Ecological Modelling, 44 (1988) 43–55
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

BIAS IN ESTIMATES OF PRIMARY PRODUCTION: AN ANALYTICAL SOLUTION

O.E. SALA [1], M.E. BIONDINI [2] and W.K. LAUENROTH [1]

[1] *Natural Resource Ecology Laboratory and Department of Range Science, Colorado State University, Fort Collins, CO 80523 (U.S.A.)*
[2] *Department of Animal and Range Science, North Dakota State University, Fargo, ND (U.S.A.)*

(Accepted 6 June 1988)

ABSTRACT

Sala, O.E., Biondini, M.E. and Lauenroth, W.K., 1988. Bias in estimates of primary production: an analytical solution. *Ecol. Modelling*, 44: 43–55.

This paper addresses the issue of the effect of random error upon the estimates of above- or belowground net primary production (NPP). We show that random errors in estimates of biomass do not compensate but always result in a positive bias in estimates of NPP. Second, we demonstrate that the larger the number of time intervals considered, the higher is the positive bias or overestimation error. An effect similar to an increase in the number of sampling periods is obtained by increasing the number of components utilized in estimating NPP. These are usually taxonomic or functional (depth, live, recent dead, etc.) components.

We calculate the magnitude of the overestimation error as a function of the size of the difference in two sequential estimates of biomass and the variability associated with them. We propose a method that uses this error to correct the estimates of NPP for the positive bias resulting from random errors and to develop confidence intervals for the corrected NPP. We suggest that this method will remove the positive bias from estimates of NPP upon which nutrient budgets, energy flow and trophic webs rely. The concepts presented will help in the design of experiments that use production as a response variable.

INTRODUCTION

Ecologists are presently facing an important controversy about the effect of random error on estimates of primary production (Singh et al., 1984; Lauenroth et al. 1986; Vogt et al., 1986). This controversy is relevant because at stake are: (1) the validity of past estimates of production upon which nutrient budgets, energy flows, and trophic webs rely; and (2) the design of future experiments that use production as a response variable.

Singh et al. (1984) demonstrated, using a simulation model, that the most commonly used techniques overestimated net root production by as much as

0304-3800/88/$03.50

700%. Overestimation occurred even when they used only significant differences in biomass to calculate production. This result contradicted previous assumptions that estimates of net root production based upon a time series of biomass always underestimated the true production because biomass peaks were missed and because of the simultaneous nature of biomass input and output processes (production and decomposition). Vogt et al. (1986) challenged the conclusions of Singh et al. (1984), suggesting that they were not generally applicable.

The overestimation question outlined by Singh et al. (1984) is not restricted to net root production but applies to any estimate of above- or belowground production based upon a time series of biomass. The objectives of this paper are: (1) to show that random error in estimates of biomass will always result in a positive bias in estimates of above- or belowground net primary production (NPP); (2) to demonstrate that the larger the number of time intervals considered, the higher the positive bias or overestimation error; (3) to estimate the magnitude of the overestimation error and to use it to correct the estimates of NPP; and (4) to develop confidence intervals for the corrected estimates of NPP.

The techniques most commonly utilized to estimate aboveground or belowground net primary production require a time series of biomass estimates. Biomass is usually estimated by the harvest technique, in which aboveground biomass is harvested from several plots or roots washed from soil cores. Biomass is sometimes divided into functional or taxonomic components (live, recent dead, old dead, depth, species, etc.). Aboveground biomass is frequently estimated by one of the several available double-sampling techniques (Wilm et al., 1944).

