*Available online a*t www.scientiaresearchlibrary.com

Scientia Research Library **ISSN 2348-0424**
USA CODEN: JETRB4

Journal of Engineering And Technology Research, 2015, 3 (4):8-20

(*http://www.scientiaresearchlibrary.com/arhcive.php*)

Development a Software Package For Remote Sensing Area Estimators Used in Image Classification

Dr. Tarek El-Damaty
Civil Engineering Dept., Faculty of Engineering, Banha University

ABSTRACT
Accurate estimation of areas covered by land use classes is an important factor to many resource management and monitoring programs, crop yield forecasting, forest and environmental management. This paper is focusing on developing a software package for area estimation techniques that used in remote sensing classifications and applying it. Soft classified image with a reference map are used. A comparative study has been done using the developed software. The results show that all areas which are estimated by based confusion matrix area estimators are more accurate and closer to the true areas than that by proportional counting estimator and the results show that; the developed software is an effective tool in supporting area estimation techniques.

Keywords : Area estimation, Remote sensing, Image classification, Land use.

INTRODUCTION

Accurate area estimation is one of the important applications of remote sensing for different studies as crop areas or forest management strategies etc. The usual approach of classifying all pixels and counting or proportioning the pixels per class are rather inaccurate that it is often necessary to make calibrations on the direct counts in order to obtain better estimates for the marginal areas (Dymond , 1992, Schriever and Congalton, 1995). The calibrations of the marginal area estimates are based on the utilization of the sample confusion matrices.

Pixel counting as an area estimator is often proposed in remote sensing projects run by the private sector for public administrations, mainly in developing countries. The estimates are acceptable only if spectral signatures are clearly discriminated and image classification is very accurate (Gallego, 2004). However, because of classification error, the area derived from pixel counting is usually biased (Gallego, 2004, Stehman, 2005). Because the pixel counting is based on a complete census of the region, the bias of this pixel count area is viewed as a “measurement bias” rather than as an “estimator bias” (Stehman, 2005). A confusion matrix provides the classification error information that allows for adjusting the area obtained from pixel counting to account for this measurement bias (Stehman, 2009).

The main objective of this study is to develop a software package for producing land cover area estimation using conventional techniques (pixel counting, proportion counting) and based error matrix techniques. For a software package to be useful, it is desirable that it should be inter active with the user and help him at various steps of program. Further, the results of any data selected should be shown immediately as and when desired by the user. Then the developed software package applied on two classified remotely sensed images with five land cover categories.

MATERIAL METHOD

AEMRSD Software

The software package Area Estimation Methods via Remotely Sensed Data (AEMRSD) has been developed in C sharp language(C #). Language has been preferred primarily because it's modular, object oriented, various graphs used for displaying outputs charts. Beside that it is very useful for coding engineering, scientific problems and provides an advanced code editor, convenient user interface designers, integrated debugger, and many other tools to make it easier to develop applications. C# is an elegant and type-safe object-oriented language that enables developers to build a variety of secure and robust applications that run on the .NET Framework. C# can be used to create traditional Windows client applications, database applications, and much more. The C sharp compiler is required to run and executed the portion of program written in C sharp. This software can be successfully run on PCs with minimum 512 MP RAM, processor 1.2 MH/sec and 16 MB graphics card. Windows 98 or later version is required.

AEMRSD Description

The following part describes briefly how to use the software (AEMRSD) to estimate different areas. Also estimating bias and dispersion criteria are described. Figurer (1) includes introduction to the program, designer and supervision committee names. Tool pars, menus and submenus for (AEMRSD) also included.

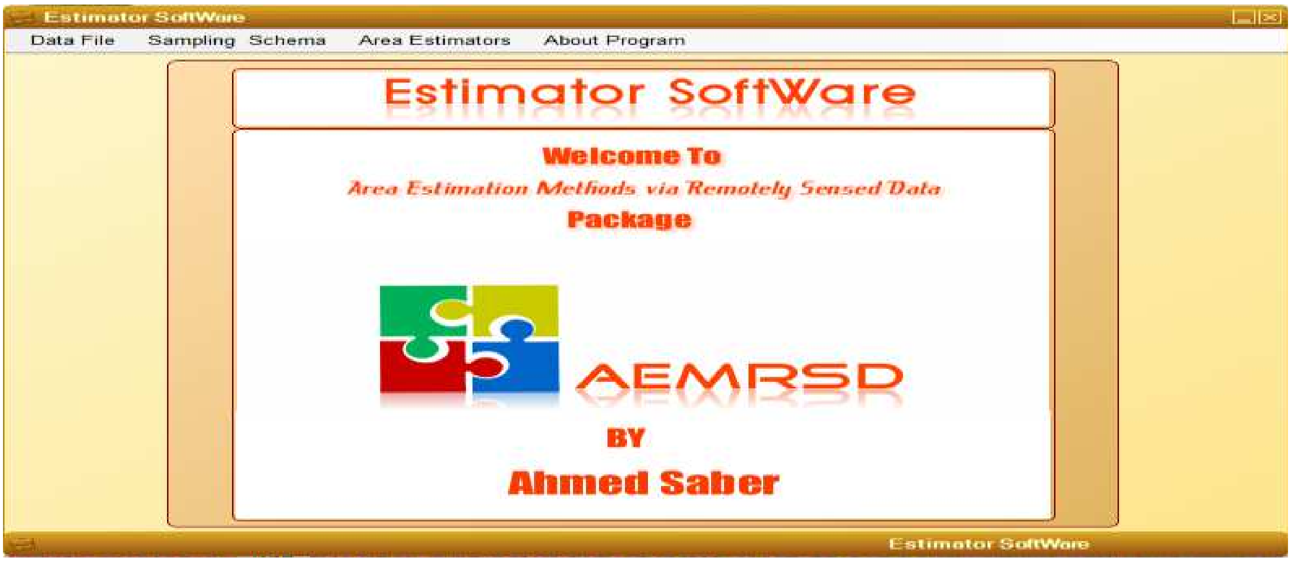

Figure (1) menus and sub menus layout for software

The software (AEMRSD) contains three main menus:

i) Data file.

ii) Sampling schema.

iii) Area estimators.

