2013

ESTIMATION AND MODELING OF FOREST ATTRIBUTES ACROSS LARGE SPATIAL SCALES USING BIOMEBGC, HIGH-RESOLUTION IMAGERY, LIDAR DATA, AND INVENTORY DATA

Jordan Seth Golinkoff
The University of Montana

ESTIMATION AND MODELING OF FOREST ATTRIBUTES ACROSS LARGE SPATIAL SCALES USING BIOMEBGC, HIGH-RESOLUTION IMAGERY, LIDAR DATA, AND INVENTORY DATA.

By

JORDAN SETH GOLINKOFF

Bachelor of Arts Computer Science, Yale University, New Haven, CT, USA 2001
Masters of Forestry, Duke University, Durham, NC, USA 2006

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Forestry

The University of Montana
Missoula, MT

May, 2013

Approved by:

Sandy Ross, Associate Dean of The Graduate School
Graduate School

Dr. Steven Running, Chair
Department of Forestry and Conservation

Dr. Cory Cleveland
Department of Forestry and Conservation

Dr. Anna Sala
Department of Biological Sciences

Dr. Jon Graham
Department of Mathematical Sciences

Dr. John Bardsley
Department of Mathematical Sciences

ABSTRACT

Golinkoff, Jordan, PhD, Spring 2013 Forestry

Estimation and modeling of forest attributes across large spatial scales.

Chairperson: Dr. Steven W. Running

The accurate estimation of forest attributes at many different spatial scales is a critical problem. Forest landowners may be interested in estimating timber volume, forest biomass, and forest structure to determine their forest's condition and value. Counties and states may be interested to learn about their forests to develop sustainable management plans and policies related to forests, wildlife, and climate change. Countries and consortiums of countries need information about their forests to set global and national targets to deal with issues of climate change and deforestation as well as to set national targets and understand the state of their forest at a given point in time.

This dissertation approaches these questions from two perspectives. The first perspective uses the process model Biome-BGC paired with inventory and remote sensing data to make inferences about a current forest state given known climate and site variables. Using a model of this type, future climate data can be used to make predictions about future forest states as well. An example of this work applied to a forest in northern California is presented. The second perspective of estimating forest attributes uses high resolution aerial imagery paired with light detection and ranging (LiDAR) remote sensing data to develop statistical estimates of forest structure. Two approaches within this perspective are presented: a pixel based approach and an object based approach. Both approaches can serve as the platform on which models (either empirical growth and yield models or process models) can be run to generate inferences about future forest state and current forest biogeochemical cycling.

Table of Contents

Introduction

The goal of this work is to develop methods to estimate forest attributes across large spatial scales. The spatial scale can range from the ownership level to the full globe although this work is primarily focused on the ownership level. The motivating factor that drives this work is to develop methods that allow landowners and policy makers to more accurately understand the stocks and fluxes of forests. Using this information, policy makers and landowners can be better informed designers and participants in the new and developing market for forest carbon offsets. Unlike traditional markets for forest products (e.g., dimensional lumber), this new market for forest carbon offsets values trees that are standing.

Forest carbon offsets serve as one tool to help mitigate climate change. A forest acts as a sponge that absorbs CO_2 – one of the most important greenhouse gases – from the atmosphere and sequesters this gas as woody tissue in trees. For policy makers, accurate estimates of current carbon stocks in forests as well as reasonable predictions of the future are important components of setting reasonable baselines and in evaluating how given climate change policies are performing. For landowners, these tools are critical components of deciding whether to engage in the forest carbon offset market and evaluating the expected returns of enrolling a forest in a forest carbon offset program.

To address this need, two distinct approaches to estimating the state and change of forests were examined. The first approach using the BiomeBGC process model was examined to understand its effectiveness at predicting forest stocks and fluxes and how to best combine this model with other data sources. BiomeBGC (BBGC) is a mechanistic model that is used to estimate the state and fluxes of carbon (C), nitrogen (N), and water (H_2O) into and out of an ecosystem. In conjunction with these nutrient cycles, BBGC models the physical processes of

radiation and water disposition. BBGC partitions incoming radiation and precipitation and treats the excess/unused portions as outflows. The primary physiological processes modeled by BBGC are photosynthesis, evapotranspiration, respiration (autotrophic and heterotrophic), decomposition, the final allocation of photosynthetic assimilate, and mortality. Chapter one provides the theoretical background for this model by describing the general model framework, comparing it to other model types such as growth and yield models and gap models, and detailing the physiological basis of how this model "grows" forests (Golinkoff 2010).

Chapter two extends this discussion of model dynamics by applying BBGC to estimate forest stocks and growth for a forest in northern California. Modeling ecosystems' productivity with process models allows hypothetical scenarios to be tested and can also help constrain claims that landowners and governments make about the carbon they are sequestering and storing. Generating reasonable estimates of potential productivity is difficult both because of a lack of data and because of future climate change. However, by using inventory measurements and flux data to calibrate process models, specific locations across the globe can be accurately represented. A process model parameterized in this way can then be used to expand estimates of productivity across space when paired with remote sensing data. This integration of multiple data sources at multiple scales can provide flexibility in estimating ecosystem state and allow for estimates to vary based on different future climate scenarios. In chapter 2, the BBGC model was run within pixels across the full ownership where pixels vary based on their underlying soil properties and their spatially inferred climate variability (Golinkoff and Running 2013). Results indicate that BBGC may be a poorly constrained inversion problem and that running separate models for all pixels across a landscape may not be the most effective means to simulate ecosystem productivity.

After studying how Biome-BGC could be used to estimate forest stocks and fluxes, a second approach that combined remote sensing and ground-based inventory data was examined to see if improvements to traditional forest estimation methods could be made. This second phase was motivated by two critical gaps. The first gap is the need to develop more accurate estimates of forest stocks that can better incorporate inventory and remote sensing data. The second gap is the need to develop forest stratification approaches that could serve as better platforms on which to stage model runs similar to those examined in the first phase of study. Chapter three proposes a new pixel-based approach to stratifying a forest and structuring a forest inventory that leverages high-resolution color-infrared imagery, high-density Light Detection and Ranging (LiDAR) data, and ground inventory plots (Golinkoff et al. 2011). This chapter specifically addresses how to develop inventory systems that comply with voluntary forest carbon offset protocols. Monitoring, reporting, and verifying carbon stocks and fluxes at a project level is the single largest direct cost of a forest carbon offset project. There are now many methods for estimating forest stocks with high accuracy that use both Airborne Laser Scanning (ALS) and high-resolution optical remote sensing data. However, many of these methods are not appropriate for use under existing carbon offset standards and most have not been field tested. To bridge this implementation gap, a new, forest stratification and sampling method that meets the requirements of the Climate Action Reserve (CAR) Forest Project Protocol was designed and applied to a verified and registered carbon project in California. This approach meets the requirements of the CAR standard while reducing the costs of inventory and increasing the accuracy of estimates of carbon stocks and basal area. Results indicate that while this method achieves improvements in inventory accuracy and reductions in cost, it is not optimal for future modeling or management planning.

To address the shortcomings of the previous forest inventory stratification system, a new method that uses a similar set of remote sensing and ground inventory data, but instead uses an object-based approach to generate forest stands of similar characteristics, was developed (Golinkoff 2013). Well defined forest stands and forest strata allow managers to accurately estimate current stocks of timber and carbon as well as plan for future harvests effectively. As was shown in chapter 3, information from remotely sensed high-resolution imagery and LiDAR can provide a powerful and data rich environment to help make inferences about forests. However, these data are often too complicated for forest managers to work with and available methods may be too rigid to provide output products that are well suited to managers' needs. In chapter 4, a new Area Dependent Region Merging (ADRM) method is outlined that uses LiDAR data in conjunction with expert knowledge to develop forest stands and strata based on user supplied constraints. This method uses an area-dependent scale parameter that allows for different merging criteria based on the size of the objects being merged. The method was applied in several different forests located in Mendocino County, CA. Results were used to develop a new forest inventory that showed improved accuracy with significantly fewer field plots. Results also showed that compared to non-area-dependent region merging approaches, the area-dependent scale parameter was more effective at reducing the within stand variability and matched more closely with a reference manual stand delineation. The use of an object-based approach augments the work done in chapter three by producing a final stand delineation and stratification that is more easily understood and useable by managers and landowners. It also does a better job of representing the features of interests in forests – e.g. clearcuts, openings, etc.

The results of these studies taken together allow for three broad conclusions that further the field. First, this works shows the power of using models of different types combined with

remote sensing and field inventory data to generate accurate estimates of forest carbon, timber, and structure. Second, this work has explored several ways that remote sensing and field data can be used to partition a forest or landscape and thereby serve as a platform for future modeling of forest growth and change. Last, these methods have not only been shown to improve inventory accuracy, reduce field inventory cost, and provide a framework for future forest modeling but they have also been successfully verified and registered as approved methods in the voluntary forest carbon offset arena. The implementation of these methods provides a blueprint for other landowners and project operators to follow and hopefully can help to increase the participation in this new market for forest ecosystem services.

References

Golinkoff, J. 2010. Biome BGC version 4.2: The Theoretical Framework. Page 71. Numerical Terradynamic Simulation Group, College of Forestry and Conservation, University of Montana, Missoula, MT. Retrieved from http://ntsg.umt.edu/sites/ntsg.umt.edu/files/project/biome-bgc/Golinkiff_BiomeBGCv4.2_TheoreticalBasis_1_18_10.pdf.

Golinkoff, J. 2013. Area Dependent Region Merging: A Novel, User-Customizable Method to Create Forest Stands and Strata. European Journal of Remote Sensing – In Press.

Golinkoff, J., M. Hanus, and J. Carah. 2011. The use of airborne laser scanning to develop a pixel-based stratification for a verified carbon offset project. Carbon Balance and Management 6:9. doi: 10.1186/1750-0680-6-9.

Golinkoff, J. S., and S. W. Running. 2013. A system to integrate multi-scaled data sources for improving terrestrial carbon balance estimates. Pages 259–287 in D. G. Brown, D. T. Robinson, N. H. F. French, and B. C. Reed, editors. Land Use and the Carbon Cycle: Advances in Integrated Science, Management, and Policy. Cambridge University Press, New York, NY, USA. Retrieved December 26, 2012, from http://www.cambridge.org/us/knowledge/isbn/item6962587/?site_locale=en_US.

Chapter 1

A Theoretical Framework of Biome-BGC version 4.2

Abstract:

BiomeBGC (BBGC) is a mechanistic model that is used to estimate the state and fluxes of carbon (C), nitrogen (N), and water (H_2O) into and out of an ecosystem. BBGC is actively used in institutions around the globe and its most recent release is version 4.2. As mentioned above, the 3 primary biogeochemical cycles represented in BBGC are the C, N, and H_2O cycles. In conjunction with these cycles, BBGC models the physical processes of radiation and water disposition. BBGC partitions incoming radiation and precipitation and treats the excess/unused portions as outflows. The primary physiological processes modeled by BBGC are photosynthesis, evapotranspiration, respiration (autotrophic and heterotrophic), decomposition, the final allocation of photosynthetic assimilate, and mortality. To model these processes, BBGC first models the phenology of the systems based on the input meteorological data.

This description of BBGC below will attempt to follow the order and structure of BBGC as it is implemented to best represent the flow of information through the model system. A general discussion of the model flow and required inputs will be given first, followed by a broad outline of the model processes and assumptions. BBGC will be compared to Forest Gap Models and Growth and Yield Models. Lastly, a detailed description of each of the BBGC's processes will be presented (Peter Thornton's thesis was an essential reference in understanding this model (Thornton 1998)).

1: General Model Flow

Figure 1 shows the general flow of the BBGC model. The first step in any BBGC model run is a spinup to bring the model into equilibrium. It is common for ecosystem models to require a steady state initial condition so as to insure that there is a balance between input and output fluxes and that the system has equilibrated to the environmental and site forcings (Thornton and Rosenbloom 2005). In the current version of BBGC, this means that the difference between the annual average daily soil carbon stocks must be less than a specified spinup tolerance value (SPINUP_TOLERANCE = 0.0005 $kg/m^2/yr$).

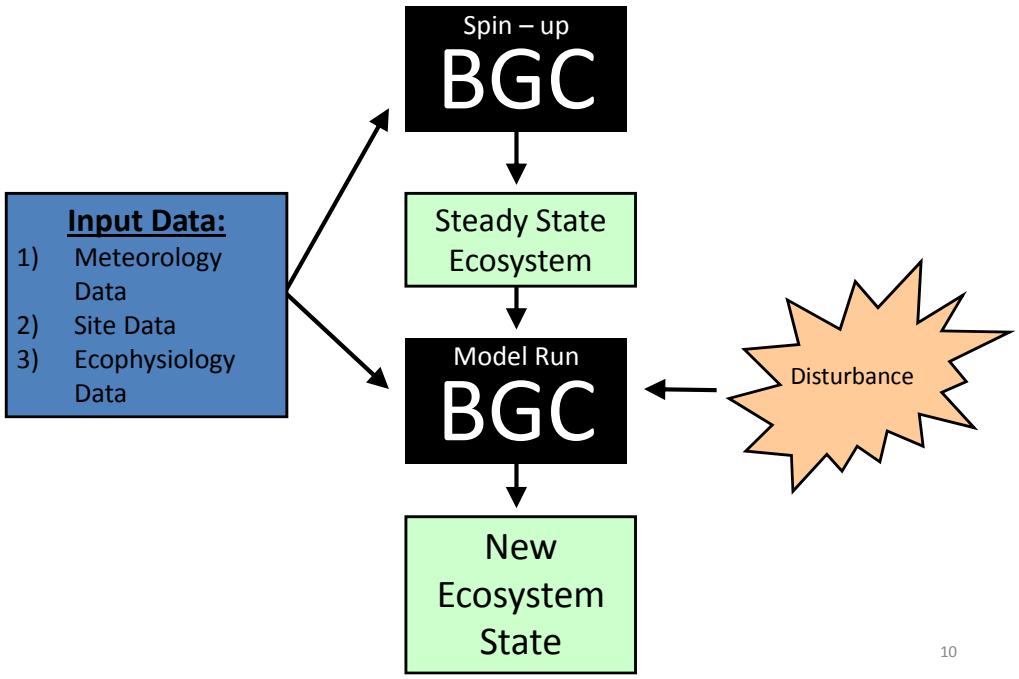

Figure 1: Conceptual diagram showing BiomeBGC general model structure

As seen in figure 1, any model run (spinup or otherwise) requires a certain set of input data. BBGC requires meteorological (met), physical (ini), and ecophysiological (epc) data for each site. Appendix A details the inputs required for each of these categories. Every model run then produces a set of data that can be outputted for the user to analyze. Appendix B lists the output variables users can request (in either binary or text form). These variables include all of the C, N, and H_2O fluxes and pools that BBGC tracks as well as summary variables (e.g. -Net Ecosystem Exchange (NEE) or Net Primary Productivity (NPP)) at daily, monthly, or annual time scales. BBGC can be run to a spinup steady state and then forward in time, or it can accept as an input the ending model state of a previous model run (a restart file) and run from this point forward with a new set of model assumptions if desired.

2: Model Overview

Broad Conceptual Basis and Critical Assumptions:

BBGC is a one-dimensional model meaning that it represents a point in space with all fluxes and stocks scaled to a per square meter basis (Thornton 1998). When run in a spatial context over a landscape, each cell is a distinct model run and does not interact with other cells. This rules out the use of BBGC to examine competitive dynamics across space such as shading from differing height growth. It also prevents more detailed analysis of the impact of vegetation on the hydrological flow across a landscape. That said, models with this spatial awareness do exist and BBGC could be modified to account for spatial interactions, but this is not pursued in this work. Given BBGC's spatial perspective, it is helpful to think of this model as an estimate of stand level processes that have been aggregated and averaged to a per unit area basis. This scale is an appropriate framework as BBGC does not attempt to represent individual trees or even individual species but rather the dynamics at a point of a plant functional type (PFT) – e.g. – evergreen needleleaf forest, or deciduous broadleaf forest, or C3 grassland (Waring and Running 2007).

Another critical abstraction BBGC makes is to ignore successional dynamics within its spatial context. BBGC is parameterized by a user to grow a given PFT for the full span of its model run. Ignoring plant succession also allows BBGC to ignore competition between PFTs that is mediated by different adaptive strategies and growth traits. As an example of where this abstraction is used, all of BBGC's pools are dimensionless and can better be thought of as buckets for storage rather than actual plant structures with known height, width, and lengths. Some variants of BBGC have attempted to remove this abstraction to model competition between PFTs (Korol, Running et al. 1995; Bond-Lamberty, Gower et al. 2005). The only exception in BBGC to the use of dimensionless pools is the treatment of leaf carbon.

To model the process of photosynthesis, BBGC converts leaf C into an equivalent leaf area (LA) based on user defined Specific Leaf Area (SLA) parameters. SLA is a measure of the thickness of a leaf and its units are area per unit mass (i.e. - m^2/kgC). BBGC further partitions leaf C and LA into sun and shade leaves. All photosynthetic, respiration, and transpiration processes are then carried out for both the sun and shade leaf components of the system. This two leaf model is more accurate than simple big-leaf models (one big leaf) and doesn't sacrifice much accuracy when compared to more complicated multi-layer approaches (De Pury and Farquhar 1997). This approach to modeling canopy dynamics is also able to capture some of the known variability of SLA through a tree crown (Koch, Sillett et al. 2004; Thornton and Zimmermann 2007). For example, it has been observed that leaves exposed to full sun usually have lower SLA than those in the shade on the same tree.

Another abstraction made in the implementation of BBGC is the chosen temporal resolution. BBGC uses both a daily and an annual timestep. Most processes are applied on a daily basis with some pool updating occurring annually (Thornton, Law et al. 2002). Despite this model time scale, many of the actual processes that occur within plants adjust rapidly to changes in the environment that happen on a sub-daily basis (Lambers, Chapin et al. 2008). However, accurate measurements at this time scale are much more difficult to obtain and in many cases are unavailable. Therefore, using a daily time-step, while not capturing some of the true ecosystem dynamics (e.g. – sun spots, clouds, wind gusts, etc), allows for a more broadly usable model. Furthermore, some of these sub-daily phenomena likely average out when looking at the daily rates.

The last two major assumptions built into BBGC concern growing ecosystems without knowing future conditions. Because BBGC is a prognostic model (it is not constrained by

diagnostic observations over time but rather builds a given system from a series of first principles), some look-ahead logic must be used to help constrain the model as it grows into the future. The first instance of this is the model's phenological approach. This approach uses a critical soil temperature constraint (and moisture constraints for grasslands) to estimate the start of growing season (and the start of senescence for deciduous systems) (White, Thornton et al. 1997). However, this requires looking ahead at the input climate data to calculate the appropriate onset and senescence dates rather than allowing the system to prognostically determine these dates on the fly. The second look ahead approach is used to prevent the model from developing a large C or N deficit. BBGC allocates newly assimilated carbon first to a carbon pool that can then be used over the course of a growing season when conditions for growth become stressful. This mimics a plant's ability to store carbon for stressful times and negates the model's need to look ahead and estimate respiration demand based on future climate (Thornton and Rosenbloom 2005).

Physical Model Processes

Radiation:

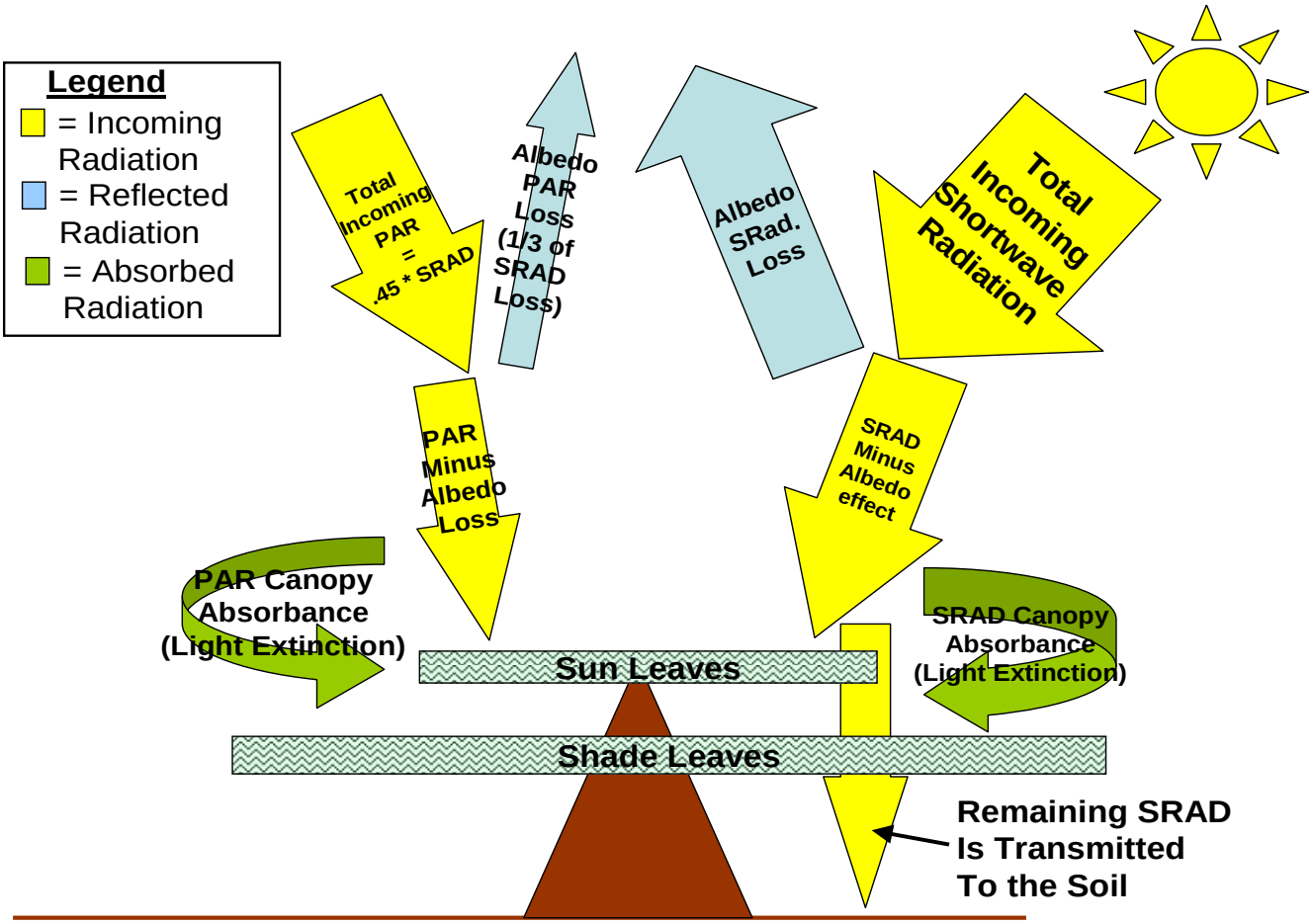

Figure 2: BiomeBGC radiation partitioning

Once model phenology is defined as described above, the first step is to account for the disposition of incoming shortwave radiation. This is done for each day the model is run. An estimate of incoming shortwave radiation is one of the required daily inputs in the meteorological data (see Appendix A). Figure 2 shows a conceptual diagram outlining how radiation is partitioned by the model. As can be seen, the proportion of radiation absorbed by the canopy depends on the sun and shade leaf LA. Therefore, prior to the radiation partitioning, the leaf C pool paired with the SLA of shade and sun leaves is used to determine the total leaf area and the sun and shade leaf proportions of this. The incoming shortwave radiation, converted first to Photosynthetically Active Radiation (PAR ~ 400 to 700 nm), is then absorbed by the canopy following Beer's Law of light attenuation (Nobel 1991; Jones 1992). The partitioned radiation is

one of the inputs then used to drive canopy evapotranspiration, photosynthesis, and soil evaporation.

Precipitation and H_2O Cycle:

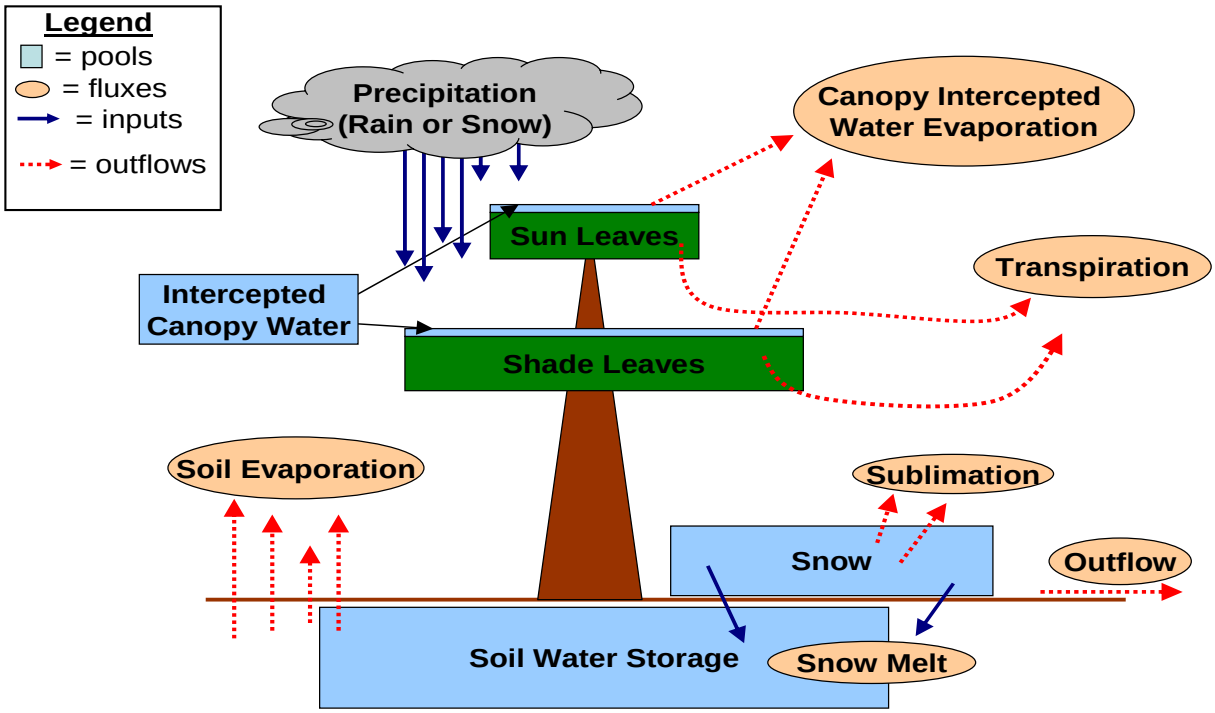

Figure 3: BiomeBGC water pools and fluxes.

Once the radiation budget for a day is calculated, the water state variables can be addressed. The only input of water into the system occurs through precipitation either as rain or snow. Daily precipitation is also one of the required daily meteorological input variables (see Appendix A). This precipitation is then routed to several potential pools. Figure 3 outlines the H_2O pools and fluxes. The first resting place for incoming precipitation is the canopy intercepted rainwater pool. The amount of intercepted rainwater is a function of a user defined canopy interception coefficient, the amount of rainwater, and the Leaf Area Index (LAI – a unitless value that is the area of all leaves per unit ground area – m^2/m^2). The model assumes no snow interception. Snow accumulates in a snow water pool when the temperature is below

freezing and melts when the temperature is warmer than freezing. Snow also can sublimate when the temperature is less than freezing based on the amount of incoming solar radiation it receives.

If there is more than enough water to fill the canopy interception pool, the remaining water enters the soil water pool. The current soil water matric potential (MPa) is a function of the water in the soil now in relation to the soil's saturated water holding capacity. Saturated soil water and field capacity soil water holding are defined based on the soil texture and depth (as specified in the site initialization file (see Appendix A – percentage sand, silt, clay, and depth) (Cosby, Hornberger et al. 1984; Saxton, Rawls et al. 1986). The current soil water matric potential (ψ_{soil}) is then determined by removing the calculated evaporation from the soil and the addition of water to the soil pool from precipitation (and snowmelt if there is any). All evaporative processes (canopy evaporation of intercepted water, transpiration during photosynthesis, and soil evaporation) are calculated using a modified Penman-Monteith Equation – PME (McNaughton and Jarvis 1983; Waring and Running 2007; Monteith and Unsworth 2008). This equation calculates an evaporation rate that is a function of incoming radiation, vapor pressure deficit (VPD), and the conductances associated with the evaporation surface.

Physiological Model Processes: C and N Cycle – Pools and Fluxes

Throughout the general discussion of the C and N cycle's pools and fluxes, refer to figures 4 and 5 for a schematic representation of these pool and fluxes in the ecosystem. Most broadly, the C cycle consists of all of the pools seen in figure 4. The only addition of C to the system occurs through the photosynthesis process. C is removed from the system during all of the respiration processes: autotrophic (maintenance and growth) and heterotrophic

(decomposition). C is also lost from the system during a fire or harvest disturbance event. In the case of fire, C pools are moved to an atmospheric pool and are not tracked by the model.

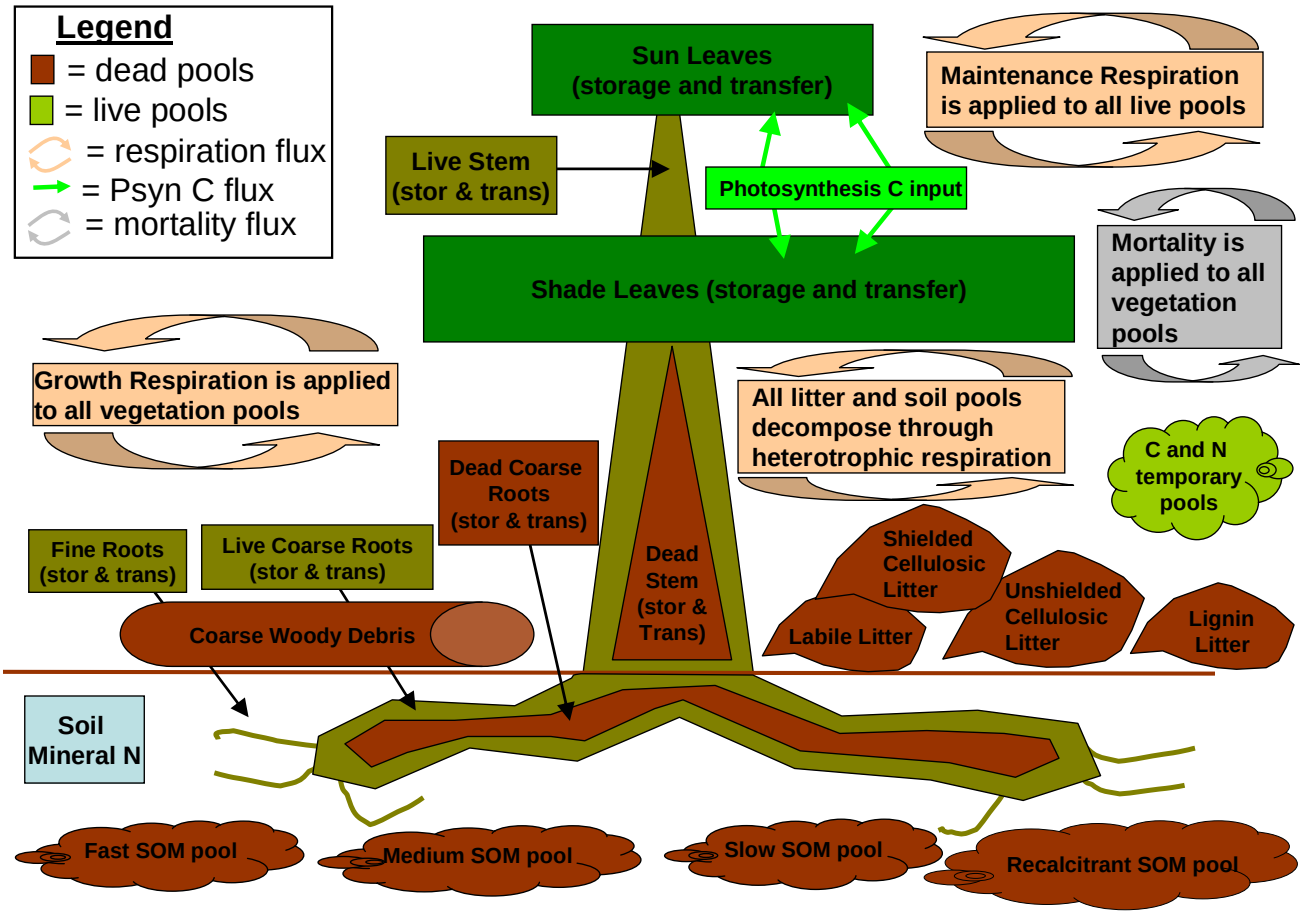

Figure 4: BiomeBGC C and N pools.

In general and as seen in figure 5, the N cycle in BBGC consists of all of the plant pools as well as a soil mineral N pool and a plant retranslocated N pool. N retranslocation occurs based on the phenology of the system as tissues turnover during the growing season. When plants lose their leaves, some of the leaf N is reabsorbed by the plant for future use. Soil mineral N is added to the system in only three ways: mineralization from the slowest soil organic matter (SOM) pool, N wet and dry deposition (Ndep) from the atmosphere, and N fixation (Nfix) (Ndep and Nfix are both user defined rates found in the .ini file – see Appendix A). Mineralized N is lost from the system either through leaching when there is H_2O outflow or through bulk

denitrification (N volatilization) where both leaching and volatilization are assumed to occur at constant rates.

Maintenance Respiration:

Once the water state and the radiation partitioning are known, BBGC enters into the main C and N cycle calculations. The first step of this process is to calculate maintenance respiration (MR) of all living tissues. This is done before photosynthesis as the MR of leaves is needed in the carbon assimilation calculation. MR in BBGC is a Q_{10} function of temperature and a linear function of the N content. A Q_{10} function is an exponential function where a 10°C increase in temperature relates to a Q_{10} factor change in the rate of respiration.

Photosynthesis:

As discussed above, the BBGC photosynthesis model uses a two-leaf representation of the canopy to model all canopy photosynthesis. All photosynthesis calculations are performed separately for sun and shade leaves. The details of the model implementation of photosynthesis are based on Farquhar et al. (1980) and will be further discussed in the detailed model description section. The photosynthesis model is based on the enzymatic kinetics of Rubisco in relation to temperature, the availability of CO_2 and the rate of Rubisco regeneration. Photosynthesis is the only process in BBGC that provides an input of C into any pool. All C comes from the C assimilated during this process. Initially, this assimilate is placed into a temporary storage pool (cpool) where it is then portioned to future growth, storage, and current growth. Before the assimilate can be allocated however, the microbial demand for N from decomposition must be derived to determine if N will limit the allocation of assimilated C.

Decomposition:

As can be seen in figure 4, BBGC has several pools that store the C and N of dead and decaying wood and leaves. The coarse woody debris (CWD) pool is the first pool that dead coarse roots and dead stem wood enter when they die. This pool then fragments into the litter pools over time. The rate of fragmentation is dependent on the moisture and temperature of the site. As opposed to coarse woody material, fine roots and leaves directly enter the litter pools when they die. The defragmented CWD and the leaves and fine roots are partitioned into specific litter pools depending on the relative amounts of carbon found in labile, cellulose, or lignin forms (user defined constants in .epc file – see Appendix A). These litter pools then decompose and enter into the soil organic matter (SOM) pools.

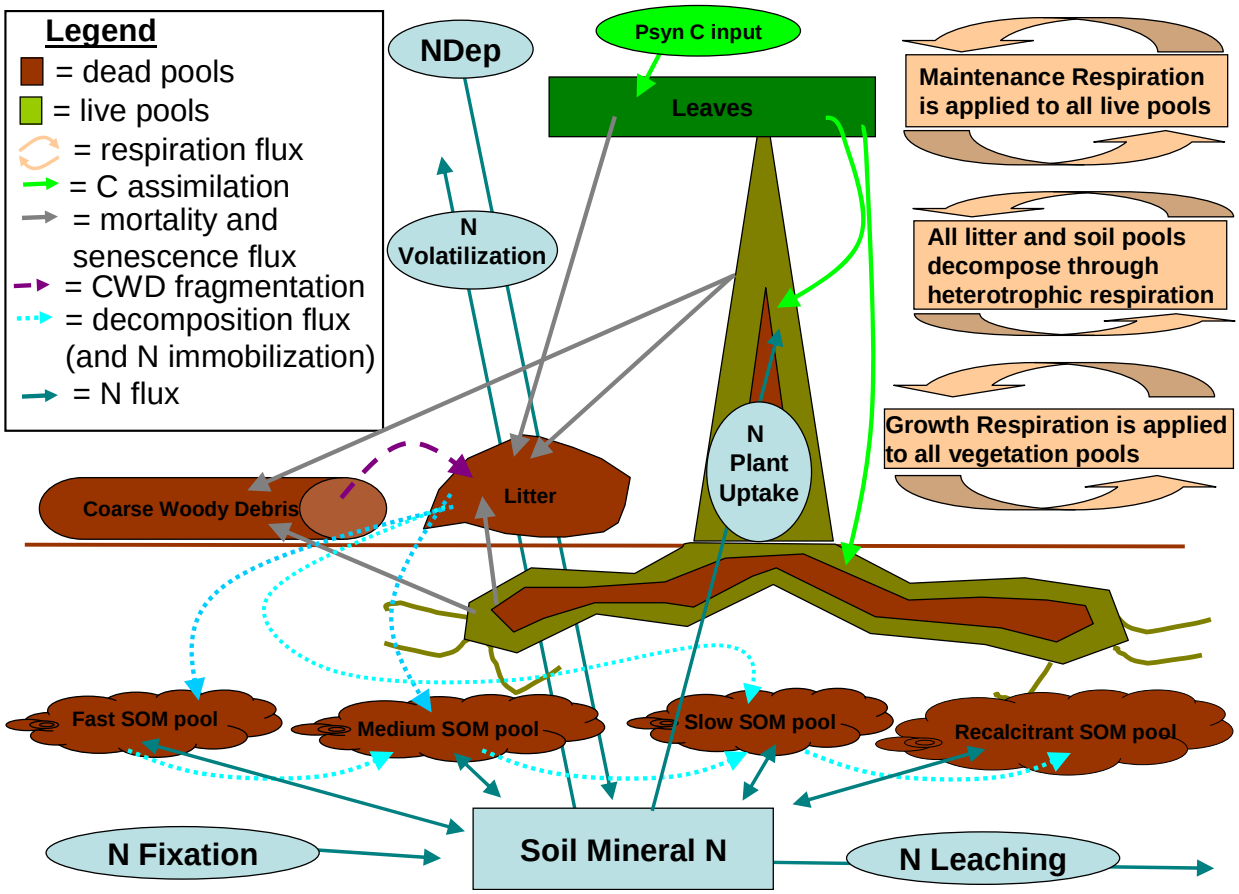

Figure 5: BiomeBGC C and N fluxes.

The SOM pools also undergo decomposition constrained by soil water and temperature. As SOM decomposes and N is immobilized by microbes, the SOM is transferred into

successively slower decomposing pools. Figure 5 shows the fluxes from CWD and litter and between the SOM pools. BBGC calculates the non-N limited rates of decay and stores these rates until the plant's N demand is calculated. The potential plant C allocation and the potential decomposition are scaled by the total N limitation of the system. This framework makes several key assumptions. As mentioned above, BBGC assumes that resolving plant N demand competition with microbial N demand at a daily timestep can appropriately represent what occurs at a sub-daily time scale. Second, BBGC assumes that microbes and plants have equal weight when competing for soil N. BBGC also assumes constant C:N ratios for soil pools regardless of PFT as well as constant decomposition rates of the litter and soil pools regardless of the PFT.

Allocation:

The allocation of assimilated C, and the actual decomposition that occurs, are all calculated after photosynthesis has found the potential assimilation and decomposition has calculated the potential decay. BBGC scales the actual allocation and decomposition based on the availability of N – both soil mineral N and retranslocated N found in the plant as storage. The core of BBGC's allocation scheme uses a set of fixed fractions for all plant structures (user defined in the .epc file – see Appendix A) to apportion C once the N limits are considered. BBGC also sets aside a fixed percentage (again user defined) of the assimilated carbon as storage for next year's growth and a fixed percentage (30%) for GR (growth respiration). When allocating this year's growth to different tissues, BBGC scales all allocation in relation to leaf carbon allocation while maintaining the user defined proportions in every pool (Waring and Pitman 1985; Waring and Running 2007; Wang, Ichii et al. 2009). All of the allocation proportions are assumed constant over the life of the ecosystem. Furthermore, although there is

explicit N limitation built into the photosynthesis calculation, if there is further N limitation during allocation, BBGC reduces shade and sun leaf assimilate proportionally to reflect this limitation.

Growth Respiration:

Growth respiration (GR) is assumed to be a constant proportion of all new tissue growth (30% of new tissue is respired - (Larcher 2003)). GR is accounted for during the allocation of assimilate to new tissue.

Mortality:

BBGC uses a user defined (epc file) fixed mortality fraction that is applied each day. BBGC also has a fixed user defined fire mortality fraction that behaves in the same way but moves the C and N to an atmospheric pool rather than into decomposing pools.

Principle of the Conservation of Energy and Mass:

BBGC's fundamental principle is that incoming energy radiation, C, N, and H_2O must all be in balance at any given time (Thornton 1998). In practice, this means that at the end of each day BBGC updates each state variable and checks for balance. For the four elements listed above to be "in balance", the incoming quantities minus the outgoing quantities must be equal to the storage in the model. After all of the processes described above are modeled, BBGC checks this condition.

3: Comparison of BBGC with Gap Models and Growth and Yield Models

There are many models used to represent forest ecosystem dynamics. These models can broadly be put into two large categories: mechanistic/process/physiological based models or empirical models. Process models, like BBGC, attempt to model and explain ecosystem function

by modeling the mechanisms within plants that cause them to grow, breathe, die, and decay (prognostic). Empirical models use measurements of ecosystems to generate relationships between critical ecosystem variables (e.g. height growth and age) and then use these measured relationships to model how ecosystems will change (diagnostic). Vanclay (1994) also makes the distinction between "models for understanding" and "models for prediction". In this framework, some models (i.e. process models and FGMs) have been developed to help improve our understanding of ecosystem function and explain the dynamics observed in natural systems. Other models (i.e. forest growth and yield models) have been developed to predict future ecosystem states for more applied purposes such as forest management or timber harvest income stream prediction. Although many process modelers would take issue with this distinction (clearly process models are used to predict future ecosystem state as well) and there have been numerous applications of "models of understanding" to forest management (e.g. - (Harmon and Marks 2002; Pietsch and Hasenauer 2002; Shugart 2002; Thornton, Law et al. 2002; Schmid, Thürig et al. 2006), in general it is still true that the vast majority of forest management occurs using empirical based models.

Despite the differences in the application of these different model types, there is less of a dichotomy and more of a continuum of model types in between pure process based approaches and pure empirical methods of understanding forest change and stocks. Figure 6 is a conceptual diagram showing the continuum between the underlying model basis as well as the model's spatial scale. Figure 6 also shows another set of color ramped axes that could be used to color each model oval: whether they focus on one species/PFT or many and whether they model mixed-age systems or even-age systems.

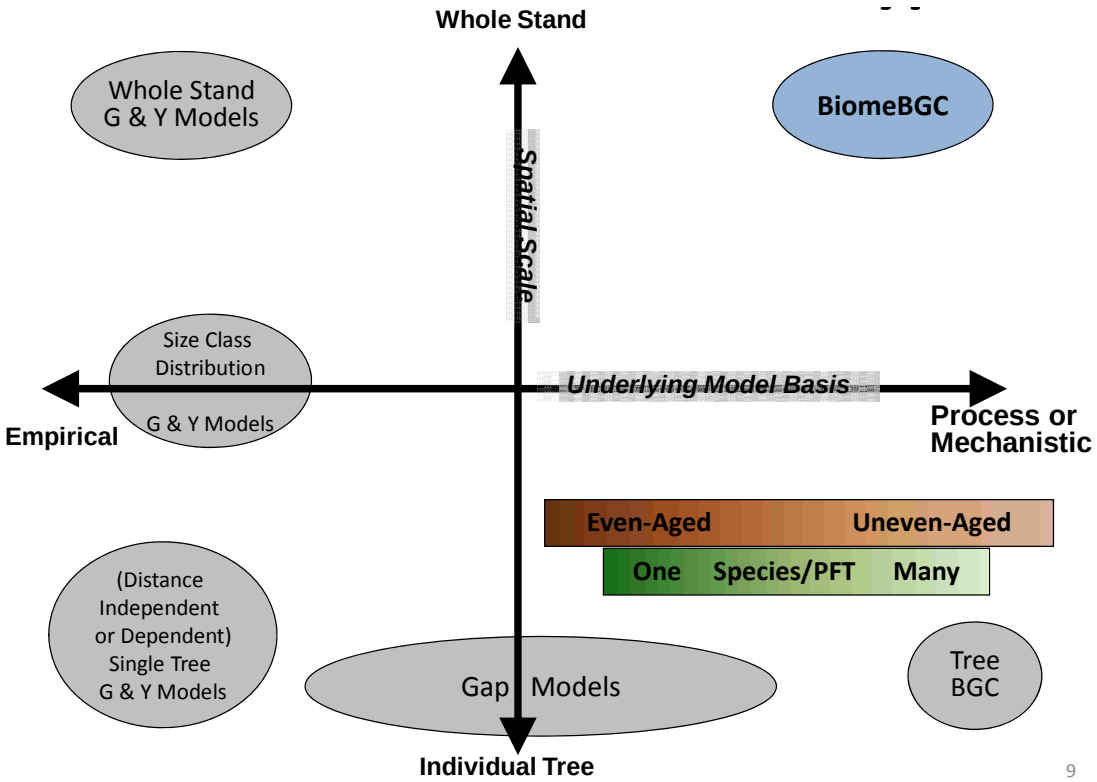

Figure 6: Conceptual ecosystem modeling continuums.
1) Model Basis vs. Spatial Scale and 2) Age Structure vs. Species/PFT composition

Forest Gap Models:

Forest Gap Models (FGMs) can be thought of as falling somewhere in between empirical approaches of forest modeling and mechanistic approaches. JABOWA, the seminal FGM, was developed to predict successional change in a New Hampshire forest (Botkin, Janak et al. 1972). Hence, FGMs are also known as successional models. Like all broad model categories, there are many variants of FGMs that have been developed over the years. Despite the wide range of FGMs that have been developed, there are some overarching characteristics that define this approach.

To begin, the scale that FGMs focus on are gaps in the forest created when large trees die and there is resulting competition between trees in these openings. As a result, most FGMs focus on forest patches between 100 and 1000 m^2 (the size of the crown of one or two dominant trees).

Second, by definition, FGMs model individual trees within each patch. Originally, FGMs were distance independent and were not spatially explicit when considering the location of individual trees. Later FGMs introduced this spatial explicitness within patches and other models developed nearest neighbor relationships between patches (e.g. ZELIG) (Bugmann 2001). The growth of trees in FGMs is driven in most cases by empirical relationships between age, height, and density. However, many FGMs also scale growth rates by site conditions such as nutrient supply and climate forcings. Additionally, all FGMs attempt to understand the dynamics of succession (mediated by shading) and in these senses they also model some of the mechanisms of forest growth and change (Shugart 2002). Most broadly, FGMs track individual tree growth, individual tree mortality, patch density, regeneration and recruitment to help explain competition and succession. Figure 7, from Solomon and Bartlein (1992), summarizes the different common components of most FGMs.

In comparison to BBGC, FGMs are focused on individual tree dynamics. Some FGMs use a stochastic approach to seed dispersal and mortality. In these cases, many FGM runs will be used to generate stand level estimates of the forest state rather than individual tree estimates (however the model itself still grows individual trees). Over the years, more and more physiology has been added to FGMs as an attempt to better model the growth of trees (e.g. the FIRE-BGC and HYBRID FGMs). In many ways, these FGMs use similar approaches to modeling growth as BBGC uses. For example, these models use a Q_{10} respiration model and the Farquhar Photosynthesis Model to estimate growth (Farquhar, Caemmerer et al. 1980). To incorporate this physiology, the temporal scale of FGMs has been changed from annual time steps to daily time steps. Incorporating these processes into a FGM at a tree level then allows forward looking projections that take into account changing climate and CO_2 levels. From the

opposite perspective, FOREST-BGC (BBGC's predecessor, see Running and Gower (1991)) was modified to incorporate some of FGMs' logic in estimating stand density and competition (Korol, Running et al. 1995). In conclusion, BBGC's process level approach of focusing on pools and fluxes at a stand level makes it substantially different than FGMs empirically driven focus on individual tree competition dynamics.

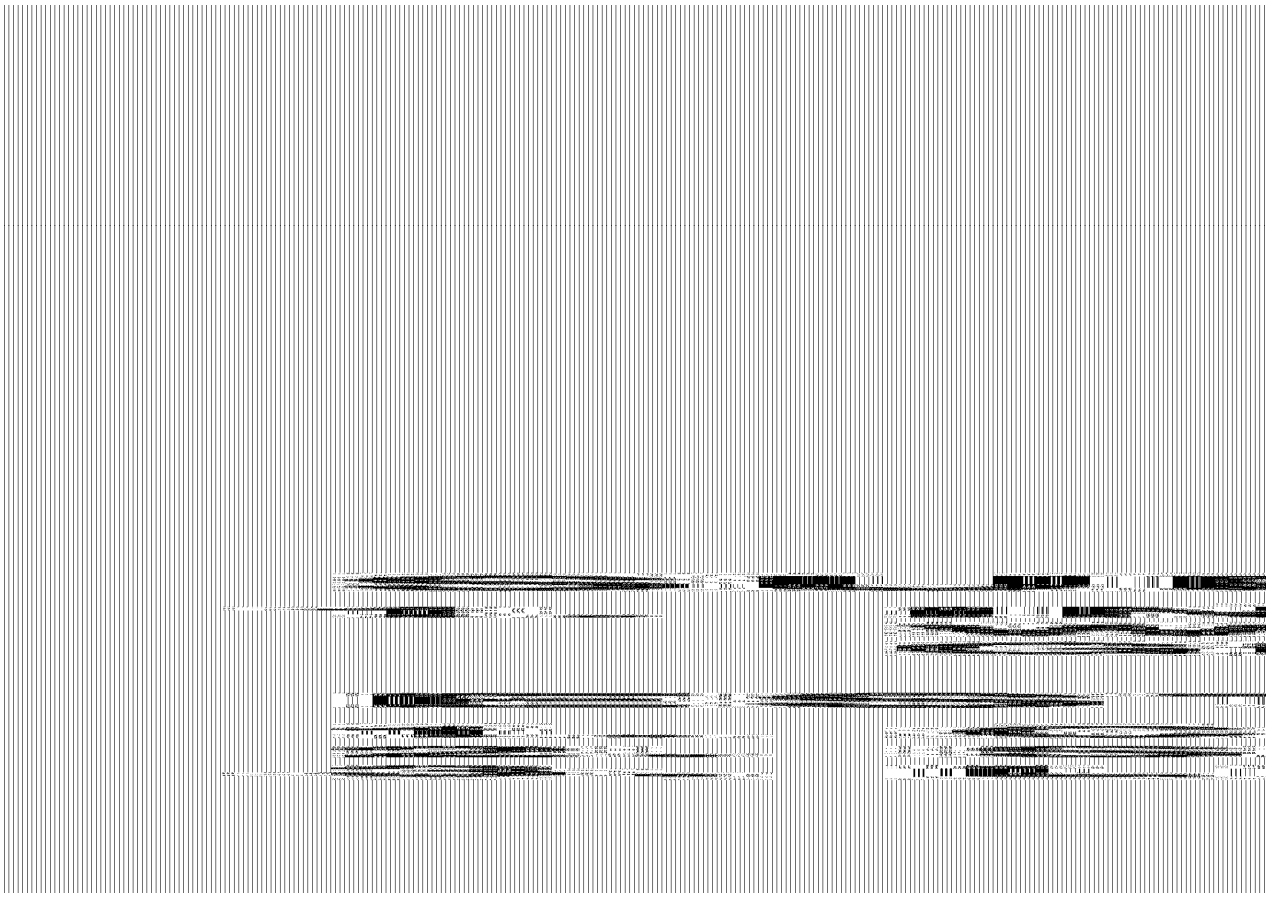

Figure 7: Figure from Soloman and Bartelein (1992) showing the FGM components.

Forest Growth and Yield Models:

As mentioned above, a growth and yield model (GYM) is an example of an empirical model. These models can take many forms and, as seen in figure 6, run the gamut between focusing on whole stand modeling to individual tree modeling. GYMs are predominantly used by field foresters to predict future forest states on the time scale of a rotation (i.e. – 20 to 50

years). Yield in these models refers to the total volume of timber available for harvest at any given time. Growth is defined as the rate that yield accumulates and is the first derivative of the yield function (Avery and Burkhart 1975). The field of GYM is quite old and foresters have used models of this sort for over 250 years beginning with simple density independent stand volume tables (Porté and Bartelink 2002).

At their core, GYMs are built from empirically derived relationships between stand characteristics such as density, height, diameter distributions, age, and site class against stand volume. Individual tree GYMs use a similar approach but relate these stand characteristics to individual tree growth. The data used to drive these models can come from long term permanent plots showing forest development over time or can be taken from many different forests of different ages, site conditions, and stocking rates to build the appropriate relationships. Because GYMs use data from past forest growth, GYMs implicitly assume that past drivers of growth such as climate and CO_2 levels will not change enough to dramatically impact the growth dynamics of forests in the future. For short time scales, this assumption may be valid but for longer timescales, this is probably an inappropriate assumption. Furthermore, although GYMs can model stands while considering different nutrient constraints, these models do not model the impact of management on the nutrient cycle and hence may overlook the impact of changes in N deposition rates or the impact of removing live trees, litter, and CWD (and their associated N pools) on plant growth.

In comparison to BBGC, most GYMs do not employ process logic to estimate ecosystem state but rather rely on observations of similar ecosystems to make predictions about stocks and change. With that said, some GYMs have incorporated some scaling logic to account for the impact of changes in the drivers to growth on the predicted future forest state. Furthermore,

most GYMs track only those variables that are important to forest managers and have explicit dimensional representations of the trees modeled. In contrast to this, BBGC accounts for all fluxes into and out of most of the pools found within a stand and does not account for tree dimensions. Because of these differences, GYMs and BBGC are at opposite ends of the modeling basis continuum seen in figure 6.