There are several ways of estimating NPP from a time series of biomass (Singh et al., 1975). The simplest way equates NPP with standing crop at the end of the growing season. The more complicated methods take into consideration the different functional (Wiegert and Evans, 1964) and specific components (Sala et al., 1981) of biomass. These methods attempted to correct the errors leading to underestimation (ELU) of NPP. ELU arise because of a difference between the concept of net primary production and the method used to calculate it from data. Net primary production is defined as the difference between the energy fixed by autotrophs and their respiration (Odum, 1971). Most commonly, NPP is calculated from biomass data, and production is equated with increments in biomass and ELU result from such things as missing peaks in biomass and the simultaneous nature of increments and losses. Consider $\text{NPP} = B_2 - B_1$ where B_i is biomass of time t_i. There is always the possibility that biomass reached a value higher than B_2 during the time interval (t_1, t_2) and therefore the estimated value of NPP is smaller than the true value of NPP. The difference $B_2 - B_1$ is a result of the balance of production and decomposition or production and senescence if we are considering only green biomass. Decomposition can occur simultaneously with production and can mask production, making the estimated NPP smaller than the true NPP. For a species the two processes may dominate in different phases of its life cycle. For plant communities made up of several species where life cycles are not in synchrony, this source of error is important.

Singh et al. (1984) demonstrated that there are other kinds of errors, those leading to overestimation (ELO) of NPP. Random errors in estimates of biomass do not compensate but accumulate leading to positive bias in estimates of NPP. Random errors do not compensate because of the definition of production. Production is estimated by the increment in biomass during a period of time. A negative difference in biomass yields a production value of zero. Consider the case in which the true increment in biomass is zero. Because of the variability associated with the estimates of biomass at time 1 (B_1) and at time 2 (B_2); when this system is sampled, B_2 will sometimes appear smaller than B_1 and the estimated production for this period will be zero. Sometimes the reverse will occur; B_2 will appear greater than B_1, and this positive difference will be accepted as production. When random error results in a negative difference, it is ignored; and when it results in a positive difference, it is accepted as production. Singh et al. (1984) demonstrated that applying statistical constraints designed to accept only 'significant' increases will reduce but not eliminate this problem. This paper will address the occurrence and magnitude of ELO as well as the different ways to correct NPP estimates.

RANDOM ERROR AND THE OVERESTIMATION OF NPP

Define B_1 and B_2, the biomass at time 1 and time 2, as the random variables associated with *sampling error* in two consecutive sampling periods. Furthermore, let's assume that both have a normal distribution with mean μ_i, standard deviation $\sigma_i = 1, 2$ [$B_i \sim N(\mu_i, \sigma_i)$]. Without loss of generality, we also assume that $B_2 > B_1$. The random variable representing the difference between these two variables is also normally distributed:

$$D = (B_2 - B_1) \sim N\left[\mu = \mu_2 - \mu_1,\ \sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - \mathrm{Cov}(B_1, B_2)}\right]$$

NPP is not the difference between B_2 and B_1 (D) but it is only the positive values of this difference. Production is estimated as the increment in biomass, and ranges between 0 and infinity. Therefore, the random variable NPP is defined as:

$$\mathrm{NPP} = \begin{cases} 0 \text{ when } (B_2 - B_1) \leq 0 \\ (B_2 - B_1) \text{ when } (B_2 - B_1) > 0 \end{cases}$$

The distribution function of NPP [F_{NNP}(NPP)] is *not normal* but a combination of a *discrete* random variable with mass at 0 and a *continuous* random variable with a truncated normal distribution:

$$F_{NNP}(\text{NPP}) = q + p\, G_{CNPP}(\text{NPP}) \tag{1}$$

where $p = P[D > 0]$; $q = 1 - p$; and:

$$G_{CNPP}(\text{NPP}) = \frac{1}{\sqrt{2\pi}\,\sigma p} \int_0^{\text{NPP}} e^{-\frac{1}{2}((x-\mu)\sigma)^2}\, dx \tag{2}$$

The mean of our estimate of production [E(NPP)] and its variance [Var(NPP)] then are:

$$E(\text{NPP}) = p\mu + \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mu/\sigma)^2} \tag{3}$$

$$\text{Var}(\text{NPP}) = p\mu^2(1-p) + p\sigma^2 + \frac{\mu\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mu/\sigma)^2}(1-2p) - \frac{\sigma^2}{2\pi} e^{-(\mu/\sigma)^2} \tag{4}$$

See Korn and Korn (1968) for a proof.

