ELSEVIER

A Neural Network Method for Efficient Vegetation Mapping

Gail A. Carpenter, Sucharita Gopal,† Scott Macomber,† Siegfried Martens,* Curtis E. Woodcock,† and Janet Franklin‡*

This article describes the application of a neural network method designed to improve the efficiency of map production from remote sensing data. Specifically, the ARTMAP neural network produces vegetation maps of the Sierra National Forest, in Northern California, using Landsat Thematic Mapper (TM) data. In addition to spectral values, the data set includes terrain and location information for each pixel. The maps produced by ARTMAP are of comparable accuracy to maps produced by a currently used method, which requires expert knowledge of the area as well as extensive manual editing. In fact, once field observations of vegetation classes had been collected for selected sites, ARTMAP took only a few hours to accomplish a mapping task that had previously taken many months. The ARTMAP network features fast online learning, so that the system can be updated incrementally when new field observations arrive, without the need for retraining on the entire data set. In addition to maps that identify lifeform and Calveg species, ARTMAP produces confidence maps, which indicate where errors are most likely to occur and which can, therefore, be used to guide map editing.

INTRODUCTION: VEGETATION MAPPING FROM REMOTE SENSING DATA

Vegetation maps serve a wide range of functions in the management of natural resources, including inventory assessment, fire control, wildlife habitat characterization, and water quality monitoring. In this context, a number of federal and state agencies as well as private companies with large land holdings currently use vegetation maps derived from satellite-based remote sensing (e.g., Aspinall and Veitch, 1993; Bauer et al., 1994; Cohen et al., 1995; Congalton et al., 1993; Franklin and Wilson, 1991). In one such application domain, Region 5 (California) of the U.S. Forest Service (USFS) has, for the past two decades, used Landsat sensor imagery for mapping vegetation in its 20 National Forests. Over that time, the demand for vegetation maps has increased, even as sensor technology and methods for deriving information from remote sensing images have continued to improve. The result is ongoing pressure to refine the knowledge derived from remote sensing, leading to new explorations of map production methods. This article reports on findings concerning the utility of one such new method, the ARTMAP neural network. The study compares ARTMAP capabilities with those of a conventional method on the benchmark problem of mapping vegetation in the Sierra National Forest.

The vegetation maps employed for management of National Forests in California identify basic lifeforms such as *conifer*, *hardwood*, *water*, and *barren.* Many of these lifeforms are further subdivided by species associations, and, when appropriate, by tree size and cover. The present analysis considers the problem of mapping lifeforms and species associations, with species labeled according to the California vegetation, or *Calveg*, classification system (Matyas and Parker, 1980). The mapping methodology that Region 5 of the USFS currently applies to this problem has evolved over the years into a rather cumbersome system which uses two separate processing streams (Woodcock et al., 1994) (Fig. 1a). The first stream produces lifeform maps from Landsat Thematic Mapper (TM) input data and a terrain-derived image of local solar zenith angle.

*Center for Adaptive Systems and Department of Cognitive and Neural Systems, Boston University

†Center for Remote Sensing and Department of Geography, Boston University

‡Center for Earth Systems Analysis Research and Department of Geography, San Diego State University

Address correspondence to Gail A. Carpenter, Dept. of Cognitive and Neural Systems, 677 Beacon St., Boston Univ., Boston, MA 02215. E-mail: gail@cns.bu.edu

Received 31 December 1998; revised 14 May 1999.

REMOTE SENS. ENVIRON. 70:326–338 (1999)

