Ecological Applications, 2(1), 1992, pp. 55–70

AGGREGATING FINE-SCALE ECOLOGICAL KNOWLEDGE TO MODEL COARSER-SCALE ATTRIBUTES OF ECOSYSTEMS[1]

EDWARD B. RASTETTER
Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543 USA

ANTHONY W. KING
Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831 USA

BERNARD J. COSBY[2]
School of Forestry and Environmental Studies, Duke University, Durham, North Carolina 27706 USA

GEORGE M. HORNBERGER
Department of Environmental Sciences, Clark Hall, University of Virginia, Charlottesville, Virginia 22903 USA

ROBERT V. O'NEILL
Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831 USA

JOHN E. HOBBIE
Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543 USA

Abstract. As regional and global scales become more important to ecologists, methods must be developed for the application of existing fine-scale knowledge to predict coarser-scale ecosystem properties. This generally involves some form of model in which fine-scale components are aggregated. This aggregation is necessary to avoid the cumulative error associated with the estimation of a large number of parameters. However, aggregation can itself produce errors that arise because of the variation among the aggregated components. The statistical expectation operator can be used as a rigorous method for translating fine-scale relationships to coarser scales without aggregation errors. Unfortunately this method is too cumbersome to be applied in most cases, and alternative methods must be used. These alternative methods are typically partial corrections for the variation in only a few of the fine-scale attributes. Therefore, for these methods to be effective, the attributes that are the most severe sources of error must be identified a priori. We present a procedure for making these identifications based on a Monte Carlo sampling of the fine-scale attributes. We then present four methods of translating fine-scale knowledge so it can be applied at coarser scales: (1) partial transformations using the expectation operator, (2) moment expansions, (3) partitioning, and (4) calibration. These methods should make it possible to apply the vast store of fine-scale ecological knowledge to model coarser-scale attributes of ecosystems.

Key words: aggregation; aggregation error; calibration; coarse-scale modeling; ecological modeling; error propagation; lumped models; model aggregation; scale corrections; scaling; scaling error; transmutation.

INTRODUCTION

The focus of ecological research is expanding from relatively small plots of land to large landscapes and regions, and has now grown to include the ecology of the globe as a whole. However, these regional and global systems are much larger than those from which basic knowledge of ecological processes is generally derived. The ability to test that knowledge through experimental manipulations is also limited to much smaller systems. If this knowledge gained and tested on small systems is to be applied to larger ones, rules must be found by which the fine-scale information (e.g., ecophysiology) can be scaled and applied to coarser-scale phenomena (e.g., global carbon balance).

This scaling can be accomplished by modeling the interactions among fine-scale components to predict coarser-scale properties of the aggregate. For example, interactions among individual plants (fine scale) might be modeled to predict the behavior of forests (coarse scale). However, there is a limit to the amount of re-

[1] Manuscript received 2 February 1990; revised 21 February 1991; accepted 1 May 1991.
[2] Present address: Department of Environmental Sciences, Clark Hall, University of Virginia, Charlottesville, Virginia 22903 USA.

duction in scale that can be incorporated in a model. Incorporating interactions among many components is generally not possible simply because of the large number of interactions involved (Beven 1989). For example, it would be impossible to simulate photosynthesis for every leaf within a forest to make predictions of forest productivity. Yet the ability to apply the knowledge gained by studying fine-scale components (like individual leaves) to predict coarse-scale phenomena (like forest productivity) is extremely desirable.

Because of the impracticality of handling large numbers of fine-scale components individually, they are generally lumped into an aggregated component and treated collectively. For example, the leaves in a canopy might be lumped and treated as a single "big leaf" (e.g., Sinclair et al. 1976, Rastetter et al. 1991) or whole plants lumped to represent regional vegetation cover (e.g., Sellers et al. 1989, Raich et al. 1991). The fine-scale equations are often applied directly, or with minor changes, to describe the properties of this coarse-scale aggregate. The problem with this approach is that the aggregate does not generally behave the same way as the fine-scale components from which it is constituted (O'Neill 1979); the "big leaf" may be inaccurately modeled with the same equations that describe the behavior of the individual leaves within the canopy. However, can the information contained in the fine-scale equations be translated so it can be applied to the aggregate?

Several authors have investigated errors associated with this translation. In general, the fine-scale components have high-frequency, transient behaviors that are damped in the aggregate (Caughley 1982, Cale and Odell 1980). If the fine-scale equations are used to model aggregate behavior, the model may be inconsistent with direct measurements of the aggregate on short time scales (Cale and O'Neill 1988). Similarly, variation among fine-scale components is subsumed in the aggregate. This subsuming of variance can produce error if the fine-scale equations are applied directly to the aggregate (O'Neill 1979). Thus, a coarse-scale model, assembled from fine-scale relationships, can be inaccurate even when the underlying, fine-scale processes are well understood and can themselves be adequately modeled. However, this loss of accuracy introduced by aggregation must be balanced against a loss in precision through the accumulation of errors associated with the estimation of a large number of parameters in complex, non-aggregated models (O'Neill 1973).

Not all aggregations produce errors. No error will result from the aggregation of components with only linear properties. Components with nonlinear properties can also be aggregated without error if the nonlinearity is only in the interactions among the aggregated components themselves and if the relationships of all aggregated components to the environment outside the aggregate are linear and identical (Cale et al. 1983). Similarly, no aggregation error will result if the attributes of the components being aggregated maintain a constant proportionality (Cale et al. 1983). Aggregation errors can be minimized if the property being scaled is highly correlated among the aggregated components (Bartell et al. 1988), if the time constants of the aggregated components are similar (O'Neill and Rust 1979, Gardner et al. 1982), or if a very large component is aggregated with a very small one (O'Neill and Rust 1979). Aggregation error tends to be less problematic if the aggregated components perform parallel functions (e.g., leaves in a canopy) rather than interact in series (e.g., populations along a trophic chain, Gardner et al. 1982).

Our first goal in this paper is to convey an understanding of what causes the aggregation problem. We relate the aggregation problem to the "fallacy of the averages" (Wagner 1969: 658) and demonstrate that rigorous analytical solutions to the problem are not always desirable because they can lead to excessively complex equations that are difficult to use and are heuristically unenlightening. Our analysis is confined to aggregations that lump fine-scale components performing similar, parallel functions (e.g., aggregating leaves within a canopy or individuals within a population). Several individual sources of error, each arising from different properties of these fine-scale components, can simultaneously contribute to the aggregation problem. Correction for all of these sources is usually impossible, but corrections for a few of the most serious sources can substantially reduce error. Our second goal is to present a procedure for identifying these most severe sources of aggregation error. Our third goal is to present several alternative procedures for reducing aggregation error. No one procedure will be best in all cases; a choice among them will usually depend upon the practicality of their application rather than their effectiveness at reducing aggregation error. We conclude with an evaluation of the effectiveness of these procedures for a particular example.

The Nature of the Aggregation Problem

Aggregations are made to reduce model complexity. Reduced complexity may be desirable for purely heuristic reasons; simple models are usually easier to understand than more complex ones. More importantly, reduced complexity may be necessary because of an accumulation of the error associated with the estimation of parameters for each of a large number of the model components (O'Neill 1973). The error associated with each of these parameters will propagate through the model and combine, often multiplicatively, with the error associated with other parameters. This error will be especially large if the parameters are estimated independently because such an estimation does not account for the covariances among parameters. Estimation of all of the parameters simultaneously through a multivariate calibration will account for this

covariance and can reduce the propagated error substantially (Hornberger and Cosby 1985). At a minimum, the error in the model as a whole will equal the cumulative propagated error from all of the parameters. Therefore very complex models with many parameters tend to lack precision. One way to enhance precision is to reduce the number of model components, thereby reducing the number of parameters that need to be estimated. Reducing the number of model components involves some form of aggregation.

Errors associated with this aggregation are often overlooked. These errors arise when the aggregated components are inappropriately described with equations and parameters derived for fine-scale components. This inappropriate application of fine-scale relationships to aggregated components can be equated to the "fallacy of the averages" (Wagner 1969), a fallacy that has widespread occurrence in the environmental sciences (Templeton and Lawlor 1981, Beven 1989).

The problem is easily visualized for the case where only two fine-scale components are being aggregated and where the fine-scale function relating some attribute of these components ($f(X)$) to a variable X is concave in only one direction (i.e., either concave upward or concave downward throughout the range of X). One coarse-scale relationship for this system would be a function relating the mean attribute of the combined components (i.e., $\bar{f} = [f(x_1) + f(x_2)]/2$) to the mean of the two variable values (i.e., $\bar{x} = [x_1 + x_2]/2$).

As a first-order approximation, the mean of the two variables could be used in the fine-scale function to estimate the coarse-scale attribute (i.e., $\hat{f} = f(\bar{x})$) used as an estimate of $\bar{f}$). If the resulting estimate is plotted on the same graph as the fine-scale equation, the point $(\bar{x}, \hat{f})$ obviously lies along the arc of the fine-scale function itself (Fig. 1). However, if the two function values are themselves averaged and plotted on the graph, the point $(\bar{x}, \bar{f})$ lies along the straight line connecting the original two points on the function (i.e., connecting the points $[x_1, f(x_1)]$ and $[x_2, f(x_2)]$). The point representing the coarse-scale relationship $(\bar{x}, \bar{f})$ always lies on the concave side of the fine-scale function. The discrepancy between $\hat{f} = f(\bar{x})$ and $\bar{f}$ is the aggregation error for the estimate (Fig. 1).

