CHAPTER 8

BUILDING UP WITH A TOP-DOWN APPROACH:

The Role of Remote Sensing in Deciphering Functional and Structural Diversity

CAROL A. WESSMAN AND C. ANN BATESON

8.1 INTRODUCTION

> *". . . . particularity and contingency, which characterize the ecological sciences, and generality and simplicity, which characterize the physical sciences, are miscible, and indeed necessary, ingredients in the quest to understand humankind's home in the universe."*
>
> *~ John Harte, 2002*

Historically, ecologists have dealt with the complexity of ecological systems through small, place-centered studies. This approach, in controlling for environmental variation, has led to important advances in our understanding of mechanisms behind ecological phenomena, and has moved us more directly from pattern to causal processes (Harte 2002). However, the need to develop broader "universal laws" in ecology is important in the context of global-wide environmental change. While ecological "laws" may lack the exactness and universality of physical laws, they will improve our power to predict the consequences of change due to human activities and climate variation, "signposting routes to a sustainable future" (Lawton 2001).

Ecological scaling, and hence any efforts to define universalities, is challenged by inherent nonlinear synergies and heterogeneity, cross-scale processes, thresholds, and emergent properties of ecosystems (e.g., Wu 1999, Peterson 2000, Wu and Li, Chapter 2). Techniques are required that are able to translate place-centered, mechanistic understanding (the "peculiarities and contingencies" *sensu* Harte 2002) across a range of spatial and temporal scales. Remote sensing of the Earth's surface, while limited in its ability to fully address all the challenges, helps constrain the scaling problem through its synoptic view of biophysical and biochemical structure across different scales (Wessman 1992, Wessman and Asner 1998). The structure of landscapes and regions (i.e., the properties of cover types and their distribution) are

J. Wu, K.B. Jones, H. Li, and O.L. Loucks (eds.),
Scaling and Uncertainty Analysis in Ecology: Methods and Applications, 147–163.

of great importance and interest to the ecological scaling effort for two reasons. First, structure superimposes constraints on the functioning of ecological systems at broad to finer scales. Second, the structure itself is an expression of the functional properties that emerge from interactions among biological, physical and geochemical processes. Scaling in remote sensing, by its attention to surface heterogeneity and the derivation of surface parameter algorithms (Chen 1999), takes an important role in earth system science, predicated on the assumption that we recognize the appropriate features to be scaled (Wessman 1992, Wessman and Asner 1998). In many respects, remote sensing mandates a generality in our observations of the earth's surface that, welcome or not, forces a new perspective of ecological properties. Through interchange between observational-based science and ecological scaling theory, generalities in ecological dynamics will surface.

Scaling in remote sensing is controlled in two fundamental ways: (1) heterogeneity in composition and configuration of the landscape under observation relative to sensor characteristics, and (2) nonlinear functional relationships between surface radiation and ecologically relevant parameters. The objective of this paper is to explore these aspects of the remote sensing scaling problem, and the power and uncertainties they interject into the upscaling of field measurements. First, we introduce the ecological variables and relationships that the community has or is attempting to measure with remote sensing. We then review briefly the radiation and remote sensing properties that are important to upscaling. Finally, we synthesize some of the more important sources of uncertainties and error in algorithms used to scale surface parameters.

8.2 ECOLOGICAL VARIABLES RETRIEVED BY REMOTE SENSING

Extraction of variables from imagery falls into two broad categories. First, classification techniques are used to categorize the landscape into discrete recognizable units with relevance to various mapping aims. Second, continuous variables are retrieved on the basis of spectral-biophysical relationships in each pixel. In the former, within-class variance is ultimately ignored as pixels are assigned discrete values representing a given category. In the latter, pixel values will lie within the range of variability of the retrieved variable contained by the observation. Continuous variables such as leaf area index (LAI), fractional cover and fraction of absorbed photosynthetically active radiation (fAPAR) represent basic components or singular functions of the landscape and will largely scale as a function of the radiative transfer properties of landscape components, their relative dominance in the grid cell, and the linearity of the retrieval algorithm. Scaling of discrete, categorical variables will be influenced by the size and nesting of ground components detectable at diverse resolutions.

The ecological variables that take on significance as scales are increased and that can be estimated remotely are integrative in nature or represent important constraints on processes (Table 8.1). These include structural variables such as LAI, biomass, land cover, and fractional cover of landscape components (e.g., green vegetation, bare soil) that quantify, to greater or lesser degrees, the spatial heterogeneity important to extrapolation or modeling of related processes. fAPAR is one of the

few parameters that provide a direct connection between ecosystem structure and function (Asner and Wessman 1997), and it also provides a means to link to other functional attributes of ecosystems such as nitrogen use, CO_2 assimilation, and water loss (Sellers 1987, Running et al. 1994, Field et al. 1995). Remote sensing of foliar chemistry is of interest due to its role in ecosystem productivity. Estimates of spatial variation in canopy chemistry from hyperspectral imagery, while empirically derived through local to regional calibration, are valuable for landscape-level links to ecosystem processes such as productivity and decomposition (Wessman et al. 1988, Zagolski et al. 1996, Martin and Aber 1997, Smith et al. 2002).

Table 8.1. *Ecological variables commonly derived from remotely sensed imagery (after Milne and Cohen 1999).*

Variable	*Type*	*Units*	*Applications*
Land Cover	Categorical	ha, %	Ecosystem model stratification, land use change, habitat characterization, ecosystem management
LAI	Continuous	$m^2\ m^{-2}$	CO_2 and trace gas exchange models or measurement extrapolation, carbon allocation
NPP, NEP	Continuous	$g\ m^{-2}\ yr^{-1}$	Estimation of ecosystem carbon gain
Fractional cover	Continuous	%	Ecosystem model stratification, land use/land cover change, succession, biophysical land surface modeling
Canopy chemistry	Continuous	%, $g\ m^{-2}$	Ecosystem productivity models, decomposition
Canopy geometry	Continuous	m	Land-atmosphere energy flux, climate models

Each of these variables gains significance with broadened extent in either of two ways. The importance of their magnitude and distribution, alone or in combination with other remotely sensed variables, may indicate configuration and/or connectivity of landscape components undetectable at the field level. Also, certain remotely sensed variables are incorporated into process models to drive or constrain simulations of biogeochemical process, land-atmosphere energy and trace gas flux, and large-scale climate models.

