

Extracting Forest Age in a Pacific Northwest Forest from Thematic Mapper and Topographic Data

D. S. Kimes,^{*} B. N. Holben,^{*} J. E. Nickeson,[†] and W. A. McKee[‡]

The feasibility of extracting forest age of young stands (< 50 yr) in a Pacific Northwest Forest using Landsat Thematic Mapper (TM) spectral bands and topographic information was explored using a neural network approach. Understanding the changes of forest fragmentation through time are important for assessing alterations in ecosystem processes (forest productivity, species diversity, nutrient cycling, carbon flux, hydrology, spread of pests, etc.) and wildlife habitat and populations. The study area was the H.J. Andrews Experimental Forest on the Blue River Ranger District of the Willamette National Forest in western Oregon. Timber harvesting has occurred in this forest over the past 45 years and has a recorded forest management history. The study area was extracted from a georeferenced TM scene acquired on 7 July 1991. A coincident digital terrain model (DTM) derived from digital topographic elevation data was also acquired. Using this DTM and an image processing software package, slope and aspect images were generated over the study area. Sites were chosen to cover the entire range of forest stand age and slope and aspect. The oldest recorded clearcut stands were logged in 1950. A number of sites were chosen as primary forest which had no recorded history of cutting. Various feed-forward neural networks trained with back propagation were tested to predict forest age from TM data and topographic data. The results demonstrated that neural networks can be used as an initial model for inferring forest age. The best network was a 6 → 5 → 1 structure with inputs of TM

Bands 3, 4, and 5, elevation, slope and aspect. The rms values of the predicted forest age were on the order of 5 years. TM Bands 1, 2, 6, and 7 did not significantly add information to the network for learning forest age. Furthermore, the results suggest that topographic information (elevation, slope, and aspect) can be effectively utilized by a neural network approach. The results of the network approach were significantly better than corresponding linear systems.

INTRODUCTION

The extensive old-growth forests of western Oregon have been subjected to widespread cutting during the past 45 years. This has created patches of recent clearcuts, young plantations, older second-growth stands, and remnant mature and old-growth forests (Spies and Franklin, 1988). Understanding the changes of forest fragmentation through time are important for assessing alterations in ecosystem processes (forest productivity, species diversity, nutrient cycling, carbon flux, hydrology, spread of pests, etc.) and wildlife habitat and populations (Ripple et al., 1991; Iverson et al., 1994; Cohen et al., 1995). It is desirable to inventory and monitor such highly fragmented areas using a high resolution sensor such as the Landsat Thematic Mapper (TM) with a nominal resolution of 28.5 m (Iverson et al., 1994). Studies have predicted broad age classes of Pacific Northwest conifer forests (< 80 yr, 80–200 yr, and > 200 yr; Cohen et al., 1995) with accuracies on the order of 73% using TM data. Fiorella and Ripple (1993) showed strong relationships between TM band values and the age of carefully selected, homogeneous, Douglas fir stands (2–35 years old). It is desirable to obtain forest age (continuous variable) on a per pixel basis in old-growth forests of western Oregon that have undergone

^{*} NASA/Goddard Space Flight Center, Biospheric Sciences Branch, Greenbelt, MD

[†] Hughes STX Corporation, Lanham, MD

[‡] Department of Forest Science, Oregon State University, Corvallis, OR

Address correspondence to Dan S. Kimes, NASA/GSFC, Biospheric Sciences Br., Code 923, Greenbelt, MD 20771.

Received 23 June 1995; revised 25 September 1995.

widespread clearcutting and planting. Specifically, in this study the feasibility of extracting forest age (on a continuous basis) of young, managed stands (< 50 yr) using TM spectral bands and topographic information was explored.

The relationship between the age of the forest since clearcutting and the TM spectral return, elevation, slope and aspect is very complex. The TM spectral return is a function of the vegetation structure and optical properties of the vegetation and background components (Kimes et al., 1985; Kimes, 1991; Cohen et al., 1995). The vegetation structure and optical properties of the components are determined by the forest growth characteristics, which are a function of the elevation, slope, aspect, planting practices of the area logged, and other site quality characteristics. The topographic effect is much more complex than a simple correction for the incident angle (Cohen et al., 1995). For the "expert" to understand and accurately capture all of the above relationships in a physical model is difficult.

The development of physically based radiative scattering models that incorporate forest growth and topography, and that can be used to extract forest parameter in its infancy (e.g., as reviewed by Wang and Jarvis, 1990; Nikolov and Fox, 1994; Levine et al., 1993). Within the scientific community, one would ideally like to develop accurate physically based models for the specific system being studied which can serve as a hypothesis for our current understanding of the physical system and as a basis for extracting forest parameters from other readily known / measured parameters. However, accurate models that are invertible in this context are lacking. Neural networks offer many advantages to this developing field.

