SIMULATION COMPARISONS OF MODELS FOR ESTIMATING POPULATION SIZE

H. C. Wei † *and Dhanwant S. Gill* ‡

ABSTRACT

This paper compares four models used for correcting visibility bias in sample surveys of estimating the size of population. The computer procedure is implemented to simulate their performances in terms of bias and standard error. Simulation is done to compare their behaviors especially when the probabilities of objects being seen(captured) are not constant in such cases. A new model is introduced to reduce bias in estimation.

From the simulation results none of the five models compared seems to consistently provide satisfactory estimate. The correction factor proposed by Caughley and Grice(1982) offers a relatively better estimate when the probability of recording observation is very low(**p**=0.1). Petersen estimate is best for high seeing/capturing probability. Chapman's(1951) modified Petersen estimate outperforms the other four models in comparison by giving far better estimate on various combinations of low to average **p** values. A new model proposed in this paper which is based on Petersen's estimate provides best results when size of the target population is small to moderate, from 10 to 50, with combination of average and high probabilities.

1. INTRODUCTION

In estimating the size of a natural population of interest from "total counts," a broad range of modern techniques exist which are based on work done by Schnabel[8] in 1930's. Problem with such estimate has long been known inaccurate due to some systematic factors inherent to the methods. Classical examples may be found in sampling wildlife and other areas(medical screening, etc.) where the observations, while sampling, are missed because of poor sightability or lacking of visibility or low density or sometime even unskillful observers; and thus bias is introduced in estimation from such incomplete "counts." Four models:

1. Henny et al.[6],
2. Petersen estimate[8],
3. Caughley and Grice[2], and
4. Chapman[3]

representing endeavors taken from several directions are available for correcting the problem of bias.

In the popular "mark and recapture" scheme[9]; one count is made determining how many individuals are captured(marked and released) while a second independent count is made. Such procedure also provides important information, as out of the second count we can determine how many were marked previously. Based on their binomial property the estimate of the total is then obtained although these counts are nontheless quite incomplete. This procedure has lent itself to a wide array of applications which include as only a small part of wildlife management such as estimating fish population in a lake by Schnabel.

---

† Current affiliation and correspondence: AT&T Bell Laboratories, Room IHP 2F341, 200 Park Plaza, Naperville, IL 60566.

‡ Current address: California State Polytechnic University, Pamona, CA.

Another approach based on the joint probabilities of seeing in separate searches, the probability of sighting was assumed to be uniform among observers, and a binomial distribution is used to estimate the population size. With variations many survey techniques derived from this procedure have been used; most notably, aerial surveys for counting emu groups in Western Australia[2]; estimating the coast osprey population for the Mid-Atlantic[7], and the census of bald eagles[5].

The use of binomial distribution is appropriate since in a sample survey either the observer sees(captures) the object or misses it. Consider the binomial distribution Binomial(**N**, **p**), where the parameter **N** represents the number of objects in the population and the parameter **p** represents the probability of seeing(capturing) an object. When **N** is fixed in advance after observing k objects successfully, the usual problem is to estimate the probability of success **p** in the experiment. In practice, situations often arise when **N** becomes the unknown parameter of interest. If **p** is assumed to be known and k successes have been observed, the experimenter would be interested in estimating **N** instead. Examples include the estimation of the total number of a certain species of wildlife in an area with the number of individuals observed on several occasions.

The bias produced from estimating population by incomplete counts exists virtually in all the surveys just mentioned. It is to the practitioners' interest to know as such. In the literature, we notice methods for correcting the bias in numerous applications, but the relative performance of these survey models was not discussed. Gill et al.[4] compared three survey models using binomial distribution for simulating their behaviors for the situation of constant sighting(capturing) probabilities across two counts. The intent of this study is to compare relative performance in terms of bias and standard error of models for estimating counting bias including cases when the assumption of constant probability of observing among observers is relaxed. Result concludes that the behavior of the proposed model, Model 5, is bounded between Model 2, Petersen estimate, and Model 4 which is suggested by Chapman. Model 5 is most accurate when the difference in probabilities of seeing(capturing) is low and the target population is not abundant(less than 100). Finally, recommendation is given toward choosing the best model in survey under various circumstances.

