Taylor & Francis
Taylor & Francis Group

Scaling up: linking field data and remote sensing with a hierarchical model

Adam M. Wilson[a]*, John A. Silander[a], Alan Gelfand[b] and Jonathan H. Glenn[a]

*[a]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA;
[b]Department of Decision Sciences, Duke University, NC, USA*

Ecologists often seek to understand patterns and processes across multiple spatial and temporal scales ranging from centimeters to hundreds of meters and from seconds to years. Hierarchical statistical models offer a framework for sampling design and analysis that can be used to incorporate the information collected at finer scales while allowing comparison at coarser scales. In this study we use a Hierarchical Bayesian model to assess the relationship between measurements collected on the ground at centimeter scales nested within 2×3 m quadrats, which are in turn nested within much larger (0.1–12 ha) plots. We compare these measurements with the Normalized Difference Vegetation Index (NDVI) derived from radiometrically and geometrically corrected 30-m resolution LANDSAT ETM+ data to assess the NDVI–Biomass relationship in the Cape Floristic Region of South Africa. Our novel modeling approach allows the data observed at submeter scales to be incorporated directly into the model and thus all the data (and variability) collected at finer scales are represented in the estimates of biomass at the LANDSAT scale. The model reveals that there is a strong correlation between NDVI and biomass, which supports the use of NDVI in spatiotemporal analysis of vegetation dynamics in Mediterranean shrubland ecosystems. The methods developed here can be easily generalized to other ecosystems and ecophysiological parameters.

Keywords: biomass; remote sensing; ground truthing; scale; Hierarchical Bayes

1. Introduction

Ecologists often seek to understand patterns and processes across a broad range of spatial scales (Levin 1992) including biomass, canopy structure, species composition, and carbon flux (Nightingale *et al.* 2004). The challenges of accomplishing this are not only collecting data at a range of spatial scales from disparate sources but relating these data analytically in a coherent manner while quantifying the uncertainty (Cressie *et al.* 2009). Satellite remote sensing, for example, offers a tantalizing quantity of data that are useful in the study of vegetation dynamics across space and through time (e.g., Kerr and Ostrovsky 2003, Murwira 2005, Murwira and Skidmore 2006). Analysis of these data can provide insight into phenology, succession, fire, and other biophysical properties such as primary productivity and biomass in ecosystems (e.g., Song and Woodcock 2003, Hoare and Frost

*Corresponding author. Email: adam.wilson@uconn.edu

2004, Diaz-DelGado *et al.* 2003, Cohen *et al.* 2006). See Lu (2006) for a particularly thorough review of various methods of estimating above-ground biomass with remotely sensed data. In that review, Lu recognized the challenges of relating relatively coarse remotely sensed data with finer ground measurements and reported that a 'synthetic analysis of multiscale data with a combination of different modelling approaches may be needed' for accurate biomass estimation.

There is often a disconnect between the scale of data collection for 'ground-truthing' exercises (often centimeters to meters) and the resolution of satellite data (tens to hundreds of meters) (Chambers *et al.* 2007). Traditional statistics (i.e., OLS regression analysis) are ill suited for data collected at different scales and so require some sort of aggregation to allow comparison with coarse resolution data (Gotway and Young 2002). Summary statistics (such as the mean) or qualitative comparisons are often used for this purpose. For example, Milich and Weiss (2000) 'ground-truthed' data from the Advanced Very High Resolution Radiometer (AVHRR) by navigating to the center of each AVHRR pixel along several transects in the Sahel along the southern boundary of the Sahara Desert. Video images and qualitative descriptions of the vegetation at these centroids were considered useful indicators of the entire 58 km^2 pixel. Samimi and Kraus (2004) used average biomass measured at several points within 120×120 m plots to compare with several remotely sensed spectral indices. Huang *et al.* (2009) used a series of separate regressions to scale up from field observations to AVIRIS pixels (3.6 m resolution) and from AVIRIS pixels to LANDSAT (30 m) pixels. Baccini *et al.* (2007) recently considered two options for scaling up from 0.6 ha plots to 1 km^2 MODIS data: (1) landscape stratification and (2) averaging of fine spatial resolution (LANDSAT) maps. However, using an average value as if it were an observed and representative value will likely lead to an overestimate of the certainty in the results by disregarding the variability observed at finer scales (Gotway and Young 2002). We present a third option that does not require aggregation of the data collected at finer scales before comparison with coarser resolution satellite data. Hierarchical Bayesian (HB) modeling offers a framework to explicitly incorporate data collected at different scales without losing information by aggregation (Agarwal *et al.* 2005). This is accomplished with the introduction of *latent variables* (Clark and Gelfand 2006b). These variables represent unobserved (and practically unobservable) quantities (such as the biomass of a 30 m × 30 m LANDSAT pixel) and are estimated using data that were observed at different scales. In this study we provide an example of a HB statistical model that integrates data collected at several scales (from centimeters to hundreds of meters) and different types (from satellite imagery to biomass measurements). This allows the direct fusion of remotely sensed data with finer measurements of biomass and the uncertainty present at each spatial scale is passed up to the estimates at coarser scales.