In the field of remote sensing, neural networks have been used recently for classification (and other spatiotemporal relationships), predicting continuous parameters of ground cover, and for fast and stable inversions of relatively complex physical models (e.g., Liu and Xiao, 1991; Key et al., 1989; Hewitson and Crane, 1994; Tsang et al., 1992; Heermann and Khazene, 1992; Bischof et al., 1992; Dawson et al., 1994; Paola and Schowengendt, 1994; Pierce et al., 1993). These studies show that neural network approaches often perform significantly better than traditional techniques. Neural networks have the ability to learn patterns or relationships given training data, and to generalize or abstract results from imperfect data (e.g., data with noise or data with a few incorrect values; Anderson and Rosenfeld, 1988; Wasserman, 1989; Zornetzer et al., 1990).

Neural networks are employed in this study because they offer the following advantages. First, neural networks can be used as an initial model. If accuracy is the major concern, a neural network may be entirely adequate and desirable. A neural network can model the forest system on the basis of a set of encoded input / output examples of the system. The network maps in-

puts to the desired output by learning the mathematical function underlying the system. Using this approach, input and output variables can be related without any knowledge or assumptions about the underlying mathematical representation.

Secondly, networks can be used to define a set of variables that are relevant to the desired parameter(s) to be inferred. If the mapping of a network is not accurate, this may indicate that some input variables are missing. Also an input variable is relevant to the problem only if it significantly increases the network's performance. Thus, networks can be used to identify relevant variables for physically-based modeling development (Fu, 1994).

Finally a network can be used as a baseline control while developing physically based models (Fu, 1994). The physical processes embedded in the model being developed can be improved if model results are not more accurate than the neural network results. The network does not assume any initial information and is viewed as a black box. Thus, according to Fu (1994), the network provides a good performance standard for evaluating a physically based model.

In this study, the feasibility of using neural networks as a method to predict forest age of young stands from TM and topographic data was tested. TM data with 28.5 m resolution is suited for the assessment of forest cover in highly fragmented landscapes (Iverson, 1994). Input variables that included TM Bands 1-7, elevation, slope, and aspect of each pixel were used to predict forest age. Feed-forward networks with a layer of input nodes, one layer of hidden nodes, and a single output node were tested. The significant input variables were identified. The resulting networks serve as a baseline control to other models being developed.

METHODS

Forest Data

The study area was the H. J. Andrews Experimental Forest on the Blue River Ranger District of the Willamette National Forest in western Oregon. The Forest is primarily in the Western hemlock (*Tsuga heterophylla*) and Pacific silver fir (*Abies amabilis*) vegetation zones (Franklin and Dyrness, 1973) and the major tree species are Douglas fir (*Pseudotsuga menziesii*), Western hemlock (*Tsuga heterophylla*), Pacific silver fir, and Western red cedar (*Thuja plicata*). The climate is maritime with wet, mild winters and dry, warm summers. The dominant species at elevations below 1000 m are Douglas fir, Western hemlock, and Western red cedar, with noble fir (*Abies procera*), Pacific silver fir, and mountain hemlock (*Tsuga mertensiana*) becoming important at higher elevations. Timber harvesting has occurred in this Forest over the past 45 years.

The Landsat (V) TM scene, acquired on 7 July

Table 1. Range of Each Variable for Training and Testing Data

Variable	Training Data	Testing Data
Year logged	1945-1989	1945-1989
Elevation	482-1524 m	482-1537 m
Slope	0-41°	0-41°
Aspect	0-358°	0-358°

1991 (WRS 46/29), was precision terrain corrected by Hughes STX's Satellite Mapping Technologies. A precision corrected product is one that has been geocorrected using ground control points coupled with orbit modeling software to remove systematic distortion and achieve geodetic alignment. A precision *terrain* corrected product uses an additional input of 3-arc-second digital topographic elevation data (DTED) to remove image parallax and achieve a product with orthoimage properties. Both the TM bands and a digital terrain model were resampled to a 25-m UTM projection. The digital terrain model, derived from a spline interpolation of a 3-arc-second DTED product was used with an image processing software package to generate slope and aspect images over the study area.

Sites were chosen to cover the entire range of forest stand age and slope and aspect. The number of sites for training and testing were 67 and 40, respectively. A number of contiguous pixels interior to each site were selected. The resulting number of pixels for training and testing were 6593 and 3555, respectively. The range of stand age, elevation, slope, and aspect are shown in Table 1. The oldest recorded clearcut stands were cut in 1950. A number of sites which were recently clearcut (< 10 years) were assigned a cut age of 1989 because the exact dates of these cuts were not recorded in the data. A number of sites were chosen in primary forest which has no recorded history of cutting. These mature and old-growth stands were arbitrarily assigned a age of 1945 for the purpose of training and testing the net.

