

## ASSESSING THEMATIC MAP USING SAMPLING TECHNIQUE

Ali A. Darvishsefat

University of Tehran, Faculty of Natural Resources, Karaj-IRAN  
E-Mail: [adarvish@chamran.ut.ac.ir](mailto:adarvish@chamran.ut.ac.ir), Fax: +98 21 8007988

### ABSTRACT

The estimation of accuracy of thematic maps derived from remotely sensed data used in conjunction with a geographic information system is presented. A systematic random sampling technique is proposed to assess the accuracy. The main objective of this work is the determination of the optimal sampling intensity and the appropriate sample size used to assess the accuracy. The data set used is a Landsat-TM scene from three different regions in Mazandaran and Gilan provinces of Iran. The regions are classified using 9 different approaches and subsequently used as thematic maps. The overall accuracy and kappa coefficient of these results are compared with a complete (100%) ground-truth from these regions that are used as a comparison reference.

87 sampling networks with intensities ranging from 0.5 - 6% and the systematic random sample plots, each measuring between 900–32'400 square meters are prepared in raster format. These networks are overlaid with the thematic maps and the reference ground-truth to obtain the final results. These results are then compared using matrices which contain information such as the overall accuracy and the kappa coefficient. The estimated sample accuracy is then compared with the real accuracy.

The results of this research indicate that through sample ground-truth with 2% intensity in a systematically random technique and appropriate sample size, considering diversity and uniform phenomena in a region, we can estimate relatively accurate ( $\pm 4-6\%$ ) the accuracy of thematic maps, prior to use in GIS-database. PC Arc/Info and IDRISI are the main data analysis packages used for this work.

in a correct estimation of thematic maps accuracy.

### 1 INTRODUCTION AND PURPOSE

The main purpose of data acquisition and input into a GIS-database is using them for analysis and decision making. Any analysis based on thematic maps with an unknown or low accuracy would distract the estimation process and would lead to misplanning. The necessity of having updated information on one hand, and the high cost and considerably long acquisition time required for fieldworks on the other hand, introduces a gradual information lack contained in satellite data compared to fieldwork. Further on, optimal information extraction from remotely sensed data is highly dependent on the capability of the individual interpreter, the methods used, the site assessed, and its characteristics. Therefore the accuracy of these maps must be estimated. In pure research work, such as examining the capabilities in small areas, a complete ground-truth (covering the whole region) is used for comparison and determining the accuracy (Darvishsefat, 1996 and Itten, et al, 1992). But since providing a complete ground-truth demands very high costs and takes a long time, only parts of a thematic maps can be checked in form of points or in sample plots (Gerado and Valenzuela, 1991). Regarding the nature of thematic maps, the procedure of checking shall be made on sample plots with proper random distribution, adequate sample plots (intensity of the sampling), suitable sample size and characteristics of the region.

The aim of this paper is to determine the optimal sampling intensity and derive the appropriate sample size for providing a sample ground-truth that will be used

### 2 DATABASE AND METHODOLOGY

This investigation is performed in three test-sites in northern Iran using Landsat-TM data. Each of the test-site is described individually in the following sections:

#### 2.1 Nour National Park Region

This investigated test-site is part of a forested area near Nour city in Mazandaran province. The relief is mostly flat and the altitude ranges between -25 and 40 m above sea level. In addition to the forest, there are other uniform and big landuse classes present in this region such as sea, farmland and urban areas. Geometric registration of the satellite image is performed using a ground control point method (resulting RMS = 19m). The test-site covers 366\*316 pixel equalling 10409 ha. The TM data is classified into 3 classes (forest, water and other) using a supervised method with Maximum Likelihood (ML). In order to have a better estimation of accuracy, the classification is repeated using different band compositions and training areas.

The base map used as ground reference including sea and forest boundaries dated from 1993 is updated to the situation of the forest stands in the year of the satellite data acquisition (1998). The updated map is digitized and converted to raster format and subsequently used as ground-truth to calculate the real accuracy and an estimation of the accuracy of the different classifications using sampling techniques. The real accuracy of the 6 classified images is 91.2%, 89.4%, 86.4% 85.7%, 84.4% and 50.3% respectively.

## 2.2 Experimental Forest, Patom District

This test-site (765 ha) is located in the mountainous forest of Noshahr in Mazandaran province. The elevation of the forest ranges from 30 to 1290 m above sea level. The satellite data is geometrically corrected using a ground control point method and a digital elevation model (DEM). The achieved RMSE is less than a TM-pixel (25.8m). Two forest stand classifications (4 and 6 stands) are performed using a ML-classifier. A forest stand map at a scale of 1:20'000 is digitized and used as ground-truth for accuracy assessment. The accuracy of the two classified images is 43.7% and 36.7%.