2. MODELS

2.1 Notations

In order to compare across different models, following unified notations are adopted:

- **S1**: the number of objects seen by the first observer but missed by the second
- **S2**: the number of objects seen by the second observer but missed by the first
- **S**: the total number seen by only one observer or the other
- **B**: the number of objects seen by both
- **M**: the number of objects missed by both

2.2 Model 1

Henny et al.[6] used the following model to estimate the population of Mid-Atlantic Ocean Coast osprey. Based on the joint probabilities of observing individuals in separate searches; the ability of observing an individual object is assumed the same among observers. The estimator for the constant probability p, is therefore based on the product of two binomials, is defined as:

$$\hat{\mathbf{p}} = \frac{2\mathbf{B}}{2\mathbf{B} + \mathbf{S}}$$

This gives the estimate of total population size as:

$$\hat{\mathbf{N}} = \frac{\mathbf{S}}{\hat{\mathbf{p}}}$$

2.3 Model 2

This model is best known as Petersen estimate[8]; the estimator is given by

$$\hat{\mathbf{N}} = \frac{\mathbf{n}_1 \mathbf{n}_2}{\mathbf{B}}$$

where $\mathbf{n}_i$ is the number of individual objects seen by observer i, $i = 1, 2$. Using the existing notations this model can well be translated into

$$\hat{N} = \frac{(B + S_1)(B + S_2)}{B}$$

since we can establish the relationship among these variables by

$$n_1 = S_1 + B; \text{ and}$$
$$n_2 = S_2 + B.$$

This model does not require the assumption of constant probability among observers.

2.4 Model 3

Caughley and Grice[2] proposed a correction factor, C,

$$\hat{C} = \frac{1}{\hat{P}}$$

as the multiplier for observed density S, estimated total is thus given by:

$$\hat{N} = \frac{S}{\hat{P}}$$

in which

$$\hat{P} = \frac{\hat{p}_1 + \hat{p}_2}{2}; \text{ and}$$

$$\hat{p}_1 = \frac{B}{(B + S_2)}$$

being the estimated probability of an object being seen by the first observer and

$$\hat{p}_2 = \frac{B}{(B + S_1)}$$

being the estimated probability of an object being seen by the second.

As we can see that the probabilities of the object been seen(captured) by the observers are estimated respectively, i.e., $\hat{p}_1$ and $\hat{p}_2$, there is no restriction on the constant probabilities for this model.

2.5 Model 4

Chapman[4] made a different kind of correction to Petersen estimate, Model 2, and gave

$$\hat{N} = \frac{(n_1 + 1)(n_2 + 1)}{(B + 1)} - 1$$

Applying our notations the estimate becomes

$$\hat{N} = \frac{(B + S_1 + 1)(B + S_2 + 1)}{(B + 1)} - 1$$

This gives an unbiased estimate when $S_1 + S_2 + 2B \geq N$.

2.6 Model 5

In an attempt to minimize the over-adjustment made by model 4 when either p_1 or p_2 is small; we propose to use $\hat{P}$ as the correction instead of 1 as been used in Model 4:

$$\hat{N} = \frac{(B + S_1 + \hat{P})(B + S_2 + \hat{P})}{(B + \hat{P})} - \hat{P}$$

The definition of $\hat{P}$ is the same as in Model 3.

3. SIMULATIONS

We assume an observer at each instance(trial) succeeds with probability p or fails with probability $1 - p$ to observe an individual of a natural population of size N. Further, trials are independent and the random variable Y is defined as

$$Y = y_1 + y_2 + \ldots\ldots + y_N$$

in which each y_i is an independent Bernoulli trial with probability p, and it is known that

$$Y \sim \mathit{Binomial}(N,p),$$

where ~ denotes "is distributed as."

Monte Carlo method is used to generate independent random samples from Binomial populations. We simulate two observers' behaviors by giving probabilities p_1 and p_2 with trial size N which generates 2 random variables, Y_1 and Y_2 separately from binomial distribution representing the number of individuals seen(captured) by these two observers. The third random variable, Y_3 is also obtained with probability $p_1 p_2$(the multiplication of p_1 and p_2) being the number of individuals seen by both observers.

Simulation is executed with the combination of $\mathbf{p}_1$ and $\mathbf{p}_2$ with 8 different trial sizes(N): 10, 20, 30, 40, 50, 100, 150 and 200. For each combination 1000 samples are generated to calculate the bias and standard error due to each of the five models. Bias is defined as the estimated number minus the actual population size, **N**. Standard error is defined as:

$$\textit{Standard Error} = \sqrt{\frac{1}{(\mathbf{T}-1)} \sum_{i=1}^{\mathbf{T}} (\hat{\mathbf{N}}_i - \mathbf{N})^2}$$

where **T** is the total number of estimated $\hat{\mathbf{N}}_i$ of the population size **N**.