Two important results are apparent from this graphical analysis. First, aggregation error will increase as the depth of the concavity in the fine-scale function increases. Thus, aggregation error will tend to be severe for deeply curved functions but negligible for nearly flat ones. Second, aggregation error will increase as the spread between the two operand values (x_1 and x_2) widens, that is, as the variance of the operands increases. As long as the function is concave in only one direction throughout the range of X, both of these results will generalize to the case where there are >2 fine-scale components being aggregated, although this case is more difficult to visualize graphically.

Because the coarse-scale values always fall within the concavity of the fine-scale function, the relationship

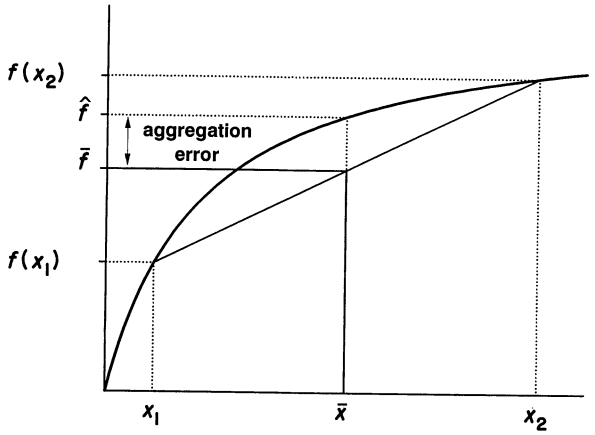

FIG. 1. Error resulting from the aggregation of two fine-scale components. The heavy line represents the relationship between some variable (X) and some attribute ($f(X)$) of two independent, fine-scale components. The two fine-scale components differ from one another in their respective values of X. Individually, the attributes for the two components are $f(x_1)$ and $f(x_2)$, respectively. Together, their mean attribute is $\bar{f}$. However, placing the mean value of X in the fine-scale relationship yields a biased estimate $\hat{f}$. The error tends to increase as the distance between x_1 and x_2 increases (i.e., the variance) and as the concavity of the fine-scale relationship deepens.

for coarse-scale attributes will tend to be flatter than the fine-scale relationship. In most cases, this tendency toward flattening will also be true for fine-scale functions that are concave upward for parts of their range and downward for others (e.g., sigmoid curves), although the effect will be small near the inflection points where concavity changes. To incorporate this flattening in the coarse-scale functions, the form of the fine-scale function must be transformed during the aggregation process (called "transmutation" by O'Neill [1979]). This transformation is usually more than just a simple change in the parameter values; the form of the equation itself must change.

One way to transform fine-scale functions into coarse-scale functions is to quantify the variation among fine-scale components using a probability density function (see also Cale 1988) and use the statistical expectation operator to derive the expected behavior of the aggregate ($E[f(X)]$, Yeh 1973: 145).

$$F = E[f(X)] = \int_{-\infty}^{+\infty} f(x)\rho(x)\,dx, \qquad (1)$$

where F is the aggregated, coarse-scale relationship (e.g., a full-canopy photosynthesis equation), $f(X)$ is a fine-scale relationship (e.g., individual-leaf photosynthesis equation), X is some variable in the fine-scale relationship that varies among the individual components (e.g., irradiance), and $\rho(X)$ is the probability density function describing how X is distributed among the fine-scale components.

Because Eq. 1 is a definite integral, the variable X does not appear in the aggregated, coarse-scale function

(F). Instead, it is replaced by the parameters in $\rho(X)$, which describe the distribution of X. There are usually several parameters in $\rho(X)$; consequently, the coarse-scale function will tend to have more parameters than the original, fine-scale function. The propagated error associated with estimating these additional parameters can be substantially smaller than the propagated error associated with independent estimates of X for each of the fine-scale components, especially if there are a large number of X values to estimate and the variability among these X values is large.

Normally there is >1 property that varies among the fine-scale components (e.g., irradiance, chlorophyll content, vapor pressure). The aggregation error in this case is equal to $\boldsymbol{E}[f(X_1, X_2, X_3, \ldots)] - f(\boldsymbol{E}[X_1], \boldsymbol{E}[X_2], \boldsymbol{E}[X_3], \ldots)$, where the X_i are the various properties that vary among fine-scale components (not individual values of a single property). The expectation procedure, applied sequentially for all sources of variation, would constitute a "perfect aggregation" (Iwasa et al. 1987, 1989) or a "strict aggregation" (Hirata and Ulanowicz 1986) because it introduces no error. To apply the procedure, all forms of variability among the components being aggregated must be fully characterized with a joint, multivariate probability density function (i.e., $\rho[X_1, X_2, X_3, \ldots]$) and incorporated into the aggregated equations.

A full characterization of all forms of variability is usually not possible, and evaluation of the integral can be difficult, making a full implementation of the procedure impractical in many instances. As we point out in the next section, there are also analytical reasons why a full implementation of this procedure is impractical. Nevertheless, partial implementations are often useful, and the procedure does provide a convenient theoretical basis with which to illustrate the consequences of aggregation and from which to compare other aggregation procedures.

An Example of Aggregation

To illustrate the consequences of aggregation, we develop corrected photosynthesis equations for an idealized canopy in which the leaves are oriented horizontally and distributed homogeneously throughout the canopy. We also assume that light hits this canopy vertically and that the leaves are small enough that the extinction of light down through the canopy is closely approximated by a continuous decay function. Individual leaves are assumed to vary only in the amount of light they intercept and in the efficiency with which they use that light. These photosynthesis equations will be used in later sections to illustrate various strategies for developing coarse-scale models and to illustrate the errors associated with these strategies. The equations are therefore developed in some detail.

Assume that the rate of photosynthesis (CO_2 uptake) per unit area for individual leaves can be modeled as a hyperbolic function of irradiance (Cosby 1984, Landsberg 1986: 80):

$$P = \frac{P_m E_0 I}{P_m + E_0 I}, \tag{2}$$

where P is the photosynthetic rate per unit area of leaf, P_m is the maximum photosynthetic rate per unit area of leaf, E_0 is the light-use efficiency, and I is the irradiance (photosynthetic photon flux density) on the leaf. To calculate the average photosynthetic rate per unit of leaf area for the whole canopy ($\bar{P}$), this equation could be applied to all the leaves in the canopy and the results averaged. However, to make this calculation, the values of I and E_0 would first need to be estimated for each leaf. Not only would these estimations be tedious, the cumulative error associated with them could be large.

Alternatively, the leaves could be sampled to estimate mean values for I and E_0 and these values could be used in the fine-scale equation (Eq. 2) as a first-order approximation of $\bar{P}$:

$$\bar{P}_u = \frac{P_m \bar{E}_0 \bar{I}}{P_m + \bar{E}_0 \bar{I}}, \tag{3}$$

where the u subscript indicates that the estimate is uncorrected, that is, that the form of the equation is identical to that of the fine-scale equation (Eq. 2).

To correct for the variability in irradiance, we need a probability distribution for I on individual leaves. This distribution would be difficult to estimate, so we make the simplifying assumption that the canopy can be partitioned into thin layers where the variability in irradiance is small. The distribution of irradiance on these thin layers will depend upon the distribution of leaf area in the canopy. We assume a canopy in which the cumulative overlying leaf area per unit of ground surface area increases linearly and continuously from 0, at the top of the canopy, to L_0, at the bottom of the canopy. We further assume that the errors associated with this continuous approximation to a discrete distribution are negligible. The cumulative overlying leaf area per unit of ground surface area at a randomly selected level in the canopy will therefore take on any value between 0 and L_0 with equal probability. The probability density function for cumulative, overlying leaf area per unit of ground surface area is therefore:

$$\rho(L) = \frac{1}{L_0} \qquad \text{for } 0 < L < L_0$$

$$= 0 \qquad \text{elsewhere,} \tag{4}$$

where L is the cumulative, overlying leaf area per unit of ground surface area above a randomly selected level in the canopy.

Next we assume that the irradiance within the canopy is distributed according to Beer's Law (Waring and Schlesinger 1985: 12):

$$I = I_T e^{-kL} \text{ for } 0 < L < L_0, \tag{5}$$

where I is the irradiance at the level in the canopy where the cumulative, overlying leaf area per unit of ground surface area equals L, I_T is the irradiance at the top of the canopy, and k is the Beer's Law extinction coefficient.

Because of the functional relationship between I and L (Eq. 5), the probability density function for I can be calculated from the probability density function for L as follows (Yeh 1973: 103):

$$\rho(I) = \frac{\rho(L)}{\left|\frac{dI}{dL}\right|}$$

$$= \frac{1}{kL_0 I} \qquad \text{for } I_T e^{-kL_0} < I < I_T$$

$$= 0 \qquad \text{elsewhere.} \tag{6}$$

Finally, substituting Eqs. 2 and 6 into the expectation operator (Eq. 1), the mean canopy photosynthetic rate per unit of leaf area, corrected for variation in irradiance ($\bar{P}_I$), can be calculated as follows:

$$\bar{P}_I = \frac{P_m}{kL_0} \ln\left\{\frac{P_m + E_0 I_T}{P_m + E_0 I_T e^{-kL_0}}\right\}. \tag{7}$$

This equation is similar to one derived by France and Thornley (1984) for crop canopies.

