Merlewood Research and Development Paper
Number 44

BIAS IN THE ESTIMATION OF GROUND FLORA PRODUCTION

J. N. R. Jeffers and J. M. Sykes

R & D 71/44

Introduction

A common method of estimating the net primary production of perennial ground flora species in ecological research is to calculate the difference between the seasonal maximum and seasonal minimum mean biomass for each species. Thus, if the biomass per plant is measured on p occasions during the year, the estimated annual production for any species is given by:

$$P = N\,(\mathrm{Max}\ b_i - \mathrm{Min}\ b_i)$$

where N is the plant density per unit area, b_i is the mean individual plant biomass on the ith occasion and Max and Min are taken over the range $i = 1$ to $i = p$.

Intuitively, it seems likely that this estimate of production would be an overestimate of the actual production, despite the fact that the difference is calculated between means rather than single observations. Nevertheless, it is claimed that the estimate could equally well be an underestimate. Any dry matter produced, but which has been eaten or has fallen off before the apparent maximum biomass has been reached, will be expected to displace the apparent maximum from the real maximum biomass. Additionally, it is only by chance that the timing of the sampling would coincide with the time at which the real maximum biomass would be reached; a similar argument can be advanced for the minimum biomass. Whilst the magnitude of the underestimate resulting from these sources is unknown, such a bias could be regarded as counteracting any overestimate resulting from the computation of production from maximum and minimum mean biomass with which this note is concerned.

This note gives the results of a preliminary exploration of the magnitude of the bias which would result from the estimation of production by the standard method applied to known populations.

A model population

As a basis for investigation of the magnitude of the bias likely to arise from the usual method of estimation, an idealized population was first constructed. The essential parameters of this population are given in Figure 1, and are summarised in Table 1. The model assumes a relatively low biomass during the winter months, followed by rapid growth to flowering in July, a further increase in biomass during fruiting in August, and a rapid decline to the winter state by November. It is not suggested that this is a model for any given species, merely a convenient starting point for investigation of the properties of the estimating procedure. The annual production of the model is 95.0g.

At each of the monthly sampling times, it was assumed that a sample of 30 plants was taken and that each plant was dried and weighed before calculating the mean biomass per plant. This operation was assumed to correspond to taking a random sample of 30 values from a normal distribution with the appropriate mean and standard deviation. Table 1 shows four sub-models, the standard deviations being proportional to the mean biomass or to fractions of the mean biomass ranging from 1/2 to 1/10. The minimum mean biomass calculated in this way was subtracted from the maximum mean biomass to provide an estimate of the annual production and the whole process was repeated to represent 30 years' measurements, i.e. giving 30 estimates of annual production based on samples of 30 plants taken at each of 12 occasions.

Table 1. Parameters of population model 1

Month	M Mean biomass (g)	Standard deviation (g) M	M/2	M/5	M/10
Jan	5.0	5.0	2.5	1.0	0.5
Feb	5.0	5.0	2.5	1.0	0.5
Mar	5.0	5.0	2.5	1.0	0.5
Apr	10.0	10.0	5.0	2.0	1.0
May	30.0	30.0	15.0	6.0	3.0
Jun	60.0	60.0	30.0	12.0	6.0
Jul	90.0	90.0	45.0	18.0	9.0
Aug	100.0	100.0	50.0	20.0	10.0
Sept	60.0	60.0	30.0	12.0	6.0
Oct	20.0	20.0	10.0	4.0	2.0
Nov	5.0	5.0	2.5	1.0	0.5
Dec	5.0	5.0	2.5	1.0	0.5

A BASIC program for carrying out this sampling procedure is given in Figure 2.

The results of these sampling experiments are given in Table 2. In all cases, the estimate of annual production was significantly positively biased. Where the standard deviation was taken as being proportional to the mean biomass, the bias was as much as 10 per cent, but for standard deviations proportional to one half, one third, or one tenth of the mean biomass, the overestimate was reduced to approximately 3 per cent. Surprisingly, although the standard errors of the production estimates were markedly reduced by making the standard deviations proportional to increasingly smaller fractions of the mean biomass, the effect of these changes on the extent of the bias was negligible.