1.1. *Study region*

In this study we illustrate a hierarchical approach to integrating data collected at different scales by bringing together satellite and field measurements in a Mediterranean shrubland ecosystem. The Cape Floristic Region (CFR) of South Africa (Figure 1) is an internationally recognized hotspot of floral biodiversity and is home to approximately 9000 plant species, 69% of which are endemic (Goldblatt and Manning 2000). The CFR experiences a Mediterranean climate with hot, dry summers and cool, wet winters in the western half that transitions to more even precipitation seasonality in the east. Mean annual rainfall ranges from 60 mm to over 3000 mm (Schulze 2007). There is evidence that short-interval fires have become more frequent over the past few decades in some areas of the CFR
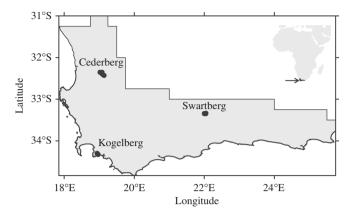
Figure 1. Map illustrating the location of the Cape Floristic Region of South Africa and the three regions visited for this study. The plots are shown as black polygons and slightly enlarged to be visible at this scale.

(Forsyth and van Wilgen 2008) and that fire probability in the region increases in hotter, drier conditions (Wilson *et al.* 2010). Thus, it is likely that warmer, drier conditions in the future will lead to an enhanced fire regime. However, fire probability in Mediterranean ecosystems is also dependent on biomass accumulation rates and their sensitivity to climate change (Mouillot *et al.* 2002). To make predictions about the impact of climate change on the fire regime, we also need to understand how biomass accumulation rates may change in the future. Unfortunately, collection of fuel load/biomass data is destructive and time-consuming and therefore limited to relatively small areas. The Normalized Difference Vegetation Index (NDVI) has been used as a proxy for many vegetation characteristics including biomass and burned area detection (e.g., Gerber 2000, Diaz-Delgado *et al.* 2002), but the challenges of calibrating it to a specific ecosystem are not trivial (Song and Woodcock 2003).

2. Methods

2.1. Data

2.1.1. Field data

The fynbos landscape is highly heterogeneous due largely to the regular disturbance of wildfire (Cowling and Lombard 2002). Areas of equal age (time since fire) and similar species composition range in size from less than a hectare to several hectares or more. We used a hierarchical sampling scheme with large (0.1–12 ha) plots with 5–10 (depending on plot size) randomly selected 2 m $\times$ 3 m (6 m^2) quadrats, each with a grid of 10 subquadrat points (Figure 2). Plot size was determined by the size of the homogeneous area. The plots were selected by exploring the region and identifying areas with relatively homogeneous vegetation structure, community composition, and time since fire. To capture the range of possible biomass values, we selected plots that ranged from virtually zero biomass (recently burned) to over 10 kg m^{-2}. The plots were mapped with a GPS (Trimble GeoExplorer 3) and the resulting data were differentially corrected.

We then navigated to the randomly located quadrats and collected hyperspectral reflectance from 380 to 1000 nm using a portable spectrometer (OceanOptics USB4000, OceanOptics, Inc., Dunedin, FL, USA) with a 2-m fiber-optic cord mounted to a pole
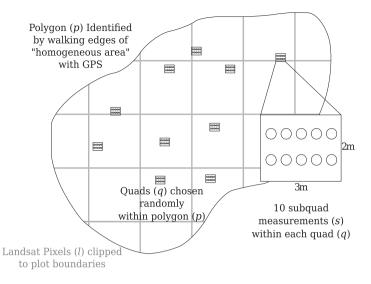
Figure 2. Schematic of the hierarchical sampling scheme. Ten subquadrat measurements (*s*) were nested within 2 m × 3 m quadrats (*q*), which were nested within 0.1–12 ha plots (*p*). Plots were selected to be the same age (time since fire) and relatively homogeneous in community composition. LANDSAT 7 ETM+ data with 30 m × 30 m resolution (shown in figure as a gray grid) were extracted for each plot. We collected data from 16 plots in three regions (see Figure 1). For clarity, schematic is not drawn to scale.