655 Avenue of the Americas, New York, NY 10010

0034-4257/99/$–see front matter
PII S0034-4257(99)00051-6

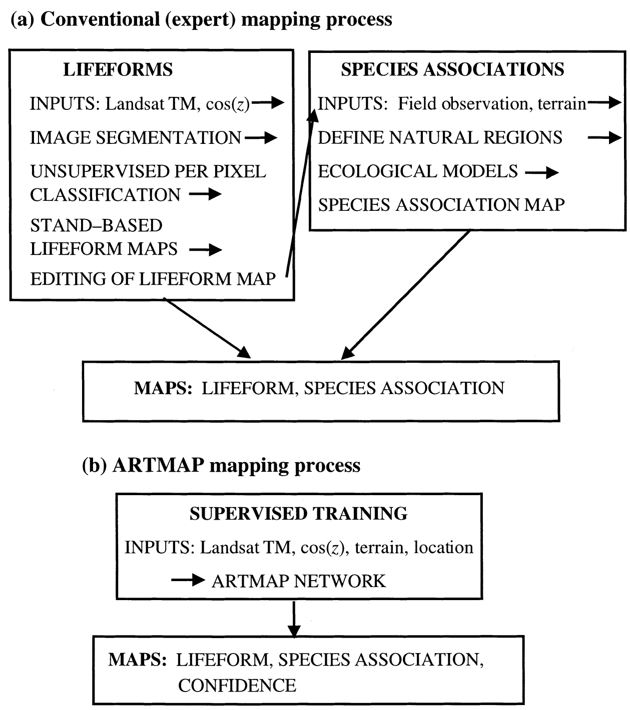


Figure 1. Mapping methods process streams comparing a) conventional (expert) and b) ARTMAP systems.

This stream includes a conventional unsupervised clustering method (the Ustats algorithm), maximum likelihood classification, and analyst labeling. The second processing stream uses field observations and terrain data (slope, aspect, elevation) to develop predictive ecological models of species associations within lifeform classes (Franklin et al., 1986). For most of the National Forests in California, application of this method has also required identification of natural regions, followed by individual calibration within each natural region of the rules that relate species associations to terrain variables.

Although the current mapping approach has been applied successfully (Woodcock et al., 1980; Franklin et al., 1986), it has several disadvantages. In particular, the unsupervised classification algorithm, which often requires several iterations, is time-consuming and inefficient; and defining the ecological models typically calls for months of expert labor. In addition, the polygon-based lifeform maps require extensive manual editing to achieve an adequate level of accuracy. In the Sierra National Forest, editing was based on both photographic interpretation of the area and field inspections, and the editing process required several additional months of work in order to produce the final expert map. In fact, labels on as many as 80,000 of the 250,000 polygons were changed during the editing phase.

The several stages of the expert mapping method have been introduced over time as prior methods, using Landsat spectral data alone, have proved inadequate for the species association task. However, the expert systems approach is intrinsically limited, since derived rules describing relationships among terrain variables and species associations are necessarily too broadly defined. In addition, this method does not use Landsat spectral data directly in mapping species associations (Fig. 1a), thus ignoring useful information in the Landsat signal. What has been lacking is an effective and computationally tractable way of combining both spectral and terrain variables for accurate, efficient image classification. The ARTMAP neural network provides that capability (Fig. 1b). This method thus allows for a greatly simplified approach to mapping lifeforms and species associations, producing accurate maps with significant savings in time, effort, and cost. The ARTMAP method, as applied to the Sierra National Forest vegetation mapping problem, will now be described.

ARTMAP NEURAL NETWORKS

Introduced relatively recently, the ARTMAP neural network (Carpenter et al., 1991; 1992) is already being used in a variety of application settings, including industrial design and manufacturing, robot sensory motor control and navigation, machine vision, and medical imaging, as well as in remote sensing (Carpenter et al., 1997; Gopal and Fischer, 1997; Gopal et al., 1999). ARTMAP belongs to the family of adaptive resonance theory (ART) networks, which are characterized by their ability to carry out fast, stable learning, recognition, and prediction, with a training procedure that requires only one pass through the data. These features differentiate ARTMAP from the family of feedforward multilayer perceptrons (MLPs), including backpropagation, which typically require slow learning and repeated presentations of the data set. MLP systems are also subject to catastrophic forgetting, whereby earlier memories are erased by subsequent training data. The inherent instability of MLP learning may make such a system unsuitable for unconstrained mapping problems with many input components or map pixels. ARTMAP systems self-organize arbitrary mappings from input vectors, representing features such as spectral values and terrain variables, to output vectors, representing predictions such as vegetation classes or environmental indices. Internal ARTMAP control mechanisms create stable recognition categories of optimal size by maximizing code compression while minimizing predictive error (Carpenter et al., 1991).