Table 2. Estimated mean production and bias corresponding to sub-models of model 1 (N = 30)

Standard deviation	Estimated mean production (g)	Percentage bias	Student's t
M	105.7 ± 2.83	11.3 ± 2.97	3.78***
M/2	97.5 ± 1.06	2.64 ± 1.11	2.36**
M/5	97.8 ± 0.76	2.92 ± 0.80	3.65***
M/10	98.1 ± 0.40	3.25 ± 0.41	7.78***

A second series of sampling experiments was carried out on the model of Figure 1, in which the number of plants sampled on each occasion was reduced to 10, as it seemed likely that a reduction in the precision with which the mean biomass was estimated would have a marked effect on the extent of the bias in the calculation of annual production. The results of this second series of experiments are given in Table 3. Again, there was significant bias in all the estimates of annual production, ranging from 35 per cent where the standard deviation was assumed to be proportional to the mean biomass to about 10 per cent where the standard deviation was assumed to be proportional to one tenth of the mean. Changes in the assumed relationship between the mean and standard deviation of the biomass had a marked effect on both the mean annual production and its standard error when the number of plants sampled on each occasion was reduced to 10.

Table 3. Estimated mean production and bias corresponding to sub-models of model 1 (N = 10)

Standard deviation	Estimated mean production (g)	Percentage bias	Student's t
M	128.7 ± 3.98	35.5 ± 4.18	8.47***
M/2	111.5 ± 2.53	17.4 ± 2.66	6.53***
M/5	105.8 ± 1.11	11.4 ± 1.16	9.75***
M/10	104.2 ± 0.67	9.7 ± 0.70	13.85***

Alternative model

Again, intuitively, it seemed likely that the extent of the bias in the estimation of annual production would be related to the shape of the seasonal biomass curve, particularly as this would influence the probability of a mean biomass estimate becoming a maximum for the year. A second model, with the parameters given in Figure 3 and Table 4, was constructed. The intention of this model was to contrast with the first by having six months at a low biomass and six months at a relatively high biomass, the annual production of the model being 50g.

Table 4. Parameters of population model 2

Month	M Mean biomass (g)	Standard deviation (g) M	M/2	M/5	M/10
Jan	10.0	10.0	5.0	2.0	1.0
Feb	10.0	10.0	5.0	2.0	1.0
Mar	10.0	10.0	5.0	2.0	1.0
Apr	60.0	60.0	30.0	12.0	6.0
May	60.0	60.0	30.0	12.0	6.0
Jun	60.0	60.0	30.0	12.0	6.0
Jul	60.0	60.0	30.0	12.0	6.0
Aug	60.0	60.0	30.0	12.0	6.0
Sept	60.0	60.0	30.0	12.0	6.0
Oct	10.0	10.0	5.0	2.0	1.0
Nov	10.0	10.0	5.0	2.0	1.0
Dec	10.0	10.0	5.0	2.0	1.0

The results of experiments with 30 and 10 plants taken as samples each month are summarised in Tables 5 and 6, respectively. As was expected, the bias was consistently greater in model 2 than in model 1, although the general relationships between the amount of bias and the numbers of sample plants taken each month, and between the amounts of bias and the assumed standard deviations, were similar.

Table 5. Estimated mean production and bias corresponding to sub-models of model 2 (N = 30)

Standard deviation	Estimated mean production (g)	Percentage bias	Student's t
M	67.2 ± 1.27	34.5 ± 2.54	13.55***
M/2	59.3 ± 0.67	18.5 ± 1.35	13.77***
M/5	54.7 ± 0.26	9.5 ± 0.52	18.39***
M/10	53.2 ± 0.11	6.5 ± 0.22	28.91***

Table 6. Estimated mean production and bias corresponding to sub-models of model 2 (N = 10)

Standard deviation	Estimated mean production (g)	Percentage bias	Student's t
M	83.1 ± 2.07	66.1 ± 4.14	15.99***
M/2	69.7 ± 1.08	39.3 ± 2.16	18.22***
M/5	64.6 ± 0.42	23.1 ± 0.84	27.47***
M/10	58.7 ± 0.31	17.4 ± 0.61	28.38***

Preliminary conclusions

The experiments with the two models of plant biomass confirm that the method of estimating annual production which has come to be regarded as "standard" is likely to overestimate the actual production. The extent of the over-estimate is dependent upon the relationship between mean biomass and the standard deviation of biomass of individual plants at different times of the year, the number of plants sampled on each occasion, and the shape of the biomass curve through the year. The effect of the various factors is summarised in Figure 4.