Data File

The data file menu figure (2) contains two submenu; Input Data and Close. Input Data submenu contains two input images first image is *classified image* which has been classified by soft or hard methods of classification. Second one is a *reference image*. This two input images must be in text format as illustrated in table (1) for soft data and (2) for hard data. The program gives chose to select the path of classified and refere nce image figure (3).after selecting images the program is ready to start.

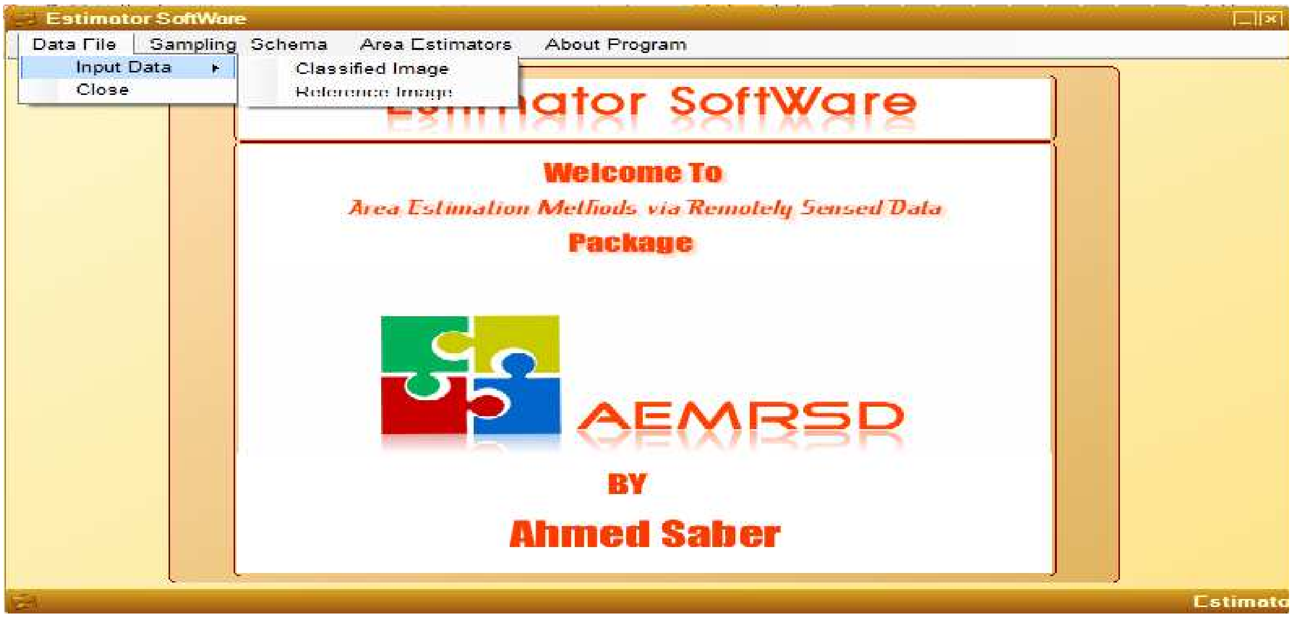

Figure (2) Input data menus and sub menu (classified and reference image)

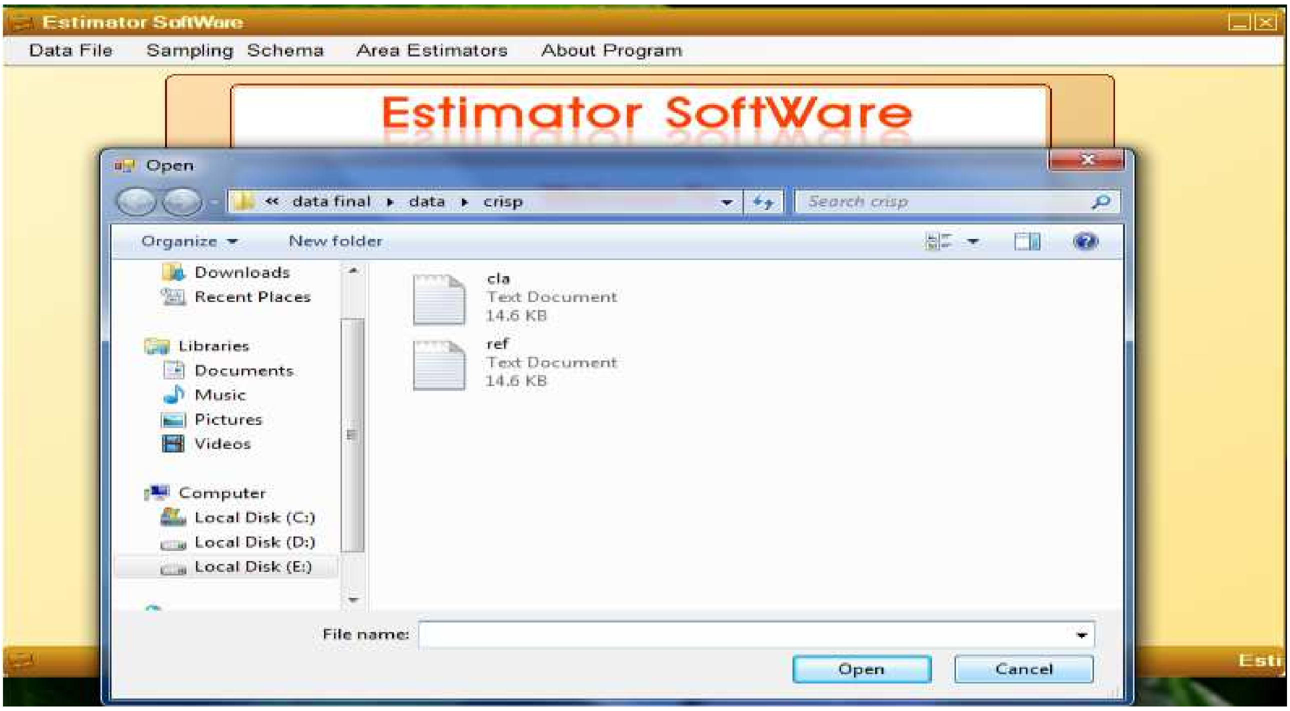

Figure (3) Chose menu to select classified and reference image.