During the course of generating binomial random variables observations with any of the following conditions were discarded and replaced by a new set of $\mathbf{Y}_i$'s:

1. when any of the $\mathbf{Y}_i$ is zero; or
2. when the number of distinct individuals observed, $\mathbf{Y}_1 + \mathbf{Y}_2 - \mathbf{Y}_3$, exceeds **N**.

4. COMPARISON AND DISCUSSION

4.1 $\mathbf{p}_1 \neq \mathbf{p}_2$

When both $\mathbf{p}_1$ and $\mathbf{p}_2$ are small and **N** is small the total population sizes were under-estimated and the order in which the models performed in the increasing magnitude of bias is as 3, 1, 2, 5 and 4. For example, Table 1 shows the estimated bias (upper entry in each cell) and standard error (lower entry) when $\mathbf{p}_1 = 0.1$ and $\mathbf{p}_2 = 0.2$:

TABLE 1. Estimated bias and standard error when $\mathbf{p}_1 = 0.1$ and $\mathbf{p}_2 = 0.2$

	N							
Model	10	20	30	40	50	100	150	200
1	-4.99	-5.50	2.55	13.52	29.75	134.12	247.34	308.43
	6.67	17.76	24.05	39.82	60.15	214.05	396.09	490.02
2	-6.08	-11.29	-12.46	-12.83	-10.16	15.14	44.38	50.67
	6.57	11.31	17.05	21.79	26.87	77.20	146.23	180.34
3	-3.95	-5.27	1.08	9.25	23.26	117.26	219.06	274.87
	5.96	18.30	22.14	35.19	53.71	190.47	353.51	441.02
4	-6.46	-12.97	-17.02	-20.68	-22.50	-21.82	-12.17	-6.78
	6.71	11.21	18.42	23.40	26.37	44.09	71.08	93.68
5	-6.24	-11.89	-13.80	-14.83	-12.93	-9.78	37.92	44.44
	6.63	11.20	17.36	21.91	26.38	72.31	138.92	173.45

Model 3 has smaller standard error when **N** is small, 10, and grows rapidly as **N** increases. In Table 2, as one of the $\mathbf{p}_i$, $i=1,2$ increased or **N** increased, models 3 and 1 over-estimated but difference in bias as well as standard error as compared to other models became less and models 2, 4 and 5 began performing slightly better.

TABLE 2. Estimated bias and standard error when $\mathbf{p}_1 = 0.2$ and $\mathbf{p}_2 = 0.3$

	N							
Model	10	20	30	40	50	100	150	200
1	0.03	10.05	26.94	47.44	62.41	114.97	89.62	158.04
	8.45	25.82	53.54	95.08	118.07	234.62	261.69	150.81
2	-3.12	-1.94	3.03	10.00	14.03	25.53	13.62	18.12
	5.64	12.33	23.23	43.95	50.23	102.48	109.80	59.35
3	1.28	11.52	29.20	50.72	66.42	127.42	132.51	188.64
	8.37	25.92	53.67	99.62	116.26	235.00	268.42	194.53
4	-4.32	-6.39	-5.83	-3.47	-1.94	5.67	6.51	5.49
	5.34	9.64	14.17	23.74	26.49	58.34	74.47	54.41
5	-3.57	-3.20	1.02	7.42	11.28	22.77	12.12	15.59
	5.51	11.40	21.26	40.98	46.75	98.53	106.79	58.51

The gains for models 2, 4 and 5 noticed here began increasingly considerable as the $\mathbf{p}_i$, $i=1,2$ increased to medium values and also as **N** increased.

Models 3 and 1 performed very poorly in comparison to others for medium $\mathbf{p}_i$'s for all values of N considered in this study. Tables 3 and

4 record such behaviors.