Admittedly, the experiments so far carried out do not test the possible reduction of the estimate of production when the time of sampling did not coincide with the time of maximum biomass, but, except in the most favourable situation of a marked seasonal maximum, samples of at least 30 plants, and relatively small increases in standard deviation with increases in mean biomass, the reduction would need to be greater than 10 per cent to counteract the bias arising from the use of the usual method of estimating production. Similarly, losses through the herbivores or handling would need to be relatively large to counterbalance the bias introduced by the calculation of production from minimum and maximum measures of biomass.

Sampling of ground flora species

It is of some interest to consider the possible extent of the bias for individual species studied in a major study of a woodland ecosystem at Meathop Wood, Grange-over-Sands, Lancashire.

The relationships between the observed means and variances for each of the species were examined by plotting the variances against the means on logarithmic-scaled graph paper. Table 7 summarises the observed relationships, and the numbers of plants examined on each sampling occasion. For all species, except Arum maculatum, Potentilla sterilis, and Endymion non-scriptus, the standard deviation was approximately proportional to one half of the mean biomass. For Potentilla sterilis and Endymion non-scriptus, the standard deviation was approximately proportional to the square of the mean, while, for Arum maculatum, the relationship was ambiguous, and the standard deviation could be described as being proportional to either one half of the mean biomass or to the square of one half of the mean biomass.

Table 7. Observed relationships between mean and standard deviation of plant biomass

Species	Standard deviation as function of mean: M	M/2	M/5	M^2	$(M/2)^2$	Number of plants
Viola riviniana		✓				30
Fragaria vesca		✓				30
Fraxinus excelsior (seedlings)		✓				30
Oxalis acetosella		✓				30
Sanicula europaea		✓				30
Geum urbanum		✓				20
Primula vulgaris		✓				10
Arum maculatum		✓			✓	10
Potentilla sterilis				✓		30
Endymion non-scriptus				✓		30

Given these relationships, and the number of plants sampled on each occasion, sampling experiments similar to those undertaken on the model populations might be expected to estimate the extent of the bias. However, information on the shape of the seasonal biomass curve would be necessary, because of the dependence of the bias on the shape of this curve. The only practical evidence of the shape of the curves available is the mean biomass calculated for each sampling of the individual species, summarised in Table 8. By using these means as approximations, the results of Table 9 have been obtained.

Four of the species, i.e. Fragaria vesca, Oxalis acetosella, Sanicula europaea and Geum urbanum, had no significant bias in their estimates of annual production. All the remaining species had significant overestimates of annual production, except for Potentilla sterilis, where there was a significant underestimate. Apart from Arum maculatum, where the overestimate was at least 45 per cent, the general level of bias was approximately 10 per cent.

Table 9. Estimated bias for annual production of individual species

Species	SD	Estimated mean production	Percentage bias	Student's t
Viola riviniana	M/2	0.223 ± 0.0044	11.54 ± 2.19	5.26***
Fragaria vesca	M/2	0.485 ± 0.0112	-1.05 ± 2.29	0.45
Fraxinus excelsior	M/2	0.337 ± 0.0078	9.08 ± 2.53	3.59***
Oxalis acetosella	M/2	0.103 ± 0.0022	-1.91 ± 2.13	0.89
Sanicula europaea	M/2	1.223 ± 0.0269	1.95 ± 2.24	0.84
Geum urbanum	M/2	1.530 ± 0.0411	1.31 ± 2.72	0.48
Primula vulgaris	M/2	6.95 ± 0.229	10.9 ± 3.65	2.98***
Arum maculatum	M/2	3.39 ± 0.095	45.1 ± 4.08	11.03***
" "	$(M/2)^2$	3.96 ± 0.187	69.6 ± 8.01	8.67***
Potentilla sterilis	M^2	0.773 ± 0.0201	-7.57 ± 2.40	-3.15***
Endymion non-scriptus	M^2	0.576 ± 0.0119	11.54 ± 2.31	4.99***

