

Modelling forest cover attributes as continuous variables in a regional context with Thematic Mapper data

W. B. COHEN†‡, T. K. MAIERSPERGER§, T. A. SPIES† and
 D. R. OETTER§

† Forestry Sciences Laboratory, Pacific Northwest Research Station, USDA
 Forest Service, 3200 SW Jefferson Way, Corvallis, OR 97331, USA

§ Department of Forest Science, Forestry Sciences Laboratory 020, Oregon
 State University, Corvallis, OR 97331, USA

(Received 9 March 1999; in final form 22 March 2000)

Abstract. We modelled forest vegetation attributes as continuous variables across western Oregon using a multi-image mosaic of Thematic Mapper (TM) data. Four specific attributes were modelled using regression analysis: percent green vegetation cover, percent conifer cover, conifer crown diameter, and conifer stand age. Reference data for the cover and diameter attributes were derived from airphotos, and existing agency polygon databases were used for stand age. We developed and applied a new method for regional mapping called *applied radiometric normalization*. The method involved development of a set of models for a centrally located ‘source’ scene which were then extended to ‘destination’ scenes (neighboring scenes in the TM mosaic). Use of airphotos and existing digital databases in combination with applied radiometric normalization translates to a cost-effective procedure for regional mapping with TM data. Modelling forest attributes as continuous variables enables creation of a flexible forest cover information base, containing important fundamental building blocks for a variety of related classification schemes.

1. Introduction

Management of forest resources occurs at a number of spatial scales. For example, at a stand level, variations in tree sizes contribute to structural diversity (Spies and Franklin 1991). Over a landscape, a mixed distribution of stand ages is essential for balancing wood production and maintenance of habitat for endangered and threatened species of wildlife (Lehmkuhl *et al.* 1991). Regionally, policies are established to protect diminishing wildlands and threatened wildlife populations (Tuchmann *et al.* 1996). At continental to global scales, strategic recommendations are made concerning the effects of forest management on sequestration of atmospheric carbon and climate change (Krankina *et al.* 1998). To adequately manage forest resources at landscape scales and above, maps of forest cover are needed. Remote sensing has played, and will continue to play, a vital role for mapping forest resources. In this study, we focus on the use of Landsat TM data to map land cover at a regional level. Our purpose is to provide spatially detailed information that can

‡ Corresponding author. E-mail: cohen@fsl.orst.edu

be used in a variety of modelling studies addressing regionally important issues, such as the preservation of fisheries and wildlife habitat (Gregory *et al.* 1998, Hemstrom *et al.* 1998) and carbon flux (Cohen *et al.* 1996). The overall goal is to provide forest cover information across the 5 million ha forest land base in western Oregon that is consistent across all land ownership categories (e.g. private non-industrial, private industrial, public), flexible in terms of definition of cover classes, and both repeatable and cost-effective to produce.

Mapping land cover with Landsat data is as old as the Landsat program itself (see Goward and Williams 1997 and Landgrebe 1997). Most studies have been conducted at a relatively local level to determine the efficacy of the technology for providing specific types of land cover information (e.g. Franklin and Moulton 1990, Apan 1997). By comparison, published regional land cover mapping efforts (e.g. $\geq 50\,000\text{ km}^2$) have only sparsely appeared in the literature (Fuller *et al.* 1994, Homer *et al.* 1997). Although use of even a single Landsat image is accompanied by a host of attendant issues such as geometric rectification, a need for reference data (commonly referred to as 'ground truth'), and accuracy assessment, use of multiple images significantly compounds these issues. The most prevalent use of multiple Landsat images is for change detection, in which a single path/row of data from two separate dates are compared (Singh 1989). Some studies have used seasonal images of a single path/row for land cover classification (e.g. Grignetti *et al.* 1997). For these temporal applications, recommended protocol involves radiometric normalization, multi-temporal reference data, and a multi-temporal land cover or cover change algorithm (Coppin and Bauer 1996).

Regional-level mapping has similar challenges as those involving single path/row multi-date imagery plus some that are relatively unique to this particular application. Images to be spatially mosaicked must be co-referenced geometrically and radiometrically normalized to permit use of a single set of mapping algorithms across image boundaries or otherwise merged to minimize mosaic seamlines. A more spatially extensive reference database must be compiled. Depending on the size and complexity of the land area under study, the level of sophistication in the mapping algorithm may need to be greater to properly address the increased variability within a land cover type and among the number of land cover types encountered. When developing a regional land cover map one must recognize a potentially wider audience and user-base, as the map product is likely to be considered for incorporation into a larger number of resource assessments than maps developed over localized areas.

1.1. Objectives

We sought to develop a cover map for the five million hectares of forest of western Oregon with the following characteristics.

- (1) *Cost-effective production.* Given the high rates of forest harvest and subsequent regrowth in this intensively managed forest region (Cohen *et al.* 1998) and the politically sensitive nature of regional forest management practices (Tuchmann *et al.* 1996), there is a need to repeat the mapping process every 5 to 10 years (Mulder *et al.* 1999). Thus, the methods for map production should be relatively cost-effective, dictating the use of existing reference data rather than initiating new field surveys specifically to support the mapping effort.
- (2) *Flexible information base.* Potential users of forest maps often have unique

and specific criteria in terms of information content and forest attribute class definitions (Morrison *et al.* 1991, Congalton *et al.* 1993). Consequently, our aim was to provide a flexible information base that permits maximum utility within the context of widely varying subsequent studies.

- (3) *Seamless across image boundaries.* Although seamlines in a multi-image mosaic can have little to do with map accuracy, they detract from overall map quality (Wadsworth *et al.* 1997). Our goal was to produce a seamless representation of forest cover across the numerous image boundaries within the study area.

1.1.1. *Cost-effective mapping*

This first objective is facilitated by using large-scale aerial photographs and existing digital stand age data to develop the reference database. Airphotos are commonly used in collecting reference data for training and testing algorithms for land cover mapping with Landsat data, but their use is primarily limited to general assistance in interpreting images, in locating field plots where the actual reference data are collected, or in extending the utility of field plot data in a double sampling context (Bauer *et al.* 1994, Vogelmann *et al.* 1998). Cost-effectiveness is also greatly aided by use of a process we call *applied radiometric normalization*, which is described more fully in §2.4. This process enables the bulk of an assembled reference database to be used in characterizing the errors in final map products (testing map accuracy), as opposed to the more common practice of using most reference data for training the mapping algorithm (Knick *et al.* 1997, Peddle *et al.* 1997).

1.1.2. *Flexible forest vegetation information base*

Our experience suggests a need for forest cover information that is not fixed into a pre-specified set of classes, but rather a system for cover estimation that permits a user-specified set of classes for a given application in a subsequent modelling study. For example, if an objective is to map forest age we might define a set of classes that can be predicted with reasonable accuracy, say < 80 years, 80–200 years, and > 200 years, as was presented in Cohen *et al.* (1995). Some prospective user of this map might need an additional class, or perhaps this user needs just two classes, but with the break at 40 years. If the underlying model to estimate forest age is *continuous*, there is flexibility to redefine the age class data layer in the manner needed. Additionally, over landscapes containing significant proportions of naturally-regenerated forest systems many forest attributes are continuous rather than discrete (e.g. tree size). As such, modelling forest attributes as continuous variables can result in maps that more accurately represent the true spatial frequency and contrast in forest cover.

Algorithms for estimating vegetation attributes as continuous variables are not new. For example, mixture models are used to map fractions of basic vegetation cover building blocks such as photosynthetic and non-photosynthetic vegetation, soil, and shadow fractions (Smith *et al.* 1990, Mustard 1993). Numerous studies have used regression modelling approaches to estimate such characteristics as leaf area index, tree size, and biomass (Butera 1986, Peterson *et al.* 1987).

We use single attribute regression modelling, following our initial work in this area (Cohen and Spies 1992, Cohen *et al.* 1995, Maiersperger *et al.* 2001). The forest attributes we model are: (1) percent green vegetation cover, (2) percent conifer cover, (3) visible tree crown diameter, and (4) local stand age. The first variable is a simple estimate of total vegetation cover that is photosynthetically active, and the

second is a measure of the degree to which the primary morphological vegetation type of the Pacific Northwest (PNW) region of the United States actually occupies this highly disturbed landscape. The third and fourth variables are commonly used measures of forest structure and condition. For the purpose of developing a reference database, the first three variables are directly interpretable on airphotos. Use of the adjective 'visible' for crown diameter is important because photointerpreted conifer crown size is commonly 15% to 20% less than true crown diameter (Dillworth 1956, Paine 1981). Stand age data are available from various forest management agencies as polygon attributes in a geographic information system (GIS) database. The term 'local' implies that stand age only exists at a grain size larger than a single TM pixel (i.e. it is a forest patch-level construct), but that we are modelling it here as a spatially continuous phenomenon at the grain size of TM data.

1.1.3. *Seamless cover mapping*

Use of multiple Landsat images requires consideration of radiometric normalization (Hall *et al.* 1991b). Although normalization is most commonly done for change detection using multitemporal data of a single ground scene, variants of this procedure has been used to minimize seamlines in the construction of single-date, aerial multi-image mosaics (Neale and Crowther 1994, Pickup *et al.* 1995). Landsat TM data have also been mosaicked using a radiometric normalization procedure. However, due to the path/row structure of Landsat image acquisitions, ephemeral cloudiness, and other constraints, images from several different dates within a year or two period must normally be used to construct a TM mosaic to represent a 'single' time period (e.g. Homer *et al.* 1997).

Much of the rich body of radiometric normalization literature is concerned with absolute calibration of individual images, whereby sensor characteristics and atmospheric and illumination angle conditions are explicitly modelled or otherwise accounted for (Chavez 1989, Vermote and Kaufman 1995). The implication is that if multiple images acquired under differing atmospheric and illumination angle conditions are absolutely calibrated, they will inherently be normalized with respect to each other. Our own experience and that of others (C. Woodcock, 1999, personal communication) is that this is not necessarily the case. Furthermore, for most users of image data, ground and atmospheric data required for absolute calibration to be effective are often absent or difficult and costly to obtain (Eckhardt and Verdin 1990, Hall *et al.* 1991b, Coppin and Bauer 1996). An alternative approach, relative calibration, is a procedure whereby individual scenes are radiometrically normalized with respect to each other, using little or no ancillary data (Schott *et al.* 1988, Hall *et al.* 1991b, Henebry and Su 1993, Elvidge *et al.* 1995, Narayana *et al.* 1995). For the construction of image mosaics, relative calibration is the common choice (Neale and Crowther 1994, Pickup *et al.* 1995, Homer *et al.* 1997). Heretofore, radiometric normalization has involved explicit radiometric adjustment of image data prior to translation of those data into land cover or cover change information (e.g. Vogelmann 1988, Hall *et al.* 1991a, Jensen *et al.* 1995). Applied radiometric normalization (described more fully below), not only has benefits in cost-effectiveness, but also is a more direct attempt to minimize seams across image boundaries by normalizing the image data in terms of the cover information of interest.

2. Methods

2.1. *Study area*

Western Oregon is a topographically complex area (figure 1(a)) that, outside of the major valley systems, is predominantly covered with forest vegetation. Franklin

and Dyrness (1988) provide the most comprehensive ecological and climatic description of the area, which is among the more diverse regions of North America in its environment and vegetation. The physiographic and geographic provinces included in this study are the Coast Ranges, the Western Cascades, and the Klamath Mountains. The Willamette Valley Province is primarily agricultural; thus forest attributes are not extensively modelled there. Climatically the study area varies greatly, with annual precipitation ranging from as low as 60 cm in the eastern portion of the Klamath Mountains to over 300 cm in the northern Coast Range. Mean monthly minimum temperatures range from -5°C in the eastern parts of the Western Cascades and Klamath Mountains to 5°C in the western part of the Klamath Mountains. Mean monthly maximum temperatures range from 31°C in the southern Klamath Mountains to 20°C along the coast from the northern Coast Ranges to the southern Klamath Mountains. The Klamath Mountains Province is the most climatically diverse portion of the study region. These forest zones also differ in fire and management history which create additional diversity in canopy structure.