4: BiomeBGC – Utility and Applications

Based on BBGC's model framework and the assumptions that underlie this framework as well as the comparison of BBGC with FGMs and GYMs, it is possible to appreciate BBGC's utility and optimal application. As a "model of understanding", BBGC is used in studying the underlying mechanisms that have caused an ecosystem to look and behave as it does. However, these mechanisms are restricted to systems with one primary plant functional type and few successional dynamics. Furthermore, given BBGC's treeless and density-less abstraction, BBGC cannot give insight into the inter-tree competitive processes at play at a location. Despite this, BBGC has been applied in many systems to help understand the drivers of growth and decay.

Because of the spinup process BBGC uses, BBGC can be used to estimate the old-growth (steady-state) outcome of systems. With realistic fire mortality parameters, BBGC can help to understand the steady state stocks and fluxes of systems that undergo periodic disturbances as well. Because BBGC accounts for changing climate and CO_2 levels in the atmosphere, BBGC can be used to predict ecosystem states and fluxes given changing climate (Vetter, Churkina et al. 2008). With some minor modifications, BBGC can also be used to model human disturbance such as harvest as well as natural disturbances (Thornton, Law et al. 2002). BBGC has also been modified to allow for modeling successional change between PFTs

at a given location by taking into account the height growth of PFTs (Bond-Lamberty, Gower et al. 2005).

Two areas that limit BBGC's utility are its spatial approach and its parameterization. Although BBGC has been used to create gridded spatial runs, it is important to understand that neighboring cells do not interact in any way. This does not prevent BBGCs use in a spatial setting, but it does limit some of the inferences that can be made from these runs. Second, although when parameterized well BBGC can accurately represent many biome types across the globe, the amount of physiological detail required to adequately initialize the model can make it prohibitively difficult to use in some cases (see Appendix A). In particular, forest managers often have information about merchantable forest stocks, the dimension of trees or the density of trees. However, they do not often have information about leaf chemistry or assimilate allocation fractions. There have been several large scale attempts to provide sources for BBGC's parameterization, however in some cases these documents still might not provide enough information to adequately proceed (White, Thornton et al. 2000; Pietsch, Hasenauer et al. 2005). Furthermore, because the outputs BBGC provides are not commonly used by managers, they have less utility in the field. One challenge moving forward is to try to modify BBGC in ways that make it more broadly useful to managers of systems rather than just academics.

Lastly, as with any representation of a complex natural system, BBGC's modeling assumptions may or may not be appropriate to represent a given system. In some cases this means that some systems require specific variants of the model (e.g. addressing the stomatal uptake of fog water in Redwood trees). In other cases, the model's logic may not be a correct representation of how systems actually work. For example, Wang et al. (2009) found that BBGC

may not be modeling enough MR and GR to accurately estimate NPP and modified the model to address this problem.

5: Detailed Model Description

This section will outline all of the major processes and equations that BBGC uses to model an ecosystem and will be faithful to the actual model structure referring to the individual model functions when appropriate. Figure 8 is a detailed flow chart of all of the conceptual components of BBGC. Throughout this discussion, reference will be made to several standard constant values. These values are defined in the bgc_constants.h file and are included here as Appendix C.

Precalculations:

Before entering into the main daily loop of BBGC, all of the input files are read, the data structures are initialized, and several calculations are performed. The soil texture information (% sand, silt, and clay) is used to find the saturated soil conditions. The following equations are used from Cosby et al. (1984):

(1) Soil Saturated Volumetric Water Content (SatVWC) = $\frac{50.5 - 0.142 * sand - 0.037 * clay}{100}$

(2) Soil Saturated Matric Potential (SatPSI) =

$$-\exp((1.54 - 0.0095 * sand + 0.0063 * silt) * \log(10) * 9.8E - 5)$$

(3) Soil Field Capacity Volumetric Water Content (FcVWC) =

$$SatVWC * \left[\left(\frac{-0.015}{SatPSI} \right)^{\left(\frac{1}{-(3.10 + 0.157 * clay - 0.003 * sand)} \right)} \right]$$

Incoming shorwave radiation (SRAD) is converted to incoming PAR by multiplying by 0.45 (Nobel 1991). The atmospheric pressure is calculated based on the elevation and using several atmospheric constants (Iribane and Godson 1981) –

$$(4) \qquad \text{AtmPres (Pa)} = PSTD * \left\{ 1 - \left[\frac{(LRSTD * elev)}{TSTD} \right] \right\}^{\left[\frac{GSTD}{\left(LRSTD * \frac{R}{MA} \right)} \right]}$$

where PSTD = standard pressure (Pa) at 0m elevation, LRSTD = standard temperature lapse rate (-K/m), TSTD = standard temperature (K) at 0m elevation, GSTD = gravitational acceleration (m/s^2), R = gas law constant (m^3*Pa/mol K), and MA = molecular weight of air (kg/mol).

Lastly, the shielded and unshielded fractions of the cellulose litter pool are calculated in the epc initialization routine. The model for this is found in several studies that outline litter decomposition rates based on lignin ratios in litter (Berg, Ekbohm et al. 1984; Berg and McClaugherty 1989; Donnelly, Entry et al. 1990; Taylor, Prescott et al. 1991; Stump and Binkley 1993). If the ratio of lignin to cellulose is less than or equal to .45, then there is no shielded cellulosic pool. If the ratio is between .45 and .7, then the shielded cellulosic component is:

$$(5) \qquad \text{Shielded cellulose fraction} = \text{cellulose fraction} * \left[\left(\frac{ligninFrac}{celluloseFrac} - 0.45 \right) * 3.2 \right]$$

(6) Unshielded cellulose fraction = cellulose fraction – shielded cellulose fraction

If the ratio is greater than .7, the shielded cellulosic component is 80% of the cellulose fraction and the unshielded component is the remaining 20% of the cellulose fraction of the litter.

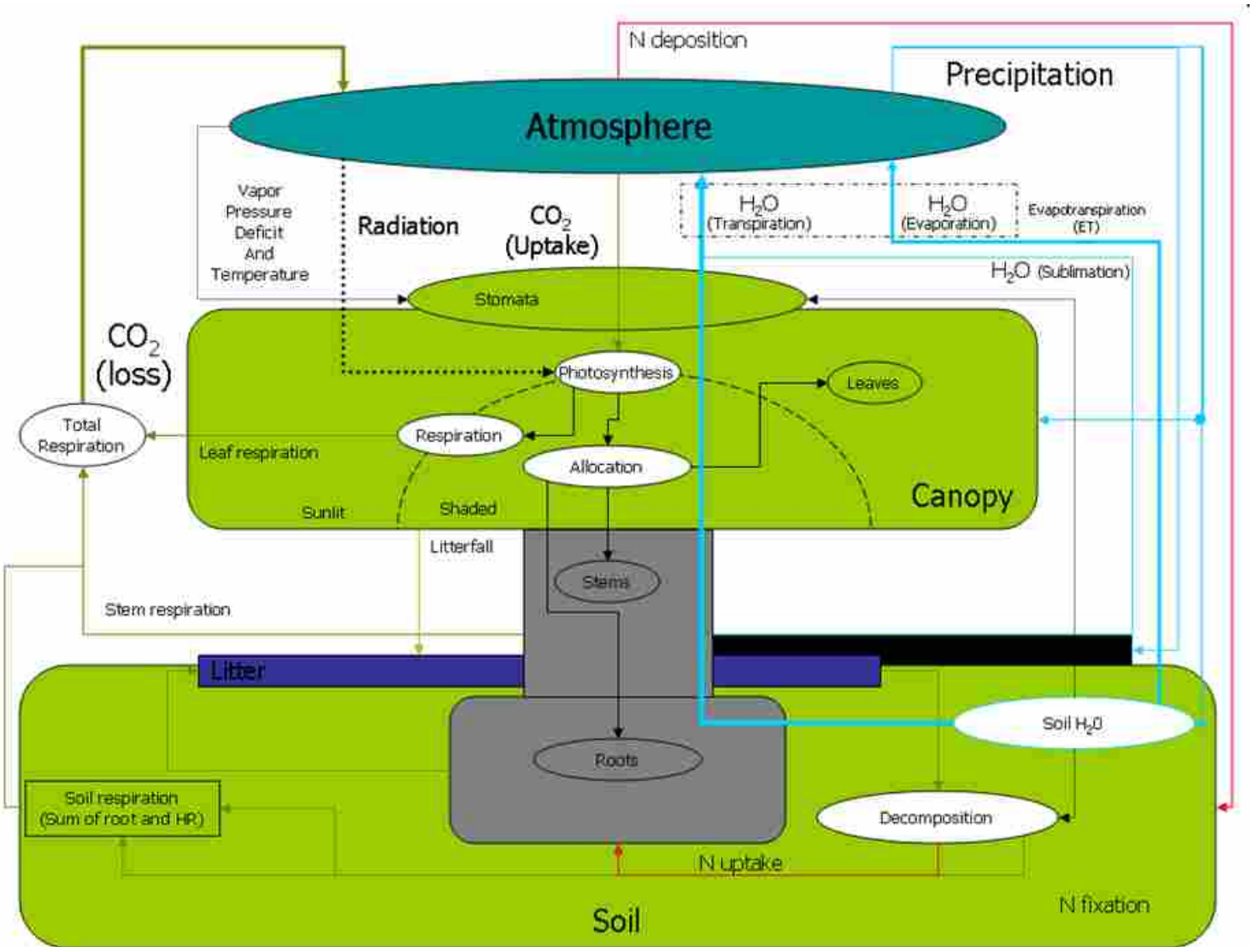

Figure 8: BiomeBGC detailed model flow chart. See: http://www.ntsg.umt.edu/models/bgc/index.php?option=com_content&task=view&id=23&Itemid=27

Prephenology (prephenology.c):

The phenology model used by BBGC is described in White et. al (White, Thornton et al. 1997) and all of the constants below can be found in that paper. BBGC's phenology can also be user specified if the user has information about the onset of growing season (i.e. – bud break) and beginning of senescence in deciduous systems. The White et. al model specifies separate phenologies for woody plants (i.e. trees and brush) versus grasses. For evergreen systems, it is assumed that the growing season is all year long.

For deciduous woody plants, leaf onset begins when the running sum of the daily average soil temperatures (when the average soil temp is above 0°C) is above a critical value defined by:

(7) $TcritSum_{woody} = \exp(4.795+0.129*Tavg)$

The model also specifies that the day length must be longer than 10 hours and 55 minutes for leaf out to occur (39300 seconds). For grasses, the leaf onset is controlled by both temperature and water availability in a similar fashion. The critical temperature running sum value for grasses is defined as:

(8) $$TcritSum_{grass} = 481 * \left[\frac{e^{32.9*(Tavg-9)} - 1}{e^{32.9*(Tavg-9)} + 1} \right] + 900$$

where Tavg is the mean daily average temperature over the full meteorological input record. The critical precipitation value is defined as:

(9) $PrcpCritSum_{grass} = AvgAnnPrcp * 0.15.$

When both the summed soil temperatures and the summed precipitation values are greater than or equal to the grass critical values, leaf onset begins. The actual leaf onset day is 15 days prior to this calculated date to estimate the start of the growing season.

The beginning of leaf senescence is also separately defined for woody and grass species. For deciduous woody PFTs, senescence begins if it is past July 1st and the day length is less than the critical day length described above and the soil temperature is less than the average fall soil temperature (Sept. and Oct. in the northern hemisphere) OR if the soil temperature ever drops below 2°C. For grasses, senescence begins under two conditions. First, if there has been less than 1.14 cm of rain in the last 30 days and there is less than .97 cm of rain in the coming 7 days and the current maximum temperature is greater than 92% of the maximum annual temperature, leaf senescence begins. Second, if it is past the middle of the year (day 182) and the three day

average minimum temperature is less than the annual average minimum temperature, senescence will also begin.

Once the prephenology is defined, the daily looping through the meteological data begins.

Daily Meteorology and Soil Temperature (daymet.c):

The daily meteorology routine populates the daily meteorology structure and also calculates several new variables from the input met data. Soil temperature is assumed to be the 11 day running weighted average of daily average temperature. The daytime and nighttime average temperatures are calculated as (Running and Coughlan 1988):

(10) tday = .045 * (tmax – tavg) + tavg

(11) tnight = (tday + tmin) / 2.

Soil temperature is then further corrected using the difference between the day's soil temperature and the average air temperature for the full met data record such that if there is snow water present:

(12) tsoil = tsoil + [.83 * (TavgAirTotal – tsoil)]

and if there is no snow water then:

(13) tsoil = tsoil + .2 * (TavgAirTotal – tsoil).

This correction is applied as snow will insulate the soil and help it to retain heat and in general soil retains more heat than the air even in the absence of snow.

Apply Prephenology to Daily Fluxes (phenology.c):

The daily phenology routine transfers C and N from transfer pools into new tissue pools if the current day is in the growing season. Litterfall is allocated differently for different PFTs. Evergreen PFTs have litterfall everyday of the year. Deciduous PFTs litterfall occurs with a linearly ramping rate starting at 0 such that all live fine roots and leaves are removed by the end of the litterfall period. The litterfall routine moves C and N from the fine roots and leaves to the four litter compartments in the proportions specified in the .epc file at the rates as defined above (see Appendix A). Based on the leaf litter C to N ratio, this routine also calculates the amount of retranslocated N that is removed from leaves before they senesce. Live and dead stem and live and dead coarse roots also have daily turnover rates as defined in the epc file.

Partition Leaf C into Sun and Shade LAI and Partition Incoming Radiation (radtrans.c):

The first step in partitioning incoming radiation is to partition the leaf carbon into sun and shade leaves. First, the whole canopy projected LAI is calculated using the user defined average SLA multiplied by the leaf C. The all-sided LAI is then found by multiplying the user supplied all-to-projected LAI ratio by the calculated projected LAI. The projected LAI for the sun and shade leaves are calculated based on Jones (1992) assuming only horizontally oriented leaves:

(14) $SunPLAI = 1 - e^{-TotalPLAI}$

(15) $ShadePLAI = TotalPLAI - SunPLAI$

Thornton (1998) explains that it is appropriate to ignore leaf angle at a daily time scale as it is an approximate integration over the full day. The SLA for sun and shade leaves follows using the user supplied ratio of shaded to sunlit SLA.

With the sun and shade leaves defined, the calculation of SRAD and PAR absorption can proceed. The logic of this process is outlined in figure 2. First the albedo effect (from the ini

file) is deducted from the incoming SRAD. Then, using Beer's Law, the SRAD absorbed by the canopy is calculated as:

(16) SWabs = (SRAD – albedo effect) * $(1 – e^{-k})$

where k is the user supplied canopy light extinction coefficient. The SRAD transmitted through the canopy (not absorbed) is simply:

(17) SWTrans = (SRAD - albedo effect) – SWabs.

The absorbed PAR is calculated similarly except that albedo is 1/3 as large for PAR because less PAR is reflected than SRAD (Jones 1992). PAR and SRAD absorbed are then further partitioned into the absorption by sun and shade leaves as:

(18) SWabsSun = k * (SRAD-albedo effect) * SunPLAI

(19) SWabsShade = SWabs – SWabsSun

These quantities are then scaled by the PLAI (projected LAI) of sun and shade leaves to get per PLAI values. PAR absorbed by sun and shade leaves is calculated similarly. The final step is to convert the radiation values that are in W/m^2 into $umol/m^2/s$ so that they can be used in later model steps.

Precipitation Routing (prcp_route.c):

With all-sided LAI known from the previous function, the canopy interception of precipitation can be calculated. The interception is a simple function of the incoming precipitation multiplied by the user defined interception rate and the all-sided LAI. No snow fall interception is modeled. Non-intercepted water is added to the soil water pool as as it is considered through fall.

Snow Water (snow_melt.c:)

As mentioned above, there is no canopy snow interception. There are several sources that help define the amount of melted snow each day as well as the amount of sublimated snow (Running and Coughlan 1988; Marks, Dozier et al. 1992; Coughlan and Running 1997). If the average daily temperature is greater than 0°C, then:

(20) $$\text{Snowmelt} = .65 * \text{tavg} + \frac{incident\ radiation\,(kJ / m^2 / day)}{latent\ heat\ of\ \ fusion\,(kJ / kg)}$$

where the latent heat of fusion is 335. If the temperature is less than 0°C, then:

(21) $$\text{Snow Sublimation} = \frac{incident\ radiation\,(kJ / m^2 / day)}{latent\ heat\ of\ \ sub\lim ation\,(kJ / kg)}$$

where the latent heat of sublimation is 2845.

Penman-Monteith Equation (within canopy_et.c):

One of the most important processes in ecosystems, and hence in BBGC, is the evaporation of H_2O. Evaporation occurring from within leaves into the atmosphere through stomata is referred to as transpiration. The sum of ecosystem evaporation and transpiration is collectively called evapotranspiration. The Penman-Monteith equation (PME) is a general equation that relates the incoming radiation, vapor pressure deficit (VPD), the density of air, the specific heat of air, and the resistances to sensible heat flux and water vapor flux to the loss of latent heat by evaporation (Waring and Running 2007; Monteith and Unsworth 2008). For plant leaves, the Penman-Monteith equation considers the leaf level conductance to water vapor which is based on the stomatal conductance to water and the leaf conductance to sensible heat which is equal to the boundary layer conductance.

Essentially, the PME uses characteristics of a particular surface (e.g. – surface resistances) and current meteorological data (e.g. – wind, incoming radiation, VPD, air temperature, air pressure) to calculate an instantaneous heat balance of an object. This heat balance is a rate of heat loss or gain. For wet objects, the rate of heat loss can be used to find the rate of evaporation. The rate of evaporation is equal to the incoming heat radiation minus the loss of sensible heat by convection (long-wave radiation).

(22) $\lambda E = R_n - C$ (Monteith and Unsworth 2008)

For transpiration in leaves, the rate of evaporation is also a function of the amount of coupling between the canopy resistance to water vapor flux and VPD. This coupling varies based on different individual species' responses to water stress. When the canopy is highly coupled to VPD (i.e. there is strong stomatal control of transpiration), the rate of evaporation is governed by the boundary layer conductance. When the canopy is not coupled or loosely coupled to atmospheric VPD (i.e. there is not a strong stomatal control of transpiration), the rate of evaporation is controlled more by stomatal conductance.

In BBGC, the PME is used to calculate soil H_2O evaporation, the evaporation of canopy intercepted water, and the transpiration of water from leaves. The PME is found in the penmon function within the canopy_et.c subroutine. The PME is called by the soil evaporation routine and the canopy evapotranspiration routine. Using the results from the PME calculations, BBGC also is able to compute the stored soil water and the water that leaves the system as outflow.

The PME equation uses many parameters. Some are user supplied, some are assumed constant, and some are allowed to vary. The actual evaporation is also a function of the input meteorological data. The PME used in BBGC requires the following inputs: 1) air temperature (°C), air pressure (Pa), VPD (Pa), incident radiant flux density (W/m^2), resistance to water vapor

flux (s/m), and resistance to sensible heat flux (s/m) (resistance = 1/conductance). The air temperature, VPD, and air pressure are input meteorological data or derived from the constant site parameters and the meteorological data. The incident radiant flux density (incoming radiation) is calculated for whatever surface the PME is being used on to calculate evaporation. This is a function of the shade and sun leaf LAI and the partitioning of incoming radiation. Hence, this varies with user supplied parameters that specify the sun to shade specific leaf area (SLA) ratio, the light extinction coefficient, the average SLA, and the model calculated leaf C. The resistance to water vapor and sensible heat are dependent on the location of evaporation. For water evaporating off of leaves, the resistance to sensible heat and the resistance to water vapor are both equal to the leaf boundary layer resistance. For transpiration, the resistance to water vapor is a function of the boundary layer, cuticular, and stomatal conductances while the resistance to sensible heat is the boundary layer resistance. For soil water evaporation, the sensible heat and water vapor resistances are both equal to a temperature and pressure corrected constant bare soil evaporation resistance based on data collected over bare soil in south-west Niger (Wallace and Holwill 1997). The boundary layer, cuticular, and stomatal conductances are all user specified parameters. Transpiration's water vapor resistance is different than just the boundary layer resistance because water vapor must pass through the stomata and/or cuticle for evaporation of internal leaf water to occur.

McNaughton and Jarvis (1983) modified this equation to account for the coupling effects that climate variables like wind, VPD and canopy architecture have on evaporation. The modified Penman-Monteith equation requires the following inputs: 1) air temperature (°C), air pressure (Pa), VPD (Pa), incident radiant flux density (W/m^2), resistance to water vapor flux

(s/m), and resistance to sensible heat flux (s/m) (resistance = 1/conductance). The penman function found within canopy_et.c first calculates the density of air as a function of temperature:

(23) $\rho = 1.292 - (0.00428 * \text{tair (°C)})$

The resistance to radiative heat transfer through the air (r_R) is then calculated:

(24) $$r_R = \frac{\rho * c_p}{4 * SBC * tempK^3}$$

where c_p is the specific heat of air (J/kg/°C) and SBC is the Stefan-Boltzmann constant ($W/m2*K^4$). The combined resistances to convective and radiative heat in parallel are then calculated as:

(25) $$r_{HR} = \frac{r_H * r_R}{r_H + r_R}$$

where r_H is the resistance to convective heat transfer (i.e. - resistance to sensible heat flux, boundary layer resistance in leaves). The latent heat of vaporization is calculated as:

(26) lhvap = 2.5023E6 - 2430.54 * tair (°C)

The next step is to find the rate of change (slope) of the saturation vapor pressure with temperature ($(\partial e_s(T)/\partial T$). This is done to find the approximate relationship between saturation vapor pressure and the unknown temperature at the site of evaporation (Monteith and Unsworth 2008). In BBGC this is done by first estimating the saturation vapor pressures at two temperatures ±0.2 °C from tair.

(27) SVP1 = 610.7 * exp(17.38 * (tair + .2) / (239.0 + (tair + .2));

(28) SVP2 = 610.7 * exp(17.38 * (tair - .2) / (239.0 + (tair - .2));

(29) s = slope = (SVP1 – SVP2)/[(tair + .2) – (tair - .2)]

With these quantities calculated the final evaporation rate (W/m^2/sec) can be calculated:

$$(30) \quad e = \frac{s * RAD + \left(\frac{\rho * c_p * VPD}{r_{HR}} \right)}{\left(\frac{\rho * c_p * r_V}{AirPa * \varepsilon * r_{HR}} \right) + s}$$

where ε is the ratio of molecular weights of water vapor and air (0.622) and r_V is the resistance to water vapor flux (r_V varies based on the surface being evaporated from. For leaf intercepted water and soil water, r_V is the boundary layer resistance. For transpiration, r_V is a function of stomatal, cuticular, and boundary layer resistance.). This evaporation rate is then multiplied by a time to find the quantity of evaporated water.

Soil Water Evaporation (baresoil_evap.c):

Soil water evaporation is calculated by scaling the potential evaporation, as calculated by the PME, by a quantity determined by the days since the last rain event. The logic behind this is that the soil will more tightly hold onto water as water becomes more scarce and therefore less evaporation will be possible (Taiz and Zeiger 2006). As opposed to leaves that have separate resistances to convective and water vapor flux due to stomata, the soil resistance to vapor flux is the same as the resistance to sensible heat flux. The first step in finding this resistance is calculating a factor to correct conductance to sensible heat based on temperature and pressure (Jones 1992):

$$(31) \quad r_{corr} = \frac{1}{\left\{ \left[(tday + 273.15) / 293.15 \right]^{1.75} \right\} * \frac{101300}{AirPa}}$$

A reference soil resistance taken over bare soil in the tiger bush of Niger is then scaled by this factor to find the boundary layer resistance to sensible heat flux (Wallace and Holwill 1997). The PME is then used to calculate the potential evaporation.

If at least as much precipitation reaches the soil as the potential evaporation, then the actual evaporation is 60% of the calculated potential evaporation. Otherwise, the days since the last rain event (DSR) counter is incremented. The realized proportion of evaporation is then calculated as:

(32) Actual Evaporation = $\frac{0.3}{DSR^2}$ * potential evaporation

Soil Matric Potential (soilpsi.c):

The soil matric potential is a function of the soil volumetric water content (VWC) and the saturated soil matric potential:

(33) psi_actual = psi_sat * (VWC / VWC_sat) ^ b

where VWC_sat and psi_sat were precalculated based on soil texture and b, also precalculated is:

(34) b = -3.1 + 0.157*clay – 0.003*sand

This essentially scales the known soil saturation potential by the ratio of current volumetric water capacity to saturation water capacity. The current volumetric soil water is calculated as:

(35) VWC = soilW (kg/m^2) / (1000 * soilDepth)

where 1000 is the density of water (kg/m^3).

Maintenance Respiration (maint_resp.c):

Maintenance Respiration (MR) in BBGC uses a Q_{10} relationship with temperature as well as the N content of tissues to estimate this rate. A Q_{10} relationship means that for every 10°C change in temperature, there is a Q_{10} factor change in respiration. The Q_{10} relationship assumes a reference temperature of 20°C and therefore all temperatures are scaled by this value. The N relationship is linear with MR being 0.218 kgC/kgN/day (Ryan 1991). MR is calculated for sun

and shade leaves and partitioned into night and day respiration since daytime respiration is needed to calculate assimilation. MR is then calculated for fine roots, live stem, and live coarse roots –

(36) $MR = 0.218 * N * Q_{10}^{(temp-20)/10}$

Once the total day and night MR is calculated for sun and shade leaves, these rates are scaled to be per PLAI and per second.

Canopy Evapotranspiration (canopy_et.c):

The canopy evapotranspiration (ET) routine calculates both canopy intercepted water evaporation and the leaf transpiration. To make the transpiration calculation, this function also calculates the leaf level conductance and the stomatal conductance. Cuticle and boundary layer conductances are user defined in the epc file as well as the maximum rate of stomatal conductance (see Appendix A). To simulate the drivers of stomatal closure, BBGC scales the maximum stomatal conductance by a series of multipliers between 0 and 1 for: 1) photosynthetic photon flux density, 2) soil water potential, 3) minimum temperature, and 4) VPD. As with the soil evaporation function, the first step is to calculate a conductance correction factor for the current air pressure and temperature (see equation 30 except that conductance is not a quotient dividing 1 but only the divisor). The equations to calculate the four multiplier functions are below and figure 9 shows these relationships graphically (Korner 1995; Conklin and Neilson 2005). The limiting values are drawn from the default evergreen needleleaf epc file –

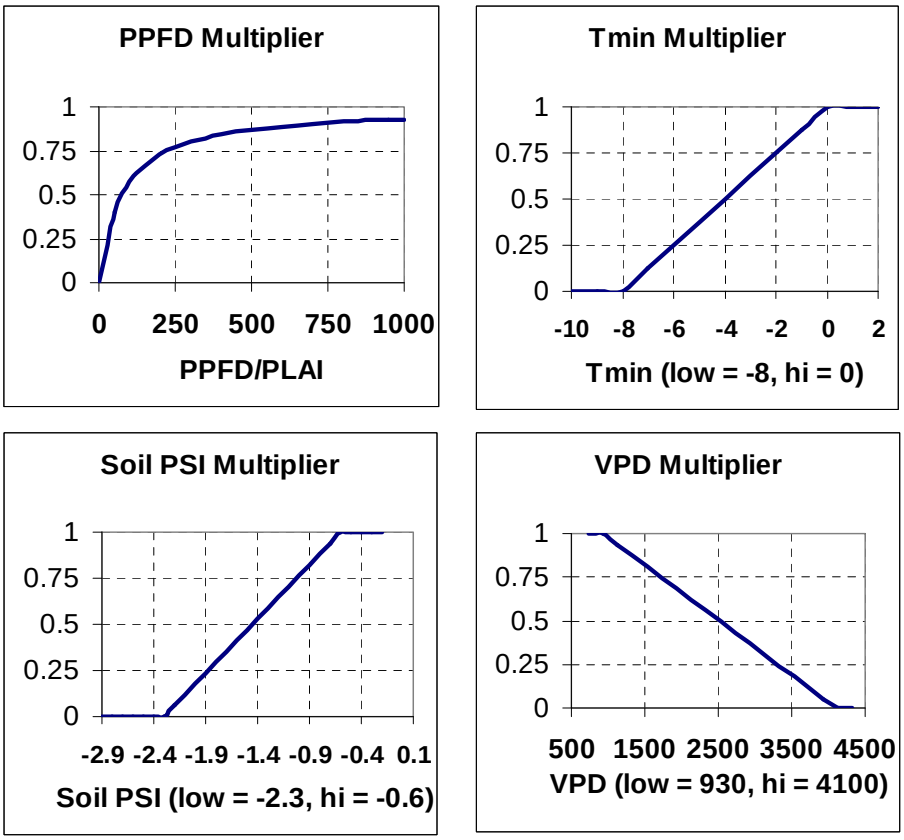

Figure 9: Graphical representations of the multiplier functions for stomatal conductance for a typical evergreen needleleaf forest.

(37) $$\text{M\_PPFD} = \frac{PPFD / PLAI}{75 + PPFD / PLAI}$$

(38) $$\text{M\_TMIN, M\_VPD, and M\_SOILPSI} = \frac{current\,value - MinSOE}{MaxSOE - MinSOE}$$

where MinSOE (minimum stomatal opening endpoint) represents the lower limit below which there is full stomatal closure and MaxSOE is the upper limit above which there is maximum stomatal opening.

The final multiplier is the product of these four multipliers. This final multiplier and the conductance correction factor are then applied to the maximum stomatal conductance supplied by the user. Leaf conductance to water vapor is assumed equal to the boundary layer conductance as is the leaf conductance to sensible heat. The total leaf conductance to water

vapor is then calculated by combining the stomatal, boundary, and cuticle level conductances in parallel for both sun and shade leaves.

(39) $$G_{T\_WV} = \frac{g_{bl}(g_s + g_c)}{g_{bl} + g_s + g_c}$$

Once all of these values are found, this function passes the leaf level information to the PME to calculate the evaporation rate of the canopy intercepted water (CIW). For the CIW, there is only boundary layer and sensible heat resistances, not a stomatal component. The resulting rate is then divided into the amount of CIW to calculate the amount of time required to evaporate this pool. The remaining time (if there is any) is then used to calculate the amount of transpiration. All calculations are done separately for sun and shade leaves. The sun and shade leaf calculated conductances are also stored for later use in the photosynthesis routine.

There are several assumptions that are made by BBGC when using the PME. First, one evaporation rate is calculated per day and expanded by multiplying by the appropriate amount of time of evaporation. This daily scale evaporation does not account for the changes that occur within a day in terms of incident radiation, VPD, and temperature. Another set of assumptions are the shapes of the scaling factor curves seen below in **Figure 9**.

The multipliers could all follow non-linear curves that define the stomatal opening. For example, it may be that stomatal conductance remains mostly constant through a range of VPDs and then stomata rapidly close as a limiting VPD is reached. Lastly, this approach assumes that all plants will eventually close their stomata given some set of limiting environmental conditions. However, as McDowell et al. (2008) showed, some anisohydric species will retain stomatal opening even under great stress.

These scaling factors control the coupling of transpiration between the canopy and VPD. The scaling of stomatal conductance is a critical assumption that regulates transpiration and also

C fixation. Based on the extensive history of use and validation of BBGC, this scaling approach to stomatal conductance seems justified. However, some species have developed alternative mechanisms of water intake and stomatal control (e.g. – Redwoods absorb fog water through their stomata). In these cases, the model logic may not be appropriate. Additionally, as mentioned above, applying the stomatal conductance scaling factors to anisohydric species may also incorrectly represent the behavior of these plants' stomatal opening. Therefore, if using BBGC to represent a specific plant community, it is important to consider these assumptions and the underlying model logic when parameterizing the model and interpreting the model results.

Photosynthesis (photosynthesis.c):

The photosynthesis function in BBGC is the single most important part of the model in that it mechanistically represents the ecosystem addition of C. The basis of this photosynthesis code is the DePury and Farquhar two-leaf model of photosynthesis (Farquhar, Caemmerer et al. 1980; De Pury and Farquhar 1997). Additionally, the enzyme kinetics built into this model are based on Woodrow and Berry (1988). The rate of photosynthesis is sensitive to the N content of leaves, the portion of N in Rubisco, and the temperature as this controls the enzyme kinetics. Photosynthesis also depends on the amount of absorbed PAR, the calculated MR, and the difference between the internal and external partial pressure of CO_2. As always, all steps are done separately for sun and shade leaves.

Photosynthesis is the process where CO_2 and H_2O molecules are combined using energy from the sun to generate simple sugars. The general reaction is:

$6CO_2 + 6H_2O + \text{Light Energy} \rightarrow C_6H_{12}O_6 + 6O_2$ (Taiz and Zeiger 2006)

BBGC represents this reaction using three separate equations to represent three different controls on the rate of photosynthesis. The first equation (equation (41)), represents the CO_2 diffusion constraints of photosynthetic rate. This is the rate that CO_2 can enter the leaf and is a function of stomatal opening and the difference between the atmospheric CO_2 pressure and the leaf internal CO_2 pressure. As described above, the stomatal opening is a function of several scaling factors that are both user-defined and model constants.

The second equation used to constrain the rate of photosynthesis represents the carboxylation rate control of the photosynthesis reaction (equation (42)). Carboxylation is the process where 3 CO_2 molecules are fixed to a carbon skeleton by the Rubisco enzyme (Taiz and Zeiger 2006). The rate at which this occurs depends on the enzyme kinetics that governs how fast CO_2 can be bound to RuBP (the carboxylation substrate) by the Rubisco enzyme. When the rate of photosynthesis is carboxylation limited, this means that the availability of the RuBP substrate is not limiting and instead photosynthesis is limited by the concentration of CO_2 (Lambers, Chapin et al. 2008). This also depends on the leaf internal pressure of CO_2 and in this way is related to equation (41) discussed below. The Rubisco enzyme is also sensitive to the amount of O_2 in the cell as Rubisco can also bind to O_2 instead of CO_2. This also constrains the rate of carboxylation. The carboxylation rate is also dependent on the temperature as all enzyme activity varies with temperature. Lastly, the assimilation is governed by the amount of leaf N in Rubisco as this determines the quantity of Rubisco available to catalyze the carboxylation reaction. The amount of Rubisco is therefore dependent on the user-defined fraction of leaf N in Rubisco and the user-defined C:N ratio of leaves (both defined in the epc file).

The third and last equation used to constrain the rate of photosynthesis represents the electron transport limitation of RuBP regeneration (equation (43)). When the leaf concentration

of CO_2 is not limiting the rate of photosynthesis (i.e. – there is enough CO_2), the assimilation rate is governed by how fast RuBP can be regenerated to fix new CO_2 molecules. This is the electron transport limitation. The figure below is taken from the Lambers et al. (1998) book. It shows the RuBP saturated (CO_2 limited) carboxylation limited portion of the assimilation curve (A(c) – A(v) in equation (42) below) and the RuBP limited due to the rate of electron transport RuBP regeneration portion of the assimilation curve (A(j) in the figure below and equation (43) below). The rate of assimilation is the minimum of these two equations (the solid lines). In BBGC, equation (41) is solved for C_i and substituted into equations (42) and (43) to create two quadratic equations that can both be solved. The smaller of the two resulting solutions is then used as the rate of C assimilation.

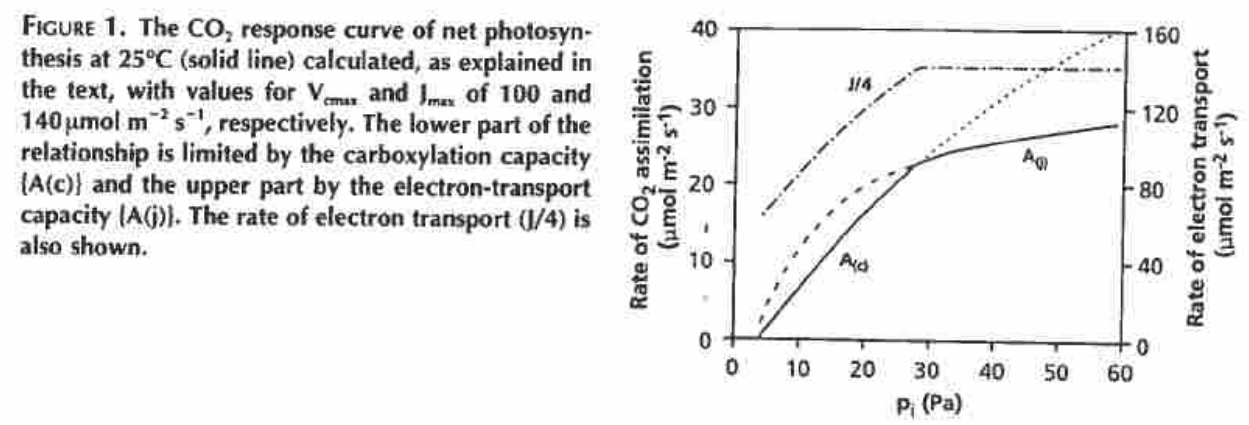

FIGURE 1. The CO_2 response curve of net photosynthesis at 25°C (solid line) calculated, as explained in the text, with values for V_{cmax} and J_{max} of 100 and 140 μmol m^{-2} s^{-1}, respectively. The lower part of the relationship is limited by the carboxylation capacity {A(c)} and the upper part by the electron-transport capacity {A(j)}. The rate of electron transport (J/4) is also shown.

The enzyme kinetics built into this model are based on Woodrow and Berry (1988). The rate of photosynthesis is sensitive to the N content of leaves, the portion of N in Rubisco, and the temperature as this controls the enzyme kinetics. Photosynthesis also depends on the amount of absorbed PAR, the calculated MR, and the difference between the internal and external partial pressure of CO_2. As always, all steps are done separately for sun and shade leaves.

The first step in this calculation is to convert the already calculated stomatal conductance to water vapor to a conductance for CO_2 and to convert this into the units used by the

photosynthesis (PSYN) submodel (m/s to umol/m^2/s/Pa). This conversion is seen below (Nobel 1991; Jones 1992):

(40) $$g_{mTc} = \frac{10^6 * g_{Tv}}{1.6 * R(Tday + 273.15)}$$

where R is the universal gas constant, g_{Tv} is the leaf scale conductance to transpired water, tday is the daytime temperature, and 1.6 is the ratio of the molecular weights of water vapor to CO_2.

Once the leaf level conductance to CO_2 is known, the main PSYN routine is begun. The core logic of the PSYN routine consists of three main equations (Farquhar, Caemmerer et al. 1980):

(41) $$A_{(v \text{ or } j)} = g_{mTc} * (C_a - C_i)$$

(42) $$A_v = \frac{V_{c\max}(C_i - \Gamma^*)}{C_i + K_c * \left(1 + \frac{O_2}{K_o}\right)} - MR_{leafday}$$

(43) $$A_j = \frac{J * (C_i - \Gamma^*)}{4.5 * C_i + 10.5 * \Gamma^*} - MR_{leafday}$$

where C_a is the atmospheric concentration of CO_2 (Pa) and C_i is the intercellular concentration of CO_2 (Pa), Γ^* (Pa) is the CO_2 compensation point in the absence of leaf MR, K_c and K_o are the kinetic constants for rubisco carboxylation and oxygenation scaled by the temperature using a Q_{10} relationship, O_2 is the atmospheric concentration of O_2 (Pa), $MR_{leafday}$ is the daytime leaf maintenance respiration on a PLAI basis, and J is the maximum rate of electron transport. Each of these variables will be discussed in more detail below.

Equation (40) represents the diffusion limitation of CO_2 on assimilation. Equation (41) represents the carboxylation limitation on the rate of assimilation. Equation (42) represents the electron transport rate of substrate regeneration limitation on assimilation. Thornton (1998)

explains that by solving equation 40 for C_i and then substituting this value back into equations 41 and 42, two quadratic equations are created that can be solved. The smaller of the two results when solving both equations is then used as the actual assimilation rate.

There are several steps required before the quadratic roots are ready to be calculated. To begin, J must be calculated. J is a function of the maximum rate of carboxylation (V_{cmax}) (Wullschleger 1993):

(44) $J_{max} = 2.1 * V_{cmax}$

where V_{cmax} is a function of the N per unit PLAI in the shade and sun leaves as well as the fraction of leaf N in rubisco and the activation potential of rubisco as defined by the Woodrow and Berry (1988):

(45) $V_{cmax} = N_{sun\ or\ shade\ leaves} *$ fraction of leaf N in rubisco * 7.16 * ACT

The N content of sun and shade leaves is a function of the user defined ratio of C:N in leaves:

(46) $N_{sun\ of\ shade\ leaves} = \frac{1/(C:N\ leaf)}{SLA_{sun\ or\ shade}}$

The fraction of leaf N in rubisco is a user supplied parameter. 7.16 is the weight proportion of rubisco relative to its N content (Kuehn and McFadden 1969; Kuehn and McFadden 1969; Fasman 1976), and ACT is the activity of rubisco scaled by temperature, $[O_2]$ and $[CO_2]$. The first step in calculating the enzyme kinetics of rubisco is to calculate the $[O_2]$ assuming it is 21% of the atmosphere by volume:

(47) $O_2 = 0.21 *$ atmospheric pressure (Pa)

Next, the rubisco activity can be calculated following its Michealis-Menten dynamics for O_2 and CO_2 (the CO_2 calculations vary depending on whether it is <= or > 15°C). All variables with a 25 subscript are the constant values at 25°C that are being scaled based on temperature.

The different Q_{10} values relate to the different Q_{10} relationships with temperature for each of these reactions.

(48) $K_o = K_{o25} * Q_{10K_o}^{\left(\frac{Tday-25}{10}\right)}$

(49) If tday > 15°C:

a. $K_c = K_{c25} * Q_{10K_c}^{\left(\frac{Tday-25}{10}\right)}$

b. $ACT = ACT_{25} * Q_{10ACT}^{\left(\frac{Tday-25}{10}\right)}$

(50) If tday <= 15°C:

a. $K_c = \frac{K_{c25} * 1.8Q_{10K_c}^{\left(\frac{Tday-15}{10}\right)}}{Q_{10K_c}}$

b. $ACT = \frac{ACT_{25} * 1.8Q_{10ACT}^{\left(\frac{Tday-15}{10}\right)}}{Q_{10ACT}}$

With ACT, K_c, and K_o calculated, Γ^* (Pa) (the CO_2 compensation point in the absence of leaf MR) and V_{cmax} can now be found. Γ^* is:

(51) $\Gamma^* = 0.105 * K_c * \frac{O_2}{K_o}$

All of these calculations are done for both sun and shade leaves and at the end of the day the total assimilate is the sum of the sun and shade leaf assimilation. The assimilated C that is calculated is actually a rate of C assimilation per second scaled by PLAI. This value must be multiplied by the length of daylight and then multiplied by the sun or shade PLAI to find the total C assimilated. This assimilate is converted from umol/m^2/sec to kg/m^2/day and then placed into the Cpool for future allocation.

Given these assumptions and the model logic described above, users can have a high degree of confidence in model runs with the major caveat that the user specified parameters need to be appropriate for a given system. As stated before, when correctly parameterized to represent a system, this model has been validated in many diverse systems across the globe. However, this parameterization can at times feel quite arbitrary and there are a few parameters that really are more "tuning dials" than real physiological system components (e.g. – soil depth). In this sense, the model has made some conceptual abstractions to allow a user to represent a given system with more certainty despite the underlying lack of clarity in the model representation of actual physiological processes. This is always a tradeoff that any model is forced to make. In practice, what this means is that BBGC can be used across many diverse systems with appropriate calibration and in this sense, with a broad enough perspective, it is very valuable. However, to explore specific physiological concepts that relate to specific plant communities, model modifications may be required to more accurately represent the underlying mechanisms at work.

N Deposition and Fixation (within bgc.c):

The N deposition and N fixation are added to the soil mineral N pool.

H_2O outflow (outflow.c):

Outflow occurs in two cases. First, for fast outflow, if the incoming precipitation causes the soil water to exceed the saturated soil water capacity, outflow is the difference between the soil water capacity and the remaining water over capacity. Second, for slow outflow, if the soil water is above field capacity, then the outflow occurs at an exponentially decaying rate each day

with half of the water that is above field capacity draining. If the soil water is less than field capacity, there is no outflow.

Decomposition (decomp.c):

The BBGC decomposition model is similar to the canopy evapotranspiration model logic in that both use multiplicative scalars to adjust maximum rates. BBGC uses a compartment system to represent the litter pools and the soil organic matter pools. The four litter pools are broken out based on the dominant material found in them – labile carbon in the form of basic sugars, starches, and proteins, unshielded cellulose, shielded cellulose with some lignin fraction, and a lignin pool. A CWD pool accepts inputs from the stem and coarse root pools and subsequently adds to the litter pools after it is broken into smaller pieces over time. Only fine roots and leaves initially enter the litter pools. The fractions of leaf and fine root litter, as well as the fractions of dead wood that are cellulose and lignin, are all user supplied and also are used to calculate the shielded and unshielded fractions of the cellulose litter pool.

Taken from Thornton and Rosenbloom (2005). Changes reflect BiomeBGC v4.2 code.

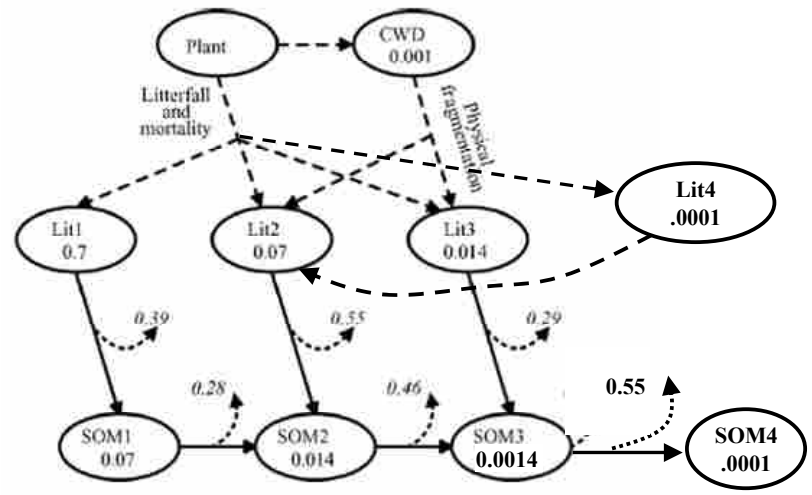

Fig. 1. Converging cascade model of litter and soil organic matter decomposition. The model includes three litter pools (Lit1, Lit2, and Lit3, see text) and three soil organic matter pools (SOM1, SOM2, and SOM3). Solid arrows indicate decomposition pathways. Curved dashed arrows indicate heterotrophic respiration fluxes associated with each decomposition pathway, with the respiration fraction (*r*, see text) for each pathway shown in italics. Straight dashed arrows indicate the physical transfer of plant biomass to litter pools and to the CWD pool, as well as the transfers out of the CWD pool into Lit2 and Lit3 as the result of physical fragmentation. Base decomposition rates, or physical fragmentation rate in the case of CWD (*k*, 1/day) are shown below the name for each pool.

In this diagram, Lit4 is the shielded cellulosic litter. It is broken down at the same rate as the lignin pool (Lit3) and is transferred to the unshielded cellulosic litter pool (Lit2) as it decomposes. SOM4 is the recalcitrant soil organic matter pool (humus).

Figure 10: BiomeBGC litter and soil pools and fluxes

Upon decomposition, the litter pool fractions enter into the soil organic matter pools. The maximum rate constants for decomposition and biomass loss through heterotrophic respiration (HR) are all defined as constants in BBGC and were defined based on a literature review of C^{14} decomposition studies by Thornton (1998). Figure 10 is a detailed diagram of the decomposition dynamics in BGC (Thornton and Rosenbloom 2005). The rates seen in this diagram are adjusted based on the temperature (Lloyd and Taylor 1994) and the availability of water in the soil as seen in figure 11 (Orchard and Cook 1983; Andren and Paustian 1987). There is no decomposition if the temperature is below -10°C.

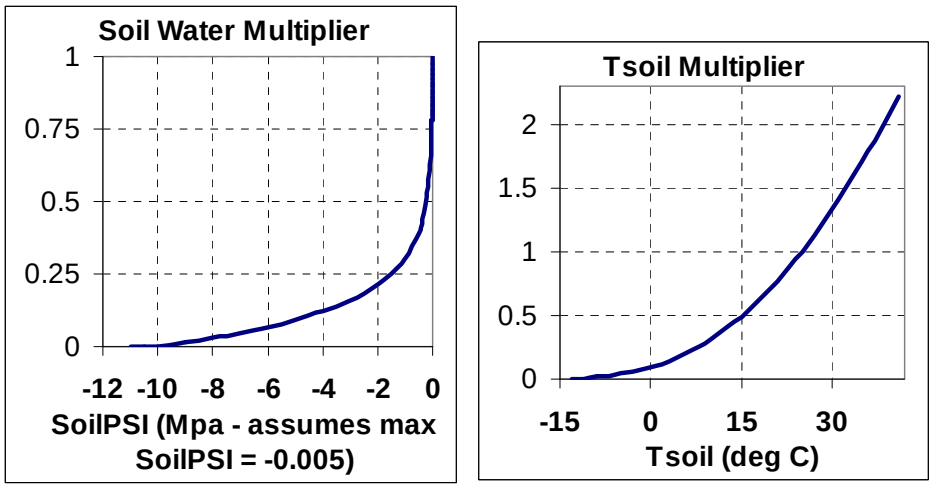

Figure 11: Decomposition rate scalar function graphs for temperature and water availability.

(52) $$t\_scalar = e^{4.344692-\frac{1}{soilTK-227.13}}$$

(53) $$w\_scalar = \frac{\ln(\min soilpsi / current\ soilpsi)}{\ln(\min soilpsi / \max soilpsi)}$$

where soilTK is the soil temperature in Kelvin. The final scalar is the product of the water and temperature scalar.

Once the rate scalar has been calculated, all of the litter compartments C:N ratios are calculated based on the known C and N amounts in each litter pool. The constant decomposition (and fragmentation for CWD) rates are then adjusted by multiplying them by the rate scalar. The adjustment is used since the original rates were found under well watered conditions with a soil temperature of 25°C. With these adjusted decomposition rates, the first step is to fragment the CWD and apportion it to the litter pools (except for the labile pool).

The non-N limited decomposition of the litter and soil compartments are then calculated for each pool. These are all potential decomposition rates and the actual decomposition will be scaled based on the competing plant N demand during the allocation routine. Each pool has a similar set of calculations applied to it. First, the potential C loss is found by multiplying the

adjusted decomposition rate by the amount of C in the litter pool. The ratio of the C:N ratio of the given soil pool that accepts the decomposed matter to the C:N ratio of the litter pool that is losing C is then found. The same step is performed for the soil to soil fluxes as well. The potential N immobilization by microbial decomposition is then calculated from these data:

$$\text{(54)} \quad \text{potential N immobilization} = \text{potential C loss} * \frac{1 - HRfrac - \frac{C:N\ donor\ pool}{C:N\ accepting\ pool}}{C:N\ accepting\ pool}$$

As a last step, the sum of all of the potential N immobilization for the fluxes between the pools is calculated. In some cases, given the C:N ratios of the pools, this may be a negative number in which case this is considered potential N mineralization.

Allocation (daily_allocation.c):

With the amount of potential C assimilate known from the photosynthesis process, and the amount of microbial N demand known from the decomposition process, it is now possible to allocate C assimilate and account for N limitation. The BBGC allocation logic is built around the concept that all new allocation is constrained by leaf C allocation (Waring and Pitman 1985; Waring and Running 2007). There are four user defined allocation ratios for woody plants: 1) new fine root C to new leaf C, 2) new stem C to new leaf C, 3) new live wood C to new total wood C, and 4) new coarse root C to new stem C. There is also one user defined proportion that defines how much of the assimilated C should be set aside for next year's growth. These ratios then are used to define how C is allocated throughout the plant.

The first step in the allocation is to find the C available to allocate. This is Gross Primary Production (GPP) minus the MR of all live tissues. In some cases, the cpool variable that holds the daily assimilated C and acts as a bank account for plant C demand can be negative. This

could occur, for example, if growing conditions prevented assimilation but MR was still occurring. In these cases, a rate of repayment of this cpool deficit is calculated such that the deficit is gone in one year and the available C for allocation is first allocated to alleviate the deficit at this rate. Any leftover C can then be allocated to plant growth.

The next step is to calculate the amount of C needed per unit of leaf C growth in all of the other pools based on the allocation fractions described above as well as the amount of C needed for growth respiration based on this allocation. BBGC assumes a constant rate of GR for all tissue growth = 30% of the total C used for new tissue. The C allometry calculation is followed by calculating the associated N needed to grow the plant based on this C allometry and the respective C:N ratios of the different plant pools.

(55) C needed / unit leaf C = $(1 + GR\%) * \left(1 + \frac{frC}{leafC} + \frac{stemC}{leafC} * \left(1 + \frac{crootC}{stemC}\right)\right)$

(56) N needed / unit leaf N =

$$\frac{1}{LeafC:N} - \frac{\frac{frC}{leafC}}{frC:N} + \frac{\frac{stemC}{leafC} * \frac{NLwoodC}{NwoodC} * \left(1 + \frac{crootC}{stemC}\right)}{LiveWoodC:N} + \frac{\frac{stemC}{leafC} * \left(1 - \frac{NLwoodC}{NwoodC}\right) * \left(1 + \frac{crootC}{stemC}\right)}{DeadWoodC:N}$$

where NLwood is new live wood and Nwood is total new wood both live and dead and croot is coarse root.

Dividing equation 54 by equation 55 gives the C:N ratio of newly allocated tissue. Multiplying this by the C available for allocation tells us the plant N demand given the potential C to allocate. The total N demand in the system is also known since the potential microbial N demand from the decomposition function is now known. If the total system N demand is less than the available soil mineral N (SMN) pool actual allocation is equal to potential allocation and actual decomposition is also potential decomposition. If the available retranslocated N can

satisfy the plant N demand, the plant N demand quantity is shifted from the retranslocation pool to the npool to be allocated at the end of the daily loop. If not, the total daily N retranslocated is applied to plant N demand and the additional N needed is removed from the SMN pool. Half of any excess N is assumed mineralized each day and is added to a bulk denitrification flux (volatilization) to be deducted from the SMN pool.