TABLE 3. Estimated bias and standard error when $p_1 = 0.3$ and $p_2 = 0.7$

Model	N 10	20	30	40	50	100	150	200
1	9.76	22.12	27.95	31.19	37.84	55.95	69.44	85.40
	16.56	39.25	48.64	49.70	56.80	78.38	98.51	115.05
2	2.10	5.86	6.75	6.94	8.03	7.80	5.10	4.21
	6.17	14.40	18.47	18.57	20.70	26.43	33.57	37.84
3	10.03	23.15	30.84	36.84	44.86	73.28	97.26	124.33
	15.09	34.88	45.88	49.88	58.19	87.95	116.21	143.37
4	0.27	2.90	4.00	4.77	5.98	6.33	3.84	3.05
	3.54	8.72	12.73	15.17	17.80	24.80	32.34	36.91
5	1.49	5.08	6.06	6.36	7.47	7.37	4.73	3.86
	5.32	13.11	17.33	17.79	20.01	26.01	33.25	37.59

TABLE 4. Estimated bias and standard error when $p_1 = 0.4$ and $p_2 = 0.6$

Model	N 10	20	30	40	50	100	150	200
1	5.48	11.64	12.96	11.14	14.64	14.51	19.02	27.48
	13.49	29.97	35.97	32.32	38.57	49.14	59.43	73.79
2	1.71	4.20	4.55	3.25	4.62	2.57	2.69	4.98
	6.14	13.69	17.22	15.65	18.80	25.04	29.63	36.23
3	8.78	18.99	24.89	27.47	35.37	57.42	83.61	113.95
	14.50	31.97	41.03	40.38	50.02	74.62	100.97	133.12
4	-0.01	1.72	2.38	1.70	3.01	1.35	1.54	3.82
	3.68	8.80	12.60	13.57	16.61	23.91	28.79	35.40
5	1.14	3.55	3.99	2.80	4.16	2.20	2.33	4.61
	5.35	12.61	16.32	15.14	18.28	24.75	29.40	36.01

However, performance of models 3 and 1 improves considerably as **N** increases in size. Among models 2, 4 and 5, in general, for medium values of p_i the standard error was smaller (although difference was not much sometimes) for model 4 but for bias each performed better than the others for some combination of p_i's. As **N** increases biases for models 2, 4 and 5 become positive but very small in magnitude.

As both or one of the p_i's increase(s) model 4 seems to have the smallest standard error. The bias for models 2, 4 and 5 become very small but are positive indicating a slight over-estimation of the population size. In general the model 4 seems to be the best but closely followed by models 5 and 2. Performance of models 3 and 1 does improve as **N** increases but they still lag behind considerably. Very rarely they come to a little close to models 4 , 5 and 2.

When the difference between p_1 and p_2 becomes large as shown in Table 5.

TABLE 5. Estimated bias and standard error when $p_1 = 0.1$ and $p_2 = 0.9$

Model	N 10	20	30	40	50	100	150	200
1	28.00	105.88	182.97	259.64	305.39	500.70	701.60	879.92
	30.34	118.62	217.23	316.18	378.10	567.35	754.72	925.45
2	1.15	8.14	14.07	21.82	26.15	32.57	40.14	43.17
	4.07	15.16	25.40	39.51	49.84	59.05	68.26	72.58
3	9.82	32.16	52.20	75.85	92.41	145.63	201.23	247.74
	12.24	40.73	66.92	100.15	125.14	175.32	228.79	273.04
4	0.17	3.61	7.30	12.65	16.52	26.45	35.27	39.23
	2.23	7.69	13.59	22.25	29.50	46.65	59.62	66.56
5	0.80	6.84	12.32	19.72	24.11	31.25	39.04	42.24
	3.41	13.13	22.59	36.00	46.20	58.84	66.63	71.30

Model 4 gives the best result; followed closely by models 5 and 2 in this order, models 3 and 1 again are far from satisfactory.

In the situation when both p_1 and p_2 are as high as one of 0.7, 0.8 or 0.9, model 2(Petersen estimate) performs the best. The order for these models' performance is clear judging from both bias and standard error, which is 2, 5, 4, 3 and 1. Tables 6 and 7 reveal such truth.