In order to satisfy our first objective we show in Appendix A that the expected value of NPP [E(NPP)] is greater than the true increment in biomass between t_1 and $t_2(\mu)$. The difference between E(NPP) and μ, is then called the overestimation error (OE):

$$\text{OE} = \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mu/\sigma)^2} - q\mu \tag{5}$$

where $\mu = \mu_2 - \mu_1$ is the true difference in biomass between t_1 and t_2; $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - \text{Cov}(B_1, B_2)}$ is the standard deviation of the difference; and q is the probability of $B_2 - B_1 \leq 0$. The overestimation error is a function of the variability in the estimate of the difference in biomass between t_1 and t_2 and the magnitude of the true difference. The larger the variability associated with the estimate of B_2 and B_1, the higher will be the OE. The larger the true increment in biomass (μ) the smaller will be the OE.

We have chosen an example in which μ is the same for two cases. The standard deviation of the difference (μ) in the first case is half of σ in the second case (Table 1). The probability p of $B_2 - B_1 > 0$ was calculated as the probability of a normal distribution with mean μ and standard deviation σ. We estimated OE using equation (5). The increase in variability from case 1 to case 2 resulted in an increase in the probability of obtaining a value smaller than 0. It, in turn, resulted in an increase in the mean of NPP E(NPP). Therefore, the difference between the means of the normal and NPP distributions increased. This is the difference between the expected value of NPP and the true increment in biomass, the overestimation error, which increased from 11% to 46%.

We can now look at a second example to demonstrate the effect of the true difference upon the overestimation error. We compared case 2 of the previous example with a case (3) which has the same σ but a smaller true difference (Table 1). The decrease in the magnitude of the true difference resulted in an increase in the OE from 4.6 g/m^2 to 6.5 g/m^2. The decrease in the true difference determined an increase in the probability of obtaining a value smaller than 0. It forces the mean of the NPP distribution away from the mean of the normal distribution. This difference between the means of both distributions is the overestimation error, which increased from case 2 to case 3 from 46% to 130%.

NUMBER OF TIME INTERVALS AND THE OVERESTIMATION ERROR

The second objective of this paper is to demonstrate that the larger the number of time intervals considered, the higher will be the overestimation error. Let's assume again that B_1 and B_2 are the random variables associated with the sampling errors of biomass estimates at time t_1 and t_2. We will first look at a numerical example in which B' (the random variable associated with sampling errors at t') has a mean (μ') which is half way between μ_2 and μ_1; second, we will prove that the same increase of OE occurs regardless of the location of μ' as long as $\mu_1 < \mu' < \mu_2$. We calculated, using the data of case 2 (Table 1), the OE associated with the estimate of NPP$_{(2)}$ for two sampling dates (t_1, t_2), $\text{OE}_{t_1,t_2} = 4.6$ g/m^2. The errors associated with the sampling periods t_1, t' and t', t_2 ($\text{OE}_{t_1,t'} = 6.5$ g/m^2; $\text{OE}_{t',t_2} = 8.9$ g/m^2) were larger than the t_1, t_2 error because the difference in biomass decreased. The error in the first sampling period was smaller than

TABLE 1

Examples that demonstrate the effect of the variability associated with the difference in biomass between t_1 and t_2 and the magnitude of the true difference upon the overestimation error

	Case 1	Case 2	Case 3
Biomass at time 1, g/m^2 (B_1)	110	110	110
Biomass at time 2, g/m^2 (B_2)	120	120	115
$B_2 - B_1$ (μ)	10	10	5
Standard deviation of B_1 (σ_1)	5	10	10
Standard deviation of B_2 (σ_2)	10	20	20
Standard deviation of $B_2 - B_1$ (σ)	11	22	22
Standard normal deviate (z)	−0.91	−0.45	−0.23
Probability of $B_2 - B_1 > 0$ (p)	0.82	0.68	0.59
Overestimation error, g/m^2 (OE)	1.1	4.6	6.5

in the second because the variance was smaller in the former and the difference in biomass was the same. The total overestimation error associated with the estimate of $\mathrm{NPP}_{(3)}$ using three sampling dates was 15.4 g/m^2, which was almost three times larger than the error for two sampling dates 4.6 g/m^2.