## 2.3 Lahijan Region

This test-site (36000 ha) is located near Lahijan city in Gilan province in northern Iran. The relief is mostly flat and covered with forest, orchard, urban area and sea. In contrast to the first test-site, the landcover is very mixed and diverse. A landcover classification (5 types) is done using geocoded satellite images and again a ML classifier. A landcover map obtained from the Ministry of Agriculture is digitized and used as ground-truth. The real overall accuracy of the classified image is 67.4%.

## 3 ACCURACY ESTIMATION USING SAMPLING TECHNIQUE

As already mentioned, the main purpose of this study is to present a suitable sampling method (in terms of intensity of the sampling) for the estimation of the accuracy of thematic maps used in a GIS-database. In other words, the main purpose is to provide an optimal sampled ground-truth for estimating the accuracy of thematic maps. By providing a sampled ground-truth, the three most important factors are the intensity of the sampling, the number of samples collected (as well as their sizes) and their distribution. Given the advantages of random systematic sampling method and its easy implementation, it has been selected as the reference method to be used in this study.

The determination of the optimal sampling intensity (e.g. what percentage of the area is to be sampled), the number of sample plots are determined from the following relation for different intensities ranging from 0.5 % to 6 % . Further on, the size of the sample plots are tested in proportion to the total area of each test-site ranging from 900 square meters (1 TM pixel) to 32'400 square meters (6 x 6 TM pixels).

$$N = I * S/s$$

Where N is the number of sample plots, I the intensity, S the area of the region, and s the area of sample plot.

For all 9 classifications performed in the three regions, the number of sample plots in relation to the different plot size for different intensities are calculated. Using IDRISI for WINDOWS and PC Arc/Info software,

sampling networks are prepared for all above states and overlaid on classified images and the ground-truth of three regions. The extraction is made in the locality of the sample plots (Fig. 1). The results, i.e. the classified images and ground-truth maps of the samples are compared with each other and error matrices subsequently prepared. Using these matrices, accuracy criteria such as the overall accuracy and the Kappa coefficient are calculated as an estimation of the accuracy of the thematic maps resulting from the classification. The results of sampling and the real accuracies are listed in Table 1. Further on, for the final discussion and the conclusions, parts of the results have been depicted ( Fig. 2 ).

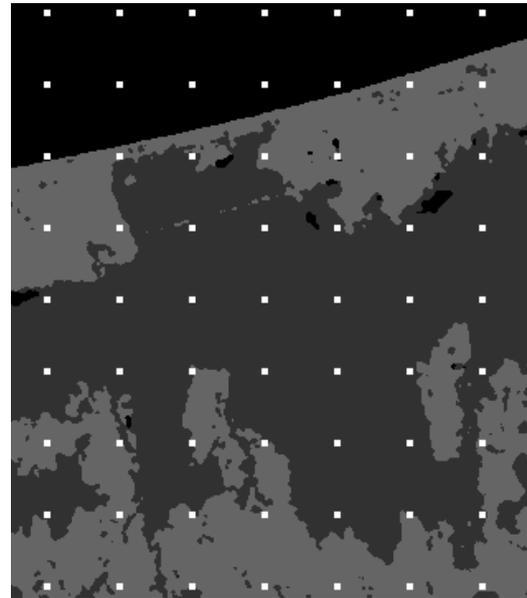


Figure 1: Random systematic network with intensity of 2% and 6\*6 pixel plot size, overlaid on classified image in the Nour region.