TABLE 6. Estimated bias and standard error when $p_1 = 0.7$ and $p_2 = 0.9$

Model	N 10	20	30	40	50	100	150	200
1	-5.88	-12.22	-18.34	-24.83	-30.96	-61.22	-90.18	-119.02
	6.30	12.69	18.87	25.34	31.56	61.96	91.14	120.05
2	-0.33	-0.72	-1.10	-1.41	-1.67	-2.98	-3.42	-3.73
	1.01	1.71	2.36	2.90	3.44	5.84	7.80	9.25
3	1.66	3.07	4.60	6.03	7.68	16.07	25.97	36.07
	2.50	4.30	6.11	7.68	9.56	18.56	29.05	39.32
4	-0.38	-0.76	-1.14	-1.45	-1.71	-3.02	-3.46	-3.77
	0.97	1.69	2.36	2.90	3.44	5.84	7.80	9.25
5	-0.35	-0.74	-1.12	-1.43	-1.69	-3.00	-3.44	-3.74
	0.99	1.70	2.36	2.90	2.44	5.84	7.80	9.25

TABLE 7. Estimated bias and standard error when $p_1 = 0.8$ and $p_2 = 0.9$

Model	N 10	20	30	40	50	100	150	200
1	-7.01	-14.80	-22.47	-30.11	-37.58	-74.42	-111.18	-147.25
	7.22	15.04	22.72	30.42	37.91	74.88	111.72	147.89
2	-0.33	-0.86	-1.28	-1.67	-2.06	-3.45	-4.89	-5.52
	0.84	1.58	2.16	2.77	3.27	5.56	7.69	9.26
3	1.13	1.17	2.45	3.24	4.11	9.26	14.42	20.73
	1.84	2.99	4.01	5.22	6.25	12.25	18.03	24.77
4	-0.35	-0.88	-1.29	-1.68	-2.07	-3.47	-4.91	-5.54
	0.83	1.58	2.16	2.77	3.27	5.56	7.70	9.27
5	-0.34	-0.87	-1.28	-1.67	-2.06	-3.46	-4.90	-5.53
	0.84	0.58	2.16	2.77	3.27	5.56	7.70	9.26

Although constantly and slightly under-estimated, models 2, 4 and 5 all provide good and close estimate, the difference among these three models is very small(off by 1 to 3 percent). Model 3 displays a unique behavior of over-estimating the total by less than 10 percent for all N's while all other four models under-estimate. For high probabilities, model 1 is the worst performer which consistently under-estimates the total by approximately over 50 percent.

4.2 $p_1 = p_2$

For very low probability, 0.1 for example(Table 8), model 3 gives best results in comparing with the other four on small to medium sizes(10 to 50). Model 2 does well for large sizes, 100 and 150, and model 5 displays far superiority for very large size, 200.

TABLE 8. Estimated bias and standard error when $p_1 = p_2 = 0.1$

Model	N 10	20	30	40	50	100	150	200
1	-6.49	-12.42	-14.43	-14.05	-12.78	41.85	132.66	219.86
	7.12	14.08	19.22	25.84	29.50	98.06	222.61	358.19
2	-6.84	-14.52	-20.11	-24.39	-28.24	-22.56	2.02	24.91
	7.07	15.02	21.30	27.00	31.56	50.26	89.34	142.88
3	-5.31	-11.17	-13.13	-12.55	-11.06	45.31	139.49	230.39
	6.14	13.16	18.48	25.26	29.05	99.22	225.41	362.60
4	-7.07	-15.35	-22.32	-28.56	-34.56	-50.08	-52.78	-52.37
	7.20	15.58	22.78	29.42	35.58	55.38	68.99	85.96
5	-6.95	-14.85	-20.87	-25.63	-29.96	-27.28	-5.00	16.65
	7.12	15.24	21.80	27.73	32.61	50.27	85.23	136.13

When both the probabilities improve from low to middle as shown in Tables 9 and 10, model 4 becomes the best performer, and the order of their performance is 4, 5, 2, 1, and 3. Model 3 consistently over-estimated the total by more than 50 percent and its large standard errors demonstrates the instability of such estimate under the circumstances.

TABLE 9. Estimated bias and standard error when $p_1 = p_2 = 0.4$

Model	N 10	20	30	40	50	100	150	200
1	5.35	16.81	21.57	23.72	23.87	36.75	43.77	52.29
	13.19	38.73	57.34	61.98	58.21	73.81	92.18	97.52
2	0.90	5.40	6.76	7.00	6.06	7.62	6.11	5.28
	6.34	18.32	27.57	29.93	27.97	34.49	43.01	43.94
3	8.02	23.11	31.91	38.30	42.45	75.49	102.50	130.91
	14.42	41.48	61.58	68.95	68.31	99.92	131.74	155.89
4	-1.19	0.58	1.65	2.39	2.30	4.54	3.37	2.71
	4.12	10.35	16.86	20.77	21.89	31.31	40.44	42.16
5	0.20	4.24	5.67	6.01	5.18	6.83	5.39	4.58
	5.58	16.63	25.81	28.45	26.91	33.83	42.49	43.54