We will now show the generality of the relationship between the number of sampling dates and the overestimation error. The random variable associated with the sampling errors at t' is $B' \sim N(\mu', \sigma')$ and lies within the μ_1, μ_2 range. $\mathrm{NPP}_{(3)}$ is the random variable associated with an estimate of net primary production using three sampling dates (t_1, t', t_2) and $\mathrm{NPP}_{(2)}$ is the random variable associated with the estimate of net primary production obtained using only two sampling dates (t_1, t_2). We proved in Appendix B that:

$$E\left[\mathrm{NPP}_{(3)}\right] > E\left[\mathrm{NPP}_{(2)}\right] \qquad \text{if} \quad \mathrm{CV}_{B_1} = \mathrm{CV}_{B_2} = \mathrm{CV}_{B'} = b$$

where μ is the true NPP and CV_{B_i} are coefficients of variation associated with biomass estimates at time t_i.

Our results indicate that ELO will be increased by increasing the number of sampling periods. An effect similar to an increase in the number of sampling periods is obtained by increasing the number of components utilized in estimating NPP. We can demonstrate using a similar analysis that the positive bias, resulting from random error, in root production estimated by layers is always greater than when estimated as differences in total biomass as a consequence of the increase of ELO. The overestimation error associated with estimating aboveground production as the sum of the production by individual species is higher than when production is estimated as differences in total biomass.

Ecologists have traditionally interpreted estimates of net primary production in a manner that emphasized the errors leading to underestimation. This has resulted in higher estimates being considered better than lower ones. There are several good reasons why this has been true. The most important of these, for our purposes, is that although the concept of net primary production is simple, it cannot be measured directly. Consequently, it must be estimated using components of the ecosystem which can be measured. Which components should be measured? From a conceptual point of view the answer is simple. Measure as many components as the budget will allow. The more components included in the calculation, the closer the method of calculation will be to the concept of net primary production. This is the rule that has guided ecologists for the past 25 years of estimating primary production.

Odum (1960) argued that the sum of the peak standing crops of individual species was greater than the peak standing crop of the community and therefore provided a better estimate of net primary production. Aber et al. (1985) compared estimates of net root production using two methods and concluded that the one which contained the larger number of components was the best. Dickerman et al. (1986) compared seven techniques for calculating net primary production for a wetland dominated by *Typha latifolia*. They concluded that most techniques underestimated production by failing to include such things as early shoot mortality, leaf turnover, and losses of portions of individual leaves. In each case the authors are correct in their statements that accounting for more processes results in conceptually more correct estimates of net primary production and lower underestimation errors. What is missing in each case is an appreciation of the effect of the uncertainty associated with each component in the calculation of the estimate of net primary production.

In general, the greater the uncertainty associated with the components, the greater the uncertainty associated with the results (O'Neill, 1973). In the case of production, increased uncertainty leads to increased overestimation. The reason for this is that values less than zero are rejected. Singh et al. (1984) introduced the idea of errors leading to overestimation. Our analysis has provided an explanation for why random errors in estimates of biomass, or any other component used in calculation of net production, always lead to overestimates of the true difference between samples collected on two dates.

How does one choose the best way to estimate production? It is clear that estimates with few components and few sample dates will have the lowest overestimation error. It is equally clear that estimates with the largest number of components and the most frequent sampling will have the lowest underestimation error. The answer to the question of which estimate is best depends upon the relationship between the objectives of the project and the system studied. The objectives place boundaries on the total error that will be acceptable and the system being studied determines the relationship between the overestimation and underestimation errors. Within the window of acceptable error one must evaluate the tradeoff between the importance of the closeness of the calculation procedure to the concept of NPP, and the size of the overestimation error. If it were possible to correct for part of all of the overestimation error, we could break out of this tradeoff situation. We will explore this possibility in the next section.