and held 1 m from the canopy. The fiber has a 25° field of view that resulted in each measurement viewing a circle of radius 0.22 m. We also measured the quantity of photosynthetically active radiation intercepted by the canopy (iPAR) with an AccuPAR ceptometer (Decagon Devices, Inc., Pullman, WA, USA). Ambient light was measured above the canopy and the percent light intercepted was calculated after measuring the light penetrating to the ground through the canopy. Both the iPAR and the reflectance were measured at each of the 10 subquadrat points described above. The iPAR was collected as a proxy for biomass, with lower light penetration (higher iPAR) in quadrats with high biomass. The linear relationship between ln(biomass) and iPAR is estimated within the model. We also estimated the percent cover of four plant functional types (PFTs) at the quadrat scale: (1) restio/graminoid, (2) ericoid/other woody sub-shrubs, (3) proteoid shrubs, and (4) forbs following van Wilgen (1982) using the Braun-Blanquet scale (Mueller-Dombois and Ellenberg 2002). Aboveground biomass was measured within 1 or 2 quadrats for each plot by cutting a proportion of the total standing biomass and grouping the vegetation PFTs. Biomass samples were oven dried (∼70°C) for at least 14 hours for herbaceous samples or 24 hours for shrubs and until weight change was insignificant with further drying (up to 60 hours). Total aboveground dry biomass from the quadrat for each PFT was calculated accounting for the moisture content of the vegetation as follows (where *B* is biomass in g/m^2):

$$B_{\text{total}}^{\text{dry}} = B_{\text{total}}^{\text{wet}} * \frac{B_{\text{sample}}^{\text{dry}}}{B_{\text{sample}}^{\text{wet}}} \qquad (1)$$

Table 1. Region, coordinates, number of plots, and dates of field work and satellite imagery used in the analysis. All dates are from 2008.

Region	Lat	Lon	Number of plots	Field Data	LANDSAT Image
Swartberg	−33.3	22.0	5	June 8–13	June 5
Cedarberg	−32.4	19.1	6	June 19–26	July 21
Kogelberg	−34.3	19.0	5	July 9–12	July 21

Our biomass estimates are similar to van Wilgen's (1982) estimates of 670 g/m^2 (4-year-old stand), 5100 g/m^2 (21-year-old stand), and 7600 g/m^2 (37-year-old stand).
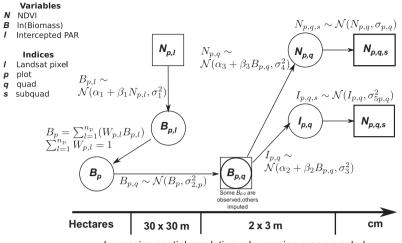
2.1.2. Satellite data

The field data were compared with 30 m resolution satellite data from the LANDSAT 7 ETM+ sensor. We acquired LANDSAT images that were as close as possible to the dates of the field sampling from the USGS Global Visualization Viewer (http://glovis.usgs.gov/)(Table 1). We used the LANDSAT Level 1T (L1T) data that have systematic radiometric and geometric correction by incorporating ground control points and a Digital Elevation Model (DEM) for topographic accuracy. In our region of interest, the scenes generally had over 100 ground control points available. LANDSAT has a 16-day interval between images and the region is often cloudy, which led to a maximum difference of 12 days between the field work and the image. LANDSAT images were converted to top of atmosphere reflectance and topographically corrected using the calibration values contained in the metadata and GRASS GIS processing routines (GRASS Development Team 2008). NDVI was calculated as NDVI $= \frac{IR-Red}{IR+Red}$. Because our sampling dates occurred after the scan line corrector (SLC) on LANDSAT 7 failed in 2003, there were gaps in the data (Howard and Lacasse 2004). Plots were selected to be near the scene centers to minimize the effect of data loss on this analysis. The LANDSAT pixels contained by the plot boundaries were extracted and given an areal weight according to the percentage of the pixel that was in the plot (i.e., pixels with less area in the plot were down-weighted). The relatively large size of the plots reduces the error due to uncertainty in the exact location of the pixels as we compare all pixels in the plot to all group measurements from the plot rather than a single ground measurement to a single (potentially misaligned) pixel.

