International Journal of Remote Sensing

Predicting forest successional stages using multitemporal Landsat imagery with forest inventory and analysis data

W. Liu [a]; C. Song [b]; T. A. Schroeder [c]; W. B. Cohen [d]

[a] Department of Geography and Planning, The University of Toledo, Toledo, OH 43606, USA [b] Department of Geography, CB# 3220, 205 Saunders Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA [c] Department of Forest Science, Oregon State University, Corvallis, OR 97331, USA [d] Forestry Sciences Laboratory, Pacific Northwest Research Station, USDA Forest Service, Corvallis, OR 97331, USA

PLEASE SCROLL DOWN FOR ARTICLE

Predicting forest successional stages using multitemporal Landsat imagery with forest inventory and analysis data

W. LIU*†, C. SONG‡, T. A. SCHROEDER§ and W. B. COHEN¶

†Department of Geography and Planning, The University of Toledo, Toledo, OH 43606, USA

‡Department of Geography, CB# 3220, 205 Saunders Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

§Department of Forest Science, Oregon State University, Corvallis, OR 97331, USA

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

Forest succession is an important ecological process that has profound biophysical, biological and biogeochemical implications in terrestrial ecosystems. Therefore, information on forest successional stages over an extensive forested landscape is crucial for us to understand ecosystem processes, such as carbon assimilation and energy interception. This study explored the potential of using Forest Inventory and Analysis (FIA) plot data to extract forest successional stage information from remotely sensed imagery with three widely used predictive models, linear regression (LR), decision trees (DTs) and neural networks (NNs). The predictive results in this study agree with previous findings that multi-temporal Landsat Thematic Mapper (TM) imagery can improve the accuracy of forest successional stage prediction compared to models using a single image. Because of the overlap of spectral signatures of forests in different successional stages, it is difficult to accurately separate forest successional stages into more than three broad age classes (young, mature and old) with reasonable accuracy based on the age information of FIA plots and the spectral data of the plots from Landsat TM imagery. Given the mixed spectral response of forest age classes, new approaches need to be explored to improve the prediction of forest successional stages using FIA data.

1. Introduction

Forest succession is an important ecological process that has profound biophysical, biological and biogeochemical implications in terrestrial ecosystems. Changes in canopy structure associated with forest succession can regulate the amount of available sunlight reaching the forest floor, which has been used as a primary factor in simulating species composition with forest succession models (Botkin *et al.* 1972, Shugart and West 1977). Both the amount of solar radiation (Song and Band 2004) and the amount of carbon assimilated by a forest stand depend on its current successional stage (Law *et al* 2001, Chen *et al.* 2002). Although the biophysical mechanisms causing net primary production (NPP) to decrease with successional stage are not well agreed upon (Yoder *et al.* 1994, Mencuccini and Grace 1996,

*Corresponding author. Email: dliu@utnet.utoledo.edu

Ryan and Yoder 1997, Hunt *et al.* 1999), the occurrence of decreasing NPP with age has been observed consistently from stand to biome scales (Brown and Lugo 1990, Murty *et al.* 1996, Pregitzer and Euskirchen 2004). As 80% of above- and 40% of below-ground carbon is located in forest ecosystems (Dixon *et al.* 1994), improving our ability to accurately predict spatial distributions of forest successional stages at the landscape scale is important for enhancing our understanding of the role that forest ecosystems play in the global carbon cycle (Birdsey *et al.* 1993, Turner *et al.* 1995, Goodale *et al.* 2002, Song and Woodcock 2003*a*).

A traditional approach to determining a forest's successional stage is through field-based mensuration. One of the drawbacks of the fieldwork-based approach is that it can only provide successional stage information for a limited number of stands. Combining field plots with remotely sensed data, however, provides an opportunity to extrapolate forest successional stage information across a continuous landscape. In fact, predicting forest successional stage distributions from remote sensing, particularly with Landsat Thematic Mapper (TM) imagery, is a very active research topic (Peterson and Nilson 1993, Cohen *et al.* 1995, Jakubauskus 1996, Kimes *et al.* 1996, Song *et al.* 2002). Earlier studies have shown that reasonable accuracy can be obtained using a single satellite image to separate forests into a few broad successional stages with regression analysis (Cohen *et al.* 1995), traditional image classifications (Hall *et al.* 1991, Jakubauskus 1996) and neural network (NN) models (Kimes *et al.* 1996). Although multitemporal satellite images have been used to predict forest successional stages (Foody *et al.* 1996, Lucas *et al.* 2002), the images in these studies were used independently. Song *et al.* (2002) provided a theoretical basis that multitemporal Landsat imagery enhances the accuracy in mapping forest successional stages compared to a single imagery. More recently, Song *et al.* (2007) demonstrated that multiple Landsat TM images can be used simultaneously with multiple regression analysis to improve the accuracy of predicting forest successional stages in western Oregon. As the benefit of using multiple images with other predictive approaches is unclear, we aimed to address this issue in our current study.

An additional difficulty in extracting forest successional stages from remotely sensed data is the limited availability of ground-collected forest inventory data. One of the most comprehensive sources of forest inventory data is collected by the US Forest Services' Forest Inventory and Analysis (FIA) programme. Collecting tree and plot attributes over a systematic grid of field plots, FIA data are available for nearly all forest lands found in the conterminous USA. Tree- and plot-level data from the FIA are available in a database form that contains detailed forest successional stage information. The potential of combining FIA plot data with multitemporal Landsat imagery to map forest successional stages is not well understood. Thus, extrapolating forest successional stage information from FIA plots to large areas with remote sensing would be of great value. Recently, Song *et al.* (2007) showed that multiple regression analysis using up to four Landsat images as explanatory variables can be used to predict FIA successional stage classes in western Oregon. Although significant predictive models were created, the overall R^2 values, an indicator of how well a regression model fits the data, were relatively low. During the past decade, NNs and decision trees (DTs) have been widely used for predicting complex, nonlinear relationships between forest attributes and remotely sensed images (Fang and Liang 2003). Compared to

conventional image classification approaches, NNs and DTs do not require data to be normally distributed and can reduce the impact of unwanted noise to maximize the generalization ability of the predictive model (Quinlan 1993, Haykin 1998). Researchers have successfully used NN and DT models for remote sensing image classification, as well as land-use/land-cover change detection (Benediktsson *et al.* 1990, Heemann and Khazenie 1992, Carpenter *et al.* 1997, Friedl and Brodley 1997, Huang and Jensen 1997, Friedl *et al.* 1999, Gopal *et al.* 1999, Liu *et al.* 2001, Seto and Liu 2003, Liu *et al.* 2004). This study explored the utility of three types of widely used models, linear regression (LR), DTs and NNs, for predicting FIA forest successional stage classes with multitemporal Landsat TM imagery.