Several different major forest vegetation zones exist in the study area (Franklin and Dyrness 1988). The Sitka Spruce Zone is only a few kilometers in width, occurring along the full length of the Oregon coast. The Western Hemlock Zone is the most extensive vegetation zone in the study area, existing in two major segments: the majority of both the Coast Ranges and the Western Cascades Provinces. The Mixed Conifer/Mixed Evergreen Zone occurs primarily in the Klamath Mountains Province, the Subalpine Forest Zone occurs in the high elevation portion of the Western Cascades Province, and in the far southeast part of the study area is a small portion of the Ponderosa Pine Zone. Across these forest zones, represented are 22 coniferous tree species in 9 different genera and numerous hardwood tree species in numerous genera. Dominant trees are typically conifers, except in riparian areas and dry valley margins where hardwood trees can dominate. Throughout the study area, many different herbaceous and both evergreen and deciduous shrub species exist that may be visible to an aerial or satellite sensor.

2.2. Image and elevation data

The images used in this study are all from the same year (1988), but they vary across the full growing season from April to August (figure 1(a), table 1). Each individual image was transformed into Tasseled Cap brightness, greenness, and wetness indices (Crist *et al.* 1986) prior to use and interpretation. This was done to compress the six bands of spectral data into the three physically meaningful spectral axes (Crist *et al.* 1984). Cohen *et al.* (1995) and Collins and Woodcock (1994) have demonstrated the utility of these axes for coniferous forest systems. As only three color layers can be displayed on a computer screen at a single time, these axes have the additional advantage of capturing most of the spectral variability in TM data for visual display (Cohen *et al.* 1995).

Our earlier work in western Oregon indicates that changes in brightness are associated with soil and litter color, varying proportions of vegetation cover, and of hardwood versus conifer cover (Cohen *et al.* 1995). Greenness is associated with changes in vegetation cover (similar to NDVI) and to some extent the proportion of conifer versus hardwood. Wetness is the primary variable for estimating forest structure in closed-canopy stands (Cohen and Spies 1992, Cohen *et al.* 1995, Collins and Woodcock 1996). In both Cohen and Spies (1992) and Cohen *et al.* (1995) we provide a mix of demonstration and hypothesis for a close mechanistic relationship

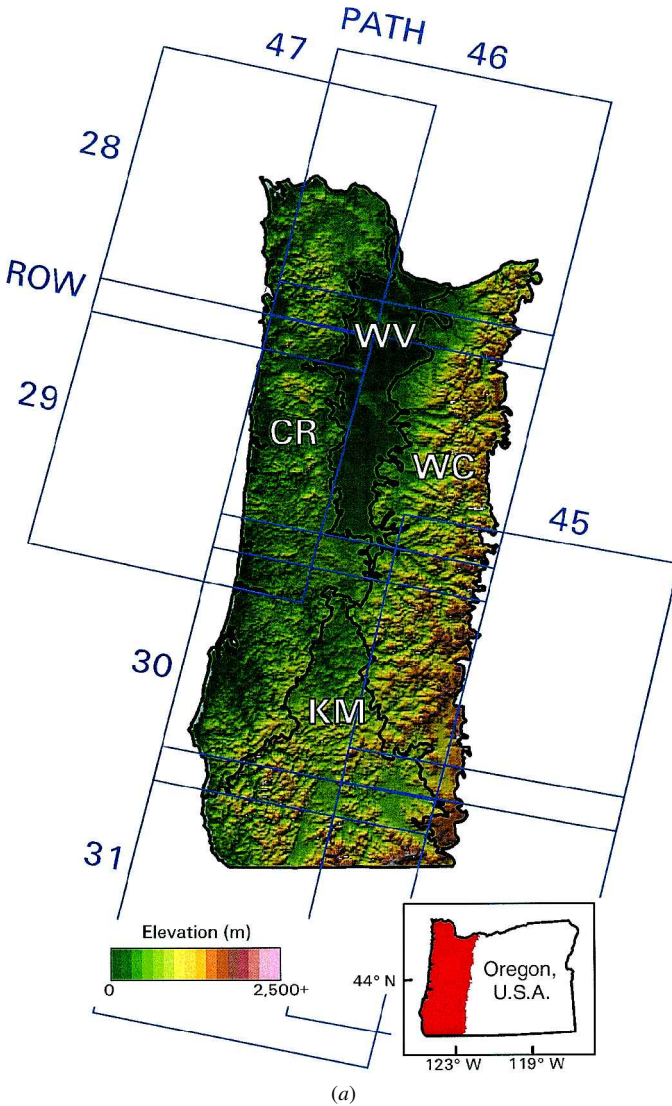


Figure 1. Study area characteristics and related information: (a) location, elevation, province boundaries (CR = Coast Ranges, WC = Western Cascades, KM = Klamath Mountains, WV = Willamette Valley), and Landsat TM scenes used, by path and row; (b) ownership patterns and distribution of reference plots (cover plot symbols are smaller to allow for spatial overlap). Total study area equals 7 275 713 ha.

between wetness and stand structure. In both of these papers we focused our interpretations of wetness almost exclusively on closed conifer forest conditions. Here, we extend our analysis of wetness into a much fuller range of forest conditions present in the PNW region.

Digital elevation model (DEM) data were used along with brightness, greenness, and wetness as part of the independent variable dataset in regression modelling. The data used are the 1:250 000 series available from the United States Geological Survey.

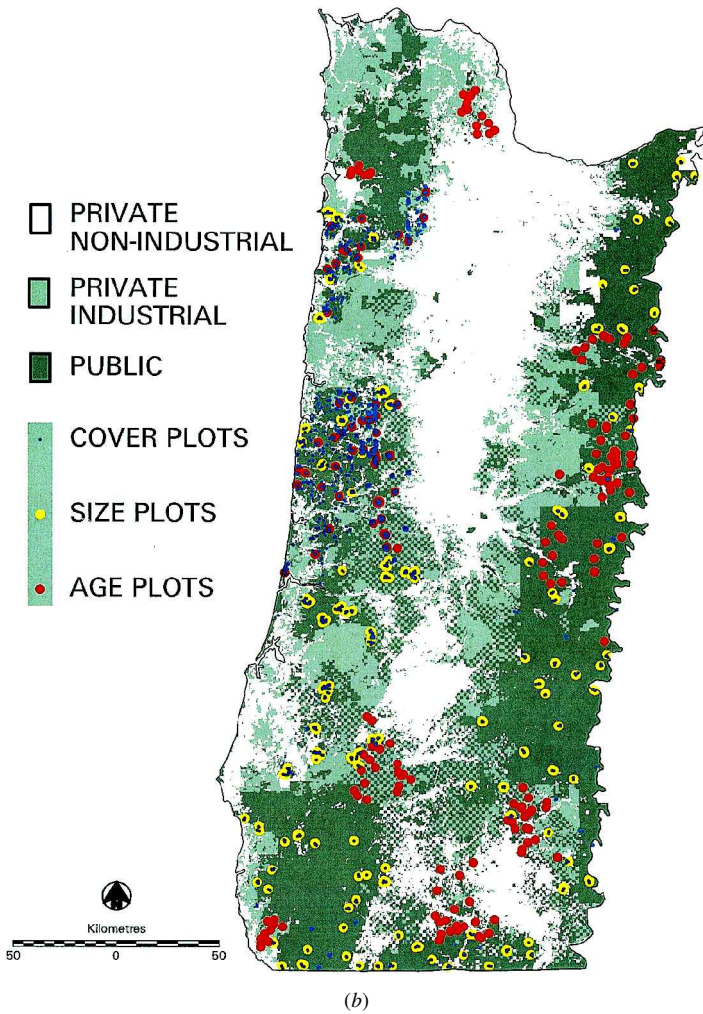


Table 1. Landsat imagery used.

Path/Row	Date of acquisition	Context
45/30	8 August 1988	Destination
45/31	8 August 1988	Destination
46/28	31 August 1988	Source
46/29	31 August 1988	Source
46/30	30 July 1988	Destination/secondary source
46/31	9 April 1988	Secondary destination
47/28	19 June 1988	Destination
47/29	21 July 1988	Destination

From these, we derived cosine of the effective solar incidence angle values ($\cos i$) using the standard physical equation given in Smith *et al.* (1980).

2.3. Reference data

2.3.1. Cover and visible crown diameter

Standard 9' by 9' colour aerial photographs varying in nominal scale from 1:12000 to 1:18000 were used. All photos were acquired within two years of the 1988 TM imagery, and were obtained from federal and state field offices in western Oregon. Photo plots were interpreted for proportions of several scene components including conifer tree canopy, hardwood tree canopy, grass/forb/herb/shrub canopy, tree gap shadow, and open/litter. Additionally, visible conifer crown diameter was interpreted for a sample of plots that were dominated by closed-canopy conifer trees. Photos were sampled and interpreted as follows.

- Photo triplets were selected at random in geographic strata defined by photo missions, with the intent being to sample an adequate number of geographically dispersed missions (figure 1(b)). Within a mission we sought a good representation of stand structure and composition variability.
- For each photo triplet, the center photo's effective area was identified and a four-point grid overlain on that area. Each point represented a potential plot center, with the distribution of plot centers being near the center of each quadrant of the effective area. Photo plot size was approximately 2 ha and a plot was accepted if it did not include high contrast boundaries such as recent clearcut and adjacent forest, roads, streams, or other non-vegetation components, or fall within a deep topographic shadow. Since we were interested in modelling mixtures of hardwood and conifer tree crowns, we accepted mixed stand conditions but sought a fine-grained mix of hardwood and conifer crowns rather than a coarse-grained mix or edge situation. To avoid pseudo-replication associated with spatial autocorrelation, a plot was rejected if one of the other three plots in the effective area was of the same stand condition. Total number of plots varied by stand attribute modelled, with over 1000 plots interpreted for green vegetation and conifer cover and nearly 250 for size. Plot locations were digitized in the TM data after ocularly locating them in the imagery.
- Plots were interpreted using a mirror stereoscope with 3× magnification. To determine percent cover of each scene component within a plot, a 20-point sampling grid was overlain and the number of occurrences of each cover type tallied. This permitted proportional estimates at 5% intervals. Crown size was estimated at each relevant grid-point with a crown diameter scale in thousands of an inch (Avery 1978) for a sample of all plots dominated by conifer crowns. Visible crown diameter can be greatly affected by local variations in photo scale, which occur when terrain elevation is variable as was the case in our mountainous study area. To calculate local photo scale the camera focal length and flying height (obtained from flight log data) and mean plot elevation derived from a DEM were combined using a standard formula (Paine 1981). The crown diameter scale measurement and the local scale calculation were used to convert each tree crown measurement into a visible crown diameter estimate (Paine 1981). Then, mean visible crown diameter was calculated for each plot.

2.3.2. Local stand age

Stand age is not an attribute that can be directly photointerpreted. Thus for this variable we used digital data files available through the United States Forest Service

(USFS) and Bureau of Land Management (BLM). These files consisted of vectorized polygons representing different forest stands/management units, ranging in size from 2 ha to 20 ha, each attributed with a stand age value. USFS and BLM holdings are well distributed throughout the study area, enabling us to extract a representative sample both geographically (figure 1(b)) and across the stand age spectrum. The age attribute in these digital files represents mean stand age of dominant and codominant trees. To avoid edges along stand boundaries, selected polygons were reduced in size by eliminating a 50 m buffer at stand edges.