2.2. Modeling

We constructed a HB statistical model that links the data collected in the field to the scale of the LANDSAT data (Figure 3). This approach allows the incorporation of diverse sources of information, can account for unknown (or unknowable) influences, and can use information from large numbers of latent variables (Clark 2005, Gelfand *et al.* 2005, Latimer *et al.* 2006, Clark and Gelfand 2006b). The subquadrat measurements (*s*) are grouped into quadrats (*q*), which are contained by plots (*p*). These plots, which were identified on the ground *a priori* to be homogeneous, thus contain multiple samples of both field and satellite (LANDSAT) data (*l*).

The measurements of iPAR and NDVI at the smallest scale (subquadrat) are treated as random draws from an unobserved quadrat scale NDVI and iPAR. These quadrat scale values are considered random variables with a full probability distribution (we used a normal distribution, but any distribution could be substituted). These variables are then

Figure 3. Directed acyclic graph (Clark and Gelfand 2006a) of the HB model structure used to scale up biomass measurements to compare with LANDSAT-NDVI data. The variables are indicated by capital letters: N = NDVI, B = ln(biomass), and I = intercepted PAR. The indices denote the following measures of scale: l = LANDSAT, p = plot, q = quadrat, s = subquadrat. The boxes and circles represent nodes in the model, which are either observed data (boxes) or unobserved, latent variables (circles) that are estimated during model fitting. The arrows represent the links between the various nodes and the directions of the arrows illustrate the conditional relationships. The text defines the relationship between the nodes, with linear regressions and nesting of samples within parent distributions. The spatial resolution of each node is coarsest at the plot level (on the left) and increases to the right to the subquadrat measurements. Intercepted photosynthetically active radiation (PAR) was used as a proxy for biomass in some quadrats.

used in a regression with the observed (harvested) biomass measurements to estimate the biomass at the unharvested quadrats. Thus, for the quadrats where we did not directly measure biomass, we have a full probability distribution of the biomass that incorporates the variability in the subquadrat NDVI and iPAR measurements and the uncertainty of the regression parameters fit using the other, observed quadrats. These quadrat-level biomass estimates are in turn considered observations of a larger, unobserved plot-level biomass. So the uncertainty inherent in the quadrat estimates is included in the biomass estimate for the entire plot. Rather than simply average the subquadrat measurements, and then average again to get a plot-level value, we consider all unobserved variables to be inexact and therefore have some quantifiable uncertainty associated with them. We can take advantage of the nested scales (i.e., points within quads within plots) in our study to use information collected at finer scales to infer biomass estimates at the coarser scales. The model is essentially a series of regressions and nesting, with the important distinction of passing the uncertainty at finer scales to the coarser scales. So we can treat biomass at the plot level as a random variable from which we have drawn samples (the quadrat and LANDSAT data).

The model structure can be written as follows:

$$B_{p,l} \sim \mathcal{N}(\alpha_1 + \beta_1 N_{p,l}, \sigma_1^2) \qquad (2)$$

$$B_p = \sum_{l=1}^{n_p} \left(W_{p,l} B_{p,l} \right) \qquad (3)$$

$$\sum_{l=1}^{n_p} W_{p,l} = 1 \tag{4}$$

$$B_{p,q} \sim \mathcal{N}(B_p, \sigma_{2,p}^2) \tag{5}$$

$$I_{p,q} \sim \mathcal{N}(\alpha_2 + \beta_2 B_{p,q}, \sigma_3^2) \tag{6}$$

$$N_{p,q} \sim \mathcal{N}(\alpha_3 + \beta_3 B_{p,q}, \sigma_4^2) \tag{7}$$

$$I_{p,q,s} \sim \mathcal{N}(I_{p,q}, \sigma_{5,p,q}^2) \tag{8}$$

$$N_{p,q,s} \sim \mathcal{N}(N_{p,q}, \sigma_{6,p,q}^2) \tag{9}$$

where B is ln(biomass), I is intercepted PAR, N is NDVI, $W_{p,l}$ represents the areal weight of each pixel (l) in plot p (which is constrained to sum to 1), and n_p is equal to the number of LANDSAT pixels in each plot. Because the measured biomass values ranged from 1 to well over 10,000 g/m^2, we used a log transform to better fit the data. The model integrates the data collected at finer scales to estimate biomass at the LANDSAT (30 m) scale. This allowed the regression of LANDSAT-NDVI values against biomass estimates in Equation (2). The next stage in the model estimated biomass for each of the LANDSAT pixels within each plot. This was accomplished by setting the overall plot biomass to be an areal weighted average of the biomass of each LANSDAT pixel within it, using Equations (3) and (4). The advantage of using relatively large plots was that we reduce the uncertainty from georeferencing errors. Instead of comparing one small field plot to one large pixel (as is typically done in ground- truthing excercises), we compared several field plots that each contain many pixels from an area of known homogeneity. The sampling design was constructed so the quadrat-level biomass estimates were random samples from the larger plot scale (see Equation (5)). As explained above, biomass estimates for some of the quadrats were directly measured, but others need to be estimated using the NDVI and iPAR data collected at all quadrats. To accomplish this, we conducted two regressions between the quadrat-level biomass and ground NDVI ($N_{p,q,s}$) and iPAR ($I_{p,q,s}$) in Equations (6) and (7). But it was not possible to measure NDVI and iPAR directly at the quadrat level. So these values were estimated using the subquadrat (s) iPAR and NDVI measurements (which are independent samples drawn from each quadrat) in Equations (8) and (9).

The model was specified and fit using OpenBUGS (Thomas *et al.* 2006) software. Convergence of the Markov chain Monte Carlo (MCMC) chains were assessed using the Gelman–Rubin convergence statistic (Gelman and Rubin 1992) and visual inspection of the chains. The model was run for 25,000 iterations, the first 5000 of which were discarded and the remaining samples were thinned by 100 to reduce autocorrelation. Due to the hierarchical nature of the model, it is not possible to calculate traditional fit metrics such as the R^2.

3. Results

The iPAR was useful as a proxy for biomass in the unharvested quadrats (Figure 4). The regression between the log of the observed biomass values and quadrat iPAR values reveals a strongly significant relationship (Table 2). This suggests that this measurement, with
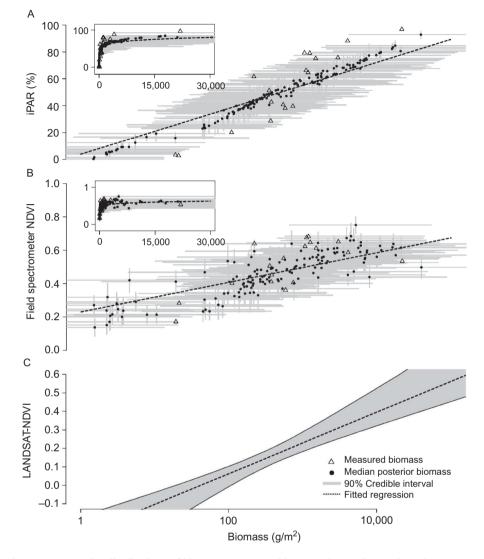
Figure 4.   Posterior distributions of biomass, NDVI, and iPAR at the quadrat scale and LANDSAT scale. Each black point represents the median posterior value for one of the 167 quadrats, with the 90% credible intervals represented with in gray. The triangles are the biomass values that were measured directly in the field. The dashed line is the fitted regression (plotted with the median posterior samples of slope and intercept). Note that the *x*-axis is log transformed. The inset plots show the same data with a linear *x*-axis for comparison. Panel (a) has biomass and iPAR at the quadrat scale, (b) biomass and NDVI as measured with the field spectrometer at the quadrat scale, and (c) biomass and LANDSAT-NDVI. Biomass was not directly observed at the LANDSAT pixel scale, so the median regression line and credible intervals are shown. See Table 2 for regression coefficients.

calibration, can be used as an efficient proxy for biomass estimation in shrublands. The relationship between biomass and NDVI at the quadrat scale was also significant, although with larger uncertainty (Table 2). At the quadrat scale, there is substantial variation in vegetation cover, with bare ground in some areas and 2 m tall shrubs in others. The increased

Table 2. Posterior estimates of regression parameters (median and 95% CIs in parenthesis).

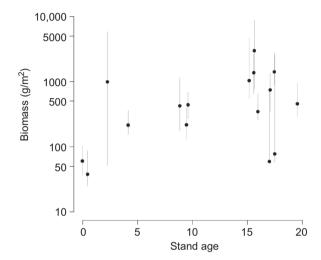| Scale	Regression	α (Intercept)	β (Slope)	σ (SD)
Quadrat	I_q vs. B_q	3.59 ($-5.43, 12.97$)	7.45 (5.86, 8.99)	10.46 (8.27, 13.1)
Quadrat	N_q vs. B_q	0.23 (0.12, 0.34)	0.04 (0.02, 0.06)	0.26 (0.24, 0.29)
LANDSAT	B_l vs. N_l	3.71 (2.89, 4.79)	13.89 (8.47, 18.53)	1.16 (0.81, 1.67)

Figure 5. Posterior biomass estimates for the LANDSAT pixels within each plot. The point is located at the median value for each plot and the gray bars show the 10–90% quantiles. Use of stand age maps as the sole predictor of biomass and fire risk could lead to inaccurate fire risk estimates.

variability is likely driven by the heterogeneity of vegetation at submeter scales and the sensitivity of NDVI to heterogeneity in the vegetation cover.