MAGNITUDE OF THE OVERESTIMATION ERROR AND THE CORRECTED NPP

Knowledge of the magnitude of the overestimate will not only aid in the selection of the appropriate experimental design but also will be instrumental in correcting estimates of NPP. We have shown in equations (3) and (4) the relationship between the true value of NPP (μ) and its variance (σ^2) and the expected value of the estimated NPP [E(NPP)] and the sample variance. If p, the proportion of cases in which $B_2 - B_1 > 0$, can be estimated, we will have reduced our problem to a system of two equations with two unknowns, which can be solved numerically. We estimate p using the estimated D ($D = B_2 - B_1$) and its standard deviation (SD). We first calculate the normal deviate z:

$$z = \frac{-D}{\mathrm{SD}}$$

and, second, using a table for the standard normal curve, we look for the probability $P\ (Z > z)$ that is the value of p, the probability of $B_2 - B_1 > 0$.

The corrected value of NPP ($\hat{\mathrm{NPP}}$) can be obtained using the estimate of p, the sample mean and the sample variance for a pair of biomass estimates. We include in Appendix C a simple algorithm to solve this system of two equations with two unknowns. The algorithm can be easily implemented in a hand-held programmable calculator. If more than two estimates of biomass are used in estimating production, $\hat{\mathrm{NPP}} = \hat{\mathrm{NPP}}_{\mathrm{i}} + \ldots + \hat{\mathrm{NPP}}_n$, where $\hat{\mathrm{NPP}}_i$ is the corrected estimate of production between i and $i+1$.

CONFIDENCE INTERVALS FOR $\hat{\mathrm{NPP}}$

The direct approach for estimating confidence intervals for $\hat{\mathrm{NPP}}$ would imply finding the distribution of $\hat{\mathrm{NPP}}_i$. However, this is very difficult, since $\hat{\mathrm{NPP}}_i$ is obtained from a nonlinear system of equations with no analytical solution. Therefore, the best way to obtain confidence intervals is to use the central limit theorem (Chow and Teicher, 1978). The distribution of $\hat{\mathrm{NPP}}$ can be approximated by the normal distribution, with:

$$\hat{\mathrm{NPP}} = \sum_{i=1}^{n} \hat{\mathrm{NPP}}_i$$

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \sigma_i^2$$

so the confidence interval CI ($\hat{\mathrm{NPP}}$) is

$$\mathrm{CI}(\hat{\mathrm{NPP}}) = \hat{\mathrm{NPP}} \pm t_{\alpha/2}(n-1)\hat{S} * N^{-1/2}$$

where

$$\hat{S} = \left[\sum_{i=1}^{n} \hat{\sigma}_i^2\right]^{1/2}$$

$\hat{\sigma}_i^2$ are the estimated sample variances, and α is the desired level of type 1 error.

ACKNOWLEDGEMENTS

This paper is a contribution of the shortgrass steppe Long Term Ecological Research program and reports on research funded by the National Science Foundation grant BSR-8114822. We would like to thank R. Leiva for his input during the early development of the ideas which led to this paper and H.W. Hunt and D.M. Swift for their valuable suggestions.

APPENDIX A

Proof that E(NPP) is greater than the true NPP (μ) and calculation of the overestimation factor. If $D = (B_2 - B_1) \sim N(\mu, \sigma)$, where μ and σ are as defined in Appendix A, then:

$$\mu = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\,\sigma} d\, \mathrm{e}^{-\frac{1}{2}((d-\mu)/\sigma)^2}\, \mathrm{d}d + \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} d\, \mathrm{e}^{-\frac{1}{2}((d-\mu)/\sigma)^2}\, \mathrm{d}d$$

$$\leq \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} d\, \mathrm{e}^{-\frac{1}{2}((d-\mu)/\sigma)^2}\, \mathrm{d}d$$

Then:

$$\mu \leq \frac{1}{\sqrt{2\pi}} \int_{-\mu/\sigma}^{\infty} (\mu + w\sigma)\, \mathrm{e}^{-\frac{1}{2}w^2}\, \mathrm{d}w = p\mu + \frac{\sigma}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}(\mu/\sigma)^2} = E(\text{NPP})$$

where $w = (d - \mu)/\sigma$ and p is the probability of $D > 0$. The overestimation factor (OE) is:

$$\text{OE} = \left| \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\,\sigma} d\, \mathrm{e}^{-\frac{1}{2}((d-\mu)/\sigma)^2}\, \mathrm{d}d \right|$$

$$= \frac{\sigma}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}(\mu/\sigma)^2} - q\mu$$

where q is the probability of $D < 0$.