Table (1): input image (soft data) text format

0.267, 0.118, 0.080, 0.470, 0.065
0.188, 0.270, 0.231, 0.129, 0.183
0.130, 0.071, 0.049, 0.712, 0.038
0.077, 0.044, 0.029, 0.828, 0.022
0.199, 0.084, 0.047, 0.639, 0.031
0.013, 0.020, 0.008, 0.952, 0.007

Table (2): input image (hard data) text format

0, 1, 0, 0, 0
1, 0, 0, 0, 0
1, 0, 0, 0, 0
0, 0, 0, 1, 0
0, 1, 0, 0, 0
1, 0, 0, 0, 0

Sampling Schema

After selecting the input classified and reference images in this case the software is ready to generate sample error matrices. Selecting sampling schema menu is generating these matrices. Sampling scheme describes the way in which sample pixels are selected from the image in order to characterize the thematic classes of interest. As shown in figure (A.4) two sampling schemes are available in software package. Simple random sampling is a method of selecting n units out of the N such that every unit of the N units has an equal chance of being selected. Systematic sampling this method supposes that the (N) units in the population are numbered from (1) to (N). In order to select a sample of n units, the first unit is selected randomly, and other sampling units are then selected at fixed interval (k).

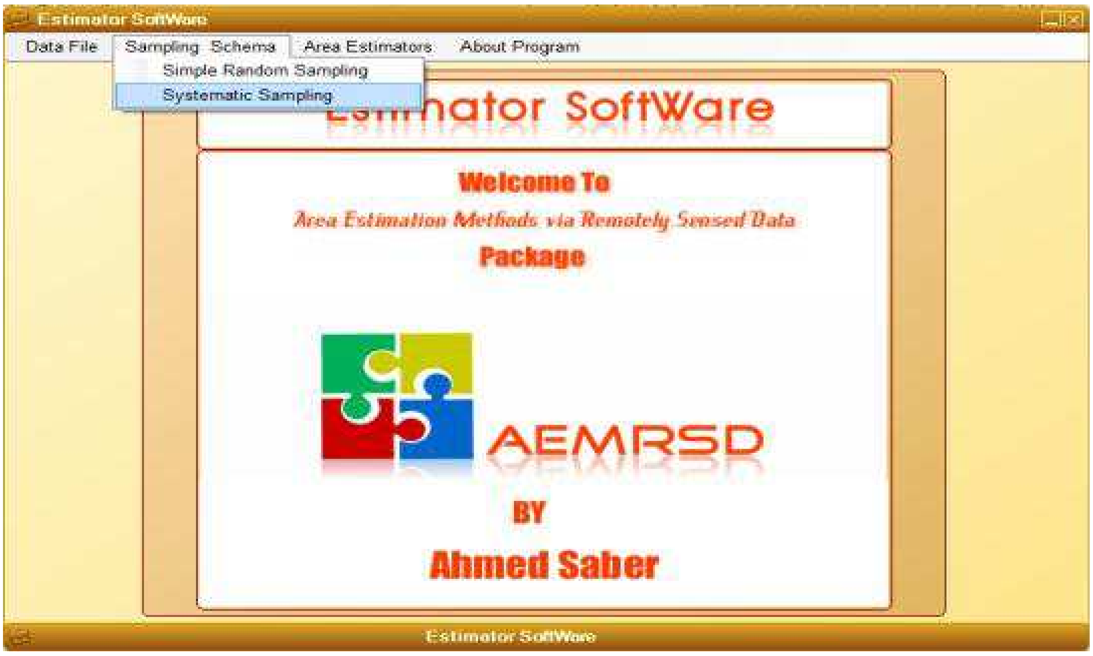

Figure (4) Sampling menu (random & systematic sampling)

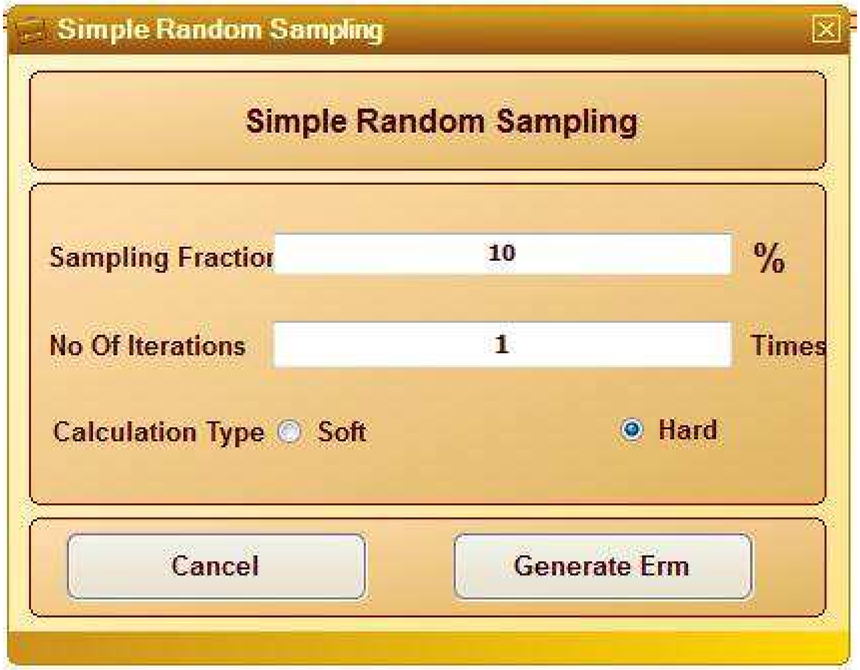

Figure (5) Dialog Box of Sampling

Dialog Box of Sampling figure (5); sampling fraction is a sample size. No. of iterations is No. of estimated sample error matrices at same sample size. Soft or hard in case of input data.
For example output of sample error matrix (3) and (4).