TABLE 10. Estimated bias and standard error when $p_1 = p_2 = 0.5$

Model	N							
	10	20	30	40	50	100	150	200
1	4.32	7.44	5.89	8.62	7.23	5.67	7.09	11.75
	13.28	25.58	29.43	32.73	33.91	44.09	54.29	65.31
2	1.49	3.18	2.21	3.78	3.08	2.22	2.86	5.44
	6.48	12.66	14.75	16.99	17.87	24.09	29.35	35.49
3	8.30	16.58	19.81	27.80	31.42	54.81	81.14	111.04
	14.86	29.30	34.74	42.51	46.28	71.38	98.58	129.73
4	-0.26	1.03	0.50	2.12	1.63	1.04	1.72	4.28
	3.92	8.55	11.42	14.51	16.06	23.15	28.55	34.70
5	0.92	2.60	1.74	3.31	2.66	1.85	2.50	5.07
	5.68	11.74	14.08	16.44	17.43	23.83	29.14	35.27

Model 1 improves its performance greatly when **N** grows although not as good as models 2, 4, or 5. Models 2, 4 and 5 differ slightly among themselves especially when sizes are very large which indicates the adjustment made respectively by models 4 and 5 to Petersen estimate is most helpful for middle seeing(capturing) probabilities.

As both p_i's increase from middle toward high, we notice there is one transition point, that is when $p_1 = p_2 = 0.6$:

TABLE 11. Estimated bias and standard error when $p_1 = p_2 = 0.6$

Model	N							
	10	20	30	40	50	100	150	200
1	0.73	-2.31	-5.15	-7.77	-8.70	-17.79	-24.87	-35.70
	9.43	12.91	15.29	18.36	21.26	35.42	46.41	55.79
2	0.86	0.27	-0.19	-0.57	0.11	0.58	2.00	1.72
	4.63	6.52	7.71	9.06	10.57	17.13	22.02	24.12
3	6.05	8.95	12.07	15.40	20.62	41.54	64.41	83.73
	10.88	15.55	19.07	23.11	28.80	52.62	76.56	95.23
4	-0.03	-0.39	-0.74	-1.08	-0.40	0.08	1.49	1.23
	3.04	5.48	7.17	8.63	10.15	16.79	21.70	23.87
5	0.56	0.05	-0.39	-0.75	-0.07	0.40	1.81	1.54
	4.12	6.23	7.55	8.92	10.44	17.02	21.91	24.03

As shown in Table 11, model 4 no longer dominates the most desirable performance, instead; as **N** grows to medium sizes (20 to 50) model 2 or model 5 scores the top performer alternately. This phenomenon ceases when **N** becomes very large, such as 100 to 200, model 4 again gives the best estimates.

The family of Petersen estimates, models 2, 4 and 5 provide a superb mean for estimation when both p_i's are equal and high(0.7 to 0.9, Tables 12, 13 and 14.) Models 2, 4 and 5 all slightly under-estimate the total and the difference among these three models is trivial. However, the amount of under-estimation is lower in comparing to the cases of unequal but high probabilities(Section 4.1). For high p_i's the ordering of their performance is 2, 5, 4, 3 and 1. This observation is in consistency with the unequal and high probabilities cases.

TABLE 12. Estimated bias and standard error when $p_1 = p_2 = 0.7$

Model	N							
	10	20	30	40	50	100	150	200
1	-3.95	-9.28	-13.94	-18.32	-22.81	-43.34	-62.37	-81.63
	6.14	11.19	16.18	21.04	25.77	47.53	67.21	86.79
2	-0.42	-1.18	-1.59	-1.86	-2.17	-2.47	-1.95	-1.25
	2.30	3.53	4.75	6.02	7.02	11.39	14.54	16.80
3	2.53	4.11	6.37	8.91	11.35	25.79	41.79	57.87
	5.12	7.53	10.56	13.96	16.88	33.07	49.61	65.77
4	-0.64	-1.35	-1.75	-2.03	-2.34	-2.64	-2.13	-1.43
	1.95	3.42	4.68	5.95	6.95	11.33	14.48	16.73
5	-0.51	-1.25	-1.65	-1.93	-2.24	-2.54	-2.02	-1.32
	2.17	3.49	4.72	6.00	6.99	11.37	14.52	16.78