The variability among leaves can also come from sources other than incident light (such as nitrogen content or different species). This variability can appear in several places in the equations. It might involve differences in the independent variables, such as irradiance in the example above. Additional variability might also arise due to differences in what would normally be viewed as constant model parameters, such as the light-use efficiency (E_0). In this context, E_0 is a spatially distributed variable. Here we assume that the light-use efficiencies among leaves are distributed independently of the irradiance and uniformly between the values of E_{01} and E_{02} (i.e., $\rho(E_0) = 1/\Delta E_0$; where $\Delta E_0 = E_{02} - E_{01}$). Following the procedure we used for irradiance, the mean photosynthetic (CO_2 uptake) rate per unit of leaf area, corrected for $E_0(P_E)$ is:

$$\bar{P}_E = P_m - \frac{P_m^2}{I\,\Delta E_0} \ln\left\{\frac{P_m + E_{02} I}{P_m + E_{01} I}\right\}, \tag{8}$$

where I is the irradiance at the level in the canopy where $\bar{P}_E$ is calculated.

Finally, we simultaneously correct for both irradiance and light-use efficiency. Because we have assumed that I and E_0 are distributed independently of one another, this correction can be made using the expectation operator (Eq. 1) with either Eq. 7 and $\rho(E_0)$ or Eq. 8 and $\rho(I)$. The resulting mean photosynthetic rate per unit of leaf area is:

$$\bar{P}_{EI} = \frac{P_m}{kL_0\,\Delta E_0}\left\{\frac{P_m}{I_T}\ln\left(\frac{P_m + E_{02} I_T}{P_m + E_{01} I_T}\right) - \frac{P_m}{I_T e^{-kL_0}}\ln\left(\frac{P_m + E_{02} I_T e^{-kL_0}}{P_m + E_{01} I_T e^{-kL_0}}\right) + E_{02}\ln\left(\frac{P_m + E_{02} I_T}{P_m + E_{02} I_T e^{-kL_0}}\right) - E_{01}\ln\left(\frac{P_m + E_{01} I_T}{P_m + E_{01} I_T e^{-kL_0}}\right)\right\}. \tag{9}$$

For the sake of illustration, we assume that the simplifying assumptions we had to make to arrive at this equation contribute negligibly to the error. We therefore assume that this equation is fully corrected.

The particular forms of the three corrections (Eqs. 7, 8, and 9) are less important to our purposes than their qualitative relation to the fine-scale equation from which they were derived (Eq. 2). First, to make the transformations, a complete characterization of the variability among the fine-scale components was needed (i.e., $\rho(I)$ and $\rho(E_0)$). As mentioned above, characterization of the variability among fine-scale components can be difficult and may require prohibitively intense sampling. To make the derivations illustrated above, we had to make simplifying assumptions to overcome this problem. Second, with each successive correction (Eq. 3 to Eqs. 7 and 8 to Eq. 9), the relationship became more complex, gained parameters, and was heuristically more difficult to grasp. Third, the uncorrected (Eq. 3) and partially corrected (Eqs. 7 and 8) relationships are biased estimators of the fully corrected relationship (Eq. 9). In particular, the partially corrected relationships lie on the concave side of the uncorrected relationship, and the fully corrected relationship lies on the concave side of both the uncorrected and partially corrected relationships (Fig. 2).

To remove all aggregation error, transformations would have to be made to account for all forms of variability. For example, corrections might also be made for the temporal variation in I_T to calculate a mean daily photosynthetic rate (e.g., Gross 1982), or corrections might be made for lateral variation in leaf area (spatial variations in L_0). However, the increased complexity with each successive transformation will make this more and more difficult, as should be obvious from an examination of the successive increase in complexity from Eq. 3 to Eqs. 7 and 8 to Eq. 9.

The problem becomes even worse when all the processes occurring in an ecosystem are considered, each of which must be corrected for all forms of variability. It would be impossible to get an adequate characterization of all these forms of variability (i.e., ρ for each). Even if it were possible, the derivation of the corrected relationships might prove tedious and unenlightening.

Nevertheless, correcting for *some* of the sources of variability is usually worthwhile. For example, Eqs. 7

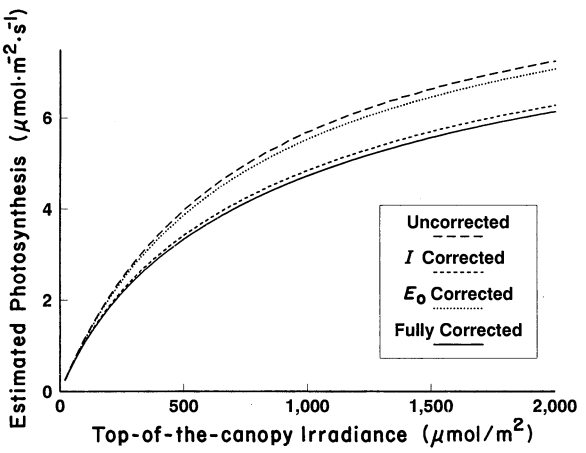

FIG. 2. Estimates of mean canopy photosynthesis (CO_2 uptake) at varying top-of-the-canopy irradiance (photosynthetic photon flux density) based on three transformations of the fine-scale relationship. The uncorrected curve is for the uncorrected function (Eq. 3). The remaining curves are corrections for irradiance (Eq.7:I), light-use efficiency (Eq. 8:E_0), and for both (i.e., fully corrected, Eq. 9). The curves are for a canopy with a leaf area index (L_0) = 6.

and 8 are less-biased estimators of canopy photosynthesis than is Eq. 3. Corrections for other sources of variability might be inconsequential. Among the many possible corrections that could be made, which will be most profitable? In the following section, we examine ways to rank the relative contributions of various sources of aggregation error to help guide such decisions.

Ranking Aggregation Error from Various Sources

Ranking the relative contributions of the fine-scale variables to aggregation error requires that these contributions be quantified. One way to make this quantification is to use the corrections based directly on the expectation operator. Each of the partial corrections for individual variables (e.g., Eqs. 7 and 8) can be compared to the full (or at least most complete) correction for all variables (e.g., Eq. 9). Assuming that the improved accuracy resulting from the corrections is not offset by a decrease in precision resulting from additional parameters, the partial correction that most closely matches the full correction will be the most profitable one to employ. However, this approach can only be implemented if the expectation operator can be integrated for the full array of variables. These integrations are often difficult. In any case, it would be convenient to know which corrections will be most beneficial before having to go through the effort of actually making the corrections.

Simpler approaches to this quantification are hampered by the nonlinearity of the fine-scale equation. Simple analytical approaches can be devised for particular classes of nonlinear equations, but we know of none that are general for all nonlinear equations. Therefore, a ranking of the relative contributions of the variables to the aggregation error must rely on an intuitive understanding of the aggregation problem and on less rigorous, non-analytical approaches to the quantification.

A simple qualitative assessment of the contributions to aggregation error can be made using the graphical analysis employed earlier (Fig. 1). By plotting the fine-scale function against each variable, the potential for error can at least be visualized. Variables for which these plots are deeply concave will tend to be associated with large portions of the aggregation error. However, caution must be taken with this approach because it is not only the range of variability in the variables that is important; the distributions within these ranges are also important. It is difficult to assess the ramifications of these distributions from a simple graph. Nevertheless, this graphical analysis is a good first step for assessing aggregation error and will thus help guide further analysis.

One important piece of information that can be determined by plotting the fine-scale functions against each of the variables is the *direction* of aggregation error. This can be an important factor to consider before making corrections. For example, consider the function

$$f(X, Y) = X^2 - Y^2 + 1, \tag{10}$$

where both X and Y vary uniformly between -0.5 and $+0.5$. From plots of this function against X and Y (Fig. 3), it is easy to see that errors associated with the two variables are in opposite directions. The function is concave upward with respect to X, resulting in a negative error. It is concave downward with respect to Y, resulting in a positive error. Because the errors are in opposite directions, they will tend to cancel each other. Thus, a correction for only one of the variables can actually make the overall error worse.

This result is easily confirmed using the expectation operator to make the actual corrections. The corrections for X and Y, respectively, are:

$$\bar{F}_x(Y) = \frac{13}{12} - Y^2 \tag{11}$$

and

$$\bar{F}_y(X) = \frac{11}{12} + X^2. \tag{12}$$

Using the respective means ($\bar{x} = 0$ and $\bar{y} = 0$) to generate estimates of the aggregated value of the function, the uncorrected equation (Eq. 10) produces an estimate of 1, the correction for X (Eq. 11) produces an estimate of 13/12, and the correction for Y (Eq. 12) produces an estimate of 11/12. The true, fully corrected value is $\boldsymbol{E}[f(X, Y)] = 1$. Because the errors associated with the two variables exactly canceled each other, the uncorrected estimate is better than either of the two cor-

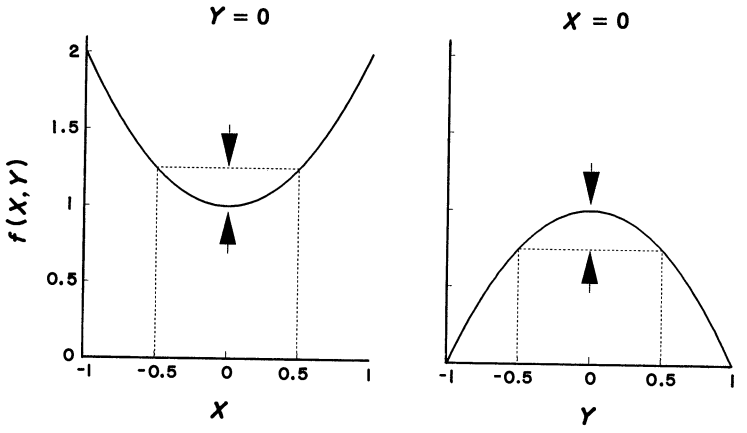

FIG. 3. An assessment of the aggregation error for variables X (left) and Y (right) for the fine-scale equation $f(X, Y) = X^2 - Y^2 + 1$. The error associated with X is negative while that for Y is positive. Assuming that X and Y are uniformly distributed between -0.5 and $+0.5$, the aggregation error for each must lie somewhere within the arrows in the respective graphs.

rected ones. Therefore, care must be taken to ensure that corrections for one variable do not enhance the manifestation of errors associated with other variables.