Table (3) output traditional error matrix for sampling fraction 10 %, one iteration. (Simple random sampling)

```
-------------------------
  Estimation Module
-------------------------
                                                        ------------------------------
                                                              Random  Sampling
                                                        ------------------------------
--------------------------------
                step1
---------------------------------
```

	R1	R2	R3	R4	R5	Total
C1	29	4	0	2	0	35
C2	4	31	12	0	2	49
C3	0	2	26	0	4	32
C4	1	0	0	10	2	13
C5	0	0	2	0	5	7
Total	34	37	40	12	13	136

Table (4) output soft error matrix for sampling fraction 10 %, one iteration. (Random sampling)

```
-------------------------
  Estimation Module
-------------------------
                                                        ------------------------------
                                                              Random  Sampling
                                                        ------------------------------
--------------------------------
                step1
---------------------------------
```

	R1	R2	R3	R4	R5	Total
C1	16.21	10.76	4.31	2.96	1.59	35.82
C2	9.21	18.19	7.91	1.22	2.44	38.98
C3	2.75	10.73	10.12	1.23	4.29	29.12
C4	3.06	2.14	1.87	6.44	1.42	14.93
C5	1.38	4.88	5.85	1.23	3.83	17.17
Total	32.61	46.69	30.05	13.09	13.57	136.01

Area Estimators

In this menu a choice estimator which will estimate deferent class area are available. Two methods conventional and error matrix based technique are available. in case of selecting the main method will estimate area the submenu illustrate choice of estimator are available. The conventional technique contains pixel and proportional counting estimator figure (6). The other menu error matrix based technique contains; additive, direct, map marginal proportion, bias removal and inverse estimators figure (7). By selecting estimator dialog box (8) for estimator are available. In this box choice of sample size, no. of iteration, sampling step and kind of sampling is available. To check bias and dispersion click show chart Patton figure (9).

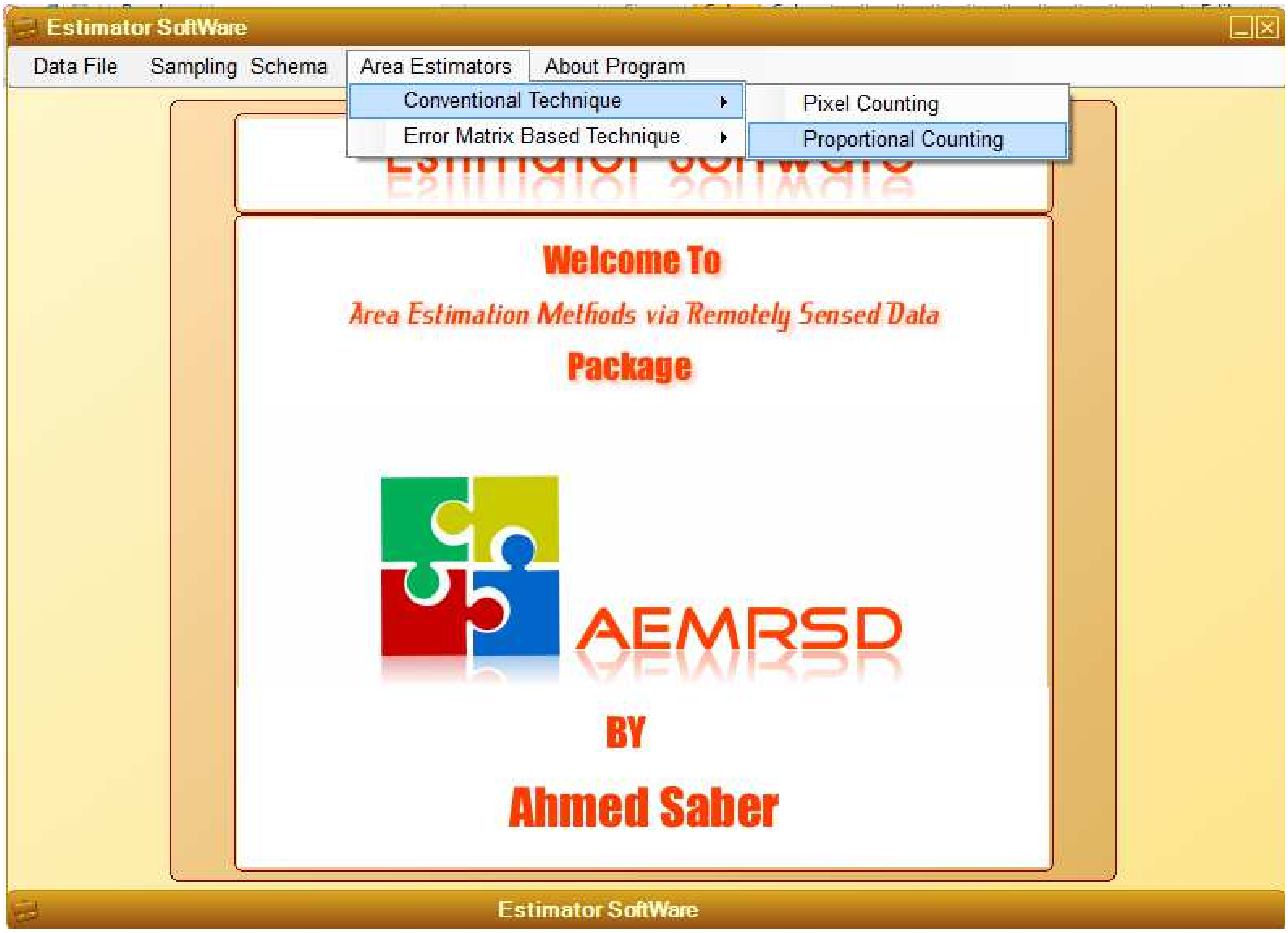

Figure (6) Area estimators menu (Conventional Technique)