8.3 THE RETRIEVAL OF ECOLOGICAL VARIABLES

Remote sensing is a valuable means of upscaling ecological variables due to multiple scales of observations and stability of satellite platforms in space and time. Repetitive acquisitions at temporal scales of interest (e.g., diurnal, seasonal,

interannual, duration of transient weather systems) enable analyses of change to answer questions about trends and cycles. But, fundamental to the retrieval and interpretation of ecological variables from remote sensing platforms are the reflectance characteristics of the observed surface and its components. Aside from some types of mapping and sophisticated radiative transfer methods, remote sensing does not replace any widely used ground measurement in ecology (Prince 1999). Most parameters derived from remote sensing data require a certain level of calibration with ground conditions (van Leeuwen et al. 1997, Qi et al. 2000). Each technique for the inference of an ecological variable must be understood in terms of the factors contributing to the measurement of reflectance integrated over the pixel.

The reflectance signal measured by the sensor is the integrated outcome of a complex interaction of surface scattering properties, including single and multiple scattering of photons, and solar and sensor viewing geometry. For example, vegetation reflectance is primarily a function of tissue (wood, green, senesced) optical properties, canopy structure and geometry (leaf and stem area and orientation, foliage clumping), soil reflectance, illumination conditions, and viewing geometry (Ross 1981, Myneni et al. 1989). The structural attributes of ecosystems (e.g., species composition, vertical structure, canopy closure) determine the relative contributions of tissue, canopy and landscape factors that drive the variation in a reflectance signal (Asner 1998). Sorting out these factors requires an understanding of the sources of variation at each scale (which is ecosystem dependent) as well as an adequate sampling (spectral, angular, and temporal) of the spectrum.

Spectral algorithms like vegetation indices (VI) that are relatively simple and are composed of few spectral bands are widely used to monitor vegetation dynamics and to infer biophysical properties such as leaf area index (LAI) and fAPAR. The main shortcoming of vegetation indices is the lack of functional relationships with biophysical parameters (van Leeuwen et al. 1997, Qi et al. 2000). Empirical or modeled relationships with variables such as LAI, fAPAR, and biomass can be developed on a site-specific basis, but these relationships are subject to changes in vegetation properties, soil background, atmosphere and the sun-surface geometry. Validation efforts for the MODIS sensor are aimed at testing the accuracy of VI products via multi-scaled analyses in order to gain an understanding of the causes of errors for potential improvement in future algorithms (Tian et al. 2002a).

Radiative transfer models provide a means to assess the canopy radiation regime from a physical and mechanistic basis, and model inversions derive variables that are more directly related to biophysical properties of vegetation. Current models are based on the physics of photon scattering, and range in complexity from one-dimensional (vertical profile) algorithms to complex three-dimensional landscape simulations (Liang and Strahler 1993, Myneni and Asrar 1993, Kuusk 1995, Asner and Wessman 1997). The models that include scaled scattering characteristics of leaves, canopies, and soil can be used to explore the radiation regime in landscapes. For example, understanding the relative influence of structure at different scales (tissue, canopies, landscapes) on the fAPAR of an ecosystem helps determine what variables need to be accounted for and which can be ignored or held at a constant value in future studies (Asner et al. 1998b, Asner et al. 1998c). The anisotropic scatter of radiation by vegetation canopies has been exploited through measurement

and modeling of the bidirectional distribution function (BRDF) to retrieve LAI and canopy geometry (Li and Strahler 1992, Privette et al. 1994, Braswell et al. 1996, Asner et al. 1997, Asner 2000). This continues to be a strong area of research with new operational BRDF instruments (e.g., MISR, POLDER) (Jin et al. 2002, Chopping et al. 2003). Less quantitative, but analogous to these modeling approaches, spectral mixture analysis decomposes the reflectance signal into the fundamental contributing components of the landscape (e.g., soil, green foliage, senesced vegetation) and simplifies the interpretable connection to biophysical parameters (Wessman et al. 1997, Sabol et al. 2002).

In this paper, we focus on the upscaling of biophysical variables through remote sensing. However remote sensing methods in land cover classification and feature mapping are very important aspects of remote sensing scaling and need to be noted here. Classification is a well-established approach to map land-cover types that represent composites such as vegetation and land use, habitat, or ecosystem types. Land cover maps are used widely as one means to quantify landscape heterogeneity and parameterize the biophysical properties of plant canopies in models of climate and biogeochemical processes. Efforts continue to improve accuracies and quantify errors associated with classification algorithms (e.g., Hlavka and Dungan 2002), particularly in the interest of scaling to continental and global scales (e.g., Friedl et al. 2002, Lotsch et al. 2003). The ability to describe spatial patterns and the underlying processes that generate them is largely determined by the relationships between the objects in the scene and the scales at which we observe them. A significant literature exists on scaling issues specific to feature extraction, spatial structures, and the spatial variation in remote sensing imagery (Marceau and Hay 1999). Multiscale approaches to upscaling and feature extraction are being developed to contend with the multi-scaled and spatially distributed objects in a landscape (Hay et al. 1997, 2001).

8.4 QUANTIFYING BIAS AND ERROR

There are several sources of error and bias in analyses based on remote sensing that are caused by the indirect relationship between reflected radiation and surface characteristics. The most obvious source of error is the intervening atmosphere between the sensor and the reflecting surface. Although methods have been developed to remove atmospheric effects from imagery (Gao et al. 1993, Qu et al. 2000), these methods do not duplicate ground based spectra and can themselves delete spectral features critical for detecting the presence of ground components in the scene (Kruse and Dwyer 1993).