2.4. Applied radiometric normalization

Applied radiometric normalization is a procedure developed to minimize boundaries in a map of vegetation cover that requires spatial mosaicking of more than one image to construct the map. Unlike more standard relative normalization procedures, applied normalization does not involve cross-calibration of the radiometric data in adjoining images (or in the more common case: a temporal image sequence of the same ground scene (e.g. Hall *et al* 1991a)). Rather, a centrally located (*source*) image is first translated into vegetation cover information derived from statistical modelling of the relationship between image spectral data and vegetation cover reference data (figure 2). Then, the vegetation cover information from the source scene is directly 'extended' to neighboring (*destination*) scenes using a new set of statistical models. These destination models are similar to the source models in terms of the spectral variables used, but the model coefficients differ according to the degree of radiometric difference between adjacent scenes.

As for any mapping project, one using the applied radiometric normalization procedure requires two sets of reference data: one for model building and an independent one for model testing or accuracy assessment. With applied normalization however, this reference data split is relevant only for the source scene. For the destination scenes, *all* reference data are used to test model performance. This provides a major cost-savings, since only a portion of the reference data otherwise needed to build and test models for all scenes in the spatial mosaic are required. To build and test models, mean pixel-level spectral responses were extracted from the image data for all plots. Likewise, mean $\cos i$ for each plot were extracted from the DEM data. For the source models, the reference data were stratified into model building (2/3) and model testing (1/3) sets. All reference data for the destination scenes were used to characterize errors in the destination models and maps.

2.4.1. Source models

When selecting an image to represent the source scene, several factors must be considered. The image should be of high radiometric quality (e.g. minimal atmospheric effects); it should have been acquired at a season that is consistent with the destination images; and to the extent possible, it should be centrally located to minimize the number of model extensions required to cover the area of interest. If several images were acquired during a single pass of the satellite, then this image swath is likely an excellent candidate for a source scene, assuming other considerations do not limit its use as such. For this study, the images from Path 46, Rows 28 and 29 were selected as the source scene for model building (table 1). As these images were acquired on the same date through a clear atmosphere they can be treated as a single radiometric entity, and thus, a single scene. They also constitute a larger

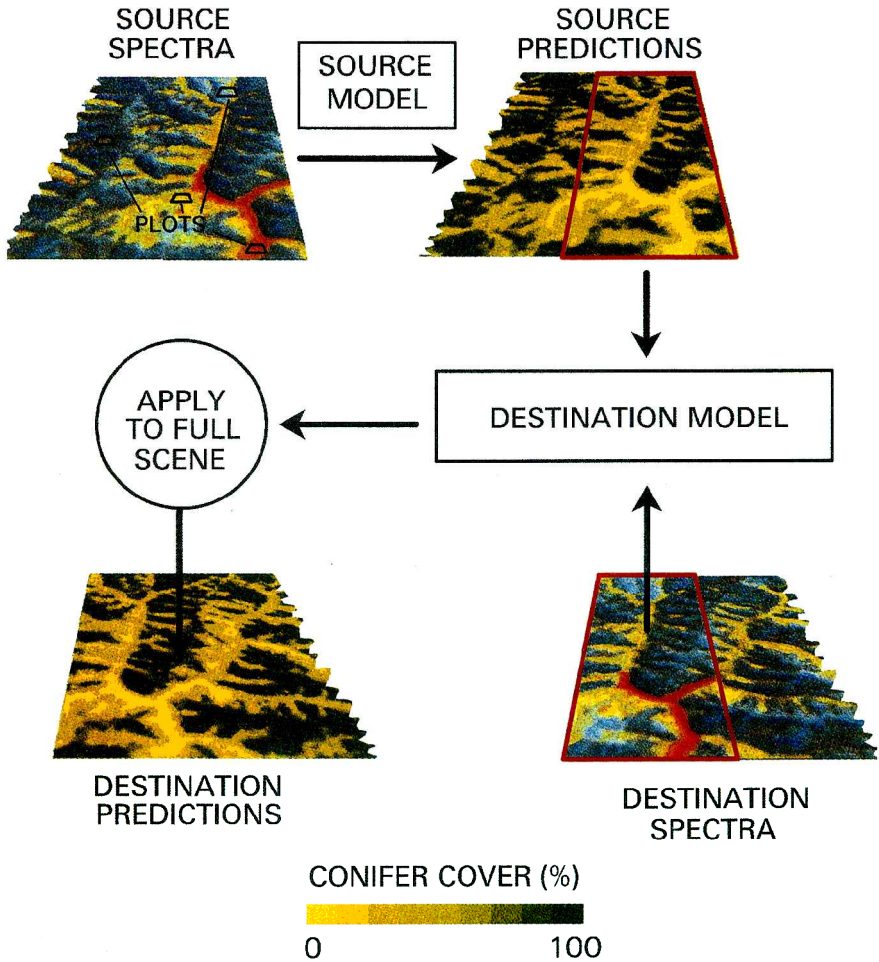


Figure 2. Concept of applied radiometric normalization.

proportion of the total study area than any other image used to construct the mosaic (figure 1(a)), and were acquired in summer (as were most other images).

Prior to developing regression models for the forest attributes of interest, forest had to be stratified from nonforest. For this we used a combination of masking and unsupervised classification. Masking involved the use of a land-use GIS layer acquired from the Oregon State GIS Services Center to eliminate agricultural and other nonforest land use designations. Unsupervised classification was used to distinguish such things as clouds, cloud and topographic shadows, and snow/ice.

Summaries of the reference data set show that a wide range of forest conditions were modelled, from open clearcuts to closed-canopy nonconifer stands to conifer stands having a wide range of tree sizes and ages (table 2). These data were used with the plot-level mean spectral and $\cos i$ values in a stepwise multiple regression analysis to derive source models. Model building commenced with graphing of basic relationships between dependent and independent variables. Several relationships were curvilinear, thus standard transformations (e.g. log, exponent, inverse, square terms) were applied to both the independent and dependent variables for use in

Table 2. Reference data statistics.

Attribute	Minimum	Maximum	μ	σ	n
Model building set					
Green vegetation cover (%)	0	100	91.8	19.7	559
Conifer cover (%)	0	100	59.6	36.1	559
Visible crown diameter (m)	1.4	10.6	5.26	2.71	46
Local stand age (years)	15	650	120.6	113.0	112
Model/map testing set					
Green vegetation cover (%)	0	100	88.0	23.4	672
Conifer cover (%)	0	100	61.7	36.6	672
Visible crown diameter (m)	1.7	13.2	5.78	2.77	186
Local stand age (years)	10	700	145.5	114.7	161

stepwise regression. We also allowed for selection of interaction terms (e.g. brightness \times wetness) in the models. A pool of potential models was generated for each attribute of interest, based on achieving the highest coefficients of determination with five or less independent variables. A final source model for each attribute (table 3) was selected based on analysis of residuals, and a desire to not overparameterize the models (i.e. in no case did a four or five variable model greatly reduce the sums of squares of error over three variable models).

For green vegetation cover and conifer cover, the full range of conditions from 0 to 100% was modelled. However, for the size and age attributes, acceptable coefficients of determination (>0.65) could be achieved only if the range of prediction was limited to conditions where both green vegetation cover and percent conifer cover were greater than 70%. Thus, all models were applied individually at the pixel-level to each pixel that was not masked by the land-use data layer or otherwise labeled nonforest during unsupervised classification. For size and age, models were applied only to pixels predicted as $>70\%$ of both green vegetation cover and conifer cover.

2.4.2. Destination models

Applied radiometric normalization was realized through the development and application of destination models. These models were constructed for the scene area within the overlap of adjacent images (figure 1(a)) in a several step process.

- The overlap area was analyzed to detect areas that have undergone substantial change between the two dates of imagery and to remove those pixels from consideration. Examples of change include new clearcuts, clouds in one date but not in another, and major phenological differences for specific areas within the overlap. We performed a simple Tasseled Cap greenness-difference thresholding to detect these changes (after Cohen *et al.* 1998).
- A statistical relationship was then developed using the predictions from the source models as dependent variables and the spectral responses of the destination scene as independent variables. This is akin to using the source predictions as a 'training set' for the models based on destination scene spectral properties. For the green vegetation and conifer cover models the full range of predictions was used. To simplify this step, all pixels predicted as a given integer (i.e. 1%, 2%, 3%, etc.) were used as a mask for that integer to select pixels in the destination image from which a mean spectral response was calculated for that

Table 3. Regression models used to map forest attributes. Note that the source scene for path/row 46/31 was 46/30.

Source models for 46/ 28–29 (Path/Row)						
Forest attribute	Intercept	Independent variable coefficients		Model r^2	RMSE	
Green vegetation cover (%)	-1161	9.90 (Wetness)	7.11 (Greenness)	0.71	559 10.7	
Conifer cover (%)	-201	0.042 (Wetness ²)	1.52 (Brightness)	0.76	559 17.8	
ln {Visible crown diameter (m)}	16.1	-0.13 (Wetness)	-0.0007 (Brightness × Greenness)	0.75	46 2.2 ^a	
ln {Local stand age (years)}	56.0	-0.42 (Wetness)	-1.06 (Brightness)	0.80	112 77.7 ^a	
Scene-specific destination models and overlap difference statistics for each forest attribute (on left side of forward slash is the value prior to overlap matching, on right side is value after post-matching)						
Green vegetation cover (%)						
46/30	-1139	9.73 (Wetness)	6.71 (Greenness)	-0.053 (Wetness × Greenness)	0.5	1
46/31	-1597	0.105 (Wetness ²)	20.36 (Greenness)	-0.160 (Wetness × Greenness)	10.9	8
45/30–31	1126	-0.065 (Wetness ²)	-16.25 (Greenness)	0.129 (Wetness × Greenness)	1.1	0
47/28	-950	8.24 (Wetness)	5.34 (Greenness)	-0.042 (Wetness × Greenness)	0.7	1
47/29	-1070	9.08 (Wetness)	5.94 (Greenness)	-0.046 (Wetness × Greenness)	0.7	1
Conifer cover (%)						
46/30	-1100	13.60 (Wetness)	2.73 (Brightness)	-0.034 (Wetness × Greenness)	1.8	1
46/31	-1206	16.31 (Wetness)	2.63 (Brightness)	-0.048 (Wetness × Greenness)	1.5	1
45/30–31	-1527	19.13 (Wetness)	0.050 (Brightness ²)	-0.048 (Wetness × Greenness)	-2.9	-2
47/28	-1045	12.84 (Wetness)	2.14 (Brightness)	-0.031 (Wetness × Greenness)	4.3	3
47/29	-803	10.05 (Wetness)	1.36 (Brightness)	-0.024 (Wetness × Greenness)	-1.1	-2

Table 3. (Continued).