Neural Networks

A backpropagation network was used in this study. It is a multilayer feed-forward network with a learning rule known as backpropagation as described in detail by Fu (1994). A network with an input layer, hidden layer, and output layer is shown in Figure 1. In this study the input layer was connected to a hidden layer which was connected to a single output node. The input nodes are fully connected to the nodes in the hidden layer, which in turn are fully connected to the output node. All processing flows from the input nodes through the hidden layer to the output node. The network structure is denoted in this article as: #input-nodes → hidden-nodes → output-nodes (i.e., a 4 → 3 → 1 network would have 4 input nodes, 3 hidden nodes, and 1 output node). No feedback processes are allowed. The backpropagation algorithm (McClelland and Rumelhart, 1988) was

used to train the weights of the network. This three-layered network was chosen because in most cases this structure is sufficient to solve the problem of inferring a continuous parameter (Fu, 1994).

A significant amount of time was spent in selecting a set of input variables in order to achieve a high accuracy of the net. The available input variables in this study were TM Bands 1-7, elevation, aspect, and slope. The output node was the year that the pixel was logged. In addition, information was available on the planting history of the area logged. In general, the site preparation was broadcast burning, and the areas were planted with Douglas fir and a few areas were planted with noble fir. On some areas two or three replantings were performed. From a remote sensing perspective, this information would not be known and thus was not used as an input variable in the initial networks. However, various networks were explored later using some of this planting information as an input variable to see if it explained some of the error in the performance of the networks. A planting index was used in this study to capture part of this information. If the area was planted, then the year of planting minus the year of logging was recorded. The first planting date was always used.

A summary of the equations used for feed-forward networks with the backpropagation learning algorithm follows. The nomenclature is similar to that of Sigillito and Hutton (1990). The three layer network is described here (Fig. 1). The network has n_i input nodes, n_h hidden nodes, and n_o output nodes. The training set has p input-target pairs, (X^p, T^p) , $p = 1, \dots, n_p$. The p th input-target pair is represented as the vectors $(X^p = x_1^p, \dots, x_{n_i}^p)$ and $(T^p = t_1^p, \dots, t_{n_o}^p)$. The input, hidden, and output levels are referred to levels 0, 1, and 2, respectively. Equations for inputs, outputs, and weights use the symbols I , O , and w , respectively. Superscripts refer to the level as well as the pair number.

The input nodes pass their inputs to the nodes above them with no computations. Thus, the output of the i th input node when the p th input vector is introduced to the network is $O_i^{p,0} = x_i^p$.

The input to the j th hidden node is

$$I_j^{p,1} = \sum_{i=1}^{n_i+1} w_{j,i}^1 O_i^{p,0} \quad (1)$$

The $n_i + 1$ rather than just n_i accounts for the bias term associated with the j th hidden node. The bias is represented as another weight w_{j,n_i+1}^1 that is learnable. The input to this fictitious bias node is always 1.0. The output of the j th hidden node is

$$O_j^{p,1} = f(I_j^{p,1}) \quad \text{or} \quad O_j^{p,1} = f\left(\sum_{i=1}^{n_i+1} w_{j,i}^1 O_i^{p,0}\right) \quad (2)$$

The f is the activation function. In this study, the function is the sigmoid or logistic function