Another source of error that has been studied extensively in the literature and is algorithmic in nature involves extrapolating relationships between remotely sensed parameters and surface attributes from the scale for which they were developed to a coarser resolution. For example, a functional relationship between LAI and the NDVI, whose defining constants were found by relating the NDVI values at a 30 m pixel size to ground values adequately sampled on 30 m plots, may not be the correct relationship for predicting LAI from AVHRR or MODIS at 1 km pixels. In order for a quantitative algorithm to be scale invariant with respect to a particular

landscape, either the algorithm must implement a linear function between the parameters or ecological parameter values within the larger pixel must be constant (Hu and Islam 1997). Hence, for quantitative algorithms, heterogeneity and non-linearity are the two factors determining the magnitude of scaling errors and bias. Since most radiance-to-ground relationships are developed on a small scale, methods are needed to measure the potential aggregation error and ideally add a correction factor. We will review several approaches for estimating large-scale quantities using smaller scale, field-acquired measurements. Many of these methods require pixel-level knowledge of fractional coverage by ground components. We investigate spectral mixture analysis (SMA), which inverts a mixture model to retrieve cover fractions, as a tool for measuring heterogeneity that can be incorporated into scaling methods.

Most but not all algorithms bridging remote sensing and surface attributes require adequate ground sampling for calibration and validation. Misregistration of image pixels with their ground locations is common and has an impact relative to the scene heterogeneity. Spatial statistics has provided tools such as the variogram, local variance and kriging to help in designing efficient sampling schemes by detecting spatial correlations and consequently appropriate lags between samples to achieve a set of statistically independent values. We will review some of these methods and issues related to error in their application.

8.4.1 Aggregation Error

We first present a theoretical framework for understanding the issues involved in applying a functional relationship from the scale for which it is correct to an aggregate. Let R be a function between the bands of a remote sensing instrument with spatial resolution L and a surface parameter P. That is,

$$R(b_1^L(i), ..., b_m^L(i)) = P^L(i) \tag{8.1}$$

where $b_j^L(i)$ is reflectance of the i^{th} pixel in band j and $P^L(i)$ is the ground parameter value in pixel i. Suppose we aggregate the pixels of the instrument into superpixels of size nL. We can compute two quantities:

$$\frac{1}{n^2}\sum_{i=1}^{n^2} R(b_1^L(i), ..., b_m^L(i)) \tag{8.2}$$

and

$$R(\frac{1}{n^2}\sum_{i=1}^{n^2} b_1^L(i), ... \frac{1}{n^2}\sum_{i=1}^{n^2} b_m^L(i)) \tag{8.3}$$

An algorithm implementing Equation 8.2 will be described as parameter-aggregated since it is the mean of the ground parameter $P^L(i)$ over the subpixels that comprise the superpixel. Algorithms implementing Equation 8.3, on the other

hand, will be described as band-aggregated since the bands are aggregated into a superpixel before the function R is applied. R is scale invariant if and only if Equation 8.2 and Equation 8.3 are equal and these two quantities are identically equal if and only if R is a linear function of its variables. However, for a particular scene, if objects such as forests or grasslands have a typical size at least as large as the size of the superpixel, then all subpixels will be similar and hence their bands values and means will be approximately equal and scale invariance will persist for any function. Equality of band values over the subpixels, of course, will not happen in a real scene, but band variation over an object often has a negligible effect on the equality between Equation 8.2 and Equation 8.3 above (e.g., within forest variance in LAI produces small scaling errors). Since R was developed for resolution L, Equation 8.2 is the correct value for the ground parameter computed for the superpixel. However, for large-scale imagery such as AVHRR, we usually do not have subpixel information and ground sampling to produce relationships is not feasible. We look at error associated with using Equation 8.3 to estimate LAI from the normalized difference vegetation index (NDVI). But, first, we investigate the nonlinearity of the NDVI itself.

8.4.1.1 NDVI

The NDVI exploits differences in vegetation reflectance response between the red and NIR to detect ground vegetation properties and is computed by

$$NDVI = (NIR - red)/(NIR + red) \tag{8.4}$$

However, since the NDVI is not a linear function of its variables (NIR and red reflectance), the NDVI of a superpixel need not equal the mean NDVI of its subpixels (i.e., the parameter-aggregated NDVI doesn't equal the band-aggregated NDVI). This discrepancy will introduce error in scaling up a functional relationship R between the NDVI and a ground parameter even when R is linear provided that the lower resolution pixels are not homogeneous. Hu and Islam (1997) investigated the effects of landscape heterogeneity on scaling errors with respect to the NDVI and reported the following:

a) The relative difference between the parameter-aggregated and band-aggregated NDVI (i.e., relative scaling error) is a function of within-superpixel variance of red and NIR reflectances and within superpixel covariance between the two bands. If the distribution of red and NIR reflectances among the subpixels is too highly heterogeneous, then the error will be too great to approximate the parameter-aggregated NDVI with the band-aggregated NDVI algorithm.
b) There was significant scaling error in a hypothetical example when the band-aggregated algorithm was used to estimate NDVI.
c) A correction term CT to account for within-superpixel heterogeneity can be computed from information found on the superpixel level and in the hypothetical example band-aggregated NDVI + CT gave a good

approximation to the parameter-aggregated NDVI. The term *CT* is computed through functional relationships parameterizing within-superpixel variance and covariance between NIR and red bands with reflectances acquired by the sensor.

The NDVI error measurement and correction algorithm in Hu and Islam (1997) has not been tested on remotely sensed data from two different instruments at two different scales. Since its assumptions are simplistic (e.g., all vegetation has the same reflectance and likewise for background) and only first and second order correction terms were considered, the result remains preliminary until such testing is performed. Aman et al. (1992) degraded SPOT and TM data collected over tropical sites in West Africa and agricultural fields in France from 20 and 30 m respectively, to resolutions ranging from 200-1000 m, which are more suitable for global vegetation studies. They found a significant linear correlation between parameter-aggregated and band-aggregated NDVI for resolutions below 1000 m with slope and intercept close to 1 and 0 respectively. Hence, they concluded, on the basis of their samples, that the parameter-aggregated NDVI can be estimated from the band-aggregated NDVI with acceptable errors. That is, the errors are less than the uncertainties relating the high resolution NDVI and ground parameters and the errors resulting from radiometric corrections. However, similar experiments with other landscape types need to be performed to determine the domain of acceptability for the band-aggregated algorithm.