In{ Visible crown diameter (m)}			Mean difference (m)	Median difference (m)	Mode difference (m)
46/30	16.8	-0.13 (Wetness)	0.0006 (Wetness × Brightness)	0.5	0
		-0.0004 (Brightness × Greenness)			
46/31	19.0	-0.14 (Wetness)	-0.0005 (Wetness × Brightness)	0	0
		0.0001 (Brightness × Greenness)			
45/30-31	17.9	-0.13 (Wetness)	-0.0002 (Wetness × Brightness)	0.2/0	0/0
		0.0002 (Brightness × Greenness)	2.05/-0.01		
47/28	19.0	-0.14 (Wetness)	0.0003 (Wetness × Brightness)	0	0
		-0.0002 (Brightness × Greenness)	0.18		
47/29	17.9	-0.14 (Wetness)	-0.0008 (Wetness × Brightness)	0	0
		0.0005 (Brightness × Greenness)	0.11		

In{ Local stand age (years)}			Mean difference (years)	Median difference (years)	Mode difference (years)
46/30	73.5	-0.56 (Wetness)	0.010 (Wetness × Brightness)	-3/0	-4/0
46/31	76.5	-0.59 (Wetness)	0.013 (Wetness × Brightness)	2/0	0/0
45/30-31	62.7	-0.47 (Wetness)	0.008 (Wetness × Brightness)	9/0	0/0
47/28	86.4	-0.66 (Wetness)	0.012 (Wetness × Brightness)	1/0	2/2
47/29	69.8	-0.53 (Wetness)	0.010 (Wetness × Brightness)	0	0
		-1.31 (Brightness)	-12.0/4.5		
		-1.59 (Brightness)	11.9/-0.4		
		-0.99 (Brightness)	25.2/0		
		-1.52 (Brightness)	6.8/-0.6		
		-1.27 (Brightness)	1.9		

^a Back-transformed into linear units.

integer of prediction. This was repeated for each integer of prediction over the full range of predictions. Because the source prediction models were not constrained between 0 and 100%, some predictions outside that range were included in the destination models. However, the models were weighted by the number of predictions at each interval so that they were not unduly influenced by these outlier predictions. Weighting the model in this way provides a model with the same coefficients that would be obtained if the spectral properties of individual pixels were used rather than the mean responses per interval. For the size and age attributes an additional thresholding was imposed by limiting the selection of destination model building pixels to those that were predicted as at least 70% green vegetation cover and 70% conifer cover.

- Destination models were developed in a manner consistent with that of the source models. We used the same basic independent variable set as those selected for the source models. However, to allow for nonlinearities in the relationship between spectral properties of adjacent scenes, the form of the independent variable set was permitted to vary (e.g. wetness was permitted to be replaced by the square of wetness, $wetness^2$) (table 3).
- The final step in applied radiometric normalization is to apply the destination model to the full destination scene. Prior to model application, the destination image is masked using the land use data layer and then subjected to an unsupervised classification to eliminate nonforest pixels. Similarly, the visible crown diameter and local stand age models are only applied to pixels that are predicted as being at least 70% green vegetation and conifer cover. Path/row 46/31 did not physically overlap the source scene, thus destination models for this scene were developed using 46/30 as a secondary source scene (table 1), constituting a double model extension.

2.5. Evaluation and post-modelling adjustment

If applied radiometric normalization was a perfect procedure there would be no difference in the predictions for any co-registered pixel from the two models (source and destination) within the overlap area. Given the statistical nature of the process, however, inconsistencies are inevitable. In fact, one means of determining the efficacy of applied radiometric normalization is to compare source and destination predictions within the overlap area. This is done by constructing a frequency histogram of prediction differences, after subtracting the destination predictions from the source predictions. Since the prediction models are unconstrained with respect to prediction range, it is quite possible to have rather unrealistic predictions for certain pixels. This normally occurs only for pixels having spectral responses that vary considerably in the overlap region between the adjacent images. Commonly, such a condition occurs because of the imperfect nature of geometric co-registration compounded by the presence of a high contrast edge (e.g. an old forest stand bordering a new clearcut), or an imperfect screening for change detection in the overlap region. When comparing prediction-difference histograms, it is useful to look not just at the mean prediction difference, but also the mode and median. Due to some extremely high or low difference values it is possible that the mean will differ from the median and mode, limiting the importance of the mean difference in evaluating this method.

Another way to determine how well the normalization procedure worked is to view images of the predictions themselves, where predictions from adjacent scenes

have been mosaicked. Although this is a qualitative view of any mismatch, it is a powerful means of determining how well the procedure minimized seams and gets directly at the objective of minimizing seams. Should there be a significant qualitative or quantitative mismatch, it can be ignored or minimized with a post-modelling adjustment. If warranted, we used a standard histogram matching algorithm that involves development of a look-up table to adjust predictions in the destination image, based on matching the cumulative frequency histograms of the source and destinations predictions (Mather 1987).

2.6. Map generation

For practical reasons, to develop maps of forest cover, predictions for green vegetation cover and conifer cover were constrained by truncation to between 0 and 100%. Because the size and age models are logarithmic, the minimum prediction is greater than zero. However, the upper ranges of predictions were constrained to 15 m diameter and 800 years, respectively, which are near the maximum values in reference data set and in the bulk of the population being modelled. To demonstrate how a single aggregate map can be constructed from the multiple continuous data layers, we combined three of those layers: green vegetation cover, conifer cover, and local stand age. First, three broad forest classes were determined, including open (<30% green vegetation cover), semi-open (30–70% green vegetation cover), and closed (>70% green vegetation cover). Second, within the closed class three subclasses were defined as follows: needleleaved (>70 conifer cover), broadleaved (if green vegetation cover minus conifer cover was >70%), and mixed (if the other two conditions were not met). Finally, within the conifer subclass, three age classes were defined: <80 years, 80–200 years, and >200 years.

2.7. Error characterization

To model forest attributes we used regression analysis, involving mean spectral values from each of numerous plots. Each of the selected models was evaluated in terms of measures such as the coefficient of determination and standard error of the estimate. A more powerful performance measure involved application of the models to an independent data set. For this we used the plot-level mean spectral values from the model testing portion of the reference data set. This resulted in a new coefficient of determination, calculated for the relationship between predicted and observed values. The slope and intercept of the predicted versus observed regression equation provided information on whether prediction biases were evident.

In addition to testing model performance, we demonstrate how the model testing data can be used to characterize errors in the aggregate map. Using the same classification logic as that imposed on the continuous data layers, the plot-level reference data were assigned class labels based on an aggregation of interpreted attributes. Then, to characterize map errors mean plot-level map predictions were compared against the aggregated reference data. Both standard and weighted error matrices were constructed.

3. Results

3.1. Source models

Models for the source scene (46/28–29) all contain three independent variables (table 3). Each model contains wetness, with the model for conifer cover relying on the square of that index. Consistent with our earlier work, wetness decreases with

increases in crown diameter and age, and for green vegetation cover and conifer cover wetness increases (Cohen and Spies 1992, Cohen *et al.* 1995). The green vegetation cover model includes greenness, which is not surprising given the strong relationship that its kindred index, NDVI, has with increases in cover over the relatively low leaf area index range from about $0\text{ m}^2\text{ m}^{-2}$ to $3\text{ m}^2\text{ m}^{-2}$ (Turner *et al.* 1999). The conifer cover model includes brightness rather than greenness. This model is not precisely the complement of hardwood cover, given that at low levels of green vegetation cover, the open component is a potential scene element. However, as will be shown later, the study area is occupied largely by closed vegetation conditions. This suggests that brightness quality of green vegetation cover, perhaps driven by amount of canopy shadow, is more important for distinguishing between hardwood and conifer canopies than is greenness quality of foliage. The local stand age model for conifers is similarly driven by brightness, probably a strong indication that amount of shadow is an important for distinguishing among conifer ages. The remainder of the terms in all four models are interaction terms, with visible crown diameter being a function of both brightness \times greenness and wetness \times brightness, whereas the other models rely on wetness \times greenness or wetness \times brightness.

The coefficients of determination for all four models are relatively high, at 0.71 to 0.80 (table 3). Although the r^2 for the conifer cover model is higher than for the green vegetation cover model, its root mean square error (RMSE) is considerably greater. Given the ranges over which the models for local stand age and visible crown diameter were developed, the RMSEs for these models are acceptable for most applications.

3.2. Destination models

For all destination models, the independent variable set used was the same as for the corresponding source model with a few exceptions (table 3). Although the wetness term is not squared in the source model for green vegetation cover, for two of the destination models wetness is squared, and for one the sign is reversed (45/30–31). Interestingly, the sign is also reversed for the other two terms of this destination model. These sign reversals are likely a function of a change in vegetation conditions relative to the rest of the study area, as Path/Row 45/30–31 cover the driest climate condition and highest elevations. This part of the region is more akin to the High Cascades Province and east side of the Cascade Range where major shifts in tree species occur relative to the remainder of the study area (Franklin and Dyrness 1988), which could lead to a shift in the relationship between spectral data and fundamental land cover characteristics. When constructing the map mosaics, Path 45 was overlain by Path 46 such that only the extreme southeast portion was actually mapped using the models that had reversed signs. For all of the conifer cover destination models wetness is not squared, even though it is squared in the source model. Additionally, for the 45/30–31 conifer cover destination model the brightness term is squared even though it was not squared for the source model. For some of the visible crown diameter destination models there were sign reversals for the second (Path 45) and third variables. Sign reversals for the third variable (wetness \times brightness) are associated with three separate model extensions. This, in conjunction with the low coefficient values for this variable, might indicate that these models were overfit.

The source/destination overlap diagnostics indicate that most destination models did not require post-modelling adjustment (table 3). Exceptions are four of the five

local stand age destination models and one visible crown diameter model. An example of the overlap diagnostics is shown in figure 3, where the mean difference before and after adjustment was -12 years and 4.5 years, respectively. The median and mode went from -3 and -4 to zero.

For the double model extension (46/30–46/31, table 1) of green vegetation cover, the mean difference in predictions was nearly 11% and the median difference was 8%. Although these are greater differences than were generally considered acceptable, we did not perform a post-modelling adjustment for this model. The destination image, in this case, was acquired in April, early in the growing season. As such, green vegetation cover was almost certainly lower than it would have been later in the growing season, as it was for the 46/30 image (30 July) from which the green vegetation cover model was being extended. Thus, we accepted this more realistic, early spring estimate of green vegetation cover for the 46/31 scene, rather than forcing it to have a higher mean green vegetation cover by performing a post-modelling adjustment.

3.3. Error characterizations

3.3.1. Models

Model predictions compared against plot-level observations indicate fairly strong predictive strength for all four forest attributes (figure 4, table 4). Green vegetation cover is a particularly effective model ($r^2=0.74$, RMSE=12%), however there is a tendency for the model to overpredict if actual cover is less than about 70%. This is evident both from the regression line (figure 4) and the slope and intercept values of the predicted versus observed regression equation (table 4). The conifer cover model is a somewhat less powerful predictor than the green vegetation cover model ($r^2=0.69$, RMSE=23%), and also has a positive bias at low conifer cover levels ($<30\%$), with a slight underprediction bias above 80%. The size and age models perform less effectively than the cover models ($r^2=0.51$, RMSE=2.1 m; $r^2=0.56$, RMSE=107 years, respectively). The size model tends to overpredict at values below 5 m and underpredict above that value. The age model exhibits a slight positive bias

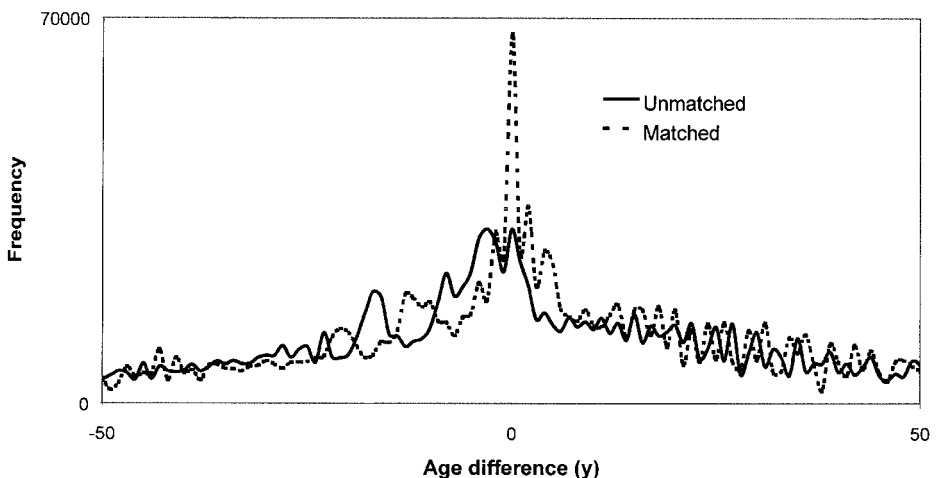


Figure 3. Frequency distribution of local stand age prediction differences (source minus destination) in the overlap of path/row 46/28–29 (source) and 46/30 (destination).