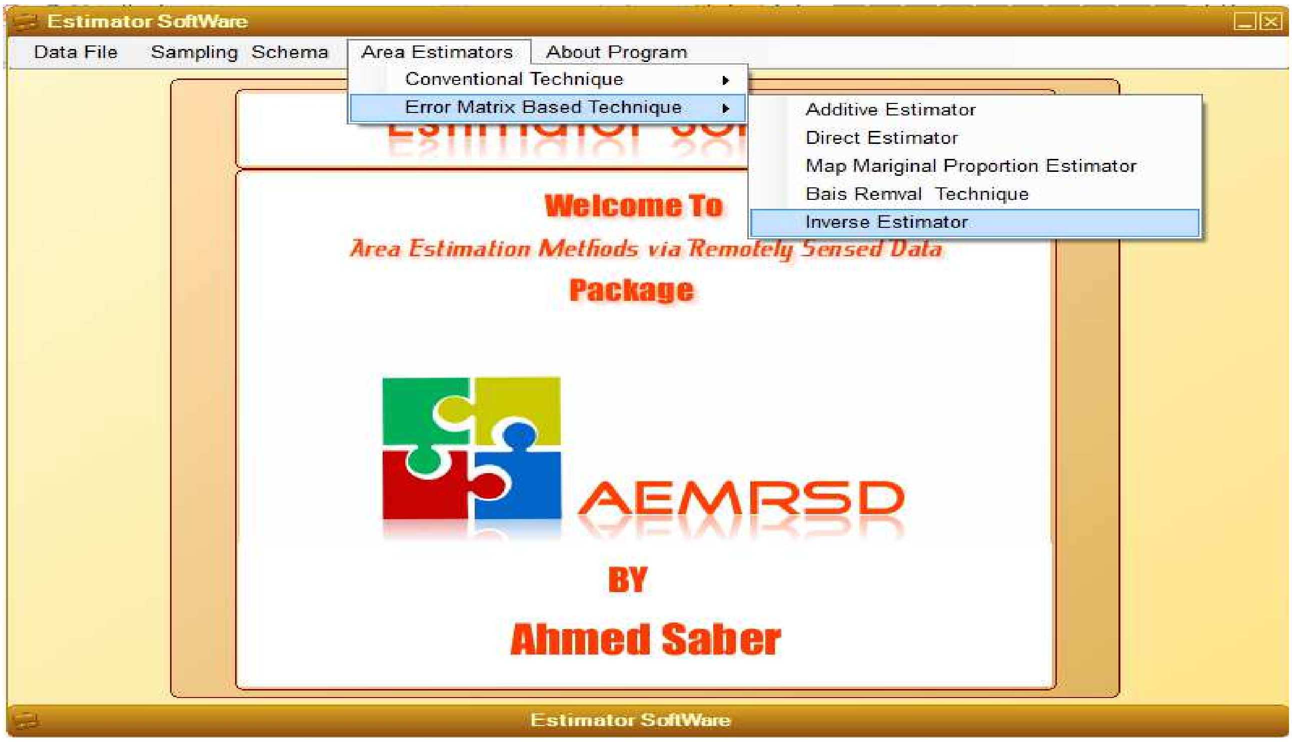

Figure (7) Area estimators menu (Error Matrix Based Technique)

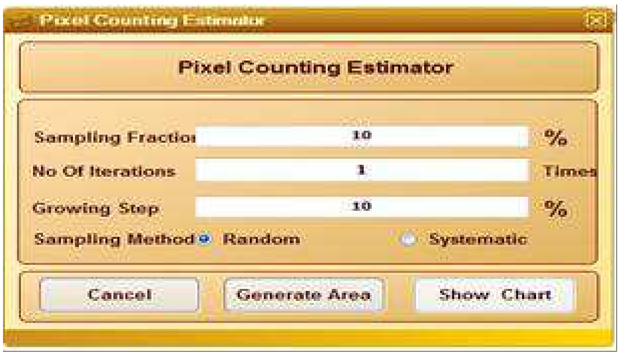

Figure (8): Dialog Box of Sampling

Table (5) illustrate sample of output area for different classes.

Estimator Module --- Random Sampling

Addtitive Estimator

n=10%	n=20%	n=30%	n=40%	n=50%	n=60%	n=70%	n=80%	n=90%	n=100%
321.72	316.86	319.72	320.22	318.61	323.67	323.41	320.98	321.93	324.39
382.99	395.4	386.7	392.23	390.42	387.25	386.54	390.08	391.61	389.04
412.36	404.16	408.26	400.98	405.58	406.24	407.27	404.99	401.82	401.53
116.79	113.64	117.72	116.36	116.99	112.94	114.73	115.32	114.92	115.54
127.14	130.94	128.61	131.22	129.4	130.89	129.04	129.64	130.72	130.5

Area Percentage

n=10%	n=20%	n=30%	n=40%	n=50%	n=60%	n=70%	n=80%	n=90%	n=100%
23.64	23.28	23.49	23.53	23.41	23.78	23.76	23.58	23.65	23.83
28.14	29.05	28.41	28.82	28.69	28.45	28.4	28.66	28.77	28.58
30.3	29.7	30	29.46	29.8	29.85	29.92	29.76	29.52	29.5
8.58	8.35	8.65	8.55	8.6	8.3	8.43	8.47	8.44	8.49
9.34	9.62	9.45	9.64	9.51	9.62	9.48	9.53	9.6	9.59