If there is less SMN than total demand, the potential N immobilization flux (the potential microbial decomposition) is scaled based on the ratio of the potential immobilization to the total system N demand. The fraction of potential N immobilization (FPI) is then:

$$\text{(57)} \quad \text{FPI} = \frac{SMN * \frac{potential\ N\ immobilization}{total\ N\ demand}}{potential\ N\ immobilization} = \frac{actual\ N\ immobilization}{potential\ N\ immobilization}$$

Once the microbial N demand is partially satisfied in this way, the plant N demand is addressed by first looking at the available retranslocated N pool. If there is enough retranslocated N and SMN available, plant C allocation will not be limited by N and will proceed at its potential rate. If there is not enough retranslocated N to meet the plant N demand, the C assimilate available for allocation is proportionally reduced (for both shade and sun leaves – Wang et al. noted this and postulated that this should be accounted for by greater respiration costs instead of scaling the amount of assimilated C (Wang, Ichii et al. 2009)). The actual C allocated to new growth is then the actual N available for plants * the N:C allometry ratios defined in equations (54) and (55). The excess C is removed from the assimilated C for sun and shade leaves in proportion to their size relative to total GPP and this serves as another limit on plant photosynthetic capacity.

With the actual allocation amounts now known, BBGC moves the assimilated C and the associated N into the different tissue pools and storage pools for next year's growth. This function also then scales each potential decomposition flux based on the FPI calculated above.

One other point to note is that during spin-up runs, BBGC adds additional N to the system if N is limiting to speed the systems' attainment of its steady-state result.

Growth Respiration (growth_resp.c):

The GR function simply allocates the C to the GR pools and the storage GR pools for next year's growth. These quantities were already calculated in the daily allocation routine.

Update C, N, and H_2O state (state_update.c):

After the functions above have been completed, BBGC moves the fluxes identified to and from the C, N, and H_2O state pools. All of the functions above do not modify the actual states but rather update flux variables that are then modified in the state_update routine. On the last day of the year, the storage C and N pools are moved to become transfer pools that can then be used for the next year's growth.

N Leaching (nleaching.c):

N leaching only occurs if there was water outflow. If there was outflow, then 10% of the SMN is removed and considered leached.

Mortality (mortality.c):

The daily mortality fraction is user defined in the epc file and is applied to all plant pools both live and dead as well as the transfer and storage pools. This function partitions the mortality into the appropriate CWD and litter pools. The mortality function also applies any user specified fire mortality to plant C and N pools moving the specified daily proportion from the

pools into a fire sink pool that represents the loss of C and N to the atmosphere. The fire routine also assumes there will be some recruitment of CWD and therefore only applies 30% of the fire mortality loss to the CWD pool.

Mass Balance Check (check_balance.c):

Once all of the state variables have been updated by both the state update routine and the mortality function, the check balance function is called to insure the Principle of the Conservation of Mass. Table 1 shows all of the sources, outflows/removals, and storage pools for C, N, and H_2O that the model tracks and balances by testing whether the balance is <= 0.00000001. The daily balance is defined as:

(58) Balance = in – out – stores

The only balance that is not checked (but probably should be) is the radiation energy balance.

Table 1: BiomeBGC inputs, outputs, and storage for mass balance check.

	In	**Out**	**Storage**
H_2O (see figure 3)	Precipitation	Outflow Soil evaporation Snow sublimation Canopy evaporation Leaf transpiration	Soil water Snow water Canopy water
C (S and T refer to Storage and Transfer Pools) (see figure 4)	Sun leaf psyn Shade leaf psyn	Leaf MR and GR Fine root MR and GR Live stem MR and GR Live coarse root MR and GR Dead stem GR Dead coarse root GR Labile litter (litter 1) HR Cellulose litter (litter 2) HR Lignin litter (litter 4) HR Fast Soil (SOM1) HR Medium Soil (SOM2) HR Slow Soil (SOM3) HR Recalcitrant Soil (SOM4) HR Fire	Leaf C and S and T Fine root and S and T Live stem and S and T Dead stem and S and T Live Coarse Root and S and T Dead Coarse Root and S and T CWD Litter Pools 1-4 SOM Pools 1-4 Cpool

<u>N</u> (S and T refer to Storage and Transfer Pools) (see figure 4)	N fixation N Deposition	N leaching N volatilization (denitrification) Fire loss N	Leaf N and S and T Fine root and S and T Live stem and S and T Dead stem and S and T Live coarse root and S and T Dead coarse root and S and T CWD Litter pools 1-4 SOM pools 1-4 Soil Mineral N Retranslocated N Pool Npool

Appendix A: Required BiomeBGC Inputs

Ecophysiology Input File (.epc file):

```
ECOPHYS       ENF-cool (evergreen needleleaf forest - cool climate)
1             (flag)    1 = WOODY             0 = NON-WOODY
1             (flag)    1 = EVERGREEN         0 = DECIDUOUS
1             (flag)    1 = C3 PSN            0 = C4 PSN
1             (flag)    1 = MODEL PHENOLOGY   0 = USER-SPECIFIED PHENOLOGY
0            *(yday)    yearday to start new growth  (when phenology flag =
0)
0            *(yday)    yearday to end litterfall  (when phenology flag = 0)
0.3          *(prop.)   transfer growth period as fraction of growing season
0.3          *(prop.)   litterfall as fraction of growing season
0.25          (1/yr)    annual leaf and fine root turnover fraction
0.70          (1/yr)    annual live wood turnover fraction
0.005         (1/yr)    annual whole-plant mortality fraction
0.005         (1/yr)    annual fire mortality fraction
1.0           (ratio)   (ALLOCATION) new fine root C : new leaf C
2.2           (ratio)   (ALLOCATION) new stem C : new leaf C
0.1           (ratio)   (ALLOCATION) new live wood C : new total wood C
0.3           (ratio)   (ALLOCATION) new croot C : new stem C
0.5           (prop.)   (ALLOCATION) current growth proportion
42.0          (kgC/kgN) C:N of leaves
93.0          (kgC/kgN) C:N of leaf litter, after retranslocation
42.0          (kgC/kgN) C:N of fine roots
50.0          (kgC/kgN) C:N of live wood
729.0         (kgC/kgN) C:N of dead wood
0.32          (DIM)     leaf litter labile proportion
0.44          (DIM)     leaf litter cellulose proportion
0.24          (DIM)     leaf litter lignin proportion
0.30          (DIM)     fine root labile proportion
0.45          (DIM)     fine root cellulose proportion
0.25          (DIM)     fine root lignin proportion
0.76          (DIM)     dead wood cellulose proportion
0.24          (DIM)     dead wood lignin proportion
0.041         (1/LAI/d) canopy water interception coefficient
0.5           (DIM)     canopy light extinction coefficient
2.6           (DIM)     all-sided to projected leaf area ratio
12.0          (m2/kgC)  canopy average specific leaf area (projected area
basis)
2.0           (DIM)     ratio of shaded SLA:sunlit SLA
0.04          (DIM)     fraction of leaf N in Rubisco
0.003         (m/s)     maximum stomatal conductance (projected area basis)
0.00001       (m/s)     cuticular conductance (projected area basis)
0.08          (m/s)     boundary layer conductance (projected area basis)
-0.6          (MPa)     leaf water potential: start of conductance reduction
-2.3          (MPa)     leaf water potential: complete conductance reduction
930.0         (Pa)      vapor pressure deficit: start of conductance
reduction
4100.0        (Pa)      vapor pressure deficit: complete conductance
reduction
```

Meteorology Input File (.met file format):

```
Missoula, 1950-1993 : Sample input for MTCLIM v4.1
MTCLIM v4.1 OUTPUT FILE : Tue Aug 25 10:15:00 1998
  year  yday    Tmax    Tmin    Tday    prcp      VPD     srad  daylen
             (deg C) (deg C) (deg C)    (cm)     (Pa)  (W m-2)     (s)
  1950     1   -3.90  -13.90   -6.65    0.10   158.19   123.31   30229
  1950     2   -7.80  -21.70  -11.62    0.00   136.27   183.78   30284
  1950     3  -16.10  -23.30  -18.08    0.00    53.36   140.67   30344
```

Physical Site Input File (.ini file):

```
Biome-BGC v4.1.2 example : (normal simulation, Missoula, evergreen
needleleaf)

MET_INPUT     (keyword) start of meteorology file control block
metdata/miss5093.mtc41  meteorology input filename
4             (int)     header lines in met file

RESTART       (keyword) start of restart control block
0             (flag)    1 = read restart file     0 = don't read restart file
0             (flag)    1 = write restart file    0 = don't write restart
file
1             (flag)    1 = use restart metyear   0 = reset metyear
restart/enf_test1.endpoint    input restart filename
restart/enf_test1.endpoint    output restart filename

TIME_DEFINE   (keyword - do not remove)
44            (int)       number of meteorological data years
44            (int)       number of simulation years
1950          (int)       first simulation year
0             (flag)      1 = spinup simulation    0 = normal simulation
6000          (int)       maximum number of spinup years (if spinup
simulation)

CLIM_CHANGE   (keyword - do not remove)
0.0           (deg C)   offset for Tmax
0.0           (deg C)   offset for Tmin
1.0           (DIM)     multiplier for Prcp
1.0           (DIM)     multiplier for VPD
1.0           (DIM)     multiplier for shortwave radiation

CO2_CONTROL   (keyword - do not remove)
0             (flag)    0=constant 1=vary with file 2=constant, file for Ndep
294.842       (ppm)     constant atmospheric CO2 concentration
xxxxxxxxxxx   (file)    annual variable CO2 filename

SITE          (keyword) start of site physical constants block
1.0           (m)       effective soil depth (corrected for rock fraction)
30.0          (%)       sand percentage by volume in rock-free soil
50.0          (%)       silt percentage by volume in rock-free soil
20.0          (%)       clay percentage by volume in rock-free soil
977.0         (m)       site elevation
46.8          (degrees) site latitude (- for S.Hem.)
0.2           (DIM)     site shortwave albedo
0.0001        (kgN/m2/yr) wet+dry atmospheric deposition of N
0.0004        (kgN/m2/yr) symbiotic+asymbiotic fixation of N
```

```
RAMP_NDEP     (keyword - do not remove)
0             (flag) do a ramped N-deposition run? 0=no, 1=yes
2099          (int)  reference year for industrial N deposition
0.0001        (kgN/m2/yr) industrial N deposition value

EPC_FILE      (keyword - do not remove)
epc/enf.epc   (file) evergreen needleleaf forest ecophysiological constants

W_STATE       (keyword) start of water state variable initialization block
0.0           (kg/m2)   water stored in snowpack
0.5           (DIM)     initial soil water as a proportion of saturation

C_STATE       (keyword) start of carbon state variable initialization block
0.001         (kgC/m2)  first-year maximum leaf carbon
0.0           (kgC/m2)  first-year maximum stem carbon
0.0           (kgC/m2)  coarse woody debris carbon
0.0           (kgC/m2)  litter carbon, labile pool
0.0           (kgC/m2)  litter carbon, unshielded cellulose pool
0.0           (kgC/m2)  litter carbon, shielded cellulose pool
0.0           (kgC/m2)  litter carbon, lignin pool
0.0           (kgC/m2)  soil carbon, fast microbial recycling pool
0.0           (kgC/m2)  soil carbon, medium microbial recycling pool
0.0           (kgC/m2)  soil carbon, slow microbial recycling pool
0.0           (kgC/m2)  soil carbon, recalcitrant SOM (slowest)

N_STATE       (keyword) start of nitrogen state variable initialization block
0.0           (kgN/m2)  litter nitrogen, labile pool
0.0           (kgN/m2)  soil nitrogen, mineral pool

OUTPUT_CONTROL   (keyword - do not remove)
outputs/oth     (text) prefix for output files
1   (flag)  1 = write daily output   0 = no daily output
1   (flag)  1 = monthly avg of daily variables  0 = no monthly avg
1   (flag)  1 = annual avg of daily variables   0 = no annual avg
1   (flag)  1 = write annual output  0 = no annual output
1   (flag)  for on-screen progress indicator

DAILY_OUTPUT     (keyword)
23     (int) number of daily variables to output
20     0 ws.soilw
21     1 ws.snoww
38     2 wf.canopyw_evap
40     3 wf.snoww_subl
42     4 wf.soilw_evap
43     5 wf.soilw_trans
44     6 wf.soilw_outflow
70     7 cs.cwdc
509    8 epv.proj_lai
528    9 epv.daily_net_nmin
620    10 summary.daily_npp
621    11 summary.daily_nep
622    12 summary.daily_nee
623    13 summary.daily_gpp
624    14 summary.daily_mr
625    15 summary.daily_gr
626    16 summary.daily_hr
```

```
627    17 summary.daily_fire
636    18 summary.vegc
637    19 summary.litrc
638    20 summary.soilc
639    21 summary.totalc
579    22 psn_sun.A

ANNUAL_OUTPUT    (keyword)
6               (int)   number of annual output variables
545     0 annual maximum projected LAI
636     1 vegetation C
637     2 litter C
638     3 soil C
639     4 total C
307     5 soil mineral N

END_INIT      (keyword) indicates the end of the initialization file
```

Appendix B: BiomeBGC Output Map (taken from output_map_init.c)

```
/* daily meteorological variables */
output_map[0] = &metv->prcp;
output_map[1] = &metv->tmax;
output_map[2] = &metv->tmin;
output_map[3] = &metv->tavg;
output_map[4] = &metv->tday;
output_map[5] = &metv->tnight;
output_map[6] = &metv->tsoil;
output_map[7] = &metv->vpd;
output_map[8] = &metv->swavgfd;
output_map[9] = &metv->swabs;
output_map[10] = &metv->swtrans;
output_map[11] = &metv->swabs_per_plaisun;
output_map[12] = &metv->swabs_per_plaishade;
output_map[13] = &metv->ppfd_per_plaisun;
output_map[14] = &metv->ppfd_per_plaishade;
output_map[15] = &metv->par;
output_map[16] = &metv->parabs;
output_map[17] = &metv->pa;
output_map[18] = &metv->co2;
output_map[19] = &metv->dayl;

/* water state variables */
output_map[20] = &ws->soilw;
output_map[21] = &ws->snoww;
output_map[22] = &ws->canopyw;
output_map[23] = &ws->prcp_src;
output_map[24] = &ws->outflow_snk;
output_map[25] = &ws->soilevap_snk;
output_map[26] = &ws->snowsubl_snk;
output_map[27] = &ws->canopyevap_snk;
output_map[28] = &ws->trans_snk;

/* water flux variables */
output_map[35] = &wf->prcp_to_canopyw;
```

```
output_map[36] = &wf->prcp_to_soilw;
output_map[37] = &wf->prcp_to_snoww;
output_map[38] = &wf->canopyw_evap;
output_map[39] = &wf->canopyw_to_soilw;
output_map[40] = &wf->snoww_subl;
output_map[41] = &wf->snoww_to_soilw;
output_map[42] = &wf->soilw_evap;
output_map[43] = &wf->soilw_trans;
output_map[44] = &wf->soilw_outflow;

/* carbon state variables */
output_map[50] = &cs->leafc;
output_map[51] = &cs->leafc_storage;
output_map[52] = &cs->leafc_transfer;
output_map[53] = &cs->frootc;
output_map[54] = &cs->frootc_storage;
output_map[55] = &cs->frootc_transfer;
output_map[56] = &cs->livestemc;
output_map[57] = &cs->livestemc_storage;
output_map[58] = &cs->livestemc_transfer;
output_map[59] = &cs->deadstemc;
output_map[60] = &cs->deadstemc_storage;
output_map[61] = &cs->deadstemc_transfer;
output_map[62] = &cs->livecrootc;
output_map[63] = &cs->livecrootc_storage;
output_map[64] = &cs->livecrootc_transfer;
output_map[65] = &cs->deadcrootc;
output_map[66] = &cs->deadcrootc_storage;
output_map[67] = &cs->deadcrootc_transfer;
output_map[68] = &cs->gresp_storage;
output_map[69] = &cs->gresp_transfer;
output_map[70] = &cs->cwdc;
output_map[71] = &cs->litr1c;
output_map[72] = &cs->litr2c;
output_map[73] = &cs->litr3c;
output_map[74] = &cs->litr4c;
output_map[75] = &cs->soil1c;
output_map[76] = &cs->soil2c;
output_map[77] = &cs->soil3c;
output_map[78] = &cs->soil4c;
output_map[79] = &cs->cpool;
output_map[80] = &cs->psnsun_src;
output_map[81] = &cs->psnshade_src;
output_map[82] = &cs->leaf_mr_snk;
output_map[83] = &cs->leaf_gr_snk;
output_map[84] = &cs->froot_mr_snk;
output_map[85] = &cs->froot_gr_snk;
output_map[86] = &cs->livestem_mr_snk;
output_map[87] = &cs->livestem_gr_snk;
output_map[88] = &cs->deadstem_gr_snk;
output_map[89] = &cs->livecroot_mr_snk;
output_map[90] = &cs->livecroot_gr_snk;
output_map[91] = &cs->deadcroot_gr_snk;
output_map[92] = &cs->litr1_hr_snk;
output_map[93] = &cs->litr2_hr_snk;
output_map[94] = &cs->litr4_hr_snk;
output_map[95] = &cs->soil1_hr_snk;
```

```
output_map[96] = &cs->soil2_hr_snk;
output_map[97] = &cs->soil3_hr_snk;
output_map[98] = &cs->soil4_hr_snk;
output_map[99] = &cs->fire_snk;

/* carbon flux variables */
output_map[120] = &cf->m_leafc_to_litr1c;
output_map[121] = &cf->m_leafc_to_litr2c;
output_map[122] = &cf->m_leafc_to_litr3c;
output_map[123] = &cf->m_leafc_to_litr4c;
output_map[124] = &cf->m_frootc_to_litr1c;
output_map[125] = &cf->m_frootc_to_litr2c;
output_map[126] = &cf->m_frootc_to_litr3c;
output_map[127] = &cf->m_frootc_to_litr4c;
output_map[128] = &cf->m_leafc_storage_to_litr1c;
output_map[129] = &cf->m_frootc_storage_to_litr1c;
output_map[130] = &cf->m_livestemc_storage_to_litr1c;
output_map[131] = &cf->m_deadstemc_storage_to_litr1c;
output_map[132] = &cf->m_livecrootc_storage_to_litr1c;
output_map[133] = &cf->m_deadcrootc_storage_to_litr1c;
output_map[134] = &cf->m_leafc_transfer_to_litr1c;
output_map[135] = &cf->m_frootc_transfer_to_litr1c;
output_map[136] = &cf->m_livestemc_transfer_to_litr1c;
output_map[137] = &cf->m_deadstemc_transfer_to_litr1c;
output_map[138] = &cf->m_livecrootc_transfer_to_litr1c;
output_map[139] = &cf->m_deadcrootc_transfer_to_litr1c;
output_map[140] = &cf->m_livestemc_to_cwdc;
output_map[141] = &cf->m_deadstemc_to_cwdc;
output_map[142] = &cf->m_livecrootc_to_cwdc;
output_map[143] = &cf->m_deadcrootc_to_cwdc;
output_map[144] = &cf->m_gresp_storage_to_litr1c;
output_map[145] = &cf->m_gresp_transfer_to_litr1c;
output_map[146] = &cf->m_leafc_to_fire;
output_map[147] = &cf->m_frootc_to_fire;
output_map[148] = &cf->m_leafc_storage_to_fire;
output_map[149] = &cf->m_frootc_storage_to_fire;
output_map[150] = &cf->m_livestemc_storage_to_fire;
output_map[151] = &cf->m_deadstemc_storage_to_fire;
output_map[152] = &cf->m_livecrootc_storage_to_fire;
output_map[153] = &cf->m_deadcrootc_storage_to_fire;
output_map[154] = &cf->m_leafc_transfer_to_fire;
output_map[155] = &cf->m_frootc_transfer_to_fire;
output_map[156] = &cf->m_livestemc_transfer_to_fire;
output_map[157] = &cf->m_deadstemc_transfer_to_fire;
output_map[158] = &cf->m_livecrootc_transfer_to_fire;
output_map[159] = &cf->m_deadcrootc_transfer_to_fire;
output_map[160] = &cf->m_livestemc_to_fire;
output_map[161] = &cf->m_deadstemc_to_fire;
output_map[162] = &cf->m_livecrootc_to_fire;
output_map[163] = &cf->m_deadcrootc_to_fire;
output_map[164] = &cf->m_gresp_storage_to_fire;
output_map[165] = &cf->m_gresp_transfer_to_fire;
output_map[166] = &cf->m_litr1c_to_fire;
output_map[167] = &cf->m_litr2c_to_fire;
output_map[168] = &cf->m_litr3c_to_fire;
output_map[169] = &cf->m_litr4c_to_fire;
output_map[170] = &cf->m_cwdc_to_fire;
```

```
output_map[171] = &cf->leafc_transfer_to_leafc;
output_map[172] = &cf->frootc_transfer_to_frootc;
output_map[173] = &cf->livestemc_transfer_to_livestemc;
output_map[174] = &cf->deadstemc_transfer_to_deadstemc;
output_map[175] = &cf->livecrootc_transfer_to_livecrootc;
output_map[176] = &cf->deadcrootc_transfer_to_deadcrootc;
output_map[177] = &cf->leafc_to_litr1c;
output_map[178] = &cf->leafc_to_litr2c;
output_map[179] = &cf->leafc_to_litr3c;
output_map[180] = &cf->leafc_to_litr4c;
output_map[181] = &cf->frootc_to_litr1c;
output_map[182] = &cf->frootc_to_litr2c;
output_map[183] = &cf->frootc_to_litr3c;
output_map[184] = &cf->frootc_to_litr4c;
output_map[185] = &cf->leaf_day_mr;
output_map[186] = &cf->leaf_night_mr;
output_map[187] = &cf->froot_mr;
output_map[188] = &cf->livestem_mr;
output_map[189] = &cf->livecroot_mr;
output_map[190] = &cf->psnsun_to_cpool;
output_map[191] = &cf->psnshade_to_cpool;
output_map[192] = &cf->cwdc_to_litr2c;
output_map[193] = &cf->cwdc_to_litr3c;
output_map[194] = &cf->cwdc_to_litr4c;
output_map[195] = &cf->litr1_hr;
output_map[196] = &cf->litr1c_to_soil1c;
output_map[197] = &cf->litr2_hr;
output_map[198] = &cf->litr2c_to_soil2c;
output_map[199] = &cf->litr3c_to_litr2c;
output_map[200] = &cf->litr4_hr;
output_map[201] = &cf->litr4c_to_soil3c;
output_map[202] = &cf->soil1_hr;
output_map[203] = &cf->soil1c_to_soil2c;
output_map[204] = &cf->soil2_hr;
output_map[205] = &cf->soil2c_to_soil3c;
output_map[206] = &cf->soil3_hr;
output_map[207] = &cf->soil3c_to_soil4c;
output_map[208] = &cf->soil4_hr;
output_map[209] = &cf->cpool_to_leafc;
output_map[210] = &cf->cpool_to_leafc_storage;
output_map[211] = &cf->cpool_to_frootc;
output_map[212] = &cf->cpool_to_frootc_storage;
output_map[213] = &cf->cpool_to_livestemc;
output_map[214] = &cf->cpool_to_livestemc_storage;
output_map[215] = &cf->cpool_to_deadstemc;
output_map[216] = &cf->cpool_to_deadstemc_storage;
output_map[217] = &cf->cpool_to_livecrootc;
output_map[218] = &cf->cpool_to_livecrootc_storage;
output_map[219] = &cf->cpool_to_deadcrootc;
output_map[220] = &cf->cpool_to_deadcrootc_storage;
output_map[221] = &cf->cpool_to_gresp_storage;
output_map[222] = &cf->cpool_leaf_gr;
output_map[223] = &cf->transfer_leaf_gr;
output_map[224] = &cf->cpool_froot_gr;
output_map[225] = &cf->transfer_froot_gr;
output_map[226] = &cf->cpool_livestem_gr;
output_map[227] = &cf->transfer_livestem_gr;
```

```
output_map[228] = &cf->cpool_deadstem_gr;
output_map[229] = &cf->transfer_deadstem_gr;
output_map[230] = &cf->cpool_livecroot_gr;
output_map[231] = &cf->transfer_livecroot_gr;
output_map[232] = &cf->cpool_deadcroot_gr;
output_map[233] = &cf->transfer_deadcroot_gr;
output_map[234] = &cf->leafc_storage_to_leafc_transfer;
output_map[235] = &cf->frootc_storage_to_frootc_transfer;
output_map[236] = &cf->livestemc_storage_to_livestemc_transfer;
output_map[237] = &cf->deadstemc_storage_to_deadstemc_transfer;
output_map[238] = &cf->livecrootc_storage_to_livecrootc_transfer;
output_map[239] = &cf->deadcrootc_storage_to_deadcrootc_transfer;
output_map[240] = &cf->gresp_storage_to_gresp_transfer;
output_map[241] = &cf->livestemc_to_deadstemc;
output_map[242] = &cf->livecrootc_to_deadcrootc;
output_map[243] = &cf->cpool_leaf_storage_gr;
output_map[244] = &cf->cpool_froot_storage_gr;
output_map[245] = &cf->cpool_livestem_storage_gr;
output_map[246] = &cf->cpool_deadstem_storage_gr;
output_map[247] = &cf->cpool_livecroot_storage_gr;
output_map[248] = &cf->cpool_deadcroot_storage_gr;

/* nitrogen state variables */
output_map[280] = &ns->leafn;
output_map[281] = &ns->leafn_storage;
output_map[282] = &ns->leafn_transfer;
output_map[283] = &ns->frootn;
output_map[284] = &ns->frootn_storage;
output_map[285] = &ns->frootn_transfer;
output_map[286] = &ns->livestemn;
output_map[287] = &ns->livestemn_storage;
output_map[288] = &ns->livestemn_transfer;
output_map[289] = &ns->deadstemn;
output_map[290] = &ns->deadstemn_storage;
output_map[291] = &ns->deadstemn_transfer;
output_map[292] = &ns->livecrootn;
output_map[293] = &ns->livecrootn_storage;
output_map[294] = &ns->livecrootn_transfer;
output_map[295] = &ns->deadcrootn;
output_map[296] = &ns->deadcrootn_storage;
output_map[297] = &ns->deadcrootn_transfer;
output_map[298] = &ns->cwdn;
output_map[299] = &ns->litr1n;
output_map[300] = &ns->litr2n;
output_map[301] = &ns->litr3n;
output_map[302] = &ns->litr4n;
output_map[303] = &ns->soil1n;
output_map[304] = &ns->soil2n;
output_map[305] = &ns->soil3n;
output_map[306] = &ns->soil4n;
output_map[307] = &ns->sminn;
output_map[308] = &ns->retransn;
output_map[309] = &ns->npool;
output_map[310] = &ns->nfix_src;
output_map[311] = &ns->ndep_src;
output_map[312] = &ns->nleached_snk;
output_map[313] = &ns->nvol_snk;
```

```
output_map[314] = &ns->fire_snk;

/* nitrogen flux variables */
output_map[340] = &nf->m_leafn_to_litr1n;
output_map[341] = &nf->m_leafn_to_litr2n;
output_map[342] = &nf->m_leafn_to_litr3n;
output_map[343] = &nf->m_leafn_to_litr4n;
output_map[344] = &nf->m_frootn_to_litr1n;
output_map[345] = &nf->m_frootn_to_litr2n;
output_map[346] = &nf->m_frootn_to_litr3n;
output_map[347] = &nf->m_frootn_to_litr4n;
output_map[348] = &nf->m_leafn_storage_to_litr1n;
output_map[349] = &nf->m_frootn_storage_to_litr1n;
output_map[350] = &nf->m_livestemn_storage_to_litr1n;
output_map[351] = &nf->m_deadstemn_storage_to_litr1n;
output_map[352] = &nf->m_livecrootn_storage_to_litr1n;
output_map[353] = &nf->m_deadcrootn_storage_to_litr1n;
output_map[354] = &nf->m_leafn_transfer_to_litr1n;
output_map[355] = &nf->m_frootn_transfer_to_litr1n;
output_map[356] = &nf->m_livestemn_transfer_to_litr1n;
output_map[357] = &nf->m_deadstemn_transfer_to_litr1n;
output_map[358] = &nf->m_livecrootn_transfer_to_litr1n;
output_map[359] = &nf->m_deadcrootn_transfer_to_litr1n;
output_map[360] = &nf->m_livestemn_to_litr1n;
output_map[361] = &nf->m_livestemn_to_cwdn;
output_map[362] = &nf->m_deadstemn_to_cwdn;
output_map[363] = &nf->m_livecrootn_to_litr1n;
output_map[364] = &nf->m_livecrootn_to_cwdn;
output_map[365] = &nf->m_deadcrootn_to_cwdn;
output_map[366] = &nf->m_retransn_to_litr1n;
output_map[367] = &nf->m_leafn_to_fire;
output_map[368] = &nf->m_frootn_to_fire;
output_map[369] = &nf->m_leafn_storage_to_fire;
output_map[370] = &nf->m_frootn_storage_to_fire;
output_map[371] = &nf->m_livestemn_storage_to_fire;
output_map[372] = &nf->m_deadstemn_storage_to_fire;
output_map[373] = &nf->m_livecrootn_storage_to_fire;
output_map[374] = &nf->m_deadcrootn_storage_to_fire;
output_map[375] = &nf->m_leafn_transfer_to_fire;
output_map[376] = &nf->m_frootn_transfer_to_fire;
output_map[377] = &nf->m_livestemn_transfer_to_fire;
output_map[378] = &nf->m_deadstemn_transfer_to_fire;
output_map[379] = &nf->m_livecrootn_transfer_to_fire;
output_map[380] = &nf->m_deadcrootn_transfer_to_fire;
output_map[381] = &nf->m_livestemn_to_fire;
output_map[382] = &nf->m_deadstemn_to_fire;
output_map[383] = &nf->m_livecrootn_to_fire;
output_map[384] = &nf->m_deadcrootn_to_fire;
output_map[385] = &nf->m_retransn_to_fire;
output_map[386] = &nf->m_litr1n_to_fire;
output_map[387] = &nf->m_litr2n_to_fire;
output_map[388] = &nf->m_litr3n_to_fire;
output_map[389] = &nf->m_litr4n_to_fire;
output_map[390] = &nf->m_cwdn_to_fire;
output_map[391] = &nf->leafn_transfer_to_leafn;
output_map[392] = &nf->frootn_transfer_to_frootn;
output_map[393] = &nf->livestemn_transfer_to_livestemn;
```

```
output_map[394] = &nf->deadstemn_transfer_to_deadstemn;
output_map[395] = &nf->livecrootn_transfer_to_livecrootn;
output_map[396] = &nf->deadcrootn_transfer_to_deadcrootn;
output_map[397] = &nf->leafn_to_litr1n;
output_map[398] = &nf->leafn_to_litr2n;
output_map[399] = &nf->leafn_to_litr3n;
output_map[400] = &nf->leafn_to_litr4n;
output_map[401] = &nf->leafn_to_retransn;
output_map[402] = &nf->frootn_to_litr1n;
output_map[403] = &nf->frootn_to_litr2n;
output_map[404] = &nf->frootn_to_litr3n;
output_map[405] = &nf->frootn_to_litr4n;
output_map[406] = &nf->ndep_to_sminn;
output_map[407] = &nf->nfix_to_sminn;
output_map[408] = &nf->cwdn_to_litr2n;
output_map[409] = &nf->cwdn_to_litr3n;
output_map[410] = &nf->cwdn_to_litr4n;
output_map[411] = &nf->litr1n_to_soil1n;
output_map[412] = &nf->sminn_to_soil1n_l1;
output_map[413] = &nf->litr2n_to_soil2n;
output_map[414] = &nf->sminn_to_soil2n_l2;
output_map[415] = &nf->litr3n_to_litr2n;
output_map[416] = &nf->litr4n_to_soil3n;
output_map[417] = &nf->sminn_to_soil3n_l4;
output_map[418] = &nf->soil1n_to_soil2n;
output_map[419] = &nf->sminn_to_soil2n_s1;
output_map[420] = &nf->soil2n_to_soil3n;
output_map[421] = &nf->sminn_to_soil3n_s2;
output_map[422] = &nf->soil3n_to_soil4n;
output_map[423] = &nf->sminn_to_soil4n_s3;
output_map[424] = &nf->soil4n_to_sminn;
output_map[425] = &nf->sminn_to_nvol_l1s1;
output_map[426] = &nf->sminn_to_nvol_l2s2;
output_map[427] = &nf->sminn_to_nvol_l4s3;
output_map[428] = &nf->sminn_to_nvol_s1s2;
output_map[429] = &nf->sminn_to_nvol_s2s3;
output_map[430] = &nf->sminn_to_nvol_s3s4;
output_map[431] = &nf->sminn_to_nvol_s4;
output_map[432] = &nf->sminn_leached;
output_map[433] = &nf->retransn_to_npool;
output_map[434] = &nf->sminn_to_npool;
output_map[435] = &nf->npool_to_leafn;
output_map[436] = &nf->npool_to_leafn_storage;
output_map[437] = &nf->npool_to_frootn;
output_map[438] = &nf->npool_to_frootn_storage;
output_map[439] = &nf->npool_to_livestemn;
output_map[440] = &nf->npool_to_livestemn_storage;
output_map[441] = &nf->npool_to_deadstemn;
output_map[442] = &nf->npool_to_deadstemn_storage;
output_map[443] = &nf->npool_to_livecrootn;
output_map[444] = &nf->npool_to_livecrootn_storage;
output_map[445] = &nf->npool_to_deadcrootn;
output_map[446] = &nf->npool_to_deadcrootn_storage;
output_map[447] = &nf->leafn_storage_to_leafn_transfer;
output_map[448] = &nf->frootn_storage_to_frootn_transfer;
output_map[449] = &nf->livestemn_storage_to_livestemn_transfer;
output_map[450] = &nf->deadstemn_storage_to_deadstemn_transfer;
```

```
output_map[451] = &nf->livecrootn_storage_to_livecrootn_transfer;
output_map[452] = &nf->deadcrootn_storage_to_deadcrootn_transfer;
output_map[453] = &nf->livestemn_to_deadstemn;
output_map[454] = &nf->livestemn_to_retransn;
output_map[455] = &nf->livecrootn_to_deadcrootn;
output_map[456] = &nf->livecrootn_to_retransn;

/* phenological variables */
output_map[480] = &phen->remdays_curgrowth;
output_map[481] = &phen->remdays_transfer;
output_map[482] = &phen->remdays_litfall;
output_map[483] = &phen->predays_transfer;
output_map[484] = &phen->predays_litfall;

/* ecophysiological variables */
output_map[500] = &epv->day_leafc_litfall_increment;
output_map[501] = &epv->day_frootc_litfall_increment;
output_map[502] = &epv->day_livestemc_turnover_increment;
output_map[503] = &epv->day_livecrootc_turnover_increment;
output_map[504] = &epv->annmax_leafc;
output_map[505] = &epv->annmax_frootc;
output_map[506] = &epv->annmax_livestemc;
output_map[507] = &epv->annmax_livecrootc;
output_map[508] = &epv->dsr;
output_map[509] = &epv->proj_lai;
output_map[510] = &epv->all_lai;
output_map[511] = &epv->plaisun;
output_map[512] = &epv->plaishade;
output_map[513] = &epv->sun_proj_sla;
output_map[514] = &epv->shade_proj_sla;
output_map[515] = &epv->psi;
output_map[516] = &epv->vwc;
output_map[517] = &epv->dlmr_area_sun;
output_map[518] = &epv->dlmr_area_shade;
output_map[519] = &epv->gl_t_wv_sun;
output_map[520] = &epv->gl_t_wv_shade;
output_map[521] = &epv->assim_sun;
output_map[522] = &epv->assim_shade;
output_map[523] = &epv->t_scalar;
output_map[524] = &epv->w_scalar;
output_map[525] = &epv->rate_scalar;
output_map[526] = &epv->daily_gross_nmin;
output_map[527] = &epv->daily_gross_nimmob;
output_map[528] = &epv->daily_net_nmin;
output_map[529] = &epv->m_tmin;
output_map[530] = &epv->m_psi;
output_map[531] = &epv->m_co2;
output_map[532] = &epv->m_ppfd_sun;
output_map[533] = &epv->m_ppfd_shade;
output_map[534] = &epv->m_vpd;
output_map[535] = &epv->m_final_sun;
output_map[536] = &epv->m_final_shade;
output_map[537] = &epv->gl_bl;
output_map[538] = &epv->gl_c;
output_map[539] = &epv->gl_s_sun;
output_map[540] = &epv->gl_s_shade;
output_map[541] = &epv->gl_e_wv;
```

```
output_map[542] = &epv->gl_sh;
output_map[543] = &epv->gc_e_wv;
output_map[544] = &epv->gc_sh;
output_map[545] = &epv->ytd_maxplai;
output_map[546] = &epv->fpi;

/* photosynthesis variables */
/* sunlit canopy fraction */
output_map[560] = &psn_sun->pa;
output_map[561] = &psn_sun->co2;
output_map[562] = &psn_sun->t;
output_map[563] = &psn_sun->lnc;
output_map[564] = &psn_sun->flnr;
output_map[565] = &psn_sun->ppfd;
output_map[566] = &psn_sun->g;
output_map[567] = &psn_sun->dlmr;
output_map[568] = &psn_sun->Ci;
output_map[569] = &psn_sun->O2;
output_map[570] = &psn_sun->Ca;
output_map[571] = &psn_sun->gamma;
output_map[572] = &psn_sun->Kc;
output_map[573] = &psn_sun->Ko;
output_map[574] = &psn_sun->Vmax;
output_map[575] = &psn_sun->Jmax;
output_map[576] = &psn_sun->J;
output_map[577] = &psn_sun->Av;
output_map[578] = &psn_sun->Aj;
output_map[579] = &psn_sun->A;

/* photosynthesis variables */
/* shaded canopy fraction */
output_map[590] = &psn_shade->pa;
output_map[591] = &psn_shade->co2;
output_map[592] = &psn_shade->t;
output_map[593] = &psn_shade->lnc;
output_map[594] = &psn_shade->flnr;
output_map[595] = &psn_shade->ppfd;
output_map[596] = &psn_shade->g;
output_map[597] = &psn_shade->dlmr;
output_map[598] = &psn_shade->Ci;
output_map[599] = &psn_shade->O2;
output_map[600] = &psn_shade->Ca;
output_map[601] = &psn_shade->gamma;
output_map[602] = &psn_shade->Kc;
output_map[603] = &psn_shade->Ko;
output_map[604] = &psn_shade->Vmax;
output_map[605] = &psn_shade->Jmax;
output_map[606] = &psn_shade->J;
output_map[607] = &psn_shade->Av;
output_map[608] = &psn_shade->Aj;
output_map[609] = &psn_shade->A;

/* carbon budget summary output variables */
output_map[620] = &summary->daily_npp;
output_map[621] = &summary->daily_nep;
output_map[622] = &summary->daily_nee;
output_map[623] = &summary->daily_gpp;
```

```
output_map[624] = &summary->daily_mr;
output_map[625] = &summary->daily_gr;
output_map[626] = &summary->daily_hr;
output_map[627] = &summary->daily_fire;
output_map[628] = &summary->cum_npp;
output_map[629] = &summary->cum_nep;
output_map[630] = &summary->cum_nee;
output_map[631] = &summary->cum_gpp;
output_map[632] = &summary->cum_mr;
output_map[633] = &summary->cum_gr;
output_map[634] = &summary->cum_hr;
output_map[635] = &summary->cum_fire;
output_map[636] = &summary->vegc;
output_map[637] = &summary->litrc;
output_map[638] = &summary->soilc;
output_map[639] = &summary->totalc;
output_map[640] = &summary->daily_litfallc;
output_map[641] = &summary->daily_et;
output_map[642] = &summary->daily_outflow;
output_map[643] = &summary->daily_evap;
output_map[644] = &summary->daily_trans;
output_map[645] = &summary->daily_soilw;
output_map[646] = &summary->daily_snoww;
```

Appendix C: BBGC Constants (from bgc_constants.h)

```
/* atmospheric constants */
/* from the definition of the standard atmosphere, as established
by the International Civil Aviation Organization, and referenced in:

Iribane, J.V., and W.L. Godson, 1981. Atmospheric Thermodynamics. 2nd
       Edition. D. Reidel Publishing Company, Dordrecht, The Netherlands.
       (pp 10,167-168,245)
*/
G_STD    9.80665         /* (m/s2) standard gravitational accel. */
P_STD    101325.0        /* (Pa) standard pressure at 0.0 m elevation */
T_STD    288.15          /* (K) standard temp at 0.0 m elevation  */
MA       28.9644e-3      /* (kg/mol) molecular weight of air */
W       18.0148e-3      /* (kg/mol) molecular weight of water */
CP       1010.0          /* (J/kg K) specific heat of air */
LR_STD   0.0065          /* (-K/m) standard temperature lapse rate */
R        8.3143          /* (m3 Pa/ mol K) gas law constant */
SBC      5.67e-8         /* (W/(m2 K4)) Stefan-Boltzmann constant */
EPS      0.6219          /* (MW/MA) unitless ratio of molec weights

/* ecosystem constants */
RAD2PAR     0.45     /* (DIM) ratio PAR / SWtotal  */
EPAR        4.55     /* (umol/J) PAR photon energy ratio */
SOIL1_CN    12.0     /* C:N for fast microbial recycling pool */
SOIL2_CN    12.0     /* C:N for slow microbial recycling pool */
SOIL3_CN    10.0     /* C:N for recalcitrant SOM pool (humus) */
SOIL4_CN    10.0     /* C:N for recalcitrant SOM pool (humus) */
GRPERC      0.3      /* (DIM) growth resp per unit of C grown */
GRPNOW      1.0      /* (DIM) proportion of storage growth resp at fixation*/
PPFD50      75.0              /* (umol/m2/s) PPFD for 1/2 stomatal closure */
```

```
DENITRIF_PROPORTION  0.01     /* fraction of mineralization to volatile */
MOBILEN_PROPORTION   0.1      /* fraction mineral N avail for leaching */

/* respiration fractions for fluxes between compartments (unitless) */
RFL1S1          0.39    /* transfer from litter 1 to soil 1 */
RFL2S2          0.55    /* transfer from litter 2 to soil 2 */
RFL4S3          0.29    /* transfer from litter 4 to soil 3 */
RFS1S2          0.28    /* transfer from soil 1 to soil 2 */
RFS2S3          0.46    /* transfer from soil 2 to soil 3 */
RFS3S4          0.55    /* transfer from soil 3 to soil 4 */

/* base decomposition rate constants (1/day) */
KL1_BASE        0.7     /* labile litter pool */
KL2_BASE        0.07    /* cellulose litter pool */
KL4_BASE        0.014   /* lignin litter pool */
KS1_BASE        0.07    /* fast microbial recycling pool */
KS2_BASE        0.014   /* medium microbial recycling pool */
KS3_BASE        0.0014  /* slow microbial recycling pool */
KS4_BASE        0.0001  /* recalcitrant SOM (humus) pool */
KFRAG_BASE      0.001   /* physical fragmentation of coarse woody debris */
```

References

Andren, O. and K. Paustian (1987). "Barley straw decomposition in the field: A comparison of models." Ecology **68**(5): 1190-1200.

Avery, T. E. and H. E. Burkhart (1975). Forest Measurements, 5th Edition. New York, NY, McGraw-Hill Higher Education.

Berg, B., G. Ekbohm, et al. (1984). "Lignin and holocellulose relations during long-term decomposition of some forst litters. Long-term decomposition in a Scots pine forest. IV." Canadian Journal of Botany **62**: 2540-2550.

Berg, B. and C. McClaugherty (1989). "Nitrogen and phosphorous release from decomposing litter in relation to the disappearance of lignin." Canadian Journal of Botany **67**: 1148-1156.

Bond-Lamberty, B., S. T. Gower, et al. (2005). "Reimplementation of the Biome-BGC model to simulate successional change." **25**: 413.

Botkin, D. B., J. F. Janak, et al. (1972). "Some Ecological Consequences of a Computer Model of Forest Growth." Journal of Ecology **60**(3): 849-872.

Bugmann, H. (2001). "A Review of Forest Gap Models." Climatic Change **51**(3): 259.

Conklin, D. and R. Neilson (2005). Biomap Model Description. Corvallis, OR, US Forest Service**:** 120.

Cosby, B. J., G. M. Hornberger, et al. (1984). "A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils." Water Resour. Res. **20**(6): 671-681.

Coughlan, J. C. and S. W. Running (1997). "Regional ecosytem simulation: A general model for simulating snow accumulation and melt in mountainous terrain." Landscape Ecology **12**: 119-136.

De Pury, D. G. G. and G. D. Farquhar (1997). "Simple scaling of photsynthesis from leaves to canopies without the errors of big-leaf models." Plant, Cell and Environment **20**: 537-557.

Donnelly, P. K., J. A. Entry, et al. (1990). "Celluslose and lignin degradation in forest soils: Response to moisture, temperature, and acidity." Microbial Ecology **20**: 289-295.

Farquhar, G. D., S. Caemmerer, et al. (1980). "A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species." Planta **149**(1): 78-90.

Fasman, G. D. (1976). Handbook of Biochemistry and Molecular Biology. Proteins, III. Cleveland, OH, CRC Press.

Harmon, M. E. and B. Marks (2002). "Effects of silvicultural practices on carbon stores in Douglas-fir - wester hemlock forests in the Pacific Northwest, U.S.A.: results from a simulation model." Canadian Journal of Forest Research **32**: 863-877.

Iribane, J. V. and W. L. Godson (1981). Atmospheric Thermodynamics. Dordrecht, Netherlands, D. Reidel Publishing Company.

Jones, H. G. (1992). Plants and Microclimate. Cambridge, Cambridge University Press.

Koch, G. W., S. C. Sillett, et al. (2004). "The limits to tree height." Nature **428**(6985): 851-854.

Korner, C. (1995). Leaf diffusive conductances in the major vegetation types of hte globe. Ecophysiology of Photosynthesis. E. D. Schulze and M. M. Caldwell. New York, Springer-Verlag**:** 463-490.

Korol, R. L., S. W. Running, et al. (1995). "Incorporating intertree competition into an ecosystem model." Canadian Journal of Forest Research **25**(3): 413-424.

Kuehn, G. D. and B. A. McFadden (1969). "Ribulase 1,5-diphosphate carboxylase from Hydrogenomonas eutropha and Hydrogenomona facilis. I. Purification, metallic ion requirements, inhibition, and kinetic constants." Biochemistry **8**(6): 2394-2402.

Kuehn, G. D. and B. A. McFadden (1969). "Ribulose 1,5-diphosphate carboxylase from Hydrogenomonas eutropha and Hydrogenomonas facilis. II. Molecular weight, subunits, composition, and sulfhydryl groups." Biochemistry **8**(6): 2403-2408.

Lambers, H., F. S. I. Chapin, et al. (2008). Plant physiological ecology. New York, NY, Springer.

Larcher, W. (2003). Physiological plant ecology: ecophysiology and stress physiology of functional groups. New York, Springer.

Lloyd, J. and J. A. Taylor (1994). "On the Temperature Dependence of Soil Respiration." Functional Ecology **8**(3): 315-323.

Marks, D., J. Dozier, et al. (1992). "Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada. I. Meteorology measurements and monitoring." Water Resources Research **28**(11): 3029-3042.

McDowell, N., W. T. Pockman, et al. (2008). "Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought?" New Phytologist **178**(4): 719-739.

McNaughton, K. G. and P. G. Jarvis (1983). Predicting effects of vegetation changes on transpiration and evaporation. Water Deficits in Plant Growth. T. T. Kozlowski. London, Academic Press. **7:** 1-47.

Monteith, J. L. and M. H. Unsworth (2008). Principles of Environmental Physics, 3rd Edition. Burlington, MA, Academic Press.

Nobel, P. S. (1991). Phsiochemical and Environmental Plant Physiology. San Diego, CA, Academic Press.

Orchard, V. A. and F. J. Cook (1983). "Relationship between soil respiration and soil moisture." Soil Biology and Biochemistry **15**(4): 447-453.

Pietsch, S. A. and H. Hasenauer (2002). "Using mechanistic modeling within forest ecosystem restoration." Forest Ecology and Management **159**(1-2): 111.

Pietsch, S. A., H. Hasenauer, et al. (2005). "BGC-model parameters for tree species growing in central European forests." Forest Ecology and Management **211**(3): 264.

Porté, A. and H. H. Bartelink (2002). "Modelling mixed forest growth: a review of models for forest management." Ecological Modelling **150**(1-2): 141-188.

Running, S. W. and J. C. Coughlan (1988). "A general model of forest ecosytem processes for regional applications I. Hydrologic Balance, Canopy Gas Exchange and Primary Production Processes." Ecological Modelling **42**: 125-154.

Running, S. W. and S. T. Gower (1991). "FOREST-BGC, A general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets." Tree Physiol **9**(1-2): 147-160.

Ryan, M. G. (1991). "Effects of Climate Change on Plant Respiration." Ecological Applications **1**(2): 157-167.

Saxton, K. E., W. J. Rawls, et al. (1986). "Estimating Generalized Soil-water Characteristics from Texture." Soil Sci Soc Am J **50**(4): 1031-1036.

Schmid, S., E. Thürig, et al. (2006). "Effect of forest management on future carbon pools and fluxes: A model comparison." Forest Ecology and Management **237**(1-3): 65.

Shugart, H. H. (2002). Forest Gap Models. Encyclopedia of Global Environmental Change. H. A. Mooney and J. G. Canadell. Chichester, John Wiley & Sons, Ltd. **2: The Earth system: biological and ecological dimensions of global environemental change.:** 316-323.

Solomon, A. M. and P. J. Bartlein (1992). "Past and furture climate change: response by mixed deciduous-coniferous forest ecosystems in northern Michigan." Canadian Journal of Forest Research **22**(11): 1727-1738.

Stump, L. M. and D. Binkley (1993). "Relationships between litter quality and nitrogen availability in Rocky Mountain forests." Canadian Journal of Forest Research **23**: 492-502.

Taiz, L. and E. Zeiger (2006). Plant Physiology. Sunderland, MA, Sinauer Associates, Inc., Publishers.

Taylor, B. R., C. E. Prescott, et al. (1991). "Substrate control of litter decomposition in four Rocky Mountain coniferous forests." Canadian Journal of Botany **69**(10): 2242-2250.

Thornton, P. E. (1998). Regional Ecosystem Simulation: Combining Surface- and Satellite-Based Observations to Study Linkages between Terrestrial Energy and Mass Budgets. College of Forestry. Missoula, MT, The University of Montana. **Doctor of Philosophy:** 288.

Thornton, P. E., B. E. Law, et al. (2002). "Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests." Agricultural and Forest Meteorology **113**: 185-222.

Thornton, P. E. and N. A. Rosenbloom (2005). "Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model." Ecological Modelling **189**(1-2): 25-48.

Thornton, P. E. and N. E. Zimmermann (2007). "An Improved Canopy Integration Scheme for a Land Surface Model with Prognostic Canopy Structure." Journal of Climate **20**(15): 3902.

Vanclay, J. K. (1994). Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests. Wallingford, UK, CAB International.

Vetter, M., G. Churkina, et al. (2008). "Analyzing the causes and spatial pattern of the European 2003 carbon flux anomaly using seven models." Biogeosciences **5**(2): 561.

Wallace, J. S. and C. J. Holwill (1997). "Soil evaporation from tiger-bush in south-west Niger." Journal of Hydrology **188-189**: 426-442.

Wang, W., K. Ichii, et al. (2009). "A hierarchical analysis of terrestrial ecosystem model Biome-BGC: Equilibrium analysis and model calibration." Ecological Modelling **220**(17): 2009-2023.

Waring, R. H. and G. B. Pitman (1985). "Modifying Lodgepole Pine Stands to Change Susceptibility to Mountain Pine Beetle Attack." Ecology **66**(3): 889-897.

Waring, R. H. and S. W. Running (2007). Forest Ecosystems: Analysis at Multiple Scales. San Francisco, CA, Elsevier Academic Press.

White, M. A., P. E. Thornton, et al. (1997). "A Continental Phenology Model for Monitoring Vegetation Responses to Interannual Climatic Variability." Global Biogeochemical Cycles **11**(2): 217-234.

White, M. A., P. E. Thornton, et al. (2000). "Parameterization and Sensitivity Analysis of the BIOME-BGC Terrestrial Ecosystem Model: Net Primary Production Controls." Earth Interactions **4**: 1-85.

Woodrow, I. E. and J. A. Berry (1988). "Enzymatic Regulation of Photosynthetic CO2, Fixation in C3 Plants." Annual Review of Plant Physiology and Plant Molecular Biology **39**(1): 533-594.

Wullschleger, S. D. (1993). "Biochemical Limitations to Carbon Assimilation in C3 Plants--A Retrospective Analysis of the A/Ci Curves from 109 Species." J. Exp. Bot. **44**(5): 907-920.

Chapter 2

A system to integrate multi-scaled data sources for improving terrestrial C balance estimates

Abstract

Climate change and the policy responses for mitigation and adaptation have necessitated a better understanding of the carbon cycle and land use change dynamics. One approach to spatially quantifying ecosystem function is estimating the potential productivity of different land use types both on a year-to-year basis and from a maximum carbon storage perspective. Modeling ecosystems productivity with process models allows hypothetical scenarios to be tested and can also help constrain claims that landowners and governments make about the carbon they are sequestering and storing. Generating reasonable estimates of potential productivity is difficult both because of a lack of data and because of future climate change. However, by using inventory measurements and flux data to calibrate process models, specific locations across the globe can be accurately represented. A process model parameterized in this way can then be used to expand estimates of productivity across space when paired with remote sensing data. This integration of multiple data sources at multiple scales can provide flexibility in estimating ecosystem state and allow for estimates to vary based on different future climate scenarios.

1. Introduction

1.1. Motivation and Applications

Between 6-17% of the total annual anthropogenic CO_2 emissions come from terrestrial ecosystem degradation or loss making up the second largest source of greenhouse gases in the world after fossil fuel emissions (Van der Werf et al. 2009). Because of this, many policy makers have focused on reducing emissions from terrestrial ecosystems as one way to help mitigate climate change. The estimation of carbon stocks and fluxes that result from land use change and ecosystem management is critical for policies that attempt to incentivize increased ecosystem sequestration or reduced emissions of greenhouse gases (GHGs). Furthermore, estimates of potential sequestration rates in optimal conditions, as well as anticipated changes in growth rates due to climate change, are important in understanding how emissions reductions or sequestration fit within a broader understanding of terrestrial carbon exchange.