2. Study area and data

The study area was defined as Landsat WRS-2, path 46 row 29, which covers portions of the western Cascades and Coastal Range mountains of western Oregon. Spectral data from four near-anniversary Landsat 5 TM images, collected on 4 August 1984, 7 July 1991, 31 July 1994 and 23 July 1997, were used in the study. Noise effects due to differences in sun angle and phenology were minimal as the images were acquired close to the anniversary date (Song and Woodcock 2003b). The atmospheric effects on images were removed with the modified dense dark vegetation (MDDV) approach (Song *et al.* 2001), which is a modified version of the DDV approach developed by Liang *et al.* (1997). It was found to be an effective approach for removing the atmospheric effects in this area due to the abundance of mature, dark vegetation (Song and Woodcock 2003b). As there are significant information redundancies in the original six reflective bands, all TM images were transformed to brightness, greenness and wetness using the tasseled cap transformation coefficients of Crist (1985).

Forest stand age classes were derived from the Pacific Northwest (PNW)-FIA Integrated Database version 1.4 (Hiserote and Waddell 2004). In western Oregon, the FIA ground data were collected between 1995 and 1997. Stand age was coded into 22 classes (table 1). Because FIA plot locations are not in the public domain, the spectral data for each plot within our study area were extracted through special arrangement with the PNW Research Station, USDA Forest Service. The digital numbers (DNs) of Landsat images for each FIA plot in our study area were extracted using the average of a 3×3 pixel window, with the centre of the window corresponding to the reference coordinate. Because the enhancement of multi-temporal Landsat imagery in mapping forest successional stages was initially proposed based on conifer stands (Song *et al.* 2002), we limited our analysis here to only conifer dominated plots. Although a total of 2441 FIA plots fell within our Landsat path/row 46/29 study area, only 1317 of those plots were classified as coniferous forest. We regrouped the 22 age classes from table 1 into broader groups of three [young (age $\leqslant$ 49 years), mature (age 50−149 years) and old (age $\geqslant$ 150 years)] and five [very young (age $\leqslant$ 19 years), young (age 20−49 years), mature (age 51−149 years), old (age 150−299 years) and very old (age $\geqslant$ 300 years)] successional stage classes. As forest succession is a continuous process, the separation of the continuous process into discrete stages is somewhat subjective. For this reason, we based our successional stage classes on age groupings that have been successfully classified in previous studies conducted in western Oregon (Cohen *et al.* 1995, Song *et al.* 2002, 2007).

Table 1. Forest stand age classes as coded in the integrated database version 1.4 compiled by the PNW FIA programme. Ground data were collected during 1995–7 for western Oregon.

Age class	Stand ages (years)
1	0–9
2	10–19
3	20–29
4	30–39
5	40–49
6	50–59
7	60–69
8	70–79
9	80–89
10	90–99
11	100–109
12	110–119
13	120–129
14	130–139
15	140–149
16	150–159
17	160–169
18	170–179
19	180–189
20	190–199
21	200–300
22	300+

3. Method

3.1 *Predictive models*

3.1.1 Linear regression (LR). A general multiple LR model was used to study the relationship between the spectral response of FIA plots and their corresponding successional stages as:

$$y = b_0 + b_1 x_1 + \ldots + b_n x_n + \varepsilon \tag{1}$$

where b_0, b_1, …, b_n are regression coefficients that describe the rate and direction of change for a plot in spectral space. The independent variables, x_1, …, x_n, include the tasseled cap brightness, greenness and wetness indices and the changes in these indices between 1984 and 1997. The independent variables may be from a single image or multitemporal images. The regression error is ε. The dependent variable, y, is stand age class. Theoretically, y has to be a continuous numerical variable. Although stand age in the real world is a continuous variable, it is recorded in the FIA dataset as an incremental age class. As the actual age is approximately the age class times 10, the age class label is a numerically meaningful variable, thus it can be treated as a continuous variable in equation (1).

3.1.2 Multilayer perceptron neural network (MLP NN). The MLP NN, one of the most popular and successful NN architectures, is suited to a wide range of applications such as classification, pattern recognition, interpolation, prediction, forecasting, and process modelling (Haykin 1998). An MLP NN comprises a number of identical units organized in layers, with those on one layer connected to those on the next layer so that the outputs of one layer are fed-forward as inputs to the next layer.

The usage of MLP NN includes two steps: training and prediction. MLP NNs are typically trained using a supervised training algorithm known as 'backpropagation'. It is an error-based learning process. The learning process with the backpropagation algorithm has three phases. In the first phase, an input vector is presented to the network, which leads to the activation of the network as a whole. This generates a difference (error) between the output of the network and the desired output. The second phase computes the error between the output value and the desired value for each output unit and propagates it successively back through the network (error backward pass). The last phase computes the changes for the connection weights by feeding the sum of squared errors from the output layer back through the hidden layers to the input layer. This process continues until the connection weights in the network have been adjusted so that the error between output value and target value converges to an acceptable level. The second stage is to apply the trained network to new data to make a prediction.

For an MLP NN, many parameters need to be set up to obtain an appropriate performance. To simplify this process, we designed a three-layer-perceptron network for all neural models in this study (figure 1), each of which has a different number of hidden neurons. We adapt the MLP NN implemented in the WEKA data mining system (Witten and Frank 2005) for model learning and testing. In this study, the input is a vector that consists of the tasseled cap brightness, greenness and wetness indices for the FIA plots and the output is the corresponding forest successional stages.

3.1.3 Decision trees (DTs). DT algorithms were originally developed by Quinlan (1993) and have been widely used in many different applications, including remote sensing image processing (Friedl and Brodley 1997, Huang and Jensen 1997, Friedl *et al.* 1999). A DT is a logical model represented as a binary (two-way split) tree that shows how the value of a target (dependent) variable can be predicted by using the values of a set of predictor (independent) variables. The structure of a typical DT is shown in figure 2. The rectangular boxes shown in the tree are called 'nodes', each of
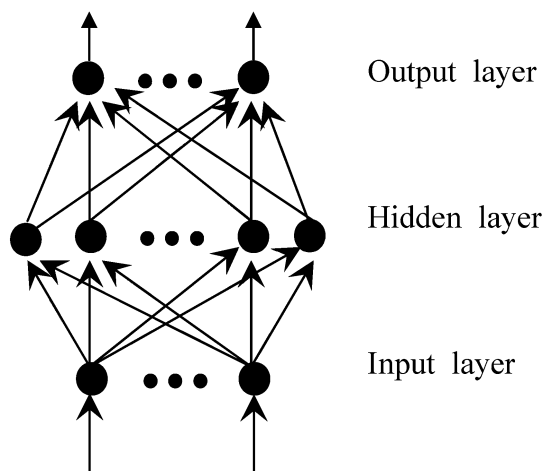