8.4.1.2 NDVI and LAI

The infeasibility of collecting LAI values in the field on a scale required for global and regional vegetation studies mandates efforts to compute LAI through functional relationships with vegetation indices derived from remotely sensed imagery or through inversions of radiative transfer models. Chen and Cihlar (1996) found the nonlinear relationship

$$NDVI = 0.5520 * LAI^{0.1844} \tag{8.5}$$

between the NDVI computed from TM imagery and LAI values collected in boreal conifer forests in the Boreal Ecosystem-Atmosphere Study (BOREAS) site with a plant canopy analyzer (*LAI-2000*, *Licor*).

The scaling error incurred in upscaling this algorithm from the TM 30 m pixel resolution to 1 km AVHRR pixels was investigated by first degrading the TM imagery to the AVHRR resolution (Chen 1999). Study areas of 990 m × 990 m with mixtures of vegetation and water were selected from the imagery to give a range of water coverage from 0% to 93%. Coexistence of these two highly contrasting surfaces in the same low resolution pixel was expected to produce large scaling errors since the LAI retrieval algorithm is nonlinear. The correct computation of LAI is the parameter-aggregated algorithm which first computes LAI from the TM-derived relationship at each subpixel and then averages LAI over the TM subpixels of the AVHRR pixel. However, typically subpixel values would not be available and

Chen (1999) investigates the error in scaling up the TM-derived relationship Equation 8.5 via the band-aggregated algorithm, which uses the NDVI of the AVHRR pixel as input into Equation 8.5. Note that in the band-aggregated algorithm, non-linearity appears both in the computation of the NDVI from the NIR and red reflectance bands and in the computation of LAI from the NDVI. A theoretical error analysis based on mixed vegetation and water pixels derived a relative error that only depends on knowledge of the water fraction w within each pixel and the scaling exponents relating AVHRR NDVI and TM NDVI to LAI. The latter exponent ($TM\ NDVI = cL^b$) corresponds to 0.1844 in Equation 8.5. Constants of the AVHRR power law ($AVHRR\ NDVI = c_0 L^{b0}$) can be estimated if subpixel water fractions are known since AVHRR NDVI at the BOREAS site has a strong power law relationship with (1-w) and LAI in the mixed pixel is proportional to (1-w), with an unknown constant of proportionality equal to LAI of 100% vegetation, i.e., LAI(100%). It is shown in Chen (1999) that the relative error for a pixel is given by

$$\frac{(\mathrm{LAI}_{(\text{parameter - aggregated})} - \mathrm{LAI}_{(\text{band - aggregated})})}{\mathrm{LAI}(100\%)} = (1-w) + (1-w)^{b_0/b} \qquad (8.6)$$

where w is the fraction of the pixel covered by water. The scaling exponent b for AVHRR pixels over the BOREAS site had a value of 0.68; by differentiating Equation 8.6 with respect to w, the maximum relative error is 0.44 and occurs for $w = 0.384$. Computations of relative errors for the 990m × 990m study sites fell encouragingly close to the theoretical predictions derived from Equation 8.6.

Chen (1999) also revealed problems with linear algorithms. Although changing densities within the pixel do not introduce error into linear algorithms, scaling errors do occur when a linear relationship changes over diverse landscape components (e.g., vegetation versus water) and discontinuities are introduced. For example, the simple ratio (SR) scaled with LAI as $SR = 2.78+0.824*LAI$ at the BOREAS sites and a bias was introduced since SR over water ($LAI = 0$) is approximately 1 and not 2.78. Scaling errors are easier to derive for the nonlinear algorithm, since in the linear case error derivation requires knowledge of LAI (100%) as well as the water fraction.

With both the linear and the nonlinear algorithms, negative biases occur when an algorithm from a finer resolution is used to estimate LAI at coarser resolutions. For a pure pixel (all vegetation), the bias with the nonlinear algorithm in Chen (1999) was less than 2% and there was no bias for the linear algorithm. Hence, errors with pure pixels could be ignored. When water and vegetation were both present, negative biases occurred in Chen (1999) close to 40% for the linear algorithm and exceeded 44% for the nonlinear algorithm. Errors increased with increased heterogeneity.

Another technique for estimating LAI per pixel is based on inversion of a radiative transfer model that produces top-of-canopy reflectance in terms of leaf tissue and soil radiative properties, LAI and leaf angle distribution (LAD). From remotely sensed data, top-of-canopy reflectance is known and other parameters such as LAD and leaf optical properties may be estimated from other sources or constrained to lie within a realistic range of values. Top-of-canopy reflectance

viewed at different, but known, sun sensor geometries gives several different constrained equations which can be inverted to retrieve values of parameters like LAI. More complexity is added to the model when other elements besides vegetation canopies are included in the pixel. In this case of pixel heterogeneity, the one-dimensional (1-D) model is supplanted with a three-dimensional (3-D) model equation, which accounts for horizontal transport of photons potentially interacting with more than one element type within the pixel. Tian et al. (2002b) retrieved LAI values from AVHRR 1 km data aggregated to several coarser resolutions (8, 16, 32, and 64 km). Their imagery was acquired over North America and classified into 6 vegetation biomes and bare soil. A lookup table was used in their model equation to associate with each biome its single scattering albedo, which measures the total scattering of energy per unit volume of the biome canopy. Coarser scales result in more mixed pixels. However, equating reflectance of mixed pixels with subpixel reflectance averages in the model equation neglects the effects of heterogeneous scattering elements within the pixel on the radiative regime and can lead to significant errors in the retrieval of LAI (Tian et al. 2003). Tian et al. (2003) found large LAI errors when forests were minority biomes within non-forest pixels and developed a spatial resolution-dependent radiative transform formulation. In this formulation, the single scattering albedo is adjusted to become a weighted average of the single scattering albedos of the six biomes with weights equal to the fractional cover within the pixel. Hence, again, knowledge of fractional cover is a requirement for estimating LAI.

8.4.2 Spectral Mixture Analysis

During the 1980's, researchers began to examine spectral mixture analysis (SMA) as a means to characterize subpixel heterogeneity by modeling a pixel's reflectance as a linear combination of the reflectance spectra of ground components (e.g., soil, green vegetation, dead vegetation, rock, etc.), called endmembers (Adams and Adams 1984, Adams et al. 1986, Smith et al. 1990). The coefficients in the linear model should lie between 0 and 1 and may be constrained to sum to 1. Physically, they correspond to the fractional coverages of the ground components in the pixel. Consequently, SMA is a promising tool for providing heterogeneity parameters needed to extend algorithms from finer to coarser resolution.