The biomass–LANDSAT-NDVI regression also revealed a significant relationship (Table 2 and Figure 4c). The uncertainty is because biomass was not directly observed at the scale of a LANDSAT pixel and thus had to be estimated from the rest of the model. More importantly, however, is that the regression parameters are well defined: $\alpha_1 = 3.71$ (95% CI: 2.89–4.79) and $\beta_1 = 13.89$ (95% CI: 8.47–18.53). When the plot-level biomass estimates (with uncertainty) are plotted against stand age, increasing biomass in the years following fire is apparent (Figure 5).

4. Discussion

This model was not constructed to test whether there is a relationship between NDVI and biomass. The existence of such a relationship has been known for decades (e.g., Pearson and Miller 1972, Tucker 1977). Instead, our objective with this model is to understand the biomass–NDVI relationship spanning a range of spatial scales by integrating the types of data that ecologists are able to collect in the field with the data provided by remote sensing. The HB model structure allows this comparison and provides full uncertainty for

the model parameters (Table 2). Using this model, we have shown that LANDSAT-NDVI is a useful proxy for biomass in the CFR and quantified the uncertainty of the relationship. The novelty of this approach is that the data observed at the quadrat scales were incorporated directly into the model and thus all the data (and variability) collected at finer scales are represented in the estimates of biomass at the LANDSAT scale. Unlike many similar approaches, our method does not require any aggregation of the small-scale data before fitting the model. This represents a departure from most ground-truthing studies that often average across and disregard the observed variability at finer scales. This modeling approach will facilitate the use of LANDSAT-NDVI data to map biomass at regional scales and allow comparison with coarser resolution satellite data. In addition, we have shown that measuring light interception (iPAR) is an efficient and accurate way to estimate biomass in shrubland ecosystems. More and larger ground observations would always be useful. But with limited time and resources, this approach allows us to use as much of the information contained in the data as possible, while quantifying the uncertainty of predictions. Our modeling approach should be broadly applicable to other systems and ecological questions. Because of the asymptotic relationship between biomass and NDVI, our specific application of the model may work better and show lower variance for shorter stature vegetation, such as grasslands and shrublands, than forested systems.

This work supports the use of vegetation indices as a proxy for biomass and provides a framework for data collection and analysis that can integrate data from disparate scales. This analysis provides evidence that remotely sensed information can be useful to monitor biomass (and fuel load) and post-fire vegetation 'recovery.' This is important in many regions to monitor fire risk across large regions where field sampling is difficult. In the CFR, the relationship between stand age and biomass (van Wilgen 1982) is often used to define fire risk. For our plots, the relationship between biomass and age is quite noisy, especially in older plots (Figure 5). This is understandable, given all the other variables that contribute to biomass accumulation rates. However, the results presented here reveal that relying solely on time since fire can be misleading because site productivity is also likely to be affected by aspect, elevation, soil moisture, and other factors. Remotely sensed data offer another source of information on fuel-load accumulation across large regions and through time.

A challenging aspect of using remotely sensed vegetation indices to monitor biomass accumulation is the saturation and increasing uncertainty of the estimates at higher NDVI and biomass levels (Figure 4c). Both NDVI and iPAR saturate at higher biomass levels. This has been observed in other systems (e.g., Steininger 2000, Lu 2005). Thus, remotely sensed biomass estimates will be most sensitive and accurate for areas and times with lower biomass. As biomass increases the estimates will become increasingly, but quantifiably, uncertain.

Calibration of remotely sensed vegetation indices to vegetation attributes opens the possibility to parameterize biomass accumulation models (with uncertainty) and to study how biomass accumulation changes in response to fire and weather. The observed relationship is likely to hold also for other, coarser resolution satellites, such as MODIS (250 m) and AVHRR (1 km) (Brown *et al.* 2006). Brown *et al.* found that NDVI records from AVHRR, SPOT-Vegetation, SeaWiFS, MODIS, and Landsat ETM+ Sensors are similar enough to allow intercomparison and the development of long-term time series. LANDSAT is very useful for high spatial resolution monitoring, but its relatively infrequent imaging (16-day intervals between images) and the potential for clouds complicate its use in ecological studies that require high temporal resolution. Post-2003 LANDSAT 7 images also

suffer from missing data due to failure of the SLC (Howard and Lacasse 2004). Thus data from the coarse-grain but more frequent sensors are also useful for spatiotemporal studies of vegetation dynamics.