More quantitative assessments of aggregation error can be made using Monte Carlo simulations. The aim of these simulations is to sample the range of fine-scale responses to the variability in the fine-scale variables. To do this, sets of fine-scale variables would be repeatedly generated at random from their respective distributions and used in the fine-scale function. The mean of this sample of fine-scale responses is an estimate of the value of the aggregated, coarse-scale response. Sequential estimates of this mean will converge to a fixed value once the sample size is large enough (see example below, Fig. 4). An estimate of the fully corrected relationship can be made by sampling with random values for all variables simultaneously. An estimate of the partial correction for any one variable can be made by sampling with random values for only that variable and leaving the rest fixed at their respective mean values. Similarly, higher order, partial corrections for ≥ 2 variables can be estimated by allowing only them to vary in the sampling. An assessment of the best partial correction can then be made by comparing the estimates of each of the partial corrections to the estimate of the full correction. Again, assuming that the propagated error due to additional parameters is not prohibitive, the partial correction that most closely matches the full correction will be the most profitable one to employ.

To illustrate this, consider the idealized canopy discussed above (see *An example of aggregation*) with $I_T = 1000\ \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, maximum photosynthetic rate $P_m = 10\ \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, $k = 0.5\ \text{m}^2/\text{m}^2$, light-use efficiency in CO_2 uptake E_0 distributed uniformly between $E_{01} = 0.018$ and $E_{02} = 0.065\ \mu\text{mol}/\mu\text{mol}$, and I distributed according to Eq. 6. When the sample means are plotted for this example (Fig. 4), they converge only after a sample size of ≈ 200. With a leaf area index (L_0)

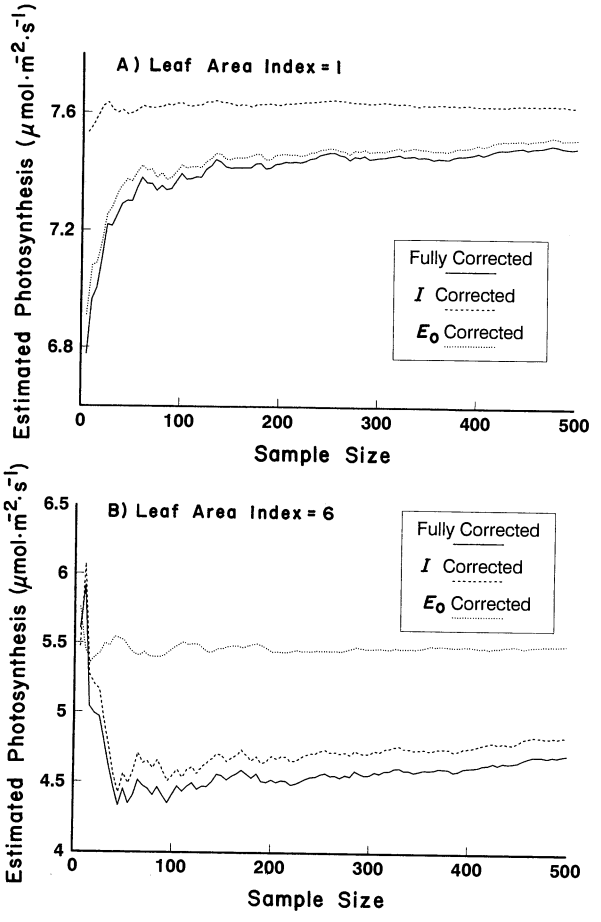

FIG. 4. Estimated mean canopy photosynthesis (CO_2 uptake) at 1000 μmol/m^2 top-of-the-canopy irradiance (photosynthetic photon flux density) based on a Monte Carlo sampling of the variables in the fine-scale photosynthesis equation (Eq. 2). Values for within-canopy irradiance (I) and light-use efficiency (E_0) were generated from their respective probability density functions. For the fully corrected curve, both I and E_0 varied in the sampling. For the I-corrected curve, I varied and E_0 was set at its mean value. For the E_0-corrected curve, E_0 varied and I was set at its mean value.

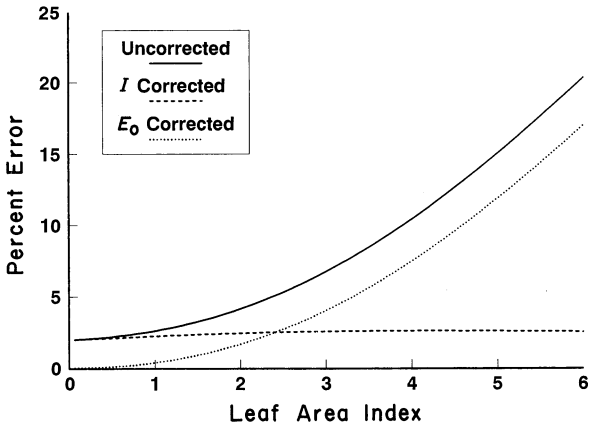

FIG. 5. Error as a percentage of the fully corrected photosynthesis equation (Eq. 9) for the uncorrected (Eq. 3), the irradiance (I)-corrected (Eq. 7), and the light-use efficiency (E_0)-corrected (Eq. 8) equations under varying values of the leaf area index. Evaluations were made at a top-of-the-canopy irradiance of 1000 μmol/m^2.

of 1 (Fig. 4A), the estimate of the correction for E_0 matches the estimate of the full correction more closely than the correction for I; thus corrections for E_0 would be expected to work better in this case. However, when the leaf area index is 6 (Fig. 4B), corrections for I appear to be the more effective.

This shift in the relative importance of the two variables is explained by changes in the variability of light within the canopy as total leaf area changes. Variability among fine-scale components is one of the major factors contributing to aggregation error. The variability in light intensity within a canopy increases as the canopy thickens (Eq. 5). This increase in variability means that aggregation error associated with I in Eqs. 3 and 8 will increase as L_0 increases. On the other hand, variability in E_0 is independent of L_0, so aggregation error associated with E_0 should be less sensitive to changes in L_0. (The small changes in aggregation error associated with E_0 are due to changes in the concavity Eq. 3 as L_0 and therefore $\bar{I}$, changes.)

These results can be confirmed using the actual corrections based on the expectation operator (Eqs. 7, 8, and 9). In canopies with a leaf area index less than ≈ 2, corrections for E_0 are more effective than corrections for I (Fig. 5). In canopies with a leaf area index >3, corrections for I are more effective than corrections for E_0. This confirms that most of the aggregation error is associated with E_0 in thin canopies, and with I in thick canopies. It is also apparent from this analysis that the overall aggregation error increases as the leaf area index increases.

For the example discussed above, the two variables entered into the fine-scale equation (Eq. 2) in exactly the same form. That is, Eq. 2 is a hyperbolic function of both I and E_0. In fact, everywhere either of these variables enters into the equation, it is multiplied by the other. Because of this symmetry, their relative contributions to the aggregation error can be attributed to differences in the shapes of their respective probability distributions; if I and E_0 had identical distributions, their respective contributions to the aggregation error would be identical.

This degree of symmetry will not generally be the case. Variables are likely to enter into the fine-scale equation in different forms. For example, the fine-scale equation might be hyperbolic with respect to one variable and logarithmic with respect to the other. In the general case, the relative contributions of the variables to the aggregation error will depend on both their probability distributions and on the form in which they enter into the fine-scale equation. As an extreme example, if a fine-scale equation of two variables is linear with respect to one of the variables and nonlinear with respect to the other, all of the aggregation error will be attributable to the nonlinear variable even if the two probability distributions are identical. Regardless of the degree of symmetry among variables, the methods presented in this paper are applicable.

USING FINE-SCALE RELATIONSHIPS TO BUILD AGGREGATED, COARSE-SCALE MODELS

From the example of aggregation discussed above, it is clear that there will be errors associated with the application of fine-scale equations to describe coarse-scale phenomena. In other contexts, errors (e.g., reasonable sampling errors) are generally acceptable as long as they are not overly large. Will a similar approach work for aggregation? That is, short of performing the tedious transformations based on the expectation operator, are there ways to reduce aggregation errors that still yield reliable model predictions?

In this section we discuss four such means of reducing aggregation error: (1) *partial transformations* using the expectation operator (Eq. 1) to correct for only the most severe sources of aggregation error; (2) *moment expansions* using a truncated Taylor series expansion of the expectation operator to approximate partial transformations; (3) *partitioning* of the coarse-scale aggregate into a manageably small number of appropriately chosen medium-scale components that are less severely aggregated; and (4) *calibration* of the fine-scale relationship to coarse-scale data. The equations derived using these four techniques will be very different, but their shapes are similar and approximate that of Eq. 9 (Figs. 2, 6, and 7).