## 4 DISCUSSION AND CONCLUSIONS

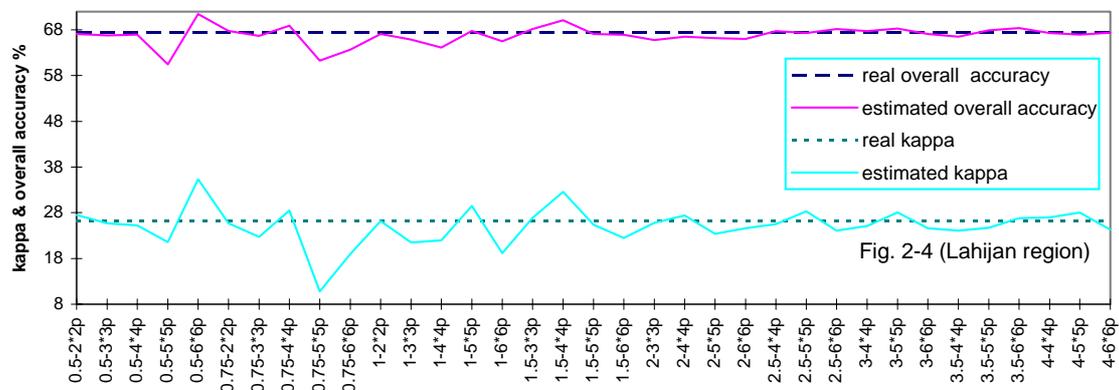
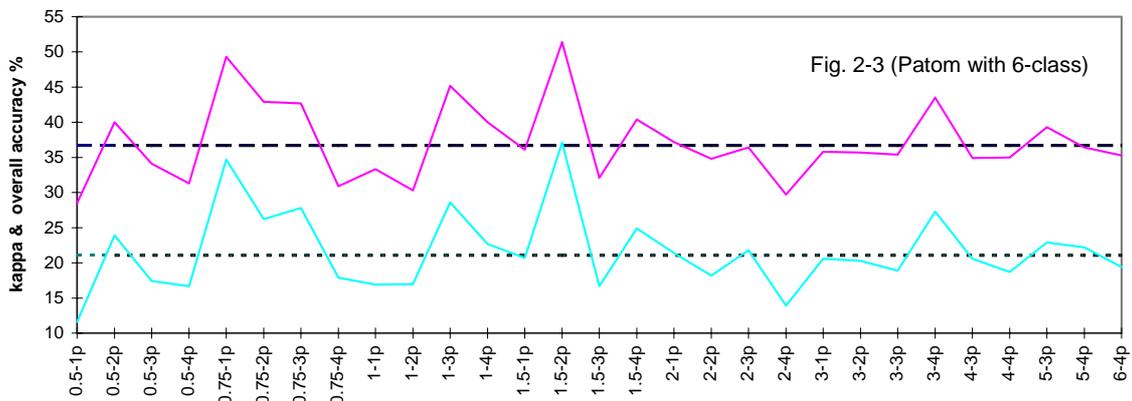
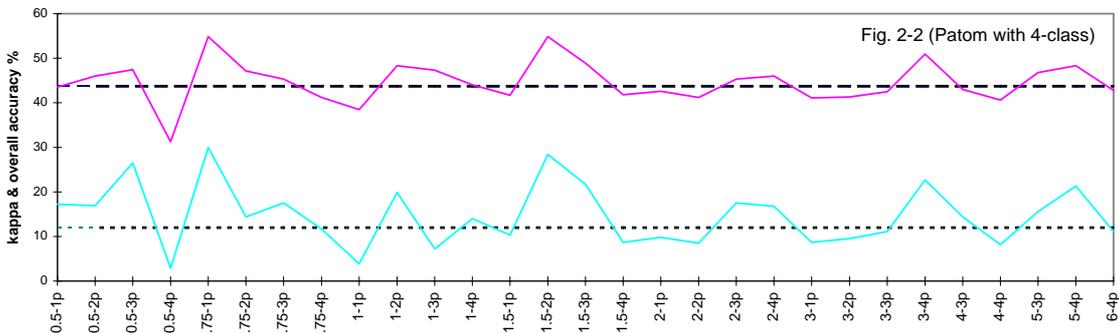
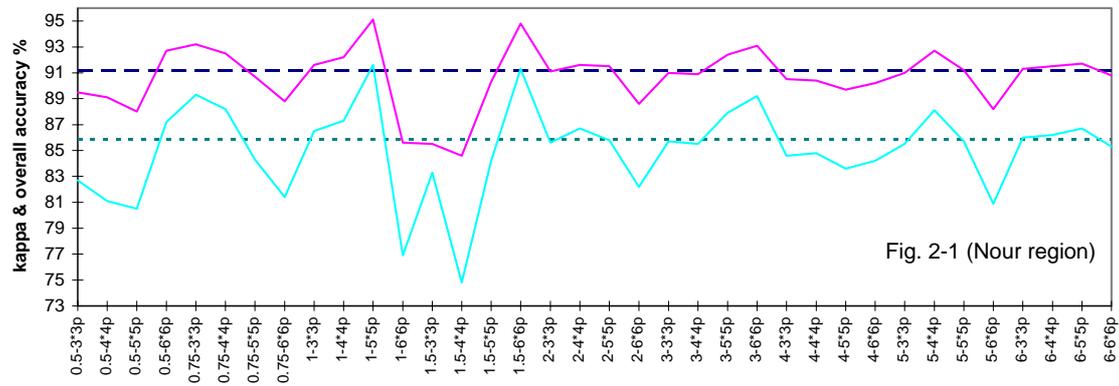
Figure 2-1 represents the range of variations of accuracies estimated with different intensities and the size of the plots related to one of the classified images of Nour. They are compared with their real accuracy. As the figure indicates, the range of variations in intensities lower than 2 % is clearly higher than the range of variation in intensities higher than 2 %. This is true for all classifications performed in this area (6 classifications). The figures related to the 4- and 6- class classifications of Patom Forest (Figures 2-2 and 2-3), confirm the results obtained from Nour area, i.e. clear decreasing of the range of variation of estimated accuracies (against the real accuracy) in intensities higher than 2 % . Based on these results and compared with the estimated accuracies of Lahijan area (Figure 2- 4), it may also be said that the turning point is at about 2 % intensity and aiming at 1.5 % intensity. The descending the optimal intensity may be explained with the big size of one of the