Estimation of bias in the hectare

If the calculations of percentage bias in Table 9 can be regarded as reasonable estimates, the values can be used to derive the approximate percentage bias for the annual production of ground flora per hectare. Table 10 gives the estimated numbers of each species occurring in the sample hectare.

Combining the percentage bias, number of plants per hectare, and estimated production for each species gives an overall percentage bias of 8.59 ± 1.27 per cent for the sample hectare.

Table 10. Estimated numbers of plants of the various species in the sample hectare

Species	Number per hectare
Viola riviniana	286,520
Fragaria vesca	29,040
Fraxinus excelsior	483,480
Oxalis acetosella	481,440
Sanicula europaea	33,880
Geum urbanum	13,200
Primula vulgaris	920
Arum maculatum	1,080
Potentilla sterilis	31,000
Endymion non-scriptus	616,200
TOTAL	1,976,760

Table 8. <u>Mean biomass for monthly samples of ground flora species</u>

Species	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
	Mean biomass per plant (g)											
Viola riviniana	0.129	0.229	0.178	0.186	0.271	0.224	0.329	0.195	0.209	0.309	0.139	0.198
Fragaria vesca	0.164	0.172	0.127	0.106	0.323	0.596	0.310	0.317	0.365	0.152	0.153	0.157
Fraxinus excelsior	0.342	0.291	0.216	0.188	0.138	0.157	0.208	0.447	0.435	0.301	0.371	0.289
Oxalis acetosella	0.048	0.039	0.030	0.061	0.058	0.056	0.076	0.065	0.135	0.054	0.038	0.036
Sanicula europaea	0.344	0.304	0.325	0.339	0.761	1.505	1.014	0.842	0.533	0.322	0.360	0.436
Geum urbanum	0.274	0.282	0.306	0.326	0.803	1.784	0.995	1.037	0.857	0.931	0.375	0.289
Primula vulgaris	2.88	2.91	6.83	4.34	5.50	4.66	3.09	9.15	6.46	5.45	6.75	5.73
Arum maculatum	1.798	1.742	2.097	2.401	3.831	4.029	4.077	2.367	2.952	1.862	1.779	2.449
Potentilla sterilis	0.440	0.946	0.565	0.497	1.196	0.849	0.537	0.524	0.603	0.518	0.360	0.669
Endymion non-scriptus	0.192	0.237	0.251	0.310	0.641	0.613	0.708	0.456	0.301	0.316	0.233	0.310