TABLE 13. Estimated bias and standard error when $p_1 = p_2 = 0.8$

Model	N							
	10	20	30	40	50	100	150	200
1	-6.31	-13.35	-20.21	-27.60	-34.03	-65.55	-96.83	-127.65
	6.70	13.80	20.74	28.22	34.78	66.62	98.07	129.03
2	-0.43	-1.10	-1.67	-2.58	-2.78	-3.72	-4.39	-5.09
	1.16	2.13	3.08	4.19	4.90	7.68	10.00	12.04
3	1.38	2.19	3.18	3.58	5.17	13.47	22.16	31.05
	2.47	4.03	5.70	7.01	9.03	18.30	27.57	36.88
4	-0.48	-1.14	-1.72	-2.62	-2.82	-3.77	-4.45	-5.15
	1.12	2.13	3.08	4.19	4.90	7.68	10.00	12.04
5	-0.45	-1.12	-1.69	-2.60	-2.80	-3.74	-4.42	-5.12
	1.14	2.13	3.08	4.19	4.90	7.68	10.00	12.04

TABLE 14. Estimated bias and standard error when $p_1 = p_2 = 0.9$

Model	N							
	10	20	30	40	50	100	150	200
1	-8.02	-16.71	-25.38	-33.93	-42.77	-84.91	-126.69	-168.38
	8.10	16.80	25.50	34.08	42.93	85.14	127.02	168.76
2	-0.20	-0.65	-1.10	-1.44	-2.02	-3.34	-4.53	-5.47
	0.57	1.14	1.72	2.19	2.84	4.75	6.53	7.99
3	0.77	0.98	1.20	1.58	1.58	4.19	7.10	10.32
	1.21	1.89	2.60	3.36	3.84	7.31	11.11	14.71
4	-0.21	-0.66	-1.11	-1.44	-2.03	-3.35	-4.54	-5.48
	0.57	1.14	1.72	2.19	2.84	4.75	6.53	7.99
5	-0.20	-0.66	-1.10	-1.44	-2.03	-3.35	-4.54	-5.47
	0.57	1.14	1.72	2.19	2.84	4.75	6.53	7.99

4.3 Model 5's Behavior

The consistent performance ordering we experienced from our study shows either 4 5 2 or 2 5 4, which is understandable since that the correction proposed by model 5 is robust with given probabilities as well as bounded by models 2 and 4. From Table 15, Model 5 does show improvement over model 4 for low to medium sizes under the combination of middle and high probabilities:

TABLE 15. Estimated bias and standard error when $p_1 = 0.5$ and $p_2 = 0.9$

Model	N							
	10	20	30	40	50	100	150	200
1	-1.44	-4.12	-6.17	-8.20	-10.74	-21.96	-33.99	-44.54
	5.58	7.57	9.76	12.14	14.48	26.78	39.70	50.64
2	0.07	-0.04	0.04	0.04	-0.08	-0.29	-0.62	-0.43
	1.80	2.37	2.93	3.70	4.07	7.11	9.74	12.08
3	3.85	7.16	10.91	14.58	17.88	35.45	52.56	70.88
	5.38	8.65	12.40	16.33	19.58	37.98	55.73	74.42
4	-0.12	-0.19	-0.10	-0.09	-0.21	-0.41	-0.74	-0.54
	1.39	2.20	2.80	3.57	3.97	7.02	9.66	12.01
5	0.00	-0.10	-0.01	0.00	-0.13	-0.33	-0.67	-0.47
	1.65	2.31	2.88	3.65	4.03	7.08	9.71	12.05

5. CONCLUSIONS AND RECOMMENDATIONS

No model seems to perform consistently better than the others. For small values of p(close to 0.1 or 0.2), although none of the models give very precise estimates, model 3 appears to do better, closely followed by model 1, for small N. Thus model 3 or model 1 seems to be appropriate models for correcting the visibility bias for the estimation of the population size N when N is small, i.e. the individuals in a particular study are rare to find and the visibility conditions are very poor, i.e. p_i's are very low. However, as N increases in size the models 4, 5 and 2 perform better and eventually outperform the models 3 and 1. In other words, for individuals who are not rare, even under poor conditions models 4, 5 and 2 seems to be suitable ones. In general the model 4 seems to do better, so far as standard error is concerned, but occasionally, model 5 or 2 performs slightly better than model 4, as far as the bias is concerned. The difference in bias becomes very small as N increases.