classes (farmland) in this area. Relying on the results obtained so far, it may be said that providing a sample ground-truth of only 2 % of a region and comparing it with thematic maps demonstrates the accuracy of these maps with a tolerance of 4 - 6 % compared with a real accuracy. Since the sampling intensity is expressed in percentage of the area of the region, it is clear that in a certain intensity a more extended part of vast regions may be sampled but the number of sample plots is determined by the size of the sample plots. In regions with smaller areas, or a higher diversity and more distributed classes (such as the Patom region), more sample plots shall be selected. On the contrary, in vast areas such as Nour region, comparatively larger plots may be used (this facilitates the implementation). In this concern, the distributed pattern of the classes in thematic maps (their expansion and massiveness), too, may be taken into account. In areas where the classes are small and

distributed, the sample plots may be smaller but more numerous. Based on the conclusions of this study, the accuracy of the thematic maps can be estimated by designing and implementation of a sampling network being proper and suitable in respect of the size of the plots and using random systematic technique in an area approximately equal to 2 % of the region. Being aware of the accuracy of thematic maps, they may be used in a GIS database and further analysis will certainly reduce wrong decisions that result from using unknown and less accurate maps. Raster based GIS are highly capable for designing sample networks and comparing the resulting sample ground-truth with thematic maps. Using accuracy criterias, the Kappa Coefficient, which considers and calculates omission and commission parts along with correct parts, is the best suited method to estimate accuracies for comparison (Congalton et al, 1983).