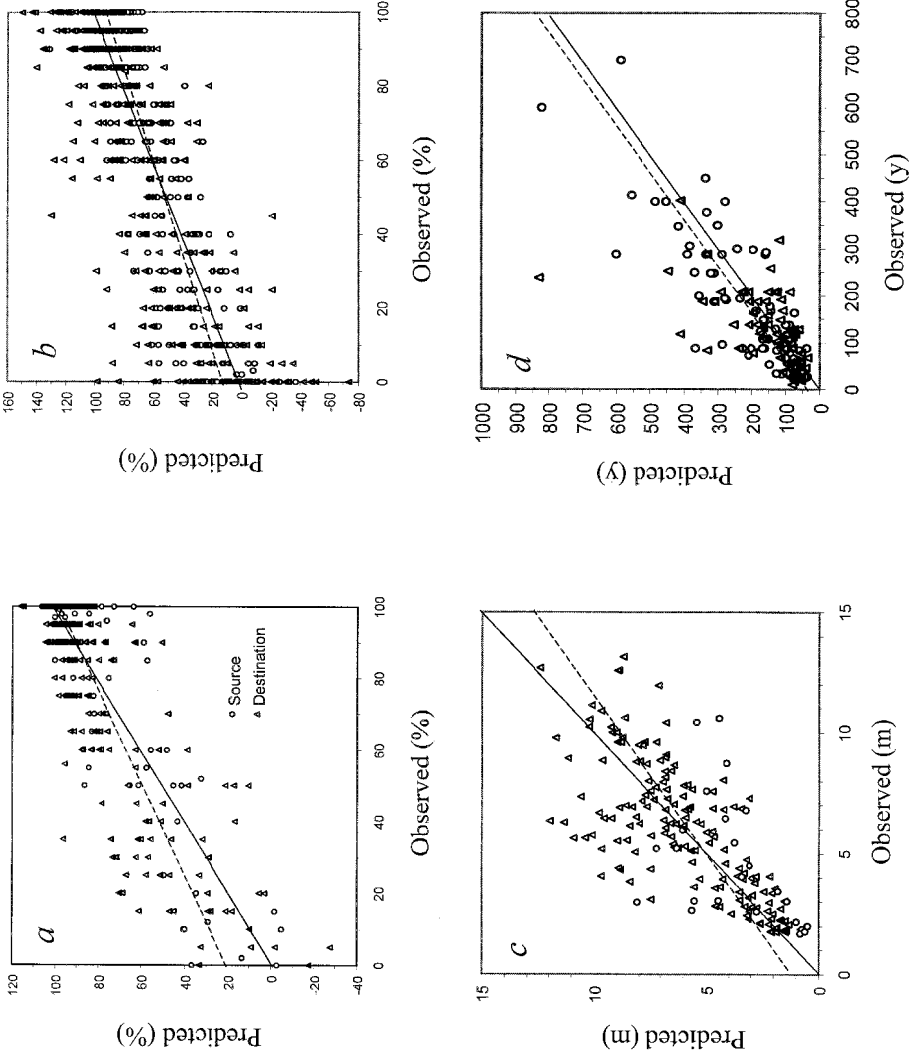


Figure 4. Relationship between model predictions and the model testing reference data set (source = 46/28–29; destination = all other images): (a) green vegetation cover (%), (b) conifer cover (%), (c) visible crown diameter, and (d) local stand age. The 1:1 line is solid and the regression line is dashed.

Table 4. Description of relationships between model predictions and independent observations.

Forest attribute	Intercept	Slope	r^2	n	RMSE
Green vegetation cover (%)	21.1	0.77	0.74	672	11.9
Conifer cover (%)	16.4	0.82	0.69	672	22.5
Visible crown diameter (m)	1.4	0.73	0.51	186	2.8
Local stand age (years)	30.9	1.01	0.56	161	107

throughout the full range of prediction. Unlike the cover models, the size and age models have a tendency to be significantly more variable at higher values (figure 4). None of the prediction models incorporated the $\cos i$ variable, and regression prediction residuals plotted against $\cos i$ revealed that there was minimal bias in the predictions associated with that variable.

3.3.2. Aggregate map

The class map has an overall accuracy of 66% for the cover attributes and 72% for three classes of local stand age (tables 5 and 6). We do not combine these two error matrices because conifer cover and age are not mutually exclusive, we have no independent affirmation that all stands for which we have reference ages are in fact closed conifer by our definition, and other related reasons. Consistent with the continuous data error assessment (figure 4), there was a strong tendency toward over prediction at the low end of the cover gradient, as witnessed by the open category being misclassified as semi-open and the semi-open category misclassified as closed. If a stand was closed it was generally predicted as such; however, there was a tendency to slightly underpredict the proportion of conifer cover in the closed condition. Unlike for the cover variables, the predictions of stand age are fairly well balanced; although we do note a tendency to overpredict stand age.

All types of misclassification are not ecological equivalent. For example, classifying open as closed conifer is considerably more troublesome than classifying closed mix as closed conifer when one is attempting to model a population of fauna dependent on old-growth conifer forest. With this concept in mind, we constructed a similarity matrix to define three relative weights of ecological significance (table 7): zero is used for unacceptable misclassifications; 0.5 is for acceptable misclassifications; and 1 represents correct classification. Considering the order of the class structure from open to closed and from young to old as being closely aligned with successional gradients, acceptable misclassifications might be predictions within one class of truth. The weight table for local stand age is not given in table 7 as it is obvious. Using this weighting system, the cover estimates are 82% correct and the age estimates are 86% accurate.

To determine if map errors are distributed differently among provinces, error matrices were assembled by province. For the cover attributes there were no major differences among province accuracies—57% for the Western Cascades, 69% for the Coast Ranges, and 61% for the Klamath Mountains. For local stand age however, the Cascades (79%) and Coast (86%) had significantly higher accuracies than the Klamath (51%).

3.4. Regional vegetation characteristics

Based on model predictions, in 1988 the forested portions of western Oregon were largely in a closed canopy condition, with only about 15% of the area having

Table 5. Error matrix for percent green vegetation cover and percent conifer cover ($n=672$).

Predicted	Observed					Per cent correct	
	Open	Semi-open	Closed broadleaved	Closed mixed	Closed conifer	Absolute	Weighted
Open	13	4	0	0	0	76.5	88.2
Semi-open	23	23	3	5	5	39.0	69.5
Closed broadleaved	1	6	34	9	1	66.7	81.4
Closed mixed	5	17	50	58	52	31.9	64.6
Closed conifer	1	4	8	36	314	86.5	91.7
	Per cent correct						
Absolute	30.2	42.6	35.8	53.7	84.4	65.8	
Weighted	57.0	69.4	63.7	76.9	92.1		81.5

Table 6. Error matrix for local stand age ($n=161$).

Predicted	Observed			Per cent correct	
	< 80 years	80–200 years	> 200 years	Absolute	Weighted
< 80 years	38	6		86.4	93.2
80–200 years	14	49	9	68.1	84.0
> 200 years		16	29	64.4	82.2
	Per cent correct				
Absolute	73.1	69.0	76.3	72.0	
Weighted	86.5	84.5	88.2		86.0

Table 7. Similarity matrix for green vegetation cover and percent conifer cover.

Predicted	Observed				
	Open	Semi-open	Closed broadleaved	Closed mixed	Closed conifer
Open	1	0.5	0	0	0
Semi-open	0.5	1	0.5	0.5	0.5
Closed broadleaved	0	0.5	1	0.5	0
Closed mixed	0	0.5	0.5	1	0.5
Closed conifer	0	0.5	0	0.5	1

less than 70% green vegetation cover (figure 5). The region is also largely coniferous, with less than 30% of the area containing less than 50% conifer, and about 50% of the area consisting of greater than 70% conifer forest. In closed-canopy conifer stands, visible crown diameters are highly variable over the region, with each of five tree size categories (< 3 m, 3–6 m, 6–9 m, 9–12 m, and > 12 m) containing nearly 20% of the size distribution. Local stand age is also quite variable, with 20% of conifer stands being less than 50 years old, 30% between 50 and 100 years old, 30% between 100 and 200 years old, and the remaining 20% being above 200 years old. Given the political sensitivity of forest management in the PNW region, it is worth restating that map error statistics presented earlier (table 6) indicated a tendency to overpredict age class.

Evaluation of model predictions at the province level reveals some basic provincial differences, particularly with respect to conifer size and age (figure 6). The Western Cascades tend to be dominated by medium-sized crowns (4 m to 9 m) and the Coast Ranges by smaller crowns (2 m to 6 m), whereas crown sizes in the Klamath Mountains are highly variable (2 m to 9 m). Likewise, stand ages in the Coast Ranges are predominantly less than 100 years, and although many conifer stands in the Cascades are below 100 years, the predominant age is above 100 years (with a significant proportion above 200 years). Again, the Klamath Mountains have an intermediate distribution, with ages largely ranging between 30 and 200 years. Given that the predominant condition in the region is closed-canopy forest we can assume that, in general, the absence of conifer forest equates to the presence of hardwood trees or a complex of mixed shrub and herbaceous species (as opposed to soil and litter). As such, we can see that although the distribution of conifer cover in the

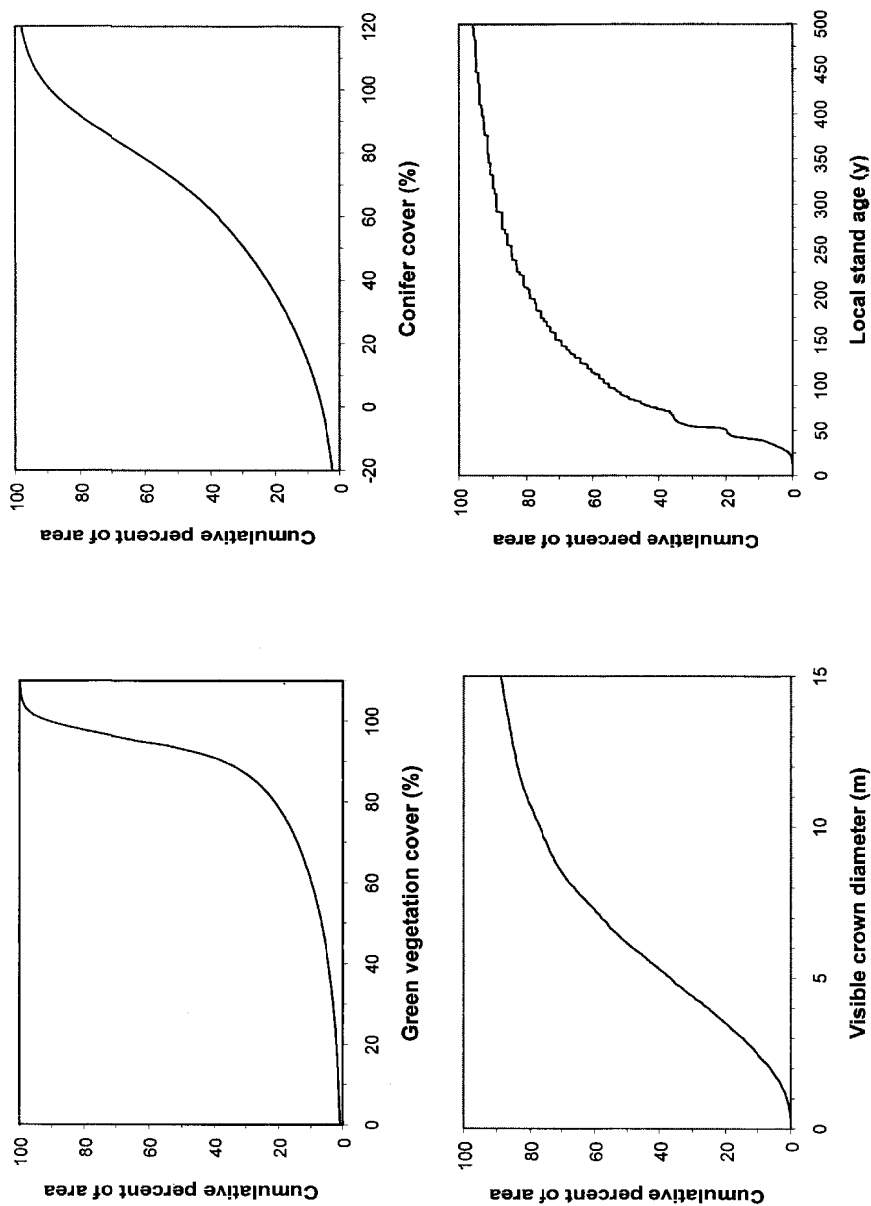


Figure 5. Cumulative frequency distributions for regional predictions of the four modelled forest attributes. The upper tails of all frequency distributions are not shown.