5. Conclusion

Issues of scale are a central subject in ecology (e.g., Levin 1992, Leibold *et al.* 2004, Beever *et al.* 2006). This work has shown that a hierarchical sampling scheme and analytical approach is useful to integrate the data collected at different scales. Using field data from our model system, the shrublands of the CFR, we have shown that it is possible to collect data at scales of centimeters to meters and to scale up to the resolution of relatively coarse-grained satellite data. Because this is a statistical scaling approach and is not dependent on understanding the mechanistic relationships, it is not limited to comparison of NDVI and biomass. This approach can be applied to virtually any process that spans multiple scales and the model framework is flexible to variations in sampling design. The number and size of plots needed for this type of satellite validation will depend on the homogeneity of the vegetation at the scale of the satellite data and the parameters of interest.

A more thorough understanding of the strength and nature of climatic controls on vegetation dynamics is vital to predicting the ecological impacts of climate change. Remote sensing offers an abundance of relevant data, but without calibration it is difficult to interpret for specific ecosystems. Decision makers (reserve managers, conservation biologists, and policy makers) need reliable information and models to develop effective management practices. This is especially important in the context of a changing environment, as managers must make decisions based on predictions of future changes.

References

Agarwal, D., *et al*., 2005. Tropical deforestation in Madagascar: analysis using hierarchical spatially explicit, Bayesian regression models. *Ecological modelling*, 185 (1), 105–131.

Baccini, A., *et al*., 2007. Scaling field data to calibrate and validate moderate spatial resolution remote sensing models. *Photogrammetric Engineering and Remote Sensing*, 73 (8), 945.

Beever, E., Swihart, R., and Bestelmeyer, B., 2006. Linking the concept of scale to studies of biological diversity: evolving approaches and tools. *Diversity and Distributions*, 12 (3), 229–235.

Brown, M., *et al*., 2006. Evaluation of the consistency of long-term NDVI time series derived from AVHRR, SPOT-Vegetation, SeaWiFS, MODIS, and Landsat ETM+ Sensors. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (7), 1787.

Chambers, J., *et al*., 2007. Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends in Ecology & Evolution*, 22 (8), 414–423.

Clark, J., 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters*, 8 (1), 2–14.

Clark, J.S. and Gelfand, A.E., 2006a. A future for models and data in environmental science. *Trends in Ecology & evolution*, 21 (7), 375–380.

Clark, J.S. and Gelfand, A.E., 2006b. *Hierarchical modelling for the environmental sciences: statistical methods and applications*. New York: Oxford University Press.

Cohen, W.B., *et al.*, 2006. MODIS land cover and LAI collection 4 product quality across nine sites in the western hemisphere. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (7), 1843–1857.

Cowling, R. and Lombard, A., 2002. Heterogeneity, speciation/extinction history and climate: explaining regional plant diversity patterns in the Cape Floristic Region. *Diversity & Distributions*, 8 (3), 163–179.

Cressie, N., *et al.*, 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19 (3), 553–570.

Diaz-DelGado, R., Lloret, F., and Pons, X., 2003. Influence of fire severity on plant regeneration by means of remote sensing imagery. *International Journal of Remote Sensing*, 24 (8), 1751.

Diaz-Delgado, *et al.*, 2002. Satellite evidence of decreasing resilience in Mediterranean plant communities after recurrent wildfires. *Ecology*, 83 (8), 2293–2303.

Forsyth, G. and van Wilgen, B., 2008. The recent fire history of the Table Mountain National Park, and implications for fire management. *Koedoe - African Protected Area Conservation and Science*, 50 (1), 3–9.

Gelfand, A., *et al.*, 2005. Modelling species diversity through species level hierarchical modelling. *Journal of the Royal Statistical Society. Series C*, 54, 1–20.

Gelman, A. and Rubin, D., 1992. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7 (4), 457–472.

Gerber, L., 2000. Development of a ground truthing method for determination of rangeland biomass using canopy reflectance properties. *African Journal of Range and Forage Science*, 17 (1), 93–100.