Test-site 1, Nour region													Test-site 2, Patom				Test-site 3 Lahijan			
int.- nxn	class. 1		class. 2		class. 3		class. 4		class. 5		class. 6		int.- nxn	4 class		6 class		int.- nxn	5 class	
	est. ov. acc. (real = 86 %)	est. kapp a (real = 77.8 %)	est. ov. acc. (real = 84 %)	est. kapp a (real = 74.6 %)	est. ov. acc. (real = 91 %)	est. kapp a (real = 85.9 %)	est. ov. acc. (real = 86 %)	est. kapp a (real = 77.1 %)	est. ov. acc. (real = 50 %)	est. kapp a (real = 23.4 %)	est. ov. acc. (real = 89 %)	est. kapp a (real = 83.0 %)		est. ov. acc. (real = 44 %)	est. kapp a (real = 12.0 %)	est. ov. acc. (real = 37 %)	est. kapp a (real = 21.1 %)		est. ov. acc. (real = 67.4 %)	est. kapp a (real = 26.2 %)
0.5-3*3	81.6	68.6	80.6	66.0	89.5	82.7	82.9	71.5	51.0	24.2	83.7	72.4	0.5-1	43.5	17.2	28.5	11.6	0.5-2*2	67.1	27.5
0.5-4*4	85.4	74.5	80.4	65.0	89.1	81.1	91.5	85.4	50.0	20.7	85.4	74.4	0.5-2	46.0	16.9	40.0	23.9	0.5-3*3	66.8	25.7
0.5-5*5	85.0	75.6	85.0	75.6	88.0	80.5	84.8	75.6	55.4	28.9	90.6	84.8	0.5-3	47.4	26.5	34.1	17.4	0.5-4*4	67.0	25.2
0.5-6*6	79.9	67.0	72.4	56.5	92.7	87.2	89.9	82.0	51.9	8.2	82.1	70.3	0.5-4	31.3	2.9	31.3	16.7	0.5-5*5	60.5	21.6
0.75-3*3	89.1	82.9	85.6	77.1	93.2	89.3	87.7	80.8	50.3	25.3	91.1	86.1	0.75-1	54.8	30.0	49.3	34.7	0.5-6*6	71.5	35.3
0.75-4*4	86.1	78.0	86.2	78.1	92.5	88.2	84.8	76.2	44.3	13.2	91.9	87.2	0.75-2	47.1	14.4	42.9	26.2	0.75-2*2	67.8	25.7
0.75-5*5	87.6	79.1	84.3	73.2	90.7	84.3	89.7	82.9	47.2	16.3	86.8	77.9	0.75-3	45.3	17.5	42.7	27.8	0.75-3*3	66.7	22.7
0.75-6*6	85.7	76.2	84.4	73.6	88.8	81.4	88.1	80.3	56.2	29.6	89.1	81.9	0.75-4	41.2	11.7	30.9	17.9	0.75-4*4	68.9	28.5
1-3*3	87.7	80.2	86.3	77.6	91.6	86.5	85.4	76.6	49.1	21.2	92.4	87.8	1-1p	38.5	3.9	33.3	16.9	0.75-5*5	61.3	10.8
1-4*4	85.6	76.4	86.5	77.7	92.2	87.3	86.4	77.8	54.8	27.9	90.0	83.8	1-2p	48.3	19.9	30.3	17.0	0.75-6*6	63.7	19.0
1-5*5	90.7	83.5	86.8	76.1	95.1	91.6	87.7	78.5	46.5	17.8	93.0	87.6	1-3p	47.3	7.2	45.2	28.6	1-2*2	67.1	26.2
1-6*6	84.4	74.9	82.6	71.2	85.6	76.9	83.2	72.8	50.0	23.7	84.9	75.7	1-4p	44.0	14.0	40.0	22.7	1-3*3	65.9	21.5
1.5-3*3	85.0	75.8	83.5	73.1	85.5	83.3	82.7	72.6	50.4	23.9	89.7	83.5	1.5-1p	41.7	10.3	36.1	20.7	1-4*4	64.1	22.0
1.5-4*4	80.0	67.1	79.8	65.8	84.6	74.8	83.0	72.4	50.7	24.4	84.6	74.7	1.5-2p	54.8	28.4	51.4	37.1	1-5*5	68.7	29.5
1.5-5*5	84.9	75.4	84.0	73.4	90.3	84.2	85.2	75.9	54.9	29.5	90.8	85.0	1.5-3p	48.9	21.7	32.1	16.7	1-6*6	65.5	19.2
1.5-6*6	93.5	89.0	85.5	75.2	94.8	91.3	92.7	87.8	52.2	25.2	90.8	84.8	1.5-4p	41.8	8.7	40.4	24.9	1.5-3*3	68.2	26.9
2-3*3	86.6	78.1	83.6	73.0	91.1	85.6	86.0	77.3	49.6	21.2	89.0	82.2	2-1p	42.6	9.8	37.2	21.4	1.5-4*4	70.1	32.6
2-4*4	87.5	80.1	83.8	73.8	91.6	86.7	87.6	80.3	50.9	24.5	90.6	84.9	2-2p	41.2	8.5	34.8	18.2	1.5-5*5	67.1	25.4
2-5*5	85.0	74.9	82.2	69.9	91.5	85.8	85.3	75.7	47.1	15.4	86.4	77.4	2-3p	45.3	17.5	36.4	21.8	1.5-6*6	66.9	22.5
2-6*6	84.7	75.5	83.5	73.0	88.6	82.2	84.6	75.9	55.0	33.7	89.0	82.6	2-4p	46.0	16.8	29.7	13.9	2-3*3	65.8	25.8
3-3*3	86.2	77.6	86.2	77.4	91.0	85.7	86.6	78.5	47.8	19.9	89.9	83.7	3-1p	41.1	8.7	35.8	20.6	2-4*4	66.5	27.4
3-4*4	86.4	78.1	83.2	72.5	90.9	85.5	86.3	78.0	52.6	26.8	89.8	83.7	3-2p	41.3	9.5	35.7	20.3	2-5*5	66.2	23.4
3-5*5	87.9	80.5	85.3	75.8	92.4	87.9	86.9	79.0	51.4	25.8	90.5	84.7	3-3p	42.5	11.1	35.4	18.9	2-6*6	66.0	24.6
3-6*6	88.1	81.3	85.3	76.7	93.1	89.2	88.1	81.5	50.8	25.6	89.0	82.8	3-4p	50.9	22.7	43.5	27.3	2.5-4*4	67.7	25.5
4-3*3	86.0	77.2	82.9	71.8	90.5	84.6	85.5	76.5	49.5	21.2	89.0	82.1	4-3p	43.0	14.4	34.9	20.6	2.5-5*5	67.3	28.3
4-4*4	83.9	74.2	83.3	73.0	90.4	84.8	85.2	76.7	52.6	27.3	87.5	80.0	4-4p	40.6	8.2	35.0	18.7	2.5-6*6	68.2	24.1
4-5*5	84.4	75.6	83.3	73.1	89.7	83.6	83.5	73.9	50.1	22.8	89.1	82.7	5-3p	46.8	15.6	39.3	22.9	3-4*4	67.7	25.1
4-6*6	85.6	76.5	86.3	77.4	90.2	84.2	85.1	75.9	49.4	21.4	91.2	85.7	5-4p	48.3	21.3	36.4	22.2	3-5*5	68.3	28.1
5-3*3	86.3	77.7	83.4	72.7	91.0	85.5	85.9	77.3	50.7	23.6	88.0	80.6	6-4p	42.8	11.1	35.3	19.4	3-6*6	67.1	24.6
5-4*4	88.3	80.7	86.0	76.6	92.7	88.1	86.4	77.7	48.3	18.6	90.4	84.3						3.5-4*4	66.5	24.1
5-5*5	86.4	77.4	82.8	71.2	91.2	85.7	86.4	77.8	50.4	22.7	89.0	81.9						3.5-5*5	67.9	24.7
5-6*6	82.7	71.6	81.6	69.0	88.2	80.9	83.2	72.8	50.9	26.0	86.7	78.2						3.5-6*6	68.4	26.8
6-3*3	86.8	78.8	83.8	73.6	91.3	86.0	86.6	87.6	51.6	25.5	89.1	82.5						4-4*4	67.3	27.0
6-4*4	87.0	87.6	84.6	74.4	91.5	86.2	86.5	78.1	49.6	21.1	89.9	83.5						4-5*5	67.0	28.0
6-5*5	86.0	77.5	84.3	74.3	91.7	86.7	85.9	77.5	49.2	21.5	88.5	81.6						4-6*6	67.4	24.3
6-6*6	85.5	76.7	83.3	72.7	90.8	85.3	83.6	73.9	49.5	22.9	89.5	83.2								