For more than a decade, decision makers have debated how to best include terrestrial ecosystems in policy approaches to mitigate and adapt to climate change. Recently, these discussions have focused on REDD (Reducing Emissions from Deforestation or Ecosystem

Degradation) at a country scale. A REDD policy framework therefore requires credible estimates of the state of ecosystem carbon stocks and fluxes at a country scale as well as some understanding of how ecosystems have changed in the past and might change in the future at this scale. As policy makers in the international arena struggle to craft national REDD policies, the voluntary carbon market has rapidly evolved to fill the void left by the absence of international climate change policy. The voluntary carbon offset market relies on a suite of different carbon offset standards (CCBA 2008, VCS 2008, CAR 2010) that provide guidance on how to monitor, report, and verify carbon sequestration activities at much smaller scales (e.g. – project level vs. country scale). Given the state of REDD policy and the voluntary carbon market, the purpose of this chapter will be to explore a technical approach that can be used to generate credible estimates of carbon stocks and fluxes and to constrain the claimed benefits of carbon projects or policies at multiple scales.

There are several methods using a variety of data sources to arrive at reasonable estimates of ecosystem growth or carbon storage at reasonable cost. A data assimilation approach may best leverage existing data sets and improve the precision of estimates across space and time. These estimates could then be applied at varying scales to provide an independent assessment of claimed climate change mitigation benefits. Today there still is a critical need for defensible, consistent, and understandable estimates of spatially explicit ecosystem carbon stocks and fluxes as well as the continued need for modeling scenarios that simulate the outcomes of policy decisions on the future state of ecosystems. These models will allow policy makers and land managers to better understand the implications of new policies and the role of ecosystems in climate change mitigation.

1.2. Constraining Mitigation Claims and Future Growth

In addition to accurate estimates of current ecosystem carbon stocks and fluxes, an understanding of the potential uptake or emission of CO_2 is necessary to define the bounds of the direct impacts ecosystems can have on the climate system. In most cases, ecosystem sequestration is measured against a Business as Usual (BAU) baseline scenario. The BAU scenario is a hypothetical counterfactual description of how an ecosystem would change without implementing a carbon project or emissions reduction policy. For example, the BAU baseline for a country like Brazil might be the average rate of forest loss over the past 20 years (Ewers et

al. 2008) extended into the future to serve as the predicted rate of deforestation. The activity that occurs after the policy or project begins is, however, constrained more by the potential productivity of the system. Therefore, when a country or project proponent claims a climate benefit, these claims must reside within the limits of how fast an ecosystem can grow and sequester carbon dioxide.

----------SIDEBAR----------

Chapter Terminology and Conceptual Underpinnings (Sidebar)

For the purposes of this chapter, we will focus on Net Ecosystem Production or NEP. NEP is Gross Primary Production minus autotrophic and heterotrophic respiration – NEP = GPP – R_A – R_H – and represents the net CO_2 sequestered by an ecosystem. For a more complete discussion of NEP and other carbon cycle concepts see Chapin III et al. (2006) or Waring and Running (2007). This understanding of potential productivity helps to inform both regional scale carbon estimates and a forest carbon offset project's long-term climate impact. However, many projections of ecosystem change assume constant climate conditions. To more completely understand the climate benefits of that proposed climate policy or a single forest carbon offset project create, understanding ecosystems' responses to climate change should be incorporated into the modeling of future ecosystem growth and change. Both of these needs (constraining current claimed climate benefits and defining the future ecosystem dynamics) require not only the best suite of data products such as forest inventory data, flux measurements, and satellite observations, but also incorporating some process modeling that can capture the range of productivity and the impacts of a changing climate.

Forest ecosystems sequester carbon dioxide (CO_2) and emit oxygen as photosynthesis occurs. Photosynthesis produces sugar which is converted into starch and other carbon based molecules that trees and other plants then store and use for growth and maintenance. The CO_2 sequestered by trees is stored in their woody biomass and as individual trees grow and die, there is a cycle of sequestration and decay (Larcher 2003, Lambers et al. 2008). At a landscape scale however, the individual tree dynamics in most cases combine to form a saturating dynamic. Over time, as the forest ages, tree mortality due to disease, age, or disturbance creates openings where new trees grow. At this scale, there is a theoretical sigmoidal increase in the carbon stored in a forest over time asymptotically approaching a maximum stored biomass (Waring and Running 2007).

---

2. Data Sources and Carbon Cycle Background

There are three quantities that are typically estimated to assess an ecosystem: its stocks or current state, its fluxes or rates of change, and its future state based on a set of assumptions about stocks and fluxes in the future. Predicting the carbon stored at a given point in time and space, as well as over time, has been done by using process based physiological models and empirical growth and yield models (Vanclay 1994, Thornton et al. 2002, Turner et al. 2004, Arney et al. 2007, Randerson et al. 2009, Shoch et al. 2009, Dixon 2010). Establishing the state of a forest ecosystem in the present is done by measuring the current forest using field-based plots to estimate stocks or flux-towers to estimate fluxes. In addition to these ground based measurements, remotely sensed images of ecosystems can provide valuable information that can be used to infer some of the ground based parameters across broad spatial extents.

2.1. Estimation of Ecosystem Carbon Stocks

2.1.1. Forest inventory and forest growth and yield models

Forest inventory and forest growth and yield models have been used by foresters for over 100 years to estimate the volume of timber found in a given area and to predict the timber yields into the future. Many growth and yield tables developed in the 1950s and 1960s are still the primary source of information when predicting ecosystem changes over time and are still used today (e.g.,(King 1966)). Measuring the biomass in a forest involves installing plots on the ground and measuring trees – both live and dead, dead material, and the soil to estimate the conditions of a forest. Traditional forest inventories were used to estimate the volume of merchantable board feet and to understand how much a forest was worth in terms of its timber value.

Because of climate change policies and the voluntary carbon market, forest inventory data is now also being used to estimate the stocks of carbon in ecosystems. Land being managed to produce timber most likely has been extensivly inventoried, If no inventory data exists, there are many manuals that describe procedures to collect inventory data and the rationale for what data to sample (Shiver and Borders 1996, Avery and Burkhart 2002, Law et al. 2008, GOFC-GOLD 2009). If collected over time, forest inventory data can be used to characterize ecosystem change. This stock change approach to estimating change is a common alternative to direct measurement of fluxes.

Forest inventory data collected by private landowners while valuable, can be difficult to obtain because it is usually proprietary. Also, data collected by private land owners are often not spatially extensive enough to inform larger, country or regional scale assessments of stocks. Many nations have developed nationwide inventory systems to meet the data needs for forest monitoring and carbon stock accounting (Gillis 2001, Kitahara et al. 2009). In the United States of America, this system is called the Forest Inventory and Analysis and there is roughly one plot every 2500 hectares across the US (Bechtold and Patterson 2005). These data are collected over time and in some places there may be re-measurements of plots. Data of this sort are critical both to have means to assess the state of ecosystems but also to validate and verify estimates of ecosystem state that are derived from other products such as remote sensing or models.

2.2. Estimation of Ecosystem Carbon Fluxes

There are several methods to measure how ecosystems are changing – either growing and sequestering carbon or losing carbon through disturbance and decay. The direct measurement of the changes in ecosystems is accomplished using either flux towers to directly measure the gas exchange that occurs in the eddies that form in and around vegetation canopies or using smaller, distributed networks of below canopy gas exchange measurement devices. Leveraging these spatially disjoint but locally intensive measurements is then accomplished using remote sensing data. In particular, large scale remote sensing products are critical for two elements of ecosystems analysis: 1) detecting ecosystem disturbance and 2) estimating ecosystem growth in the form of Net and Gross Primary Productivity (NPP and GPP).

2.2.1. Flux Towers

Flux towers are located around the globe and usually are large towers that are taller than the tallest member of the ecosystem they are studying which enables them to measure the exchange of gases into and out of the ecosystem (Munger and Loescher 2006). Flux towers are a relatively recent addition to the field-based data collection methods. Flux towers measure the exchange of gases over an ecosystem. Over the course of a day, as ecosystems

Figure 1: The Metoliuos Flux Tower located in a Ponderosa Pine Forest in Easter Oregon ("Carbon Uptake by Forests" n.d.)

photosynthesize and respire, vapor, CO_2, O_2, and water vapor are transferred to and from the ecosystem and to and from the atmosphere. Flux towers, using a suite of technologies, measure the exchange of these gases. The results are then used to estimate the total ecosystem flux of both water and carbon. The area that a flux tower measures is referred to as its "footprint" and varies in size in relation to the height of the tower. Typically, the tower footprint is 10 times the height of the sensor, however for very tall towers, (e.g. the 450m tall Chequamegon Ecosystem-Atmosphere Study Flux Tower located in northern Wisconsin) the tower footprints can be 100 times larger than the sensor height (Davis et al. 2003). Unfortunately, flux towers are expensive to build and maintain (about $100,000 a year for the minimal installation and equivalent costs for maintenance and operation of the tower) so there are relatively few of them and they cover an exceedingly small proportion of the total landscape.

Figure 2: The ChEAS Flux Tower is an example of a very tall tower with a large footprint.

2.2.2. Distributed sensor networks (DSNs):

DSNs are another technology that is even less mature than flux towers but holds some promise in measuring ecosystem change across space. Distributed sensor networks (DSNs) incorporate some of the same equipment that is used to measure instantaneous change on flux towers. The sole difference between DSNs and flux towers is that DSNs have sensors distributed across space as opposed to being anchored to a flux tower at a single point. In most installations, these sensors are placed below the canopy, but there is nothing (beyond cost concerns) to prevent a network of above canopy sensors similar to a grouping of flux towers. These distributed networks of sensors are expensive relative to traditional inventory but are less expensive than a full flux tower installation. By deploying a network of sensors that are connected in real time to a larger network, researchers can remotely monitor instantaneous changes in ecosystem fluxes as well as the drivers of these fluxes (Hart and Martinez 2006, Rundel et al. 2009, Yang et al. 2009). Sensors of this type are used to measure CO_2 and water vapor exchange above the soil, flow rates in streams, soil temperature and moisture, as well as microclimate data below the canopy and across space. They can be thought of as miniature flux towers placed below the canopy.

2.3. Remote Sensing Data Products

Remote sensing data can be acquired either from satellites or aircraft observations. There are many different types of remotely sensed data defined by the scale of the data received. The spatial scale is defined both by the resolution or pixel size of the imagery generated and the area covered by the data (the spatial extent). The temporal scale is defined by the time interval over which data are collected. Generally, today, the finer the spatial reolution the less frequently are the observations collected. Remotely sensed data can also be broadly split into two groups: passively acquired observations or active response observations.

2.3.1. Passive remotely sensed observations

Passive remotely sensed observations are acquired by sensors that measure either reflected solar radiation or emitted terrestrial radiation. For example, Moderate Resolution Imaging Spectroradiometer satellite data (MODIS) is a passive sensor system that acquires measurements in the visible, near infrared, shortwave infrared and thermal infrared portions of the EM spectrum. MODIS observations are collected gobally on a near-daily basis, with 250m, 500m, or 1km pixel sizes. Landsat Thematic Mapper observations are collected at a 30m spatial resolution with a global return frequency of 16 days. The earlier Landsat Multispectral Scanner observations (1972-1992) were collected at 80m spatial resolution and 18 day repeat cycles ("Landsat Missions" n.d., "MODIS Website" n.d., "The Landsat Program" n.d., Lillesand et al. 2004).

Remotely sensed data are useful because they can be used in difficult to reach areas and can provide wall-to-wall coverage of areas over time at scales where this density of data would be cost prohibitive to collect in any other way. Given the features of remotely sensed data described above, they are also a critical tool in monitoring and identifying when major changes occur in an ecosystem. For example, changes in canopy cover or leaf area as inferred from greenness and the fraction of absorbed radiation can be used to estimate annual NPP or GPP (Running, Nemani et al. 2004). When compared over time, remote sensing products can highlight areas of ecosystems that have experienced disturbance (Goward et al. 2008, Mildrexler et al. 2009, Huang et al. 2010) and these products can then be used to estimate the flux of carbon loss to the atmosphere.

2.3.2. LIDAR and RADAR Active Remote Sensing

Active remotely sensed observations are acquired by generating a pulse of energy and then measuring features of the response at the surface as this energy pulse is reflected. The characteristics or signature of this backscatter at different locations can then be mapped to physical attributes of the Earth's surface. LiDAR (Light Detection and Ranging) is one form of an active remote sensing system. The other commonly used is radar (radio detection and ranging). Due to the characteristics of these active remote sensing systems, LiDAR and radar can provide an assessment of the three-dimensional structure of vegetation canopies.

Because ecosystem biomass and carbon storage are in many cases closely related to height, and these active remote sensing approaches can estimate height, they can help to define the biomass storage of ecosystems across space. As such, radar and LiDAR have been used to estimate forest height and biomass (Patenaude et al. 2004, Akay et al. 2009, Collins et al. 2009, Goetz et al. 2009, García et al. 2010). Two important caveats to this are: 1) radar signals may saturate in ecosystems that store large amounts of carbon (Kasischke et al. 1997) and 2) both of these data sources require ground data to calibrate and validate their results. Radar is particularly useful for forest sensing as the wavelengths it uses can penetrate cloud cover and allow monitoring of ecosystems that are consistently covered in clouds such as moist tropical forests (Kasischke et al. 1997, Sãnchez-Azofeifa et al. 2009).

In addition to the remote sensing of forest structure, remotely sensed data can be used to make inferences about ecosystem change (both disturbance and flux). As discussed above, LIDAR and RADAR can both be used to estimate the stock of specific ecosystem variables (e.g. carbon). As with the measurement of carbon stocks using ground based forest inventory, flux towers and distributed sensor networks are the only direct way to measure ecosystem fluxes over a short time scale. Like forest inventory data, these flux measurements serve a critical need by providing data that can be used to validate and train ecosystem models and remotely sensed estimates of fluxes. The estimation of forest stocks and fluxes using remote sensing data draws on an extensive body of research modeling the processes that take place in ecosystems as well as empirical relationships between the observed remotely sensed data and the measured stocks or fluxes. However, to predict the capacity of a system in the present, or the potential storage or sequestration of a system in the future, process models are needed.

2.4. Terrestrial Ecosystem Carbon Models (TECMs)

2.4.1. Process based physiological models

Process based physiological models use our understanding of the workings of photosynthesis, respiration, the physics of water movement and state change, and decomposition to estimate the growth dynamics of an ecosystem. The scale of a given model determines what processes are included and how the model behaves. For example, BiomeBGC works at a single point in space and uses both daily and annual time steps to estimate leaf-level photosynthesis. This model works at the leaf level to model photosynthesis, rather than using a simpler light use efficiency (LUE) model (a LUE model would apply a light conversion efficiency factor to incoming solar radiation to estimate the amount of fixed carbon dioxide). BiomeBGC accepts meteorological, soil, ecosystem type, and atmospheric CO_2 concentration inputs and uses these variables to drive the model. Other models such as C-Fix (Maselli et al. 2008) or the CASA model (Potter et al. 1993, 2003) work at slightly larger scales and, rather than focus on within-leaf physiology, use the fraction of absorbed photosynthetically active radiation (FPAR) to estimate the photosynthesis of leaves. Other models such as ED (Albani et al. 2006) are demographic or gap models that model tree growth and competition based on tree size and forest structure. Another example of a process model that works at a higher level of abstraction is the 3-PG model. This model estimates growth and storage based on light use efficiency and scaling values based on water availability, nutrient availability, and a suite of other constraining variables (Landsberg and Waring 1997, Sands and Landsberg 2002, Landsberg et al. 2003).

2.4.2. Empirical Growth and Yield Models

In addition to the process models described above, empirical models such as forest growth and yield models (GYMs) can help to estimate the potential of a system in the present and into the future. At their core, GYMs are built from empirically derived relationships between stand (a stand is a contiguous forest area with similar conditions) characteristics such as density, height, age, and site class against stand volume or biomass (Avery and Burkhart 2002). Individual tree GYMs use a similar approach as stand GYMs but relate stand characteristics to individual tree growth as opposed to overall stand growth (Porté and Bartelink 2002). The data used to drive these models can come from long term permanent plots showing forest development over time or can be taken from many different forests of different ages, site conditions, and stocking rates to build the appropriate relationships. Because GYMs use data from past forest growth, GYMs implicitly assume that past drivers of growth such as climate and

atmospheric CO_2 levels will not change enough to dramatically impact the growth dynamics of forests in the future. For short time scales, this assumption may be valid but for longer timescales, this is probably an inappropriate assumption. Because of these shortcomings, process models should be used when the questions of interest relate to how changing environmental conditions will effect forest growth.

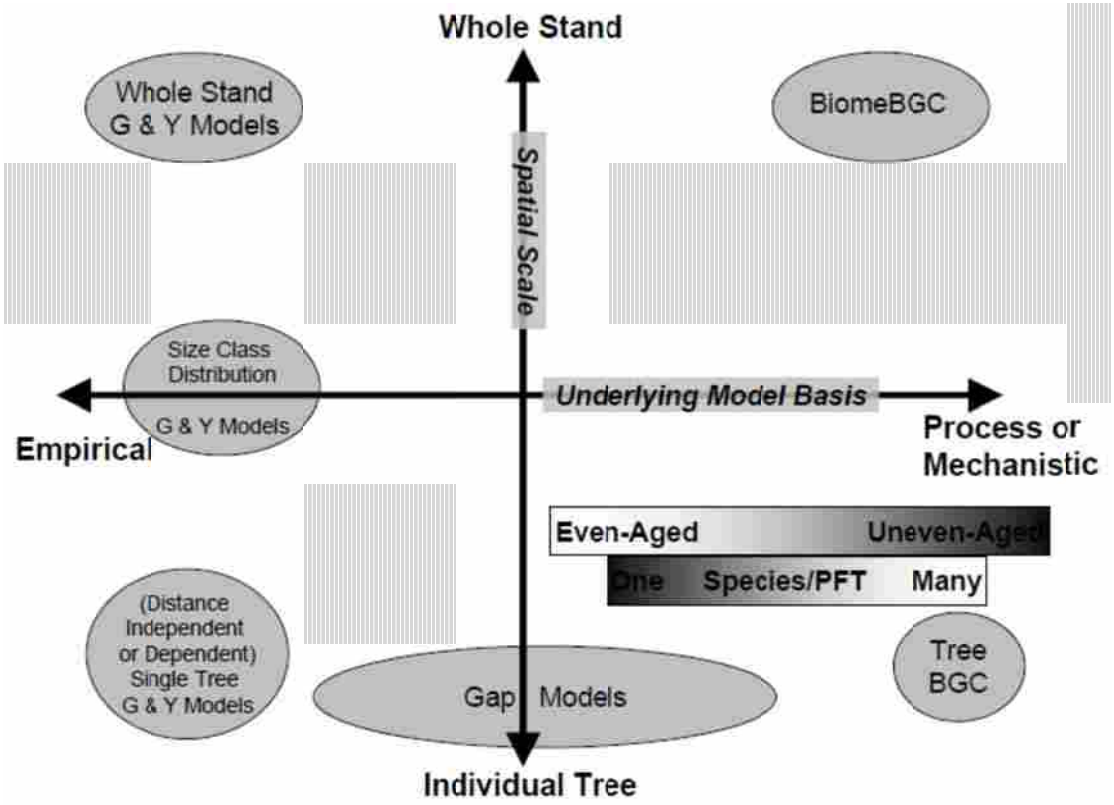

Figure 3: Conceptual ecosystem modeling continuums. 1) Model Basis vs. Spatial Scale and 2) Age Structure vs. Species/PFT composition – Different model logic based on spatial scale: individual tree to stands – temporal scale: hourly to annually – and the abstractions they choose to make. This diagram also shows two sets of grayscale ramped axes that could be used to color each model oval depending on the model type: one species/plant function type or many and mixed-age systems or even-aged systems.

The background above details the data sources that are available to make estimates of the stocks and fluxes of ecosystems. Forest inventory data forms the core of how scientists and managers have traditionally understood forest systems. Direct measurements of fluxes either at the micro scale with individual sensor installations or across space with distributed sensor networks or with larger flux towers are the backbone of the validation data available for models that predict ecosystem fluxes and need flux data both for parameterization and to constrain their results. Using both forest inventory data and flux data, remotely sensed data and process models can be parameterized, constrained and validated to accurately represent the state and changes of

ecosystems across space (e.g. - (Pietsch et al. 2005, Heinsch et al. 2006, Randerson et al. 2009)). The power of remote sensing data is that it allows the earth system to be observed across large spatial extent at relatively frequent intervals. Once remotely sensed data are paired with models and measurement to generate accurate estimates of the current stocks and fluxes, the next step is to use models to help predict the future ecosystem state in terms of the ecosystems' potential to grow and store carbon.

3. Example – Using BiomeBGC to estimate ecosystem states and fluxes across space

Given the broad range of data sources and models available to estimate carbon stocks and fluxes, it is helpful to consider an example system to elucidate some of the principles that will be discussed in this chapter. With the goal of illustrating both the state of the science and some of the shortcomings of current approaches to the estimation of stocks and fluxes across space, the BiomeBGC model will be used to estimate the state of a forest located in Mendocino County, CA. The Garcia River Forest is a moist temperate rainforest dominated by Redwood (S*equoia sempervirens*) and Douglas-Fir (*Pseudotsuga menziesii*) that is about 10,000 hectares in size. This forest is actively managed and there are relatively accurate biomass estimates across the full 10,000 hectare extent from forest inventory data.

3.1. BiomeBGC Model Background

BiomeBGC (BiomeBGC) is a mechanistic model that is used to estimate the state and fluxes of carbon (C), nitrogen (N), and water (H_2O) into and out of an ecosystem. BiomeBGC is actively used in institutions around the globe and its most recent release is version 4.2. In addition to the C, N, and H_2O cycles, BiomeBGC models the physical processes of radiation and water disposition. BiomeBGC partitions incoming radiation and precipitation and treats the excess/unused portions as outflows. The primary physiological processes modeled by BiomeBGC are photosynthesis, evapotranspiration, respiration (autotrophic and heterotrophic), decomposition, the final allocation of photosynthetic assimilate, and mortality. To model these processes, BiomeBGC first models the phenology of the systems based on the input meteorological data (Thornton 1998, Thornton and Running 2002, Golinkoff 2010).

The general flow of the BiomeBGC model is based on an abstraction of how natural ecosystem operate (Figure 4 and Table 2).

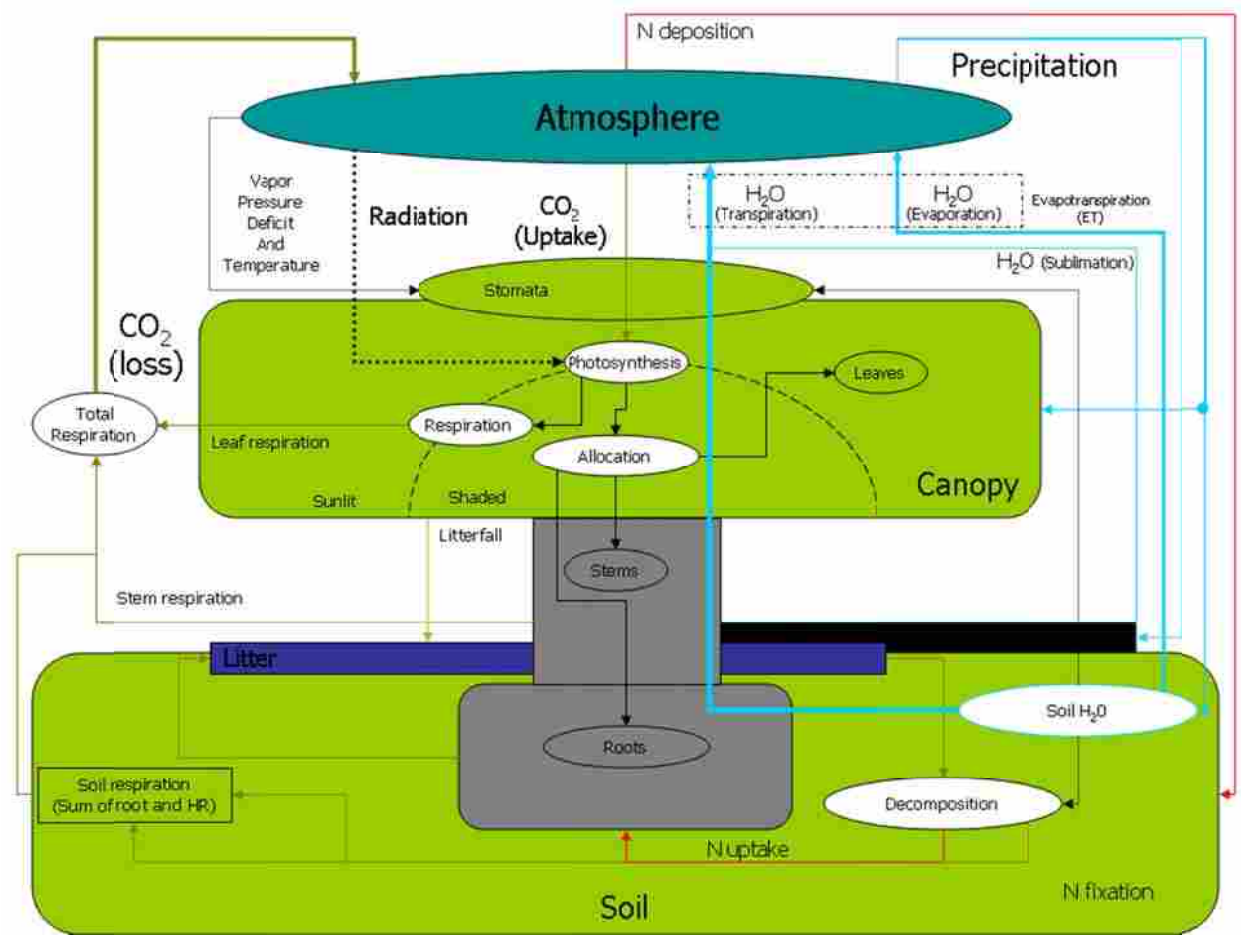

Figure 4: Conceptual diagram showing BiomeBGC general model structure ("Biome-BGC - Conceptual diagram" 2005)

Any model run requires a certain set of input data. BiomeBGC requires meteorological, physical, and ecophysiological data for each site. Every model run then produces a set of data that can be outputted for the user to analyze. These variables include all of the C, N, and H_2O fluxes and pools that BiomeBGC tracks as well as summary variables (e.g. -Net Ecosystem Exchange (NEE) or Net Primary Productivity (NPP)) at daily, monthly, or annual time scales. BiomeBGC can be run to a spinup steady state and then forward in time, or it can accept as an input the ending model state of a previous model run (a restart file) and run from this point forward with a new set of model assumptions if desired.

Table 2: A partial list of inputs that must be provided to the BiomeBGC model.

Data Type	Name	Units	Description
Climate Data	Tavg	Degree C	Average Daily Temperature
	Tmin	Degree C	Minimum Daily Temperature
	Prcp	cm	Daily precipitation
	VPD	Pascals	Average daily Vapor Pressure Deficit
	SRAD	W/m^2	Daily Solar Radiation
Ecophysiology Data	C:N of leaves	kgC/kgN	Carbon to Nitrogen ratio of leaf biomass
	Annual Mortality	1/yr	Annual whole plant mortality fraction
	SLA	m^2/kgC	Canopy average specific leaf area (projected area basis)
	FLNR	no units	fraction of leaf N in Rubisco
	VPAstart	Mpa	The vapor pressure deficit where leaf conductance begins to be reduced.
	VPAcomplete	Mpa	The vapor pressure deficit where leaf conductance is zero.
Site Data	CO_2	ppm	A constant or a file with changing yearly atmospheric CO_2 concentrations
	SoilDepth	m	Effective soil depth (corrected for rock fraction)
	% silt, sand, clay	%	Percentage silt, sand, and clay in soil
	Elev	m	Site elevation
	Latitude	degrees	Site latitude
	Ndep	$kgN/m^2/yr$	Wet and dry atmospheric deposition of N
	Nfix	$kgN/m^2/yr$	Symbiotic and asymbiotic fixation of N

BiomeBGC is run at one point in space and for estimation of ecosystem states across space, the point returns are simply gridded to create a spatial estimate. However, because of the structure of BiomeBGC, BiomeBGC does not incorporate cell-to-cell interactions or flows of nutrients or water between cells. Given BiomeBGC's point-based perspective, it is helpful to think of this model as an estimate of stand level processes that have been aggregated and averaged to a per unit area basis. In general, this model divides photosynthesis between shade leaves and sun leaves. The carbon fixed by these leaves is then partitioned to other organs with the theoretical tree as well as into soil carbon pools. Carbon is also modeled as lost to respiration both for maintenance and growth. A full discussion of the details of how BiomeBGC works is beyond the scope of this chapter but several references have been included here to aid the reader

should they want more information (De Pury and Farquhar 1997, White et al. 1997, 2000, Thornton et al. 2002, Koch et al. 2004, Thornton and Rosenbloom 2005, Thornton and Zimmermann 2007, Golinkoff 2010).

3.2. BiomeBGC Model Parameterization

Figure 5: The Redwood Region of California in dark grey (see Johnstone and Dawson 2010)

To parameterize BiomeBGC for a given location, there are many physiological, phenological, and site characteristics that are needed. Without these data, there already exists a set of default values for different plant functional types (e.g. evergreen needleleaf trees or deciduous broadleaf trees) for BiomeBGC within the ecophysiological initialization file. However, these defaults may not adequately capture the details of a specific site. The parameterization of BiomeBGC for the Garcia River Forest uses data from several published studies from the Redwood Region. The Redwood Region is a thin strip of land that runs along the coast of Northern California and is often shrouded in fog (Johnstone and Dawson 2010) (Figure 5). The default ecophysiological parameters were taken from White et al. (2000) with some parameter modifications based on research done in the Redwood region (

Table 3).

Table 3: Modified BiomeBGC ecophysiology parameters

EPC Variable	Value	Units	Study Authors
Specific Leaf Area	13.8	m^2/kgC	(Ambrose et al. 2009)
C/N Litter	118	kgC/kgN	(Zinke and Crocker 1962)
Canopy H_20 interception	0.0041	1/LAI/day	(Van Pelt and Franklin 2000, Berrill and O'Hara 2003, Ewing et al. 2009)
Annual Mortality Rate	0.0017	1/yr	(Busing and Fujimori 2005)

3.3. BiomeBGC Input Data – Gridding Climate and Soil Driving Data

Once the model parameters have been established to reflect the ecosystem, the driving data for the model must be interpolated across space at an appropriate scale to represent the main topographical drivers of ecosystem productivity. The topographical drivers are slope, aspect, and elevation and impact the temperature and water availability at a given site. Additionally, the soil data must also be converted to a grid of the same extent and resolution as the driving meteorological data. The creation of the daily climate data used for this particular model run is a bit more complicated owing to the fact that redwood trees are known to absorb fog moisture through their needles (Weathers 1999, Burgess and Dawson 2004, Ewing et al. 2009, Simonin et al. 2009). However, the BiomeBGC model structure uses only soil water holding capacity when determining the moisture limitations of growth. Therefore, to address this "missing source" of plant available moisture, fog precipitation was added to the soil water at regular intervals across the year based on the measured amounts of fog water by month reported by Dawson (1998). The raw meteorological data were generated across the forest extent at 250 meter resolution using DAYMET (http://www.daymet.org/, accessed May, 2009). The DAYMET algorithms are based on the logic used by the MT-CLIM program (Thornton et al. 1997, 2000, Thornton and Running 1999, Hasenauer et al. 2003). The soil data used were percent sand, silt, and clay and this information was taken from the Natural Resource Conservation Service (NRCS) Soil Survey Geographic Database (SSURGO – see ("Soil Data Mart - Home." n.d.) http://soildatamart.nrcs.usda.gov/ - accessed May, 2009).

Once each 250m grid cell has the appropriate site (soil and topography) and climate driving data, an initialization file is created for each grid cell. This initialization file directs BiomeBGC to use the climate and site data provided, along with the physiology data defined as above to grow the ecosystem. The ecosystem is then grown until it reaches a steady state (i.e. – an old growth state). At this point, 95% of the aboveground biomass pools were removed to simulate the almost complete harvest of this area by the late 1950s. The harvested ecosystem was then grown for 50 years to simulate the average age of the forest today.

3.4. BiomeBGC Model Results

After parameterizing the BiomeBGC model and creating a set of gridded input data, the model was run as described in section 3.3 and the final model results were collected and summarized across space. BiomeBGC can generate many outputs but for this discussion we will focus on NPP and total carbon stocks (Figure 6). The results in this example were then validated at a property scale to the average carbon stocks measured using over 1000 forest inventory plots. The forest inventory plots used were variable radius plots that used a basal area factor prism or relaskope to determine which trees to measure. Variable radius plots use a probability proportional to size sampling approach that makes it more likely to measure larger trees (Avery and Burkhart 2002). The average carbon stocks on this property in 2005 based on these plots is about 130 Mg/ha. The average growth on a per year basis (NPP) is between 6.6 and 10 Mg/ha according to BiomeBGC. Based on the Forest Projection and Planning Systems (FPS) growth and yield model (Arney et al. 2007) as well as plots installed over time, the average growth rate is about .54 kg/m^2/year. At a property level, it seems that the process model results are reasonably close to the estimates of carbon stocks and growth as found by plot measurements and local growth and yield models.

The lower estimate derived from the growth and yield model and measurements reflects ongoing harvests on this property, which were not modeled using BiomeBGC. This validation of the model results is done at a very coarse scale (average across a 10,000 ha property) and not on a cell-by-cell basis. However, the rough agreement in the model and measurement is encouraging because it suggests that the BiomeBGC model logic is appropriately capturing some of the biophysical and physiological processes in this ecosystem and producing results that fall within the natural range of ecosystem variability. Furthermore, areas with the highest forest carbon stocks are found in stream bottoms as would be expected. Unfortunately, the annual NPP metrics should also correlate with the terrain in a similar way and they seem to have the opposite pattern with the highest productivity areas occurring near ridge tops and the lowest productivity areas in the stream bottoms. This result probably stems from the fact that the model is inadequately representing the moisture limitations experienced by the trees in this ecosystem (see section 3.3 about Redwood trees unique fog water uptake).

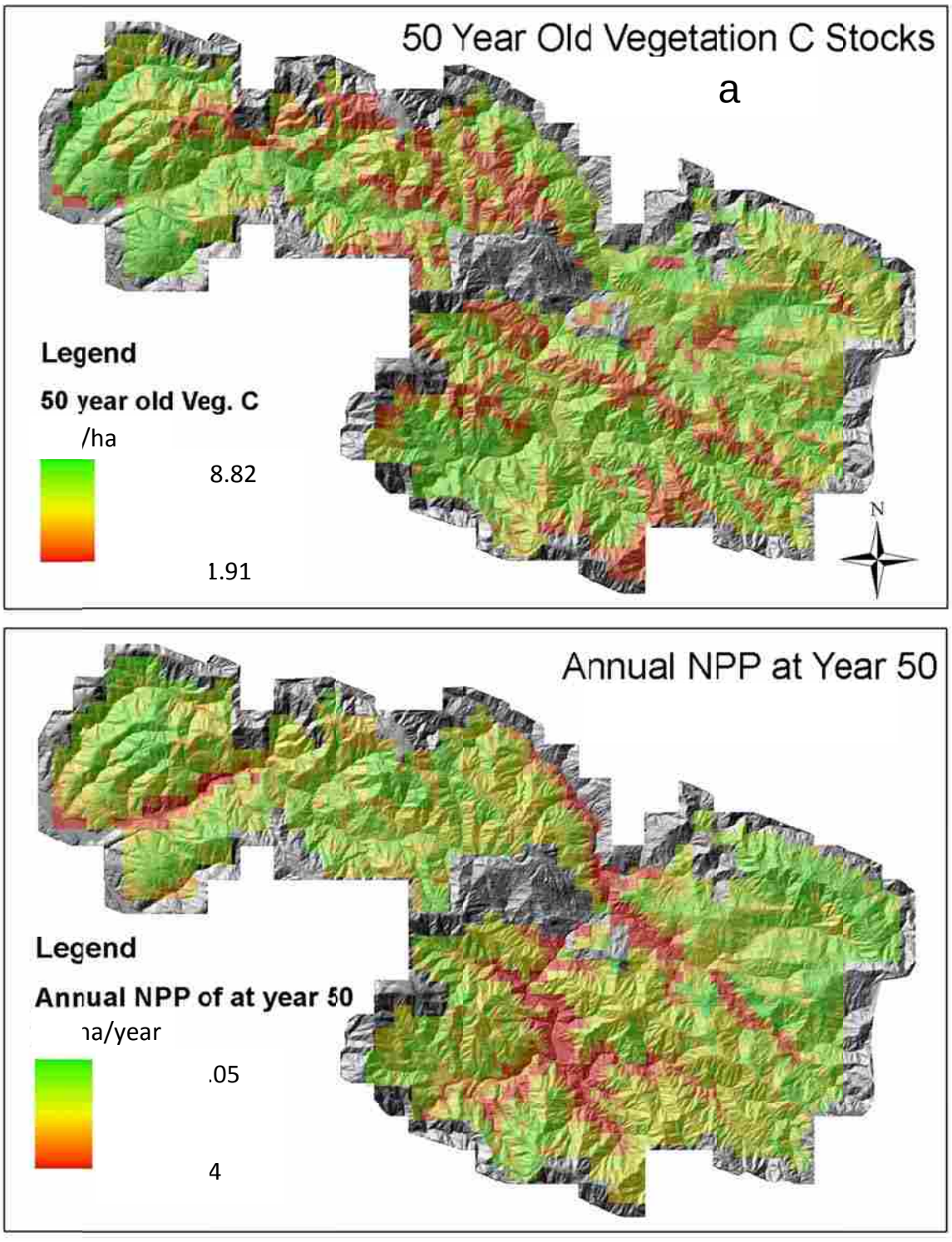

Figure 6: BiomeBGC Outputs Showing Vegetation C stored as well as Annual NPP draped over a Hillshade Image for the Garcia River Forest in Mendocino County, CA

3.5. Discussion – Shortcomings and Potential Directions

As shown above, as a result of parameterizing this model to accurately represent this site, and adjusting the precipitation to account for fog water use, the BiomeBGC model does a reasonable job of estimating average ecosystem states and fluxes when aggregated across a 10,000 ha extent. There are several problems with this approach however that suggest improvements that are needed. These are:

1. Incorporate the impacts of harvest on ecosystems into the BiomeBGC model.
2. Incorporation of estimates of the ecosystem age. These data are critical constraint when estimating the stocks and fluxes of an ecosystem.
3. Improvements to the model to more adequately represent the ecosystem physiology and structure. For example, in the case of redwoods, the BiomeBGC model does not adequately capture the foliar uptake of water and therefore a work-around approach to water availability must be employed.
4. Improvements in techniques to estimate certain parameters that are difficult if not impossible to know with any certainty. For example, an accurate estimate of soil depth across space often does not exist. However, the soil depth parameter in BiomeBGC is critical when determining when moisture availability will become limiting to plant growth.

Based on these concerns, a more general and consistent approach to parameterizing and using the BiomeBGC model spatially is needed. Specifically, the improvements listed above should be considered when considering how models like BiomeBGC might be applied to on-the-ground questions regarding the capacity of ecosystems to mitigate and adapt to climate change.

Given the errors seen in this particular modeling exercise, it is clear that models should be carefully evaluated and validated before they are applied in a policy setting. More importantly, the discrepancies discussed above illustrate the fact that although no modeling exercise is perfect, the results taken as a whole demonstrate that ecosystem function can be accurately represented. In any model run applied across space, it is possible to find areas of agreement and disagreement. Although ideally we would like to accurately model ecosystem states and fluxes at many different scales with little or no error, in practice this is unlikely if not impossible both because models rely on many types of input data that have their own errors and uncertainty (e.g. – soil data, past management data, etc.) and because of errors in model logic.

However, the purpose of models is to bridge the gaps of each individual data source and to better generate scalable estimates of ecosystems.

It is difficult if not impossible to adequately measure an entire watershed's carbon stocks and fluxes. Wall to wall measurements become even more infeasible at a country scale due to time and cost constraints. However, a model system can incorporate sparse measurement data as well as remotely sensed data to generate more accurate wall-to-wall estimates of ecosystem state. While no model is perfect, the ability to generate estimates of an ecosystem state and fluxes at multiple scales both temporally and spatially provides a strong rationale for model use despite the inevitable errors found at particular locations or points in time.

4. Estimating Potential Productivity

The productivity of an ecosystem is generally thought of as the rate at which an ecosystem can sequester carbon. Productivity can also be thought of in relation to the maximum ecosystem storage of carbon. Younger forests for example can be thought of as highly productive when they are rapidly adding biomass and sequestering CO_2. The coastal redwood forests in California can also be considered some of the most productive forests in the world given that in their climax state, they can store more carbon than any other ecosystem (Busing and Fujimori 2005). The potential productivity of a given site can therefore be thought of as either the maximum rate that biomass or carbon is accumulated or it can be thought of as the maximum amount of carbon stocks that the system can eventually store given the forcing variables of the site conditions and the climate. These two concepts can be distinguished as a rate potential and a state potential. Formally, the rate potential of a given system is the maximum possible rate of ecosystem carbon uptake ($NEP = GPP - R_A - R_H$) given the climate and site constraints. The state potential is the maximum possible biomass storage at a late successional state given the climate and site constraints ($state_{\max} = \int_0^{SS} NEP$, where SS is the steady state climax state defined by a little or no change in soil carbon stocks) . The difference between the rate and state potential scenarios and the actual measured or observed scenarios is therefore the influence of human disturbance and/or management. Conceptually, the rate and state potential are valuable for constraining claimed climate mitigation benefits as well as for better understanding of the impact of land-use change on the carbon cycle.

Because of the theoretical nature of potential productivity, process models are required to estimate values for rate and state potentials. Furthermore, process models are essential when considering the impacts of future climate change on potential rates and states. However, as discussed above in section 2, before a process model can be applied to a single point in space, let alone across a large spatial extent, there are many data sources that must be collected, organized, and processed both to drive the model and to parameterize the model to accurately apply to the location it is intended to model (see

Table 3). This is a non-trivial exercise and there is a need for a streamlined approach to parameterizing and applying model logic across space.

4.1 Data Assimilation – How Process Models Can Incorporate Measurements

As described above, process models use our understanding of how terrestrial ecosystems work to model how ecosystems grow and change over time. There are many different models and model types and each model's focus and purpose will in some way dictate how it is designed and what processes have been incorporated into the model logic. Regardless of the model used, data measured at the site to be modeled can be used both to parameterize the model to better estimate the site and to validate the results of the model runs. This process of incorporating a variety of data sources (model results, model structure, measurements at a given location, remote sensing data) is broadly described as data assimilation.

Data assimilation (DA) has been used extensively in many fields. Within the earth sciences, DA is most developed within the atmospheric and oceanographic communities and is used to estimate large scale atmospheric transport of gases constrained by point measurements of gas concentrations from flasks or flux towers (Evensen 2003, Mathieu and O'Neill 2008, Reichle 2008). However, the idea to leverage multiple data sources to better estimate Earth system processes and the application of these methods has become state of the art in the terrestrial ecosystem modeling community as well (Running et al. 1999, Knorr and Kattge 2005, Thum et al. 2007, Wang et al. 2007, 2009).

The general idea of DA is a model-data fusion (MDF) whereby a model is constrained and parameterized by the available data to generate model outcomes that are closer to data observations. One approach to this MDF for simple systems that can be represented in closed-form equations is simply to invert the model given the measurements to "solve" for the

parameters. For example, if an ecosystem model could be represent as a linear system that consisted of a set of operators that mapped some input variables to an output set of variables, this system could be formally written as (Equation 1):

Equation 1: $y = Z * \beta$

Where y is a vector of the ecosystem output state, Z is a matrix of predictor values, and β is a vector representing the model parameters. Using simple linear algebra to solve this system given known inputs and measured ecosystem state variable, the model structure could be calculated (Johnson and Wichern 2002) (Equation 2).

Equation 2: $\beta = Z^{-1} * y$ (or if Z is not invertible $\beta = (Z' * Z)^{-1} * Z' * y$)

Most Terrestrial Ecosystem Models (TECMs) however are too complicated to be represented in this way both because they are non-linear and because their form prevents a simple representation and therefore other approaches are necessary to help parameterize them. Regardless of how the system is represented, the basic structure of a DA approach is to consider the forcing variables that drive the model behavior, the model structure – i.e. the parameters and logic – as a function, a set of initial conditions, and the output state of the system. Model systems can be represented using either a continuous form (Equation 3) or adiscrete form (Equation 4).

Equation 3: $$\frac{dx}{dt} = f(x, u, p) + noise$$

Equation 4: $$x^{n+1} = x^n + \Delta t f(x^n, u^n, p) + noise$$

where x is a vector of the state variable, u is a vector of forcing variables, and p is a vector of the model parameters. f is the model logic that is applied to these inputs and results in a new set of state variables defined by the rate of change to the system dx/dt (Raupach et al. 2005).

The observed data can also be considered as a function. In the case where the observed data exactly matches the variables in the state vector generated by the model, no model is needed and the observations alone are used. However, in many cases, the measured variables need to be converted to analogues of the model outcomes both from a scale perspective and in that the observed data may be surrogates of the actual quantities that are modeled as opposed to the variables themselves (e.g. – we may measure standing volume of a forest but the TECM predicts the carbon content of an ecosystem on a per area basis). The desired latent variable can then be modeled as a function of the observed data (Equation 5):

Equation 5: $z^n = h(x^n,u^n) + noise$

where x is a vector representing the measured data at time n, u is again the vector forcing variables, z is the desired latent variable, and h is a function to convert the data and forcing variables to a set of data to constrain or parameterize the model (Raupach et al. 2005).

Once the model and data are represented in this manner, the DA method proceeds to estimate a set of target values. These target values can be parameters of the model, outputs of the model (e.g. – state variables), or even the error structure itself. These target values are the values that the DA attempts to constrain and refine. With the target values defined, the final step in the DA process is to estimate the target values by minimizing a cost function that considers the data values as well as the uncertainty of the data values. The uncertainty of the model is considered the representation error. This includes the uncertainty of the parameters as well as any uncertainty associated with the model logic. The uncertainty of the observations is the natural variability of the estimates as well as the error associated with the measurements of these data. In most cases, the representation uncertainty should be larger than the observation uncertainty (Raupach et al. 2005). An optimization approach is used to find a global minimization of the cost function and by doing so to generate estimates of the target values (Wang et al. 2009).

DA methods can be broadly separated into sequential or non-sequential methods. Sequential methods consider new data over time and use these observations to constrain multiple timesteps of a model. Non-sequential methods, or batch methods, consider all of the observational data and model outputs at one time when estimating target values (Raupach et al. 2005, Wang et al. 2009). Non-sequential approaches are often used for parameterizations that then guide model runs given a set of initial observations. Sequential approaches are best used when the data observations occur over time and the model states also occur at more than one point in time. Non-sequential approaches are powerful in that they use all of the available data at once to estimate the target values. However, this approach can also be problematic from a computational capacity perspective with extremely large data sets. Conversely, sequential approaches effectively break problems into smaller, more manageable pieces and allow for changing model states over time.

The summary of DA presented above is a broad overview of how this process works. One important aspect of DA is that the final estimates of the target values are largely determined

by the uncertainty associated with the model and the observations. More certain quantities will be weighted more heavily and will therefore have more influence on the final outcome. As noted above, the uncertainty of the observed data may be large but in most cases should be less than the uncertainties associated with the model. In many cases, the model uncertainty can be highly subjective and is a subject of expert opinion and qualitative analysis. Because the outcome of a DA is highly dependent on the uncertainties of the data sources used to constrain the model, it is important to accurately and consistently collect uncertainty data whenever possible.

Although the DA method has the potential to improve parameterizations of models and also improve the model's estimates of ecosystem states, these approaches are not perfect and there are several caveats to consider when using DA:

1. Fox et al. (2009) have shown that when DA methods were applied to synthetic results that had noise added to them, many DA results failed to adequately estimate model parameter values. If DA cannot estimate parameters from a system where the true parameter values are known, it is possible that DA will fail to adequately capture the dynamics of natural systems.
2. There is often a mismatch between the scale and intensity of observed data and the model outputs (Raupach et al. 2005). Converting the observed data to equivalent scales (both spatial and temporal) is both a sampling problem and a modeling exercise and has the potential to introduce new and large uncertainties to the observed data (see Equation 5 above).
3. Most DA techniques assume unbiased error structures. In the presence of biases, DA could result in biased estimates of the target values.

Despite these hurdles, DA has been used to successfully constrain ecosystem modeling exercises and informs current research efforts in this field. For the purposes of land use change and carbon cycle modeling, these methods are particularly helpful because they allow the multiple observational data sources outlined above (forest inventory, flux towers, distributed sensor networks, and remote sensing data) to be effectively combined and used to constrain TECMs. Another critical need these methods meet is the ability to automate some of these calibration processes so that researchers do not need to parameterize each model location individually but instead can use an automated process. Using DA in multiple phases can also allow for many data sources to be successfully integrated into the final model structure (Zhu et

al. 2009). Lastly, pairing process models with sequential DA approaches and remotely sensed disturbance indices could allow for real-time adjustments of model results to estimate carbon stocks and fluxes from parameterized models.

4.2. Expanding Models across Space

As discussed above, process models can be important tools for constraining the impacts ecosystems can have in mitigating climate change. However, to serve this purpose effectively, it is imperative that accurate and defensible estimates of both actual and potential ecosystem state can be generated across space and into the future. This is a difficult task considering that there is still significant uncertainty in the estimates of current carbon stocks and fluxes at large scales (Van der Werf et al. 2009). Despite the difficulties, there has been much progress in estimating current ecosystem stocks and fluxes by combining remotely sensed data, forest inventory data, and TECMs (Turner et al. 2004, 2007, Houghton et al. 2007, Potter et al. 2007a, 2008, Saatchi et al. 2007, Baccini et al. 2008, Blackard et al. 2008, Goetz et al. 2009, Paivinen et al. 2009). The studies above show that using existing inventory paired with both remotely sensed data and TECMs, it is possible to generate estimates of current carbon stocks.

There are several approaches that these studies use to generate spatially explicit estimates of ecosystem stocks and fluxes. Some studies use an empirical approach that relates remotely sensed grid-cell level characteristics to the available estimates of stocks from scattered inventory estimates across the study region (Houghton et al. 2007, Muukkonen and Heiskanen 2007, Baccini et al. 2008, Paivinen et al. 2009, Powell et al. 2010). Other approachs use simple allometric models that relate the remotely sensed leaf area to other structural ecosystem components (e.g. bole biomass) (Zhang and Kondragunta 2006). Other approaches combine several remote sensing products and generate classes of cells. These strata are then related to the inventory data found within them (Saatchi et al. 2007, Blackard et al. 2008, Wulder et al. 2008). The most complex approaches use process models that have been calibrated using remote sensing products and/or inventory and flux data to estimate ecosystem stocks and fluxes (Nemani et al. 2003, Potter et al. 2007a, 2008, Turner et al. 2007). This last approach is most similar to the approach needed to estimate the potential productivity of a site as this potential can only be generated using process models.

The basic outline of the approach to estimate the potential rate and state of ecosystems across space would be similar to the approach outlined above. First, a suite of remote sensing products would be combined to create strata or land cover types across a region in space. At a grid-cell level, a process model would be calibrated using the remote sensing data paired with existing inventory data and flux data at that grid cell. The DA methods outlined above would be used to constrain the process model results given these data sources. This tuning process would automate the model calibration and allow each grid cell or stratum to have a unique model representing it. Once the process model was calibrated for cells with inventory and flux data, cells of the same stratum could be estimated using the process model and validated using other inventory and flux data not used in the model calibration. Finally, the process model would be run to a steady state using different emissions and climate scenarios to estimate future rate and state potentials. Using this run, the maximum flux rates (NPP, GPP, or NEP) could also be found over the full length of the model run. Each of these data points would then be expanded to cells of a similar stratum to generate a spatially explicit map of potential ecosystem rates and states.

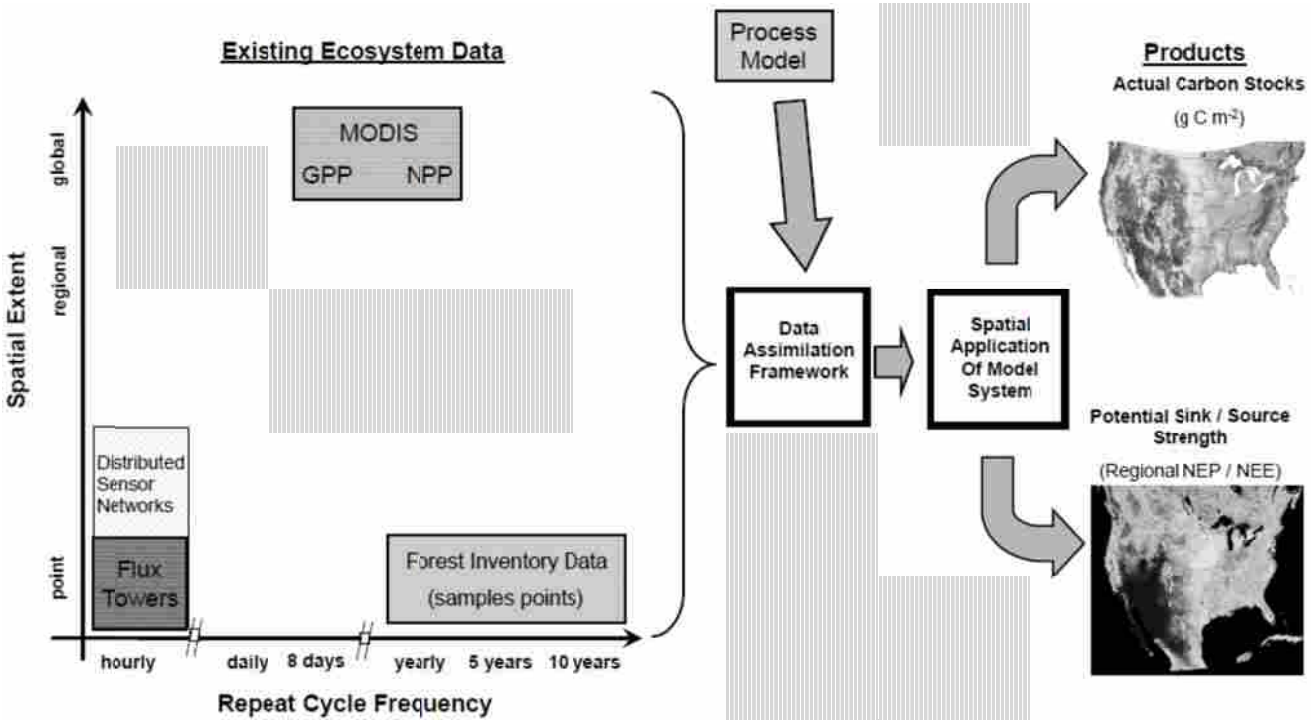

Figure 7: Schematic of available data, the scale of the data, and the results of a data assimilation approach

This automated modeling approach, while theoretically appealing, has several problems that must be addressed. First, collecting and organizing all of the data that may help to constrain the results is no small feat as there is no central clearing house for this sort of data. Second, the forcing data (e.g. climate data, soil data) may be sparse and have large uncertainties. Third, the TECM chosen will have a large impact on the final estimates (Cramer et al. 1999, Randerson et al. 2009). Given that these models produce representations of the true ecosystem structure, there may be large uncertainties in the final estimates generated from this approach. Lastly, running TECMs at a grid-cell level across large spatial extents presents major computational demands and may in fact make such an effort difficult. This last concern may be partially mitigated by using a stratification system as opposed to individual grid-cell level models.