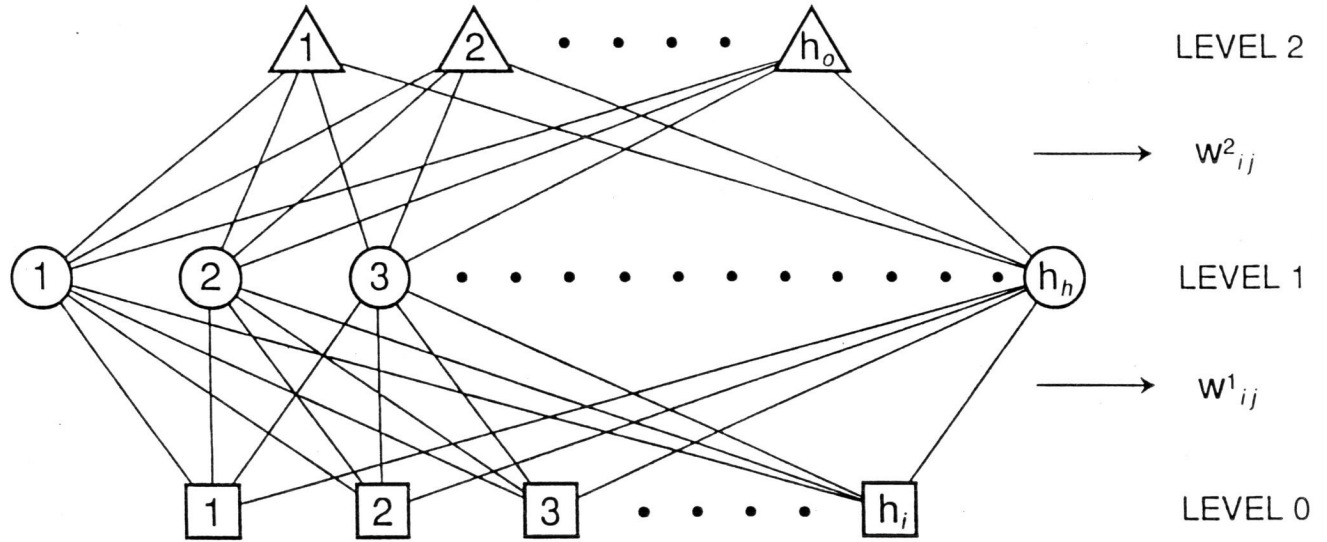


Figure 1. Neural network with three layers where the squares represent n_i input nodes, the circles represent n_h hidden nodes, and the triangles represent n_o output nodes, w^2_{ij} represents the weight matrix applied to the inputs to the output nodes and the w^1_{ji} represents the weight matrix applied to the inputs to the hidden nodes.

$$f(x) = \frac{1}{1 + e^{-x}}$$

The same type of equations are applied to the next level. The input to the j th output node is

$$I_j^{p,2} = \sum_{i=1}^{n_h+1} w_{ji}^2 O_i^{p,1} \quad (3)$$

The $n_h + 1$ rather than just n_h accounts for the bias term associated with the j th output node. The bias is represented as another weight w_{j,n_h+1}^1 that is learnable. The input to this fictitious bias node is always 1.0. The output of the j th output node is

$$O_j^{p,2} = f(I_j^{p,2}) \quad \text{or} \quad f\left(\sum_{i=1}^{n_h+1} w_{ji}^2 O_i^{p,1}\right) \quad (4)$$

Backpropagation is used to derive the weights and biases of the network. The back propagation equations were popularized by Rumelhart et al. (1986). The error term is

$$E = \frac{1}{2} \sum_{p=1}^{n_p} \sum_{i=1}^{n_o} (T_i^p - O_i^{p,2})^2$$

Differentiating E with respect to w_{ji}^2 and w_{ji}^1 gives us a component of the gradient of E . To minimize the error, the following learning equations are used in practice for the weights at level 2 and 1 (t is the number of iterations):

$$w_{ji}^2(t+1) = w_{ji}^2(t) + \eta \sum_{p=1}^{n_p} \delta_j^{p,2} O_i^{p,1} + \alpha \Delta w_{ji}^2(t) \quad (5)$$

where

$$\delta_j^{p,2} = (T_j^p - O_j^{p,2}) O_j^{p,2} (1 - O_j^{p,2}),$$

and

$$w_{ji}^1(t+1) = w_{ji}^1(t) + \eta \sum_{p=1}^{n_p} \delta_j^{p,1} O_i^{p,0} + \alpha \Delta w_{ji}^1(t) \quad (6)$$

where

$$\delta_j^{p,1} = O_j^{p,1} (1 - O_j^{p,1}) \sum_{m=1}^{n_o} w_{mj}^2 \delta_m^{p,2}$$

Neural networks can be trained to both underfit and overfit the data as with traditional methods of interpolation. High training error (or inability to converge) is characteristic of underfitting. High testing error is characteristic of overfitting. One would like a close correspondence between the training and testing errors. The size of the network (number of hidden nodes in this case) was adjusted until the two errors were similar.

Each element of the input and output vectors for all samples were scaled (between 0 and 1). This normalized the inputs and outputs to certain dynamic range for transfer functions used in training the net.

In general, 10 runs each with a different random number were used to populate the initial weights. The different initial weights showed very little difference in the stability of the final results. Various α (momentum term) and η (learning rate) values were tested. Generally, 200,000–300,000 epochs were run until the errors converged to a relatively constant value.

RESULTS AND DISCUSSION

A few networks tested for inferring the forest age are shown in Table 2. Of the networks tested, the one that was chosen as the best (6 → 5 → 1 network in Table 2) had the minimum number of nodes, close correspon-

Table 2. Results of Various Net Structures Using the Six Input Variables (TM Bands 3, 4, 5, Elevation, Slope and Aspect)^a

Net Structure	rms (years)		R ²	
	Train	Test	Train	Test
6 → 7 → 1	7.08	7.79	0.777	0.595
6 → 6 → 1	5.06	5.68	0.795	0.673
6 → 5 → 1**	5.12	5.59	0.788	0.686
6 → 4 → 1	5.10	5.76	0.782	0.652
6 → 3 → 1	6.56	9.02	0.726	0.492

^a The number of pixels used for training and testing were 6,593 and 3,555, respectively. The root mean square (rms) values in years and the R² values for the predicted forest age are reported.