Goldblatt, P. and Manning, J., 2000. *Cape plants : a conspectus of the Cape flora of South Africa*. Pretoria: National Botanical Institute.

Gotway, C. and Young, L., 2002. Combining incompatible spatial data. *Journal of the American Statistical Association*, 97 (458), 632–648.

GRASS Development Team, 2008. *Geographic Resources Analysis Support System (GRASS GIS) Software*, Open Source Geospatial Foundation, USA.

Hoare, D. and Frost, P., 2004. Phenological description of natural vegetation in southern Africa using remotely-sensed vegetation data. *Application of Vegetation Science*, 7 (1), 19–28.

Howard, S. and Lacasse, J., 2004. An evaluation of gap-filled landsat SLC-off imagery for wildland fire burn severity mapping. *Photogrammetric Engineering & Remote Sensing*, 70 (8), 877–880.

Huang, C.Y., *et al.*, 2009. Multiscale analysis of tree cover and aboveground carbon stocks in pinyon-juniper woodlands. *Ecological Applications*, 19 (3), 668–681.

Kerr, J. and Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. *Trends in Ecology & evolution*, 18 (6), 299–305.

Latimer, A., *et al.*, 2006. Building statistical models to analyze species distributions. *Ecological Applications*, 16 (1), 33–50.

Leibold, M.A., *et al.*, 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, 7 (7), 601 – 613.

Levin, S., 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur Award lecture. *Ecology*, 73 (6), 1943–1967, Landscape Ecology Week 2.

Lu, D., 2005. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *International Journal of Remote Sensing*, 26 (12), 2509–2526.

Lu, D., 2006. The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing*, 27 (7), 1297–1328.

Milich, L. and Weiss, E., 2000. GAC NDVI interannual coefficient of variation (CoV) images: ground truth sampling of the Sahel along north-south transects. *International Journal of Remote Senstive*, 21 (2), 235–260.

Mouillot, F., Rambal, S., and Joffre, R., 2002. Simulating climate change impacts on fire frequency and vegetation dynamics in a Mediterranean-type ecosystem doi:10.1046/j.1365-2486.2002.00494.x. *Global Change Biology*, 8 (5), 423–437.

Mueller-Dombois, D. and Ellenberg, H., 2002. *Aims and Methods of Vegetation Ecology*. Caldwell, NJ. Blackburn Press.

Murwira, A. and Skidmore, A., 2005. The response of elephants to the spatial heterogeneity of vegetation in a Southern African agricultural landscape. *Landscape ecology*, 20 (2), 217–234.

Murwira, A. and Skidmore, A., 2006. Monitoring change in the spatial heterogeneity of vegetation cover in an African savanna. *International Journal of Remote Sensing*, 27 (11), 2255–2269.

Nightingale, J., Phinn, S., and Held, A., 2004. Ecosystem process models at multiple scales for mapping tropical forest productivity. *Progress in Physical Geography*, 28 (2), 241.

Pearson, R.L. and Miller, L.D., 1972. *Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie*. Ann Arbor, MI: Willow Run Laboratories, Environmental Research Institute of Michigan, 1355.

Samimi, C. and Kraus, T., 2004. Biomass estimation using Landsat-TM and-ETM+. Towards a regional model for Southern Africa? *GeoJournal*, 59 (3), 177–187.

Schulze, R.E., ed., 2007. The South African Atlas of Agrohydrology and Climatology. Technical report WRC Report 1489/1/06, Water Research Commission, Pretoria, South Africa.

Song, C. and Woodcock, C.E., 2003. Monitoring forest succession with multitemporal Landsat images: factors of uncertainty. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (11), 2557–2567.

Steininger, M., 2000. Satellite estimation of tropical secondary forest above-ground biomass: data from Brazil and Bolivia. *International Journal of Remote Sensing*, 21 (6–7), 1139–1157.

Thomas, A., *et al*., 2006. Making BUGS Open. *R News*, 6 (1), 12–17.

Tucker, C., 1977. Use of Near Infrared/Red Radiance Ratios for Estimating Vegetation Biomass and Physiological Status. Technical report X-923-109, NASA Goddard Space Flight Center 41 p.

van Wilgen, B., 1982. Some effects of post-fire age on the above-ground plant biomass of Fynbos(Macchia) vegetation in South Africa. *Journal of Ecology*, 70 (1), 217–225.

Wilson, A.M., *et al*., 2010. A Hierarchical Bayesian model of wildfire in a Mediterranean biodiversity hotspot: implications of weather variability and global circulation. *Ecological Modelling*, 221, 106–112.