To illustrate these techniques, we again consider the idealized canopy discussed above (see *An example of aggregation*) in which only the irradiance (I) and light-use efficiency (E_0) vary among leaves. Again assume that $P_m = 10$ μmol·m^{-2}·s^{-1}, that $k = 0.5$ m^2/m^2, and that E_0 is distributed uniformly between $E_{01} = 0.018$ and $E_{02} = 0.065$ μmol/μmol. Given the leaf area index (L_0) and the top-of-the-canopy irradiance (I_T), the pho-

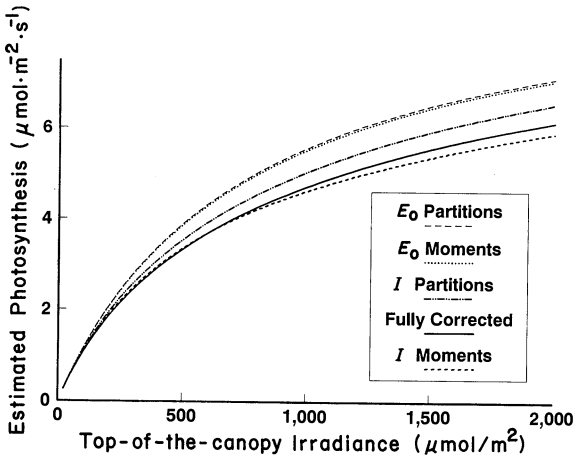

FIG. 6. Estimates of mean canopy photosynthesis (CO_2 uptake) at varying top-of-the-canopy irradiance (photosynthetic photon flux density) based on four approximations to the fully corrected equation (Eq. 9). The approximations are derived from moments expansions (Eq. 14) for irradiance (I) and light-use efficiency (E_0), and from a partitioning (Eq. 18) for I and E_0. The curves are for a canopy with a leaf area index (L_0) = 6.

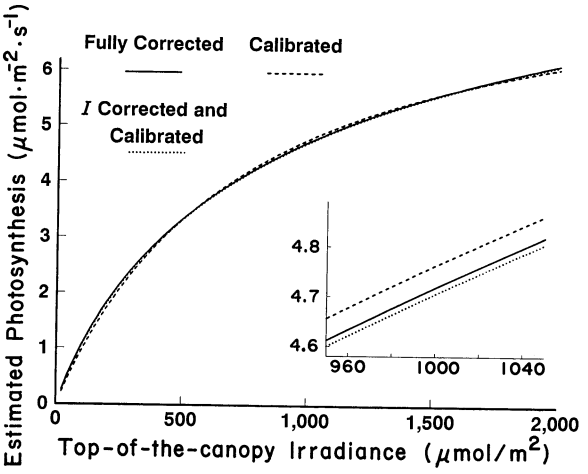

FIG. 7. Estimates of mean canopy photosynthesis (CO_2 uptake) at varying top-of-the-canopy irradiance (photosynthetic photon flux density) based on two calibrations to the fully corrected equation (Eq. 9). The fine-scale equation (Eq. 2) was calibrated directly and was corrected for irradiance (Eq. 7:I) before being calibrated. Inset is an expansion of the relationship near a top-of-the-canopy irradiance of 1000 μmol/m^2. The curves are for a canopy with a leaf area index (L_0) = 6.

tosynthetic rate for this idealized canopy can be exactly predicted from Eq. 9. Data were generated for L_0 values of 1 and 6 m^2/m^2 at 100 evenly spaced values of top-of-the-canopy irradiance (I_T) ranging from 20 to 2000 μmol·m^{-2}·s^{-1}. Uncorrected data were also generated at the same L_0 and I_T values using Eq. 3 (Fig. 2). These data sets were then used to rank the effectiveness of the four correction procedures (Table 1).

This type of ranking is typically made using three criteria. First, a model will rank highly if there is a high correlation between the target data (in this case data from Eq. 9) and model predictions (Draper and Smith 1981). Correlation was of little use in this case because all four techniques yield predictions that are highly correlated with data generated using Eq. 9 ($r > 0.999$).

Second, a model will rank highly if the residuals are serially uncorrelated and if there is no discernible relationship between the residuals and other variables (Box and Jenkins 1976, Draper and Smith 1981, Cosby 1984, Cosby and Hornberger 1984). This criterion was of little use in this case because all four techniques retain at least some of the aggregation error. Therefore, in the absence of measurement error and other noise, the residuals have a high positive serial correlation ($r > 0.99$ for all models) and have a pattern that is clearly related to I_T (Figs. 2, 6, and 7). In any real application, this high serial correlation and the relation to I_T are likely to be hidden by noise in the data if the residuals are small.

The third criterion is to evaluate the magnitude of the residuals (Mallows 1973, Draper and Smith 1981). For the idealized canopy the residuals were roughly proportional to the predicted value of photosynthesis (Figs. 2 and 6). We therefore normalized the residuals relative to photosynthesis by expressing them as a proportion of the rate of photosynthesis where the residual was measured ($[\hat{P} - P]/P$). The mean of these relative residuals squared is a measure of the overall goodness of fit of the corrected equations throughout the range of I_T. By taking the square root of this mean, the overall error can be expressed as a proportion of the predicted photosynthetic rate. If there were variability in P and the models were unbiased (i.e., $\hat{P} = E[P]$), the resulting statistic would be identical to the standard error of the relative residuals. We therefore call the statistic the standard relative error (SRE):

$$\text{SRE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\frac{\hat{P}_i - P_i}{P_i}\right)^2}{n}}, \qquad (13)$$

where $\hat{P}_i$ is the estimated photosynthetic rate and P_i is the actual photosynthetic rate as calculated with Eq. 9 at each of the n (=100) values of I_T.

Uncorrected estimates

As discussed in *Ranking aggregation error. . .*, above, uncorrected (Eq. 3) and fully corrected (Eq. 9) predictions of canopy photosynthesis differed less in the thin canopy with a leaf area index (L_0) of 1 than in the thick canopy with an L_0 of 6 (Fig. 5). The SRE for the thinner canopy was <3%, while that for the thicker canopy was >18%. In the following discussion, we compare residual errors for the four correction procedures to these two values. Thus, the residual errors will be ex-

TABLE 1. Standard relative errors (SRE) and percentage of uncorrected error remaining in estimates of canopy photosynthetic rate, based on various aggregation procedures.

Aggregation procedure	Leaf area index = 1		Leaf area index = 6	
	SRE (%)	% of uncorrected	SRE (%)	% of uncorrected
Uncorrected (Eq. 3)	2.59	100	18.38	100
Transformed				
For irradiance (Eq. 7)	2.20	85	2.35	13
For light-use efficiency (Eq. 8)	0.38	15	15.40	84
Moment expansions (Eq. 14)				
For irradiance through the				
Second moment	2.20	85	1.39	8
Third moment	2.20	85	9.02	49
Fourth Moment	2.20	85	2.51	14
For light-use efficiency through the				
Second moment	0.59	23	15.59	85
Third moment	0.59	23	15.59	85
Fourth moment	0.41	16	15.41	84
Partitioned				
By irradiance (Eq. 18)	2.30	89	6.25	34
By light-use efficiency	1.03	40	16.23	88
Calibrated	0.82	32	3.29	18

pressed both as SREs and as percentages of the SREs for the uncorrected estimates (i.e., 100 × [SRE-corrected]/[SRE-uncorrected]).

Corrections based on partial transformations

As we have already discussed, partial corrections using the expectation operator can substantially improve estimates. Corrections for both I and E_0 improved the estimates of mean canopy photosynthesis regardless of the leaf area index (Table 1). With a leaf area index of 1, corrections for E_0 (Eq. 8) reduced the error to 15% of the uncorrected error (SRE = 0. 38%), while corrections for I (Eq. 7) only reduced it to 85% of the uncorrected error (SRE = 2.20%). On the other hand, with a leaf area index of 6 (Fig. 2), correcting for I decreased error to 13% of the uncorrected error (SRE = 2.35%), while correcting for E_0 only decreased error to 84% of the uncorrected error (SRE = 15.40%).

Moment expansions

Moment expansions are derived from a Taylor series expansion of the fine-scale function in the expectation operator (Eq. 1). They should not be confused with the "method of moments" (Brunk 1975), which is a parameter estimation procedure. The moment expansion described here is a means of transforming *equations* from fine scale to coarse scale. Parameter estimation is a completely different problem, which we discuss briefly below.

If $f(X)$ is expanded about the mean of X (μ_x), Eq. 1 can be rewritten in the following form:

$$F = \sum_{i=0}^{\infty} \frac{f^{(i)}(\mu_x)}{i!} E[(X - \mu_x)^i], \qquad (14)$$

where $f^{(i)}(\mu_x)$ is the i^{th} derivative of $f(X)$ with respect to X evaluated at μ_x and $E[(X - \mu_x)^i]$, called the i^{th} moment of X about μ_x, is the expected value of $(X - \mu_x)^i$ (i.e., setting $f(X) = (X - \mu_x)^i$ in Eq. 1). This same expansion can be repeated sequentially for each variable in the fine-scale equation. These additional expansions will result in a series of interactive terms involving pairs of variables, then triplets and so forth. For the case of only two variables, the first few terms of the series are:

$$F_{xy} = f(\mu_x, \mu_y) + \frac{\sigma_x^2}{2}\frac{\partial^2 f(\mu_x, \mu_y)}{\partial x^2} + \frac{\sigma_y^2}{2}\frac{\partial^2 f(\mu_x, \mu_y)}{\partial y^2} + \sigma_{xy}\frac{\partial^2 f(\mu_x, \mu_y)}{\partial x\, \partial y} + \ldots \qquad (15)$$

where σ_x^2 and σ_y^2 are the variances of X and Y ($\sigma_x^2 = E[\{X - \mu_x\}^2]$), σ_{xy} is the covariance between X and Y, and the $\partial/\partial x$ and $\partial/\partial y$ operators signify partial differentiation with respect to X and Y.