** Best network.

dence between the training and testing accuracies, and a relatively high overall accuracy. These guidelines assure that the network neither underfits or overfits the data. The rms and R² of the predicted forest age versus the true forest for each network is shown in Table 2. The 6 → 5 → 1 network had rms values of 5.12 years and 5.59 years, respectively for the training and testing data. The R² values of this network were 0.79 and 0.69, respectively, for the training and testing data. These rms's in years corresponds to a normalized rms (divide by 44 years) of 0.117 and 0.127 for the training and testing data, respectively. The accuracy of these results are good considering that the input data is on a per pixel basis (e.g., no spatial information is used in this study).

Initially the performance of each network was poor and unstable. It was discovered that the network's ability to learn depended on the way the input data were normalized. It is common practice to normalize each input channel independently between the values of 0 and 1. This has the positive effect of each input channel having a full dynamic range between 0 and 1 relative to all other channels. Using this method of normalization, the networks had difficulty training—the networks performance was uneven and unpredictable. However, when the spectral bands were normalized as a group, the networks trained and tested well, that is, the answers were less sensitive to changes in η and α . Eberhart and Dobbins (1990) noted that in theory a network should be able to find the relationships between channels when normalization is performed independently on each channel. In practice, however, this technique of normalization sometimes makes training of the network difficult resulting in poor network performance. When the spectral bands are normalized as a group, the dynamic range of the channels is decreased; however, the relationship between two channels is maintained, and no multiplicative factors are introduced to the problem.

The behavior of the networks tested had several characteristics that are commonly reported in the literature (Eberhart and Dobbins, 1990). For example, it was

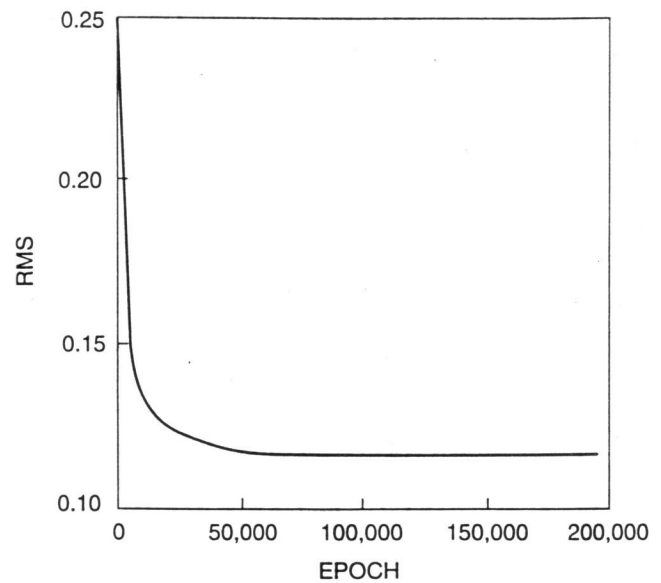


Figure 2. Root mean error square (rms) as a function of number of epochs for the 6 → 5 → 1 network in Table 2.

found that unstable answers were experienced until the η was lowered (on the order of 0.005 in this study). With the lower η the rms was stable and consistently converged to the same value (Fig. 2). The α value was set at 0.05.

Various networks were used to define a set of variables that are relevant to the problem of inferring forest age. We began with a large network that included all inputs and trained the net. The input variables were deleted one at a time. Irrelevant input variables were deleted. An input variable was considered relevant only if it significantly increased the network's performance. TM Bands 1, 2, 6, and 7 did not significantly increase the network's performance and thus were deleted as input variables. In this study the TM Band 6 (10.45–12.46 μm) did not add any significant information to the network for learning the forest age. However, Sader (1986) found that the Thermal Infrared Multispectral Scanner was capable of detecting temperature differences related to relative differences in canopy closure and green leaf area. The thermal domain is much more dependent on a multitude of environmental parameters in contrast to the optical domain, which is largely dependent only on the optical and structural properties of plant components. In this study TM Bands 3, 4, and 5 and the elevation, slope, and aspect significantly contributed to the net's performance (> 0.44 rms in years).

The number of hidden nodes were pruned until training performance began to significantly decrease. At this point further pruning of the number of hidden nodes was explored to extend the generalization of the network (Fu, 1994).

In all networks, both the training and testing statistics steadily improved with each 1000 epochs, but the