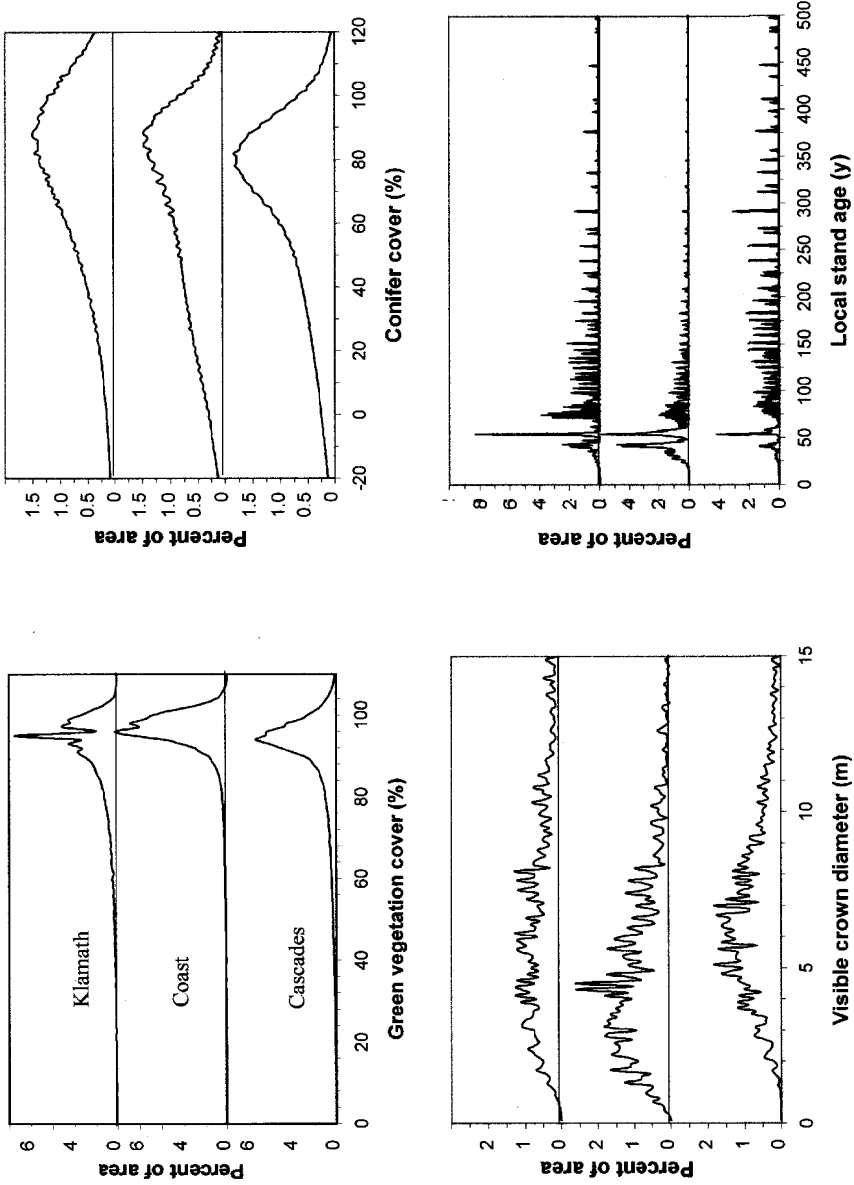


Figure 6. Frequency distributions of predictions for the four modelled forest attributes, by province. The upper tails of all frequency distributions are not shown.

Coast Ranges is skewed slightly more to the higher end than the distribution in the Cascades, the hardwood and mixed brush condition is somewhat more prevalent in the Coast than in the Cascades. Consistent with the other attributes discussed, the Klamath has an intermediate distribution.

Using model predictions to develop maps of forest conditions provides information about spatial patterns of vegetation across the region (figure 7). Noteworthy patterns include: (1) the difference in conifer age class distributions between the Coast Ranges and the Western Cascades; (2) open and semi-open areas that are distributed as scattered, recently clearcut lands and in a recently burned area in the Klamath Province; (3) the increase in conifer cover from the Coast to the Cascades; and (4) the influence of land ownership (figure 1(b)), with public lands containing

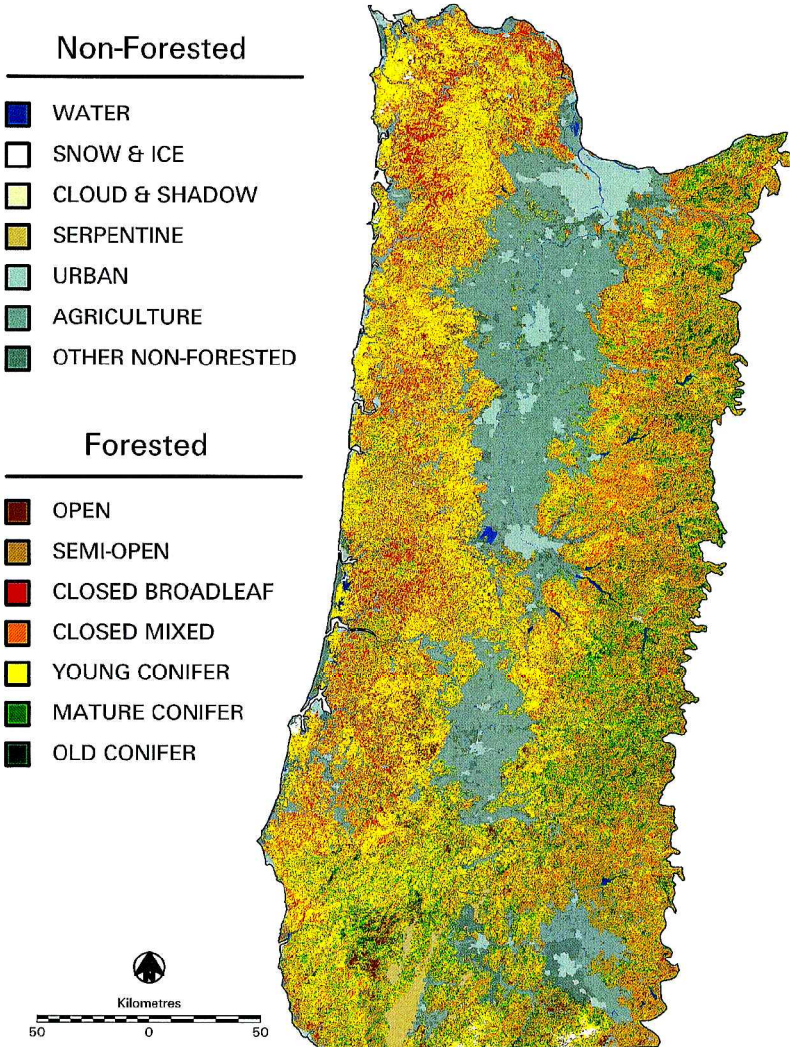


Figure 7. Land cover classification derived from overlay of predictions for three forest attributes (green vegetation cover, conifer cover, and local stand age), and the combination of unsupervised classification and the land-use mask (for non-forest classes).

most of the older conifer forest and private industrial and non-industrial lands containing a high proportion of early-successional and young conifer forest.

3.5. Efficacy of applied radiometric normalization

A regional view of the class map suggests that seam boundaries where images were mosaicked are essentially nonexistent (figure 7) and this is confirmed by a close-up view (e.g. figure 8). It is particularly important to note that if seams were present in individual data layers (green vegetation cover, conifer cover, age) they would be exacerbated by the overlay procedure used to develop the class map. Even the seam at the juncture of 46/30 and 46/31, where percent green vegetation cover was not precisely matched and where there was a double extension, is not clearly visible.

Besides examining the prediction-difference histograms, as described earlier, another powerful means of evaluating how well applied radiometric normalization worked was to examine whether there were biases in predictions across scene boundaries (figure 4). For green vegetation cover, conifer cover, and stand age we see that the relationships between model-predicted and observed were consistent for both the source and destination images, indicating the success of applied radiometric normalization for these variables. However, for visible crown diameters beyond about 8 m there is a disparity in the source and destination crown size samples, with destination samples being overpredicted and source samples being underpredicted. This may be due in part to the relatively few plots used to develop this model for the source scene (46, versus 112 for stand age) and the relatively narrow range over

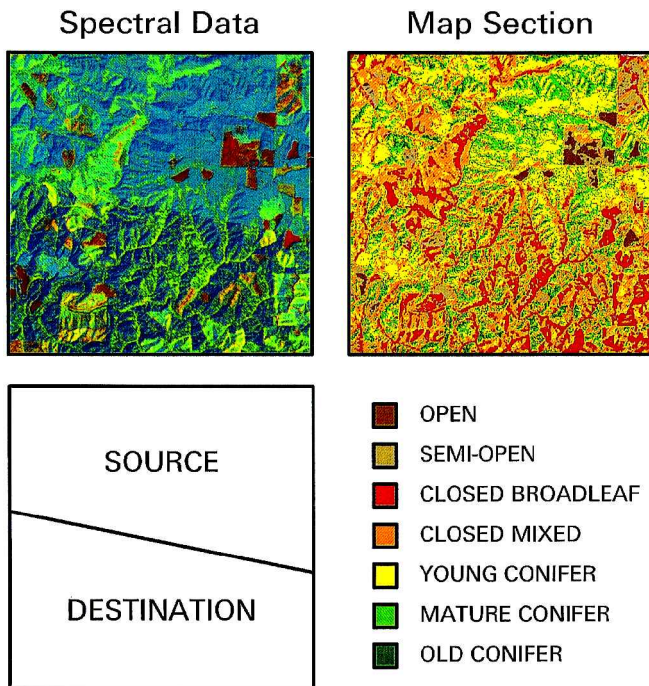


Figure 8. Example of an obvious radiometric 'seam' in the mosaicked spectral data (source = 46/29, destination = 46/30), and the lack of a seam in the associated map section. For the spectral data display, RGB = brightness, greenness, wetness.

which the model was developed (<11 m), compared to the range over which the model was applied (up to 13 m).

4. Discussion

4.1. Flexible land cover information

Land cover maps are essential when they are used to help inform decisions related to spatial allocation of resource management prescriptions or other related activities. For maximum utility, cover maps should contain information that represents the basic building blocks for user-specified classifications. There are any number of possible alternatives for this approach to land cover mapping.

Running *et al.* (1994) presents a land cover logic for use in modelling global biogeochemical cycles in which there are six types of vegetation cover specified: evergreen needleleaf forest, evergreen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, annual broadleaf nonforest, and grasses. Depending on climate, specific subtypes can be further specified, such as boreal or temperate for evergreen needleleaf forest. At a regional level, one could further subdivide such classes into species-association groups, assuming the species groups are spectrally or environmentally distinct. Although this land cover logic is clearly amenable to further user-specified class definition, the logic itself is classification based. That is, every cell in a map based on this logic will have one of the six basic cover-class labels. At the cell level, especially the 1 km or so cell size used for global mapping, two or more of these classes may be present. Missing from this logic, and thus limiting its flexibility, is a continuous estimate of proportions of these basic classes.