Figure 1. Three-layer fully connected multilayer perceptron (MLP) neural network. The number of neurons at input layer is equal to the dimension of the input vector. The network will have *n* output neurons if it is designed to distinguish *n* classes. The input layer and output layer are connected by a hidden layer where patterns in the input layer are recognized.
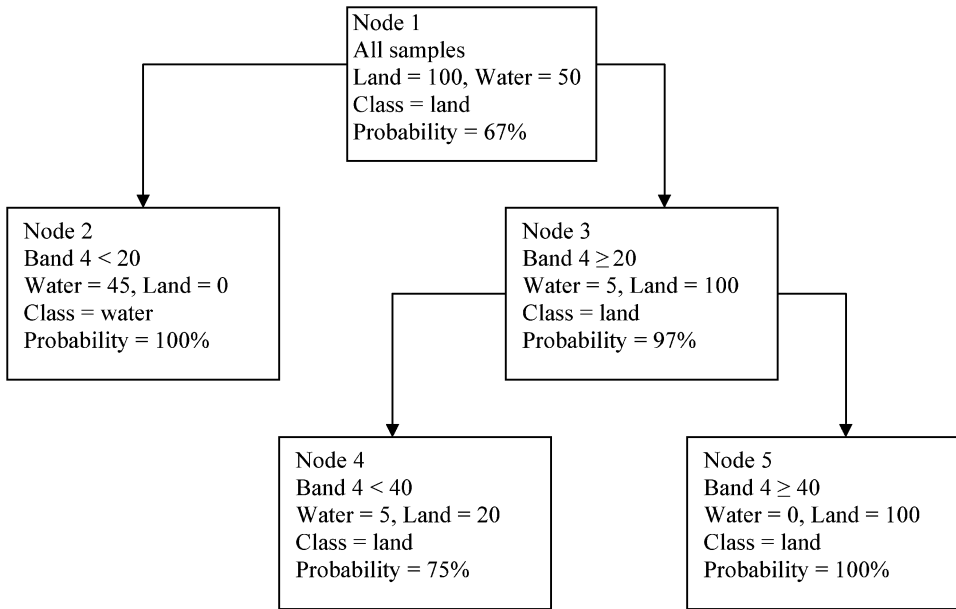
Figure 2. Classification scheme of decision trees.

which represents a set of samples from the original dataset. Nodes that have child nodes (nodes 1 and 3 in the example) are called 'interior' nodes. Nodes that do not have child nodes (nodes 2, 4 and 5 in the example) are called 'terminal' or 'leaf' nodes. The topmost node (node 1 in the example) is called the 'root' node.

The divide-and-conquer approach is widely used to construct a DT for a classification problem. First, the DT algorithm selects a variable and places it at the root node. The selected attribute is determined by a certain statistic, such as entropy. Then the DT applies a binary split that divides the samples in the root node into two groups (child nodes). The same procedure is repeated recursively to split the child nodes. The splitting process terminates when it reaches either of the two conditions: (1) a few samples remain in a node; or (2) the standard deviation of the samples in a node is just a small fraction (say, less than 5%) of that of the original sample. DT models commonly overfit the data, which can limit its predictive potential. In the current study we used a pruning technology that merges some of the lower leaves or nodes to avoid the overfitting problem. Once a DT is created, we can use it to predict the target value for specific samples where only the predictor variables are known. Similar to LR and NN models, DTs use the spectral data from Landsat to predict the FIA successional age classes.

3.2 *Feature combination with multitemporal Landsat imagery*

To fully understand how multitemporal Landsat images affect the predictive accuracy of forest successional stages, we created six feature combinations as shown in table 2. The 1997 image was used as the base image as the time of the image acquisition best matches the time of FIA data collection. Therefore, the first feature combination consists of the tasseled cap brightness, greenness and wetness indices from the 1997 image. We continued adding additional images to create new feature combinations. The last feature combination includes the tasseled cap brightness,

Table 2. List of feature combinations for training and testing in predicting forest successional stages using multiple regression, MLP NN and DT models.

Feature	Independent variables
1	BGW of 1997
2	BGW of 1997 and 1994
3	BGW of 1997, 1994 and 1991
4	BGW of 1997 and 1984
5	BGW of 1997, 1994, 1991 and 1984
6	BGW of 1997, 1994, 1991 and 1984, changes in BWG between 1984 and 1997

B, brightness; G, greenness; W, wetness.

greenness and wetness indices from all four images, as well as the changes in these indices between the 1984 and 1997 images. The feature combinations presented in table 2 were tested using all three predictive models described above.

3.3 *Model calibration and validation*

To evaluate more efficiently and accurately the performance of each model, we used a fivefold cross-validation technique, where the data are split five times, each time into a training (calibration) and testing (validation) set. Samples are used only once for training and testing purposes (Zhu and Rohwer 1996) and at no time are samples used simultaneously for training and testing. Each split is done such that 80% of the data are used for training, and the remaining 20% of the data are used for testing, allowing for a true out-of-sample evaluation of the model. When testing is complete, we base our evaluation of each model on the predictions from all five testing samples.

During the study, we found several FIA plots whose recorded successional stage appeared to be in disagreement with the observed spectral response recorded by Landsat. The primary cause of this disagreement is probably the result of changes in ground cover occurring directly before or after image acquisition. Based on past experience (Song *et al.* 2002, 2007), we manually removed plots whose positions in the spectral space were in obvious disagreement with the age class information recorded by the FIA. To fully evaluate each predictive model we report our results for both the original and cleaned datasets.

4. Results

4.1 *Analysis of FIA and remote sensing data*

The analysis of the mean tasseled cap brightness, greenness and wetness indices extracted over the FIA plots shows that none of the predictive variable has a normal distribution (table 3). The skewness indicates that the spectral variables are misaligned either to the left or right with differing magnitudes. The correlation analysis found that the relationships between successional age classes and the independent variables are weak, particularly for the 1991, 1994 and 1997 wetness, and the changes in greenness. This indicates that the LR model may have limited accuracy in predicting forest age with the spectral variables in table 2.

To better understand how forest stand age and spectral signals are related, we plotted all stands in the brightness and greenness spectral space constructed with the 1997 image (figure 3). The spectral data for each FIA sample is plotted as a symbol

Table 3. Results of statistical analysis for independent variables and their correlation coefficients with stand ages.