4.3. Scale Flexibility

One benefit of using a DA approach (assuming its successful implementation) is the flexibility it provides in terms of the scale of the questions it allows to be addressed. Some of the most difficult aspects of large scale estimates of ecosystem stocks and fluxes are the myriad different data resolutions along with the sparse availability of actual measurements. As an example, MODIS reflectance data comes in 1km to 250m grid cell sizes for the entire globe. Annual NPP and 8 day GPP are calculated using a light use efficiency model at a 1km resolution (Running et al. 2004). Flux towers or continuous forest inventory data would be the ideal calibration and/or validation data sets; however there are less than 500 flux towers world wide (FLUXNET,("FLUXNET Integrating Worldwide CO_2 Flux Measurements" 2010)) and most forest inventory datasets are not remeasured frequently enough to provide accurate data about year to year changes. In addition to the MODIS datasets, Landsat data are available at 30m resolution but have much sparser temporal resolution. For Landsat data, similar issues of the availability of calibration and validation data apply and because of the smaller grid cell size, some larger scale flux tower footprints may exceed the 30m Landsat grid cell size making inferences difficult. In all of these cases, the flexibility of a DA approach and using a TECM allows for scaleable spatial products. Once the TECM is calibrated using the available observations, it can then be regridded and run at multiple scales should the need arise. Furthermore, sparse flux and inventory data can be integrated to better constrain the model results.

5. Conclusion

The need for spatially explicit estimates of forest biomass storage and CO_2 sequestration and emissions has never been greater. This is not a new field (Running et al. 1989, Tague and Band 2004), but impending and existing climate change mitigation and adaptation policies paired with a vibrant voluntary carbon market are driving the demand for high quality, credible, consistent, and accurate estimates of carbon stocks and fluxes in ecosystems around the globe. Fortunately, there are many sources of data that can help to constrain these estimates. Field based measurements of stocks and fluxes include traditional forest inventory, flux towers, and distributed sensor networks. Remote sensing technologies using both passive and active approaches like MODIS, Landsat, LIDAR, and RADAR can provide wall to wall spatial coverage over large areas to help estimate biomass accrual in areas with sparse ground data.

Regardless of the specific datasets available, using a data assimilation approach to combine the available data maximizes the accuracy of the final estimates of ecosystem sequestration and storage. The use of a model system to assimilate multiple data sources is a classic case of the sum of the data sets being greater than the parts. Although each of the data sets mentioned above are valuable, taken alone they are not as effective at answering the questions and addressing the needs of policy makers and carbon project developers. Using a DA approach to calibrate a TECM from the available data allows for more flexibility in applying the TECM across different spatial scales calibrated based on observations. A well calibrated TECM can then be used to estimate current stocks and fluxes as well as potential stocks and fluxes to further bracket the possible climate mitigation benefits associated with any given area.

Despite the flexibility and power of this approach, there is still a high level of discomfort with using process models (or models of any sort) to establish policy baselines or to constrain the outcomes of climate mitigation projects. Therefore, in the short term more work is needed to improve the accuracy and precision of process models and to thoroughly validate their results using trusted and well-understood data sources. Once this is done in many diverse ecosystems, the potential to apply calibrated TECMs to policy questions will be possible, and will allow ecosystems to play a greater role in climate change mitigation and adaptation policy.

References

Aalde, H., P. Gonzalez, M. Gytarsky, T. Krug, W. A. Kurz, S. Ogle, J. Raison, D. Schoene, N. H. Ravindranath, N. G. Elhassan, L. S. Heath, N. Higuchi, S. Kainja, M. Matsumoto, M. J. S. Sanchez, and Z. Somogyi. 2006. Chapter 4: Forest Land. *in* E. H.S., B. L., M. K., N.

T., and T. K., editors. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. National Greenhouse Gas Inventories Programme, Iges, Japan.
Van Aardt, J. A. N., R. H. Wynne, and R. G. Oderwald. 2006. Forest Volume and Biomass Estimation Using Small-Footprint Lidar-Distributional Parameters on a Per-Segment Basis. Forest Science 52:636–649.
Akay, A., H. Oğuz, I. Karas, and K. Aruga. 2009. Using LiDAR technology in forestry activities. Environmental Monitoring and Assessment 151:117–124.
Albani, M., D. Medvigy, G. C. Hurtt, and P. R. Moorcroft. 2006. The contributions of land-use change, CO2 fertilization, and climate variability to the Eastern US carbon sink. Global Change Biology 12:2370–2390.
Altman, D. G., and P. K. Andersen. 1989. Bootstrap investigation of the stability of a Cox regression model. Statistics in Medicine 8:771–783. Retrieved March 1, 2011, .
Ambrose, A. R., S. C. Sillett, and T. E. Dawson. 2009. Effects of tree height on branch hydraulics, leaf structure and gas exchange in California redwoods. Plant, Cell & Environment 32:743–757.
Arney, J. D., K. S. Milner, and B. L. Klienhenz. 2007. Biometrics of Forest Inventory, Forest Growth, and Forest Planning.
Avery, T. E., and H. E. Burkhart. 2002. Forest Measurements, 5th Edition. McGraw-Hill Higher Education, New York, NY.
Baccini, A., N. Laporte, S. J. Goetz, M. Sun, and H. Dong. 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. Environmental Research Letters 3:9pp.
Balzter, H., C. S. Rowland, and P. Saich. 2007. Forest canopy height and carbon estimation at Monks Wood National Nature Reserve, UK, using dual-wavelength SAR interferometry. Remote Sensing of Environment 108:224.
Bechtold, W. A., and P. L. Patterson. 2005. The Enhanced Forest Inventory and Analysis Program -- National Sampling Design and Estimation Procedures. Southern Research Station.
Bell, J. F., and J. R. Dilworth. 2007. Log Scaling and Timber Cruising2007 Edition. Cascade Printing Company, Corvallis, OR.
Berrill, J.-P., and K. L. O'Hara. 2003. Predicting Multi-Aged Coast Redwood Stand Growth and Yield Using Leaf Area Allocation. Page 42pp. California Department of Forestry and Fire Protection.
Biome-BGC - Conceptual diagram. 2005, June. . Retrieved December 16, 2010, from http://www.ntsg.umt.edu/models/bgc/index.php?option=com_content&task=view&id=23&Itemid=27.
Blackard, J. A., M. V. Finco, E. H. Helmer, G. R. Holden, M. L. Hoppus, D. M. Jacobs, A. J. Lister, G. G. Moisen, M. D. Nelson, R. Riemann, B. Ruefenacht, D. Salajanu, D. L. Weyermann, K. C. Winterberger, T. J. Brandeis, R. L. Czaplewski, R. E. McRoberts, P. L. Patterson, and R. P. Tymcio. 2008. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. Remote Sensing of Environment 112:1658–1677.
Borders, B. E., B. D. Shiver, and M. L. Clutter. 2005. Timber Inventory of Large Acreages Using Stratified Two-Stage List Sampling. Southern Journal of Applied Forestry 29:152–157.

Breidenbach, J., E. Næsset, V. Lien, T. Gobakken, and S. Solberg. 2010a. Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. Remote Sensing of Environment 114:911–924. doi: 10.1016/j.rse.2009.12.004.

Breidenbach, J., A. Nothdurft, and G. Kändler. 2010b. Comparison of nearest neighbour approaches for small area estimation of tree species-specific forest inventory attributes in central Europe using airborne laser scanner data. European Journal of Forest Research 129:833–846. doi: 10.1007/s10342-010-0384-1.

Burgess, S. S. O., and T. E. Dawson. 2004. The contribution of fog to the water relations of Sequoia sempervirens (D. Don): foliar uptake and prevention of dehydration. Plant, Cell & Environment 27:1023–1034.

Busing, R. T., and T. Fujimori. 2005. Biomass, production and woody detritus in an old coast redwood (Sequoia sempervirens) forest. Plant Ecology 177:177–188.

California Department of Forestry and Fire Protectiton. 2008. Important Information for Timber Operations Proposed with the Range of Northern Spotted Owl. Page 35. Department of Forestry and Fire Protection, Sacramento, CA.

Cap and Trade | California Air Resources Board. 2011, February. . Retrieved February 9, 2011, from http://www.arb.ca.gov/cc/capandtrade/capandtrade.htm.

CAR. 2010, August 20. Forest Project Protocol, Version 3.2. Climate Action Reserve. Retrieved from http://www.climateactionreserve.org/.

CCBA. 2008. Climate, Community and Biodiversity Project Design Standards Second Edition. Retrieved from http://www.climate-standards.org/.

Chapin III, F. S., G. M. Woodwell, J. T. Randerson, E. B. Rastetter, G. M. Lovett, D. D. Baldocchi, D. A. Clark, M. E. Harmon, D. S. Schimel, R. Valentini, C. Wirth, J. D. Aber, J. J. Cole, M. L. Goulden, A. D. McGuire, J. M. Melillo, H. A. Mooney, J. C. Neff, R. A. Houghton, M. L. Pace, M. G. Ryan, S. W. Running, O. E. Sala, W. H. Schlesinger, and E. D. Schulze. 2006. Reconciling Carbon-cycle Concepts, Terminology, and Methods. Ecosystems 9:1041–1–50.

Clark, M. L., D. A. Roberts, and D. B. Clark. 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sensing of Environment 96:375–398. doi: 10.1016/j.rse.2005.03.009.

Collins, J. N., L. B. Hutley, R. J. Williams, G. Boggs, D. Bell, and R. Bartolo. 2009. Estimating landscape-scale vegetation carbon stocks using airborne multi-frequency polarimetric synthetic aperture radar (SAR) in the savannahs of north Australia. International Journal of Remote Sensing 30:1141–1159.

Coops, N. C., M. A. Wulder, D. S. Culvenor, and B. St-Onge. 2004. Comparison of forest attributes extracted from fine spatial resolution multispectral and lidar data. Canadian Journal of Remote Sensing 30:855–866.

Cramer, W., D. W. Kicklighter, A. Bondeau, B. M. Iii, G. Churkina, B. Nemry, A. Ruimy, and A. L. Schloss. 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. Global Change Biology 5:1–15. doi: 10.1046/j.1365-2486.1999.00009.x.

Davis, K. J., P. S. Bakwin, C. Yi, B. W. Berger, C. Zhao, R. M. Teclaw, and J. G. Isebrands. 2003. The annual cycles of CO2 and H2O exchange over a northern mixed forest as observed from a very tall tower. Global Change Biology 9:1278–1293.

Dawson, T. E. 1998. Fog in the California redwood forest: ecosystem inputs and use by plants. Oecologia 117:476–485.

Demaeyer, P., M. De Dapper, and Gamanya. 2007. An automated satellite image classification design using object-oriented segmentation algorithms: A move towards standardization. Expert Systems with Applications 32:616–624. doi: 10.1016/j.eswa.2006.01.055.

Denman, K. L., G. Brasseur, A. Chidthaisong, P. Ciais, P. M. Cox, R. E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. M. Jacobs, U. Lohmann, S. Ramachandran, P. L. da Silva Dias, S. C. Wofsy, and X. Zhang. 2007. Couplings Between Changes in the Climate System and Biogeochemistry. Page 90 *in* S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, editors. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Derksen, S., and H. J. Keselman. 1992. Backward, forward and stepwise automated subset selection algorithms : frequency of obtaining authentic and noise variables. British Journal of Mathematical and Statistical Psychology 45:265–282. Retrieved March 1, 2011, .

Dixon, G. E. 2010. Essential FVS: A User's Guide to the Forest Vegetation Simulator. Page 240. User's Guide, United States Department of Agriculture, Forest Service, Fort Collins, CO.

Evensen, G. 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics 53:343–367.

Ewers, R. M., W. F. Laurence, and C. M. Souza. 2008. Temporal Fluctuations in Amazonian Deforestation Rates. Environmental Conservation 35:303–310. doi: 10.1017/S0376892908005122.

Ewing, H., K. Weathers, P. Templer, T. Dawson, M. Firestone, A. Elliott, and V. Boukili. 2009. Fog Water and Ecosystem Function: Heterogeneity in a California Redwood Forest. Ecosystems 12:417–433.

FLUXNET Integrating Worldwide CO2 Flux Measurements. 2010, April 4. . Retrieved from http://www.fluxnet.ornl.gov/fluxnet/index.cfm.

Fox, A., M. Williams, A. D. Richardson, D. Cameron, J. H. Gove, T. Quaife, D. Ricciuto, M. Reichstein, E. Tomelleri, C. M. Trudinger, and M. T. Van Wijk. 2009. The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. Agricultural and Forest Meteorology 149:1597–1615.

García, M., D. Riaño, E. Chuvieco, and F. M. Danson. 2010. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. Remote Sensing of Environment 114:816.

Ghioca-Robrecht, D. M., C. A. Johnston, and M. G. Tulbure. 2008. Assessing The Use Of Multiseason Quickbird Imagery For Mapping Invasive Species In A Lake Erie Coastal Marsh. Wetlands 28:1028–1039.

Gillis, M. D. 2001. Canada's National Forest Inventory (Responding to Current Information Needs). Environmental Monitoring and Assessment 67:121–129.

Glover, F. 1990. Tabu Search: A Tutorial. Interfaces 20:74–94. Retrieved December 21, 2010, .

Glover, F., E. Taillard, and D. de Werra. 1993. A user's guide to tabu search. Annals of Operations Research 41:1–28. doi: 10.1007/BF02078647.

Gobakken, T., and E. Naesset. 2008. Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. Canadian Journal of Forest Research 38:1095–1109.

Goetz, S., A. Baccini, N. Laporte, T. Johns, W. Walker, J. Kellndorfer, R. Houghton, and M. Sun. 2009. Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. Carbon Balance and Management 4:2.

GOFC-GOLD. 2009. Reducing greenhouse gas emissions from deforestation and degradation in developing countries: a sourcebook of methods and procedures for monitoring, measuring, and reporting. (F. Achard, S. Brown, R. DeFries, G. Grassi, M. Herold, D. Mollicone, D. Pandey, and C. Souza, Eds.). GOFC-GOLD Project Office, hosted by Natural Resources Canada, Alberta, Canada, Alberta, Canada.

Golinkoff, J. 2010. Biome BGC version 4.2: The Theoretical Framework. Page 71. Numerical Terradynamic Simulation Group, College of Forestry and Conservation, University of Montana, Missoula, MT. Retrieved from http://ntsg.umt.edu/sites/ntsg.umt.edu/files/project/biome-bgc/Golinkiff_BiomeBGCv4.2_TheoreticalBasis_1_18_10.pdf.

Gonzalez, P., G. P. Asner, J. J. Battles, M. A. Lefsky, K. M. Waring, and M. Palace. 2010. Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. Remote Sensing of Environment 114:1561–1575.

Gonzalez, R. C., R. E. Woods, and S. L. Eddins. 2009. Digital Image Processing Using MATLAB, 2nd edition. Prentice Hall, Upper Saddle River, NJ.

Goward, S. N., J. G. Masek, W. Cohen, G. Moisen, G. J. Collatz, S. Healey, R. A. Houghton, C. Huang, R. Kennedy, B. E. Law, S. Powell, D. Turner, and M. A. Wulder. 2008. Forest disturbance and North American carbon flux. EOS Transactions, AGU 89:105–106.

GPSMAP 76CSx Owner's Manual. (n.d.). . Owner's Manual, Garmin, 1200 East 151st St. Olathe, Kansas 66062.

Grace, J., C. Nichol, M. Disney, P. Lewis, T. Quaife, and P. Bowyer. 2007. Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence? Global Change Biology 13:1484–1497.

Hall, R. J., R. S. Skakun, E. J. Arsenault, and B. S. Case. 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. Forest Ecology and Management 225:378–390.

Harrell, F. E. 2001. Regression Modeling StrategiesCorrected. Springer.

Hart, J. K., and K. Martinez. 2006. Environmental Sensor Networks: A revolution in the earth system science? Earth-Science Reviews 78:177.

Hasenauer, H., K. Merganicova, R. Petritsch, S. A. Pietsch, and P. E. Thornton. 2003. Validating daily climate interpolations over complex terrain in Austria. Agricultural and Forest Meteorology 119:87–107.

Hawbaker, T. J., N. S. Keuler, A. A. Lesak, T. Gobakken, K. Contrucci, and V. C. Radeloff. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. Journal of Geophysical Research 114:11 PP. doi: 200910.1029/2008JG000870.

Heinsch, F. A., M. Zhao, S. W. Running, J. S. Kimball, R. R. Nemani, K. J. Davis, P. V. Bolstad, B. D. Cook, A. R. Desai, D. M. Ricciuto, B. E. Law, W. C. Oechel, H. Kwon, H. Luo, S. C. Wofsy, A. L. Dunn, J. W. Munger, D. D. Baldocchi, L. Xu, D. Y. Hollinger, A. D. Richardson, P. C. Stoy, M. B. S. Siqueira, R. K. Monson, S. P. Burns, and L. B.

Flanagan. 2006. Evaluation of Remote Sensing Based Terrestrial Productivity from MODIS Using Regional Eddy Flux Network Observations. IEEE Transactions on Geoscience and Remote Sensing 44:1908–1925.
Hilker, T., M. A. Wulder, and N. C. Coops. 2008. Update of forest inventory data with lidar and high spatial resolution satellite imagery. Canadian Journal of Remote Sensing 34:5–12. Retrieved March 3, 2011, .
Houghton, R. A., D. Butman, A. G. Bunn, O. N. Krankina, P. Schlesinger, and T. A. Stone. 2007. Mapping Russian forest biomass with data from satellites and forest inventories. Environmental Research Letters:7.
Huang, C., S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sensing of Environment 114:183–198.
Hudak, A., J. S. Evans, N. L. Crookstone, M. J. Falkowski, B. K. Steigers, R. Taylor, and H. Hemingway. 2008a. Aggregating Pixel-Level Basal Area Predictions Derived from LiDAR Data to Industrial Forest Stands in North-Central Idaho. Page 14. USDA Forest Service.
Hudak, A. T., N. L. Crookston, J. S. Evans, M. J. Falkowski, A. M. . Smith, P. E. Gessler, and P. Morgan. 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. Canadian Journal of Remote Sensing 32:126–138. Retrieved January 23, 2011, .
Hudak, A. T., N. L. Crookston, J. S. Evans, D. E. Hall, and M. J. Falkowski. 2008b. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment 112:2232–2245. doi: 10.1016/j.rse.2007.10.009.
Ioki, K., J. Imanishi, T. Sasaki, Y. Morimoto, and K. Kitada. 2009. Estimating stand volume in broad-leaved forest using discrete-return LiDAR: plot-based approach. Landscape and Ecological Engineering 6:29–36. doi: 10.1007/s11355-009-0077-4.
Jaskierniak, D., P. N. J. Lane, A. Robinson, and A. Lucieer. 2011. Extracting LiDAR indices to characterise multilayered forest structure using mixture distribution functions. Remote Sensing of Environment 115:573–585. doi: 10.1016/j.rse.2010.10.003.
Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2003. National-Scale Biomass Estimators for United States Tree Species. Forest Science 49:12–35.
Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2004. Comprehensive Database of Diameter-based Biomass Regressions for North American Tree Species. Northeastern Research Station.
Johansen, K., N. C. Coops, S. E. Gergel, and Y. Stange. 2007. Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. Remote Sensing of Environment 110:29–44. doi: 10.1016/j.rse.2007.02.014.
Johnson, R. A., and D. W. Wichern. 2002. Applied Multivariate Statistical Analysis, Fifth Edition. Prentice-Hall, Inc., Upper Saddle River, NJ.
Johnstone, J. A., and T. E. Dawson. 2010. Climatic context and ecological implications of summer fog decline in the coast redwood region. Proceedings of the National Academy of Sciences 107:4533–4538.
Jolliffe, I. T. 2002. Principle Component Analysis, 2nd Edition, 2nd edition. Springer-Verlag. Retrieved August 20, 2010, .
Kasischke, E. S., J. M. Melack, and C. M. Dobson. 1997. The use of imaging radars for ecological applications--A review. Remote Sensing of Environment 59:141–156.

King, J. E. 1966. Site Index Curves for Douglas-fir in the Pacific Northwest. Page 49. Weyerhaeuser Forestry Research Center, Centralia, Washington.
Kitahara, F., N. Mizoue, and S. Yoshida. 2009. Evaluation of data quality in Japanese National Forest Inventory. Environmental Monitoring and Assessment 159:331–340.
Knorr, W., and J. Kattge. 2005. Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. Global Change Biology 11:1333–1351.
Koch, G. W., S. C. Sillett, G. M. Jennings, and S. D. Davis. 2004. The limits to tree height. Nature 428:851–854.
Koutsias, N., M. Tsakiri-Strati, M. Karteris, and Mallinis. 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. ISPRS Journal of Photogrammetry and Remote Sensing 63:237–250. doi: 10.1016/j.isprsjprs.2007.08.007.
Lambers, H., F. S. I. Chapin, and T. L. Pons. 2008. Plant physiological ecology. Springer, New York, NY.
Landsat Missions. (n.d.). . Retrieved December 9, 2010, from http://landsat.usgs.gov/.
Landsberg, J. J., and R. H. Waring. 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. Forest Ecology and Management 95:209–228.
Landsberg, J. J., R. H. Waring, and N. C. Coops. 2003. Performance of the forest productivity model 3-PG applied to a wide range of forest types. Forest Ecology and Management 172:199–214.
Larcher, W. 2003. Physiological plant ecology: ecophysiology and stress physiology of functional groups. Springer, New York.
Latifi, H., A. Nothdurft, and B. Koch. 2010. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. Forestry 83:395 –407. doi: 10.1093/forestry/cpq022.
Law, B. E., T. Arkebauer, J. L. Campbell, J. M. Chen, O. Sun, M. Schwartz, C. van Ingen, and S. Verma. 2008. TCO: Terrestrial Carbon Observations: Protocols for Vegetation Sampling and Data Submissions. Page 92. Global Terrestrial Carbon Observing System (GTOS).
Lefsky, M. A., W. B. Cohen, G. G. Parker, and D. J. Harding. 2002. Lidar Remote Sensing for Ecosystem Studies. BioScience 52:19–30. Retrieved March 9, 2011, .
Lefsky, M. A., A. T. Hudak, W. B. Cohen, and S. A. Acker. 2005. Patterns of covariance between forest stand and canopy structure in the Pacific Northwest. Remote Sensing of Environment 95:517–531. doi: 10.1016/j.rse.2005.01.004.
Levy, P. S., and S. Lemeshow. 2008. Sampling of Populations: Methods and Applications, 4th edition. Wiley.
Lillesand, T. M., R. W. Kiefer, and J. W. Chipman. 2004. Remote Sensing and Image Interpretation. Fith Edition. John Wiley & Sons, Inc, New York, NY.
Lindberg, E., J. Holmgren, K. Olofsson, J. Wallerman, and H. Olsson. 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. International Journal of Remote Sensing 31:1175. doi: 10.1080/01431160903380649.
Lumley, T., and A. Miller. 2009. leaps: regression subset selection. Retrieved from http://CRAN.R-project.org/package=leaps.

Magnussen, S., and P. Boudewyn. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research 28:1016–1031. doi: 10.1139/cjfr-28-7-1016.

Maltamo, M., O. M. Bollandsås, E. Næsset, T. Gobakken, and P. Packalén. 2011. Different plot selection strategies for field training data in ALS-assisted forest inventory. Forestry 84:23 –31. doi: 10.1093/forestry/cpq039.

Maltamo, M., J. Peuhkurinen, J. Malinen, J. Vauhkonen, P. Packalén, and T. Tokola. 2009. Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. Silva Fennica 43:507–521.

Maselli, F., M. Chiesi, L. Fibbi, and M. Moriondo. 2008. Integration of remote sensing and ecosystem modelling techniques to estimate forest net carbon uptake. International Journal of Remote Sensing 29:2437–2443. doi: 10.1080/01431160801894857.

Mathieu, P.-P., and A. O'Neill. 2008. Data assimilation: From photon counts to Earth System forecasts. Remote Sensing of Environment 112:1258–1267.

MATLAB Image Processing Toolbox. 2011. . Mathworks, Natick, Massachusetts, U.S.A. Retrieved from http://www.mathworks.com/products/matlab/?s_cid=global_nav.

McRoberts, R. E., M. D. Nelson, and D. G. Wendt. 2002. Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique. Remote Sensing of Environment 82:457–468. doi: 10.1016/S0034-4257(02)00064-0.

Mildrexler, D. J., M. Zhao, and S. W. Running. 2009. Testing a MODIS Global Disturbance Index across North America. Remote Sensing of Environment 113:2103–2117.

MODIS Website. (n.d.). . Retrieved December 9, 2010, from http://modis.gsfc.nasa.gov/.

Moeur, M., and A. R. Stage. 1995. Most Similar Neighbor: An Improved Sampling Inference Procedure for Natural Resource Planning. Forest Science 41:337–359(23).

Munger, J. W., and H. W. Loescher. 2006. Guidelines for Making Eddy Covariance Flux Measurements. Ameriflux.

Muukkonen, P., and J. Heiskanen. 2007. Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: A possibility to verify carbon inventories. Remote Sensing of Environment 107:617–624.

Næsset, E. 1997a. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 52:49–56. doi: 10.1016/S0924-2716(97)83000-6.

Næsset, E. 1997b. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment 61:246–253. doi: 10.1016/S0034-4257(97)00041-2.

Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80:88–99. doi: 10.1016/S0034-4257(01)00290-5.

Nemani, R. R., C. D. Keeling, H. Hashimoto, W. M. Jolly, S. C. Piper, C. J. Tucker, R. B. Myneni, and S. W. Running. 2003. Climate-Driven Increases in Global Terrestrial Net Primary Production from 1982 to 1999. Science 300:1560–1563. doi: 10.1126/science.1082750.

Nothdurft, A., J. Saborowski, and J. Breidenbach. 2009. Spatial prediction of forest stand variables. European Journal of Forest Research 128:241–251. doi: 10.1007/s10342-009-0260-z.

Packalén, P., and M. Maltamo. 2006. Predicting the Plot Volume by Tree Species Using Airborne Laser Scanning and Aerial Photographs. Forest Science 52:611–622.

Paivinen, R., J. Van Brusselen, and A. Schuck. 2009. The growing stock of European forests using remote sensing and forest inventory data. Forestry 82:479–490. doi: 10.1093/forestry/cpp017.

Patenaude, G., R. A. Hill, R. Milne, D. L. A. Gaveau, B. B. J. Briggs, and T. P. Dawson. 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. Remote Sensing of Environment 93:368–380.

Van Pelt, R., and J. F. Franklin. 2000. Influence of canopy structure on the understory environment in tall, old-growth, conifer forests. Canadian Journal of Forest Research 30:1231–1245. doi: 10.1139/cjfr-30-8-1231.

Pesonen, A., A. Kangas, M. Maltamo, and P. Packalén. 2010. Effects of auxiliary data source and inventory unit size on the efficiency of sample-based coarse woody debris inventory. Forest Ecology and Management 259:1890–1899. doi: 10.1016/j.foreco.2010.02.001.

Pietsch, S. A., H. Hasenauer, and P. E. Thornton. 2005. BGC-model parameters for tree species growing in central European forests. Forest Ecology and Management 211:264–295.

Porté, A., and H. H. Bartelink. 2002. Modelling mixed forest growth: a review of models for forest management. Ecological Modelling 150:141–188.

Potter, C., P. Gross, S. Klooster, M. Fladeland, and V. Genovese. 2008. Storage of carbon in U.S. forests predicted from satellite data, ecosystem modeling, and inventory summaries. Climatic Change 90:269–282.

Potter, C., S. Klooster, S. Hiatt, M. Fladeland, V. Genovese, and P. Gross. 2007a. Satellite-derived estimates of potential carbon sequestration through afforestation of agricultural lands in the United States. Climatic Change 80:323–336.

Potter, C., S. Klooster, A. Huete, and V. Genovese. 2007b. Terrestrial carbon sinks for the United States predicted from MODIS satellite data and ecosystem modeling. Earth Interactions 11:1–21.

Potter, C., S. Klooster, R. Myneni, V. Genovese, P.-N. Tan, and V. Kumar. 2003. Continental-scale comparisons of terrestrial carbon sinks estimated from satellite data and ecosystem modeling 1982-1998. Global and Planetary Change 39:201–213.

Potter, C., J. T. Randerson, C. B. Field, P. A. Matson, P. M. Vitousek, H. A. Mooney, and S. A. Klooster. 1993. Terrestrial Ecosystem Production: A Process Model Based on Global Satellite and Surface Data. Global Biogeochemical Cycles 7:811–841.

Powell, S. L., W. B. Cohen, S. P. Healey, R. E. Kennedy, G. G. Moisen, K. B. Pierce, and J. L. Ohmann. 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. Remote Sensing of Environment 114:1053–1068.

De Pury, D. G. G., and G. D. Farquhar. 1997. Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. Plant, Cell and Environment 20:537–557.

R Development Core Team. 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org.

Ramsey, F. L., and D. W. Schafer. 2002. The Statistical Sleuth: A Course in Methods of Data Analysis, 2nd edition. Duxbury Thomson Learning, Pacific Grove, CA, USA.

Randerson, J. T., F. M. Hoffman, P. E. Thornton, N. M. Mahowald, K. Lindsay, Y.-H. Lee, C. D. Nevison, S. C. Doney, G. Bonan, R. StÖCkli, C. Covey, S. W. Running, and I. Y. Fung. 2009. Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. Global Change Biology 15:2462–2484.

Raupach, M. R., P. J. Rayner, D. J. Barrett, R. S. DeFries, M. Heimann, D. S. Ojima, S. Quegan, and C. C. Schmullius. 2005. Model data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Global Change Biology 11:378–397.

Reichle, R. H. 2008. Data assimilation methods in the Earth sciences. Advances in Water Resources 31:1411–1418.

Rundel, P., E. Graham, M. Allen, J. Fisher, and T. Harmon. 2009. Environmental sensor networks in ecological research. New Phytologist 182:589–607.

Running, S. W., D. D. Baldocchi, D. P. Turner, S. T. Gower, P. S. Bakwin, and K. A. Hibbard. 1999. A Global Terrestrial Monitoring Network Integrating Tower Fluxes, Flask Sampling, Ecosystem Modeling and EOS Satellite Data. Remote Sensing of Environment 70:108–127.

Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. C. Reeves, and H. Hashimoto. 2004. A continuous satellite-derived measure of global terrestrial primary production. BioScience 54:547–560.

Running, S. W., R. R. Nemani, D. L. Peterson, L. E. Band, D. F. Potts, L. L. Pierce, and M. A. Spanner. 1989. Mapping Regional Forest Evapotranspiration and Photosynthesis by Coupling Satellite Data with Ecosystem Simulation. Ecology 70:1090–1101. Retrieved December 15, 2010, .

Saatchi, S. S., R. A. Houghton, R. C. D. S. AlvalÁ, J. V. Soares, and Y. Yu. 2007. Distribution of aboveground live biomass in the Amazon basin. Global Change Biology 13:816–837.

Sãnchez-Azofeifa, G. A., K. L. Castro-Esau, W. A. Kurz, and A. Joyce. 2009. Monitoring carbon stocks in the tropics and the remote sensing operational limitations: from local to regional projects. Ecological Applications 19:480–494.

Sands, P. J., and J. J. Landsberg. 2002. Parameterisation of 3-PG for plantation grown Eucalyptus globulus. Forest Ecology and Management 163:273–292.

Shiver, B. D., and B. E. Borders. 1996. Sampling techniques for forest resource inventory. John Wiley & Sons, Inc., New York, NY.

Shoch, D. T., G. Kaster, A. Hohl, and R. Souter. 2009. Carbon Storage of Bottomland Hardwood Afforestation in the Lower Mississippi Valley, USA. Wetlands 29:535–542.

Simonin, K. A., L. S. Santiago, and T. E. Dawson. 2009. Fog interception by Sequoia sempervirens (D. Don) crowns decouples physiology from soil water deficit. Plant, Cell & Environment 32:882–892.

Smith, J. E., L. S. Heath, and J. C. Jenkins. 2003. Forest Volume-to-Biomass Models and Estimates of Mass for Live and Standing Dead Trees of U.S. Forests. Northeastern Research Station.

Smith, J. E., L. S. Heath, K. E. Skog, and R. A. Birdsey. 2006. Methods for Calculating Forest Ecosystem and Harvest Carbon with Standard Estimates for Forest Types of the United States. Northeastern Research Station.

Soil Data Mart - Home. (n.d.). . Retrieved December 10, 2010, from http://soildatamart.nrcs.usda.gov/.

Song, C., M. B. Dickinson, L. Su, S. Zhang, and D. Yaussey. 2010. Estimating average tree crown size using spatial information from Ikonos and QuickBird images: Across-sensor and across-site comparisons. Remote Sensing of Environment 114:1099–1107. doi: 10.1016/j.rse.2009.12.022.

Ståhl, G., S. Holm, T. G. Gregoire, T. Gobakken, E. Næsset, and R. Nelson. 2011. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. Canadian Journal of Forest Research 41:96–107. doi: 10.1139/X10-161.

Straub, C., H. Weinacker, and B. Koch. 2010. A comparison of different methods for forest resource estimation using information from airborne laser scanning and CIR orthophotos. European Journal of Forest Research 129:1069–1080. doi: 10.1007/s10342-010-0391-2.

Tague, C. L., and L. E. Band. 2004. RHESSys: Regional Hydro-Ecologic Simulation System—An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling. Earth Interactions 8:1–42. Retrieved December 15, 2010, .

Takahashi, T., Y. Awaya, Y. Hirata, N. Furuya, T. Sakai, and A. Sakai. 2010. Stand volume estimation by combining low laser-sampling density LiDAR - data with QuickBird panchromatic imagery in closed-canopy Japanese cedar (Cryptomeria japonica) plantations. International Journal of Remote Sensing 31:1281. doi: 10.1080/01431160903380623.

The Conservation Fund. 2006. Garcia River Forest Integrated Resource Management Plan. Page 289. Management Plan, The Conservation Fund, Caspar, CA. Retrieved May 4, 2011, from http://www.conservationfund.org/sites/default/files/The%20Conservation%20Fund%20Garcia%20River%20Forest%20Integrated%20Resource%20Management%20Plan.pdf.

The Landsat Program. (n.d.). . Retrieved December 9, 2010, from http://landsat.gsfc.nasa.gov/.

Thompson, S. K. 2002. Sampling, 2nd edition. John Wiley & Sons, Inc., New York, NY.

Thornton, P. E. 1998. Regional Ecosystem Simulation: Combining Surface- and Satellite-Based Observations to Study Linkages between Terrestrial Energy and Mass Budgets. The University of Montana, College of Forestry, Missoula, MT.

Thornton, P. E., H. Hasenauer, and M. A. White. 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. Agricultural and Forest Meteorology 104:255–271.

Thornton, P. E., B. E. Law, H. L. Gholz, K. L. Clark, E. Falge, D. S. Ellsworth, A. H. Goldstein, R. K. Monson, D. Hollinger, J. C. Paw U, and J. P. Sparks. 2002. Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. Agricultural and Forest Meteorology 113:185–222.

Thornton, P. E., and N. A. Rosenbloom. 2005. Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. Ecological Modelling 189:25–48.

Thornton, P. E., and S. W. Running. 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agricultural and Forest Meteorology 93:211–228.

Thornton, P. E., and S. W. Running. 2002. User's Guide for Biome-BGC, Version 4.1.2. The University of Montana.

Thornton, P. E., S. W. Running, and M. A. White. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. Journal of Hydrology 190:214–251.

Thornton, P. E., and N. E. Zimmermann. 2007. An Improved Canopy Integration Scheme for a Land Surface Model with Prognostic Canopy Structure. Journal of Climate 20:3902–3923.

Thum, T., T. Aalto, T. Laurila, M. Aurela, P. Kolari, and P. Hari. 2007. Parametrization of two photosynthesis models at the canopy scale in a northern boreal Scots pine forest. Tellus B 59:874–890.
Tibshirani, R. 1996. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58:267–288. Retrieved August 20, 2010, .
Turner, D. P., S. V. Ollinger, and J. S. Kimball. 2004. Integrating Remote Sensing and Ecosystem Process Models for Landscape- to Regional-Scale Analysis of the Carbon Cycle. BioScience 54:573–584.
Turner, D. P., W. D. Ritts, B. E. Law, W. B. Cohen, Z. Yang, T. Hudiburg, J. L. Campbell, and M. Duane. 2007. Scaling net ecosystem production and net biome production over a heterogeneous region in the western United States. Biogeosciences 4:597–612.
UNFCCC. 1998. Kyoto Protocol to the United Nations Framework Convention on Climate Change. United Nations.
UNFCCC. 2009. The Copenhagen Accord.
United States Forest Service Pacific Northwest Research Station. 2010. Volume estimation for the PNW-FIA Database. Page 69. United States Department of Agriculture, Forest Service, USFS Pacific Northwest Research Station.
Vanclay, J. K. 1994. Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests. CAB International, Wallingford, UK.
VCS. 2008. Verified Carbon Standard - Guidance for Agriculture, Forestry and Other Land Use Projects. Retrieved from http://www.v-c-s.org/.
Wallerman, J., and J. Holmgren. 2007. Estimating field-plot data of forest stands using airborne laser scanning and SPOT HRG data. Remote Sensing of Environment 110:501–508. doi: 10.1016/j.rse.2007.02.028.
Wang, Y. P., D. Baldocchi, R. A. Y. Leuning, E. V. A. Falge, and T. Vesala. 2007. Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight FLUXNET sites. Global Change Biology 13:652–670.
Wang, Y., H. Weinacker, and B. Koch. 2008. A Lidar Point Cloud Based Procedure for Vertical Canopy Structure Analysis And 3D Single Tree Modelling in Forest. Sensors 8:3938–3951. doi: 10.3390/s8063938.
Wang, Y.-P., C. M. Trudinger, and I. G. Enting. 2009. A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales. Agricultural and Forest Meteorology 149:1829–1842.
Waring, R. H., and S. W. Running. 2007. Forest Ecosystems: Analysis at Multiple Scales. Elsevier Academic Press, San Francisco, CA.
Weathers, K. C. 1999. The importance of cloud and fog in the maintenance of ecosystems. Trends in Ecology & Evolution 14:214–215.
Van der Werf, G. R., D. C. Morton, R. S. DeFries, J. G. J. Olivier, P. S. Kasibhatla, R. B. Jackson, G. J. Collatz, and J. T. Randerson. 2009. CO2 emissions from forest loss. Nature Geoscience 2:737–738.
White, M. A., P. E. Thornton, and S. W. Running. 1997. A Continental Phenology Model for Monitoring Vegetation Responses to Interannual Climatic Variability. Global Biogeochemical Cycles 11:217–234.
White, M. A., P. E. Thornton, S. W. Running, and R. R. Nemani. 2000. Parameterization and Sensitivity Analysis of the BIOME-BGC Terrestrial Ecosystem Model: Net Primary Production Controls. Earth Interactions 4:1–85.

Winrock International. 2010, November. The American Carbon Registry Forest Carbon Project Standard, Version 2.1.
Wulder, M. A., J. C. White, R. A. Fournier, J. E. Luther, and S. Magnussen. 2008. Spatially explicit large area biomass estimation: Three approaches using forest inventory and remotely sensed imagery in a GIS. Sensors 8:529–560.
Xiao, X., S. Boles, J. Liu, D. Zhuang, and M. Liu. 2002. Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. Remote Sensing of Environment 82:335–348. doi: 10.1016/S0034-4257(02)00051-2.
Yang, J., C. Zhang, X. Li, Y. Huang, S. Fu, and M. Acevedo. 2009. Integration of wireless sensor networks in environmental monitoring cyber infrastructure. Wireless Networks 16:1091–1108. doi: 10.1007/s11276-009-0190-1.
Zhang, X., and S. Kondragunta. 2006. Estimating forest biomass in the USA using generalized allometric models and MODIS land products. Geophysical Research Letters 33:1–5.
Zheng, D., L. S. Heath, and M. J. Ducey. 2007. Forest biomass estimated from MODIS and FIA data in the Lake States: MN, WI and MI, USA. Forestry 80:265–278.
Zhu, L., J. M. Chen, Q. Qin, J. Li, and L. Wang. 2009. Optimization of ecosystem model parameters using spatio-temporal soil moisture information. Ecological Modelling 220:2121–2136.
Zinke, P. J., and R. L. Crocker. 1962. The Influence of Giant Sequoia on Soil Properties. Forest Science 8:2–11.

Chapter 3

The Use of LiDAR and High-Resolution Imagery to Develop a Pixel-Based Stratification System to Estimate Carbon Stocks for a Verified Forest Carbon Offset Project

Abstract

The voluntary carbon market is a new and growing market that is increasingly important to consider in managing forestland. Monitoring, reporting, and verifying carbon stocks and fluxes at a project level is the single largest direct cost of a forest carbon offset project. There are now many methods for estimating forest stocks with high accuracy that use both Airborne Laser Scanning (ALS) and high-resolution optical remote sensing data. However, many of these methods are not appropriate for use under existing carbon offset standards and most have not been field tested. To bridge this implementation gap, a new forest stratification and sampling method that meets the requirements of the Climate Action Reserve (CAR) Forest Project Protocol has been designed and applied to a verified and registered carbon project in California. This approach meets the requirements of the CAR standard while reducing the costs of inventory and increasing the accuracy of estimates of carbon stocks and basal area. This method also applies a unique parametric and non-parametric application of ALS data to forest carbon estimation.

The 3 goals of this paper are to 1) present a novel method that has been successfully verified and registered at the project level and can be easily understood by land managers and verifiers, 2) present a method that can determine the optimum grid cell size to aggregate remote sensing data and that can be used to find the minimum sample size needed to meet given accuracy targets, and 3) explain how to leverage the inventory data collected in this way for future management, monitoring, and carbon verifications.

1. Introduction

The world's forests are a critical sink of carbon dioxide (Denman et al. 2007).It is estimated that forest degradation or destruction results in 6 to 17% of total anthropogenic CO_2 emissions annually (Van der Werf et al. 2009). Because of the importance of forest ecosystems in adapting to and mitigating climate change, there are now many policy initiatives to preserve and restore forest ecosystems for a climate benefit (UNFCCC 1998, 2009). Despite years of discussion however, policies to reduce emissions from terrestrial ecosystems have generally not been adopted. An exception to this is California's cap and

trade system that will incorporate carbon offsets starting in 2012 (barring a legal challenge) – see ("Cap and Trade | California Air Resources Board" 2011).

In part due to the dearth of climate change policies, a vibrant voluntary carbon offset market has sprung up centered around a suite of different carbon project standards (CCBA 2008, VCS 2008, CAR 2010, Winrock International 2010), and managing forests for carbon offsets can provide an important income stream for landowners willing to undertake the costs and requirements of these standards. These standards all have slightly different requirements regarding how to quantify the amount of carbon offsets generated, but generally all require periodic ground-based installation and measurement of plots to monitor project level carbon storage. This paper will focus on the requirements of the Climate Action Reserve Forest Project Protocol as this protocol is substantially similar to what will likely be adopted by the state of California for their compliance carbon market system. The ground based inventory described here, like most traditional forest monitoring, relies on tree measurement and conversion to volume, biomass, and carbon equivalents using established species-specific regressions developed through destructive sampling of trees (Jenkins et al. 2003, 2004, Smith et al. 2003, 2006). These sample-based estimates of forest carbon storage are then extrapolated across the full project, often through a stratification approach, whereby unsampled areas receive estimates from areas with similar characteristics based on their remotely sensed attributes (McRoberts et al. 2002).

This traditional approach to estimating forest parameters has recently been supplemented and improved upon with the use of remote sensing technologies like Light Detection and Ranging data (LiDAR) paired with high resolution multi-spectral imagery. While these new technologies can accurately estimate forest carbon stocks and fluxes, some of the methods

are not easily applicable to forest carbon offset projects because of their complexity and expense. There is a need to apply these new remote sensing products in the context of the voluntary carbon market to show their usefulness at a project level in conformance with typical forest carbon project standards.

2. Background

2.1. ALS and Optical Remote Sensing

Optical remote sensing products derived from airborne and satellite-borne sensors – Landsat Thematic Mapping Imagery (Hall et al. 2006, Demaeyer et al. 2007), IKONOS imagery (Song et al. 2010), Quickbird imagery (Johansen et al. 2007, Koutsias et al. 2008, Ghioca-Robrecht et al. 2008, Song et al. 2010), SPOT HRG imagery (Xiao et al. 2002), Moderate Resolution Imaging Spectroradiometer (MODIS) (Running et al. 2004, Grace et al. 2007, Houghton et al. 2007, Potter et al. 2007b, Zheng et al. 2007, Baccini et al. 2008, Blackard et al. 2008), and others (Goetz et al. 2009, Paivinen et al. 2009) – have all been used to classify forest landscapes and in some cases to estimate standing carbon stocks. However, estimates of carbon stocks and classifications created using optical sensors alone usually have trouble differentiating areas with high carbon stocks (Lefsky et al. 2002, 2005). Synthetic Aperture Radar (SAR) sensors can help improve estimates of biomass but these sensors also saturate in high biomass systems (Balzter et al. 2007). Because of these limitations, the estimation of forest carbon stocks is often greatly improved using forest structure data and specifically forest height. Airborne Laser Scanning (ALS) provides a richer summary of forest conditions and more accurate estimates of volume and biomass due to its ability to accurately capture forest heights (LiDAR intensity values can also be used to improve estimates).

ALS paired with other optical remote sensing data is a well-established approach to spatially estimating forest attributes (Wallerman and Holmgren 2007, Hilker et al. 2008, Ioki et al. 2009, Gonzalez et al. 2010, Takahashi et al. 2010, Breidenbach et al. 2010a, Straub et al. 2010). The use of optical remote sensing data in conjunction with LiDAR data is helpful in both delineating crown boundaries and in differentiating between species (Lefsky et al. 2005, Hilker et al. 2008, Gonzalez et al. 2010, Takahashi et al. 2010, Breidenbach et al. 2010a, Straub et al. 2010). The ability to make species level distinctions is especially important when estimating merchantable timber volumes and biomass, as these parameters differ between species in trees that are the same size.

ALS data is collected from an instrument that is flown over the forest on an airplane or helicopter. Laser pulses emitted from an airborne instrument reflect off of terrain and vegetation revealing both forest structure (e.g. – height, sub-canopy elements) and a detailed digital elevation model (Magnussen and Boudewyn 1998, Akay et al. 2009). Individual laser returns can be discrete or continuous (waveform). The spatial resolution can vary from many returns per square meter to sparser returns. The coverage of the ALS can range between full coverage of a given area with no gaps to a sample of the area based on transects below the flight lines to spot samples within transects (i.e. GLAS) (Ståhl et al. 2011, Maltamo et al. 2011).

There are two broad categories of ALS data analysis approaches: area-based approaches (ABA) / statistical canopy height distribution approaches, and individual tree crown approaches (ITC). Many individual tree approaches use the cloud of LiDAR point data and their relationship to neighborhood points to build individual crown polygons and/or 3-dimensional tree profiles (Coops et al. 2004, Wang et al. 2008, Akay et al. 2009). These

individual tree records can then be aggregated to any scale required to create stand level estimates. These ITC approaches use both parametric and non-parametric approaches (Maltamo et al. 2009).

In area-based approaches, plot level data is related to remote sensing data that has been aggregated to pixel, plot, or polygon (e.g. stand) units to estimate volume, biomass, or other area based metrics. Area based approaches fall broadly into two main categories:

1) The first category relates grid-cell or stand level remote sensing data to measured plot characteristics to build parametric models to represent forest data. These models have been shown to explain the vast majority of the variation in tree height, diameter at breast height, volume, biomass, basal area, and a suite of other variables (Næsset 1997a, 1997b, 2002, Magnussen and Boudewyn 1998, Hudak et al. 2006, Ioki et al. 2009, Gonzalez et al. 2010, Takahashi et al. 2010)

2) The second broad category uses non-parametric classification or nearest neighbor methods to stratify the forest into similar groups (Packalén and Maltamo 2006, Hudak et al. 2008a, 2008b, Nothdurft et al. 2009, Breidenbach et al. 2010b, Latifi et al. 2010, Jaskierniak et al. 2011). Non-parametric approaches include k-nearest neighbor techniques (Moeur and Stage 1995) and classification algorithms such as Random Forests (Hudak et al. 2008b).

Area-based approaches and individual tree approaches to estimating forest parameters are not mutually exclusive however, and several authors have shown how area-based systems can be combined with individual tree methods (Lindberg et al. 2010, Breidenbach et al. 2010a)

2.2. ALS and Optical Remote Sensing for a Forest Carbon Offset Project

The methods outlined above all provide different approaches to using ALS data and other data sources to estimate forest parameters. There are two main hurdles in using these methods for forest carbon offset projects. First, the method must be cost-effective and must also fit within the existing management framework of the project. Second, the estimation method must meet the monitoring and verification requirements of the carbon offset protocol. These protocols require periodic inventory of the forest and the application of species-level biomass and carbon conversion equations to all inventory estimates (Aalde et al. 2006, VCS 2008, CAR 2010). For example, the Climate Action Reserve Forest Project Protocol v3.2 requires that the United States Forest Service biomass conversions are used for all trees in the project area. Using a stratified inventory approach provides an easily understandable way to generate strata-level tree lists simply from plot data and because of this is more easily verified (CAR 2010). Although it may be possible to use some of the existing approaches within a forest carbon project framework, their complexity makes them difficult to understand and potentially challenging to verify. Some approaches do not generate species specific estimates of tree size that can then be used to expand to volume and/or biomass using approved biomass regressions (e.g. - (Ioki et al. 2009)). **The primary objective of this paper will be to describe how the ALS and optical remote sensing stratification system adequately meets the requirements of forest carbon protocols while improving the accuracy of forest inventory estimates.**

In addition to describing a method for ALS and optical remote sensing data to stratify a forest ownership to meet the requirements of a carbon project protocol, this paper will also detail how and where sampling should occur. ALS and optical remote sensing data provide a wealth of information that can be used to increase the efficiency of sampling a forest. **A**

secondary objective of this paper then, is to provide a method to choose the optimal size for the units of analysis (grid-cell size) and to locate plots across the project once the grid is established. Past research has used LiDAR data to stratify an area and locate field plots but these studies have not combined both LiDAR and optical data in the stratification and plot location. These studies have shown that using LiDAR data to first stratify an area and then to locate field plots based on initial strata reduced the root mean squared error (RMSE) of predicted volume (Hawbaker et al. 2009, Maltamo et al. 2011).

The question of the optimal grid-cell size has been addressed from the opposite direction by Gobakken and Næsset (2008). They examined the optimum plot size to use to best correlate the remote sensing data with the inventory data; however their analysis only used fixed area plot designs and did not examine at what scale to aggregate the remote sensing data (i.e. – how big should the grid cells be?). Van Aardt et al. (2006) examined various sizes of stands using variable radius plots but their analysis involved the best fit when a stand could contain multiple plots and did not use a regular grid system. Therefore, this new approach will show how to find the most appropriate grid cell size that relates variable radius prism plots to remotely sensed data where each grid cell receives no more than one plot.

Although there has been ample discussion of the technical nature of ALS-assisted forest estimation, few studies move beyond the initial analysis and results with an eye to future management and monitoring. **The third and final objective of this study is to examine how to best leverage data generated by this stratification and modeling exercise for typical management purposes and how to perform inventory updates assuming regular remote sensing data acquisition is not feasible (given cost constraints).**

Using an ALS and optical remote sensing stratification system, a verified and registered carbon project in Mendocino County, California, the Garcia River Forest (GRF), was inventoried in 2010 to meet the requirements of the California Climate Action Reserve (CAR) Forest Project Protocol. Three remotely sensed image datasets – color infrared data (CIR), Red, Green, and Blue true color imagery (RGB), and LiDAR data – were used to create a canopy segment layer, a canopy height model, and a digital elevation model. These data were summarized to 20m (1/10 acre) grid cells over the property. An initial systematic random sample was then installed over the full property. The remotely sensed variables were collapsed using a principal components analysis, and combined with the canopy segment summary variables and topographic descriptors, and field survey data to explain the variation in the initial sample of basal area (BA) using a regression model (models to predict trees per hectare (TPH) and percent conifer BA were also developed). The BA model was then used to estimate the basal area for each grid-cell on the property. The BA modeled estimates were then combined with average canopy height derived from the LiDAR canopy height model and the product of basal area and canopy height was calculated as a proxy of volume. This proxy was then divided into classes using an optimal binning heuristic, to define the strata. After this final stratification was completed, a second set of plots were installed to fully inventory each strata, with the number of plots based on the variability of each strata (see Figure 1 below).