An ARTMAP Mapping Method

The ARTMAP neural network mapping method presented here automatically produces vegetation maps from spectral and terrain data. As a supervised learning system, ARTMAP is trained by example. Network performance on the Sierra National Forest mapping task was evaluated using the *cross-validation method* (Mosier, 1951), which requires that the set of testing sites be disjoint from the set of training sites. For each pixel in a training set site, the network was presented with a vector representing input variables, such as spectral band values, along with the label of the associated Calveg class of the site in which the pixel was located. During testing, the trained ARTMAP network predicted a vegetation class for each pixel. The final site-level class prediction was taken to be the one produced by the largest number of pixels in a test-set site. The seventeen Calveg classes were then merged into six lifeform classes (*conifer, hardwood, chaparral, herbaceous, water, barren*).

ARTMAP performance was compared with that of the expert method (Fig. 1a) on the Sierra National Forest mapping task. Predictive accuracy was evaluated in terms of the percentage of the test set a system classified correctly, for both lifeform and Calveg species identifications. This quantitative measure of the neural and conventional methods was augmented by visual presentations of results, namely, confusion matrices and vegetation maps. Confusion matrices summarize patterns of errors among map classes at test sites, while vegetation maps provide forest-wide spatial views of system predictions.

ARTMAP Computations

Carpenter et al. (1999) have developed an ARTMAP network for prediction of mixtures of classes, and have shown how the system is used in remote sensing applications by mapping the Plumas National Forest, in California. That article includes a self-contained ARTMAP implementation algorithm. While the general version of this algorithm predicts vegetation mixtures, the same system can predict discrete output classes, as the special case of unitary "mixtures." This classifier algorithm is the one that is applied throughout the present study.

In general, a number of ARTMAP variations have been used for solving different problems. The present system uses the following technical options for each computation: the MT– match tracking rule (Carpenter and Markuzon, 1998), a winner-take-all activation rule, a Weber law choice function, and parameter values $\bar{\rho}_a=0$ (baseline vigilance), $\varepsilon=0$ (match tracking), and $\alpha=10^{-6}$ (choice parameter). Knowing these parameter values, an investigator could readily replicate the current system by implementing the ARTMAP algorithm recently published in this journal.

THE SIERRA NATIONAL FOREST DATA SET

During training, a supervised classification system is presented with a set of input vectors and their associated output classes. For the mapping problems considered here, each input vector specifies satellite sensor data, plus terrain and geographic location information, for a given pixel; and the output class is a Calveg species label, as used by Region 5 of the USFS. During testing, a supervised classification system is required to predict output labels for inputs that were never seen during training. Labeled sites were compiled by the USFS as part of the conventional mapping process, with an initial set later augmented to cover all Calveg classes. Thus the resulting collection is not the result of a strict *a priori* sample design, but rather represents the best set of data already available at the start of the present study.

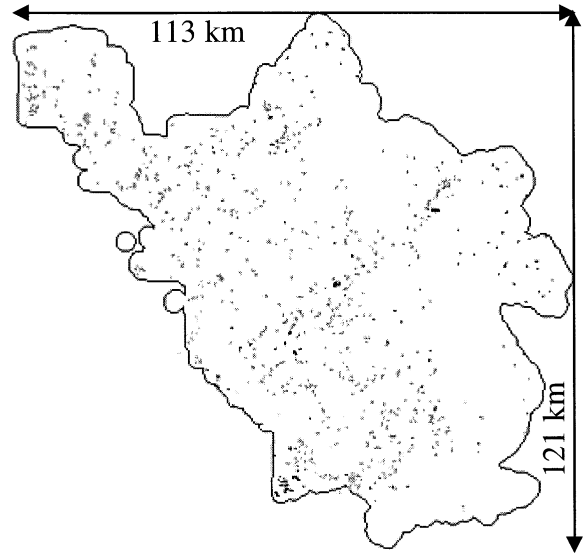


Figure 2. Location of the 1013 field observation sites in the Sierra National Forest.

Field Observation Labels

Training and test site labels specifying the vegetation (Calveg) classes were assembled by ground observation of 1013 sites in the Sierra National Forest. In all, these sites cover 59,903 pixels, which represents about 0.5% of the Forest. Spectral and terrain information for these pixels and the matching Calveg labels of the corresponding sites collectively comprise the *field observation data set.* Figure 2 indicates the location and distribution of these sites. The expert method partitioned the full map of the Sierra National Forest into a quarter million polygons and assigned a Calveg label to each. An average polygon, or site, occupied approximately 59 pixels. Thus, since Landsat pixels are 30 m×30 m, the map polygons were nominally 230 m×230 m size, on average.