The first term in the series is simply the fine-scale function evaluated at the variable means. That is, it is the uncorrected estimate (e.g., Eq. 3) of the aggregated relationship. The remaining terms are corrections to this estimate. This series will eventually converge to the true value of F. However, the number of correction terms required before the series converges to a reasonable estimate of the aggregated relationship (F) will depend upon the form of the fine-scale relationship ($f(X, Y)$) and upon the joint distribution of X and Y ($\rho(X, Y)$).

TABLE 2. Means and moments about the means for irradiance and light-use efficiency used in moment expansions.

Means

$$\bar{I} = \frac{I_T(1 - e^{-k\,L_0})}{k\,L_0}$$

$$\bar{E}_0 = \frac{E_{01} + E_{02}}{2}$$

Variance

$$\sigma_I^2 = \boldsymbol{E}[(I - \bar{I})^2] = \frac{I_T^2\,(1 - e^{-2k\,L_0})}{2k\,L_0} - \bar{I}^2$$

$$\sigma_{E_0}^2 = \boldsymbol{E}[(E_0 - \bar{E}_0)^2] = \frac{E_{02}^3 - E_{01}^3}{3\Delta\,E_0} - \bar{E}_0^2$$

Skew

$$m_{I3} = \boldsymbol{E}[(I - \bar{I})^3] = \frac{I_T^3\,(1 - e^{-3k\,L_0})}{3k\,L_0} - 3\bar{I}\sigma_I^2 - \bar{I}^3$$

$$m_{E03} = \boldsymbol{E}[(E_0 - \bar{E}_0)^3] = \frac{E_{02}^4 - E_{01}^4}{4\Delta E_0} - 3\bar{E}_0\sigma_{E_0}^2 - \bar{E}_0^3$$

Kurtosis

$$m_{I4} = \boldsymbol{E}[(I - \bar{I})^4] = \frac{I_T^4\,(1 - e^{-4k\,L_0})}{4k\,L_0} - 4\bar{I}m_{I3} - 6\,\bar{I}^2\sigma_I^2 - \bar{I}^4$$

$$m_{E04} = \boldsymbol{E}[(E_0 - \bar{E}_0)^4] = \frac{E_{02}^5 - E_{01}^5}{5\Delta E_0} - 4\bar{E}_0 m_{E03} - 6\bar{E}_0^2\sigma_{E_0}^2 - \bar{E}_0^4$$

Moment expansions are an application of a truncated form of Eqs. 14 or 15 to approximate the rigorous transformations based on Eq. 1. These equations are another way to express the expectation operator. However, they may be more useful expressions than Eq. 1 if the integration in Eq. 1 should prove to be overly cumbersome but the derivatives of the fine-scale relationship can be readily calculated. Eqs. 14 and 15 also have the advantage that the probability density function, $\rho(X, Y)$, need not be specified as long as the moments of the distribution can be estimated by sampling. For the case where the variables are independently distributed, these moments can be estimated by sampling the fine-scale variable (X) and estimating the expected values of the powers of X:

$$\hat{\boldsymbol{E}}[X^k] = \frac{\sum_{i=1}^{n} (x_i^k)}{n}, \qquad (16)$$

where $\hat{\boldsymbol{E}}[X^k]$ is the sample estimate of $\boldsymbol{E}[X^k]$, and x_i are the n sample values of x. The moments about the mean can then be calculated from a binomial expansion:

$$\hat{\boldsymbol{E}}[(X - \mu_x)^k] = \sum_{i=0}^{k} \binom{k}{i} (-\bar{x})^i \hat{\boldsymbol{E}}[X^{k-i}], \qquad (17)$$

where $\binom{k}{i} = k!/[(k - i)!\,i!]$ and $\bar{x} = \hat{\boldsymbol{E}}[X]$. Alternatively, the moments can be calculated from a known or assumed probability density function (i.e., from $\rho(X)$). We calculated the moments directly from $\rho(I)$ and $\rho(E_0)$ (Table 2).

Because of the infinite number of terms in Eq. 15, the moment expansion method can, in general, only serve as an approximation to the partial transformations. Exceptions to this, of course, are polynomial functions, where higher-order derivatives all equal zero and the series therefore terminates after a finite number of terms. Other exceptions would involve probability density functions with higher order moments equal to zero (e.g., a Gaussian distribution). In other cases, higher-order terms in the expansion may be small enough that reasonable approximations to the coarse-scale relationship can be made with only the first few terms in the series.

To illustrate the moment expansion procedure, estimates for the mean canopy photosynthetic rates were made using the second, third, and fourth moments about the mean (i.e., the variance, skewness, and kurtosis) for both irradiance and light-use efficiency (Fig. 6, Table 1). Because we only corrected for one variable at a time, interactive terms were not considered. As with the partial transformations, corrections for either source of variance improved estimates to some degree. Corrections for irradiance only reduced errors to 85% of the uncorrected error (SRE = 2.20%) in the canopy

with a leaf area index of 1, but reduced it to 14% of the uncorrected error (SRE = 2.51%) in the canopy with a leaf area index of 6. Corrections for light-use efficiency, on the other hand, worked better in the thinner canopy than in the thicker one, reducing error to 16 and 84% of the uncorrected error, respectively (SRE = 0.41 and 15.41%).

The major problem with this procedure is determining how many terms in the series must be included before a reliable estimate is achieved. An examination of the successive estimates for the moment expansions for irradiance illustrates the problem. When the leaf area index was 1, the series converged almost immediately; the third and fourth moments contributed very little to the estimate (Table 2). When the leaf area index was 6, on the other hand, there was no evidence that the estimates had converged. Indeed, the estimate including only the second-moment term was better than either of the two subsequent estimates. However, there would be no way of knowing this unless the fully corrected prediction were available.

Partitioning

An obvious way of reducing aggregation error is to reduce the variability among the components to be aggregated. This can be done by partitioning the fine-scale components into relatively homogeneous sub-aggregates. For example, the leaves within the canopy might be partitioned into two groups based upon irradiance. All the leaves with irradiance greater than the mean irradiance could go in one group (i.e., the upper canopy), and those with lower irradiance than the mean could go in the other (i.e., the lower canopy). Because the variability within these sub-aggregates is smaller than it is in the fully aggregated canopy, the aggregation error associated with applying the uncorrected equation (Eq. 3) to them will also be smaller. The full aggregation is then made by taking a weighted mean of the results from the sub-aggregates.

To illustrate the partitioning procedure, Eq. 3 was applied to each of the two layers within the canopy using the respective mean irradiance for each layer and the overall mean light-use efficiency. The two estimates were then weighted by leaf areas of the respective layers and averaged to calculate the mean canopy photosynthetic rate:

$$\bar{P}_{pI} = \frac{L_u P(\bar{I}_u, \bar{E}_0) + L_l P(\bar{I}_l, \bar{E}_0)}{L_0} \qquad (18)$$

for $$\bar{I}_u = \frac{I_T(1 - e^{-kL_u})}{kL_u}$$

and $$\bar{I}_l = \frac{I_T(e^{-kL_u} - e^{-kL_0})}{kL_l}$$

where L_u and L_l are the leaf areas per unit of ground area within the upper and lower canopy layers ($L_u + L_l = L_0$), the function $P(I, E_0)$ is Eq. 3 evaluated at the respective values of irradiance and light-use efficiency, $\bar{I}_u$ is the mean irradiance in the upper canopy layer and $\bar{I}_l$ is the mean irradiance in the lower layer. The most obvious partitioning is to split the canopy into layers with irradiance greater than and less than the canopy mean irradiance, that is, let $L_u = (\ln[I_T] - \ln[\bar{I}])/k$. Estimates based on this partitioning only reduced error to 89% of the uncorrected error (SRE = 2.30%) for the thinner canopy, but reduced it to 34% of the uncorrected error (SRE = 6.25%) for the thicker canopy (Fig. 6, Table 1).

A similar partitioning based on light-use efficiency also improves the estimate of mean photosynthetic rate. This partitioning was made by splitting the canopy into segments with E_0 ranging between 0.018 and 0.0415 and between 0.0415 and 0.065. Errors were reduced by this partitioning to 40% of the uncorrected error (SRE = 1.03%) in the thinner canopy but to only 88% of the uncorrected error (SRE = 16.23%) in the thicker canopy (Fig. 6, Table 1).

These estimates can be improved further by increasing the number of partitions and by combining partitions for irradiance with partitions for light-use efficiency. As the number of partitions increased, the estimates would converge to the value predicted by Eq. 9, the fully corrected equation. In the extreme this disaggregation essentially reverts to the simulation of individual leaves.

Calibration

Calibrations are a direct application of a fine-scale relationship to a coarse-scale phenomenon but with parameter values estimated by fitting the function to coarse-scale data. Any parameter estimation procedure can be used to make these estimations (least-squares fitting, Snedecor and Cochran 1967; method of moments, Brunk 1975; extended Kalman filter, Cosby 1984, Cosby and Hornberger 1984; recursive least squares and instrumental-variable, Young 1984; Levenberg-Marquardt fitting, Press et al. 1986; Monte-Carlo fitting, Hornberger and Cosby 1985). It is not our purpose here to evaluate the relative merits of these procedures. We assume that they will all give similar results where they are applicable.

We used a Levenberg-Marquardt routine (Press et al. 1986: 523) to estimate P_m and E_0 by fitting Eq. 2 to the data generated with Eq. 9 based on 100 evenly spaced values of I_T ranging from 20 to 2000 μmol·m^{-2}·s^{-1}. This calibration can be done using either the mean-within-canopy or the top-of-the-canopy irradiance for I in Eq. 2. However, as will become clear below, there are disadvantages to retaining the two parameters needed to calculate the mean irradiance (i.e., k and L_0). We therefore calibrated using I_T.