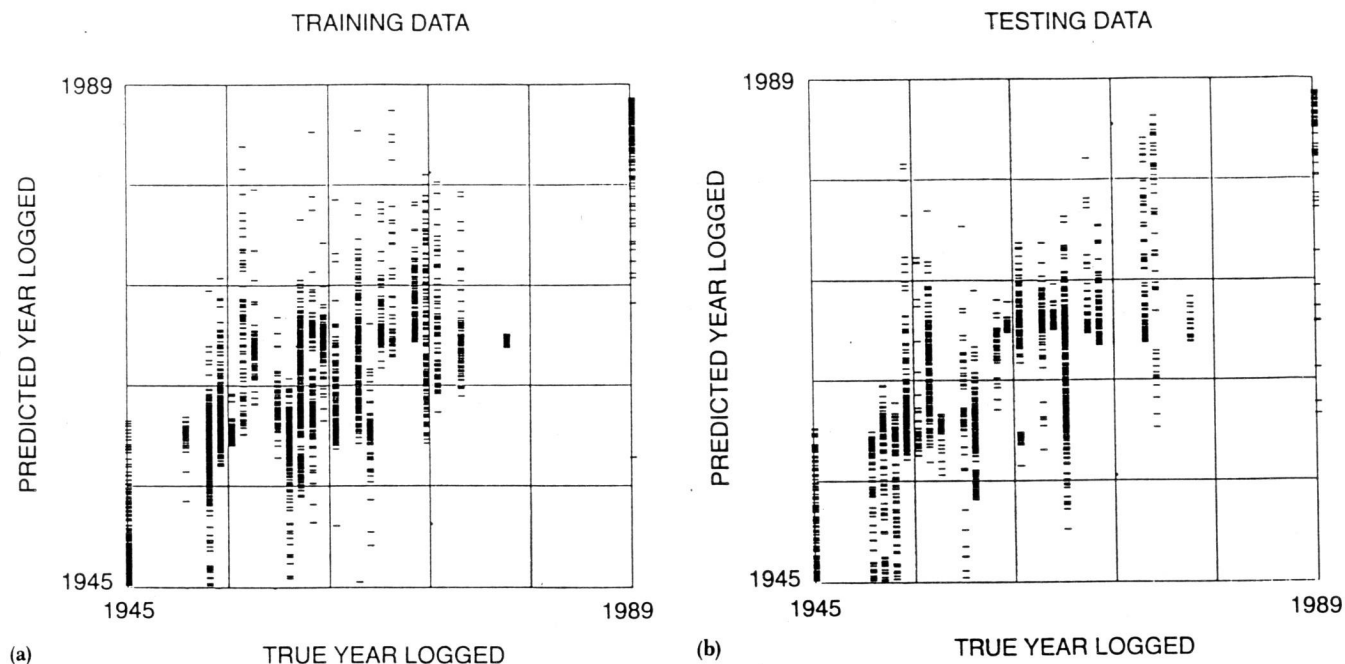


Figure 3. Predicted year logged versus the true year logged for the training (3a) and testing (3b) data. The network used was the $6 \rightarrow 5 \rightarrow 1$ network in Table 2. The number of pixels / points shown are 6593 for the training data and 3555 for the testing data. The rms and R^2 values for the training data were 5.1 and 0.80, respectively. The rms and R^2 values for the testing data were 5.6 and 0.69, respectively.

differential between the training and testing statistics always stayed relatively constant. When the number of hidden nodes were varied to try to improve generalization, the smallest differential between the training and testing data for both the rms and R^2 values occurred for the $6 \rightarrow 5 \rightarrow 1$ network (Table 2).

Figure 3 shows the predicted year logged from the $6 \rightarrow 5 \rightarrow 1$ network (Table 2) versus the true year logged from the training and testing data. The number of pixels shown for the training and testing data are 6553 and 3555, respectively. The dark areas in Figure 3 represent a relatively large number of points. It is clear from Figure 3 that the network generalizes the information it has learned from the training data to the unseen testing data relatively well.

Using only TM Bands 3, 4, and 5 with no topographic information, a $3 \rightarrow 3 \rightarrow 1$ network had a training and testing rms of 8.0 and 8.4 years, respectively. This is improved significantly, however, when the topographic information (elevation, slope, and aspect) is also used as input to create the $6 \rightarrow 5 \rightarrow 1$ network mentioned in Table 2 which had a training and testing rms of 5.1 years and 5.6 years, respectively. Thus, these results suggest that topographic information can be effectively utilized by a neural network approach.

Many studies using remotely-sensed data utilize vegetation indices. The widely used normalized difference vegetation index (NDVI) was used along with the elevation, slope, and aspect as inputs to a $4 \rightarrow 5 \rightarrow 1$ network.

NDVI is calculated as $(TM4 - TM3) / (TM4 + TM3)$. The rms for the training and testing were 6.0 years and 8.1 years, respectively. Using TM data, Sader et al., (1989) found that the NDVI was not significantly correlated with forest regeneration age classes. In comparison, the $6 \rightarrow 5 \rightarrow 1$ network (Table 2) that uses TM Bands 3, 4, 5, elevation, slope, and aspect had an rms for training and testing of 5.1 years and 5.6 years, respectively.

Multiple linear regression techniques were applied and compared to the neural network approach. The rms values for a linear system using the six variables (TM Bands 3, 4, 5, elevation, slope, and aspect) for the training and testing data were 7.1 years and 7.5 years, respectively. The respective R^2 values were 0.56 and 0.51. In this linear system the topographic information is not utilized effectively as would be expected since the topography introduces a nonlinear component to the system. For example, a linear system using only TM Bands 3, 4, and 5 and no topographic information had an rms for training and testing of 7.3 years and 7.5 years, respectively. In contrast to the rms values of the $6 \rightarrow 5 \rightarrow 1$ network (Table 2), these linear systems had rms values on the order of 35–40% higher.