Variable	Mean	Standard deviation	Skewness	Correlation with age
B84	0.2256	0.0909	2.088	−0.51
G84	0.1753	0.0634	0.691	−0.50
W84	−0.0551	0.0430	−1.991	0.33
B91	0.2620	0.0943	0.425	−0.56
G91	0.2093	0.0730	0.513	−0.56
W91	−0.0528	0.0422	−1.886	0.15
B94	0.2148	0.0777	0.775	−0.53
G94	0.1707	0.0612	0.658	−0.50
W94	−0.0455	0.0385	−2.198	0.22
B97	0.2065	0.0734	0.546	−0.55
G97	0.1782	0.0592	0.504	−0.54
W97	−0.0380	0.0327	−2.305	0.16
ΔB84−97	−0.0191	0.0650	−3.294	0.10
ΔG84−97	0.0029	0.0471	1.129	−0.02
ΔW84−97	0.0171	0.0407	0.853	−0.22

B, brightness; G, greenness; W, wetness. B84 means brightness value of year 1984; ΔB84−97 indicates change in brightness between 1984 and 1997 images; and so on.

that represents one of the three successional stages. While each stage appears to occupy a dominate region of the spectral space (e.g. lower left corner for old growth, upper right region for young stands and middle area for mature stands), it is obvious that there is not a clear boundary where each stage is distinctly separate from the others. As a result, accurate prediction of forest successional stage with spectral information alone appears to be a difficult pattern recognition problem. There are multiple sources of errors that may contribute to the noisy pattern observed in figure 3, including, but not limited to, the accuracy of the age classes reported by the FIA, the natural variation of stands with similar ages located in different geographic positions, georegistration and atmospheric variations in the Landsat images, as well as errors associated with FIA plot coordinates. All these sources of error contribute to the noisy relationship between spectral data and forest stand age, leading to erroneous predictions of the successional stage classes.

4.2 *Multitemporal image classification*

We tested the three predictive models with six feature combinations in table 2 to understand the effect of using multiple Landsat images to predict forest successional stages. To test model performance, we aggregated the 22 age classes in the FIA dataset into three classes for use with the DT and MLP NN predictive models. As regression requires the dependent variable to be continuous, we preserved the variability in the ages for the regression model. We first built an LR model with 22 forest age classes. Then the predicted age classes were aggregated into the same three age classes for comparison with DT and MLP NN. The results in figure 4 indicate that the NN models consistently outperformed the DT models, although the latter produced higher accuracy than the LR models when using all six feature combinations. For all three methods tested, the prediction of forest successional stages with a single image (year 1997) shows the lowest accuracy. The predictive accuracy increases as each additional image is added. Both the NN and LR models produce the highest accuracy (65.2% and 48.5%) when using feature 6 (table 2). One
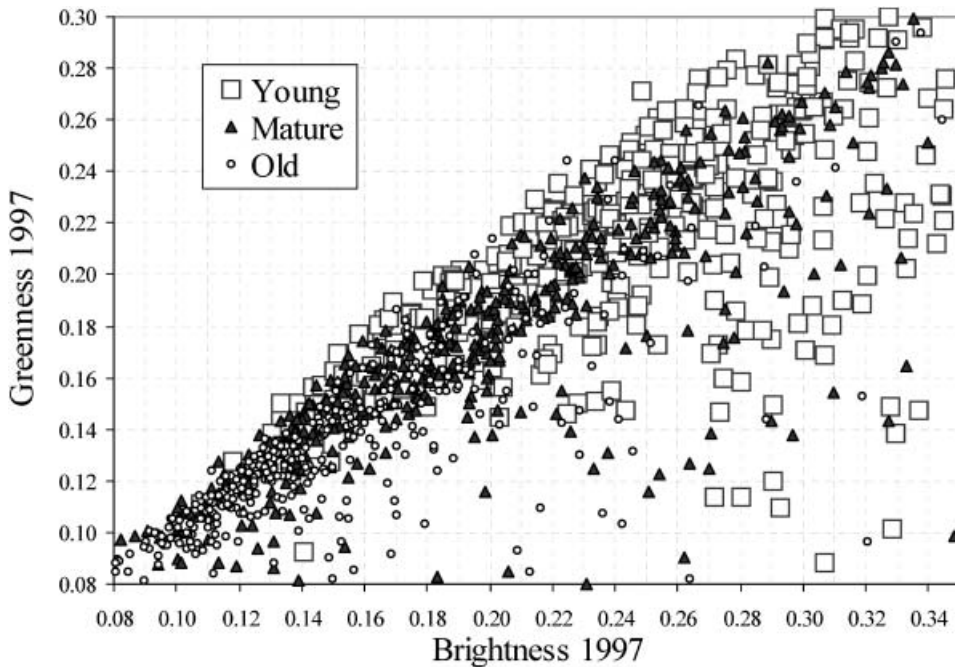
Figure 3. Distribution of FIA plots in the brightness/greenness space of 1997 image as they are aggregated to three successional stages: young (age<50 years), mature (age $51-149$ years) and old (age>150 years).

interesting finding is that the NN models built on features 4 and 5 (table 2) show similar performances, indicating that after the 1997 and 1984 images are used, the 1991 and 1994 images do not contribute much in improving the prediction of forest successional stages. Because forest succession is a gradual process, once the 1984 and 1997 images are used, the 1991 and 1994 images in the middle have little new information to add, particularly for mature and old-growth stands. Overall, the DT model using feature 4 achieved the highest predictive accuracy (61.2%) of any of the models tested. This clearly shows the enhanced value of using 1984 and 1997 images simultaneously in prediction forest successional stages compared to using any one of them alone.

4.3 Age class resolution

As multitemporal images can improve predictive performance, we used feature 6 in table 2 to assess the effects of age class resolution on prediction of forest successional stage. This analysis is similar to section 4.2, except the LR, DT and MLP NN predictive models are used with the features from table 2 to predict two aggregated groupings of the forest successional classes (i.e. the three and five class groups described earlier). The predictive accuracy of the three methods (figure 5) clearly indicates that it is not practical to separate forest successional stages into 22 classes using Landsat TM images. None of the predictive models can accurately predict all 22 successional stage classes. Not surprisingly, the accuracy improves dramatically when successional stage classes are grouped into more general categories. The highest overall accuracy was achieved when predicting three broad successional
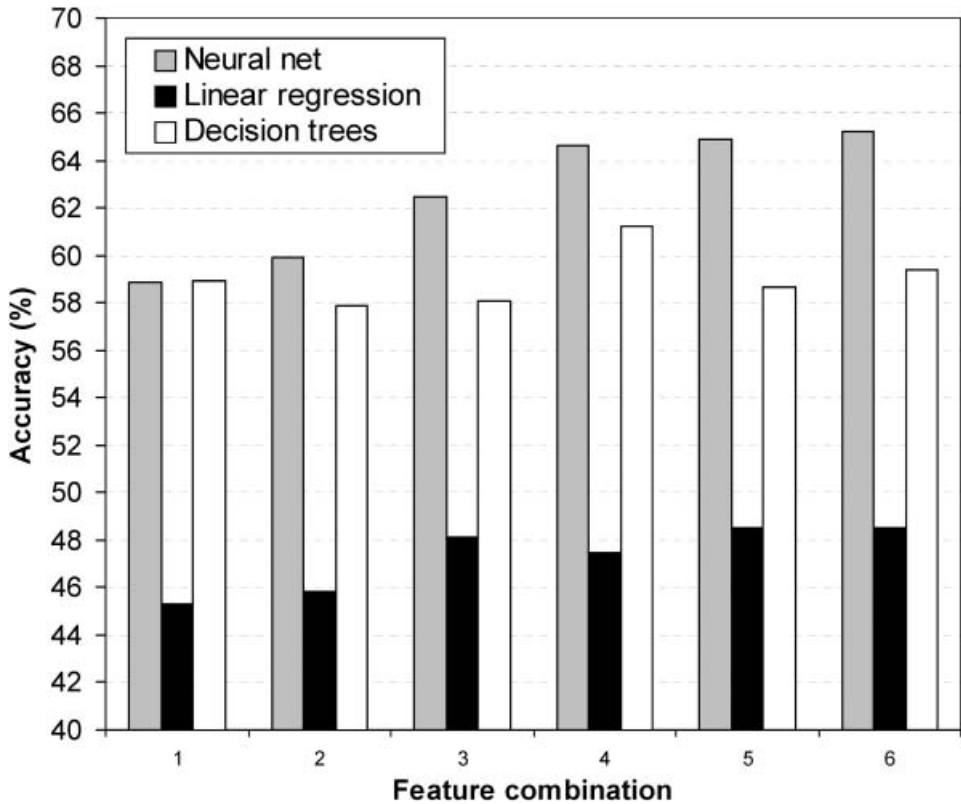
Figure 4.   Predictive accuracy of models with six feature combinations in table 2 and three successional stages: young (age<50 years), mature (age 51−149 years) and old (age>150 years).