Figure (9) Bias and dispersion charts

RESULTS AND DISCUSSION

The developed software package applied on two classified remotely sensed images with five land cover categories. First image in this study is from Indian Remote Sensing (IRS) satellite LISS II sensor at spatial resolution of 36 x 36 m used as classified map. Second image at fine spatial resolution (6 m) obtained from IRS 1C PAN sensor has been used (with topographical maps or topo-sheet (number 53G/13) at 1:50,000 scale ,1988, and existing field surveyed map at 1:1000 scale ,1992) as reference map.
All results of this application are illustrated in Table (6) and (7) and in Charts (1), (2), (3) and (4). In Tables (6) and (7) the area percentage estimated by software program is recorded. First row (Proportion in reference map) illustrate the true area percentage for deferent classes. Second row (Proportion in classified map) illustrate the area percentage classified for deferent classes. Below rows (Average area percentage estimated by deferent estimators) illustrate the average of area estimated for sampling from 10 % to 100 % by deferent estimators. In this table's total quantitative error are used to evaluate different estimators as first evaluation criteria. Charts (1) and (2) illustrate the relation between sample size and bias. Absolute bias criterion can be considered as difference between estimated proportions and their true values. Charts (3) and (4) describe the relation between sample size and average dispersion.

Software Output

In this case soft classified image and reference image are used; error matrices generated with simple random sampling scheme and the area were investigated by different estimators for different classes in Table (6).

Table (6) Average area percentage for different classes estimated by different estimators (Soft Input Data, Random Sampling)

Land cover type		Built-up land	Grass land	Trees	Agriculture	Barren land
Proportion in reference map (%)		24.358	32.189	23.802	9.471	10.18
Proportion in classified map (%)		27.455	27.33	22.518	9.787	12.911
Average area percentage estimated by deferent estimators	Additive Estimator	24.336	32.256	23.786	9.462	10.161
	Bias Removal Technique	25.676	29.237	22.982	9.63	12.058
	Direct Estimator	24.347	32.32	23.777	9.42	10.136
	Map Marginal Proportion based Estimator	23.706	33.456	23.976	9.036	9.827
	Inverse Estimator	24.4	31.829	23.998	9.476	10.296
	proportional Counting	27.51	27.381	22.478	9.758	12.875

Total quantitative error for proportional counting estimator is 12.266 %, 0.133 % for Additive Estimator, 7.127 % for Bias Removal Technique, 0.262 % for Direct Estimator, 2.881 % for Map Marginal Proportion based Estimator and 0.719 % for Inverse Estimator. The Additive Estimator has the smallest total quantitative error 0.133 % and Proportion counting estimator have largest total quantitative error 12.266 %. The results illustrate that area estimated from confusion matrix area estimators is much closer to the true value than that estimated by conventional methods (proportion counting). [4] illustrates in case of hard input data also the confusion matrix area estimator is much closer to the true value than that estimated by conventional methods (pixel counting estimator). From Table (3) The Additive Estimator and direct estimator produced the most accurate area estimates with mean values closer to the true proportions and Bias Removal Technique produced the most inaccurate area estimates.

Output Of Soft Classified Data With Systematic Sampling

In this case soft classified image and soft reference image are used; error matrix generated with systematic sampling scheme and the area is investigated by different estimators for different classes.

Table (7): Average area percentage for different classes estimated by different estimators

(Soft Data, Systematic Sampling)

Land cover type		Built-up land	Grass land	Trees	Agriculture	Barren land
Proportion in reference map (%)		24.358	32.189	23.802	9.471	10.18
Proportion in classified map (%)		27.455	27.33	22.518	9.787	12.911
Average Area Percentage Estimated By Deferent Estimators	Additive Estimator	24.318	32.179	23.793	9.5	10.211
	Bias Removal Technique	25.649	29.209	22.976	9.653	12.071
	Direct Estimator	24.316	32.193	23.827	9.455	10.21
	Map Marginal Proportion Based Estimator	23.68	33.27	24.043	9.097	9.911
	Inverse Estimator	24.409	32.035	23.693	9.551	10.31
	proportional Counting	27.51	27.303	22.528	9.76	12.9

In Table (7) the total quantitative error between the proportions of land cover categories reduced from 12.321 % with proportional counting estimator to 0.119 % in Additive Estimator, 7.17 % in Bias Removal Technique, 0.117 % in Direct Estimator, 2.643 % in Map Marginal Proportion based Estimator and 0.524% in Inverse Estimator. In Table (7) the direct estimator produced the most stable area estimates with mean values closer to the true proportions and Bias Removal Technique produced the most unstable area estimates. The different class's area estimated from confusion matrix area estimators is much closer to the true value than that estimated by conventional methods (proportion or pixel counting estimator). Charts (1) and (2) display the absolute bias while Charts (3) and (4) illustrate the average dispersion as second and third evaluation criteria.