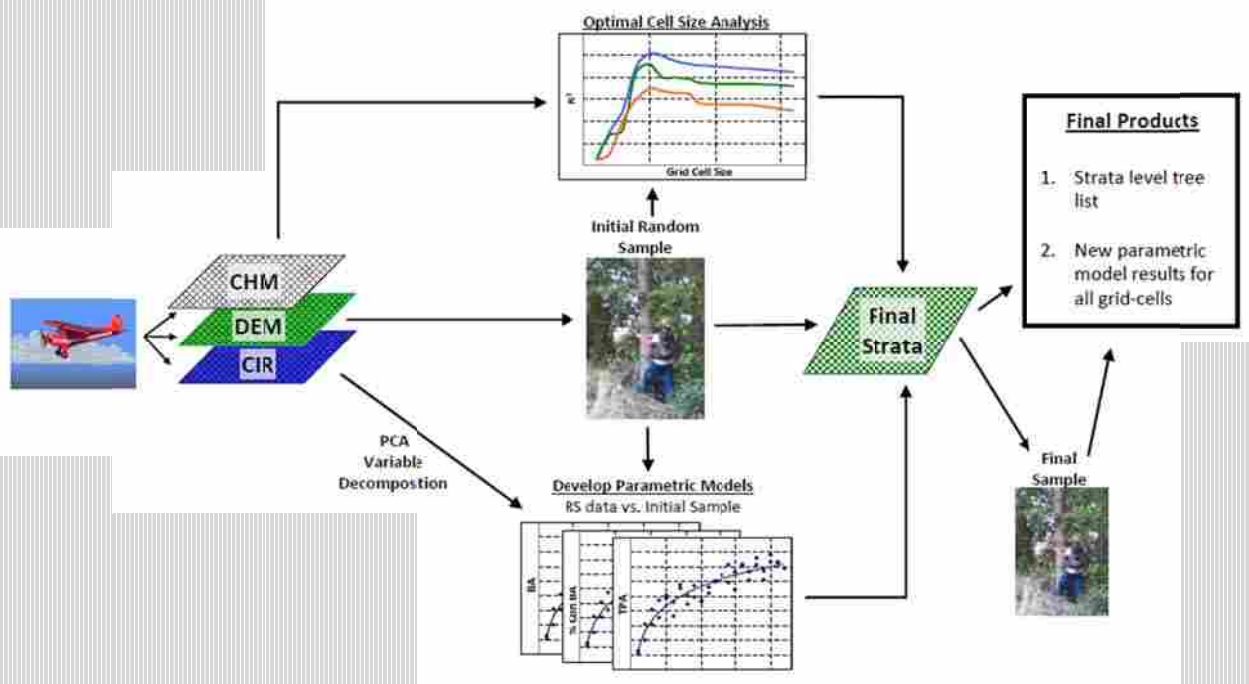

Figure 1: Outline of ALS and optical remote sensing data stratification method.

3. Results

3.1. *Traditional Stratification and Inventory and Approaches*

Traditional forest stand delineation and stratification (typing) are done by examining aerial photos of a forest and manually drawing boundaries around similar forest areas. This approach requires a forester to then place each stand into a stratum, based on their familiarity with actual forest conditions. This stratification may also use a visual check of ground data and may incorporate some plot data to inform how stands are assigned to strata (Shiver and Borders 1996, Avery and Burkhart 2002, Borders et al. 2005, Bell and Dilworth 2007).

This approach to stand-delineation and stratification is preferred to unstratified sampling designs, both because of its simplicity and its accuracy in estimating forest parameters. This approach is also preferred because knowing stand boundaries is useful for management purposes and harvest planning. The use of forest strata and stand delineation is ideal in forests with well-documented management histories and/or areas where even-age

management was used in the past. Stand boundaries are easily seen and delineated when they correspond to past management and management history can inform the typing of stands. However, in forests managed with uneven-aged silvicultural systems or without a well maintained history of past management, it can be difficult to create a stand map that accurately partitions the variability of a forest due to the relative homogeneity of the forest when observed from aerial photos. In this study, the field site fits within the second of these categories: the past management was well-documented but the uneven-aged harvests have left a forest that does not have many clear stand boundaries (see Figure 2), thus rendering the traditional stratification approach less accurate.

Using an ALS and optical remote sensing stratification (ORS) system, the 9,623 ha (23,780 acre) GRF property was divided into 36 strata (35 forested and 1 non-forested) across the property. Each stratum is at least 4.05 ha (10 acres) in size. Strata with higher numbers generally represent better stocked forest areas that have larger trees with more volume and carbon. This approach to forest stratification produces inventory estimates with more statistical confidence relative to the traditionally stand-based inventory approach using about half as many plots (see Table 1 and Table 4). Figure 2 shows a map of the strata generated by this new approach with the old stand boundaries shown in black. Except for the green areas that correspond with grassland, brush-fields, true oak woodlands, or stands treated to reduce tanoak competition most of the property has unclear stand boundaries in a traditional sense, with a high degree of variability within stands.

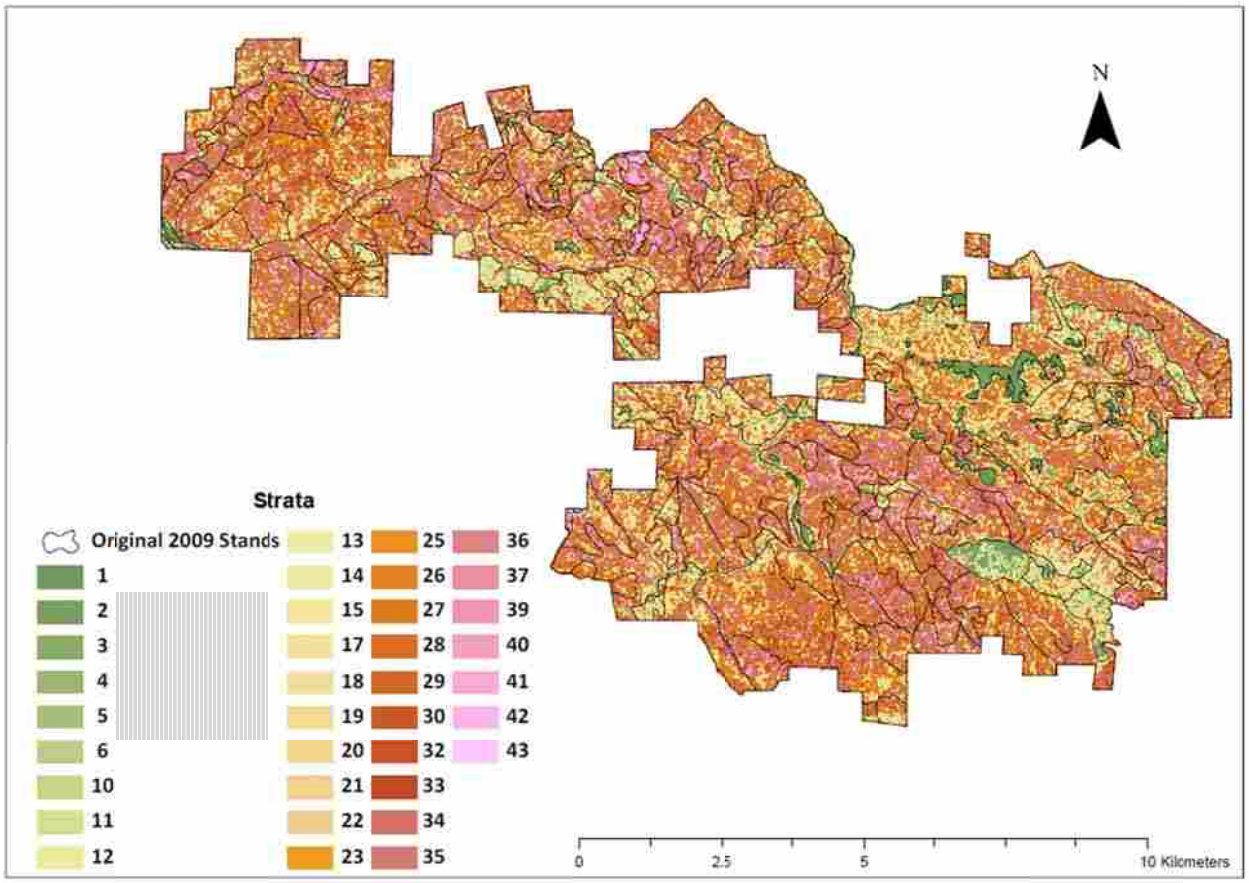

Figure 2: Overlay of 2009 Stand Layer with final stratification of the Garcia River Forest.

Table 1: Accuracy Results of Inventory – the original forest accuracy estimates are based on all plots grown forward to 2009 using the Forest Projection and Planning System growth and yield model calibrated to the Northern California redwood region. The 90% accuracy percentage is the property level standard error of the mean multiplied by the 90% t-value (1.645) divided by the mean value.

	Sample Type	Original Forest Inventory: (Multi-Stage Probability Proportional To Size Stand Based Stratification)	ALS and ORS Grid-Based Inventory: (post-stratification)
Accuracy Results	**C 90% Accuracy**	3.72%	3.42%
	BA 90% Accuracy	5.4%	3.60%
	BF 90% Accuracy	7.56%	5.30%

3.1.1. Regression Model Results from the Initial 199 Plots

The model form used to explain the variation in BA is shown below. Both the response and predictor variables have been transformed using a natural logarithm transformation.

$$Y = X\beta + \varepsilon$$

where Y is the transformed response, X is a matrix of transformed predictors identified by the Lasso method and β is the vector of least squares coefficients. The predictor variables used in these regressions are several topographic and LiDAR tree crown variables and the principle components of the color-infrared (CIR) and RGB imagery data sets as well as the PCA rotations for a suite of variables derived from the LiDAR data (the PCA rotations were used to reduce the number of parameters to analyze when building these regressions – see the Appendix for a full list of the predictor variables considered). The components of the β vector and the predictor variables (X) for the BA model are listed in Table 2. The variables are arranged such that those explaining most of the variation are listed first and those explaining the least are last. Regression relationships for trees per hectare and percent conifer BA are also shown below. These relationships were used when joining strata with less than 10 acres into other larger strata in the last step of the stratification process. A logistic model form was used for % Conifer BA as this model form results in outputs between zero and one.

As has been found in previous crown-based inventory projects, the LiDAR and CIR based variables predict the BA and TPH components best, while LiDAR and RGB variables are more helpful in predicting species composition (Lefsky et al. 2005, Gonzalez et al. 2010, Takahashi et al. 2010, Breidenbach et al. 2010a, Straub et al. 2010). The dominance of the color variables in predicting species composition is likely due to the realized species composition of the property being better represented by the image spatial domain than the image frequency domain. The spatial domain treats the image plane as a spatially related database and summarizes the pixel information in context to its neighbors. The frequency

domain works on the Fourier transformation of the pixel information. In this case texture, characterized by both grain size and arrangement was more important in discerning species composition than were the absolute color values (Clark et al. 2005). In other environments where the leaf color differs more profoundly, color has been more important than texture.

Table 2: Final Model Form and Coefficients (all coefficients are significantly different from 0 at the 95% confidence level)

BA		TPA		% Conifer BA	
Intercept	3.079788313	**Intercept**	6.19851	**Intercept**	-0.04949619
CIR3	-0.11917071	**Crown closure**	0.0006754	**LI1**	0.161971603
Average crown segment height	0.00519755	**LI6**	-0.19544	**RGB4**	0.81924046
Crown closure	0.017182801	**LI7**	0.05154	**LI2**	0.09321113
LI7	0.07755464	**LI4**	0.02984	**LI6**	-0.19769152
		CIR6	-0.11007	**RGB5**	-0.50907623
		LI2	-0.20571	**LI7**	0.294606256
		LI1	0.18478	**RGB1**	0.824221728
				RGB6	-0.42129326
				LI5	-0.50907623

Table 3: Initial Model Fit Statistics

Model	MSE	R^2	Correlation of Predicted vs Me6asured	Number of variables
BA	0.21687	0.635	0.647	4
TPA	1.46939	0.568	0.284	7
%ConBA	1.95837	NA	0.474	10

The coefficient of determination is not reported for the percent conifer BA as this statistic is not appropriate for logistic regression. Figure 3 shows the modeled versus measured BA in the original and final plots. An examination of the model fit with the original 199 plots (blue) showed that there weren't any strong trends in the residuals.

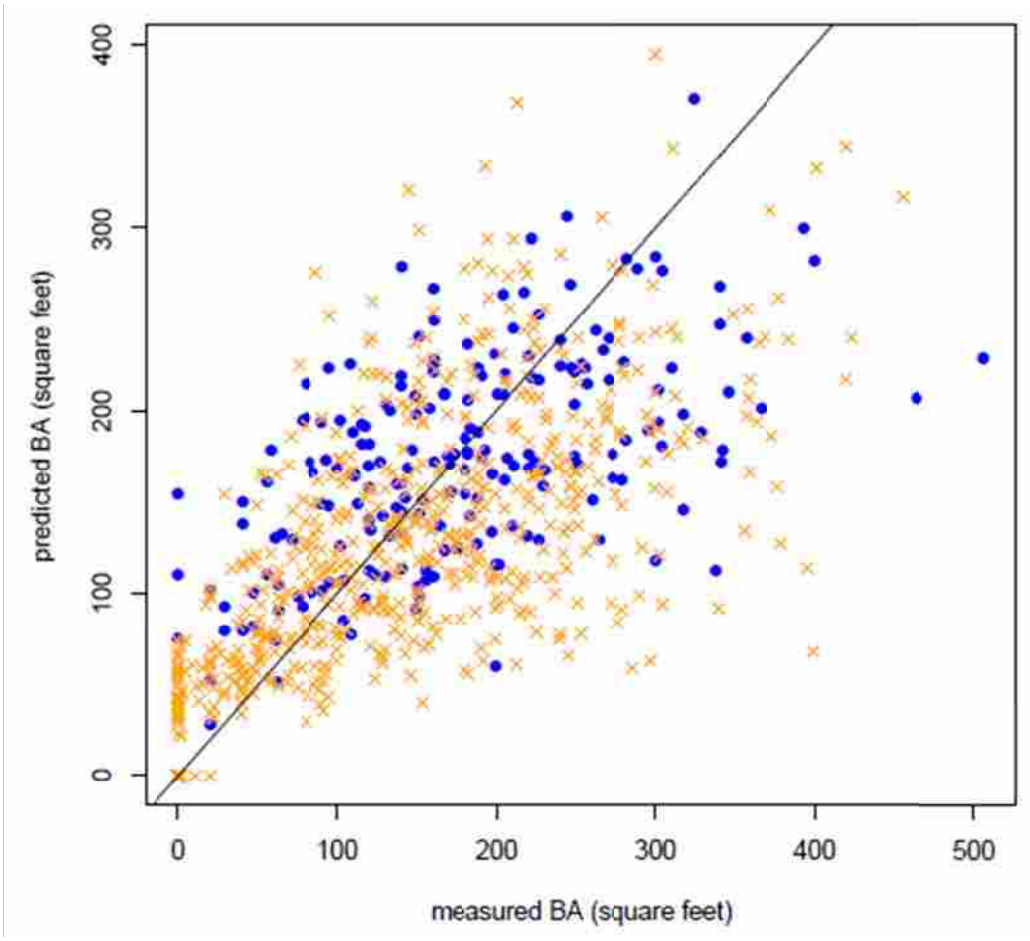

Figure 3 - Model results for BA (initial sample: blue dots, final sample: orange x's). The BA model residuals were not significantly different than a normal distribution (Pearson Chi-Square Normality Test, p-value = 0.7076)

3.1.2. Final Stratification Results

The final ALS-optical remote sensing stratification system resulted in more accurate property level estimates of live and dead carbon and basal area than the prior traditional stratification system (Table 1). An accurate stand delineation has the goal of maximizing between-stand variance while minimizing within-stand variance. To better understand the improvement this new approach to stratification provides, it is compared to the previous inventory that used a traditional stand-based stratification.

Based on the results seen using this new stratification approach there are several conclusions that can be drawn. First, with half as many plots (Table 4), we have more

statistical confidence in the inventory using this method due to the high resolution stratification derived from the remotely sensed imagery (Table 1). Second, this new stratification approach has shown that past samples most likely averaged more highly stocked riparian areas with non-riparian areas and therefore showed less volume on this property. Third, this new strata system allows for a flexible approach that can be easily leveraged when designing timber harvest plans or trying to understand the habitat of a given area. For example, accurate inventory estimates can now be made for any polygon across the full ownership simply by aggregating a set of grid cells.

Table 4: Summary and Comparison of 2009 and 2010 Stratification Systems. The 2010 "stands" are called stands as that is their closest analogue when thinking about a traditional stand-based stratified forest inventory. However, these "stands" do not correspond to management units and are therefore better thought of as pixels.

		2009	2010
Plot Data	**Total Plots**	1579	810
	Max Plots/Strata	394	40
	Min Plots/Strata	4	15
	Median Plots/Strata	45	22
	Average Plots/Strata	75	23
Stand Data	**Total Stands (Pixels)**	278	240,410
	Sampled Stands (Pixels)	170	810
	Max Stand (Pixel) Area (ha)	1,023	0.04
	Min Stand (Pixel) Area (ha)	0.8	0.04
	Median Stand (Pixel) Area (ha)	14	0.04
	Mean Stand (Pixel) Area (ha)	33	0.04
Strata Data	**Forested Strata #**	21	35
	Max Strata (ha)	1,704	1,816
	Min Strata (ha)	7.3	3.9
	Median Strata (ha)	230	76
	Average Strata (ha)	444	255

3.1.3. Old Stand Level Comparison

Visually, the strata systems are much different (see figure 2 and figure 4), as the old stand boundaries lump together many cells that are currently considered different strata. This visual comparison shows that although the old stratification and stand delineation does a reasonable job of capturing some of the differences in the stands, there are many areas where it is hard to see well defined stand boundaries.

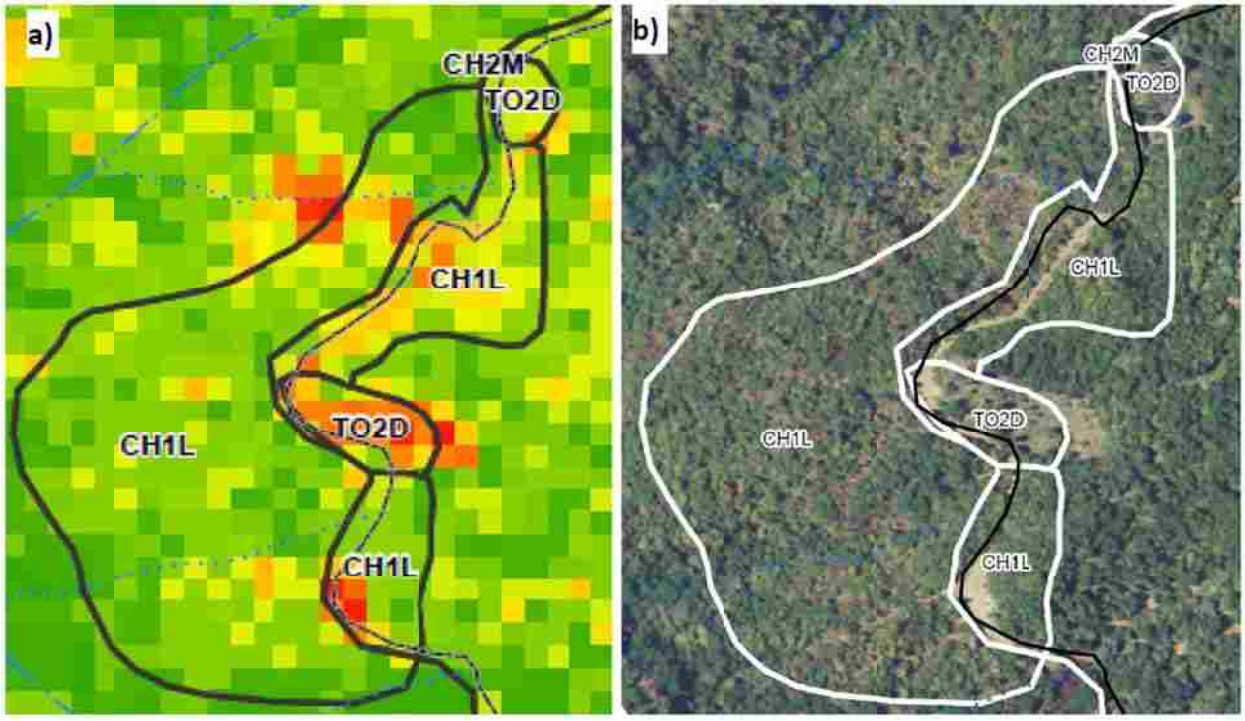

Figure 4: A visual comparison of the current stratification system versus the prior system. Figure 4a shows the current strataification system (with the old stand boundaries as well). The 0.04 ha grid cells are shaded to represent their different strata with redder cells having less volume than green cells. Figure 4b shows the prior stand delineation for this same area with the true-color imagery of the area as the base layer to show actual forest conditions. Note that the new strata grid-cells do correspond to the old stratification in areas where there are clear stand boundaries but in mixed forest conditions the new system can distinguish different forest conditions that the original strata system lumped together. This new strata system also does a much better job of mapping landings/clearings and wide road areas (most of the red and orange cells).

Table 5: Comparison of recently cruised stands using old strata system to current strata system. The estimates of stand level TPH, Board Feet (BF)/ha, and Metric Tons of Carbon (MgC)/ha showed no statistically significant difference between the past stand delineation estimate and aggregating the current stratification system to the old stand boundaries except for basal area (paired t-test p-values: BA = 0.034, TPH = 0.23, BF/ha = 0.81, metric tons Carbon/ha = 0.7).

2009 Data (2008 Plot Data Is Grown to 2009)									2010 Data				
Strata	Stand	Ha	Year Cruised	Plots	BA (m^2/ha)	TPH (>5cm)	BF per ha	C (Mg/ha - no dead)	# of 2010 Strata	BA (m^2/ha)	TPH (>5cm)	BF per ha	C (Mg/ha - no

													dead)
DR1M	**2**	53	2009	4	47.3	739.8	35,031	174.5	26	45.7	824.8	28,938	157.6
GX2D	**115**	7	2009	4	25.4	339.7	6,169	123.6	16	38.2	709.6	17,104	132.8
MH2D	**171**	35	2008	4	32.7	1,255.6	32,564	127.8	23	44.0	822.0	24,558	150.9
DR2D	**239**	13	2008	4	19.0	219.0	26,084	93.7	23	42.0	695.5	28,769	145.5
DR3D	**265**	54	2008	4	43.2	883.1	55,819	217.8	29	44.6	737.9	32,592	154.5
CH2M	**269**	183	2008	20	43.1	1,404.7	35,222	156.2	30	48.5	839.4	34,417	169.5
CH2M	**270**	138	2008	16	47.5	1,646.9	28,088	170.3	30	48.3	842.1	34,136	168.2
CH2M	**271**	131	2008	16	40.0	1,745.1	22,248	140.2	29	46.7	808.5	32,671	163.1

Another way to compare the current strata system to the prior system is to look at some well sampled stands in the prior inventory and compare those estimates to the current strata-based estimates (Table 5). Quantitatively the differences between mean estimates of stand variables are not statistically significant (except for BA – this result was also found in Hudak et al. (2008a) and they postulate that this bias is a result of the natural logarithm transformations and back transformations). These results therefore are an indication that the current stratification system, though much different than the previous system, produces estimates of stand level parameters that are similar to a traditional forest inventory (but more accurate). The advantage is that these estimates can now be found for any arbitrary polygon across the forest by grouping cells of interest and generating estimates for this group (Hudak et al. 2008a). This approach therefore presents a much more flexible set of data from which to gauge forest conditions.

4. Discussion

4.1. Selection of Grid Size

The first step in partitioning the variability of the GRF was to establish a grid across the whole property. Many LiDAR driven forest inventories in past studies have used stem-

mapped plots to correlate ground data with remote sensing data by using the actual location of trees and their crowns to build models that relate to the remotely sensed crown polygons and crown heights (Gonzalez et al. 2010). In this application however, variable radius plots were used to correlate the vegetation and the cell variability recognized by the LiDAR imagery. Stem mapping was not chosen because it would have been prohibitively expensive due to the high number of stems per ha and the steep terrain. However, because variable radius plots were used it is difficult to know the optimal size for grid-cells given that the size of the plots is variable (Avery and Burkhart 2002).

The exercise of choosing the size of the grid cells is dependent on several factors. The first consideration is the ability to accurately locate sample plots using handheld GPS units. The GPS units used by the inventory cruisers have accuracies that exceed 10m (33 feet) 95% of the time (*GPSMAP 76CSx Owner's Manual* n.d.). The second factor when choosing the grid size is finding the optimal cell size to reduce the variability between the remote sensing data and the measured plot data. Past studies have shown that it is important to choose a grid size that best matches the size of the plots installed (Magnussen and Boudewyn 1998, Næsset 2002). van Aardt et al. (2006) also explored this question using an object based approach (as opposed to pixels, objects are non-uniform areas of similar characteristics) and found only a small loss of accuracy with increasing object size. Pesonen et al. (2010) have also examined the optimal fixed area grid cell size but for that study focused on finding the optimum grid cell size when estimating coarse woody debris as opposed to standing trees.

Approaching the question of the optimum size to best relate plot data to remote sensing data, Gobokken and Næsset(2008) used a Monte Carlo analysis to explore the optimal size of fixed area plots in developing accurate forest inventory estimates. This analysis is similar to

our current question but may be difficult to implement in practice as the plots may already be measured or it may not be appropriate to change the plot design mid-sample.

In this case, a 4.6m^2/ha (20 ft^2/acre) basal area factor (BAF) prism was used on each plot. Generally, a 4.6 BAF prism samples about 0.04 ha but this will change depending on the size of the trees. To test this, the average of the limiting distances of each tree measured in all of the variable radius plots was calculated and the median plot size based on this analysis was determined to be 0.036 ha. However, larger trees would likely be outside of grid cells that are 0.4 ha or smaller. In addition, there is a greater chance that the location of the plot in the field would fall outside of the target grid cell due to the variability in the estimates of location made by the handheld GPS units. Therefore, grid cells less than 0.4 ha ($1/10^{th}$ acre) were deemed too small.

As the grid cell size increases to sizes larger than 0.4 ha, the variability of the forest within the cell (and hence the remote sensing data) increases. Because of this, it was hypothesized that any model that relates plot metrics to summarized grid cell remote sensing data will theoretically perform worse as the size of the cell increases to sizes larger than the plot. For these reasons, a 0.04 ha cell size was used as it was deemed to be the smallest cell size that would contain a 4.6 BAF plot and the location error associated with the handheld GPS units, and result in minimal within cell variability.

After further analysis following the completion of the inventory, the 0.04 ha grid cell size may have been slightly too small to create the strongest relationship between plot values (e.g. – BA, TPH, volume, carbon, etc) , topographical data (elevation, slope, aspect), and remotely sensed data (e.g. –orthophoto band intensity). The optimal grid cell analysis was undertaken after the inventory was completed as a means to assess if the pixel size used was the best size

and to inform future projects. The approach outlined below is one method that could be used to decide on the size of pixels to divide a forested area into and would ideally be used prior to the final sample. To determine the optimal grid cell size, a sample of the remotely sensed data was taken at each field plot point with a series of increasing circular areas (see figure 5a). The mean and standard deviation of all remotely sensed variables for each circular region for each data set was then calculated for each size circle. Once the remote sensing derived data had been summarized to each sample size, an exhaustive model selection routine was run to find the best model assuming the best model was defined using Bayes Information Criteria (BIC) (Lumley and Miller 2009, R Development Core Team 2011). The BIC was used as the metric of model performance because it does not assume that a relationship between explanatory and predictor values exists and has a larger penalty with larger data sets (Ramsey and Schafer 2002). Once the model with the lowest BIC was chosen for each circular area the amount of variation explained was graphed relative to each other sample size (figure 5b). In this way, an objective approach to model selection can provide a metric to judge which size grid-cell is optimal. Based on the results seen here, it seems the optimal cell size was about 0.08 ha (1/5 acre). This would be slightly larger than the cell size actually used.

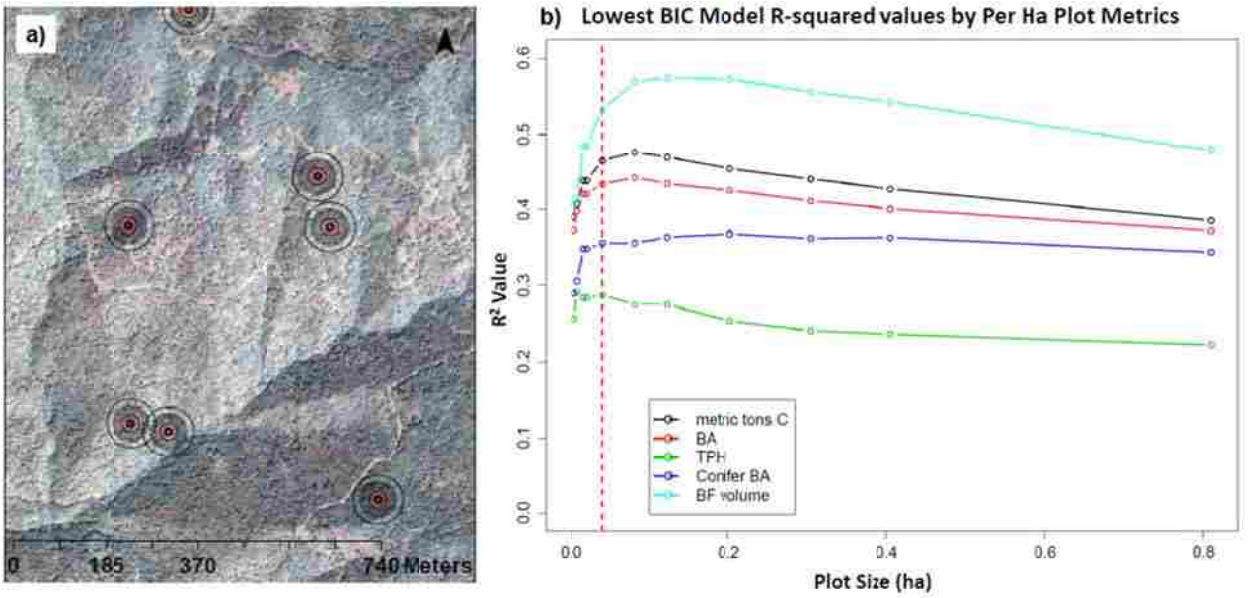

Figure 5: a) Remote sensing sample units of different size. Red circle represents 0.04 ha. b) Results of lowest BIC model selection approach using an exhaustive search of all potential model permutations. Dashed red line shows 0.04 ha size.

4.2. Sampling Intensity by Strata and Plot Location within Strata

The optimal sampling intensity of the final sample can be determined using a Neyman allocation of plots (or an optimal allocation of plots if the plots have variable costs in different strata) using the traditional approach to estimating the appropriate sample size (Shiver and Borders 1996). For strata that do not have an adequate initial sample to have confidence in the estimate of the sampling variability, an estimate of the variability of the strata can be found using the remote sensing data for those strata compared to the other strata. In this case, using the models developed from the initial strata to populate the cells of the under-sampled strata an estimate of the population variance can be found and used to calculate the optimum sample size. Plots are then randomly located within the strata.

4.3. Future Directions

4.3.1. Management Planning

Using this new approach will be a significant departure from how forest planning traditionally proceeds using a stand-based approach. Using a grid-based stratification, analysis of given forest areas in these small units can provide more fine-grained information about any given area. For example, when laying out timber harvest plan boundaries, these forest strata can be used to more accurately understand current stocking and forest conditions and allow for better layout of plan boundaries and a better description of pre-harvest conditions and habitat.

Although this stratification approach provides much higher resolution data in terms of understanding current forest conditions, there are several challenges to using this approach. To begin, this grid system does not lend itself to easy modeling of future management because the stand structure (20 m^2 pixels) does not yield logical management units. Secondly, although we have more confidence in the total volume of any given cell across the property, there may be more variation in the species composition within a stratum type. This is a result of the fact that total volume, not merchantable volume, was the variable whose variation was optimized during the creation of strata. In future efforts, both total volume and merchantable volume should be considered when creating strata boundaries.

4.3.2. Sampling of Harvest or Disturbance

As mentioned above, this strata system provides a highly flexible and accurate picture of current forest conditions. Moving forward, as areas are harvested or undergo natural disturbance, however, sampling will revert back to a more traditional harvest area (stand) based approach. The reason for this is twofold. First, the cost of collecting new remote sensing data annually prevents the collection of the necessary data to drive this

stratification process. Second, the known THP boundaries or disturbance events can be used to generate more accurate stand boundaries. Therefore, future sampling will proceed by first delineating the disturbed area and then sampling within this area to estimate the standing forest stocks post disturbance.

4.3.3. Ecological Monitoring

We anticipate that the canopy height model will be used in the future to generate a revised Northern Spotted Owl (NSO) habitat model to assist in management of the NSO. One of the benefits of this small grid system is that the final plot data can also be used to develop full parametric models for any variable of interest. In some cases (e.g. canopy cover), models are not required as the variable in question is measured directly by the LiDAR data. In this case, the canopy cover found in trees greater than 28cm (11in) DBH will be modeled to inform the classification of NSO habitat (California Department of Forestry and Fire Protectiton 2008) (traditionally this classification was based on lower resolution ocular estimates).

4.3.4. Pre-Aggregation for Process Modeling

Hawbaker et al. (2009) show that there is a need for ALS to be leveraged across larger landscapes and that ALS can help to create more accurate estimates of biophysical variables at a landscape scale by helping to better define the sampling design used. The method of sampling and stratification outlined in the following section can also be used to both validate process models and to serve as a pre-aggregation framework across a large landscape. Although this method uses ALS and optical remote sensing data with

continuous coverage across the landscape it could also be applied to larger scales using a variety of data sources with or without full coverage. Specifically, by running models based on a small set of strata instead of in each grid-cell across a region much more efficient and rapid estimates of ecosystem state can be generated.

Lefsky et al. (2005) have shown the value of using ALS combined with Landsat data to construct independent estimates of landscape net primary productivity and net ecosystem productivity to compare with light-use efficiency models or biogeochemistry models. Their work used remote sensing data collected over time to detect change. The strata system developed here will serve as the basis for future biogeochemistry model runs that will also attempt to better estimate ecosystem carbon fluxes at the GRF.

4.4. Conclusion

The method described below not only provides a cost effective and flexible approach to stratifying a forest but also has been designed and applied in the context of the requirements of existing forest carbon project protocols. This is highly valuable given that monitoring, reporting, and verifying carbon stocks and fluxes at a project level is the single largest external cost of a forest carbon offset project. Although currently the use of LiDAR approaches for smaller scales still is not cost effective, using a method like this one at scales larger than 10,000 ha (25,000 acres) may pay for themselves by reducing the cost of the field inventory required.

Additionally, the use of both parametric approaches (to develop models from the initial sample) and non-parametric approaches (to partition the variables of interest into strata) provide more power to determine the optimum sampling intensity and location across a large

ownership. Furthermore, the two-stage sample allows for the optimum grid cell size to be found.

For management decisions, this ALS and optical remote sensing stratification design and high-resolution grid allows for more accurate estimates of volume at any scale larger than a 0.04 ha grid cell (1/10 acre). This new strata layer and the data associated with it will serve as a baseline of forest conditions against which future management at the Garcia River Forest can be compared and assessed. Additionally, because of the flexibility built into this method, it can be scaled to much larger or smaller spatial extents. This is valuable for planning both local and larger scale ongoing management and monitoring activities.

5. Methods

5.1. Study Site

The Garcia River Forest (GRF) project is a 9,623 ha (23,780 acre) forest located in Mendocino County, California northwest of the town of Boonville. This forest is owned by The Conservation Fund (TCF) and is protected by a conservation easement held by the Nature Conservancy (TNC). The goals of the project are to conserve and restore highly productive and biologically diverse forests and streams, and to implement sustainable forest management practices that support the local economy (The Conservation Fund 2006). This region is historically dominated by a mix of redwood (*Sequoia sempivirens*)and Douglas-fir (*Pseudotsuga menziesii*) trees but due to decades of industrial timber management and intensive harvesting of this forest there is now a higher than natural amount of Tanoak (*Lithocarpus densiflorus*) in traditionally conifer dominated stands.

Due to the past management of the GRF, most stands have a mix of young 2nd or 3rd growth redwood and Douglas-fir trees with high proportions of tanoak. Most areas are

heterogeneous within stand boundaries and these conditions are the norm across the full ownership. Past management consisted mostly of "thinning from above" – removing the larger, better trees from most stands – and as a result most stands are made up of small, young trees.

Because of the state of the forest today, it is difficult to use a traditional stand mapping approach to delineate areas that are substantially similar. The result of applying the traditional air photo interpretation approach to stand mapping in this forest resulted in the creation of large stands that have high degrees of within stand variability and don't always relate to logical management units (see figure 2).

5.2. Field Data

5.2.1. 2009 Data (used for comparison to 2010 stratification results)

The existing inventory consisted of plots installed over several years using several different cruising protocols. Both variable radius plots and fixed area plots were installed across the property from 1999 to 2009. Most recently (2006 to 2008), all cruising occurred on a 400 by 400 meter (20 by 20 chain) grid that covered the full ownership using 4.6 Basal Area Factor prisms (Table 4). The complete inventory from 1999 to 2009 was grown forward to 2009 using the Forest Projection and Planning System growth and yield model to compare property level estimates in 2009 to the new stratification method in 2010. However, only plot data from 2008 and 2009 was used to compare individual stand level estimates to aggregated pixel estimates (see table 5).

The old stand layer was a traditional timber stand typing done by head's up digitizing stand boundaries using color imagery (acquired in 2004) of the forest. Each stand was then placed within a strata that described the dominant tree size and species based on the

professional judgment of the land manager. The old strata types had 3 fields: a 2 digit species code that described the dominant species or species mix, a 1 digit size-class code that described the dominant tree size, and a 1 digit canopy density code that described the degree of canopy closure.

5.2.2. 2010 Data (used for stratification)

The 2010 inventory data was collected between June and September of 2010. It consists of 810 variable radius plots that use a 4.6 m^2/ha (20 ft^2/acre)basal area factor (BAF) prism to measure trees at least 14 cm (5.5 inches) DBH. All plots have height measured on all trees (both live and dead) that are tallied in the variable radius plot. In addition to the trees measured in the prism plot, there is a 0.04 ha ($1/10^{th}$ acre) circular plot for understory vegetation, a 0.004 ha ($1/100^{th}$ acre) plot to measure regeneration (trees less than 14 cm DBH), and a 30.5 m (100 ft) transect to measure down woody debris. Table 4 summarizes the current inventory data and the past inventory data. The past 2009 inventory and stand layer was used as a baseline against which to compare the new 2010 ALS based stratification and inventory system.

The field sampled plots for the preliminary sample (199 plots) were a random selection of a 400m by 400m (20 by 20 chain) grid. The following table lists the summary statistics for this sample.

Table 6: Initial 199 Plot Summary Statistics

Variable	Min	Mean	Max
BA (m^2/ha)	0	40.73	116.1
TPH (Trees Per ha)	2	2,339	14,944
% conifer BA	0	56.6	100
Average height (m)	7	29	62

5.3. Remote Sensing Data

Both color-infrared imagery and LiDAR data were collected for the full property (Table 4). The color-infrared imagery has 0.6 meter (2 foot) resolution with horizontal accuracy less than 1 meter. The raw LiDAR returns range from 2.5 to 27 returns per square meter with at least 5 returns per square meter for forested areas. The LiDAR data exceeds 15cm of vertical accuracy and 50cm of horizontal accuracy. The LiDAR returns were summarized to make a 1 square meter digital elevation map and a 0.5 square meter canopy height model. The CHM is gridded to 0.5 m^2 and based on the interpolated "highest" return within each pixel. In addition to these grids, the LiDAR data were used to generated a crown polygon layer for the full GRF. The crown polygon layer was created using a watershed transformation algorithm applied to the CHM that segmented individual tree crowns that are isolated in height from adjacent regions.

Table 4: Summary of Remote Sensing Data Collected in 2009

	Color Infrared	**Light Detection And Ranging**
Acronym	**CIR**	**LiDAR**
Date Collected	7/1/2009	
Source	Fixed-wing aircraft	
Instrument	Digital Mapping Camera from Zeiss/Intergraph Imaging	ALTM Gemini from Optech Incorporated
Scale	Full ownership	
Projection	North American Datum 1983 UTM zone 10N	
Resolution	0.6 meter	5 returns / square meter, 24° field of view, 0.44 postings/square meter.
Spectrum	visible and near-infrared (380 nm to 2500nm)	near-infrared (760nm to 2500nm)
Accuracy	Horizontal accuracy sub 1 meter	Horizontal accuracy sub 50cm Vertical accuracy sub 15cm
Data Form	4 bands: red, blue, green, and near-infrared	Discrete Waveform with classified returns (ground, mid-canopy, upper-canopy)

Products	Ortho-rectified 4 band CIR	All and first return LiDAR (raw data) 1m^2 Digital Elevation Model (DEM) 0.5m^2 Canopy Height Model (CHM)Crown Polygon Layer

5.4. Technical Description of the Method

5.4.1. Data Summarization to 400 m^2 pixels

The first step before any analysis, inventory, or stratification could occur was to summarize all of the remote sensing data to the 400 m^2 grid cells. This involved finding the average and variance of all of the remote sensing data sets (e.g. CIR, RGB, , canopy height, crown polygons, topography variables – slope, aspect, elevation, and a whole suite of other variables derived from the remote sensing data in both the spatial and frequency domains). The complete set of variables used for the analysis and a brief description of them are listed in the appendix.

The source data for the cell summaries used in the stratification come from two passive image datasets and summarized LiDAR. The three image sets (CIR, RGB and CHM) were processed with MATLAB's image processing toolbox(*MATLAB Image Processing Toolbox* 2011). The image processing routines work in two domains; the spatial, and the frequency (Gonzalez et al. 2009). The pixels from the image data sets are about 0.6 meters on a side. The CHM is treated as a gray scale image where height above the ground is scaled to the gray scale.

5.4.2. Initial Plot Installation

To develop the final stratification, a set of "training" field plots were installed to find the relationships between plot data and the cell data (e.g. volume, carbon, basal area). To

do this, an initial set of 199 plots were installed across the GRF. A random sample of points located at the intersections of a 400 m by 400 m (20 by 20 chain) grid was chose to cover a broad spatial area.

5.4.3. Variable Reduction using Principle Component Analysis (PCA)

The 400 m^2cell data was summarized using principle component analysis to reduce the number of variables. Factor analysis was used to determine how many of the principle components should be retained (Jolliffe 2002). Table 8 lists the amount of variation explained by the first eight and the next eight principal components in the each of the image datasets.

Table 5: Principle Component Decomposition of the Imagery Datasets

Image set	Variance explained by first eight	Variance explained next eight
RGB	76.00%	13.40%
CIR	75.10%	13.70%
CHM	72.60%	14.70%

Based upon the reduction in explained variance and the need to keep the preliminary sample small, the first eight component vectors were selected to represent the data sets in the preliminary sample.

The original optical data consisted of 4 bands of data: blue, red, green, and NIR reflectance values. Although it would be possible to analyze this data by combining all 4 bands into one image, instead this optical data was used to create two images: a color-infrared (CIR) image and a Red-Green-Blue (RGB) image. The CIR image combines the red, green, and NIR values. There are two reasons why the red and green bands were included in both the CIR and RGB datasets: 1) to check that the atmospheric correction

was applied correctly and 2) to have finer control of the linear combination of the data when conducting the analysis.

Since two of the color bands (red and green) are present in both the CIR and RGB image data, a correlation analysis was conducted to determine the amount of overlap between the principal components of the two datasets. The following table charts the Pearson correlations with those having p-values less than 0.05 having an asterisk.

Table 6: Correlation Analysis between the CIR and RGB Principle Component Datasets

PrinComp	RGB1	RGB2	RGB3	RGB4	RGB5	RGB6	RGB7	RGB8
CIR1	0.925*	-0.128	-0.418*	0.137	-0.144*	-0.282*	-0.011	-0.08
CIR2	-0.072	0.977*	0.091	-0.067	0.064	-0.237*	-0.034	-0.048
CIR3	-0.09	-0.243*	0.158*	0.736*	-0.1	0.029	-0.252*	0.068
CIR4	-0.273*	0.068	0.891*	-0.185*	-0.189*	0.382*	-0.328*	-0.018
CIR5	-0.249*	0.101	0.371*	-0.254*	0.931*	0.123*	-0.11	0.0002
CIR6	-0.226*	-0.201*	-0.079	-0.237*	-0.206*	0.840*	-0.056	0.087
CIR7	0.052*	-0.019	-0.378*	-0.041	-0.009	-0.319*	0.890*	-0.084
CIR8	-0.170*	-0.008	0.066	-0.051	0.06	-0.177*	0.337*	0.880*

A quick scan of the table shows that as expected some of the principal components are highly correlated. This correlation reduces the efficiency of variable screening methods applied to this data, meaning that more plots will be required to achieve the same level of certainty. The impact of the correlations was examined by repeating the parameterization of the models described below with both data sets separately and then both together.

5.4.4. Parameterization of Models to Relate Remote Sensing Data to Initial Inventory

The data collected in the first 199 plots was then correlated to the reduced set of remotely sensed variables found using the PCA. Several models were built that related remotely sensed data to the measured plot data in each sampled 400 m^2cell. However,

only the BA regression model, multiplied by each cell's average canopy height, was used by the Tabu Search Algorithm to develop the initial strata. The BA model was then used to predict the BA in all of the 240,410 400 m^2 cells across the full ownership. The .5m^2 resolution Canopy Height Model (CHM) was averaged across each 20 by 20m pixel and used to estimate the average canopy height in each pixel (no model was required as this is directly measured by the LiDAR data).

Stepwise procedures have been found to produce poor variable screens (Harrell 2001). This is partially due to the repeated comparisons not representing the proper elimination probabilities (Derksen and Keselman 1992). However there are other problems with the method such as the parameter estimates being biased high, and the standard error of the estimates being too low. This results in F and chi-squared statistics not having the desired distributions (Altman and Andersen 1989). Based upon this the Lasso method (Tibshirani 1996) was used for the variable screening of the predictive models. The Lasso is a penalized least squares method which selects a set of regression coefficients (β^{Lasso}) as the coefficients that minimize the following equation:

$$\hat{\beta}^{Lasso} = argmin_{\beta} \left\{ \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{i=1}^{p} |\beta_i| \right\}$$

In the above equation, y is an n-length vector of the response variables, X is an n by p matrix of predictor variables. β_0 and β_j are the standard regression intercept and coefficient vectors while the last term is a penalty term applied to each coefficient – lambda is the penalty multiplier that is applied to each estimated coefficient.

To ensure that no single predictor swamps the effects of others, the matrix of predictors(X) is centered and scaled, and then λ is chosen by cross-validation. This

means that a portion of the plots are held back from the regression and these plots are then predicted by the resulting regression. The Lambda value is iteratively adjusted to produce the lowest prediction error of this cross-validation. The Lasso serves as a variable selection methodology by selecting few predictors thus alleviating problems attendant to having many potential predictors compared to the number of observations. Furthermore, since the Lasso tends to select only a few of a set of correlated predictors, it also helps reduce problems with spatial correlation (Tibshirani 1996).

5.4.5. Final Stratification Using Supervised Classification

Based on the predictions of the BA model described above, an optimal binning process (Glover 1990, Glover et al. 1993) was used to create bins (strata) for each cell based on the product of BA and height. The strata for each cell was determined by minimizing the amount of variation of the product of BA and height in each strata. The product of BA and height is highly correlated to volume and therefore cells within a given strata have similar volume totals. This classification method is considered supervised since it is driven by the initial inventory data collected across the GRF.

Once the supervised classification was completed, to prevent any strata from being less than 4.05 ha (10 acres) in size, an algorithm was applied to swap grid cells that were on the "edge" of each strata into neighboring strata (considering the nearness according to BA, height, Trees Per Hectare (TPH), and % conifer BA). The goal of this algorithm was to minimize the variation covered within a given strata while reducing the total number of strata.

5.4.6. Selection of Remaining 611 Sample Plots Based on Final Stratification

The final 611 plots were randomly placed within each final strata in proportion to the variability in product of BA and height. This sampling design is a classic post-stratification design and therefore uses stratified random sampling estimators (Shiver and Borders 1996, Thompson 2002).

6. Abbreviations

ALS – Airborne Laser Scanning
BA – Basal Area
BF – Board Foot Volume using the USFS board foot volume calculations (United States Forest Service Pacific Northwest Research Station 2010)
CHM – Canopy Height Model
CIR – Color Infrared optical data
GRF – Garcia River Forest
ha – hectare = 10,000 m^2
LiDAR – Light Detection and Ranging
PCA - Principle Component Analysis
RGB – Red, Green, and Blue optical data
THP - Timber Harvest Plan
TNC – The Nature Conservancy
TCF – The Conservation Fund
TPH – Trees Per Hectare

7. Acknowledgements

The Garica River Forest is a project of The Conservation Fund, in partnership with the Nature Conservancy, State Coastal Conservancy and the Wildlife Conservation Board. Funding for this work was provided by The Nature Conservancy's Priscilla Bullitt Collins Northwest Forest Fund, The Nature Conservancy, and The Conservation Fund. The authors would like to thank Evan Smith, Scott Kelly, Madison Thompson, and Holly Newberger for their logistical and technical support of this work as well as their review. We'd also like to thank Mark Reynolds, and Katie Andrews for their support of the data collection as well as Steve Running for his

review. Finally, three anonymous reviewers are thanked for their comments which notably improved the quality of the original manuscript.

8. References

Aalde, H., P. Gonzalez, M. Gytarsky, T. Krug, W. A. Kurz, S. Ogle, J. Raison, D. Schoene, N. H. Ravindranath, N. G. Elhassan, L. S. Heath, N. Higuchi, S. Kainja, M. Matsumoto, M. J. S. Sanchez, and Z. Somogyi. 2006. Chapter 4: Forest Land. *in* E. H.S., B. L., M. K., N. T., and T. K., editors. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. National Greenhouse Gas Inventories Programme, Iges, Japan.

Van Aardt, J. A. N., R. H. Wynne, and R. G. Oderwald. 2006. Forest Volume and Biomass Estimation Using Small-Footprint Lidar-Distributional Parameters on a Per-Segment Basis. Forest Science 52:636–649.

Akay, A., H. Oğuz, I. Karas, and K. Aruga. 2009. Using LiDAR technology in forestry activities. Environmental Monitoring and Assessment 151:117–124.

Albani, M., D. Medvigy, G. C. Hurtt, and P. R. Moorcroft. 2006. The contributions of land-use change, CO2 fertilization, and climate variability to the Eastern US carbon sink. Global Change Biology 12:2370–2390.

Altman, D. G., and P. K. Andersen. 1989. Bootstrap investigation of the stability of a Cox regression model. Statistics in Medicine 8:771–783. Retrieved March 1, 2011, .

Ambrose, A. R., S. C. Sillett, and T. E. Dawson. 2009. Effects of tree height on branch hydraulics, leaf structure and gas exchange in California redwoods. Plant, Cell & Environment 32:743–757.

Arney, J. D., K. S. Milner, and B. L. Klienhenz. 2007. Biometrics of Forest Inventory, Forest Growth, and Forest Planning.

Avery, T. E., and H. E. Burkhart. 2002. Forest Measurements, 5th Edition. McGraw-Hill Higher Education, New York, NY.

Baccini, A., N. Laporte, S. J. Goetz, M. Sun, and H. Dong. 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. Environmental Research Letters 3:9pp.

Balzter, H., C. S. Rowland, and P. Saich. 2007. Forest canopy height and carbon estimation at Monks Wood National Nature Reserve, UK, using dual-wavelength SAR interferometry. Remote Sensing of Environment 108:224.

Bechtold, W. A., and P. L. Patterson. 2005. The Enhanced Forest Inventory and Analysis Program -- National Sampling Design and Estimation Procedures. Southern Research Station.

Bell, J. F., and J. R. Dilworth. 2007. Log Scaling and Timber Cruising2007 Edition. Cascade Printing Company, Corvallis, OR.

Berrill, J.-P., and K. L. O'Hara. 2003. Predicting Multi-Aged Coast Redwood Stand Growth and Yield Using Leaf Area Allocation. Page 42pp. California Department of Forestry and Fire Protection.

Biome-BGC - Conceptual diagram. 2005, June. . Retrieved December 16, 2010, from http://www.ntsg.umt.edu/models/bgc/index.php?option=com_content&task=view&id=23&Itemid=27.

Blackard, J. A., M. V. Finco, E. H. Helmer, G. R. Holden, M. L. Hoppus, D. M. Jacobs, A. J. Lister, G. G. Moisen, M. D. Nelson, R. Riemann, B. Ruefenacht, D. Salajanu, D. L. Weyermann, K. C. Winterberger, T. J. Brandeis, R. L. Czaplewski, R. E. McRoberts, P. L. Patterson, and R. P. Tymcio. 2008. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. Remote Sensing of Environment 112:1658–1677.

Borders, B. E., B. D. Shiver, and M. L. Clutter. 2005. Timber Inventory of Large Acreages Using Stratified Two-Stage List Sampling. Southern Journal of Applied Forestry 29:152–157.

Breidenbach, J., E. Næsset, V. Lien, T. Gobakken, and S. Solberg. 2010a. Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. Remote Sensing of Environment 114:911–924. doi: 10.1016/j.rse.2009.12.004.

Breidenbach, J., A. Nothdurft, and G. Kändler. 2010b. Comparison of nearest neighbour approaches for small area estimation of tree species-specific forest inventory attributes in central Europe using airborne laser scanner data. European Journal of Forest Research 129:833–846. doi: 10.1007/s10342-010-0384-1.

Burgess, S. S. O., and T. E. Dawson. 2004. The contribution of fog to the water relations of Sequoia sempervirens (D. Don): foliar uptake and prevention of dehydration. Plant, Cell & Environment 27:1023–1034.

Busing, R. T., and T. Fujimori. 2005. Biomass, production and woody detritus in an old coast redwood (Sequoia sempervirens) forest. Plant Ecology 177:177–188.

California Department of Forestry and Fire Protectiton. 2008. Important Information for Timber Operations Proposed with the Range of Northern Spotted Owl. Page 35. Department of Forestry and Fire Protection, Sacramento, CA.

Cap and Trade | California Air Resources Board. 2011, February. . Retrieved February 9, 2011, from http://www.arb.ca.gov/cc/capandtrade/capandtrade.htm.