APPENDIX B

Let's assume that B_1, B_2 and B' are as defined in the main text and have a normal distribution with:

$$\mu_1 = \mathrm{E}(B_1)$$

$$\mu_2 = \mathrm{E}(B_2)$$

$$\mu' = \mathrm{E}(B') = \mathrm{E}(B_1) + a[\mathrm{E}(B_2) - \mathrm{E}(B_1)]$$

where $0 < a < 1$. Let's also assume that $\text{CV}_{B1} = \text{CV}_{B2} = \text{CV}_{B'} = b$, where CV stands for coefficient of variation. We also assume that B_1, B_2 and B' are independent since they represent the random variables associated with the errors in estimating biomasses at time t_i. We want to emphasize that we are *not* saying that the biomasses are independent (they are *not random either!*) but that the errors associated in estimating the biomasses from a limited number of sample are independent (something guaranteed by adequate randomization).

(a) We first prove that:

$$\mathrm{Var}(B' - B_1) = \sigma_1^2 > a^2 \ \mathrm{Var}(B_2 - B_1) = a^2\sigma^2$$

and

$$\mathrm{Var}(B_2 - B') = \sigma_2^2 > (1-a)^2 \ \mathrm{Var}(B_2 - B_1) = (1-a)^2\sigma^2$$

Because of our assumption of equal coefficients of variation ($\mathrm{CV}_{B_i} = b = \mathrm{CV}_{B'}$; $i = 1, 2$):

$$\mathrm{Var}(B_1) = b^2 \ \mathrm{E}(B_1)^2 = b^2\mu_1^2$$

$$\mathrm{Var}(B_2) = b^2 \ \mathrm{E}(B_2)^2 = b^2\mu_2^2$$

$$\begin{aligned}\mathrm{Var}(B') &= b^2[\mathrm{E}(B_1) + a \ \mathrm{E}(B_2 - B_1)]^2 \\ &= b^2[\mu_1 + a(\mu_2 - \mu_1)]^2\end{aligned}$$

Then:

$$\begin{aligned}\mathrm{Var}(B' - B_1) = \sigma_1^2 &= b^2\mu_1^2 + b^2[\mu_1 + a(\mu_2 - \mu_1)]^2 \text{ (because of independence)} \\ &= b^2(2\mu_1^2 + 2a\mu_1\mu_2 - 2a\mu_1^2 + a^2\mu_2^2 - 2a^2\mu_2\mu_1 + \mu_1^2a^2) \\ &= a^2b^2(\mu_1^2 + \mu_2^2) + b^2[2\mu_1^2 + 2a\mu_1\mu_2(1-a) - 2a\mu_1^2]\end{aligned}$$

Because we assume that $\mu_2 > \mu_1$, then:

$$\sigma_1^2 > a^2b^2(\mu_1^2 + \mu_2^2) + b^2[2\mu_1^2(1 - a^2)]$$

Now $b^2[2\mu_1^2(1 - \mathrm{a}^2)] > 0$ and $\sigma^2 = b^2(\mu_1^2 + \mu_2^2)$. So $\sigma_1^2 > \mathrm{a}^2\sigma^2$, which implies:

$$\frac{a}{\sigma_1} < \frac{1}{\sigma}$$

In a similar manner we can establish that $\sigma^2 > (1-a)^2\sigma^2$, which implies:

$$\frac{(1-a)}{\sigma_2} < \frac{1}{\sigma}$$

(b) We now prove that $\mathrm{ENPP}_{(3)} > \mathrm{ENPP}_{(2)}$, where $\mathrm{ENPP}_{(i)}$, $i = 2, 3$, are defined as in the main text.