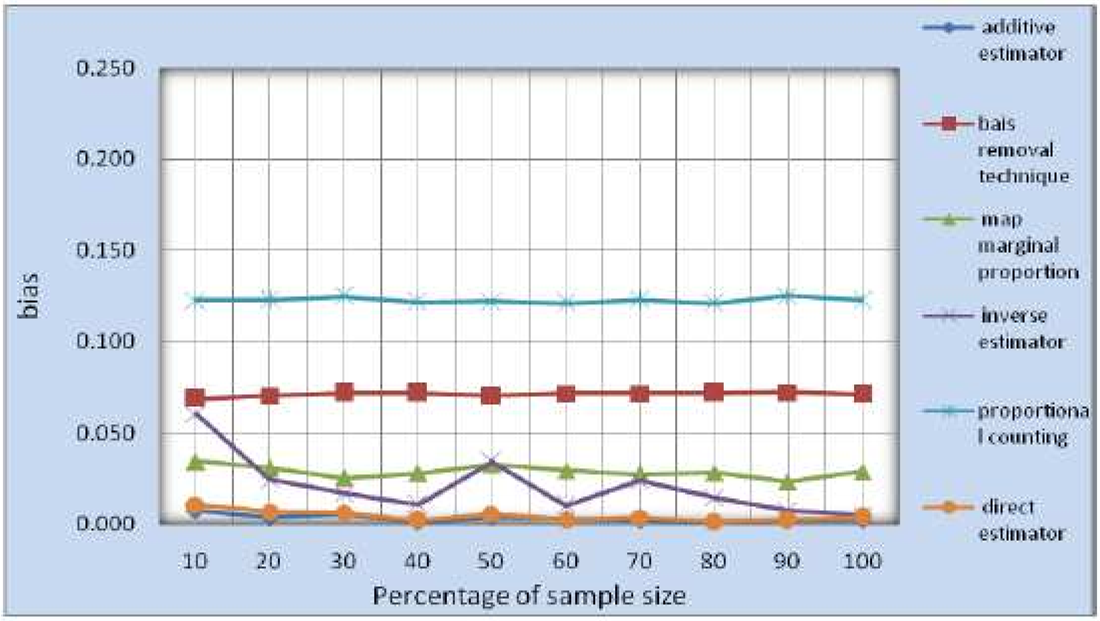

Chart (1): Bias (Soft Data with Random Sampling).

The average absolute biases for all estimators under different sampling fraction and random sampling scheme are given in chart (1).Some important observations can be made on this chart: For all sample size proportional counting estimator (conventional technique) often has the largest bias than all confusion matrix area estimators. Otherwise in confusion matrix area estimator's direct and additive estimators have the smallest bias for all samples fraction and map marginal proportion estimator often has the largest bias. Generally the direct and additive estimators have the smallest bias of all estimators under all sampling fractions.

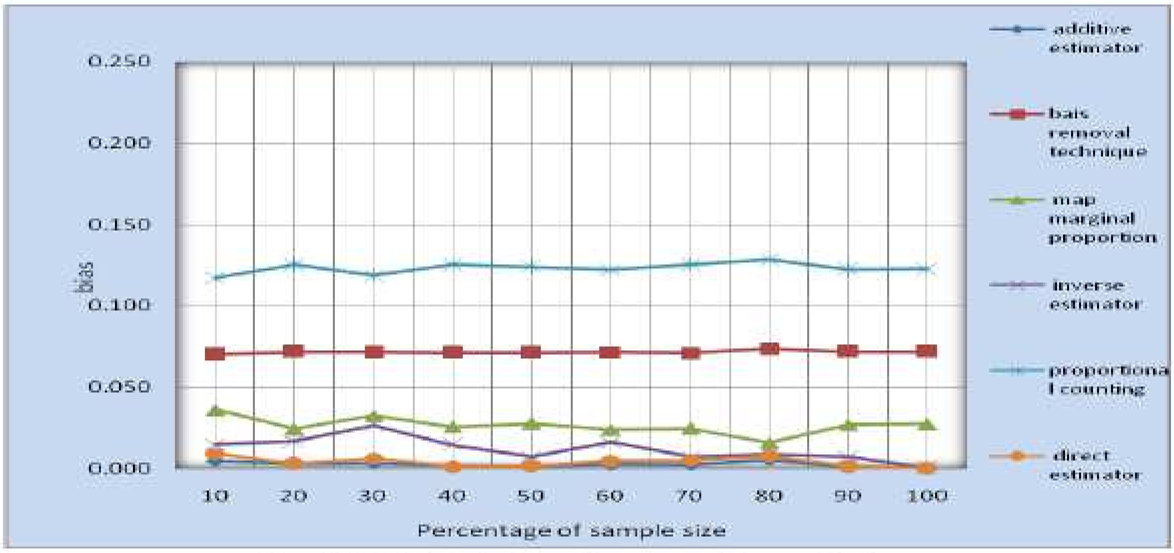

Chart (2): Bias (soft data with systematic sampling)

Chart (2) gives the same results as Chart (1) proportional counting estimator (conventional technique) often has the largest bias than all confusion matrix area estimators. Direct and additive estimators have the smallest bias for all samples.

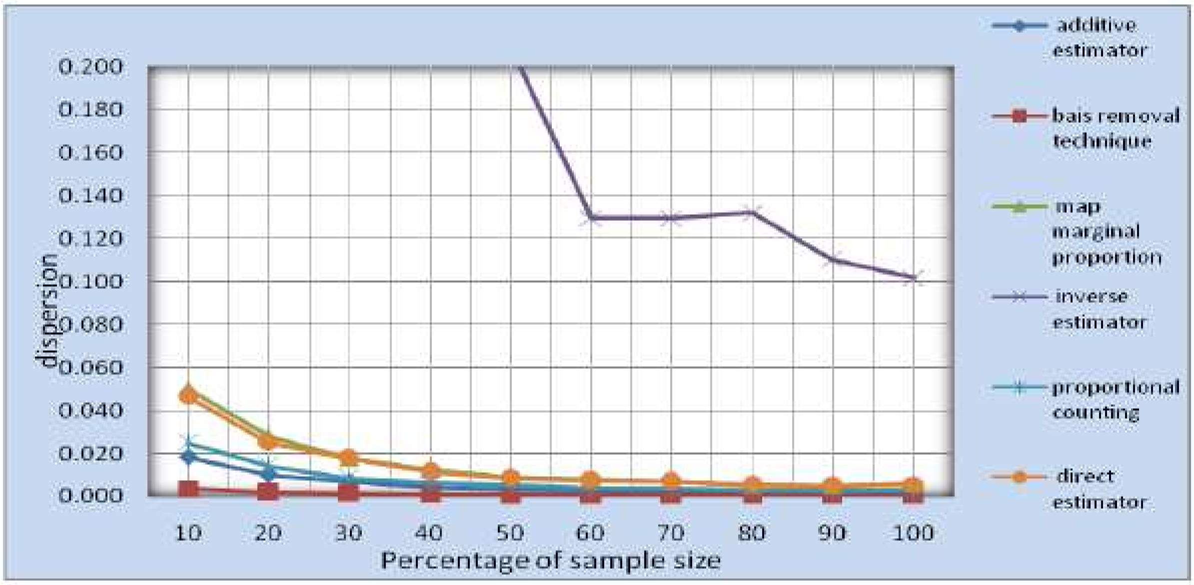

Chart (3): Dispersion (Soft Data with Random Sampling)

Chart (3) describes the average dispersion for all estimators direct observation leads to: The dispersions for all estimators approach zero as the sampling fraction increases. This means increasing in sampling fraction leads to increasing in precision of area estimation for all estimators. Inverse estimator has the largest dispersion in all estimators and bias removable technique and additive estimators have the smallest dispersion in all sampling fraction.