CAR. 2010, August 20. Forest Project Protocol, Version 3.2. Climate Action Reserve. Retrieved from http://www.climateactionreserve.org/.

CCBA. 2008. Climate, Community and Biodiversity Project Design Standards Second Edition. Retrieved from http://www.climate-standards.org/.

Chapin III, F. S., G. M. Woodwell, J. T. Randerson, E. B. Rastetter, G. M. Lovett, D. D. Baldocchi, D. A. Clark, M. E. Harmon, D. S. Schimel, R. Valentini, C. Wirth, J. D. Aber, J. J. Cole, M. L. Goulden, A. D. McGuire, J. M. Melillo, H. A. Mooney, J. C. Neff, R. A. Houghton, M. L. Pace, M. G. Ryan, S. W. Running, O. E. Sala, W. H. Schlesinger, and E. D. Schulze. 2006. Reconciling Carbon-cycle Concepts, Terminology, and Methods. Ecosystems 9:1041–1–50.

Clark, M. L., D. A. Roberts, and D. B. Clark. 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sensing of Environment 96:375–398. doi: 10.1016/j.rse.2005.03.009.

Collins, J. N., L. B. Hutley, R. J. Williams, G. Boggs, D. Bell, and R. Bartolo. 2009. Estimating landscape-scale vegetation carbon stocks using airborne multi-frequency polarimetric synthetic aperture radar (SAR) in the savannahs of north Australia. International Journal of Remote Sensing 30:1141–1159.

Coops, N. C., M. A. Wulder, D. S. Culvenor, and B. St-Onge. 2004. Comparison of forest attributes extracted from fine spatial resolution multispectral and lidar data. Canadian Journal of Remote Sensing 30:855–866.

Cramer, W., D. W. Kicklighter, A. Bondeau, B. M. Iii, G. Churkina, B. Nemry, A. Ruimy, and A. L. Schloss. 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. Global Change Biology 5:1–15. doi: 10.1046/j.1365-2486.1999.00009.x.

Davis, K. J., P. S. Bakwin, C. Yi, B. W. Berger, C. Zhao, R. M. Teclaw, and J. G. Isebrands. 2003. The annual cycles of CO2 and H2O exchange over a northern mixed forest as observed from a very tall tower. Global Change Biology 9:1278–1293.

Dawson, T. E. 1998. Fog in the California redwood forest: ecosystem inputs and use by plants. Oecologia 117:476–485.

Demaeyer, P., M. De Dapper, and Gamanya. 2007. An automated satellite image classification design using object-oriented segmentation algorithms: A move towards standardization. Expert Systems with Applications 32:616–624. doi: 10.1016/j.eswa.2006.01.055.

Denman, K. L., G. Brasseur, A. Chidthaisong, P. Ciais, P. M. Cox, R. E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. M. Jacobs, U. Lohmann, S. Ramachandran, P. L. da Silva Dias, S. C. Wofsy, and X. Zhang. 2007. Couplings Between Changes in the Climate System and Biogeochemistry. Page 90 *in* S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, editors. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Derksen, S., and H. J. Keselman. 1992. Backward, forward and stepwise automated subset selection algorithms : frequency of obtaining authentic and noise variables. British Journal of Mathematical and Statistical Psychology 45:265–282. Retrieved March 1, 2011, .

Dixon, G. E. 2010. Essential FVS: A User's Guide to the Forest Vegetation Simulator. Page 240. User's Guide, United States Department of Agriculture, Forest Service, Fort Collins, CO.

Evensen, G. 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics 53:343–367.

Ewers, R. M., W. F. Laurence, and C. M. Souza. 2008. Temporal Fluctuations in Amazonian Deforestation Rates. Environmental Conservation 35:303–310. doi: 10.1017/S0376892908005122.

Ewing, H., K. Weathers, P. Templer, T. Dawson, M. Firestone, A. Elliott, and V. Boukili. 2009. Fog Water and Ecosystem Function: Heterogeneity in a California Redwood Forest. Ecosystems 12:417–433.

FLUXNET Integrating Worldwide CO2 Flux Measurements. 2010, April 4. . Retrieved from http://www.fluxnet.ornl.gov/fluxnet/index.cfm.

Fox, A., M. Williams, A. D. Richardson, D. Cameron, J. H. Gove, T. Quaife, D. Ricciuto, M. Reichstein, E. Tomelleri, C. M. Trudinger, and M. T. Van Wijk. 2009. The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. Agricultural and Forest Meteorology 149:1597–1615.

García, M., D. Riaño, E. Chuvieco, and F. M. Danson. 2010. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. Remote Sensing of Environment 114:816.
Ghioca-Robrecht, D. M., C. A. Johnston, and M. G. Tulbure. 2008. Assessing The Use Of Multiseason Quickbird Imagery For Mapping Invasive Species In A Lake Erie Coastal Marsh. Wetlands 28:1028–1039.
Gillis, M. D. 2001. Canada's National Forest Inventory (Responding to Current Information Needs). Environmental Monitoring and Assessment 67:121–129.
Glover, F. 1990. Tabu Search: A Tutorial. Interfaces 20:74–94. Retrieved December 21, 2010, .
Glover, F., E. Taillard, and D. de Werra. 1993. A user's guide to tabu search. Annals of Operations Research 41:1–28. doi: 10.1007/BF02078647.
Gobakken, T., and E. Naesset. 2008. Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. Canadian Journal of Forest Research 38:1095–1109.
Goetz, S., A. Baccini, N. Laporte, T. Johns, W. Walker, J. Kellndorfer, R. Houghton, and M. Sun. 2009. Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. Carbon Balance and Management 4:2.
GOFC-GOLD. 2009. Reducing greenhouse gas emissions from deforestation and degradation in developing countries: a sourcebook of methods and procedures for monitoring, measuring, and reporting. (F. Achard, S. Brown, R. DeFries, G. Grassi, M. Herold, D. Mollicone, D. Pandey, and C. Souza, Eds.). GOFC-GOLD Project Office, hosted by Natural Resources Canada, Alberta, Canada, Alberta, Canada.
Golinkoff, J. 2010. Biome BGC version 4.2: The Theoretical Framework. Page 71. Numerical Terradynamic Simulation Group, College of Forestry and Conservation, University of Montana, Missoula, MT. Retrieved from http://ntsg.umt.edu/sites/ntsg.umt.edu/files/project/biome-bgc/Golinkiff_BiomeBGCv4.2_TheoreticalBasis_1_18_10.pdf.
Gonzalez, P., G. P. Asner, J. J. Battles, M. A. Lefsky, K. M. Waring, and M. Palace. 2010. Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. Remote Sensing of Environment 114:1561–1575.
Gonzalez, R. C., R. E. Woods, and S. L. Eddins. 2009. Digital Image Processing Using MATLAB, 2nd edition. Prentice Hall, Upper Saddle River, NJ.
Goward, S. N., J. G. Masek, W. Cohen, G. Moisen, G. J. Collatz, S. Healey, R. A. Houghton, C. Huang, R. Kennedy, B. E. Law, S. Powell, D. Turner, and M. A. Wulder. 2008. Forest disturbance and North American carbon flux. EOS Transactions, AGU 89:105–106.
GPSMAP 76CSx Owner's Manual. (n.d.). . Owner's Manual, Garmin, 1200 East 151st St. Olathe, Kansas 66062.
Grace, J., C. Nichol, M. Disney, P. Lewis, T. Quaife, and P. Bowyer. 2007. Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence? Global Change Biology 13:1484–1497.
Hall, R. J., R. S. Skakun, E. J. Arsenault, and B. S. Case. 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. Forest Ecology and Management 225:378–390.
Harrell, F. E. 2001. Regression Modeling StrategiesCorrected. Springer.
Hart, J. K., and K. Martinez. 2006. Environmental Sensor Networks: A revolution in the earth system science? Earth-Science Reviews 78:177.

Hasenauer, H., K. Merganicova, R. Petritsch, S. A. Pietsch, and P. E. Thornton. 2003. Validating daily climate interpolations over complex terrain in Austria. Agricultural and Forest Meteorology 119:87–107.

Hawbaker, T. J., N. S. Keuler, A. A. Lesak, T. Gobakken, K. Contrucci, and V. C. Radeloff. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. Journal of Geophysical Research 114:11 PP. doi: 200910.1029/2008JG000870.

Heinsch, F. A., M. Zhao, S. W. Running, J. S. Kimball, R. R. Nemani, K. J. Davis, P. V. Bolstad, B. D. Cook, A. R. Desai, D. M. Ricciuto, B. E. Law, W. C. Oechel, H. Kwon, H. Luo, S. C. Wofsy, A. L. Dunn, J. W. Munger, D. D. Baldocchi, L. Xu, D. Y. Hollinger, A. D. Richardson, P. C. Stoy, M. B. S. Siqueira, R. K. Monson, S. P. Burns, and L. B. Flanagan. 2006. Evaluation of Remote Sensing Based Terrestrial Productivity from MODIS Using Regional Eddy Flux Network Observations. IEEE Transactions on Geoscience and Remote Sensing 44:1908–1925.

Hilker, T., M. A. Wulder, and N. C. Coops. 2008. Update of forest inventory data with lidar and high spatial resolution satellite imagery. Canadian Journal of Remote Sensing 34:5–12. Retrieved March 3, 2011, .

Houghton, R. A., D. Butman, A. G. Bunn, O. N. Krankina, P. Schlesinger, and T. A. Stone. 2007. Mapping Russian forest biomass with data from satellites and forest inventories. Environmental Research Letters:7.

Huang, C., S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sensing of Environment 114:183–198.

Hudak, A., J. S. Evans, N. L. Crookstone, M. J. Falkowski, B. K. Steigers, R. Taylor, and H. Hemingway. 2008a. Aggregating Pixel-Level Basal Area Predictions Derived from LiDAR Data to Industrial Forest Stands in North-Central Idaho. Page 14. USDA Forest Service.

Hudak, A. T., N. L. Crookston, J. S. Evans, M. J. Falkowski, A. M. . Smith, P. E. Gessler, and P. Morgan. 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. Canadian Journal of Remote Sensing 32:126–138. Retrieved January 23, 2011, .

Hudak, A. T., N. L. Crookston, J. S. Evans, D. E. Hall, and M. J. Falkowski. 2008b. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment 112:2232–2245. doi: 10.1016/j.rse.2007.10.009.

Ioki, K., J. Imanishi, T. Sasaki, Y. Morimoto, and K. Kitada. 2009. Estimating stand volume in broad-leaved forest using discrete-return LiDAR: plot-based approach. Landscape and Ecological Engineering 6:29–36. doi: 10.1007/s11355-009-0077-4.

Jaskierniak, D., P. N. J. Lane, A. Robinson, and A. Lucieer. 2011. Extracting LiDAR indices to characterise multilayered forest structure using mixture distribution functions. Remote Sensing of Environment 115:573–585. doi: 10.1016/j.rse.2010.10.003.

Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2003. National-Scale Biomass Estimators for United States Tree Species. Forest Science 49:12–35.

Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2004. Comprehensive Database of Diameter-based Biomass Regressions for North American Tree Species. Northeastern Research Station.

Johansen, K., N. C. Coops, S. E. Gergel, and Y. Stange. 2007. Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. Remote Sensing of Environment 110:29–44. doi: 10.1016/j.rse.2007.02.014.
Johnson, R. A., and D. W. Wichern. 2002. Applied Multivariate Statistical Analysis, Fifth Edition. Prentice-Hall, Inc., Upper Saddle River, NJ.
Johnstone, J. A., and T. E. Dawson. 2010. Climatic context and ecological implications of summer fog decline in the coast redwood region. Proceedings of the National Academy of Sciences 107:4533–4538.
Jolliffe, I. T. 2002. Principle Component Analysis, 2nd Edition, 2nd edition. Springer-Verlag. Retrieved August 20, 2010, .
Kasischke, E. S., J. M. Melack, and C. M. Dobson. 1997. The use of imaging radars for ecological applications--A review. Remote Sensing of Environment 59:141–156.
King, J. E. 1966. Site Index Curves for Douglas-fir in the Pacific Northwest. Page 49. Weyerhaeuser Forestry Research Center, Centralia, Washington.
Kitahara, F., N. Mizoue, and S. Yoshida. 2009. Evaluation of data quality in Japanese National Forest Inventory. Environmental Monitoring and Assessment 159:331–340.
Knorr, W., and J. Kattge. 2005. Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. Global Change Biology 11:1333–1351.
Koch, G. W., S. C. Sillett, G. M. Jennings, and S. D. Davis. 2004. The limits to tree height. Nature 428:851–854.
Koutsias, N., M. Tsakiri-Strati, M. Karteris, and Mallinis. 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. ISPRS Journal of Photogrammetry and Remote Sensing 63:237–250. doi: 10.1016/j.isprsjprs.2007.08.007.
Lambers, H., F. S. I. Chapin, and T. L. Pons. 2008. Plant physiological ecology. Springer, New York, NY.
Landsat Missions. (n.d.). . Retrieved December 9, 2010, from http://landsat.usgs.gov/.
Landsberg, J. J., and R. H. Waring. 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. Forest Ecology and Management 95:209–228.
Landsberg, J. J., R. H. Waring, and N. C. Coops. 2003. Performance of the forest productivity model 3-PG applied to a wide range of forest types. Forest Ecology and Management 172:199–214.
Larcher, W. 2003. Physiological plant ecology: ecophysiology and stress physiology of functional groups. Springer, New York.
Latifi, H., A. Nothdurft, and B. Koch. 2010. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. Forestry 83:395 –407. doi: 10.1093/forestry/cpq022.
Law, B. E., T. Arkebauer, J. L. Campbell, J. M. Chen, O. Sun, M. Schwartz, C. van Ingen, and S. Verma. 2008. TCO: Terrestrial Carbon Observations: Protocols for Vegetation Sampling and Data Submissions. Page 92. Global Terrestrial Carbon Observing System (GTOS).
Lefsky, M. A., W. B. Cohen, G. G. Parker, and D. J. Harding. 2002. Lidar Remote Sensing for Ecosystem Studies. BioScience 52:19–30. Retrieved March 9, 2011, .

Lefsky, M. A., A. T. Hudak, W. B. Cohen, and S. A. Acker. 2005. Patterns of covariance between forest stand and canopy structure in the Pacific Northwest. Remote Sensing of Environment 95:517–531. doi: 10.1016/j.rse.2005.01.004.

Levy, P. S., and S. Lemeshow. 2008. Sampling of Populations: Methods and Applications, 4th edition. Wiley.

Lillesand, T. M., R. W. Kiefer, and J. W. Chipman. 2004. Remote Sensing and Image Interpretation. Fith Edition. John Wiley & Sons, Inc, New York, NY.

Lindberg, E., J. Holmgren, K. Olofsson, J. Wallerman, and H. Olsson. 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. International Journal of Remote Sensing 31:1175. doi: 10.1080/01431160903380649.

Lumley, T., and A. Miller. 2009. leaps: regression subset selection. Retrieved from http://CRAN.R-project.org/package=leaps.

Magnussen, S., and P. Boudewyn. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research 28:1016–1031. doi: 10.1139/cjfr-28-7-1016.

Maltamo, M., O. M. Bollandsås, E. Næsset, T. Gobakken, and P. Packalén. 2011. Different plot selection strategies for field training data in ALS-assisted forest inventory. Forestry 84:23 –31. doi: 10.1093/forestry/cpq039.

Maltamo, M., J. Peuhkurinen, J. Malinen, J. Vauhkonen, P. Packalén, and T. Tokola. 2009. Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. Silva Fennica 43:507–521.

Maselli, F., M. Chiesi, L. Fibbi, and M. Moriondo. 2008. Integration of remote sensing and ecosystem modelling techniques to estimate forest net carbon uptake. International Journal of Remote Sensing 29:2437–2443. doi: 10.1080/01431160801894857.

Mathieu, P.-P., and A. O'Neill. 2008. Data assimilation: From photon counts to Earth System forecasts. Remote Sensing of Environment 112:1258–1267.

MATLAB Image Processing Toolbox. 2011. . Mathworks, Natick, Massachusetts, U.S.A. Retrieved from http://www.mathworks.com/products/matlab/?s_cid=global_nav.

McRoberts, R. E., M. D. Nelson, and D. G. Wendt. 2002. Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique. Remote Sensing of Environment 82:457–468. doi: 10.1016/S0034-4257(02)00064-0.

Mildrexler, D. J., M. Zhao, and S. W. Running. 2009. Testing a MODIS Global Disturbance Index across North America. Remote Sensing of Environment 113:2103–2117.

MODIS Website. (n.d.). . Retrieved December 9, 2010, from http://modis.gsfc.nasa.gov/.

Moeur, M., and A. R. Stage. 1995. Most Similar Neighbor: An Improved Sampling Inference Procedure for Natural Resource Planning. Forest Science 41:337–359(23).

Munger, J. W., and H. W. Loescher. 2006. Guidelines for Making Eddy Covariance Flux Measurements. Ameriflux.

Muukkonen, P., and J. Heiskanen. 2007. Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: A possibility to verify carbon inventories. Remote Sensing of Environment 107:617–624.

Næsset, E. 1997a. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 52:49–56. doi: 10.1016/S0924-2716(97)83000-6.

Næsset, E. 1997b. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment 61:246–253. doi: 10.1016/S0034-4257(97)00041-2.

Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80:88–99. doi: 10.1016/S0034-4257(01)00290-5.

Nemani, R. R., C. D. Keeling, H. Hashimoto, W. M. Jolly, S. C. Piper, C. J. Tucker, R. B. Myneni, and S. W. Running. 2003. Climate-Driven Increases in Global Terrestrial Net Primary Production from 1982 to 1999. Science 300:1560–1563. doi: 10.1126/science.1082750.

Nothdurft, A., J. Saborowski, and J. Breidenbach. 2009. Spatial prediction of forest stand variables. European Journal of Forest Research 128:241–251. doi: 10.1007/s10342-009-0260-z.

Packalén, P., and M. Maltamo. 2006. Predicting the Plot Volume by Tree Species Using Airborne Laser Scanning and Aerial Photographs. Forest Science 52:611–622.

Paivinen, R., J. Van Brusselen, and A. Schuck. 2009. The growing stock of European forests using remote sensing and forest inventory data. Forestry 82:479–490. doi: 10.1093/forestry/cpp017.

Patenaude, G., R. A. Hill, R. Milne, D. L. A. Gaveau, B. B. J. Briggs, and T. P. Dawson. 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. Remote Sensing of Environment 93:368–380.

Van Pelt, R., and J. F. Franklin. 2000. Influence of canopy structure on the understory environment in tall, old-growth, conifer forests. Canadian Journal of Forest Research 30:1231–1245. doi: 10.1139/cjfr-30-8-1231.

Pesonen, A., A. Kangas, M. Maltamo, and P. Packalén. 2010. Effects of auxiliary data source and inventory unit size on the efficiency of sample-based coarse woody debris inventory. Forest Ecology and Management 259:1890–1899. doi: 10.1016/j.foreco.2010.02.001.

Pietsch, S. A., H. Hasenauer, and P. E. Thornton. 2005. BGC-model parameters for tree species growing in central European forests. Forest Ecology and Management 211:264–295.

Porté, A., and H. H. Bartelink. 2002. Modelling mixed forest growth: a review of models for forest management. Ecological Modelling 150:141–188.

Potter, C., P. Gross, S. Klooster, M. Fladeland, and V. Genovese. 2008. Storage of carbon in U.S. forests predicted from satellite data, ecosystem modeling, and inventory summaries. Climatic Change 90:269–282.

Potter, C., S. Klooster, S. Hiatt, M. Fladeland, V. Genovese, and P. Gross. 2007a. Satellite-derived estimates of potential carbon sequestration through afforestation of agricultural lands in the United States. Climatic Change 80:323–336.

Potter, C., S. Klooster, A. Huete, and V. Genovese. 2007b. Terrestrial carbon sinks for the United States predicted from MODIS satellite data and ecosystem modeling. Earth Interactions 11:1–21.

Potter, C., S. Klooster, R. Myneni, V. Genovese, P.-N. Tan, and V. Kumar. 2003. Continental-scale comparisons of terrestrial carbon sinks estimated from satellite data and ecosystem modeling 1982-1998. Global and Planetary Change 39:201–213.

Potter, C., J. T. Randerson, C. B. Field, P. A. Matson, P. M. Vitousek, H. A. Mooney, and S. A. Klooster. 1993. Terrestrial Ecosystem Production: A Process Model Based on Global Satellite and Surface Data. Global Biogeochemical Cycles 7:811–841.

Powell, S. L., W. B. Cohen, S. P. Healey, R. E. Kennedy, G. G. Moisen, K. B. Pierce, and J. L. Ohmann. 2010. Quantification of live aboveground forest biomass dynamics with

Landsat time-series and field inventory data: A comparison of empirical modeling approaches. Remote Sensing of Environment 114:1053–1068.
De Pury, D. G. G., and G. D. Farquhar. 1997. Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. Plant, Cell and Environment 20:537–557.
R Development Core Team. 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org.
Ramsey, F. L., and D. W. Schafer. 2002. The Statistical Sleuth: A Course in Methods of Data Analysis, 2nd edition. Duxbury Thomson Learning, Pacific Grove, CA, USA.
Randerson, J. T., F. M. Hoffman, P. E. Thornton, N. M. Mahowald, K. Lindsay, Y.-H. Lee, C. D. Nevison, S. C. Doney, G. Bonan, R. StÖCkli, C. Covey, S. W. Running, and I. Y. Fung. 2009. Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. Global Change Biology 15:2462–2484.
Raupach, M. R., P. J. Rayner, D. J. Barrett, R. S. DeFries, M. Heimann, D. S. Ojima, S. Quegan, and C. C. Schmullius. 2005. Model data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. Global Change Biology 11:378–397.
Reichle, R. H. 2008. Data assimilation methods in the Earth sciences. Advances in Water Resources 31:1411–1418.
Rundel, P., E. Graham, M. Allen, J. Fisher, and T. Harmon. 2009. Environmental sensor networks in ecological research. New Phytologist 182:589–607.
Running, S. W., D. D. Baldocchi, D. P. Turner, S. T. Gower, P. S. Bakwin, and K. A. Hibbard. 1999. A Global Terrestrial Monitoring Network Integrating Tower Fluxes, Flask Sampling, Ecosystem Modeling and EOS Satellite Data. Remote Sensing of Environment 70:108–127.
Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. C. Reeves, and H. Hashimoto. 2004. A continuous satellite-derived measure of global terrestrial primary production. BioScience 54:547–560.
Running, S. W., R. R. Nemani, D. L. Peterson, L. E. Band, D. F. Potts, L. L. Pierce, and M. A. Spanner. 1989. Mapping Regional Forest Evapotranspiration and Photosynthesis by Coupling Satellite Data with Ecosystem Simulation. Ecology 70:1090–1101. Retrieved December 15, 2010, .
Saatchi, S. S., R. A. Houghton, R. C. D. S. AlvalÁ, J. V. Soares, and Y. Yu. 2007. Distribution of aboveground live biomass in the Amazon basin. Global Change Biology 13:816–837.
Sãnchez-Azofeifa, G. A., K. L. Castro-Esau, W. A. Kurz, and A. Joyce. 2009. Monitoring carbon stocks in the tropics and the remote sensing operational limitations: from local to regional projects. Ecological Applications 19:480–494.
Sands, P. J., and J. J. Landsberg. 2002. Parameterisation of 3-PG for plantation grown Eucalyptus globulus. Forest Ecology and Management 163:273–292.
Shiver, B. D., and B. E. Borders. 1996. Sampling techniques for forest resource inventory. John Wiley & Sons, Inc., New York, NY.
Shoch, D. T., G. Kaster, A. Hohl, and R. Souter. 2009. Carbon Storage of Bottomland Hardwood Afforestation in the Lower Mississippi Valley, USA. Wetlands 29:535–542.
Simonin, K. A., L. S. Santiago, and T. E. Dawson. 2009. Fog interception by Sequoia sempervirens (D. Don) crowns decouples physiology from soil water deficit. Plant, Cell & Environment 32:882–892.

Smith, J. E., L. S. Heath, and J. C. Jenkins. 2003. Forest Volume-to-Biomass Models and Estimates of Mass for Live and Standing Dead Trees of U.S. Forests. Northeastern Research Station.

Smith, J. E., L. S. Heath, K. E. Skog, and R. A. Birdsey. 2006. Methods for Calculating Forest Ecosystem and Harvest Carbon with Standard Estimates for Forest Types of the United States. Northeastern Research Station.

Soil Data Mart - Home. (n.d.). . Retrieved December 10, 2010, from http://soildatamart.nrcs.usda.gov/.

Song, C., M. B. Dickinson, L. Su, S. Zhang, and D. Yaussey. 2010. Estimating average tree crown size using spatial information from Ikonos and QuickBird images: Across-sensor and across-site comparisons. Remote Sensing of Environment 114:1099–1107. doi: 10.1016/j.rse.2009.12.022.

Ståhl, G., S. Holm, T. G. Gregoire, T. Gobakken, E. Næsset, and R. Nelson. 2011. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. Canadian Journal of Forest Research 41:96–107. doi: 10.1139/X10-161.

Straub, C., H. Weinacker, and B. Koch. 2010. A comparison of different methods for forest resource estimation using information from airborne laser scanning and CIR orthophotos. European Journal of Forest Research 129:1069–1080. doi: 10.1007/s10342-010-0391-2.

Tague, C. L., and L. E. Band. 2004. RHESSys: Regional Hydro-Ecologic Simulation System—An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling. Earth Interactions 8:1–42. Retrieved December 15, 2010, .

Takahashi, T., Y. Awaya, Y. Hirata, N. Furuya, T. Sakai, and A. Sakai. 2010. Stand volume estimation by combining low laser-sampling density LiDAR - data with QuickBird panchromatic imagery in closed-canopy Japanese cedar (Cryptomeria japonica) plantations. International Journal of Remote Sensing 31:1281. doi: 10.1080/01431160903380623.

The Conservation Fund. 2006. Garcia River Forest Integrated Resource Management Plan. Page 289. Management Plan, The Conservation Fund, Caspar, CA. Retrieved May 4, 2011, from http://www.conservationfund.org/sites/default/files/The%20Conservation%20Fund%20Garcia%20River%20Forest%20Integrated%20Resource%20Management%20Plan.pdf.

The Landsat Program. (n.d.). . Retrieved December 9, 2010, from http://landsat.gsfc.nasa.gov/.

Thompson, S. K. 2002. Sampling, 2nd edition. John Wiley & Sons, Inc., New York, NY.

Thornton, P. E. 1998. Regional Ecosystem Simulation: Combining Surface- and Satellite-Based Observations to Study Linkages between Terrestrial Energy and Mass Budgets. The University of Montana, College of Forestry, Missoula, MT.

Thornton, P. E., H. Hasenauer, and M. A. White. 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. Agricultural and Forest Meteorology 104:255–271.

Thornton, P. E., B. E. Law, H. L. Gholz, K. L. Clark, E. Falge, D. S. Ellsworth, A. H. Goldstein, R. K. Monson, D. Hollinger, J. C. Paw U, and J. P. Sparks. 2002. Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. Agricultural and Forest Meteorology 113:185–222.

Thornton, P. E., and N. A. Rosenbloom. 2005. Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. Ecological Modelling 189:25–48.

Thornton, P. E., and S. W. Running. 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agricultural and Forest Meteorology 93:211–228.
Thornton, P. E., and S. W. Running. 2002. User's Guide for Biome-BGC, Version 4.1.2. The University of Montana.
Thornton, P. E., S. W. Running, and M. A. White. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. Journal of Hydrology 190:214–251.
Thornton, P. E., and N. E. Zimmermann. 2007. An Improved Canopy Integration Scheme for a Land Surface Model with Prognostic Canopy Structure. Journal of Climate 20:3902–3923.
Thum, T., T. Aalto, T. Laurila, M. Aurela, P. Kolari, and P. Hari. 2007. Parametrization of two photosynthesis models at the canopy scale in a northern boreal Scots pine forest. Tellus B 59:874–890.
Tibshirani, R. 1996. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58:267–288. Retrieved August 20, 2010, .
Turner, D. P., S. V. Ollinger, and J. S. Kimball. 2004. Integrating Remote Sensing and Ecosystem Process Models for Landscape- to Regional-Scale Analysis of the Carbon Cycle. BioScience 54:573–584.
Turner, D. P., W. D. Ritts, B. E. Law, W. B. Cohen, Z. Yang, T. Hudiburg, J. L. Campbell, and M. Duane. 2007. Scaling net ecosystem production and net biome production over a heterogeneous region in the western United States. Biogeosciences 4:597–612.
UNFCCC. 1998. Kyoto Protocol to the United Nations Framework Convention on Climate Change. United Nations.
UNFCCC. 2009. The Copenhagen Accord.
United States Forest Service Pacific Northwest Research Station. 2010. Volume estimation for the PNW-FIA Database. Page 69. United States Department of Agriculture, Forest Service, USFS Pacific Northwest Research Station.
Vanclay, J. K. 1994. Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests. CAB International, Wallingford, UK.
VCS. 2008. Verified Carbon Standard - Guidance for Agriculture, Forestry and Other Land Use Projects. Retrieved from http://www.v-c-s.org/.
Wallerman, J., and J. Holmgren. 2007. Estimating field-plot data of forest stands using airborne laser scanning and SPOT HRG data. Remote Sensing of Environment 110:501–508. doi: 10.1016/j.rse.2007.02.028.
Wang, Y. P., D. Baldocchi, R. A. Y. Leuning, E. V. A. Falge, and T. Vesala. 2007. Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight FLUXNET sites. Global Change Biology 13:652–670.
Wang, Y., H. Weinacker, and B. Koch. 2008. A Lidar Point Cloud Based Procedure for Vertical Canopy Structure Analysis And 3D Single Tree Modelling in Forest. Sensors 8:3938–3951. doi: 10.3390/s8063938.
Wang, Y.-P., C. M. Trudinger, and I. G. Enting. 2009. A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales. Agricultural and Forest Meteorology 149:1829–1842.
Waring, R. H., and S. W. Running. 2007. Forest Ecosystems: Analysis at Multiple Scales. Elsevier Academic Press, San Francisco, CA.

Weathers, K. C. 1999. The importance of cloud and fog in the maintenance of ecosystems. Trends in Ecology & Evolution 14:214–215.
Van der Werf, G. R., D. C. Morton, R. S. DeFries, J. G. J. Olivier, P. S. Kasibhatla, R. B. Jackson, G. J. Collatz, and J. T. Randerson. 2009. CO2 emissions from forest loss. Nature Geoscience 2:737–738.
White, M. A., P. E. Thornton, and S. W. Running. 1997. A Continental Phenology Model for Monitoring Vegetation Responses to Interannual Climatic Variability. Global Biogeochemical Cycles 11:217–234.
White, M. A., P. E. Thornton, S. W. Running, and R. R. Nemani. 2000. Parameterization and Sensitivity Analysis of the BIOME-BGC Terrestrial Ecosystem Model: Net Primary Production Controls. Earth Interactions 4:1–85.
Winrock International. 2010, November. The American Carbon Registry Forest Carbon Project Standard, Version 2.1.
Wulder, M. A., J. C. White, R. A. Fournier, J. E. Luther, and S. Magnussen. 2008. Spatially explicit large area biomass estimation: Three approaches using forest inventory and remotely sensed imagery in a GIS. Sensors 8:529–560.
Xiao, X., S. Boles, J. Liu, D. Zhuang, and M. Liu. 2002. Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. Remote Sensing of Environment 82:335–348. doi: 10.1016/S0034-4257(02)00051-2.
Yang, J., C. Zhang, X. Li, Y. Huang, S. Fu, and M. Acevedo. 2009. Integration of wireless sensor networks in environmental monitoring cyber infrastructure. Wireless Networks 16:1091–1108. doi: 10.1007/s11276-009-0190-1.
Zhang, X., and S. Kondragunta. 2006. Estimating forest biomass in the USA using generalized allometric models and MODIS land products. Geophysical Research Letters 33:1–5.
Zheng, D., L. S. Heath, and M. J. Ducey. 2007. Forest biomass estimated from MODIS and FIA data in the Lake States: MN, WI and MI, USA. Forestry 80:265–278.
Zhu, L., J. M. Chen, Q. Qin, J. Li, and L. Wang. 2009. Optimization of ecosystem model parameters using spatio-temporal soil moisture information. Ecological Modelling 220:2121–2136.
Zinke, P. J., and R. L. Crocker. 1962. The Influence of Giant Sequoia on Soil Properties. Forest Science 8:2–11.

9. Appendix

This is a listing of the variables used in the described analysis.

9.1. *Topographic Variables:*

9.1.1. Average elevation

9.1.2. Variance of the elevation of the cell.

9.1.3. Average aspect

9.1.4. Variance of the aspect of the cell.

9.1.5. Average slope

9.1.6. Variance of the slope of the cell.

9.1.7. A measure of the difference between the actual topography of the cell and a plane joining its corners.

9.2. ***Crown Segment Variables:***

9.2.1. Number of polygon centroids within a cell (pcount).

9.2.2. Average of the maximum height above the ground for the polygons (cell height).

9.2.3. Variance of the maximum height above ground for the polygons.

9.2.4. Crown closure as the percentage of the cell area covered by polygons.

9.2.5. Curvature of the cell in relation to the eight nearest neighbor cells (NLN).

9.2.6. Average LiDAR first return intensity for the cell.

9.2.7. Variance of the LiDAR first return intensity for the cell.

9.2.8. Average intensity of the infrared band of the CIR data fused to the polygons.

9.2.9. Variance of the intensity of the infrared band of the CIR data fused to the polygons.

9.2.10. Average intensity of the red band of the RGB data fused to the polygons.

9.2.11. Variance of the intensity of the red band of the RGB data fused to the polygons.

9.2.12. Average intensity of the green band of the RGB data fused to the polygons.

9.2.13. Variance of the intensity of the green band of the RGB data fused to the polygons.

9.2.14. Average intensity of the blue band of the RGB data fused to the polygons.

9.2.15. Variance of the intensity of the blue band of the RGB data fused to the polygons.

9.2.16. Ratio of the infra-red to red bands.

9.2.17. Normalized difference vegetation index(NDVI = (IR - red) / (IR + red)).

9.3. Image Variables

Image set variables consist of two types of analysis; spatial and frequency. Spatial analysis quantified the relationships between the pixels based upon their location with respect to one another. Frequency analysis characterizes the spectral characteristics of the pixels both in relation to one another and to standard frequency distributions.

There are no known relationships between these summary variables and the structural characteristics of the vegetation from which the light was reflected. This is an intriguing line of research but time has not yet been allotted for its pursuit. The CHM was treated as a greyscale image for this analysis.

9.3.1. **Spatial Domain**

9.3.1.1. Image profile analysis consisting of summaries of the eight vectors originating at the center of the image and radiating to each corner and the middle of each edge. This includes the mean, variance, median, skewness, kurtosis, entropy, mean absolute deviation, median absolute deviation of the pixels on the profile.

9.3.1.2. Image pixel analysis, the pixel based mean, variance, median, entropy, mean and median absolute deviation from a unit vector.

9.3.1.3. Histogram analysis of the image.

9.3.1.4. Sum of the Hough lines within the image. This has been used to identify plantations, and roads.

9.3.1.5. K-mean clustering of the color bands in the image.

9.3.1.6. Quadrant analysis to yield the ratio of the number of pixels two groups based upon color in each quadrant.

9.3.1.7. Number of cluster centers arising from the first group from the quadrant analysis.

9.3.1.8. The fraction of shadow.

9.3.1.9. The values of a three parameter Weibul fit to the image intensity histogram. The number of local maximum points and the location of the first three local maximums in a three dimensional histogram constructed in l, a, b color space.

9.3.1.10. The correlation, contrast, busyness, and texture strength of a neighborhood grey level difference matrix.

9.3.1.11. Neighborhood occurrence test based on eight offsets and compared with SID.

9.3.1.12. Contiguous region analysis including the average area, eccentricity, extent, orientation, and solidity of two size classes of blobs.

9.3.2. **Frequency Domain**

9.3.2.1. The ratio of the geometric mean to the arithmetic mean of the frequency space image.

9.3.2.2. Comparison of a vector of texture based properties such as contrast homogeneity correlation and energy using the gray scale co-occurrence matrix for a fixed diagonal offset on an image to a spectral information divergence.

9.3.2.3. Comparison of a vector of texture based properties such as contrast homogeneity correlation and energy using the gray scale co-occurrence matrix for a fixed diagonal offset on an image to a spectral angle measure.

9.3.3. **Reduced variable set**

9.3.3.1. CIR1-CIR8 the first eight principle components of the color infrared image

9.3.3.2. RGB1-RBG8 the first eight principle components of the true color image

9.3.3.3. LI1-LI8 the first eight principle components of the canopy height image

Chapter 4

Area Dependent Region Merging: A Novel, User-Customizable Method to Create Forest Stands and Strata.

Abstract:

Remotely sensed high-resolution imagery and LiDAR data can be used to develop stand delineations and stratifications for forest inventory and management purposes. A new Area Dependent Region Merging method is introduced that uses LiDAR data and expert knowledge to develop forest stands and strata based on user-supplied constraints. This method uses an area-dependent scale parameter that allows for different merging criteria based on the size of the objects being merged. This method was used to develop a new forest inventory that showed improved accuracy with significantly fewer field plots. Compared to non-area-dependent region merging approaches, this method more effectively reduced within stand variability and more closely matched a manual stand delineation.

1. Introduction

Global products acquired from satellites, such as the Moderate Resolution Imaging Spectroradiometer (MODIS – 250m to 1km pixels), Landsat (30m), Quickbird (1m pixels) and IKONOS (1m pixels) have all been used to make inferences about forest stocks, fluxes, and structure across both large and small scales (Running et al. 2004, Hall et al. 2006, Heinsch et al. 2006, Zhang and Kondragunta 2006, Potter et al. 2007b, Mildrexler et al. 2009, Song et al. 2010). In recent years, Light Detection and Ranging (LiDAR) data products as well as higher resolution imagery products acquired from airplanes rather than satellites have been used to make inferences about forest ecosystems at smaller spatial scales and higher resolutions (Akay et al. 2009, Asner et al. 2011). Some scientists have combined multiple remote sensing data products to generate wall-to-wall estimates (Potter et al. 2007a, Gonzalez et al. 2010, Ke et al. 2010, Golinkoff et al. 2011) or have used small samples of higher resolution remote sensing

products combined with lower resolution data to reduce costs and improve the accuracy of estimates (Næsset 2002, Wulder and Seemann 2003, Wulder et al. 2012).

As higher resolution remote sensing data products become available, it is possible and necessary to group small pixels into meaningful objects. The need to use objects as opposed to pixels arises when individual pixels are smaller than the features of interest (e.g. stands or individual tree crowns) in an image (Johansen et al. 2010). The aggregation of pixels into objects allows object level properties to be summarized efficiently while reducing the amount of data required to be stored about the image (Hofmann et al. 2011, Ali et al. 2009, Sasaki et al. 2012). Object-based analysis has also been shown to perform better than using pixel-based approaches in classifying images (Blaschke 2010). This improvement in classification using objects as opposed to pixels (the current approach) may also translate to improved performance of predictive models of future ecosystem state as well and may provide a better platform on which to run these models (Maselli et al. 2009, Golinkoff and Running 2013)

Geographic Object Based Image Analysis (GEOBIA) is a relatively new field that uses objects as the fundamental unit of analysis when interacting with geographic imagery (Hay and Castilla 2008, Blaschke 2010, Addink et al. 2012). How to best partition a landscape into objects and some of the implications of the segmentation chosen were first discussed within the context of the Modifiable Areal Unit Problem (Openshaw 1984). Openshaw explained that the location and boundaries of objects within a given area are in many cases arbitrary and that there are infinite possible combinations of non-overlapping objects, some more suitable than others, in defining reasonable divisions within the landscape. Many methods to optimally segment a landscape have been proposed to address this issue. These methods build upon years of image

segmentation work in the computer vision and medical imaging fields (Fu and Mui 1981, Pal and Pal 1993).

Many authors have reviewed segmentation methods and discussed specific forestry applications of these methods. Cheng et al. (2001) provided a general overview of the many different approaches that can be used to segment a color image. Dey et al. (2010) related these techniques to the remote sensing field and reviewed the most commonly applied methods. Broadly, the results of image segmentation can be considered to fall on a spectrum from over to under segmented. Over-segmented images have too many segments and break objects into their component parts. Under segmented images have grouped relevant objects together and image fidelity and resolution is lost as a result (Möller et al. 2007, Marpu et al. 2010). Segmentation methods can also be distinguished by those that create final objects by grouping similar pixels and/or sub-objects together versus those that create final objects by splitting larger objects apart based on discontinuities (Addink et al. 2012). This concept is similar to agglomerative and divisive hierarchical methods of cluster analysis (Mardia et al. 1979). Segmentation methods also vary based on the input data considered: spectral intensities or digital numbers of pixels, spatial attributes such as neighborhood relationships and texture, object shape and size, and prior knowledge of the image (Dey et al. 2010). Segmentation can also vary based on the degree of user guidance or supervision versus automation as well as whether a model driven versus image driven approach is used (Baatz and Schäpe 2000, Hay and Castilla 2008, Dey et al. 2010). In model driven approaches, an underlying image structure is assumed and used as a model that then drives the image segmentation.

Because of the myriad options available to segment images and the inherent subjectivity of the final segments, the method chosen depends heavily on the goals of the segmentation and

the spatial characteristics of the forest. In forestry operations, the final object of interest is often a stand. The definition of a stand may vary but in general it refers to a contiguous area of forest that is managed as a unit and that has trees that are homogenous relative to surrounding stands (Sullivan et al. 2009). Traditionally, forest stands were defined by human photo interpreters. However, as high-resolution remote sensing data and powerful microprocessors have become ubiquitous, it has become possible to remove some of the human labor and subjectivity from the stand delineation process by developing repeatable algorithms to complete this task (Leckie et al. 2003). Generally, when devising forest segmentation, it is preferable to have some limits on the sizes of polygons. Forest stands are used to manage for inventory and harvest operations and stands that are too large or too small become difficult from a data management perspective and impractical for operations.

Forest stands have been delineated using many different approaches and the method described here is an extension of some of the work that has already been done in this field. One of the most widely used approaches to the creation of objects is the eCognition program (see http://www.ecognition.com/). This commercially available software uses fuzzy logic and incorporates user-defined variables to define the importance of object shape as well as a merge stopping criteria (scale parameter) (Baatz and Schäpe 2000, Benz et al. 2004). This software has been used in many studies and has been show to be a powerful tool to segment forests into stands and into individual tree polygons (Van Aardt et al. 2006, Pascual et al. 2008, Riggins et al. 2009, Ke et al. 2010). Although powerful, this program is not freely available and requires users to iteratively choose the scale parameter that is optimal for their work. This scale parameter is a constant that will change the sizes of polygons. However, this parameter is not directly a constraint on forest stand size and instead defines a degree of dissimilarity that causes merging to

stop. This is problematic, as many managers would like area constraints to limit the results of a stand delineation method more directly.

Other researchers have designed and used different approaches to segmentation. Leckie et al. (2003) built individual tree crown polygons from high resolution multi-spectral imagery and then combined these individual crown polygons into larger stands based on crown closure, stem density, and species composition. This approach smoothed individual tree crown data and used a minimum size constraint to guide stand creation. Another approach designed by Haywood and Stone (2011) for Eucalyptus stands in Australia uses a single minimum area constraint but also included a similarity metric that can incorporate data from a user specified number of remote sensing layers. Other approaches to forest segmentation use an iterative nearest-neighbor approach that selects regions to merge in several iterations based on relaxing the difference constraint. These algorithms proceed by increasing the amount of difference between neighbors that will trigger merging to occur until either a final mean polygon size is reached or until a maximum difference in feature space is reached (Hay et al. 2005, Castilla et al. 2008, Wang et al. 2010). By using single global targets for mean polygon size and/or a constraint on minimum size, there may be small stands that are quite different from their neighbors that are merged instead of retained. For example, if a stand is smaller than the minimum size constraint and much different from its neighbors, these algorithms will force this polygon to be merged. Similarly, if a stand is larger than the minimum size constraint but more different than the feature space difference constraint, this polygon will be merged with its neighbors.

All of these methods show varying levels of success in defining stands that accurately partition forest systems. However, a forest manager or inventory planner may need more control over the final stand delineation and how important stand area should be in determining when

objects are merged. An increase in control should allow managers to specify both minimum and maximum stand sizes, stand shape, the variables used to define stand boundaries, and the differences between objects that allow for objects to remain as distinct stands. In particular, unlike existing region merging area controlled segmentation methods, an area-dependent region merging approach allows for more control over final stand size. To address these needs, an area-dependent, region-merging (ADRM) method is proposed that allows users to select relevant real world sub-object characteristics (as opposed to spectral or textural properties) and specify polygon sizes given user-defined feature space distances. To evaluate the success of this result, a novel, scaled-variability metric is introduced to compare stand delineation outcomes. This evaluation method is a fast and intuitive approach that can allow analysts to understand the effectiveness of any given stand delineation model run compared to other model runs or against a reference case.

2. Study Site / Data

2.1. Study Site

The Big River and Salmon Creek Forests are located in Mendocino County, CA near Ft. Bragg and owned by The Conservation Fund (TCF) (see location map Figure 1). TCF is a non-profit organization whose mission is to conserve threatened and important ecosystems and promote rural economies across the United States of America (see http://www.conservationfund.org). Both forests are dominated by Redwood (*Sequoia sempivirens*) and Douglas Fir (*Pseudotsuga menziesii*). The Big River forest is ~4,700ha and the Salmon Creek forest is ~1,700ha. These forests are currently managed as a unit and together are a verified and registered forest carbon offset project under the Climate Action Reserve Forest

Project Protocol version 2.1 (CAR FPP v2.1) (CAR 2007). The Big River / Salmon Creek (BRSC) ownership has been extensively managed in the past using both even- and uneven-aged harvesting. This has resulted in a patchwork of old clear-cuts dominated by one size and age class as well as mixed-size and age stands. It is currently actively managed for both timber production and carbon offsets, in addition to watershed restoration and public recreation.

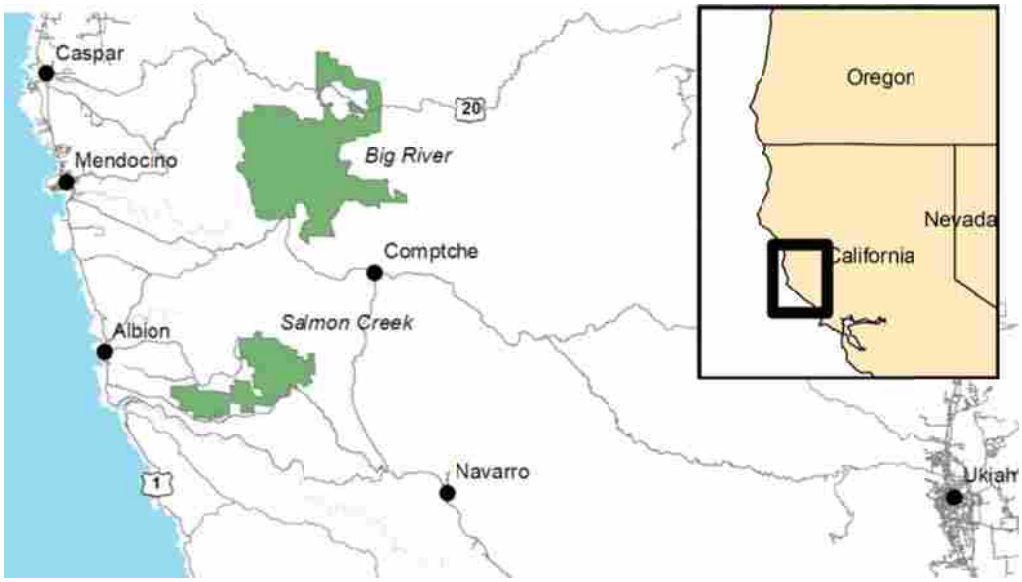

Figure 1: Area Map

Due to the history of management on this property, most stands are made up of second and third growth trees. Although there are some past clear-cuts where stand boundaries are easy to detect, much of the forest has heterogeneous stand conditions that grade into other stand conditions making traditional air photo interpretation or automated stand delineation difficult. The result of these conditions results in highly subjective stand boundary creation and large stands that make management planning difficult.

2.2. Remote Sensing Data

The BRSC forests were flown for full coverage Light Detection and Ranging (LiDAR) in August 2011. The LiDAR data was collected along east-west flight transects flown at ~900m

altitude with at least 10% overlap between flight lines. The LiDAR data was collected using an ALTM Gemini from Optech Incorporated sensor with up to 4 returns per point and on average 4 points per square meter. The data was binned into $1m^2$ pixels to create a digital elevation model (DEM) and a canopy height model (CHM). The CHM bins were based on the highest hit within the $1m^2$ pixels and the DEM was created using the ground point returns. The CHM is the total height of each pixel minus the bare earth DEM elevation and represents the top of the tree canopies. The stand delineation method described here relied primarily on the LiDAR data but could also incorporate imagery as well.

2.3. Ground Inventory Data

2.3.1. Existing Inventory

The existing inventory on the BRSC forests prior to the acquisition of the remote sensing data contained plots collected over the previous 11 years. The initial inventory sampling design was a multi-staged probability proportional to stand area list sample within a broader stratified inventory. This means that stands within each stratum were selected with replacement with probability proportional to their area (Borders et al. 2005). The plots in this inventory were variable radius basal area factor prism plots. The prism factor varied depending on the age and stocking of the stand with the target of 4 to 8 count trees included in each plot (Shiver and Borders 1996, Bell and Dilworth 2007). In 2011, the existing inventory relied on 2597 plots and resulted in 3.43% estimated inventory accuracy at the 90% confidence level (estimated inventory accuracy = (z-statistic * SE) / mean where the z-statistic for the 90% confidence level = 1.645). Estimates of carbon density are calculated using the approved biomass and carbon equations required by the CAR FPP v2.1 and most are based on national scale biomass estimators developed by the US Forest Service (Jenkins et al. 2004, CAR 2007). The stands were stratified

using three variables: dominant species (by the percentage basal area), dominant size class (by percentage basal area), and canopy closure.

As areas of the forest were harvested, these locations were delineated by hand on the stand map and re-inventoried as new stratum. Table 1 shows the summary of the plot data used. As can be seen, the majority of plots are at least 8-years-old and this, paired with changing on-the-ground conditions due to harvest and forest growth, created a need for a complete re-inventory of the entire forest area.

Table 1: Existing inventory plots used by year. As is often the case in forest inventories, plots are collected on a rolling basis in between full forest re-inventory efforts. In this case, data from plots collected over 11 years were considered for the original inventory estimates. All plot data has been grown forward for final inventory results.

Year	2011 Plots used	% of Total
2000	21	4.7%
2001	59	2.3%
2002	1271	48.9%
2003	26	1.0%
2004	371	14.3%
2005	73	2.8%
2006	0	0.0%
2007	336	12.9%
2008	102	3.9%
2009	80	3.1%
2010	146	5.6%
2011	12	0.5%
Total:	2597	

2.3.2. 2012 Inventory

The updated 2012 inventory was installed after the final stand delineation and stand stratification was completed. The sampling design used was exactly the same as the prior inventory – a multi-stage probability proportional to area within strata list design (Borders et al. 2005). All inventory plots were installed with variable radius basal area prisms as in the previous inventory. A total of 677 plots were installed in 2012.

3. Method

The method proceeds in four steps (Figure 2). In the first step, the CHM of the forest is partitioned into management compartments. By first partitioning the property in this way, the final results can be controlled more easily and will more closely relate to the management constraints of the forest (see (Leckie et al. 2003)). This step was done by a local forester and divided the forest along major roads and streams to create large areas that would share logging infrastructure. Within each compartment, small objects – microstands – are then created using an appropriate method (a watershed algorithm applied to a smoothed CHM, as detailed in section 3.1). This step is designed to create sub-objects that correspond to similarly sized clumps of trees but are smaller than stands. The third step involves the user iteratively selecting the optimal constraints for stand creation. This involves selecting and weighting the variables adopted in the region merging algorithm and selecting the stand shape and size constraints. In the final step, the stands within each compartment are merged together and stratified to create a full property level stand delineation and stratification. The input and output formats for this method are shape files and the program is written in Perl using the GDAL/OGR module (GDAL Development Team 2012) . Please contact the author for access to the source code.

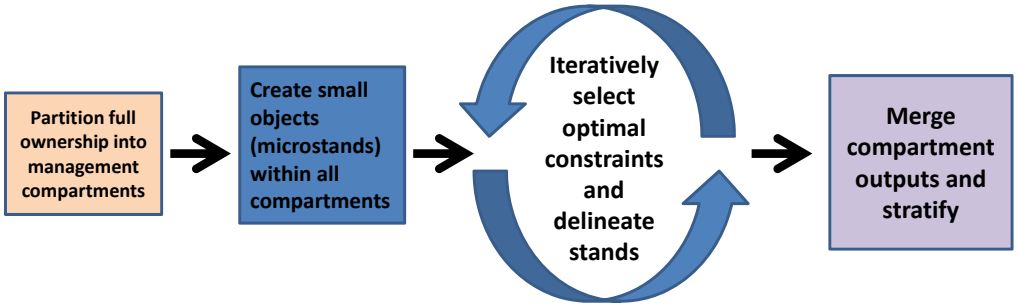

Figure 2: Broad method outline.

3.1. Microstand Creation

The first step in creating stand polygons requires creating small regions across the full forest extent. There are many methods that have been put forward to move from an initial set of

pixel-based layers to small microstand objects. For this particular case, microstands consisting of clumps of similarly sized trees were desired. To build these microstands, the $1m^2$ resolution CHM was first up-sampled to $4m^2$ cells. A 5 by 5 median filter was applied, to preserve the edges in the image, as median filters have been shown to be edge preserving smoothers (Hay et al. 2005). A morphological gradient image was then created from this up-sampled, smoothed CHM layer. The morphological gradient is a measure of local variation in an image and has edge enhancing effects. The morphological gradient image was then further smoothed using a 3 by 3 median filter. The smoothed morphological gradient image was converted to a microstand map using the watershed algorithm (Figure 3). The watershed algorithm finds areas of pixels within contours (analogous to how a watershed is defined in nature using the flow of water) (Gonzalez et al. 2009). Microstand creation was done using the MATLAB Image Processing Toolbox and ArcGis software (*MATLAB Image Processing Toolbox* 2011, *ArcMap 10.1* 2012).

