$

where   dv = relative error,
      Ai = monitoring area of the ith sample unit,
      Bi = reference area of the ith sample unit,
The average relative error is given as

$$d|v| = \frac{\sum |A_i - B_i|}{\sum B_i} \qquad (5)$$

The relative mean square root error has the form as

$$dm = \sqrt{\frac{\sum (\Delta_i - \mu_0)^2}{n-1}} \qquad (6)$$

where $\Delta_i = \dfrac{A_i - B_i}{B_i}$,

$\mu_0 = \dfrac{1}{n} \sum \Delta_i$,

i =1,2,…,n (n = the number of sample).

Therefore, the mean square root of single polygon *dm* is

$$dm = \sqrt{\frac{\sum (A_i \cdot dm_i^2)}{\sum (A_i)}} \qquad (7)$$

where single polygon area *Ai* is the sample weight here, *dmi* is the error of every area unit.

Since samples are used to replace population, we can consider that the *dmi* is the single polygon mean square root of the monitoring region stratum. We suppose that all the polygons are not correlative and they are of the same error, the relative mean squrae root of the total area of the ith stratum in monitoring region *dMi* is

$$dM_i = \frac{dm_i}{\sqrt{N_i}} \qquad (8)$$

5. EXAMPLE

The number of the change polygons is 532 in land Use monitoring by remote sensing image in ×× city. We sample 173 polygons with stratified random sampling.The accuracy assessment result is given in table 4 and table 5, Using the method we suggest in this paper.

Stratum (i)	<10	10-20	20-50	>50
The number (ni)	72	31	43	28
The total of Area (Si)	414	484	1460	4142
Relative error(dVi)	4.51%	1.76%	2.76%	5.90%
AV(d|v|)	10.4%	12.9%	15.3%	12.1%
MSR(dmi)	35.4%	16.3%	8.1%	13.0%
MSR of single polygon(dm)	14.9%			

Table4. The accuracy assessment of single polygon.

Stratum(i)	≤10	10-20	20-50	≥50
The number (ni)	259	103	89	59
The total of area (Si)	1233	1378	2793	8097
RMSR (dmi)	35.4%	16.3%	8.1%	13.0%
RMSR of total area (dMi)	2.2%	1.5%	0.8%	1.7%
MSR of total area /relative error	233 / 1.6%			

Table5. The accuracy assessment of the wholregion.

6. CONCLUSIONS

This techniques has been tested in Land use Dynamic Detection project conducted by the Ministry of Land and Resources P.R.China in the latest years.But because of the complexity of accuracy assessment, we should consider many factors, such as the hardness of field collection ,the expending of time and money,etc. The implement of accuracy assessment should depend on the cooperation of all departments and the improvement of accuracy assessment techniques.

REFERENCES

1. Siamak Khorram, 1999. Accuracy Assessment of Romote Sensing_Derived Change Detection. ASPRS. pp. 131-135.

2. Rosenfield, G.H., 1982. Sample Design for Estimating Change in Land Use and Land Cover. *PE&RS*, Vol.48, No.5, pp.793-801.

3. Stephen V. Stehman and Raymond L.Czaplewski , 1998. Design and Analysis for Thematic Map Accuracy Assessment : Fundamental Priniples, *Remote Sensing of Environment* Vol.64, No.3, pp .331-344.

4. Zhou Sheng,etc,1988.*Probability and Statistical.* Zhejiang,China.