Milne and Cohen (1999) present a method by which high spatial resolution data (e.g. Landsat TM) can be used in conjunction with a coarse spatial resolution classification (e.g. from the Moderate Resolution Imaging Spectroradiometer (MODIS) or Advanced Very High Resolution Radiometer (AVHRR)) to define mixing proportions of basic cover type classes within the coarse resolution cells. This method, called proportional reclassification, requires a classification at the finer resolution which is then overlain with the coarser grid and summarized to derive the proportion of each land cover class within each coarse resolution cell. Any fine-grained locally or regionally important classification logic can be used. One would simply characterize the proportions of regionally important cover classes contained within the globally important coarse-grained cover classes. However, this does not provide mixing proportions at the fine grain needed for regional applications.

A direct means of providing basic building blocks for user-specified classifications at a fine grain size is mixture modelling. Mixture modelling has been used to determine percent vegetation cover in natural grasslands (Pearson 1971), cropland (Quarmby 1992), young pine plantations (Elvidge *et al.* 1993), tropical forest (Cross *et al.* 1991), and semi-arid shrubland (Ustin *et al.* 1986). For vegetation mapping, most of this activity has been focused on characterizing fractions of globally important elemental scene components (endmembers) such as soil, green vegetation, non-photosynthetic vegetation, and shade (Smith *et al.* 1990, Adams *et al.* 1995). It should be possible to define additional, regionally important mixture modelling endmembers for vegetated systems, but examples of this are not prevalent in the literature.

In the work presented here, we attempt to define regionally important fine-grained elemental scene components that provide the basic building blocks for user-defined land cover classes. In the PNW region, important basic elemental scene components include proportions of conifer and broadleaved cover, stand structure

related variables such as tree size, age, biomass, volume, and downed woody debris, as well as fine-grained information such as snags. Unfortunately, it is not possible to extract accurate information concerning several of these scene components from current satellite data. However, we have demonstrated that one can reliably map several important basic scene components at the regional level using TM data. Percent green vegetation cover provides information on the recency of disturbance, percent conifer cover is a measure of the degree to which an area is occupied by the most common vegetation lifeform in the region, and visible crown diameter and local stand age are suitable surrogates for forest structure. Moreover, as class level data are commonly preferred for resource management decisions and are useful in mapping and patterns analysis, we demonstrated that one can utilize basic continuous cover information to derive a set of ecologically important forest classes for use in subsequent studies.

4.2. Applied radiometric normalization for regional mapping with Landsat data

Most regional mapping efforts rely on coarse resolution imagery. For example, in a 1994 special issue of *Photogrammetric Engineering and Remote Sensing* on large area land cover characterization, five out of the six papers that actually mapped an area greater in size than a single Landsat scene, used AVHRR data. However, because even at the regional level fine-grained land cover information is often desirable, there have been a handful of regional mapping studies based on Landsat TM. Fuller *et al.* (1994) developed a land cover map of Great Britain using TM data. The reference data they used were based on field reconnaissance and there was no attempt to radiometrically normalize the imagery. Rather, individual land cover maps were developed for each image, and after classification the maps were mosaicked using a 'blending' logic at the image seams. As a part of the Gap Analysis Program, Homer *et al.* (1997) mosaicked 14 TM images of Utah prior to developing a land cover class map. Radiometric normalization was accomplished by spectral histogram matching, where radiometrically stable targets were hand selected, with attention to adequate sampling of bright and dark targets. Airphotos were used to supplement field data for training and testing. To assemble their reference database for TM mapping of northeastern Minnesota, Bauer *et al.* (1994) conducted field surveys to supplement airphoto interpretation, but they chose to mosaic individual images after each was classified. Vogelmann *et al.* (1998) used TM and airphoto data to map a four state area in the eastern United States, and like Homer *et al.* (1997) they used histogram matching to normalize the spectral data prior to constructing the mosaic and subsequent land cover class map.

From these regional land cover mapping studies based on TM data, two consistent themes emerge. First, it is apparent that the need to assemble an adequate reference database is critical, with most efforts relying on project-specific field-based data that are sometimes supplemented by airphotos samples. Second, edge-matching at image boundaries is an important consideration for final map products. Edge-matching has been accomplished both by relative radiometrically normalization of the spectral data prior to classification and by post-classification procedures that attempt to blend independent scene-based classifications. In this study, we developed a new analysis framework for regional land-cover mapping applications of Landsat TM data. This framework is based on the use of existing aerial photography and digital stand-level data files and on a process we call applied radiometric normalization.

For obvious reasons, a reliance on existing airphotos and digital data (should

those data exist) to build the reference database is cost-effective in comparison to a new field survey effort. An additional cost-related benefit of our analysis framework—a benefit that becomes increasingly more important the less one relies on existing reference data—is that a larger proportion of the total reference data set is used for characterizing errors in the cover map than is used for training the mapping algorithm. This latter point is possible because of applied radiometric normalization, which has the added benefit of facilitating the creation of a seamless map for regional characterizations. Applied radiometric normalization differs from more standard approaches to normalization in that the spectral data themselves are not normalized. Rather, a source image is translated into land cover information and the cover information is extended to adjacent images using a statistical relationship between source predictions and extension spectral properties in the overlap region of the images.

Given that applied radiometric normalization is based on spatial extension of the relationships between vegetation attributes and spectral properties, it is important to ask ‘What are the spatial limits of extensibility?’ For variables like percent green vegetation cover and percent conifer cover, there may be no obvious spatial limit within a specific biome type, such as needleleaf evergreen forest. For example, in this study, both of these variables were successfully extended over several different distinct climatic and vegetation zones consisting of significant variation in a large number of dominant plant species and growth forms. The only troublesome extension was the double extension from 46/30 to 46/31 for green vegetation cover. The percent cover mismatch that occurred here, however, was largely due to the early April date of the 46/31 image compared to the late July 46/30 image. Double extension was probably not the cause for this mismatch, as the extension of conifer cover (an attribute that is not significantly variable from April to July) to this image was quite successful. Additionally, the three provinces over which this study was conducted had similar (though not exact) distributions of both green vegetation cover and conifer cover. This undoubtedly aided the extension process.

For local stand age, applied radiometric normalization was somewhat less effective, as indicated both by the need to execute a post-modelling adjustment and by the variations in map errors among provinces. This apparent limitation may be a function of several factors. With a post-modelling adjustment there was little difference in the form and shape of the predicted versus observed relationship when comparing the source and destination results (figure 4), indicating that extension was satisfactorily accomplished. However, for mapped classes the prediction strength was significantly less for the Klamath province than for the other two provinces. Whether this is due to misapplication of applied radiometric normalization, or an inherently noisier relationship between age and spectral response in the Klamath, is unknown. Age is not a directly estimated characteristic as is greenness or vegetation cover, and additional factors likely influence the relationship between canopy structure and age.

4.3. *Management implications*

The implementation of the Northwest Forest Plan (NWFP) for federal lands in the Pacific Northwest (Tuchmann *et al.* 1996) moves federal planning and management to a regional level and degree at which it has never occurred before. Regional-scale vegetation databases such as we developed here are needed to implement monitoring components of the NWFP which focus on landscape patterns of forest

structure and habitat conditions for multiple species (Mulder *et al.* 1999). The results of this study indicate that relatively rapid, cost-effective, flexible, approaches to mapping forest conditions using TM can be developed at regional scales. The vegetation maps produced by this process can provide a basis for characterizing the general spatial distribution and pattern of old-growth forest conditions. Given the uncertainty in the needs of yet-to-be-developed landscape-scale wildlife habitat models, the continuous approach provides flexibility to model a variety of species with different cover-class requirements.

Repeated measures of regional vegetation, as called for under the NWFP, also require cost-effective, relatively rapid approaches that can be conducted within very limited agency budgets for monitoring and no funding for new field work to obtain model-building data sets. Our approach can meet these requirements. Further study is needed, however, to determine how well existing regional plot inventories can be used to build remote sensing models, how model error varies regionally, sources of error in regional models, and the minimum areal extent at which regional models can be applied to finer-scale management problems.

5. Conclusions

- Regional forest cover maps with acceptable error rates can be developed from Landsat TM data.
- Use of airphotos and existing digital forest databases facilitate cost-effective mapping.
- Forest attributes can be modelled as continuous variables that serve as building blocks for user-defined land cover classes.
- A process we call radiometric normalization enables seamless mapping across multiple image boundaries. Moreover, because this process uses reference data conservatively, it significantly reduces the cost of gathering those data for model building and testing.

Acknowledgments

This research was funded and/or supported by the Terrestrial Ecology Program and the Land-Cover Land-Use Change Program at NASA, the CLAMS study at the PNW Research Station and Oregon State University, the National Science Foundation-sponsored H. J. Andrews Forest LTER Program, and the EPA's National Health and Environmental Effects Research Laboratory. We greatly appreciate the reviews and thoughtful comments of Karin Fassnacht and Michael Lefsky and an anonymous reviewer.

References

- ADAMS, J. B., SABOL, D. E., KAPOV, V., FILHO, R. A., ROBERTS, D. A., SMITH, M. O., and GILLESPIE, A. R., 1995, Classification of multispectral images based on fractions of endmembers application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, **52**, 137–154.
- APAN, A., 1997, Land cover mapping for tropical forest rehabilitation planning using remotely-sensed data. *International Journal of Remote Sensing*, **18**, 1029–1049.
- AVERY, T., 1978, *Forester's Guide to Aerial Photo Interpretation*, Agricultural Handbook 308, United States Department of Agriculture, Washington, DC, USA.
- BAUER, M. E., BURK, T. E., EK, A. R., COPPIN, P. R., LIME, S. D., WALSH, T. A., WALTERS, D. K., BEFORT, W., and HEINZEN, D. F., 1994, Satellite inventory of Minnesota forest resources. *Photogrammetric Engineering and Remote Sensing*, **60**, 287–298.