Asner and colleagues (Asner et al. 1997, Asner et al. 1998a) combined SMA with an easily inverted 1-D model to calculate LAI for woody and herbaceous vegetation types in a complex savanna landscape. Inversion of the 1-D model for each of the cover types, in effect, accounted for the spatially heterogeneous landscape and, avoided a computationally intense inversion of a 3-D model. The study was initiated with high resolution spectral mixture analysis (Landsat TM) to compute fractional cover of trees, shade, senescent grass, bare soil and water. A suite of AVHRR images was acquired over the same area at different sun-sensor geometries. For each AVHRR image, the SMA model using the TM fraction covers was inverted to produce NIR and red reflectance values for tree, grass, shade and soil at the 1-km scale of the AVHRR pixel (Asner et al. 1997). Inconsistent shade fractions resulting from different sun-sensor geometries were corrected using a

geometrical-optical model (Li and Strahler 1992) prior to the inversions. These angular reflectances were used with a 1-D radiative transfer model to compute LAI for each vegetation type (tree and grass). Regional canopy LAI was computed by multiplying the type LAIs by their respective fractional covers. SMA in this example was critical not only because of the efficiency of inverting a 1-D radiative transfer model to compute LAI for two vegetation types (tree and grass), but also because inversion of the AVHRR SMA model produced endmember reflectance values for the AVHRR instrument, resolution and sun-sensor geometry.

The promise of SMA for solving scaling problems must be tempered, of course, by recognition of several sources of errors in the mixture model. First of all, multiple scattering of photons between different ground components can invalidate the linearity assumption. However, fraction errors from nonlinear mixing can be minimized by acquiring images from view angles close to the hot spot direction (Villeneuve et al. 1998). SMA requires the knowledge of endmembers, which are reflectance spectra of pure ground components. Selection of endmember spectra is the most difficult task in SMA and the most profound source of cover fraction errors. Perhaps, the most common methods of acquiring endmembers are collecting them from the field or picking pixels from the image that are homogeneously covered by one ground component. However, it is very difficult to align field-collected spectra with image spectra even after (or because of) atmospheric correction or conversion from radiance to reflectance (Kruse and Dwyer 1993). Moreover, remotely sensed images over arid and semi-arid landscapes may not have at their resolution pure pixels of green vegetation and in this situation using image endmembers will distort all cover fractions. A promising solution to the endmember selection problem has been methods that derive endmembers from the variance structure of the data (Boardman 1993, Bateson and Curtiss 1996) using principal component analysis. These derived endmembers do not necessarily coincide with pixel reflectances and may represent pure spectra when there are no pure pixels in the image. Moreover, since they are derived from the image, they have been subjected to all image pre-processing. In recent years, the assumption that each ground component is represented by a unique spectral signature has been questioned and new mixture models (Asner and Lobell 2000, Bateson et al. 2000) have been devised that substitute for a single endmember spectrum a bundle or collection of spectra representing endmember variability. Bundle unmixing produces ranges of possible fraction values (Bateson et al. 2000) or mean and standard deviation (Asner and Lobell 2000).

8.5 CALIBRATION AND VALIDATION

Tracking ground characteristics such as LAI, fAPAR and biomass with indices such as the NDVI derived from spectral reflectance requires ground sampling to establish functional relationships between the ground characteristics and reflectance-based indices, whose values have no direct physical interpretation. Regression modeling and curve fitting are common methods used to determine from image pixel values and corresponding ground samples the best equation to relate the imagery to the landscape. Because typical pixel sizes range from 20m to 1km and the extent of the

imagery is usually very large, efficient sampling strategies adequately representing image and ground variability are needed. Random sampling with a sample size $n = (\sigma t / e)^2$, where σ is the standard deviation, e the desired error and t the Student's t-value for a 95% confidence interval and (n-1) degrees of freedom, can result in unnecessary ground sampling since it does not consider spatial correlations. That is, sampling neighboring pixels or neighboring regions within a pixel with correlated values can introduce costly redundancies. Systematic sampling based on spatial statistics may achieve results at least as good with a fraction of the effort.

An important tool of spatial statistics is the semivariogram, which is a function $\gamma(h)$ measuring the average dissimilarity between parameter values sampled at ground locations or computed for image pixels that are h units apart. The semivariogram at lag h is mathematically defined by

$$\gamma(h) = 1/2k(h) \sum_{i=1}^{k(h)} (V(x_i) - V(x_i + h))^2 \tag{8.7}$$

where $V(x_i)$ is the value of the parameter at ground location or pixel x_i, $V(x_i+h)$ is the value at a location or pixel h units away and $k(h)$ is the number of differences at lag h. In many natural scenes, $\gamma(h)$ will increase with h since nearby locations have similar characteristics compared to those at a distance. Since semivariance in the field or in an image is computed for discrete lag values, a continuous mathematical curve C is fitted to the scatter plot of lag versus semivariance in order to analyze spatial patterns. A spherical model (Isaaks and Srivastava 1989) is most commonly used for the fit and has three properties useful for understanding spatial correlations and deriving sampling strategies. When a spherical model is used, the graph of C increases with h until it reaches a plateau P at the lag S. P (referred to as the *sill* of the semivariogram) estimates the true variance of the data and S (called the *range*) is the lag distance at which values become uncorrelated and represents the typical size of objects in the scene. The value $\lim_{h \to 0} \gamma(0) = c0$ is called the nugget and its deviation from 0 may be due to fine scale or subpixel variance, measurement error or fitting with an incorrect model (Isaacs and Srivastava 1989).

Tian et al. (2002b) devised sampling strategies for validating MODIS LAI products by decomposition of semivariograms into hierarchical components (e.g., semivariances of forest, stands and trees) to reveal the spatial pattern of different characteristic scales within the scene. Other applications of the semivariogram to ground validation and calibration can be found in (Curran 1988).