Site labels from the field observation data set were not directly used in editing the expert map. Thus performance on these sites could serve as an independent standard by which to evaluate both conventional and neural methods. As a supervised learning method, ARTMAP used a portion of the field observation data set for training each network, with the standard cross-validation method ensuring that training and test sets were disjoint.

Table 1. Vegetation Classes in the 1013 Field Observation Sites

Calveg class	*# Sites*	*Lifeform*	*# Sites*
Mixed conifer pine	122	Conifer	561
Red fir	116		
Subalpine	37		
Ponderosa pine	102		
Mixed conifer fir	121		
East pond pine	22		
Lodgepole pine	41		
Black oak	49	Hardwood	213
Canyon live oak	60		
Oak diggerpine	69		
Blue oak	35		
Mixed chaparral	19	Chaparral	38
Montane chaparral	19		
Dry grass	51	Herbaceous	101
Wet meadow grass	50		
Water	50	Water	50
Barren	50	Barren	50

Both ARTMAP and the expert system labeled each site as belonging to one of six lifeform classes: *conifer, hardwood, chaparral, herbaceous, water, barren.* These six life-forms were further subdivided into seventeen Calveg classes. Table 1 lists these classes, specifying the number of sites in the field observation data set for each species association and lifeform class. The table shows that lifeform classes were unevenly sampled in the data, given that the number of sites was roughly proportional to area for each class. Conifer, and to a lesser extent hardwood, were represented far more than other lifeforms, while chaparral was poorly sampled. Sampling of individual conifer species was also uneven. Not surprisingly, sampling density tended to correlate with ARTMAP predictive accuracy, as discussed below.

Input Data: Spectral, Terrain, and Location Variables

For each pixel in the data set, up to 12 variables were available for training the neural network. Six of these were spectral variables, namely, the original digital values of TM Bands 1–5 and 7. A digital elevation model (DEM) provided four more variables: the cosine of the local solar zenith angle [$\cos(z)$], slope, elevation, and aspect (direction of slope). Two more variables specified the location of the pixel (UTM northing and easting).

Figure 3 shows grayscale maps of the forest for nine of the 12 input components. These maps illustrate the different view provided by each individual variable. To produce a vegetation map, a given combination of these scalar inputs was presented to the ARTMAP network. The spectral, terrain, and location variables were originally in a variety of units, each spanning a different range (Table 2). Before presentation to the neural network, each variable was rescaled to the interval [0,1], based on its range of values in the data set. For example, an original Band 1 value x (which lies between 27 and 224) would be replaced, as an ARTMAP input component, by $(x-27)/197$ (which lies between 0 and 1). In general, each original input variable x was replaced by $(x-Min)/(Max-Min)$, where the minimum and maximum values of the original variables were as listed in Table 2. Apart from this scaling step, input values were not preprocessed.

ARTMAP METHODOLOGY: PREDICTIONS FOR THE FIELD OBSERVATION DATA SET

Each reported measure of ARTMAP predictive accuracy is the result of applying fivefold cross-validation (Mosier, 1951) to the field observation data set (Table 1). The field observation set was thereby partitioned into five subsets, each with approximately 200 sites. For a given network, one subset was reserved as a test set, while ARTMAP was trained on the remaining four subsets. During training, a vector of information for each pixel in each designated training site was presented to the system, along with the site's Calveg label. Test performance was evaluated only on the subset of sites not seen during training. After each pixel in a given test set site had produced an individual output, the predicted Calveg class label was taken to be the one predicted by the largest number of pixels in that site.