In order to prove that $\mathrm{E}[\mathrm{NPP}_{(3)}] > \mathrm{E}[\mathrm{NPP}_{(2)}]$, it suffices to show that $\mathrm{OE}_{(3)} - \mathrm{OE}_{(2)} > 0$, where OE_i: overestimation of true NPP, using i sampling dates. Define:

$$\Psi(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x^2} \quad \text{and} \quad \mu=B_2-B_1$$

so

$\mu'-\mu_1=a\mu$

$\mu_2-\mu'=(1-a)\mu$

Then:

$$\begin{aligned}\mathrm{OE}_{(3)}-\mathrm{OE}_{(2)}=&\left|\int_{-\infty}^{-a\mu/\sigma_1}\sigma_1 x\,\Psi(x)\,\mathrm{d}x\right|+\left|\int_{-\infty}^{-(1-a)\mu/\sigma_2}\sigma_2 x\,\Psi(x)\,\mathrm{d}x\right|\\&-a\mu\int_{-\infty}^{-a\mu/\sigma_1}\Psi(x)\,\mathrm{d}x-(1-a)\mu\int_{-\infty}^{-(1-a)\mu/\sigma_2}\Psi(x)\,\mathrm{d}x\\&-\left|\int_{-\infty}^{-\mu/\sigma}\sigma x\,\Psi(x)\,\mathrm{d}x\right|+\mu\int_{-\infty}^{-\mu/\sigma}\Psi(x)\,\mathrm{d}x\end{aligned}$$

In part (a) we establish that under the condition of equality in the coefficient of varition of biomass estimates at t_1, t_2 and t':

$$\frac{a}{\sigma_1}<\frac{1}{\sigma} \quad \text{and} \quad \frac{(1-a)}{\sigma_2}<\frac{1}{\sigma}$$

which implies that $\sigma_1+\sigma_2>\sigma$ and $-a/\sigma_1>-1/\sigma$ and $-(1-a)/\sigma_2>-1/\sigma$. Therefore:

$$\begin{aligned}&\mathrm{OE}_{(3)}-\mathrm{OE}_{(2)}\\&>\left|\int_{-\mu/\sigma}^{-a\mu/\sigma_1}\sigma_1 x\,\Psi(x)\,\mathrm{d}x\right|+\left|\int_{-\mu/\sigma}^{-(1-a)\mu/\sigma_2}\sigma_2 x\,\Psi(x)\,\mathrm{d}x\right|\\&\quad-a\mu\int_{-\mu/\sigma}^{-a\mu/\sigma_1}\Psi(x)\,\mathrm{d}x-(1-a)\mu\int_{-\mu/\sigma}^{-(1-a)\mu/\sigma_2}\Psi(x)\,\mathrm{d}x\\&>\int_{-\mu/\sigma}^{-a\mu/\sigma_1}a\frac{\mu}{\sigma_1}\sigma_1\,\Psi(x)\,\mathrm{d}x+\int_{-\mu/\sigma}^{-(1-a)\mu/\sigma_2}[(1-a)/\sigma_2]\mu\sigma_2\,\Psi(x)\,\mathrm{d}x\\&\quad-a\mu\int_{-\mu/\sigma}^{-a\mu/\sigma_1}\Psi(x)\,\mathrm{d}x-(1-a)\mu\int_{-\mu/\sigma}^{-(1-a)\mu/\sigma_2}\Psi(x)\,\mathrm{d}x=0\end{aligned}$$

So $\mathrm{OE}_{(3)}-\mathrm{OE}_{(2)}>0$, which implies that $\mathrm{E}[\mathrm{NPP}_{(3)}]>\mathrm{E}[\mathrm{NPP}_{(2)}]$.

APPENDIX C

Algorithm for solving the system of equations (3) and (4)

Define:

$$A_{ij}=\frac{\sigma_j}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}(\mu_i/\sigma_j)^2}$$

where i is the subscript for μ and j the subscript for σ.