Atkinson et al. (2000) used semivariograms to determine two scales of variations for biophysical properties of mean tree diameter at breast height, mean diameter at first leafing branch and tree density in a tropical forest of Cameroon Africa. Two sampling strategies were investigated in the analysis. One strategy maintained 1 ha subplots, while the other strategy averaged them to obtain a larger nugget area. Large nugget values relative to the sill for subplot semivariograms suggested variation at the 1 ha scale. The sill was reached at 20-25 km for all semivariograms.

From the analysis, the authors concluded that AVHRR 1 km data could capture the large-scale variance.

Atkinson et al. (2000) also evaluated within-pixel sampling strategies based on ordinary block kriging that is a best linear unbiased estimation (BLUE) method that uses the semivariogram to approximate mean values from a set of sample points. Block kriging determines the weights in the weighted average

$$\overline{V}(P) = \sum_{i=1}^{k} w_i V(x_i) \tag{8.8}$$

where $V(x_i)$ is the parameter value at point x_i and each x_i is in the ground pixel P whose mean value $\overline{V}$ is being approximated. To insure an unbiased estimate (i.e., mean error of the estimator is 0), weights w_i are constrained to sum to one. The weights are also constrained to minimize the variance σ_k of the errors. This variance is referred to as the block kriging or estimation variance. Under the BLUE constraints the estimation variance is

$$\sigma_k = \sum wi\bar{\gamma}(xi,P) + \mu - \bar{\gamma}(P,P) \tag{8.9}$$

where μ is the Lagrange parameter, $\bar{\gamma}(x_i,P)$ is the integral semi-variance between the pixel P and x_i, and $\bar{\gamma}(P,P)$ is the within pixel variance (for more details, *see* Atkinson et al. 2000, Isaaks and Srivastava 1989). Note that σ_k does not depend on the particular values of $V(x_i)$, but only on the semivariogram and the spatial pattern of x_i's. Hence, given the semivariogram, the kriging variance can be computed for any sampling strategy. Burgess et al. (1981) used block kriging to show that systematic sampling is more efficient than random sampling. Other researchers have used it to determine sample sizes needed for a specified precision (Webster et al. 1989).

Atkinson et al. (2000) examined two sampling strategies for scaling up 1 ha subplots for comparison with 1km AVHRR pixels. The subplots were arranged in an equilateral triangular pattern with each triangle constituting a plot. Although the estimation variances derived from representing each plot by the average of the three subplots were approximately 3 times less than those calculated as if a single subplot value was used to represent the triangular plot, the regression precisions obtained with the two sampling strategies were very close. They differed by factors of 1.02 (basal area), 1.03 (biomass) and 1.03 (tree density). That is, a sampling strategy based on intensive fieldwork may yield considerably better estimates of ground parameters than a less costly one without substantially improving the precision of the regression model to predict values in unknown locations. Regression modeling and prediction are, of course, the ultimate goals.

In conclusion, sampling and validation strategies based on semivariance and kriging can prove the sufficiency of smaller, more easily obtainable sample sizes.

However, even these sizes can be unnecessarily large when the purpose of the data (i.e., to develop regression equations) is ignored.

8.6 CONCLUSIONS

Biophysical variables (structural and functional) needed to track global environmental change must be collected at large scales that require the use of remotely sensed data. However, field studies are still necessary to relate remotely sensed parameters to their landscape counterparts. Consideration of the scaling problems inherent in extrapolating from the field to the image is critical to the use of remote sensing as a tool. An awareness, at the very least, of the sources of variance within a reflectance observation is important, as the structure of the canopy (leaf area, presence of senescent material, etc.) and landscape (canopy closure, background, etc.) will strongly influence the reflectance signal, and hence the biophysical interpretation.

When remote sensing algorithms are nonlinear (e.g., computing LAI from NDVI, radiative transfer inversions), errors resulting from scaling from high to low resolution are mainly due to increased mixing of ground components in the larger pixel. Maps of the landscape based on spectral mixture analysis, classification or other methods are useful in measuring and correcting for this error. However, classification accuracy is sensitive to pixel size relative to the size of objects in the scene since relative pixel size impacts within-class variance and the level of classes (i.e., tree stands versus forests) that can be mapped. Pixel size also impacts selection of endmembers for SMA when pure image endmembers are being sought, since pixel heterogeneity increases with pixel size. However, methods for constructing endmembers from the variance structure of the data are promising techniques for retrieving endmembers from the image when no pure pixels reside in the imagery. Scaling up parameter values from the ground to the image requires calibration and validation. Spatial statistics provides tools (local variance, semivariogram and kriging) for determining adequate distances between sample locations and testing efficiency versus accuracy trade-offs for various sampling strategies.

Moving from high resolution (e.g., TM) to coarse resolution imagery (e.g., AVHRR, MODIS) has scaling challenges, but it seems that, with adequate measures of surface heterogeneity through such methods as spectral mixture analysis and land cover classification, the problems are not insurmountable. The generality needed to gain perspective of the large-scale properties of ecological phenomena is attainable through remote sensing, yet we must understand the tool well enough to accurately accomplish the scaling operations we need. In concert with this, an active and reciprocal connection between remote sensing, ecological field studies, and scaling theory is important to guide scaling efforts and allow for the “surprises” which deepen our insights into the general behaviors of ecological systems.

ACKNOWLEDGEMENTS

We thank Jianguo Wu for organizing this workshop and book. We also thank NASA for funding our projects and others in order to explore large-scale ecology and ways to get there.

REFERENCES

Adams, J. B., and J. Adams. 1984. Geological mapping using Landsat MSS and TM images: removing vegetation by modeling spectral mixtures. 3rd Thematic Conference: "Remote Sensing for Exploration Geology" ERIM 2:615-622.

Adams, J. B., J. O. Smith, and P. E. Johnson. 1986. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander I site. Journal of Geophysical Research 91:8098-8112.

Aman, A., H. P. Randriamanantena, A. Podaire, and R. Frouin. 1992. Upscale integration of normalized difference vegetation index: the problem of heterogeneity. IEEE Transactions in Geoscience and Remote Sensing 30:326-338.

Asner, G. P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sensing of Environment 64:234-253.

Asner, G. P. 2000. Contributions of multi-view angle remote sensing to land-surface and biogeochemical research. Remote Sensing Reviews 18:137-162.