classes. Similarly, Cohen *et al*. (1995) concluded that categorizing forest successional stages into more than three classes significantly reduced classification accuracy.

In this sense, it is likely that accuracy is reduced given the distribution of the independent variables is not normal and the noise level in the data set is high. The adjusted R^2 value of the LR model is relatively low (0.39), indicating a weak overall relationship between forest age and the spectral information captured by Landsat. As the DT and MLP NN models do not require data to be normally distributed, they typically perform better than linear models when data are noisy. This is apparent as our results indicate that both the MLP NN and DT models do a better job predicting forest successional stages (figure 4) from the relatively noisy Landsat data.

4.4   Error analysis

One crucial reason for the relatively low accuracy in predicting forest successional stages in this study is the nature of the dataset itself. There is a significant amount of spectral overlap among the different successional stage classes (figure 3). To better understand the forest successional stage prediction capability of the MLP NN model and the nature of the spectral characteristics of the FIA dataset, we analysed the error matrix based on feature 6 in table 2. The error matrix for the NN model is
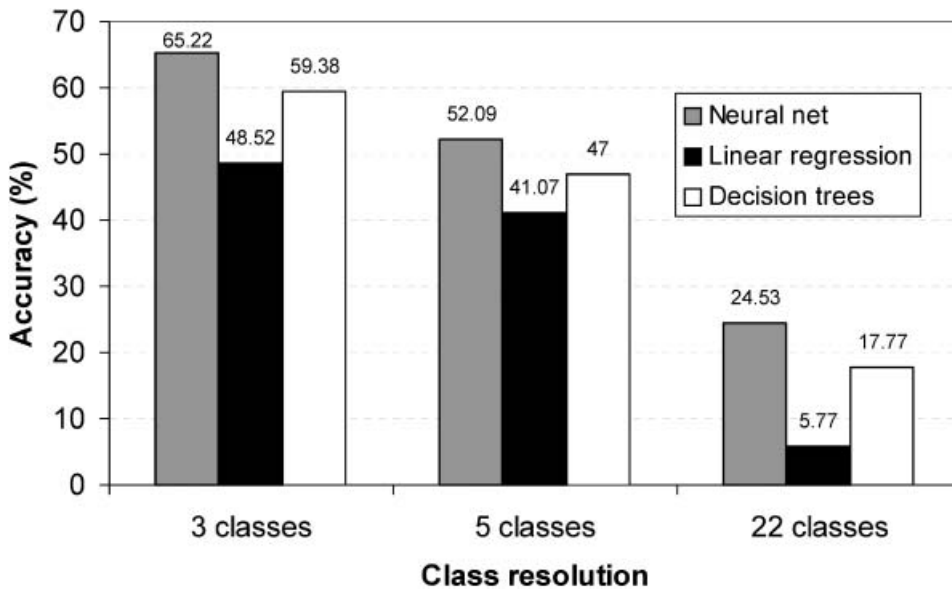
Figure 5.   Predictive accuracy for FIA plots in 3, 5 and 22 classes for feature 6 in table 2.

shown in table 4. With total accuracy of 65% and kappa statistics of 0.48, the model predicted young and old stands more accurately than mature stands. Approximately 48% of mature stands were misclassified as either young or old stands. The spectral separation between young and old stands is fairly good, resulting in fewer prediction errors from one class to another. The error matrix of DTs with feature 6 in table 2 is shown in table 5. Again we see higher user's and producer's accuracies for young and old stands, but lower accuracies for mature stands. A similar pattern is seen in the predictive results based on the LR model in table 6, but with a much lower overall accuracy and kappa statistic.

To understand the nature of the error, we analysed the prediction results at the finest class resolution (22 classes). Within the prediction results of the NN model built on feature 6, each misclassified sample was counted based on its original class label (22 classes). Figure 6 shows the distribution of the sample data through each of the 22 age classes and the corresponding percentage of misclassified samples for each class. As shown in figure 6, a higher percentage of misclassification occurs at the class transition stages (age classes 5, 6, 7 and 8 represent the transition from young to mature; age classes 14, 15 and 16 represent the transition from mature to old). Relatively low misclassification rates occur for very young and very old stands.

Table 4.   Error matrix of MLP NN prediction (three successional stages) with the feature 6 list in table 2. The total accuracy is 0.65 with the kappa statistic 0.48.

	Class 1	Class 2	Class 3	Total	Producer's accuracy
Class 1	326	113	23	462	0.71
Class 2	108	231	102	441	0.52
Class 3	44	68	302	414	0.73
Total	478	412	427	1317	
User's accuracy	0.68	0.56	0.71		

Table 5. Error matrix of DTs prediction (three successional stages) with the feature 6 list in table 2. The total accuracy is 0.59 with the kappa statistic 0.39.

	Class 1	Class 2	Class 3	Total	Producer's accuracy
Class 1	289	133	40	462	0.63
Class 2	110	212	119	441	0.48
Class 3	45	88	281	414	0.68
Total	444	433	440	1317	
User's accuracy	0.65	0.49	0.64	0.59	

Table 6. Error matrix of LR prediction (three successional stages) with the feature 6 list in table 2. The total accuracy is 0.49 with the kappa statistic 0.23.