The ARTMAP results reported here are all the product of fast learning, and each pixel was presented only once during ARTMAP training. The fast learning capability of this network has the advantage of permitting online and incremental training of large databases. However, fast learning also causes results to vary somewhat with the order of input presentation: early training set input vectors typically establish an overall internal category structure which is fine-tuned by later inputs. This feature, which might appear to be a disadvantage of fast learning, can actually be used to help boost performance, through the device of *voting.* For the present study, for each fixed training/test subset partition of the field observation data set, an ARTMAP system was trained five different times, each with a different ordering of the training set. For each ordering, the Calveg class prediction of each test site was recorded. Once the five predictions were available, the system made a site-level prediction by voting: the final ARTMAP prediction for a given test set site was taken to be the Calveg class predicted by the largest number of voting networks. For example, at a given site, if three networks chose the correct output class and each of the other two networks chose a different class (i.e., a 3–1–1 vote), the site would be correctly classified. Overall Calveg accuracy was calculated as the fraction of test sites correctly labeled by this procedure.

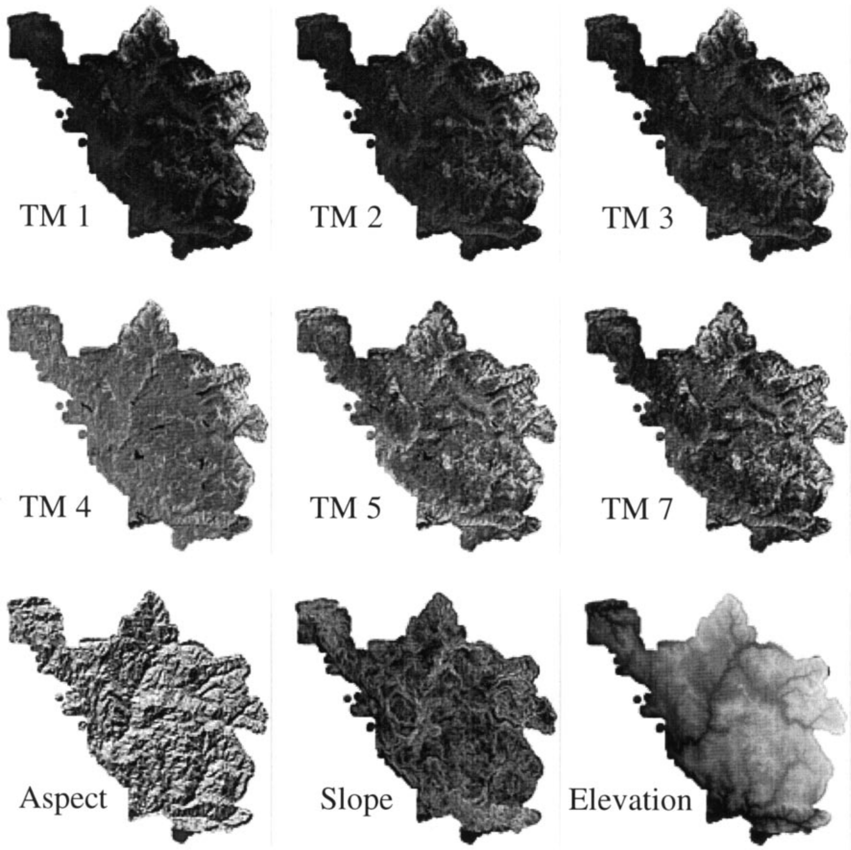


Figure 3. Values of individual input components in the Sierra National Forest data set. The top six maps show the spectral inputs and the bottom three show terrain information. Not shown are the cosine of the solar zenith angle, northing, and easting. In each map, a dark pixel corresponds to a high value of the input variable.

Note that voting could occasionally result in a tie (2–2–1 or 1–1–1–1–1) among the five networks making predictions for a given test site. In this case, if one of the (2 or 5) tied winning outputs was the actual Calveg class, the site was counted as contributing a fraction (1/2 or 1/5) to the total number of correct predictions.

Once voting was completed for a given test subset of the field observation data set, the entire procedure was repeated, in turn, for each of the five test subsets. Thus, in addition to ensuring a strict separation between training and testing sites, the cross-validation method helps eliminate spurious variations in the randomly selected training/testing subset partition. With voting further reducing variability across input orderings, and with no individual parameter selection required, reported ARTMAP results are robust. Lifeform predictions were obtained by merging all Calveg predictions for each of the six lifeform classes (Table 1).