Error was reduced to 32% of the uncorrected error (SRE = 0.82%) in the thinner canopy and to 18% (SRE = 3.29%) in the thicker canopy (Fig. 7, Table 1). In

the thinner canopy, estimates of P_m and E_0 were 9.9 μmol·m^{-2}·s^{-1} and 0.031 μmol/μmol, respectively. These estimates are close to their respective mean values for the fine-scale relationship (i.e., 10 μmol·m^{-2}·s^{-1} and 0.0415 μmol/μmol) because of the low overall aggregation error in the thinner canopy. On the other hand, in the thicker canopy, where aggregation error was more severe, the estimates of P_m and E_0 were well removed from the fine-scale values. The estimate of P_m was 8.4 μmol·m^{-2}·s^{-1}, and the estimate of E_0 was 0.011 μmol/μmol. The value for E_0 is not even within the range (0.018 to 0.065 μmol/μmol) of the fine-scale values. Such shifts in parameter values should be expected any time the same equation is applied at different scales. Parameter values consistent with data measured at one scale, therefore, may not be consistent with data measured at another scale (Cushman 1986). For example, if Eq. 2 were calibrated with data obtained by enclosing a single leaf within a cuvette, the parameter estimates would not be the same as those based on a calibration with data obtained by enclosing the entire canopy in a large cuvette (assuming such a thing could be done). By calibrating Eq. 2 to data generated with Eq. 9 we have, in essence, treated the canopy as if it were a "big leaf" enclosed in a cuvette to obtain the data needed for the calibration.

Calibrations correct for all sources of aggregation error simultaneously, including hidden and unknown sources. They are generally very good at reducing this error. However, there are three obvious disadvantages to the calibration approach. (1) Coarse-scale data are required to perform the calibration. These data are often difficult to acquire, and their absence may, in fact, be the motivation for scaling up from fine-scale data. (2) The parameter estimates are valid within the range of the calibration data, but the reliability of the calibration may decrease outside this range. (3) Only responses to changes in variables already represented in the calibrated equation can be simulated. Responses to variables that would become incorporated if the equation were further transformed through a rigorous aggregation would be lost. For example, calibrations of Eqs. 2 or 8 could not be used to simulate responses to changes in canopy leaf area because L_0 is missing in these equations. On the other hand, because L_0 is in Eq. 7, it could be calibrated and used for this purpose.

Combined approaches

The four error-reduction approaches described above can also be used in combination with one another. For example, the equation derived from a partial transformation for irradiance (Eq. 7) could be used instead of the uncorrected relationship (Eq. 3) in a partitioning for light-use efficiency. The number of combinations of error sources and corrections is large, and will not be illustrated here. However, combinations culminating with a calibration (it only makes sense to do the calibration last) are particularly powerful. For example, in the thicker canopy (L_0 = 6), the partial transformation for irradiance (Eq. 7) was the most effective approximation to the fully corrected equation (Eq. 9). This transformation not only reduces the aggregation error, it also improves the shape of the curve. Calibrations of this already partly corrected function might therefore be expected to work even better than either the transformation or the calibration alone (Fig. 7). Indeed, calibration of the irradiance transformation (Eq. 7) to data for the thicker canopy reduced error to 4% of the uncorrected error (SRE = 0.77%). This procedure is therefore about three times more effective than any of the other corrections! Because the transformed equation already approximated the full correction well, its calibration yielded estimates of P_m and E_0 that were not very different from the fine-scale values (9.8 μmol·m^{-2}·s^{-1} and 0.040 μmol/μmol, respectively). Following any of the corrections with a calibration is highly recommended.

Error Propagation

In each of the correction procedures described above, except the calibration procedure, the variability among the fine-scale components had to be characterized and somehow incorporated into the corrected equations. Incorporating this variability inevitably increases the number of parameters in the corrected equations over the number in the uncorrected equation. Each of these parameters needs to be estimated in some way. Error associated with these estimates can propagate through the model and decrease the *precision* of the predictions. This loss in precision can be rectified to some extent if all the parameters are estimated simultaneously to account for the covariance between parameters (e.g., with a Monte Carlo calibration like that of Hornberger and Cosby 1985). Most parameter estimates, however, will be independent of one another even if the parameters are not truly independent. There can be a substantial trade-off between this decrease in *precision* due to error propagation in the corrected equations and the decrease in *accuracy* associated with aggregation error in the uncorrected or partially corrected equations (O'Neill 1973).

Clearly, this trade-off between precision and accuracy depends not only upon the magnitude of aggregation errors but also upon how precisely the parameters can be estimated. As an illustration, we assume that all the parameters can be determined to within ±5% of their true values. Estimates were made for the thicker canopy (L_0 = 6) with the same 100 top-of-the-canopy irradiance values used above. Thirty-two combinations of parameter values were used in the equations by assigning P_m, E_{01}, E_{02}, k, and L_0 values either 5% above or 5% below their original values. In the case of the calibrated model, four combinations of param-

eter values were used by assigning P_m and E_0 values $\pm 5\%$ of the calibrated values. Estimates made with the altered parameters were compared directly to estimates made with the original parameters using the same equation, not to estimates using the fully corrected equation. In this way, errors caused by parameter uncertainty were isolated from those due to aggregation. Propagated parameter error was quantified by calculating the standard relative error (SRE, Eq. 13) between the estimates with and without parameter alteration. The largest of the 32 (or 4) SRE values for each correction was used as an index of the potential propagated error due to parameter uncertainty.

The propagated error was $\approx 10\%$ (range = 9.47 to 12.01%) for all but two of the corrections; the error associated with the moment expansion for irradiance was >25%, and the error associated with the calibration was only 5%. The large error for the moment expansion was the result of a high sensitivity of the derivatives and moments to the parameter perturbations. If more terms were included in the expansion, this error would be attenuated and would eventually converge to 11.98%, the propagated error for the fully corrected equation (Eq. 9). However, as with the expansion estimates themselves, it is difficult to determine a priori the number of terms required before the error converges.

The low error associated with the calibration was a direct consequence of the calibrated equation having only two parameters rather than four or five as in the other equations. Even when combined with the aggregation error (3.29%), the total error associated with the calibration was only 8.29%, while that for the fully corrected equation (Eq. 9) was 11.98%, all of which resulted from the propagation of parameter uncertainty (i.e., no aggregation error). The total error associated with the calibration could have been worse than that for the fully corrected equation if the error associated with estimating parameters for the calibrated equation had been larger (about $\pm 10\%$). A small number of parameters does not guarantee small total errors because aggregation error and the precision of parameter estimates must also be considered.

Summary and Conclusion

Most ecological knowledge is derived from small, easily measured and easily manipulated systems. However, the focus of ecology is now expanding to include far larger systems that are more difficult to measure or manipulate (e.g., global systems). Models of these large, coarse-scale systems will not be able to retain the degree of detail and resolution used in models of fine-scale, more tractable systems, and the equations derived for the fine-scale processes cannot be applied directly in coarse-scale models. Yet the ability to apply the knowledge embodied in the fine-scale equations to these coarse-scale models is extremely desirable. How can fine-scale knowledge be translated so it can be applied at coarser scales? This is the question we have addressed in this paper.

As the scale of perspective becomes coarser, and larger systems are considered, smaller components are generally aggregated to maintain manageable complexity in the models. Because of variability among the components being aggregated, any nonlinear relationship describing the behavior of the smaller components will be unreliable when directly applied to the aggregate. The error will increase with the variability among fine-scale components and with the depth of the concavity in the equations describing the fine-scale process.

Variability among fine-scale components can be associated with any of the variables in the fine-scale equation. To make effective corrections, it is often useful to identify, a priori, the sources that cause the most severe aggregation error. To this purpose, we used a Monte Carlo sampling procedure to quantify the contributions of various sources of error. This procedure can be used to rank sources of error so that corrections can first be made for the most severe ones. However, this ordering can change from situation to situation; a correction for one variable might be more effective than a correction for another variable under some conditions, but the reverse might be true under other conditions. In the canopy example used in this paper, corrections for light extinction were more effective than corrections for light-use efficiency when the canopy was thick, but in thin canopies the reverse was true.

Application of the statistical expectation operator to the fine-scale equations is an effective means of transforming fine-scale equations into coarse-scale process descriptions. Fine-scale variability is thereby incorporated into the transformed equations through a statistical characterization. In theory, this procedure could be applied sequentially to all forms of variability among the fine-scale components, thus providing a rigorous analytical transformation of the fine-scale equations so they can be used to describe coarse-scale properties of the aggregated system. However, the procedure is sometimes difficult or impossible to apply because (1) adequate statistical characterizations of the fine-scale variability are difficult to find and (2) each transformation makes the equations more complex and subsequent transformations become more and more difficult.

There is also a cost associated with such transformations. Each transformation incorporates parameters characterizing a particular aspect of the fine-scale variability into the corrected equations. The error associated with the estimation of these new parameters will be propagated through the model, and, as more parameters are added, the model predictions become less precise. Thus, although the transformations increase the accuracy of the model by reducing aggregation error, they also decrease its precision by incorporating

more error in the form of more parameters. Because of the increase in complexity of the equations, there is also a loss in heuristic value with each successive transformation. The simplicity of Eq. 3, for example, makes it far easier to comprehend than Eq. 9. The simpler equations are also easier to analyze and manipulate, making them far more useful. Nevertheless, transformations for some of the more serious sources of aggregation error are often worthwhile.