	Class 1	Class 2	Class 3	Total	Producer's accuracy
Class 1	98	359	5	462	0.21
Class 2	46	329	66	441	0.75
Class 3	15	187	212	414	0.51
Total	159	875	283	1317	
User's accuracy	0.62	0.38	0.75	0.49	

This finding matches the natural characteristics regarding the spectral response of forests. Stands with similar successional stages have similar canopy structure, and thus have similar spectral signatures. The spectral similarity of stands in different successional stage classes results in erroneous predictions.

Applying the models to the cleaned FIA dataset with feature 6 from table 2, we found that the overall prediction accuracy of the three successional stage classes
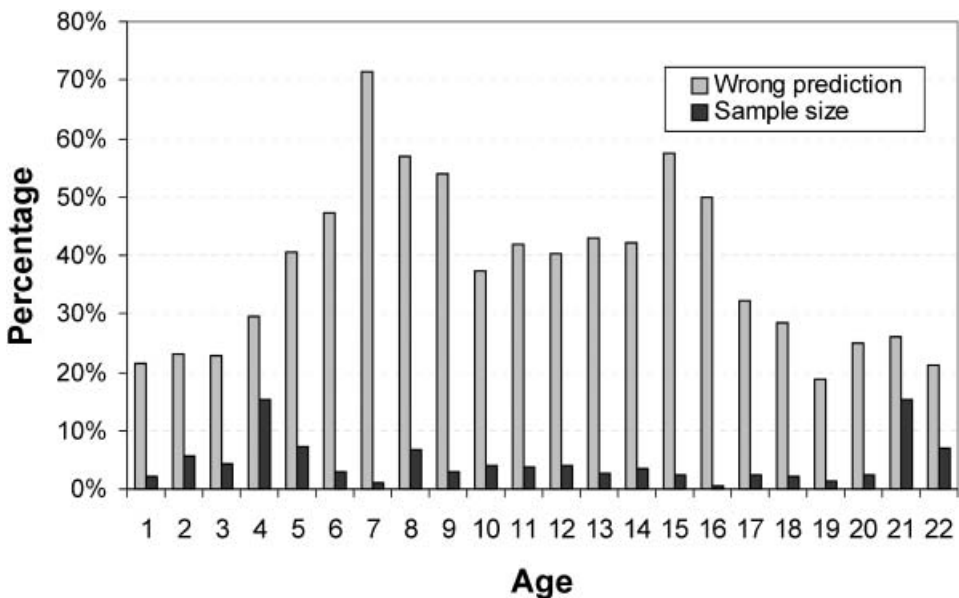

Figure 6. Distribution of errors and sample size with 22 age classes. The grey bars indicate the percentage of wrong prediction of the samples within the age class. The black bars indicate the percentage of sample size of the total plots.

improved from 65% to 71% for MLP NN (table 7), from 59% to 64% for DT (table 8), and from 49% to 55% for LR (table 9). The most dramatic improvement in both the user's and producer's accuracies was observed for young stands as the cleaning removed those plots that are labelled as mature or old stands in the FIA dataset but are young in the image as a result of change on the ground.

5. Discussion

5.1 *The importance of forest age in ecosystem functions*

Ecosystem functions are closely related to forest successional stages, particularly in the forest's ability to sequestrate carbon from the atmosphere (Harmon *et al.* 1990, Chen *et al.* 2002, Pregitzer and Euskirchen 2004), and its value in protection of other species (Dobson *et al.* 1997) and conservation of fresh water resources (Jones and Grant 1996, Watson *et al.* 1999). Young fast-growing forests, as in our study area, have high potential to sequester large amounts of carbon (Song and Woodcock 2003*a*) whereas old-growth forests typically store very large amounts of carbon. Although their capacity of additional carbon storage may be limited, disturbance of old-growth forests can lead to the rapid release of large quantities of carbon to the atmosphere that require many years to recover to equilibrium (Harmon *et al.* 1990, Korner 2003). Old-growth forests also provide suitable habitat for species such as

Table 7. Error matrix of MLP NN prediction (three successional stages) with the feature 6 list in table 2 and cleaned dataset. The total accuracy is 0.71 with the kappa statistic 0.56.

	Class 1	Class 2	Class 3	Total	Producer's accuracy
Class 1	271	64	17	352	0.77
Class 2	54	195	98	347	0.56
Class 3	12	58	263	333	0.79
Total	337	317	378	1032	
User's accuracy	0.80	0.62	0.70		

Table 8. Error matrix of DT prediction (three successional stages) with the feature 6 list in table 2 and cleaned dataset. The total accuracy is 0.64 with the kappa statistic 0.46.

	Class 1	Class 2	Class 3	Total	Producer's accuracy
Class 1	260	66	26	352	0.74
Class 2	79	157	111	347	0.45
Class 3	21	68	244	333	0.73
Total	360	291	381	1032	
User's accuracy	0.72	0.54	0.64		

Table 9. Error matrix of LR prediction (three successional stages) with the feature 6 list in table 2 and cleaned data set. The total accuracy is 0.55 with the kappa statistic 0.33.

	Class 1	Class 2	Class 3	Total	Producer's accuracy
Class 1	122	226	4	352	0.35
Class 2	31	242	74	347	0.70
Class 3	0	125	208	333	0.62
Total	153	593	286	1032	
User's accuracy	0.80	0.41	0.73		

the spotted owl and red tree vole that are listed as threatened or endangered species (Lamberson *et al.* 1992). Thus, accurate maps of forest successional stage classes over a large area would be of great value for planning environmental protection and conservation efforts. Such maps, however, are rarely available. Although the accuracy of our prediction of forest successional stages is far from satisfactory, it is a major step forwards as both the training data and the imagery are widely available.

5.2 *Complexity of the data set*

Our study revealed that noise in the training data is a major source of error in predicting forest successional stages. In spectral space, data samples of different forest ages are heavily mixed, causing difficulty in separating forests into distinct age classes. This may be caused by several factors. One might be errors in recording forest ages during field data collection. This can add noisy information to the data and further mislead models causing serious misclassification problems. A second factor is the timing of field data collection in relation to the ground-cover changes associated with forest thinning or clearcutting. Differences between significant ground-cover changes in relation to FIA and image acquisition will add noise to the spectral signature associated with forest age. In addition, there are large natural variations in the spectral response for stands of similar ages. Spectral differences arise when stands of similar age and canopy structure are located at different geographic positions (i.e. slope, aspect). These differences result in added noise, further complicating the prediction of forest successional stage classes. It is not surprising, then, that the prediction accuracy is lower for forest successional stages than for traditional land-cover classifications, which typically focus on predicting ground features with distinctly different spectral responses (e.g. forest vs. urban). Although the highest prediction accuracy in our study is around 65% (which is typically lower than most land-cover classifications), the results demonstrate the characteristics of FIA data and show the potential of mapping forest successional ages over a large area with Landsat data. Our study also demonstrates how prediction accuracy can be improved by cleaning FIA data before model development.