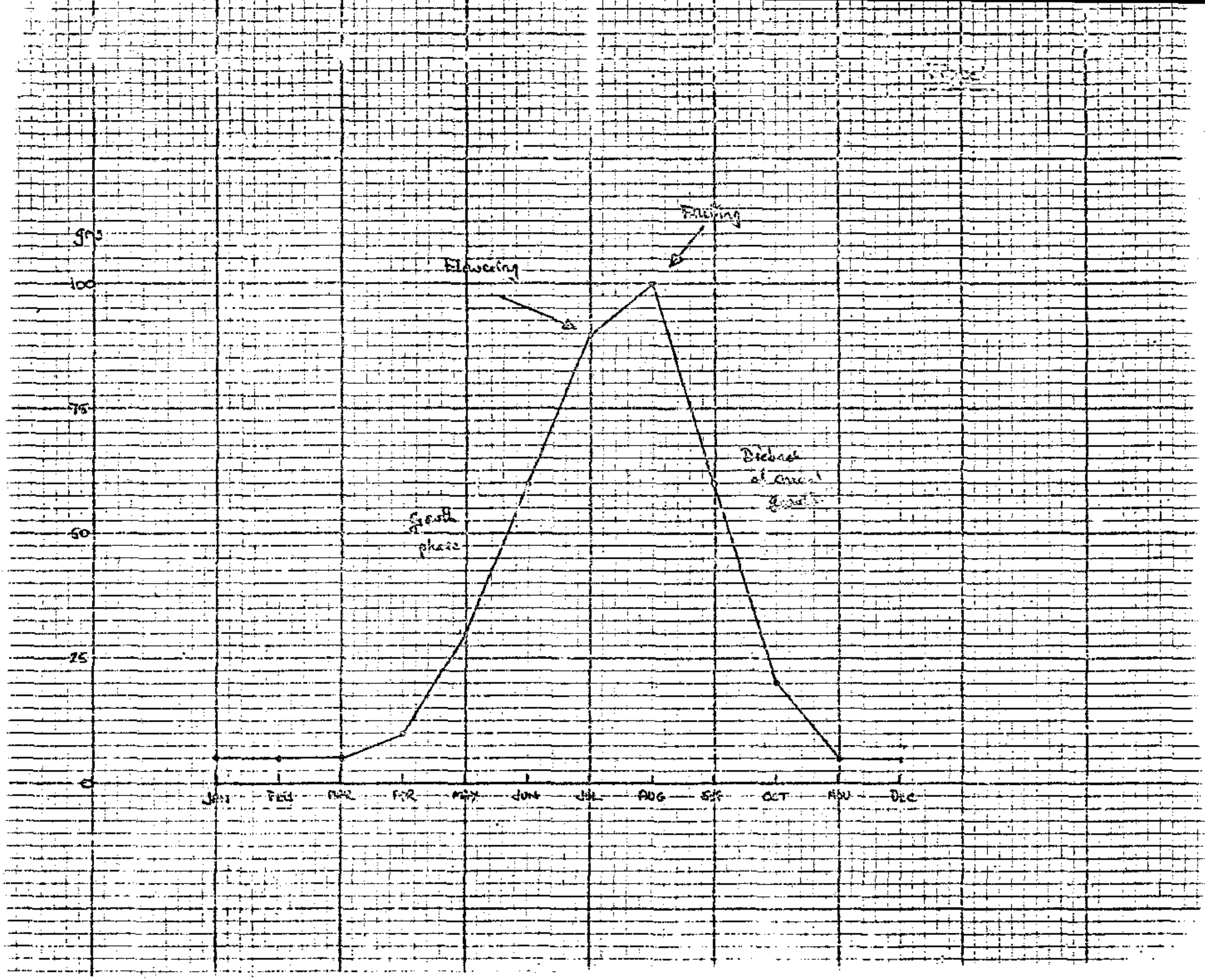

grs
100
75
50
25
0
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
DEC
Flowering
Fruiting
Growth phase
Dieback of aerial growth

Figure 2. BASIC program for sampling experiments

```
TAPE

10 REM PROGRAM TO TEST BIOMASS ESTIMATES
20 DEF FNR(X)=SQR(-LOG(RND(0)))*(COS(6.283185*X)+SIN(6.283185*X))
30 DEF FNP(M)=M+M*(FNR(RND(0)))
34 RANDOMIZE
35 LET A=0
38 LET S=0
40 FOR I=1 TO 30
50 PRINT
60 PRINT
70 LET H=0
80 LET L=100
90 LET P=0
100 FOR J=1 TO 12
110 READ M
120 FOR K=1 TO 30
130 LET E=FNP(M)
140 LET P=P+E
150 NEXT K
160 LET P=P/30
170 IF P<H THEN 190
180 LET H=P
190 IF P>L THEN 210
200 LET L=P
210 PRINT P,
220 NEXT J
230 LET B=H-L
240 PRINT " "," ",B
250 RESTORE
254 LET S=S+(B-A)*(B-A)*(1-1/I)
256 LET A=A+(B-A)/I
260 NEXT I
280 DATA 5,5,5,10,30,60,90,100,60,20,5,5
300 LET S=SQR((S/29)/30)
310 PRINT
320 PRINT "ESTIMATED MEAN BIOMASS";A;"SE";S;"T";(A-95)/S
330 STOP
340 END

READY
```