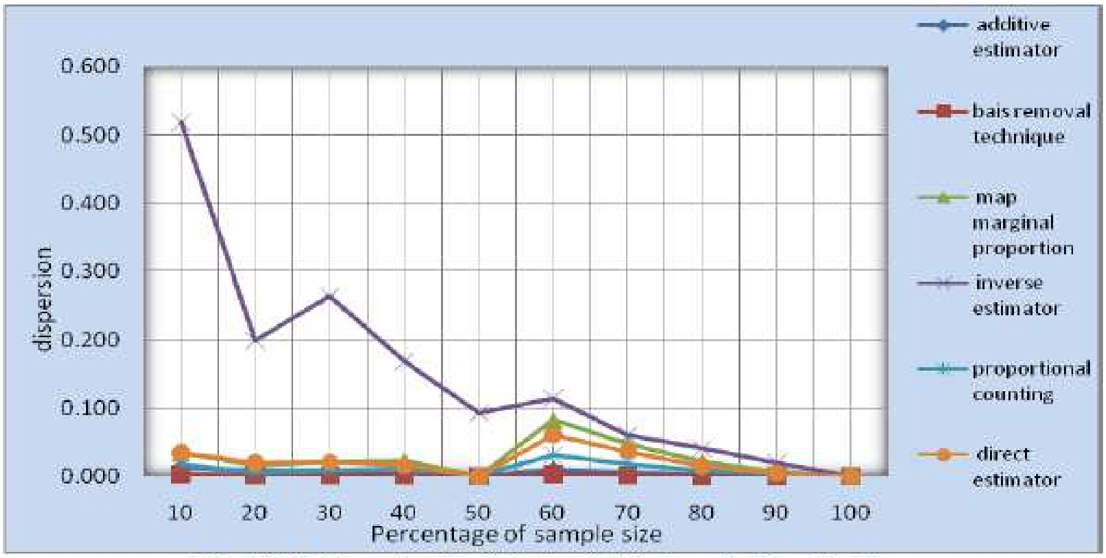

Chart (4): Dispersion (Soft Data with Systematic Sampling)

Chart (4) illustrate the average dispersion for estimators using systematic sampling comparing these results with Chart (3) Inverse estimator still has the largest dispersion in all estimators and bias removable technique and additive estimators still have the smallest dispersion in all sampling fraction.

CONCLUSION

In this paper, a software package for producing land cover area estimation using conventional techniques (pixel counting, proportion counting) and based error matrix techniques has been developed and applied. a proportional counting and five based confusion matrix area estimators for soft image classifications were studied and individual comparison has been done for estimators.
The results applying the developed software illustrate that:

- The proportional estimator is not suggested for use because of its large total quantitative error and large bias in all sample size.
- The areas estimated by all based confusion matrix area estimators are more accurate and closer to the true areas than that by proportion counting estimator. Also this is suggested in case of the input data are hard classified images.
- Direct and additive estimators produce the most accurate areas than other estimators. In case of using based confusion matrix area estimators to estimate different class area the direct and additive estimators are therefore recommended.

- Bias removal estimator is based on computation of error of omission and error of commission. If it is minimized, that would improve the area estimates from this technique. This may be happened if the classification accuracy is high.

- The developed software is an effective tool in supporting area estimation techniques.

REFERENCE

[1]. Card, D., (**1982**). "Using known map category marginal frequencies to improve estimates of thematic map accuracy". Photogrammetric Engineering & Remote Sensing, 48(3): 431−439.

[2]. Congalton, R. G., (**1991**)," A review of assessing the accuracy of classiŽ cations of remotely sensed data". Remote Sensing of Environment, 37, 35–46.

[3]. Czaplewski, R. L., & CATTS, G. P. (**1992**). "Calibration of remotely sensed proportion or area estimates for misclassification error", Remote Sensing of Environment, 39, 29−43.

[4]. Dymond, J. R., (**1992**), "How accurately do image classifier estimate area?", International Journal of Remote Sensing, 13(9):1735-1742.

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

[6]. Hay, A. M., (**1988**)," The derivation of global estimates from a confusion matrix", International Journal of Remote Sensing, 9(8): 1395–1398. Ibrahim, M. A., (2004), "Evaluation of soft classifiers for remote sensing data ", Ph.D. Thesis, Indian Institute of technology, Roorkee, India.

[7]. Priesley, P., and Smith, J. L., (**1987**), " Using Classification Error Matrices to Improve the Accuracy of Weighted Land-Cover Models ", Photogrammetric Engineering and Remote Sensing, 53, 1258–1263.

[8]. Stehman, S. V., **1997**, Selecting and interpreting measures of thematic classiŽ cation accuracy.Remote Sensing of Environment, 62, 77–89.

[9]. 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

[10]. Walsh, T. A., and Burk, T. E., (**1993**), " Calibration of satellite classification of land area ", Remote Sensing of Environment, 46, 281–290.

[11]. Woodcock, C. E., and Gopal, S., (**2000**), " Fuzzy set theory and thematic maps: accuracy assessment and area estimation ", International Journal of Geographical Information Science, 14(2), 153–172.

[12]. Yuan, D., (**1997**), " A simulation comparison of three marginal area estimators for image classification ", Photogrammetric Engineering and Remote Sensing, 63(4): 385–392.