Generally, in a remote sensing mission information on the planting history on a per pixel basis would not be available. However, planting history clearly has an effect on the forest growth characteristics and thus ultimately affects the vegetation structure and optical properties of the forest components. The details of the

planting practices were discussed in the Methods section. The influence of the planting index was explored for all pixels that were planted. The planting index is the year of planting minus the year of logging. The planting generally occurred within 2 years of logging (75% of the cases); however, the time of planting varied up to 16 years after logging. The accuracies of various networks that used this index as an additional input variable were explored. Only pixels that were planted were used. There were 5211 and 2762 pixels for training and testing, respectively. A $7 \rightarrow 5 \rightarrow 1$ network (inputs: TM Bands 3, 4, 5, elevation, slope, aspect, and planting index) had an rms of 4.0 years and 5.7 years for the training and testing data, respectively. In comparison, a $6 \rightarrow 5 \rightarrow 1$ network (inputs: TM Bands 3, 4, 5, elevation, slope, and aspect) without the planting index had an rms of 4.7 years and 6.0 years for the training and testing data, respectively. The planting variable seems to contain useful information. However, many other variables are not accounted for; for example, the number of replantings, the density of planting, the species used in planting, the growth potential for various sites due to edaphic (soils), and topographic conditions (elevation, slope, aspect), the variations in site preparations, etc.

CONCLUSIONS AND IMPLICATIONS

The results demonstrated that neural networks can be used as an initial model to extract forest age of young stands (< 50 yr) using TM spectral bands and topographic information. The best network was a $6 \rightarrow 5 \rightarrow 1$ structure with inputs of TM Bands 3, 4, 5, elevation, slope, and aspect. The rms values of the predicted forest age were on the order of 5 years. TM Bands 1, 2, 6, and 7 did not significantly add information to the network for learning forest age. Furthermore, the results suggest that topographic information (elevation, slope, and aspect) can be effectively utilized by a neural network approach. The results of the network approach were significantly better than corresponding linear systems.

The variability in site quality, site preparation, and planting practices for different stands introduces unknown variables in remote sensing missions. The variability occurs at many levels: site preparation after logging, planting or natural regeneration, number of replantings, dates of replanting(s), species of trees planted, density of planting, and growth potential of site, etc. In most remote sensing missions, gathering planting information on a per pixel basis is not possible. Thus, the accuracy of inferring forest age using a network will depend on the planting practices in the region being studied. For example, if the planting history varies widely, such as in this study, one would expect networks to perform similar to the $6 \rightarrow 5 \rightarrow 1$ network (Table 2) with rms values on the order of 5 years. However, in regions where the planting practices are relatively

uniform, one would expect a significantly lower rms error in predicting the forest age.

In this study the input data was from individual TM pixels. Much variation occurs at this scale within a training or test site. Consequently, it is believed in future efforts that averaging data from adjacent pixels may improve the training and testing accuracies. In addition, explicit calculations and direct measures of spectral texture can be introduced to the network as additional information which can improve the network's accuracy of prediction. Numerous calculations of texture exist in the literature. Paola and Schowengendt (1994) have shown that spectral texture can be entered into the network directly in the form of a 3×3 window and significantly improve the performance of the network.

Ultimately, the scientific community needs to develop physically based radiative scattering models that incorporate forest growth and topography. These models need to be accurate and invertible for the desired parameters. Until these activities mature, the neural network approach provides an initial model for predicting forest age. The network analysis in this study defined a set of variables that are relevant to modeling efforts designed to infer forest age. In addition, the study provides a performance standard for evaluating other models.

REFERENCES

- Anderson, J. A., and Rosenfeld, E., Eds. (1988), *Neurocomputing: Foundations of Research*, MIT Press, Cambridge, MA.
- 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., *Int. J. Remote Sens.* 16: 721-746.
- Bischof, H., Schneider, W., and Pinz, A. J. (1992), Multispectral classification of Landsat images using neural networks, *IEEE Trans. Geosci. Remote Sens.* 30:482-489.
- Dawson, M. S., Amar, F., Manry, M. T., Rawat, V., and Fung, A. K. (1994), Classification of remote sensing data using fast learning neural networks and topology selection algorithms, in *Proc. IGARSS '94*, California Institute of Technology, Pasadena, CA, 8-12 August, pp. 1410-1412.
- Eberhart, R. C., and Dobbins, R. W. (1990), Implementations, in *Neural Network PC Tools* (R. C. Eberhart and R. W. Dobbins, Eds.), Academic, New York, pp. 35-58.
- Franklin, J. F., and Dynnes, C. T. (1973), *Natural vegetation of Oregon and Washington*. Oregon State University Press, Oregon.
- Fiorella, M., and Ripple, W. J. (1993), Analysis of conifer forest regeneration using Landsat Thematic Mapper data, *Photogramm. Eng. Remote Sens.* 59:1383-1388.
- Fu, L. (1994), *Neural Networks in Computer Intelligence*, McGraw-Hill, New York, 460 pp.
- Heermann, P. D., and Khazene, N. (1992), Classification of multispectral remote sensing data using a back-propagation