Asner, G. P., and D. B. Lobell. 2000. A biophysical approach for automated SWIR unmixing of soils and vegetation. Remote Sensing of Environment 71:99-112.

Asner, G. P., and C. A. Wessman. 1997. Scaling PAR absorption from the leaf to landscape level in spatially heterogeneous ecosystems. Ecological Modelling 103:81-97.

Asner, G. P., C. A. Bateson, J. L. Privette, N. E. Saleous, and C. A. Wessman. 1998a. Estimating vegetation structural effects on carbon uptake using satellite data fusion and inverse modeling. Journal of Geophysical Research 103:28839-28853.

Asner, G. P., C. A. Wessman, and S. Archer. 1998b. Scale dependence of absorption of photosynthetically active radiation in terrestrial ecosystems. Ecological Applications 8:1003-1021.

Asner, G. P., C. A. Wessman, C. A. Bateson, and J. L. Privette. 1998c. Impact of tissue, canopy, and landscape factors on the hyperspectral reflectance variability of arid ecosystems. Remote Sensing of Environment 74:69-84.

Asner, G. P., C. A. Wessman, and J. L. Privette. 1997. Unmixing the directional reflectances of AVHRR sub-pixel landcovers. IEEE Transactions in Geoscience and Remote Sensing 35:868-878.

Atkinson, P. M., G. M. Foody, P. J. Curran, and D. S. Boyd. 2000. Assessing the ground data requirements for regional scale remote sensing of tropical forest biophysical properties. International Journal of Remote Sensing 21:2571-2587.

Bateson, C. A., and B. Curtiss. 1996. A manual endmember selection method and spectral unmixing. Remote Sensing of Environment 55:229-243.

Bateson, C. A., G. P. Asner, and C. A. Wessman. 2000. Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis. IEEE Transactions in Geoscience and Remote Sensing 38:1083-1094.

Boardman, J. 1993. Automated spectral unmixing of AVIRIS data using convex geometry concepts. Pages 11-14 *in* Proceedings of the 4th Annual JPL Airborne Geoscience Workshop, Jet Propulsion Laboratory, Pasadena, CA.

Braswell, B. H., D. S. Schimel, J. L. Privette, B. Moore III, W. J. Emery, E. W. Sulzman, and A. T. Hudak. 1996. Extracting ecological and biophysical information from AVHRR optical data: an integrated algorithm based on inverse modeling. Journal of Geophysical Research 101:23335-23348.

Burgess, T.M., Webster, R., and McBratney. 1981. Optimal interpolation and isarithmic mapping of soil properties. IV. Sampling strategy. Journal of Soil Science 32:643-659.

Chen, J. M. 1999. Spatial scaling of a remotely sensed surface parameter by contexture. Remote Sensing of Environment 69:30-42.

Chen, J. M., and J. Cihlar. 1996. Retrieving leaf area index of a boreal conifer forest using Landsat TM images. Remote Sensing of Environment 55:153-162.

Chopping, M. J., A. Rango, K. M. Havstad, F. R. Schiebe, J. C. Ritchie, T. J. Schmugge, A. N. French, L. H. Su, L. McKee, and M. R. Davis. 2003. Canopy attributes of desert grassland and transition communities derived from multiangular airborne imagery. Remote Sensing of Environment 85:339-354.

Curran, P. J. 1988. The semivariogram in remote sensing: an introduction. Remote Sensing of Environment 37:493-507.

Field, C. B., J. T. Randerson, and C. M. Malmstrom. 1995. Global net primary production: combining ecology and remote sensing. Remote Sensing of Environment 51:74-88.

Friedl, M. A., D. K. Melver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E. Woodcock, G. S., S. A., C. A., A. Baccini, F. Gao, and C. Schaaf. 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83:287-302.

Gao, B. C., A. F. H. Goetz, and K. B. Heidebrecht. 1993. Derivation of scaled surface reflection from AVIRIS data. Remote Sensing of Environment 44:165-178.

Harte, J. 2002. Toward a synthesis of the Newtonian and Darwinian worldviews. Physics Today 10:29.

Hay, G. J., D. J. Marceau, P. Dubê, and A. Bouchard. 2001. A multiscale framework for landscape analysis: object-specific analysis and upscaling. Landscape Ecology 16:471-490.

Hay, G. J., K. O. Niemann, and D. G. Goodenough. 1997. Spatial thresholds, image-objects, and upscaling: a multiscale evaluation. Remote Sensing of Environment 62:1-19.

Hlavka, C. A., and J. L. Dungan. 2002. Areal estimates of fragmented land cover: effects of pixel size and model-based corrections. International Journal of Remote Sensing 23:711-724.

Hu, Z., and S. Islam. 1997. A framework for analyzing and designing scale invariant remote sensing algorithms. IEEE Transactions in Geoscience and Remote Sensing 35:747-754.

Isaaks, E. H., and R. M. Srivastava. 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.

Jin, Y., F. Gao, C. B. Schaaf, X. Li, A. H. Strahler, C. J. Bruegge, and J. V. Martonchik. 2002. Improving MODIS surface BRDF/albedo retrieval with MISR multiangle observations. IEEE Transactions in Geoscience and Remote Sensing 40:1593-1604.

Kruse, F. A. and J.L. Dwyer. 1993. The effects of AVIRIS atmospheric calibration methodology on identification and quantification mapping of surface mineralogy, Drums Mountain, Utah. Pages 101-104 *in* Proceedings of the 4th Annual JPL Airborne Geoscience Workshop, Jet Propulsion Laboratory, Pasadena, CA.

Kuusk, A. 1995. A fast, invertible canopy reflectance model. Remote Sensing of Environment 51:342-350.

Lawton, J. 2001. Earth system science. Science 292:1965.

Li, X., and A. H. Strahler. 1992. Geometric-optical directional reflectance modeling of the discrete crown vegetation canopy: effect of crown shape and mutual shadowing. IEEE Transactions in Geoscience and Remote Sensing 30:276-292.

Liang, S., and A. H. Strahler. 1993. An analytical BRDF model of canopy radiative transfer and its inversion. IEEE Transactions in Geoscience and Remote Sensing 31:1081-1095.