Step 1. Make μ_j the uncorrected estimate of NPP [$\mu_i = \mathrm{E(NPP)}$], and σ_i is the uncorrected standard deviation of NPP:

$$\left[\mathrm{SD(NPP)} = \left(\sqrt{\mathrm{Var}\ B_1 + \mathrm{Var}\ B_2 - \mathrm{Cov}(B_1,\ B_2)}\right)\right]$$

Usually, covariance $(B_1,\ B_2) = 0$ because of independence. Calculate A_{ii}.

Step 2. Calculate σ_{i+1} as follows:

$$\sigma_{i+1} = \left\{\frac{1}{p}\left[\mathrm{Var(NPP)} - p\mu_i^2(1-p) - \mu_i A_{ii}(1-2p) + A_{ii}^2\right]\right\}^{1/2}$$

Step 3. Calculate $A_{i,i+1}$.

Step 4. Calculate μ_{i+1} as follows:

$$\mu_{i+1} = \frac{1}{p}\left[E(\mathrm{NPP}) - A_{i,i+1}\right]$$

Step 5. Calculate $A_{i+1,i+1}$.

Step 6. If $\mu_{i+1} - \mu_i > \mathrm{CL}$, then go to Step 2.
Else stop.

SD(NPP) and E(NPP) are the uncorrected estimated values of the mean and standard deviation of NPP, and they are fixed in the algorithm.

μ_{i+1} and σ_{i+1} are the corrected values for mean and standard deviation of NPP. CL is the convergence limit criterion. We chose $\mathrm{CL} = 0.01$.

REFERENCES

Aber, J.D., Melillo, J.M., Nadelhoffer, K.J., McClaugherty, C.A. and Pastor, J., 1985. Fine root turnover in forest ecosystems in relation to quantity and form of nitrogen availability: comparison of two methods. Oecologia Berlin, 66: 317–321.

Chow, Y.S. and Teicher, H., 1978. Probability Theory. Springer, New York, 458 pp.

Dickerman, J.A., Stewart, A.J. and Wetzel, R.G., 1986. Estimates of net annual aboveground production: sensitivity to sampling frequency. Ecology, 67: 650–659.

Korn, G.A. and Korn, T.M., 1968. Mathematical handbook for scientists and engineers. McGraw-Hill, New York, 1130 pp.

Lauenroth, W.K., Hunt, H.W., Swift, D.M. and Singh, J.S., 1986. Response to Vogt et al. Ecology, 67: 580–582.

Odum, E.P., 1960. Organic production and turnover in old field succession. Ecology, 41: 34–39.

Odum, E.P., 1971. Fundamentals of Ecology. Saunders, Philadelphia, PA, 574 pp.

O'Neill, R.V., 1973. Error analysis of ecological models. In: O.J. Nelson (Editor), Radionuclides in Ecosystems. USAEC-CONF-710501. Technical Information Division, Washington, DC, pp. 898–908.

Sala, O.E., Deregibus, V.A., Schlichter, T.M. and Alippe, H.A., 1981. Productivity dynamics of a native temperate grassland in Argentina. J. Range Manage., 34: 48–51.

Singh, J.S., Lauenroth, W.K. and Steinhorst, R.K., 1975. Review and assessment of various techniques for estimating net aerial primary production in grasslands from harvest data. Bot. Rev., 41: 181–232.

Singh, J.S., Lauenroth, W.K., Hunt, H.W. and Swift, D.M., 1984. Bias and random errors in estimation of net root production: a simulation approach. Ecology, 65: 1760–1764.

Vogt, K.A., Grier, C.C., Gowen, S.T., Sprugel, D.G. and Vogt, D.J., 1986. Overestimation of net root production: a real or imaginary problem? Ecology, 67: 577–579.

Wiegert, R. and Evans, F.C., 1964. Primary production and the disappearance of dead vegetation on an old field. Ecology, 45: 49–63.

Wilm, H.G., Costello, D.F. and Klipple, G.E., 1944. Estimating forage yield by the double sampling method. J. Am. Soc. Agron., 36: 194–203.