As both or one of the p_i's increase(s) to moderate or high values, the models 4, 5 and 2 are in general far better than the models 3 and 1. For standard error the model 4 seems to be consistently better than the others. However, the models 5 and 2 follow closely for standard error but occasionally do better than the models 3 so far bias in estimation is concerned. The difference in performance becomes small as N increases. In short, under adequate to good visibility conditions the model 4 is recommended. When the individuals are in abundance, i.e. N is large, out of models 4, 5 and 2, it is advisable to pick the one that seems to do better as suggested by the graphs for standard error and visibility bias depending upon a particular combination of p_1 and p_2 values.

Model 1 and model 3 occasionally provide good estimates only when values of p_i's and N are all small. Further, model 1 heavily depends on compatible p_i's; in our study only when $p_1 = p_2$ did model 1 give reasonable results although still not as good as the others under same circumstances.

Petersen estimate, model 2, demonstrates constant superiority for the combinations of various high p values. This behavior clearly indicates the adjustment proposed by either model 4 or model 5 is not necessary when the catching probabilities in the survey are excellent. Which also suggests that with highly skilled observers as well as high density of population the difference among p_i's does not affect Petersen estimate's usefulness.

When the difference of probabilities of seeing(capturing) among observers becomes large, the correction suggested by model 4 is mostly needed. This is especially true when the participating observers equipped with significant different levels of surveying experience.

In a more likely situation in which the sample survey is teamed up by one highly skillful observer with the other one little less experienced; then model 5 is the choice for small to moderate sizes of target population. This can also be interpreted as the situation when conducting the survey the objects are highly visible for one direction and less identifiable from the other due to some systematic factors such as the object's habitat. If the difference between $\mathbf{p}_i$'s is not great the model 5 provides the most viable solution to limit the bias.

Following table summarizes our conclusion:

TABLE 16. Recommended model for various combinations

Probabilities		**N**		
p_1	p_2	*small*	*medium*	*large*
low	low	3	3	4
	middle	2	4	4
	high	4	4	4
middle	middle	4	4	4
	high	5	5	2
high	high	2	2	2

In order to do a good job in picking an appropriate model for visibility bias correction, it is strongly recommended that for a given area with known individuals in it, a pilot study should be done to estimate the probabilities of seeing an individual by each of the observers going to participate in the main sample survey. For precise and more reliable estimates, it is also recommended that these kind of sample studies should be undertaken when the visibility conditions are good and both the observers are experienced ones, i.e. both $\mathbf{p}_i$'s are close to 1. In a particular situation, when the background provides a good contrast to observe individuals and there are fewer leaves on the trees. When there is not much choice available for the time when the sampling has to be done, one should try to increase $\mathbf{p}_i$'s by conducting flights at different times over the same area in opposite directions.

REFERENCES

[1] Caughley, G. (1974) "Bias in aerial survey", *Journal of Wildlife Management* 38(4), pp. 921-933.

[2] Caughley, G. and Grice D. (1982) "A correction factor for counting emus from the air, and its application to counts in Western Australia", *Australia Wildlife Research* 9, pp. 253-259.

[3] Chapman, D. G. (1951) "Some properties of the hypergeometric distribution with applications to zoological censuses", *University of California Pub. Stat.* 1, pp 131-160.

[4] Gill, D. S. and Janardan, K. G. (1984) "Models for estimating visibility bias in estimation of nesting bald eagles", in *Proc. Modeling and Simulation 1984*, Volume 15, Part 3, University of Pittsburgh.

[5] Grier, J. D. (1981) "Aerial-visibility bias and survey techniques for nesting bald eagles in Northwestern Ontario", *Journal of Wildlife Management* 45, pp. 83-92.

[6] Henny, C. J., Byrd, M. A., Jacobs, J. A., McLain, P. D., Todd, M. R., and Halla, B. F. (1977) "Mid-Atlantic Coast osprey population: present numbers, productivity, pollutant contamination, and status", *Journal of Wildlife Management* 41, pp. 254-265.

[7] Magnusson, W. E., Caughley, G. J., and Grigg, G. C. (1978) "A double-survey estimate of population size from incomplete counts", *Journal of Wildlife Management* 42(1), pp. 174-176.

[8] Schnabel, Z. (1938) "The estimation of the total fish population of a lake", *American Math. Monthly* 45, pp. 348-352.

[9] Seber, G. A. F. (1973) *The estimation of animal abundance and related parameters*, Griffin: London.