- neural network, *IEEE Trans. Geosci. Remote Sens.* 30:81-88.
- Hewitson, B. C., and Crane, A. G., Eds. (1994), *Neural Nets: Applications in Geography*, The GeoJournal Library Vol. 29, Kluwer Academic, Dordrecht, 195 pp.
- Iverson, L. R., Cook, E. A., and Graham, R. L. (1994), Regional forest cover estimation via remote sensing: the calibration center concept, *Landscape Ecol.* 9:150-174.
- Key, J., Maslanik, J. A., and Schweiger, A. J. (1989) Classification of merged AVHRR and SMMR Arctic data with neural networks, *Photogramm. Eng. Remote Sens.* 55:1331-1338.
- Kimes, D. S. (1991), Radiative transfers in homogeneous and heterogeneous vegetation canopies, in *Photon-Vegetation Interaction: Applications in Optical Remote Sensing and Plant Ecology*, Chap. (R. B. Myneni and J. Ross, Eds.), Springer-Verlag, New York, Chap. 11, pp. 339-388.
- Kimes, D. S., Newcomb, W. W., Nelson, R. F., and Schutt, J. B. (1985), Directional reflectance distributions of a hardwood and pine forest canopies. *IEEE Trans. Geosci. Remote Sens.* S-24:281-293.
- Levine, E. R., Ranson, K. J., Smith, J. A., et al. (1993), Forest ecosystem dynamics: linking forest succession, soil process and radiation models, *Ecol. Model.* 65:199-219.
- Liu, Z. K., and Xiao, J. Y. (1991), Classification of remotely-sensed image data using artificial neural networks, *Int. J. Remote Sens.* 12:2433-2438.
- McClelland, J. L., and Rumelhart, D. E. (1988), *Explorations in Parallel Distributed Processing*, MIT Press, Cambridge, MA, 344 pp.
- Nikolov, N. T. and Fox, D. G. (1994), A coupled carbon-water-energy-vegetation model to assess responses of temperate forest ecosystems to changes in climate and atmospheric CO₂. Part I. Model concept, *Environ. Pollut.* 83:251-262.
- Paola, J. D., and Schowengendt, R. A. (1994), Comparisons of neural networks to standard techniques for image classification and correction, in *Proc. IGARSS '94*, California Institute of Technology, Pasadena, CA, 8-12 August, 1994, 1404-1406.
- Pierce, L. E., Dobson, M. C., Wilcox, E. P., and Ulaby, F. T. (1993), Artificial neural network inversion of tree canopy parameters in the presence of diversity, in *Proc. IGARSS '93*, Houston, Texas, pp. 394-397.
- Ripple, W. J., Bradshaw, G. A., and Spies, T. A. (1991), Measuring forest landscape patterns in the Cascade range of Oregon, *Biol. Conservation* 57:73-88.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagation errors, *Nature* 323:533-536.
- Sader, S. A. (1986), Analysis of effective radiant temperatures in a Pacific northwest forest using thermal infrared Multi-spectral Scanner data, *Remote Sens. Environ.* 19:105-115.
- Sader, S. A., Waide, R. B., Lawrence W. T., and Joyce A. T. (1989), Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM Data, *Remote Sens. Environ.* 28:143-156.
- Sigillito, V. G., and Hutton, L. V. (1990), Case study II: Radar signal processing, in *Neural Network PC Tools*, Academic, New York, Chap. 11.
- Spies, T. A., and Franklin, J. F. (1988), Old growth and forest dynamics in the Douglas-fir region of western Oregon and Washington, *Nat. Areas J.* 8:190-201.
- Tsang, L., Chen, Z., Oh, S., Marks, R. J., and Chang, A. T. C. (1992), Inversion of snow parameters from passive microwave remote sensing measurements by a neural network trained with a multiple scattering model, *IEEE Trans. Geosci. Remote Sens.* 30:1015-1024.
- Wang, Y. P., and Jarvis, P. G. (1990), Description and validation of an array model MAESTRO, *Agric. For. Meteorol.* 51:257-280.
- Wasserman, P. D. (1989), *Neural Computing: Theory and Practice*, Van Nostrand Reinhold, New York, 230 pp.
- Zornetzer, S. F., Davis, J. L., and Lau, C. (1990), *An Introduction to Neural and Electronic Networks*, Academic, New York.