Lotsch, A., Y. Tian, M. A. Friedl, and R. B. Myneni. 2003. Land cover mapping in support of LAI and FPAR retrievals from EOS-MODIS and MISR: classification methods and sensitivities to errors. International Journal of Remote Sensing 24:1997-2016.

Marceau, D. J., and G. J. Hay. 1999. Remote sensing contributions to the scale issue. Canadian Journal of Remote Sensing 25:357-366.

Martin, M. E., and J. D. Aber. 1997. High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. Ecological Applications 7:431-443.

Myneni, R. B., and G. Asrar. 1993. Radiative transfer in three-dimensional atmosphere-vegetation media. Journal of Quantitative Spectroscopy and Radiative Transfer 49:585-598.

Myneni, R. B., J. Ross, and G. Asrar. 1989. A review on the theory of photon transport in leaf canopies. Agricultural Forest Meteorology 45:1-153.

Peterson, G. D. 2000. Scaling ecological dynamics: self-organization, hierarchical structure, and ecological resilience. Climatic Change 44:291-309.

Prince, S. D. 1999. What practical information about land-surface function can be determined by remote sensing? Where do we stand? *in* J. D. Tenhunen and P. Kabat, editors. Integrating Hydrology, Ecosystem Dynamics, and Biogeochemistry in Complex Landscapes. Wiley, Chichester, UK.

Privette, J. L., R. B. Myneni, C. J. Tucker, and W. J. Emery. 1994. Invertibility of a 1-D discrete ordinates canopy reflectance model. Remote Sensing of Environment 48:89-105.

Qi, J., Y. H. Kerr, M. S. Moran, M. Weltz, A. R. Huete, S. Sorooshian, and R. Bryant. 2000. Leaf area index estimates using remotely sensed data and BRDF models in a semiarid region. Remote Sensing of Environment 73:18-30.

Qu, Z., A. F. H. Goetz, and K. B. Heidebrecht. 2000. High-accuracy atmospheric correction for hyperspectral data (HATCH). Pages 373-380 *in* Proceedings of the 9th Annual JPL Airborne Geoscience Workshop, Jet Propulsion Laboratory, Pasadena, CA.

Ross, J. K. 1981. The Radiation Regime and Architecture of Plant Stands. Kluwer, Boston, MA.

Running, S. W., T. R. Loveland, and L. L. Pierce. 1994. A vegetation classification logic based on remote sensign for use in global biogeochemical models. Ambio 23:77-89.

Sabol, D. E., A. R. Gillespie, J. B. Adams, M. O. Smith, and C. J. Tucker. 2002. Structural stage in Pacific Northwest forests estimated using simple mixing models of multispectral images. Remote Sensing of Environment 80:1-16.

Sellers, P. J. 1987. Canopy reflectance, photosynthesis, and transpiration. II. The role of biophysics in the linearity of their interdependence. International Journal of Remote Sensing 6:1335-1372.

Smith, M. L., S. V. Ollinger, M. E. Martin, J. D. Aber, R. A. Hallett, and C. L. Goodale. 2002. Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. Ecological Applications 12:1286-1302.

Smith, M. O., S. L. Ustin, J. B. Adams, and A. R. Gillespie. 1990. Vegetation in desert I. A regional measure of abundance from multispectral images. Remote Sensing of Environment 31:1-26.

Tian, Y., Y. Wang, Y. Zhang, Y. Knyazikhin, J. Bogaert, and R. B. Myneni. 2003. Radiative transfer based scaling of LAI retrievals from reflectance data of different resolutions. Remote Sensing of Environment 84:143-159.

Tian, Y., C. E. Woodcock, Y. Wang, J. L. Privette, N. V. Shabanov, L. Zhou, Y. Zhang, W. Buermann, J. Dong, B. Veikkanen, T. Hame, K. Andersson, M. Ozdogan, Y. Knyazikhin, and R. B. Myneni. 2002a. Multiscale analysis and validation of the MODIS LAI product I. Uncertainty assessment. Remote Sensing of Environment 83:414-430.

Tian, Y., C. E. Woodcock, Y. Wang, J. L. Privette, N. V. Shabanov, L. Zhou, Y. Zhang, W. Buermann, J. Dong, B. Veikkanen, T. Hame, K. Andersson, M. Ozdogan, Y. Knyazikhin, and R. B. Myneni. 2002b. Multiscale analysis and validation of the MODIS LAI product. II. Sampling strategy. Remote Sensing of Environment 83:431-441.

van Leeuwen, W. J. D., A. R. Huete, C. L. Walthall, S. D. Prince, A. Begue, and J. L. Roujean. 1997. Deconvolution of remotely sensed spectral mixtures for retrieval of LAI, fAPAR and soil brightness. Journal of Hydrology 188-189:697-724.

Villeneuve, P. V., S. A. Gerstl, and G. P. Asner. 1998. Estimating nonlinear mixing effects for arid vegetation scenes with MISR channels and observation directions. Proceedings of the International Geoscience and Remote Sensing Symposium 3:31-35.

Webster, R., P. J. Curran, and J. W. Munden. 1989. Spatial correlation in reflected radiation from the ground and its implications for sampling and mapping by ground based radiometry. Remote Sensing of Environment 29:67-78.

Wessman, C. A. 1992. Spatial scales and global change: bridging the gap from plots to GCM grid cells. Annual Review of Ecology and Systematics 23:175-200.

Wessman, C. A., and G. P. Asner. 1998. Ecosystems and the problems of measurement at large spatial scales. Pages 346-371 *in* P. M. Groffman and M. L. Pace, editors. Successes, Limitations, and Frontiers in Ecosystem Ecology. Springer-Verlag, New York.

Wessman, C. A., J. D. Aber, D. L. Peterson, and J. M. Melillo. 1988. Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. Nature 335:154-156.

Wessman, C. A., C. A. Bateson, and T. L. Benning. 1997. Detecting fire and grazing pattern in tallgrass prairie using spectral mixture analysis. Ecological Applications 7:493-511.

Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.

Zagolski, F., V. Pinel, J. Romier, D. Alcayde, J. P. Gastellu-Etcheforry, G. Giordano, G. Marty, and E. Mougin. 1996. Forest canopy chemistry with high spectral resolution remote sensing. International Journal of Remote Sensing 17:1107-1128.