5.3 *Model characteristics*

LR is a popular statistical tool and can generate robust empirical predictive models. Its popularity may be attributed to the interpretability of model parameters and ease of use. However, LR models assume normal distribution of the dependent variables with respect to independent variables. This assumption may not always be true for a given dataset. In addition, LR models are not appropriate for use when relationships are nonlinear in nature. Thus, the use of both MLP NNs and DTs are appealing for a number of reasons:

- They are rich and flexible nonlinear systems that show robust performance in dealing with noisy data and have the ability to discover hidden patterns automatically from input data.
- They may be better suited than linear modelling systems to predict outcomes when the relationships between the variables are complex, multidimensional and nonlinear, as found in the complex and noisy forest successional stage dataset.
- There is no need for assumptions about data distribution, such as normality.

The construction of NN architecture, however, is a complex 'trial–test' process and there are no set methods for it. Another limitation of NN models is that the contribution of each variable to the model cannot be easily calculated and presented as they are in regression models. NN analysis generates weights that are implicitly stored inside the network architecture. The lack of interpretability at the level of individual variables (predictors) is one of the most criticized features in NN models. Compared to MLP NNs, DTs have several advantages:

- They are simple to understand and interpret.
- They use a white box model. If a given result is provided by a model, the explanation for the result is easily replicated by simple mathematics.
- Symbolic rules (IF-THEN like) can be automatically extracted from the learned trees for better explanation of the predictive outcomes and decision-making process.

All the advantages of NN and DT models favour their use over linear models for predicting forest successional stages. There is, however, no single algorithm that performs better than all other algorithms. To this end, there is room for much more work to be done before a definite conclusion can be reached.

6. Conclusions

Our study revealed the complexity of predicting forest successional stages with optical satellite imagery acquired by the Landsat sensor. While there is great potential in using FIA plot data to train models for predicting broad successional stages of conifer forests, it seems that overlap in spectral space will probably preclude the prediction of finer groupings of successional stage classes with optical imagery. Our results indicate that multitemporal remotely sensed Landsat imagery improves the predictive accuracy with all three models used in this study: LR, NN and DT. Because of the spectral overlap and the non-normal distribution of the independent variables, LR models had the lowest overall predictive accuracy. DT and MLP NN models, however, were more successful at predicting forest successional stages as they do not require data with normal distributions and can more effectively handle the noise commonly associated with the spectral response of forests sampled over broad age class continuums. In addition, cleaning data to remove plots with obvious successional stage/spectral mismatches will help to ensure that FIA data are best utilized during model development, ultimately resulting in improved predictive accuracy of forest successional stage classes.

References
BENEDIKTSSON, H., SWAIN, P. and ERSOY, O., 1990, Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, pp. 540–551.
BIRDSEY, R.A., PLANTINGA, A.J. and HEATH, L.S., 1993, Past and prospective carbon storage in United States forests. *Forest Ecology and Management*, **58**, pp. 33–40.

BOTKIN, D.B., JANAK, J.F. and WALLIS, J.R., 1972, Rationale, limitations, and assumptions of a northeastern forest growth simulator. *IBM Journal of Research and Development*, **16**, pp. 101–116.

BROWN, S. and LUGO, A.E., 1990, Tropical secondary forests. *Journal of Tropical Ecology*, **6**, pp. 1–32.

CARPENTER, G., GJAJA, M., GOPAL, S. and WOODCOCK, C., 1997, ART neural network for remote sensing: vegetation classification from Landsat TM and terrain data. *IEEE Transactions on Geoscience and Remote Sensing*, **30**, pp. 308–325.

CHEN, W.J., CHEN, J.M., PRICE, D.T. and CIHLAR, J., 2002, Effects of stand age on net primary productivity of boreal black spruce forests in Ontario, Canada. *Canadian Journal of Forest Research*, **32**, pp. 833–842.

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, USA. *International Journal of Remote Sensing*, **16**, pp. 721–746.

CRIST, E.P., 1985, A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, **17**, pp. 301–306.

DOBSON, A.P., BRADSHAW, A.D. and BAKER, A.J.M., 1997, Hopes for the future: restoration ecology and conservation biology. *Science*, **277**, pp. 515–522.

DIXON, R.K., BROWN, S., HOUGHTON, R.A., SOLOMON, A.M., TREXLER, M.C. and WISNIEWSKI, J., 1994, Carbon pools and flux of global forest ecosystems. *Science*, **263**, pp. 185–190.

FANG, H. and LIANG, S., 2003, Retrieving leaf area index with a neural network method: simulation and validation. *IEEE Transactions on Geoscience and Remote Sensing*, **40**, pp. 2052–2062.

FRIEDL, M.A. and BRODLEY, C.E., 1997, Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, **61**, pp. 399–409.

FRIEDL, M.A., BRODLEY, C.E. and STRAHLER, A.H., 1999, Maximizing land cover classification accuracies at continental to global scales. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, pp. 969–977.

GOODALE, C.L., APPS, M.J., BIRDSEY, R.A., FIELD, C.B., HEATH, L.S., HOUGHTON, R.A., JENKINS, J.C., KOHLMAIER, G.H., KURZ, W., LIU, S.R., NABUURS, G.J., NILSSON, S. and SHVIDENKO, A.Z., 2002, Forest carbon sinks in the northern hemisphere. *Ecological Applications*, **12**, pp. 891–899.

GOPAL, S., WOODCOCK, C.E. and STRAHLER, A.H., 1999, Fuzzy neural network classification of global land cover from a 1 degree AVHRR data set. *Remote Sensing of Environment*, **67**, pp. 230–243.

FOODY, G.M., PALUBINSKAS, G., LUCAS, R.M., CURRAN, P.J. and HONZAK, M., 1996, Identifying terrestrial carbon sinks: classification of successional stages in regenerating tropical forest from Landsat TM data. *Remote Sensing of Environment*, **55**, pp. 205–216.

HALL, F.G., BOTIN, D.B., STREBEL, D.E., WOODS, K.D. and GOETZ, S.J., 1991, Large-scale patterns of forest succession as determined by remote sensing. *Ecology*, **72**, pp. 628–640.

HARMON, M.E., FERRELL, W.K. and FRANKLIN, J.F., 1990, Effects on carbon storage of conversion of old-growth forests to young forests. *Science*, **247**, pp. 699–702.

HAYKIN, S., 1998, *Neural Networks: A Comprehensive Foundation* (New York, NY: Prentice Hall).

HEEMANN, P. and KHAZENIE, N., 1992, Classification of multispectral remote sensing data using a back-propagation neural network. *IEEE Transaction on Geoscience and Remote Sensing*, **30**, pp. 81–88.