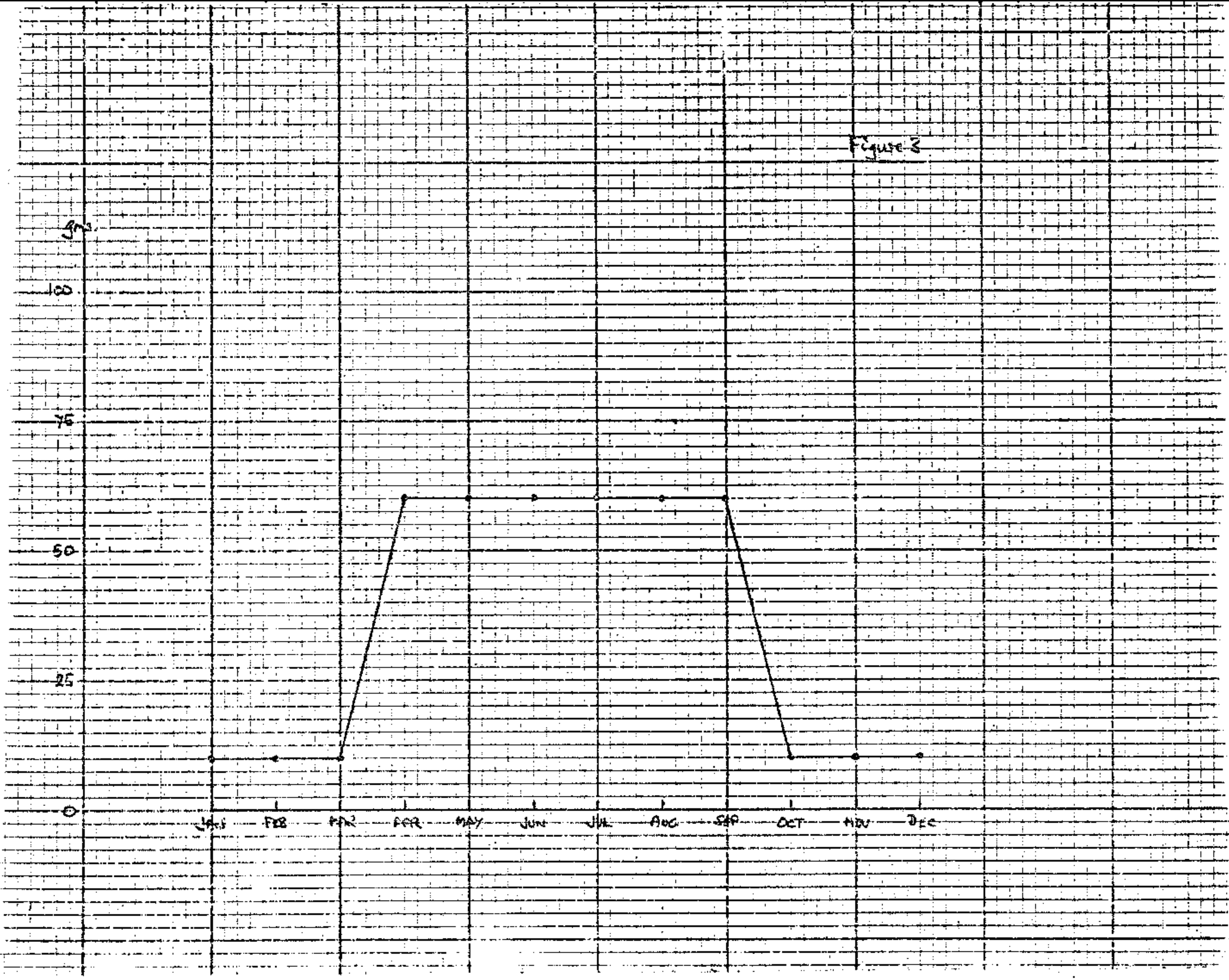
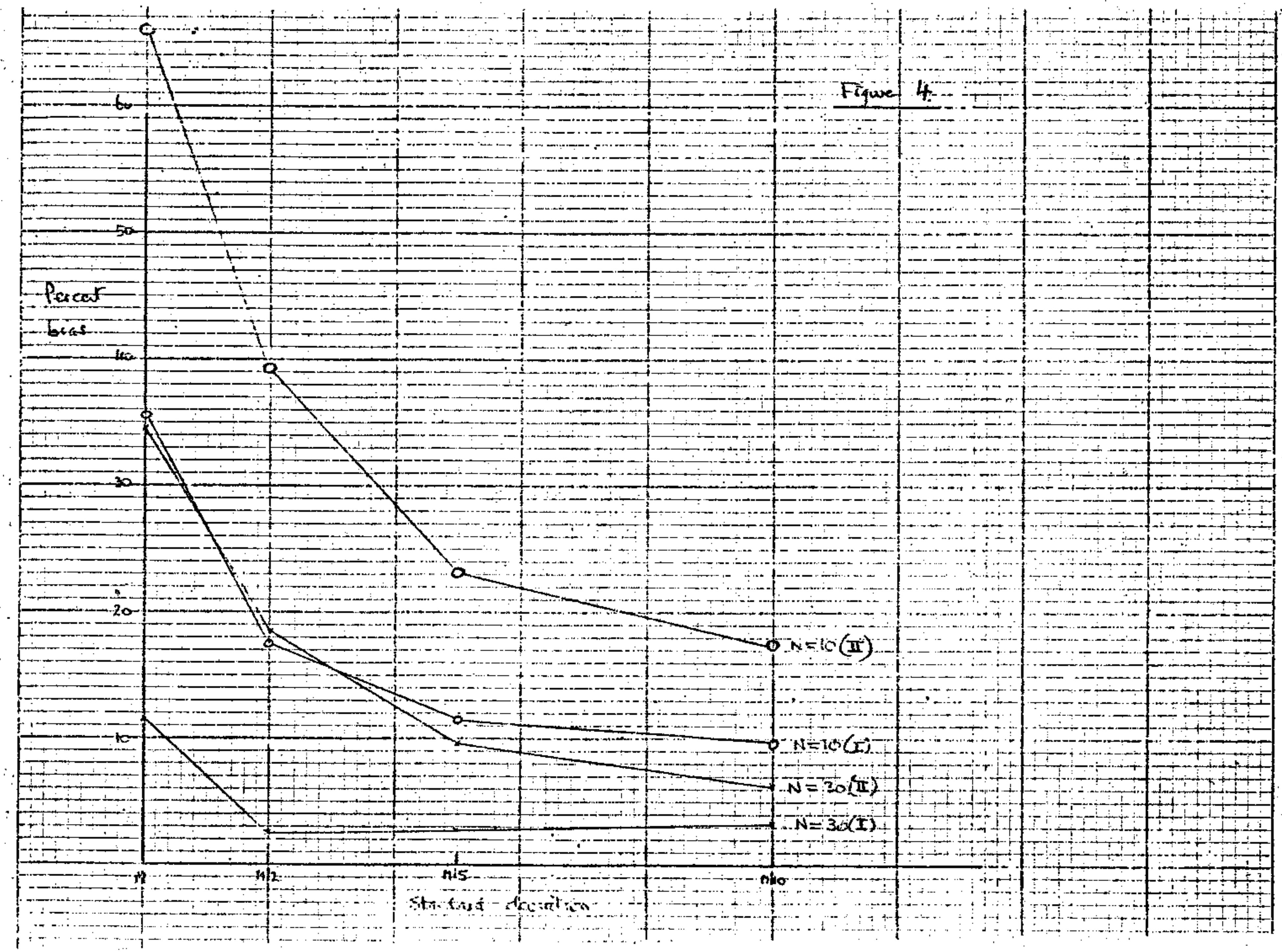


Figure 3

Figure 4.



Percent bias
60
50
40
30
20
10
N=10(II)
N=10(I)
N=30(II)
N=30(I)
Standard deviation