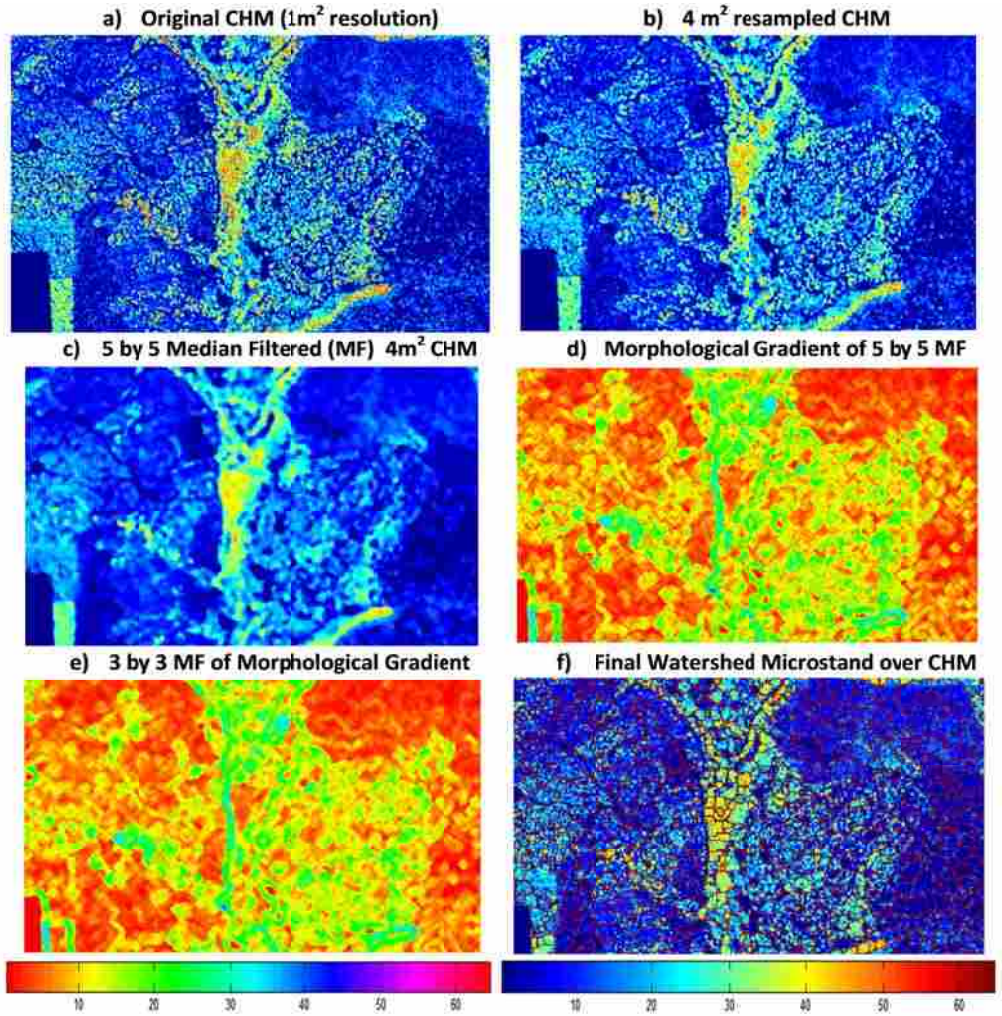

Figure 3: Microstand creation routine. MF = Median Filter. CHM = Canopy Height Model. Blue to yellow colors represent height in meters – see the legend for images a, b, c, and f under image f. Green and red color spectrum represents the gradient of heights in meters – see the legend for images d and e under image e.

The intent of this work is to highlight the region merging approach using an area-dependent scale parameter. This approach requires that small regions – microstands – are first created and then merged. As described above, a smoothed, gradient image was processed using a watershed segmentation algorithm to create the initial microstands. However, any appropriate method could be used to generate the microstand objects that will be used as the basis for the final stand creation. These microstand polygons should ideally range from a single tree crown to a collection of several tree crowns that are of similar type and species. Alternatively, microstands can be made to be plot sized and sampled prior to region merging.

3.2. Area-Dependent Region Merging (ADRM)

Once microstands are created, these regions can be merged together based on a set of user specified constraints. Regions are labeled using characteristics of importance to forest managers. Using this approach, managers have a better understanding of the factors that drive the final stand delineation and more control over the outcome. For example, using LiDAR data, each region was assigned an average canopy height defined as the mean of the maximum heights of trees based on a tree crown segmentation. The tree crown segmentation was done using the watershed method on the inverted, unsmoothed CHM layer for areas taller than 3m. Each microstand was also assigned a percent canopy cover metric defined as the percentage of LiDAR returns occurring above 2m in height. These metrics closely relate to volume and stand vigor and are therefore important when trying to map merchantable timber and carbon stocks (Nilsson 1996, Popescu et al. 2003, Ioki et al. 2009, Latifi et al. 2010). Additional metrics (e.g. species composition) can be considered on an as needed basis. However, it is recommended that at most three variables are used to avoid problems associated with high-dimensional neighborhood calculations. This "curse of dimensionality" leads to excessively large neighborhoods for each individual variable and reduces the skill of the results in predictions and classifications (Hastie et al. 2009).

Once a set of attributes is assigned to each microstand, the manager weights each attribute depending on the importance of the attribute for the final stand delineation. The manager also chooses an optimal shape weight as well. In this case, the shape constraint used was the ratio of polygon perimeter to the square root of the polygon area. This is equivalent to the compactness variable as defined by Benz et al. (2004). Other possible shape constraints that have been discussed in the literature are the simple perimeter to area ratio, the object smoothness

defined as the ratio of the polygon perimeter to the bounding box perimeter, or the object rectangularity defined as the ratio of object area to the bounding box area (Turner et al. 2001, Benz et al. 2004, Wang et al. 2010).

The next step in the ADRM process is to define the area constraints that should limit the stand creation results. The size of stand polygons is important to forest managers for several reasons. First, from an operational perspective, stands should be units that fit practical harvesting requirements, such as large enough to meet certain economies of scale (typically at least 2ha) while at the same time not being too large to be operationally infeasible to manage. Second, from a sampling perspective, variability is introduced if there is a wide range of variability in stand size. Therefore, it is important that most stands are similarly sized (Shiver and Borders 1996). The most suitable stand areas are often well known by forest managers with experience in the field. Thus, this method enables managers more control over the outcome of the stand delineation process to meet these needs.

The final user-defined parameters to set is/are the scale parameter(s) that will define the difference thresholds that drive polygon merging, as well as the type of scale parameter to use. There are three types of scale parameters that can be chosen: 1) a standard, non-area-dependent scale parameter (ADSP), 2) a stepwise, discontinuous area-dependent scale parameter, or 3) a continuous area-dependent scale parameter. A standard, non-ADSP uses one difference for all merging regardless of polygon size. A stepwise ADSP uses one or more area / scale parameter pairs to define different merging criteria given different area limits. A continuous ADSP uses a smooth boundary to define different merging criteria given different areas. A continuous ADSP is any equation that connects the points (areaMin, maxDiff) and (areaMax, minDiff) (see Figure 4 for a set of example ADSPs). Figure 5 shows a comparison between traditional region merging

approaches and the ADRM method. All microstand attributes and the shape attribute are standardized using a range standardization approach as this has been shown to result in better outcomes during clustering (Milligan and Cooper 1988). Variable standardization was used to insure that the absolute variable size was normalized so all variables might have equal weight in the analysis, allowing user defined weights to be applied appropriately.

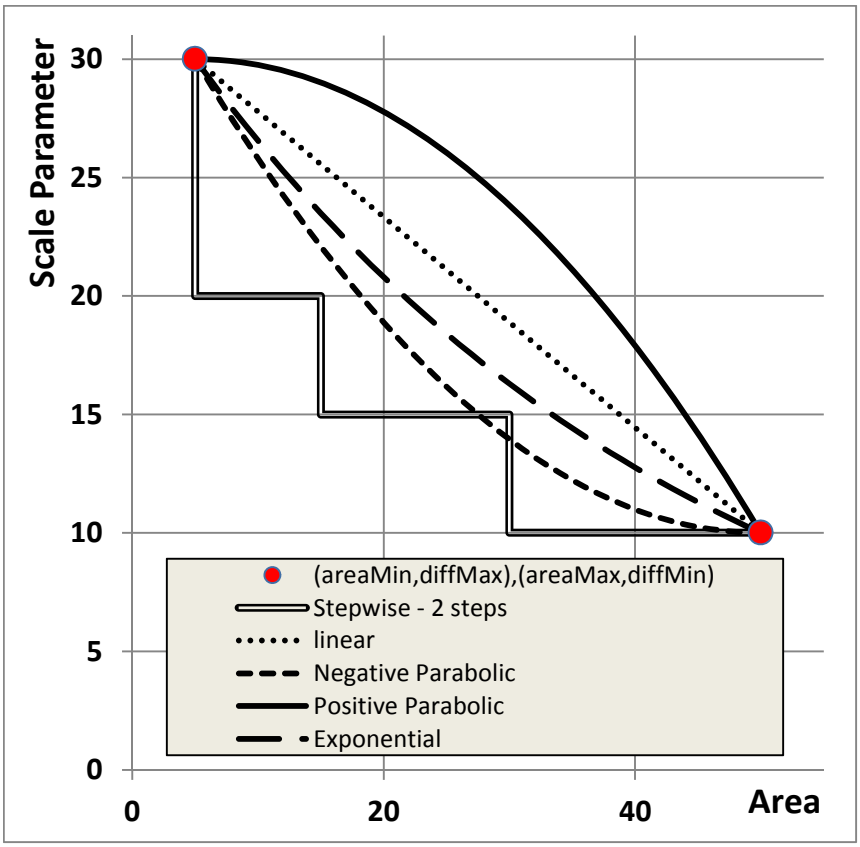

Figure 4: Example of area-dependent scale parameter for 5 different types. Any polygon and neighbor pair whose difference in feature and shape space is below these lines would be merged.

After all user-defined parameters have been chosen, the ADRM process can begin. The order of merging affects the final merging outcome and several region merging algorithms and optimizations have been discussed in the literature (Castilla et al. 2008, Wang et al. 2010, Haywood and Stone 2011). The region merging approach used in this method was an iterative relaxation of the difference constraint similar to that described in Wang et al. (2010). Merging of regions proceeds by first choosing the best neighbor if the best neighbor is less than the ADSP. A

neighbor is considered "best" if for two polygons p_i and n_i, n_i is the least different from p_i for all neighbors n_1 to n_n and p_i is the least different from n_i for all of n_i's neighbors p_1 to p_n (where difference is calculated as the sum of the weighted attribute and shape differences). "Best" merging continues until no further merges can occur. At that point, the constraint is relaxed from the best constraint to a difference criterion that is some fraction of the maximum scale parameter allowed. The fraction is defined by a user selected number of iterations that by default is 5. For each potential merge, each polygon is first checked to see if the difference between itself and its neighbor is less than the tolerance of the current iteration and that the difference is less than the ADSP. Only if both constraints are met and both polygons are smaller than the maximum area is the polygon merged with its neighbor. A random ordered list (Fisher and Yates 1948) of merge polygons is used at every iteration to avoid problems with clumping during the merging process.

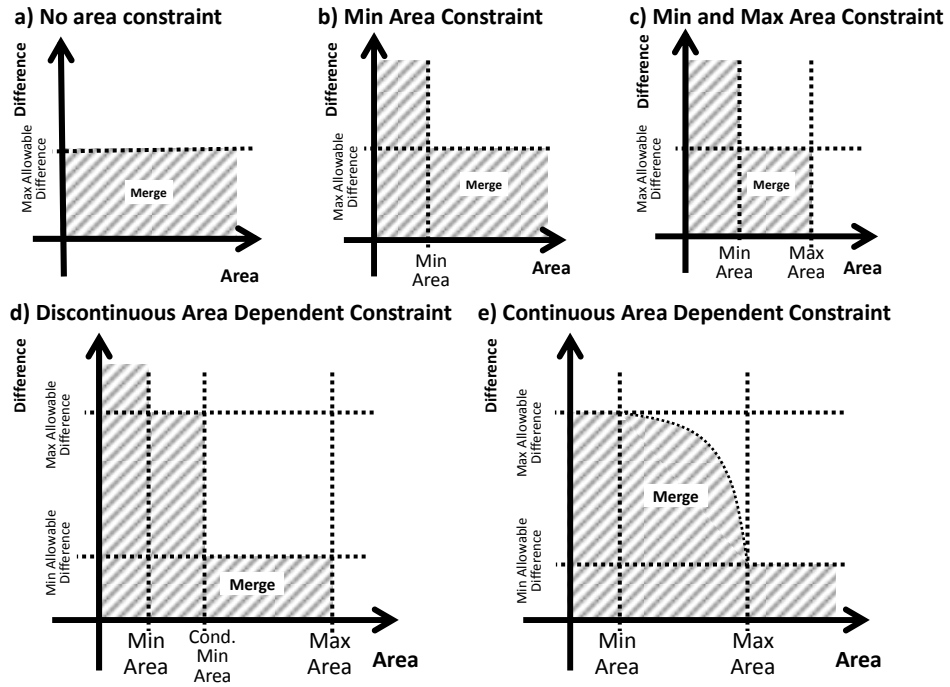

Figure 5: Conceptual framework of proposed area-dependent scale parameter to control region merging. Difference represents the scale parameter constraint.

3.3. Strata Formation (Classification)

Stratification is a well-known method to reduce sampling effort and improve inventory accuracy when estimating population characteristics (Thompson 2002). The last step in the method to develop stand delineations for a forest is to stratify these stands into similar forest types. The stratification is again user driven and is supplied by the user before running the merging tool. The strata are defined by listing the variables that will be the basis of the strata and the breaks that define the different bins within each stratum. In this way, a full set of strata can be used to partition all the stands once they are made. The strata are applied based on the attributes of each final stand polygon and these attributes are based on the original microstand polygon attributes. In the example case where canopy height and canopy cover are used as the variables of interest, the strata were defined as 7.62 m (25ft) height bins and 20% canopy cover bins (Open: <20%, Low: 20% to 40%, Medium: 40% to 60%, Dense: 60% to 80%, Extremely Dense: >80%).

A series of model runs using different parameters were experimented with to define a set of potential stand delineations. Table 2 summarizes the parameters chosen for each model run. To assess how these stand delineations and stratifications performed they were compared to a photo interpretation done by hand by the forest manager on a small portion of the Salmon Creek property. Even for a trained forester with experience in this geographic region, creating a stand map is difficult and subjective given the variability across the forest. A stratification accuracy assessment of the hand-done stratification versus a series of region merging processes was then conducted. Each model run with different input constraints was visually examined. The results were also compared to the number and the mean variability of the hand-created stand delineation.

Table 2: Summary of stand creation model runs. "None" means no area-dependent scale parameter was used, "stepwise" used a discontinuous approach, "linear" used a form diff = -a(area) + b, "exponential" used a form diff = a*exp(-b(area)), "Neg. Parab" is a downward opening parabola and "Pos. Parab" is an upward

opening parabola with equation form y = (area – a)^2/4b + c. "ShapeWt" is the user defined weight that shape parameters are given relative to the attribute similarity. "maxDiff" is the limit above which no merging between features occurs (unless the features are less than minArea). The ("areaN","diffN") pairs are used to define either stepwise or continuous ADSPs.

Run	ADSP used	maxDiff	maxArea	minArea	shapeWt	area1	diff1	area2	diff2	iters
1	**None**	30	50	3	0%					5
2	**None**	20	50	3	0%					5
3	**None**	10	50	3	0%					5
4	**None**	15	50	3	10%					5
5	**None**	15	50	3	5%					5
6	**Stepwise**	20	50	3	5%	10	15			5
7	**Stepwise**	15	50	3	5%	10	10			5
8	**Stepwise**	20	50	3	5%	30	5	10	15	5
9	**linear**	20	50	3	5%	10	15			5
10	**linear**	15	50	3	5%	10	10			5
11	**linear**	20	50	3	5%	10	5			5
12	**exponential**	20	50	3	5%	10	15			5
13	**exponential**	15	50	3	5%	10	10			5
14	**exponential**	20	50	3	5%	10	5			5
15	**Neg. Parab**	20	50	3	5%	10	15			5
16	**Neg. Parab**	15	50	3	5%	10	10			5
17	**Neg. Parab**	20	50	3	5%	10	5			5
18	**Pos. Parab**	20	50	3	5%	10	15			5
19	**Pos. Parab**	15	50	3	5%	10	10			5
20	**Pos. Parab**	20	50	3	5%	10	5			5

4. Results

The stratification accuracy of each stand delineation outcome was calculated by examining the area within each strata intersection of the manually delineated stands and the automatically delineated stands (Congalton 1991). This accuracy calculation is not a sample of classes but rather looks at the full forest area to see how many pixels were placed into the same class as the hand-delineated stand layer. It should be noted that because the hand stand delineation is highly subjective, this manual delineation serves as a benchmark to compare the performance of the ADRM method. The strata were derived for both the hand delineation and the automated delineation based on the average characteristics of all of the combined microstand polygons. In this way, these are more subjective classes rather than true forest types. Although different than a classical classification accuracy assessment, the classes of managed forests are

difficult to define and subjective themselves so this was deemed an unbiased approach to assess the similarity of the results. Lastly, it should be noted that the stratification accuracy relates only to the stratification and not the forest segmentation stand boundaries.

In addition to the stratification accuracy results calculated, a new metric was designed to provide an estimate of the average variability within the created stands. The weighted, scaled total variance (WSTV) describes the skill of the stand creation routine based on the degree to which the stands minimize within stand variability. The WSTV was calculated for each model run. This metric is based on the original microstand objects. Each microstand has attributes of interest based on the needs of the manager. These attributes are also given importance weights by the manager. Using these weights, the linear combination of the attributes is calculated for each microstand. Once final stands have been created, these linear combinations are aggregated to calculate the mean and variance of the microstands within the final stand delineations. Weighted, scaled total variance is defined as:

$$WSTV = \left(\sum_{i=1}^{n} \frac{area_i}{area_{tot}} * lcVar_i \right) * n$$

$$area_{tot} = \sum_{i=1}^{n} area_i$$

$$lcVar_i = \sum_{j=1}^{m_i} \frac{\left(lcMean_j - lcMean_i \right)^2}{m_i}$$

$$lcMean_i = \sum_{j=1}^{m_i} \frac{lcMean_j}{m_i}$$

$$lcMean_j = \sum_{k=1}^{p} w_k * a_k$$

where n is the total number of stands created, m_i is the number of microstands within each final stand i, p is the number of user-chosen merge parameters, and w_k is the user-defined weight of attribute a_k. $lcVar_i$ is the variance of the linearly combined, user-specific attributes. In this way, the WSTV represents the variability across all variables of interest and all final stands.

The estimates of stand level variability are area weighted by the stand area and summed across all stands to calculate a final metric of mean forest variability. This estimate of forest level within stand variability is highly dependent on the size of the stands. Larger stands will tend to have larger variability as they typically have more types of sub-objects. Because of this, stand creation routines that create more stands will necessarily have lower variability within stands (see Figure 6). Therefore, to assess the skill of the stand creation routine it is necessary to remove the impact of the stand number on the estimate of within stand variability. This is done by multiplying the area weighted forest variability by the number of stands and results in the scaled, area weighted estimate of forest variability. The results of each stand creation run can be seen in Table 3 and Figure 7.

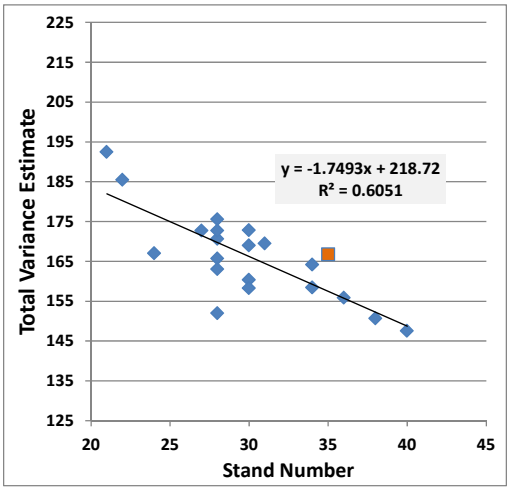

Figure 6: Forest level variance as a function of the number of stands created. The orange square is the manual reference stand layer.

Table 3: Summary of model run results. The three colored models represent some potential results with optimal stratification accuracy and weighted, scaled total variance. Please refer to figure 7 for a graphical explanation of why the three colored models were chosen.

Run	Type	Stratification Accuracy	Mean Variance	Stand Number	Scaled Variance
Reference	**by hand**	**NA**	167	**35**	5,838
Run1	**None**	44.1%	192	21	4,042
Run2	**None**	50.6%	167	24	4,009
Run3	**None**	50.4%	148	40	5,902
Run4	**None**	47.0%	152	28	4,256
Run5	**None**	53.6%	169	31	5,254
Run6	**Stepwise**	58.4%	160	30	4,811
Run7	**Stepwise**	55.7%	164	34	5,583
Run8	**Stepwise**	50.3%	169	30	5,071
Run9	**linear**	45.3%	173	27	4,663
Run10	**linear**	48.9%	156	36	5,611
Run11	**linear**	53.3%	166	28	4,640
Run12	**exponential**	49.4%	176	28	4,916
Run13	**exponential**	52.2%	171	28	4,778
Run14	**exponential**	54.9%	163	28	4,565
Run15	**Pos. Parab**	52.3%	185	22	4,080
Run16	**Neg. Parab**	48.3%	158	30	4,749
Run17	**Neg. Parab**	51.6%	173	28	4,837
Run18	**Pos. Parab**	53.3%	173	30	5,186
Run19	**upParab**	50.6%	151	38	5,725
Run20	**upParab**	55.1%	158	34	5,387

The horizontal black line in Figure 7 is the scaled variability metric for the hand delineated stands. As can be seen, the manual stand layer did not perform as well in terms of reducing within stand variability as the computer generated stands. For this reason, the stratification accuracy percentage should be viewed as a guide but not a definitive metric of the success of the stand delineation. Manual stand creation is highly subjective and may or may not represent the best partitioning of variability. The accuracy statistic does, however, allow for a comparison between the outcomes and how closely the automatic results match the manual results in terms of classification accuracy.

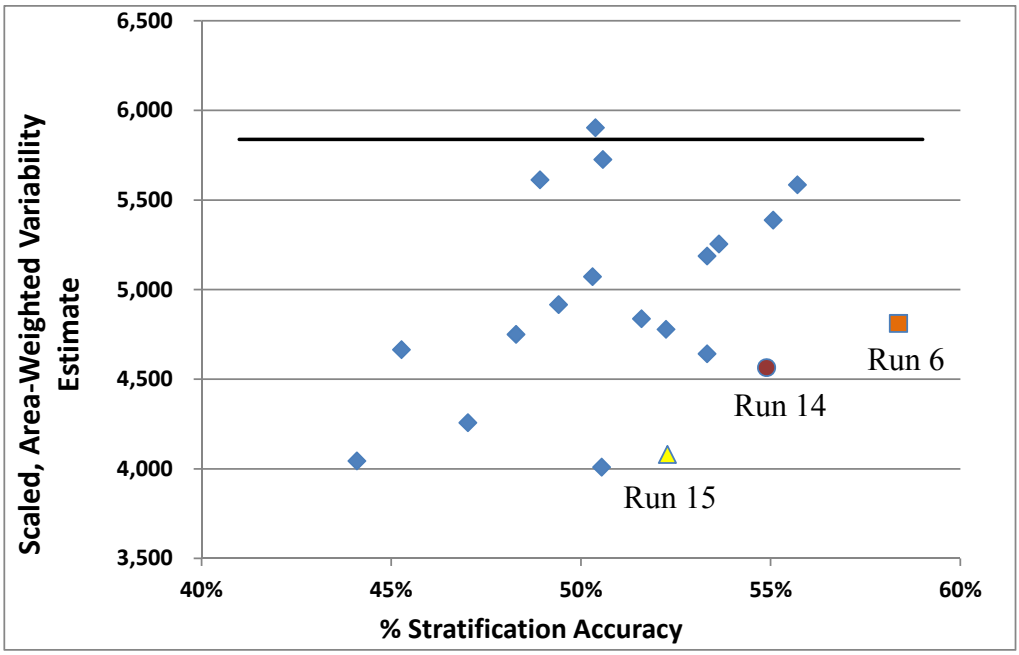

Figure 7: Scaled, area-weighted forest level variability estimate versus the stratificaiton accuracy of 20 different stand creation model runs (see table 1 for a description of the model runs). The orange square (run6), the red circle (run14), and yellow triangle (run15) represent some of the best outcomes as they show the least variability but have the greatest accuracy when compared to the reference layer. The horizontal black line shows the weighted scaled forest variability estimate calculated for the manual reference stand layer.

Figure 8 shows the manual stand delineation compared to the three highlighted runs in Figure 7. It is interesting to note that the best performing stand creation runs all used an ADSP. This speaks to the importance of area dependency in stand creation. The final stand delineations also show less compactness of stands and more complicated boundaries as they more closely follow forest features than they would in a manual delineation. This difference is due to the nature of the region merging algorithm as well as the preference of this photo-interpreter to build smoother stand boundaries.

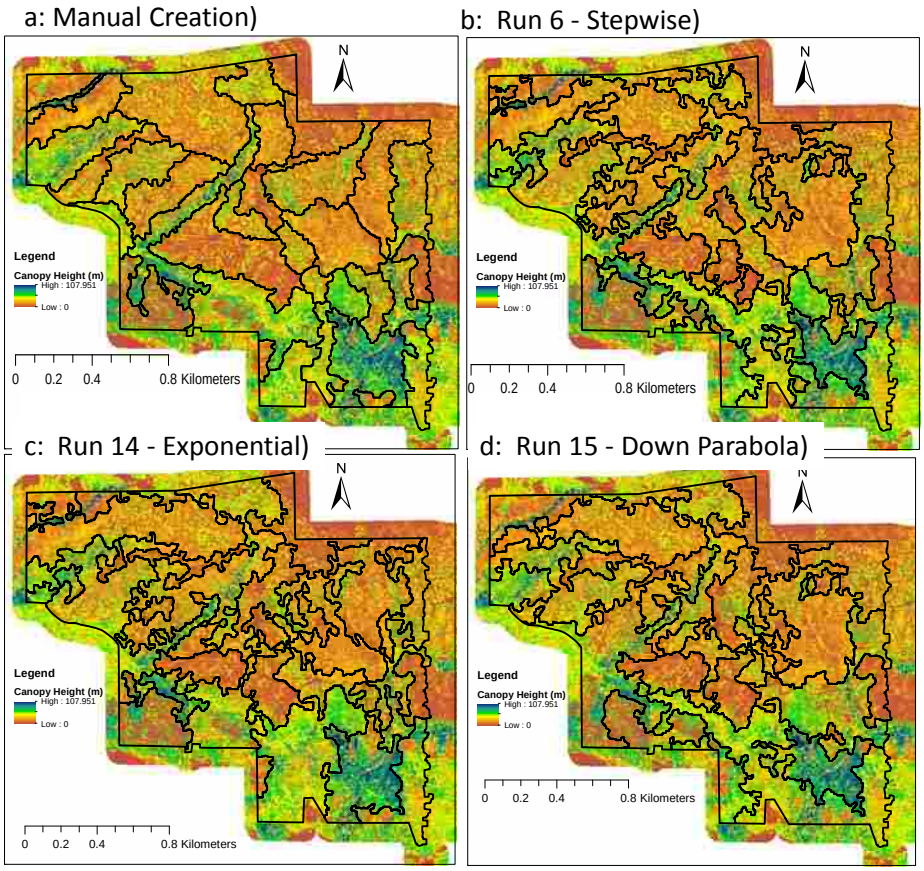

Figure 8: Comparison of manual stand layer and three automated stand creation runs.

Based on these results, a stepwise ADSP was chosen for the full forest stand delineations. The results of the estimated inventory accuracy using the new stand layer were compared with the original estimated inventory accuracy. The new inventory achieved an estimated inventory accuracy of 3.81% at the 90% confidence level using a total of 677 plots. The prior inventory had an estimated inventory accuracy of 3.43% at the 90% confidence level but used 2597 plots – almost 4 times as many.

5. Discussion

A flexible, user-customizable, area-dependent region merging stand creation method has been described above. This method allows for fine grain control of the stand delineation process using real-world attributes that forest managers can understand. At the same time, it provides

powerful, area-based merging criteria that can serve to better partition the variability of a forest. A stratification approach that allows for on-the-fly forest classification was also introduced to allow managers' input both into the stand creation process and the final stratification results. This method was compared to a manually delineated stand map used as a reference and was also evaluated using an area-weighted, scaled metric of within-stand variability that can be easily generated and provides an accessible and rapid means of assessing the skill of a given stand creation model run. A large improvement in sampling efficiency was observed using this method. Over large areas, these improvements can result in substantial cost savings as field inventory is one of the most expensive elements of forest management.

The ADRM method described here was developed for use in an active working forest to meet the needs of forest managers to create reasonable stands for operations. There are several issues managers should consider when using this method. First, from an operational perspective, the boundaries created by the ADRM method are complicated than stands created by humans. The more complicated boundaries may lead to difficulties in locating stand boundaries in the field. In general, boundaries created in an automated, rule-driven way will be more complicated that boundaries created by human interpreters. Photo-interpreters, especially those with experience managing timber harvests, are more likely to group dissimilar areas together to create smoother stand boundaries. The automated approach, even when shape constraints are imposed, is less likely to do this. In practice, this is not a major problem as the inventory estimates can be used to inform management even if the stand boundaries do not exactly line up with how a forest may be harvested in the future.

This issue of complicated boundaries may also make it difficult to be sure a given plot falls in the correct stand. For this work, it was found that plots should be located with GPS

coordinates using handheld GPS receivers to remove bias and cruiser subjectivity when locating plots. Although in some cases plots may fall in neighboring stands using this method, most plots will fall in the correct stand and the resulting inventory is still much more accurate than a traditional heads-up digitized stand layer. To further constrain the final stand layer, it may be beneficial for large forest ownerships to first manually delineate logical management compartments and then run this stand delineation within those compartments to constrain the outcome. Even with the best stand delineation (from the perspective of variability reduction), one stand may fall into multiple areas that in reality would never be managed simultaneously. To control for this, the use of management compartments to constrain the results is a critical step in this process. The use of logical management units will also reduce the complexity of stand boundaries in many areas.

In the example developed here, the stand delineation was applied first, a stratification was developed, and a field sample then proceeded within the delineated stands. However, it may be preferable to stratify and sample within microstands and then merge these microstands into larger stand units after sampling. By reversing the order, sampling may be conducted first, allowing the user to generate any number of post-sample stand layers. The inventory data may then be used to both inform the microstand attributes for merging purposes and to populate the final stand layer, which may be useful depending on monitoring or management goals. Sampling microstands first allows for flexibility in creating stand layers to meet these objectives. Any microstand layer can be used, although microstands should be at least as big as the plots that would be installed within them (see Golinkoff et al. (2011) for a discussion of finding the optimal microstand size).

This reversed order approach was done using square microstands in a pixel-based stratification done by Golinkoff et al. (2011). The original inventory was designed to minimize field effort while at the same time achieving a highly accurate inventory. However, future harvest planning was not considered during the design of this pixel-based inventory. For harvest planning, the square microstands were merged into larger stands that can now be used for management planning.

The models used in this work were chosen to illustrate some of the differences that different ADSPs produce when creating stand delinieations and stratifications. This process of iterative variable selection and ADSP parameter tuning is time consuming but is important to provide managers more control of the final stand layer that is created. The area constraints chosen will vary based on management objectives and this may be the first variables that are experimented with to determine the spatial scale of the final forest segmentation. After this step, a series of maximum attribute differences and shape constraints can be examined to further constrain the outputs. The final step in this process is to experiment with different ADSP forms and parameterizations. This general framework for stand layer creation was followed for this work and can be seen in the ordering of models in Tables 2 and 3. Generally, exponential and positive parabolic forms will result in more stands. The linear ADSP form will result in slightly more stands created. The negative parabolic ADSP form will result in the fewest stands. The stepwise ADSP is more easily understood and was chosen in this case.

ADSPs provide finer grained control for managers in determining the final outcome of stand delineation. In some senses, this can be considered a model-based segmentation as it assumes a structure to the forest variability that varies with the stand area. Results show that this area-dependent approach performed better than a single non-area-dependent scale factor.

However, it is still difficult for a forest manager to choose the optimal parameters for this method and as a result many parameters must be examined before deciding on the best parameter set and area dependency to use. Because of this, more work is needed to develop a system that automates some of the variable selection and tuning.

Another future direction for this work may be to examine the area-dependent nature of the forest as a whole and how this might provide a model of forest variability. Woodcock (1988) proposed the use of variograms to model the spatial structure of remote sensing data but this assumes that the same variogram results apply to the full forest extent. Using this area-based approach, it may be possible to generate separate variogram models for different regions of a forest that correspond to the area-dependent differences observed in the stand creation process. This would have value in partitioning an ecosystem for process-modeling purposes particularly for future predictions (see Golinkoff and Running (2013)).

The ADRM method presented here provides several improvements to existing forest segmentation results. This method also builds upon much of the work that has already been done in this field. This method has incorporated many components of other object creation algorithms (e.g. – shape constraints, size constraints, iterative merging criteria relaxation, random seed regions, weighted attributes) but it has added to these methods in several key ways. First, by defining management compartments, assigning real-world attributes to each initial microstand, allowing managers to weight these attributes and define the forest stratification classes, this stand delineation and stratification method can be better controlled and understood by the managers who will actually use the results. Second, a new, scaled, within-stand variability metric has been proposed and used to provide a measure of how well a given stand delineation model run performs. This metric, when used in combination with a reference stand delineation and

stratification, can be used to select the optimal merging parameters and structure. Last, the use of an ADSP to control the region-merging has been shown to improve the outcome of forest stand delineation. The improvement is seen in the stratification accuracy when compared to a reference stand delineation and stratification, in the scaled, within-stand variability metric, and in the drastic improvement in sampling efficiency when using this approach to guide a new forest inventory.

Acknowledgements

This work would not have been possible without the support of The Conservation Fund and particularly Evan Smith, Jena Meradith, Scott Kelly, Lynsey Kelly, Madison Thompson, and Holly Newberger. Particular thanks to Scott Kelly and Evan Smith for their review of all of the methods and recommendations for improvements. Thanks also to Dr. Steve Running and Dr. Jon Graham at the University of Montana for their support and guidance for this work. This work was funded by The Conservation Fund (TCF) and occurred on TCF owned and managed forest lands. Additional funding for imagery and analysis was provided through a grant of the Bechtel Foundation. Finally, three anonymous reviewers are thanked for their comments which notably improved the quality of the original manuscript.

References

Van Aardt, J. A. N., R. H. Wynne, and R. G. Oderwald. 2006. Forest Volume and Biomass Estimation Using Small-Footprint Lidar-Distributional Parameters on a Per-Segment Basis. Forest Science 52:636–649.

Addink, E. A., F. M. B. Van Coillie, and S. M. De Jong. 2012. Introduction to the GEOBIA 2010 special issue: From pixels to geographic objects in remote sensing image analysis. International Journal of Applied Earth Observation and Geoinformation 15:1–6. doi: http://dx.doi.org/10.1016/j.jag.2011.12.001.

Akay, A., H. Oğuz, I. Karas, and K. Aruga. 2009. Using LiDAR technology in forestry activities. Environmental Monitoring and Assessment 151:117–124. doi: http://dx.doi.org/10.1007/s10661-008-0254-1.

Ali, S. S., P. M. Dare, and S. D. Jones. 2009. A Comparison of Pixel- and Object-Level Data Fusion Using Lidar and High-Resolution Imagery for Enhanced Classification. Pages 3–

17 *in* S. Jones and K. Reinke, editors. Innovations in Remote Sensing and Photogrammetry. Springer Berlin Heidelberg. Retrieved November 18, 2012, from http://www.springerlink.com.weblib.lib.umt.edu:8080/content/t14421u508822507/abstract/.

ArcMap 10.1. 2012. . ESRI (Environmental Systems Resource Institute, Redlands, CA, USA. Retrieved from http://www.esri.com/.

Asner, G. P., R. F. Hughes, J. Mascaro, A. L. Uowolo, D. E. Knapp, J. Jacobson, T. Kennedy-Bowdoin, and J. K. Clark. 2011. High-resolution carbon mapping on the million-hectare Island of Hawaii. Frontiers in Ecology and the Environment:110301094720075. doi: http://dx.doi.org/10.1890/100179.

Baatz, M., and A. Schäpe. 2000. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. Pages 12–23 *in* J. Strobl, editor.

Bell, J. F., and J. R. Dilworth. 2007. Log Scaling and Timber Cruising2007 Edition. Cascade Printing Company, Corvallis, OR.

Benz, U. C., P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing 58:239–258. doi: 10.1016/j.isprsjprs.2003.10.002.

Blaschke, T. 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65:2–16. doi: http://dx.doi.org/10.1016/j.isprsjprs.2009.06.004.

Borders, B. E., B. D. Shiver, and M. L. Clutter. 2005. Timber Inventory of Large Acreages Using Stratified Two-Stage List Sampling. Southern Journal of Applied Forestry 29:152–157.

CAR. 2007, September. Forest Project Protocol, Version 2.1. California Climate Action Registry. Retrieved from http://www.climateactionreserve.org/.

Castilla, G., G. J. Hay, and J. R. Ruiz-Gallardo. 2008. Size-constrained Region Merging (SCRM): An Automated Delineation Tool for Assisted Photointerpretation. Photogrammetric Engineering & Remote Sensing 74:409–419. doi: 0099-1112/08/7404–0409.

Cheng, H. D., X. H. Jiang, Y. Sun, and J. Wang. 2001. Color image segmentation: advances and prospects. Pattern Recognition 34:2259–2281. doi: http://dx.doi.org/10.1016/S0031-3203(00)00149-7.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37:35–46. doi: http://dx.doi.org/10.1016/0034-4257(91)90048-B.

Dey, V., Y. Zhang, and M. Zhong. 2010. A Review on Image Segmentation Techniques with Remote Sensing Perspective. Pages 31–42 *in* W. Wagner and B. Székely, editors. ISPRS TC VII Symposium – 100 Years ISPRS. IAPRS, Vienn. Retrieved from http://www.isprs.org/proceedings/XXXVIII/part7/a/pdf/31_XXXVIII-part7A.pdf.

Fisher, R. A., and F. Yates. 1948. Statistical tables for biological, agricultural and medical research., 3rd edition. Oliver & Boyd, London.

Fu, K. S., and J. K. Mui. 1981. A survey on image segmentation. Pattern Recognition 13:3–16. doi: 10.1016/0031-3203(81)90028-5.

GDAL Development Team. 2012. GDAL - Geospatial Data Abstraction Library. Open Source Geospatial Foundation. Retrieved from http://gdal.osgeo.org.

Golinkoff, J., M. Hanus, and J. Carah. 2011. The use of airborne laser scanning to develop a pixel-based stratification for a verified carbon offset project. Carbon Balance and Management 6:9. doi: 10.1186/1750-0680-6-9.

Golinkoff, J. S., and S. W. Running. 2013. A system to integrate multi-scaled data sources for improving terrestrial carbon balance estimates. Pages 259–287 *in* D. G. Brown, D. T. Robinson, N. H. F. French, and B. C. Reed, editors. Land Use and the Carbon Cycle: Advances in Integrated Science, Management, and Policy. Cambridge University Press, New York, NY, USA. Retrieved December 26, 2012, from http://www.cambridge.org/us/knowledge/isbn/item6962587/?site_locale=en_US.

Gonzalez, P., G. P. Asner, J. J. Battles, M. A. Lefsky, K. M. Waring, and M. Palace. 2010. Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. Remote Sensing of Environment 114:1561–1575. doi: http://dx.doi.org/10.1016/j.rse.2010.02.011.

Gonzalez, R. C., R. E. Woods, and S. L. Eddins. 2009. Digital Image Processing Using MATLAB, 2nd edition. Prentice Hall, Upper Saddle River, NJ.

Hall, R. J., R. S. Skakun, E. J. Arsenault, and B. S. Case. 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. Forest Ecology and Management 225:378–390. doi: http://dx.doi.org/10.1016/j.foreco.2006.01.014.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition. Springer. Retrieved from http://www-stat.stanford.edu/~hastie/Papers/ESLII.pdf.

Hay, G. J., and G. Castilla. 2008. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. Pages 75–89 *in* T. Blaschke, S. Lang, and G. J. Hay, editors. Object-Based Image Analysis. Springer Berlin Heidelberg. Retrieved December 2, 2012, from http://link.springer.com/chapter/10.1007/978-3-540-77058-9_4.

Hay, G. J., G. Castilla, M. A. Wulder, and J. R. Ruiz. 2005. An automated object-based approach for the multiscale image segmentation of forest scenes. International Journal of Applied Earth Observation and Geoinformation 7:339–359. doi: http://dx.doi.org/10.1016/j.jag.2005.06.005.

Haywood, A., and C. Stone. 2011. Semi-automating the stand delineation process in mapping natural eucalypt forests. Australian Forestry 74:13+.

Heinsch, F. A., M. Zhao, S. W. Running, J. S. Kimball, R. R. Nemani, K. J. Davis, P. V. Bolstad, B. D. Cook, A. R. Desai, D. M. Ricciuto, B. E. Law, W. C. Oechel, H. Kwon, H. Luo, S. C. Wofsy, A. L. Dunn, J. W. Munger, D. D. Baldocchi, L. Xu, D. Y. Hollinger, A. D. Richardson, P. C. Stoy, M. B. S. Siqueira, R. K. Monson, S. P. Burns, and L. B. Flanagan. 2006. Evaluation of Remote Sensing Based Terrestrial Productivity from MODIS Using Regional Eddy Flux Network Observations. IEEE Transactions on Geoscience and Remote Sensing 44:1908–1925. doi: http://dx.doi.org/10.1109/TGRS.2005.853936.

Hofmann, P., T. Blaschke, and J. Strobl. 2011. Quantifying the robustness of fuzzy rule sets in object-based image analysis. International Journal of Remote Sensing 32:7359–7381. doi: http://dx.doi.org/10.1080/01431161.2010.523727.

Ioki, K., J. Imanishi, T. Sasaki, Y. Morimoto, and K. Kitada. 2009. Estimating stand volume in broad-leaved forest using discrete-return LiDAR: plot-based approach. Landscape and Ecological Engineering 6:29–36. doi: http://dx.doi.org/10.1007/s11355-009-0077-4.

Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2004. Comprehensive Database of Diameter-based Biomass Regressions for North American Tree Species. Northeastern Research Station.

Johansen, K., R. Bartolo, and S. Phinn. 2010. SPECIAL FEATURE – Geographic Object-Based Image Analysis. Journal of Spatial Science 55:3–7. doi: http://dx.doi.org/10.1080/14498596.2010.494653.

Ke, Y., L. J. Quackenbush, and J. Im. 2010. Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. Remote Sensing of Environment 114:1141–1154. doi: http://dx.doi.org/10.1016/j.rse.2010.01.002.

Latifi, H., A. Nothdurft, and B. Koch. 2010. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. Forestry 83:395 –407. doi: http://dx.doi.org/10.1093/forestry/cpq022.

Leckie, D. G., F. A. Gougeon, N. Walsworth, and D. Paradine. 2003. Stand delineation and composition estimation using semi-automated individual tree crown analysis. Remote Sensing of Environment 85:355–369. doi: http://dx.doi.org/10.1016/S0034-4257(03)00013-0.

Mardia, K. V., J. T. Kent, and J. M. Bibby. 1979. Multivariate analysis. Academic Press, London; New York.

Marpu, P. R., M. Neubert, H. Herold, and I. Niemeyer. 2010. Enhanced evaluation of image segmentation results. Journal of Spatial Science 55:55–68. doi: http://dx.doi.org/10.1080/14498596.2010.487850.

Maselli, F., M. Chiesi, M. Moriondo, L. Fibbi, M. Bindi, and S. W. Running. 2009. Modelling the forest carbon budget of a Mediterranean region through the integration of ground and satellite data. Ecological Modelling 220:330–342. doi: http://dx.doi.org/10.1016/j.ecolmodel.2008.10.002.

MATLAB Image Processing Toolbox. 2011. . Mathworks, Natick, Massachusetts, U.S.A. Retrieved from http://www.mathworks.com/products/matlab/?s_cid=global_nav.

Mildrexler, D. J., M. Zhao, and S. W. Running. 2009. Testing a MODIS Global Disturbance Index across North America. Remote Sensing of Environment 113:2103–2117. doi: http://dx.doi.org/10.1016/j.rse.2009.05.016.

Milligan, G. W., and M. C. Cooper. 1988. A study of standardization of variables in cluster analysis. Journal of Classification 5:181–204. doi: http://dx.doi.org/10.1007/BF01897163.

Möller, M., L. Lymburner, and M. Volk. 2007. The comparison index: A tool for assessing the accuracy of image segmentation. International Journal of Applied Earth Observation and Geoinformation 9:311–321. doi: http://dx.doi.org/10.1016/j.jag.2006.10.002.

Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80:88–99. doi: http://dx.doi.org/10.1016/S0034-4257(01)00290-5.

Nilsson, M. 1996. Estimation of tree heights and stand volume using an airborne lidar system. Remote Sensing of Environment 56:1–7. doi: http://dx.doi.org/10.1016/0034-4257(95)00224-3.

Openshaw, S. 1984. The Modifiable Areal Unit Problem. CATMOG - Concepts and Techniques in Modern Geography 38:1–41.

Pal, N. R., and S. K. Pal. 1993. A review on image segmentation techniques. Pattern Recognition 26:1277–1294. doi: http://dx.doi.org/10.1016/0031-3203(93)90135-J.

Pascual, C., A. García-Abril, L. G. García-Montero, S. Martín-Fernández, and W. B. Cohen. 2008. Object-based semi-automatic approach for forest structure characterization using lidar data in heterogeneous Pinus sylvestris stands. Forest Ecology and Management 255:3677–3685. doi: http://dx.doi.org/10.1016/j.foreco.2008.02.055.

Popescu, S. C., R. H. Wynne, and R. F. Nelson. 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Canadian Journal of Remote Sensing 29:564–577. doi: http://dx.doi.org/10.5589/m03-027.

Potter, C., P. Gross, V. Genovese, and M.-L. Smith. 2007a. Net primary productivity of forest stands in New Hampshire estimated from Landsat and MODIS satellite data. Carbon Balance and Management 2. doi: http://dx.doi.org/10.1186/1750-0680-2-9.

Potter, C., S. Klooster, A. Huete, and V. Genovese. 2007b. Terrestrial carbon sinks for the United States predicted from MODIS satellite data and ecosystem modeling. Earth Interactions 11:1–21. doi: http://dx.doi.org/10.1175/EI228.1.

Riggins, J., J. Tullis, and F. Stephen. 2009. Per-segment Aboveground Forest Biomass Estimation Using LIDAR-Derived Height Percentile Statistics. GIScience & Remote Sensing 46:232–248. doi: http://dx.doi.org/10.2747/1548-1603.46.2.232.

Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. C. Reeves, and H. Hashimoto. 2004. A continuous satellite-derived measure of global terrestrial primary production. BioScience 54:547–560. doi: http://dx.doi.org/10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2.

Sasaki, T., J. Imanishi, K. Ioki, Y. Morimoto, and K. Kitada. 2012. Object-based classification of land cover and tree species by integrating airborne LiDAR and high spatial resolution imagery data. Landscape and Ecological Engineering 8:157–171. doi: http://dx.doi.org/10.1007/s11355-011-0158-z.

Shiver, B. D., and B. E. Borders. 1996. Sampling techniques for forest resource inventory. John Wiley & Sons, Inc., New York, NY.

Song, C., M. B. Dickinson, L. Su, S. Zhang, and D. Yaussey. 2010. Estimating average tree crown size using spatial information from Ikonos and QuickBird images: Across-sensor and across-site comparisons. Remote Sensing of Environment 114:1099–1107. doi: http://dx.doi.org/10.1016/j.rse.2009.12.022.

Sullivan, A. A., R. J. McGaughey, H.-E. Andersen, and P. Schiess. 2009. Object-Oriented Classification of Forest Structure from Light Detection and Ranging Data for Stand Mapping. Western Journal of Applied Forestry 24:198–204.

Thompson, S. K. 2002. Sampling, 2nd edition. John Wiley & Sons, Inc., New York, NY.

Turner, M. G., R. H. Gardner, and R. V. O'Neil. 2001. Landscape Ecology in Theory and Practice: Pattern and Process. Springer, New York, NY.

Wang, Z., J. R. Jensen, and J. Im. 2010. An automatic region-based image segmentation algorithm for remote sensing applications. Environmental Modelling & Software 25:1149–1165. doi: http://dx.doi.org/10.1016/j.envsoft.2010.03.019.

Woodcock, C. E., A. H. Strahler, and D. L. B. Jupp. 1988. The use of variograms in remote sensing: II. Real digital images. Remote Sensing of Environment 25:349–379. doi: http://dx.doi.org/10.1016/0034-4257(88)90109-5.

Wulder, M. A., and D. Seemann. 2003. Forest inventory height update through the integration of lidar data with segmented Landsat imagery. Canadian Journal of Remote Sensing 29:536–543. doi: http://dx.doi.org/10.5589/m03-032.

Wulder, M. A., J. C. White, R. F. Nelson, E. Næsset, H. O. Ørka, N. C. Coops, T. Hilker, C. W. Bater, and T. Gobakken. 2012. Lidar sampling for large-area forest characterization: A review. Remote Sensing of Environment 121:196–209. doi: http://dx.doi.org/10.1016/j.rse.2012.02.001.

Zhang, X., and S. Kondragunta. 2006. Estimating forest biomass in the USA using generalized allometric models and MODIS land products. Geophysical Research Letters 33:1–5. doi: http://dx.doi.org/10.1029/2006GL025879.

Summary and Conclusions

The work discussed above details several approaches to estimate current forest stocks and fluxes. Chapter one summarizes the theoretical basis of the BiomeBGC process model and how it compares to other model types. Chapter two uses this model along with field inventory data to make predictions about forest productivity for a forest in northern California. Chapter three incorporates high-resolution optical remote sensing data and LiDAR data with field inventory data to generate a pixel-based stratification of a forest. This work served as the basis for a new forest inventory and will be the basis of future forest modeling. The final chapter improved upon this work by instead developing a system to create stands and strata from remote sensing and field data. This work also is the basis for a new forest inventory and future modeling but is a more easily understandable, fits within more traditional conceptions of how forest stands are defined, and creates a more efficient platform to model future forest growth and management.

All of the work detailed above is done to help demonstrate efficient and accurate methods to estimate the carbon stocks and fluxes of forests so that they may be more easily incorporated into climate change mitigation efforts. Although the pace of policy developments to address climate change and deforestation at a global scale is slow, there are already several new developments that place a value on the carbon stored in forests – most notably the California cap and trade legislation. This state level policy, along with several voluntary forest carbon offset standards, has created a need for reliable data about current and future forest conditions across large spatial scales. Policy makers are generally more interested in large scale estimates across states, nations, and the globe. Conversely, landowners are often more interested in estimates of their

individual ownership or region. Because of these different focuses, scaleable and flexible methods to estimate stocks and model future conditions are needed.

The work presented in this dissertation is motivated by this and explores scaleable methods to estimate current and future forest conditions. The Biome-BGC model was first examined as a candidate for this type of work. It was calibrated to match current conditions at a forest in northern California and run into the future to estimate Net Primary Productivity. However, after extensive work using a inversion software package (PEST: Model-Independent Parameter Estimation and Uncertainty Analysis 2013), it was determined that Biome-BGC may be a poorly constrained inversion problem. That is, there are more parameters than observations and therefore there are many possible solutions that would satisfy the inversion. Furthermore, Biome-BGC is not optimally suited to model extensively managed forest systems as it is designed to estimate steady-state ecosystem dynamics. New research has attempted to bridge this gap as there is still a need for a process model like Biome-BGC when trying to answer questions about future ecosystem state given different future climate scenarios.

Because of the limitations of Biome-BGC, a different approach was explored that uses the high resolution remote sensing data and LiDAR data. These new approaches also attempt to stratify a landscape into a set of similar types and by so doing allow for more efficient modeling work to be done. The first approach discussed in Chapter 3 used a pixel-based framework and classified 400 m2 pixels into a set of 40 strata. Although this method showed great promise in reducing inventory cost while improving inventory accuracy, it is difficult to use for forest management planning. Additionally, pixel-based classification efforts have been shown to be less effective at partitioning landscape variability than object-based methods – especially when

the objects of interest are larger than the size of the pixels. The limitations of a pixel-based approach spurred the development of the object-based approach described in Chapter 4.

The work outlined in these chapters can be considered to fall along a spectrum of methods to best estimate and model forest conditions. The Biome-BGC model was first explored to understand if this model could be applied at an ownership scale to accurately estimate current carbon stocks and fluxes. This model was then applied using a pixel-based approach where each pixel had unique soil and climate driving variables. Because of the processing intensive nature of this approach across larger scales and the difficulty in calibrating this model, new methods were needed. The next step was to develop a method using a pixel-based approach that leveraged the data found in both remote sensing products and ground inventory. After developing this method, an object-based approach was explored as this provides a more flexible, user-friendly, and powerful way to segment a forest landscape.

At each stage in this process, the products from the prior stage can be inserted into the current stage and better results obtained. For example, Biome-BGC would optimally be run at the strata level instead of the individual pixel level using either a pixel-based or object-based stratification. Furthermore, the pixel-based stratification results can be used as an input layer for the object-based region merging algorithm and will result in a new map of similar regions across a forest. The use of a pixel layer as the input to an object-based segmentation algorithm is a powerful and scaleable extension of this work. Large-scale, pixel-based, remote sensing products like MODIS or Landsat imagery can be used to generate landscape segments at the state, regional, or national scale. With these new tools, more appropriate and efficient forest modeling can occur and the results of this modeling can be used to inform individual landowner decisions or forest policies.

References

PEST: Model-Independent Parameter Estimation and Uncertainty Analysis. 2013. . Watermark Numerical Computing. Retrieved from http://www.pesthomepage.org/Home.php.