- BUTERA, M. K., 1986, A correlation and regression analysis of percent canopy closure versus TMS spectral response for selected forest site in the San Juan National Forest, Colorado. *IEEE Transactions on Geoscience and Remote Sensing*, **24**, 122–129.
- CHAVEZ, P. S., 1989, Radiometric calibration for Landsat thematic mapper multispectral images. *Photogrammetric Engineering and Remote Sensing*, **55**, 1285–1294.
- COHEN, W. B., FIORELLA, M., GRAY, J., HELMER, E., and ANDERSON, K., 1998, An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, **64**, 293–300.
- COHEN, W. B., HARMON, M., WALLIN, D., and FIORELLA, M., 1996, Two decades of carbon flux from forests of the Pacific Northwest. *BioScience*, **46**, 836–844.
- COHEN, W. B., and SPIES, T. A., 1992, Estimating structural attributes of Douglas-fir/western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment*, **41**, 1–17.
- COHEN, W. B., SPIES, T. A., and FIORELLA, M., 1995, Estimating the age and structure of forests in a multi-ownership landscape of western Oregon, U.S.A. *International Journal of Remote Sensing*, **16**, 721–746.
- COLLINS, J. B., and WOODCOCK, C. E., 1994, Change detection using the Gramm–Schmidt transformation applied to mapping forest mortality. *Remote Sensing of Environment*, **50**, 267–279.
- COLLINS, J. B., and WOODCOCK, C. E., 1996, An assessment of several linear change detection techniques for mapping forest mortality using multitemporal Landsat TM data. *Remote Sensing of Environment*, **56**, 66–77.
- CONGALTON, R. G., GREEN, K., and TEPLY, J., 1993, Mapping old growth forests on national forest and park lands in the Pacific Northwest from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, **59**, 529–535.
- COPPIN, P., and BAUER, M., 1996, Digital change detection in forested ecosystems with remote sensing imagery. *Remote Sensing Reviews*, **13**, 207–234.
- CRIST, E. P., and CICONI, R. C., 1984, A physically-based transformation of thematic mapper data—the TM tasseled cap. *IEEE Transactions on Geoscience and Remote Sensing*, **22**, 256–263.
- CRIST, E. P., LAURIN, R., and CICONI, R. C., 1986, Vegetation and soils information contained in transformed thematic mapper data. *Proceedings of IGARSS '86, 6th International Geoscience and Remote Sensing Symposium, ESA Publication Division, Zurich, Switzerland, 8–11 September 1986* (Paris: European Space Agency).
- CROSS, A. M., SETTLE, J. J., DRAKE, N. A., and PAIVINEN, R. T., 1991, Subpixel measurement of tropical forest cover using AVHRR data. *International Journal of Remote Sensing*, **12**, 1119–1129.
- DILLWORTH, J., 1956, The use of aerial photographs in cruising second-growth Douglas-fir stands. Ph. D. Thesis, University of Washington, Seattle, USA.
- ECKHARDT, D. W., and VERDIN, J. P., 1990, Automated update of an irrigated lands GIS using SPOT HRV imagery. *Photogrammetric Engineering and Remote Sensing*, **56**, 1515–1522.
- ELVIDGE, C. D., CHEN, Z., and GROENEVELD, D. P., 1993, Detection of trace quantities of green vegetation in 1990 AVIRIS data. *Remote Sensing of Environment*, **44**, 271–279.
- ELVIDGE, C. D., YUAN, D., WEERACKOON, R. D., and LUNETTA, R., 1995, Relative radiometric normalization of Landsat Multispectral Scanner (MSS) data using an automated scattergram-controlled regression. *Photogrammetric Engineering and Remote Sensing*, **61**, 1255–1260.
- FRANKLIN, J., and DYRNESS, C., 1988, *Natural Vegetation of Oregon and Washington* (Corvallis, OR: Oregon State University Press).
- FRANKLIN, S. E., and MOULTON, J. E., 1990, Variability and classification of Landsat thematic mapper spectral response in southwest Yukon. *Canadian Journal of Remote Sensing*, **16**, 2–13.
- FULLER, R. M., GROOM, G. B., and JONES, A. R., 1994, The land cover map of Great Britain: an automated classification of Landsat Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, **60**, 553–562.
- GOWARD, S., and WILLIAMS, D., 1997, Landsat and Earth System Science: development of terrestrial monitoring. *Photogrammetric Engineering and Remote Sensing*, **63**, 887–900.
- GREGORY, S. V., HULSE, D. W., LANDERS, D. H., and WHITELAW, E., 1998, Integration of

biophysical and socio-economic patterns in riparian restoration of large rivers. In *Hydrology in Changing Environments, Proceedings of the British Hydrological Society International Conference, Exeter, UK, 1998*, Vol. 1, edited by H. Wheaton and C. Kirby (Chichester: Wiley), pp. 231–247.

- GRIGNETTI, A., SALVATORI, R., CASACCHIA, R., and MANES, F., 1997, Mediterranean vegetation analysis by multi-temporal satellite sensor data. *International Journal of Remote Sensing*, **18**, 1307–1318.
- HALL, F. G., BOTKIN, D. B., STREBEL, D. E., WOODS, K. D., and GOETZ, S. J., 1991a, Large-scale patterns of forest succession as determined by remote sensing. *Ecology*, **72**, 628–640.
- HALL, F. G., STREBEL, D. E., NICKESON, J. E., and GOETZ, S. J., 1991b, Radiometric rectification: toward a common radiometric response among multirate, multisensor images. *Remote Sensing of Environment*, **35**, 11–27.
- HEMSTROM, M., SPIES, T., PALMER, C., KIESTER, R., TEPLY, J., McDONALD, P., and WARBINGTON, R., 1998, Late-Successional and Old-Growth Forest Effectiveness Monitoring Plan for the Northwest Forest Plan for the Northwest Forest Plan. Report PNW-GTR-438, USDA Forest Service, Pacific Northwest Research Station, Portland, OR, USA.
- HENEBRY, G. M., and SU, H., 1993, Using landscape trajectories to assess the effects of radiometric rectification. *International Journal of Remote Sensing*, **14**, 2417–2423.
- HOMER, C., RAMSEY, R., EDWARDS, JR, T. C., and FALCONER, A., 1997, Landscape cover-type modelling using a multi-scene Thematic Mapper mosaic. *Photogrammetric Engineering and Remote Sensing*, **63**, 59–67.
- JENSEN, J. R., RUTCHEY, K., KOCH, M. S., and NARUMALANI, S., 1995, Inland wetland change detection in the Everglades Water Conservation Area 2A using a time series of normalized remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, **61**, 199–209.
- KNICK, S., ROTENBERRY, J., and ZARRIELLO, T., 1997, Supervised classification of Landsat Thematic Mapper imagery in a semi-arid rangeland by nonparametric discriminant analysis. *Photogrammetric Engineering and Remote Sensing*, **63**, 79–86.
- KRANKINA, O. N., FIORELLA, M., COHEN, W., and TREYFELD, R. F., 1998, The use of Russian forest inventory data for carbon budgeting and for developing carbon offset strategies. *World Resource Review*, **10**, 52–66.
- LANDGREBE, D., 1997, The evolution of Landsat data analysis. *Photogrammetric Engineering and Remote Sensing*, **63**, 859–867.
- LEHMKUHL, J. F., RUGGIERO, L. F., and HALL, P. A., 1991, Landscape-scale patterns of forest fragmentation and wildlife richness and abundance in the southern Washington Cascade Range. In *Wildlife and Vegetation of unmanaged Douglas-Fir Forests*, USDA Forest Service Gen. Tech. Rep. PNW-GTR-285, Portland, OR, USA, edited by L. F. Ruggiero *et al.*, pp. 425–442.
- MAIERSPERGER, T. K., COHEN, W. B., and GANIO, L., 2001, A TM-based Hardwood-Conifer Mixture Index for forests of the Oregon Coast Range. *International Journal of Remote Sensing*, **22**, 1053–1066.
- MATHER, P., 1987, *Computer Processing of Remotely-Sensed Images* (Chichester: Wiley).
- MILNE, B. T., and COHEN, W. B., 1999, Multiscale assessment of binary and continuous land cover variables for MODIS validation, mapping, and modelling applications. *Remote Sensing of Environment*, **70**, 82–98.
- MORRISON, P. H., KLOEPFER, D., LEVERSEE, D. A., SOCHA, C. M., and FERBER, D., 1991, *Ancient Forests in the Pacific Northwest* (Washington, DC: The Wilderness Society).
- MULDER, B. S., NOON, B. R., SPIES, T. A., RAPHAEL, M. G., OLSEN, A. R., PALMER, C. J., REEVES, G. H., and HART, H. H., 1999, The Strategy and Design of the Effectiveness Monitoring Program for the Northwest Forest Plan, Report PNW-GTR-437, USDA Forest Service, Pacific Northwest Research Station, Portland, OR, USA.
- MUSTARD, J. F., 1993, Relationships of soil, grass, and bedrock over the Kaweah Serpentine Melange through mixture analysis of AVIRIS data. *Remote Sensing of Environment*, **44**, 293–308.
- NARAYANA, A., SOLANKI, H. U., KRISHNA, B. G., and NARAIN, A., 1995, Geometric correction and radiometric normalisation of NOAA AVHRR data for fisheries applications. *International Journal of Remote Sensing*, **16**, 765–771.

- NEALE, C. M. U., and CROWTHER, B. G., 1994, An airborne multispectral video/radiometer remote sensing system: development and calibration. *Remote Sensing of Environment*, **49**, 187–194.
- PAINE, D., 1981, *Aerial Photography and Image Interpretation for Resource Management* (New York: Wiley).
- PEARSON, R. L., 1971, Design of Field Spectrometer Lab. MS Thesis, Colorado State University, Fort Collins, CO, USA.
- PEDDLE, D., HALL, F., LEDREW, E., and KNAPP, D., 1997, Classification of forest land cover in BOREAS. II: Comparison of results from a sub-pixel scale physical modelling approach and a training based method. *Canadian Journal of Remote Sensing*, **23**, 131–142.
- PETERSON, D. L., SPANNER, M. A., RUNNING, S. W., and TEUBER, K. B., 1987, Relationship of thematic mapper simulator data to leaf area index of temperate coniferous forests. *Remote Sensing of Environment*, **22**, 323–341.
- PICKUP, G., CHEWINGS, V. H., and PEARCE, G., 1995, Procedures for correcting high resolution airborne video imagery. *International Journal of Remote Sensing*, **16**, 1647–1662.
- QUARMBY, N. A., 1992, Towards continental scale crop area estimation. *International Journal of Remote Sensing*, **13**, 981–989.
- RUNNING, S. W., LOVELAND, T. R., and PIERCE, L. L., 1994, A vegetation classification logic based on remote sensing for use in global biogeochemical models. *Ambio*, **23**, 77–81.
- SCHOTT, J. R., SALVAGGIO, C., and VOLCHOK, W. J., 1988, Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment*, **26**, 1–16.
- SINGH, A., 1989, Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, **10**, 989–1003.
- SMITH, J. A., LIN, T. L., and RANSON, K. J., 1980, The Lambertian assumption and Landsat data. *Photogrammetric Engineering and Remote Sensing*, **46**, 1183–1189.
- SMITH, M. O., USTIN, S. L., ADAMS, J. B., and GILLESPIE, A. R., 1990, Vegetation in deserts: I. A regional measure of abundance from multispectral images. *Remote Sensing of Environment*, **31**, 1–26.
- SPIES, T. A., and FRANKLIN, J. F., 1991, The structure of natural young, mature, and old-growth forests in Washington and Oregon. In *Wildlife and Vegetation of unmanaged Douglas-Fir Forests*, USDA Forest Service Gen. Tech. Rep. PNW-GTR-285, Portland, OR, USA, edited by L. F. Ruggiero *et al.*, pp. 90–109.
- TUCHMANN, E. T., CONNAUGHTON, K. P., FREEDMAN, L. E., and MORIWAKI, C. B., 1996, The Northwest Forest Plan. A Report to the President and Congress, USDA Office of Forestry and Economic Assistance, Portland, OR, USA.
- TURNER, D. P., COHEN, W. B., KENNEDY, R. E., FASSNACHT, K. S., and BRIGGS, J. M., 1999, Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, **70**, 52–68.
- USTIN, S. L., ADAMS, J. B., ELVIDGE, C. D., REJMANEK, M., ROCK, B. N., SMITH, M. O., THOMAS, R. W., and WOODWARD, R. A., 1986, Thematic mapper studies of semiarid shrub communities. *BioScience*, **36**, 446–452.
- VERMOTE, E., and KAUFMAN, Y. J., 1995, Absolute calibration of AVHRR visible and near-infrared channels using ocean and cloud views. *International Journal of Remote Sensing*, **16**, 2317–2340.
- VOGELMANN, J. E., 1988, Detection of forest change in the Green Mountains of Vermont using Multispectral Scanner data. *International Journal of Remote Sensing*, **9**, 1187–1200.
- VOGELMANN, J., SOHL, T., and HOWARD, S., 1998, Regional characterization of land cover using multiple sources of data. *Photogrammetric Engineering and Remote Sensing*, **64**, 45–57.
- WADSWORTH, R., COX, R., and FULLER, R., 1997, The adjustment of frequency distributions of cover types within sections of the Land Cover Map of Great Britain. *International Journal of Remote Sensing*, **18**, 3569–3582.