Tab. 1 : Real accuracy of 9 classifications in 3 test-site and their estimated accuracies (kappa & overall acc.) which calculated with sampling technique with different intensities and sample sizes (int. =intensity, n\*n=sample size in pixel).



Sampling intensity (%) and sample size (pixel)

Figure 2: Comparison of the real accuracy in 3 test-sites with their estimated accuracies which are calculated with sampling technique, using different intensities and plot sizes.

## 5 RECOMMENDATIONS

With making the results of this study more reliable and with achieving an experimental relation between the sampling intensity with factors such as area, diversity, and massiveness of classes, it is suggested that similar studies may be made in few other areas with a different variety of classes.

It is also recommended that for the determination of the intensity and correct number of sample plots, the results of the tests and different states of this study may be analysed using adequate statistical methods.

## 6 ACKNOWLEDGMENTS

The author would like to thank the University of Tehran for the financial support on this study. The assistance of Mr. Fallah Shamsii is also acknowledged.

## 7 REFERENCES

Congalton R., R. Oderwald and R. Mead, 1983: Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques. PE&RS, Vol. 49, No. 12, pp. 1571-1678.

Darvishsefat Ali A. 1996: Einsatz und Fusion von Multisensoralen Satellitendaten zur Erfassung der Waldinventuren. Proc. ISPRS 1996 Symposium, Commission VII, Vienna, Austria, pp. 186-191.

Gerardo B, and Valenzuela, 1991: Integration of GIS and Remote Sensing in Landuse and Erosion Studies. In Book: Remote Sensing and Geographic Information Systems.

Itten K., P. Meyer, T. Kellenberger, R. Leu, ST. Sandmeier, P. Bitter and K. Seidel, 1992: Correction of the Impact of Topography and Atmosphere on Landsat-TM Forest Mapping of Alpine Region. Department of Geography, University of Zurich.