HISEROTE, B. and WADDELL, K., 2004, *The PNW-FIA Integrated Database User Guide Version 1.4* (Portland, OR: Forest Inventory and Analysis Program, Pacific Northwest Research Station).

HUANG, X. and JENSEN, J.R., 1997, A machine learning approach to automated knowledge-base building for remote sensing image analysis with GIS data. *Photogrammetric Engineering and Remote Sensing*, **63**, pp. 1185–1194.

HUNT, E.R., LAVIGNE, M.B. and FRANKLIN, S.E., 1999, Factors controlling the decline of net primary production with stand age for balsam fir in Newfoundland assessed using an ecosystem simulation model. *Ecological Modelling*, **122**, pp. 151–164.

JAKUBAUSKAS, M.E., 1996, Thematic Mapper characterization of lodgepole pine seral stages in Yellowstone National Park, USA. *Remote Sensing of Environment*, **56**, pp. 118–132.

JONES, J.A. and GRANT, G.E., 1996, Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon. *Water Resources Research*, **32**, pp. 959–974.

KIMES, D.S., HOLBEN, B.N., NICKESON, J.E. and MCKEE, W.A., 1996, Extracting forest age in a pacific northwest forest from Thematic Mapper and topographic data. *Remote Sensing of Environment*, **56**, pp. 133–140.

KORNER, C., 2003, Slow in, rapid out – carbon flux studies and Kyoto targets. *Science*, **300**, pp. 1242–1243.

LAMBERSON, R.H., MCKELVEY, R., NOON, B.R. and VOSS, C., 1992, A dynamic analysis of northern spotted owl viability in a fragmented forest landscape. *Conservation Biology*, **6**, pp. 505–512.

LAW, B.E., THORNTON, P.E., IRVINE, J., ANTHONI, P.M. and VAN TUYL, S., 2001, Carbon storage and fluxes in ponderosa pine forests at different developmental stages. *Global Change Biology*, **7**, pp. 755–777.

LIANG, S., FALLAH-ADL, H., KALLURI, S., JAJA, J., KAUFMAN, Y.J. and TOWNSHEND, J.R.G., 1997, An operational atmospheric correction algorithm for Landsat Thematic Mapper imagery over the land. *Journal of Geophysical Research*, **102**, pp. 17173–17186.

LIU, W., SETO, K., WU, E., GOPAL, S. and WOODCOCK, C., 2004, ART-MMAP: a neural network approach to subpixel classification. *IEEE Transactions on Geoscience and Remote Sensing*, **42**, pp. 1976–1983.

LIU, W., GOPAL, S. and WOODCOCK, C., 2001, Spatial data mining for classification, visualization and interpretation with ARTMAP neural network. In *Data Mining for Scientific and Engineering Applications*, R. Grossman (Ed.), pp. 205–222 (The Netherlands: Kluwer Academic).

LUCAS, R.M., HONZAK, M., DO AMARAL I., CURRAN, P.J. and FOODY, G.M., 2002, Forest regeneration on abandoned clearances in central Amazonia. *International Journal of Remote Sensing*, **23**, pp. 965–988.

MENCUCCINI, M. and GRACE, J., 1996, Hydraulic conductance, light interception and needle nutrient concentration in Scots pine stands and their relations with net primary productivity. *Tree Physiology*, **16**, pp. 459–468.

MURTY, D., MCMURTRIE, R.E. and RYAN, M.G., 1996, Declining forest productivity in aging forest stands: a modeling analysis of alternative hypotheses. *Tree Physiology*, **16**, pp. 187–200.

PETERSON, U. and NILSON, T., 1993, Successional reflectance trajectories in northern temperate forests. *International Journal of Remote Sensing*, **14**, pp. 609–613.

PREGITZER, K.S. and EUSKIRCHEN, E.S., 2004, Carbon cycling and storage in world forests: biome patterns related to forest age. *Global Change Biology*, **10**, pp. 2052–2077.

QUINLAN, J., 1993, *C4.5: Programs for Machine Learning* (San Mateo, CA: Morgan Kaufmann).

RYAN, M.G. and YODER, B.J., 1997, Hydraulic limits to tree height and tree growth. *BioScience*, **47**, pp. 235–242.

SETO, K.C. and LIU, W., 2003, Comparing ARTMAP neural network with maximum-likelihood classifier for detecting urban change. *Photogrammetric Engineering and Remote Sensing*, **69**, pp. 981–990.

SHUGART, H.H. and WEST, D.C., 1977, Development of an Appalachian deciduous forest succession model and its application to assessment of the impact of the chestnut blight. *Journal of Environmental Management*, **5**, pp. 161–179.

SONG, C., WOODCOCK, C.E., SETO, K.C., PAX-LENNEY, M. and MACOMBER, S.A., 2001, Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment*, **75**, pp. 230–244.

SONG, C., WOODCOCK, C.E. and LI, X., 2002, The spectral/temporal manifestation of forest succession in optical imagery: the potential of multitemporal imagery. *Remote Sensing of Environment*, **82**, pp. 285–302.

SONG, C. and WOODCOCK, C.E., 2003a, The importance of landuse history and forest age structure in regional terrestrial carbon budgets. *Ecological Modelling*, **64**, pp. 33–47.

SONG, C. and WOODCOCK, C.E., 2003b, Monitoring forest succession with multitemporal Landsat images: factors of uncertainty. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 2557–2567.

SONG, C. and BAND, L.E., 2004, MVP: a model to simulate the spatial patterns of photosynthetically active radiation under discrete forest canopies. *Canadian Journal of Forest Research*, **34**, pp. 1192–1203.

SONG, C., SCHROEDER, T.A. and COHEN, W.B., 2007, Predicting temperate conifer forest successional stage distributions with multitemporal Landsat Thematic Mapper imagery. *Remote Sensing of Environment*, **106**, pp. 228–237.

TURNER, D.P., KOERPER, G.J., HARMON, M.E. and LEE, J.J., 1995, A carbon budget for forests of the conterminous United States. *Ecological Applications*, **5**, pp. 421–436.

WATSON, F.G.R., VERTESSY, R.A. and GRAYSON, R.B., 1999, Large-scale modelling of forest hydrological processes and their long-term effect on water yield. *Hydrological Processes*, **13**, pp. 689–700.

WITTEN, I. and FRANK, E., 2005, *Data Mining: Practical Machine Learning Tools and Techniques* (San Francisco, CA: Morgan Kaufmann).

YODER, B.J., RYAN, M.G., WARING, R.H., SCHOETLE, A.W. and KAUFMANN, M.R., 1994, Evidence of reduced photosynthetic rates in old trees. *Forest Science*, **40**, pp. 513–527.

ZHU, H. and ROHWER, R., 1996, No free lunch for cross-validation. *Neural Computation*, **8**, pp. 1421–1426.