Three less rigorous forms of transformation were also described. The first, a moment expansion, is an approximation of the rigorous transformation based on the expectation operator. The expectation operator applied to a fine-scale equation can be expressed as an infinite series. The first term in the series is the fine-scale equation evaluated at the mean value of the variable for which corrections are being made. The remaining terms are adjustments to this initial estimate using successively higher-order statistical moments (variance, skew, kurtosis, etc.) of the distribution of the variable. The series will eventually converge so that successive terms do not significantly improve the estimate. However, unless the higher-order derivatives of the fine-scale equation or the higher-order moments of the variable distributions are known to be zero, it can be difficult to determine a priori how many terms will be necessary for convergence.

Partitioning, the second of the less rigorous transformations, reduces the aggregation error by reducing the degree of aggregation. Instead of aggregating all of the fine-scale components into a single coarse-scale component, they are first aggregated into a manageably small number of intermediate-scale components. A weighted mean of their intermediate-scale behaviors is then used as an estimate of the coarse-scale relationship. This partitioning is most effective when variability within each of the intermediate-scale components is as small as possible. An optimal partitioning will therefore depend upon the distribution of the variable for which corrections are being made.

The third of the less rigorous corrections is a direct calibration of the fine-scale equation to coarse-scale data. The major advantage of this technique is that the complexity of the model can be kept to a minimum. This means that there will be fewer parameters and therefore fewer sources of error. This retention of precision can often outweigh any gain in accuracy achieved by using one of the other transformations. However, there are three important disadvantages to calibration: (1) they require coarse-scale data; (2) predictions based upon them can become less reliable outside the range of the data used for calibration; and (3) they can only be used to simulate responses to changes in variables that are already explicitly represented in the calibrated equation. Nevertheless, very few ecological models can be realistically applied without some final calibration at the coarser scale. Thus, the most important contribution of this procedure may be found in its combination with other transformation techniques. The best results will be achieved with an aggregation scheme that optimizes the balance between the gain in accuracy associated with reduced aggregation error and the loss of precision associated with increased model complexity.

Any effective aggregation of fine-scale knowledge to anticipate coarse-scale response requires some data or information from the coarser scale. Transformations using the expectation operator and the moment expansion require information on the statistical distribution of fine-scale components across the extent of the coarser scale. Calibration requires data on the aggregate response at the coarser scale, and partitioning requires data (e.g., means) from intermediate scales. Statistical description of fine-scale components may in many circumstances be easier to acquire than data on the aggregate response of coarse-scale components. The distribution of climate-growth response for individual forest stands vs. holistic measurements of the aggregate response of a forested landscape to climate change is one example. On the other hand, aggregate responses may, in some cases, be more empirically accessible than the statistical description of fine-scale components. For example, measurements of gas exchange for low-stature canopies may be more practical than gathering data on the statistical distribution of the comparable response for individual leaves and their interactions.

Thus, the availability of data from the coarser scale will, in large part, determine the choice of aggregation scheme. The most effective approach is determined by the specific application, and general recommendations are hard to make. However, our own general approach can be summarized as follows:

(1) If high quality data from the coarser scale are available, use them to calibrate either the fine-scale equation directly or the least biased of the partially corrected equations you are able to compute. It is difficult to imagine circumstances where, data permitting, it is not preferable to approach a problem directly at the scale of interest. However, it is important to remember that although a calibrated parameter may retain the name and units associated with it at finer scales, the parameter is now an "effective" parameter whose quantification is no longer bound by empirical measurements at finer scales (Cushman 1986, Beven 1989).

(2) If coarse-scale data are unavailable, or if the express purpose of the activity is to determine if or how fine-scale phenomena are expressed at coarser scales, a statistical description of the fine-scale components across the extent of the coarser scale should be acquired. The fine-scale attributes should then be ranked by their contribution to aggregation error. Transformations combining the expectation operator, moment expansions, and partitioning can then be applied to the important sources of error. Transformations can be terminated when the complexity of the transformed

model or the potential for error propagation exceeds tractable or practical limits.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation through grants to the Marine Biological Laboratory (BSR-8718426, BSR-8702328).

LITERATURE CITED

Bartell, S. M., W. G. Cale, R. V. O'Neill, and R. H. Gardner. 1988. Aggregation error: research objectives and relevant model structure. Ecological Modelling **41**:157–168.

Beven, K. 1989. Changing ideas in hydrology: the case of physically-based models. Journal of Hydrology **105**:157–172.

Box, G. E. P., and G. M. Jenkins. 1976. Time series analysis forecasting and control. Holden-Day, San Francisco, California, USA.

Brunk, H. D. 1975. An introduction to mathematical statistics. Xerox College Publishing, Lexington, Massachusetts, USA.

Cale, W. G. 1988. Characterizing populations as entities in ecosystem models: problems and limitations of mass-balance modelling. Ecological Modelling **42**:89–102.

Cale, W. G., and P. L. Odell. 1980. Behavior of aggregate state variables in ecosystem models. Mathematical Biosciences **49**:121–137.

Cale, W. G., and R. V. O'Neill. 1988. Aggregation and consistency problems in theoretical models of exploitative resource competition. Ecological Modelling **40**:97–109.

Cale, W. G., R. V. O'Neill, and R. H. Gardner. 1983. Aggregation error in nonlinear ecological models. Journals of Theoretical Biology **100**:539–550.

Caughley, G. 1982. Vegetation complexity and dynamics of modelled grazing systems. Oecologia (Berlin) **54**:309–312.

Cosby, B. J. 1984. Dissolved oxygen dynamics of a stream: model discrimination and estimation of parameter variability using an extended Kalman filter. Water Science and Technology **16**:561–569.

Cosby, B. J., and G. M. Hornberger. 1984. Identification of photosynthesis–light models for aquatic systems. I. Theory and simulations. Ecological Modelling **23**:1–24.

Cushman, J. H. 1986. On measurement, scale, and scaling. Water Resources Research **22**:129–134.

Draper, N. R., and H. Smith. 1981. Applied regression analysis. John Wiley & Sons, New York, New York, USA.

France, J., and J. H. M. Thornley. 1984. Mathematical models in agriculture. Butterworths, London, England.

Gardner, R. H., W. G. Cale, and R. V. O'Neill. 1982. Robust analysis of aggregation error. Ecology **63**:1771–1779.

Gross, L. J. 1982. Photosynthesis dynamics in varying light environments: a model and its application to whole leaf carbon gain. Ecology **63**:84–93.

Hirata, H., and R. E. Ulanowicz. 1986. Large-scale system perspectives on ecological modelling and analysis. Ecological Modelling **31**:79–104.

Hornberger, G. M., and B. J. Cosby. 1985. Selection of parameter values in environmental models using sparse data: a case study. Applied Mathematics and Computation **17**:335–355.

Iwasa, Y., V. Andreasen, and S. Levin. 1987. Aggregation in model ecosystems. I. Perfect aggregation. Ecological Modelling **37**:287–302.

Iwasa, Y., S. Levin, and V. Andreasen. 1989. Aggregation in model ecosystems. II. Approximate aggregation. IMA Journal of Mathematics Applied in Medicine and Biology **6**:1–23.

Landsberg, J. J. 1986. Physiological ecology of forest production. Academic Press, London, England.

Mallows, C. L. 1973. Some comments on C_p. Technometrics **15**:661–675.

O'Neill, R. V. 1973. Error analysis of ecological models. Pages 898-908 *in* D. J. Nelson, editor. Radionuclides in ecosystems. CONF-710501. National Technical Information Service, Springfield, Virginia, USA.

———. 1979. Transmutations across hierarchical levels. Pages 59–78 *in* G. S. Innis and R. V. O'Neill, editors. Systems analysis of ecosystems. International Co-operative, Fairland, Maryland, USA.

O'Neill, R. V., and B. Rust. 1979. Aggregation error in ecological models. Ecological Modelling **7**:91–105.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1986. Numerical recipes: the art of scientific computing. Cambridge University Press, New York, New York, USA.

Raich, J. W., E. B. Rastetter, J. M. Melillo, D. W. Kicklighter, P. A. Steudler, B. J. Peterson, A. L. Grace, B. Moore III, and C. J. Vorosmarty. 1991. Potential net primary productivity in South America: applications of a global model. Ecological Applications **1**:399–429.

Rastetter, E. B., M. G. Ryan, G. R. Shaver, J. M. Melillo, K. J. Nadelhoffer, J. E. Hobbie, and J. D. Aber. 1991. A general biogeochemical model describing the responses of C and N cycles in terrestrial ecosystems to changes in CO_2, climate, and N deposition. Tree Physiology, *in press*.

Sellers, P. J., W. J. Shuttleworth, and J. L. Dorman. 1989. Calibrating the simple biosphere model for Amazonian tropical forest using field and remote sensing data. Part 1. Average calibration with field data. Journal of Applied Meteorology **28**:727–759.

Sinclair, T. R., C. E. Murphy, Jr., and K. R. Knoerr. 1976. Development and evaluation of simplified models for simulating canopy photosynthesis and transpiration. Journal of Applied Ecology **13**:813–830.

Snedecor, G. W., and W. G. Cochran. 1967. Statistical methods. Iowa State University Press, Ames, Iowa, USA.

Templeton, A. R., and L. R. Lawlor. 1981. The fallacy of the averages in ecological optimization theory. American Naturalist **117**:390–391.

Wagner, H. M. 1969. Principles of operations research. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Waring, R. H., and W. H. Schlesinger. 1985. Forest ecosystems: concepts and management. Academic Press, New York, New York, USA.

Yeh, R. Z. 1973. Modern probability theory. Harper & Row, New York, New York, USA.

Young, P. 1984. Recursive estimation and time-series analysis. Springer-Verlag, New York, New York, USA.