A similar procedure produced the Calveg and lifeform maps of the entire Sierra National Forest, except with the ARTMAP training set consisting of all 1013 field observation sites. After pixels from these sites, in

Table 2. Range of Input Values, before Scaling

		Units	*Min*	*Max*
Spectral variables	Band 1	DN[a]	27	224
	Band 2	DN	7	113
	Band 3	DN	3	137
	Band 4	DN	1	108
	Band 5	DN	1	141
	Band 7	DN	1	83
Terrain variables	Elevation	*m*	257	4147
	Slope	%	0	100
	Aspect	Bins	1	12
	cos(z)	—	0	1
Location variables	Northing		4,068,000	4,180,500
	Easting		238,200	355,000

[a] DN=digital numbers (8 bits)

Table 3. Mapping Accuracies for ARTMAP and Conventional Methods for Various Combinations of Spectral, Terrain, and Location Inputs

Tag	*Input variables*	*Correct lifeform*	*Correct Calveg*	*No. of Input Components*	*Internal Categories*
ARTMAP					
b	TMB (TM Bands 1–5&7)	78%	45%	6	1,721
B	TMB+(TMB+cos(z))	80%	47%	7	1,702
T	Terrain (aspect, slope, elevation)	70%	40%	3	1,064
L	Location (northing, easting)	73%	49%	2	303
LT	Terrain, location	75%	51%	5	368
BT	TMB+, terrain	**83%**	54%	10	697
BL	TMB+, location	82%	53%	9	631
BLT	TMB+, terrain, location	**83%**	**57%**	12	340
Conventional methods					
Exp	TMB+	64%	N/A		
Edit	TMB+, Terrain manual editing	**86%**	**61%**		

five random orderings, were presented to the network, voting produced a Calveg label for each pixel in the complete map. In addition to improving accuracy and reducing variability, the number of voting networks that agree on the winning label provides a confidence index for each pixel-level prediction. Voting thus automatically produces an ARTMAP *confidence map*, as described below.

Comparing Input Component Combinations

In order to examine the relative contributions of various spectral, terrain, and location components, ARTMAP networks were trained using selected subsets of these variables (Table 3). For example, the combination *b* denotes an ARTMAP system where the input vector consisted of the values of the six spectral bands (TM Bands 1–5&7); combination *L* denotes a system where the input vector consisted of the two location components (northing, easting); and combination *BLT* denotes a system where the input vector consisted of all 12 input components (six spectral bands, cosine of the solar zenith angle, three terrain variables, and two location variables). Table 3 shows ARTMAP predictive accuracy, for both life-form and Calveg species associations, for eight different input combinations. ARTMAP performance is also compared with that of the conventional (expert systems) mapping method, before editing (*Exp*) and after editing (*Edit*). Because the conventional system used the edited lifeform map as a precursor of the Calveg species association map, an accuracy measure for the unedited expert map was available only for the lifeform task.

Table 3 also specifies the median numbers of internal categories, or "rules," in the trained ARTMAP networks. In the course of learning, an ARTMAP system adds category nodes incrementally, in response to predictive errors, using the minimum number of nodes needed for accurate performance. Since ARTMAP features fast online learning, a noisy or inconsistent input set tends to cause many predictive errors, and hence may produce networks with large numbers of internal category nodes. Systems with fewer nodes, or greater *code compression*, reflect a more orderly construction of the internally defined rules which the network self–organizes during training. Such systems often exhibit better test set accuracy, or generalization, on data not seen during training. In addition, the more nodes, the slower the algorithm's execution time. Thus the number of category nodes is an important index of ARTMAP performance.

In Table 3, the combination *b*, which uses spectral data only, represented the baseline case. The combination *B* adds the cosine of the solar zenith angle [cos(z)], which boosted ARTMAP predictive accuracy for lifeforms and Calveg classes by 2%. Including location as an input vector improved both performance and code compression. In fact, networks trained with location inputs (*LT*, *BL*, *BLT*) created fewer than half as many internal categories as systems presented with all the same inputs minus location (*T*, *B*, *BT*). Location alone (*L*) yielded maps which were fairly accurate at test sites—in fact, considerably more accurate than the uncorrected expert map (Table 3)—but which were nonetheless of dubious utility (Fig. 4): The prediction for a test pixel was determined primarily by the vegetation class label of the nearest training site. Adding either location or terrain data (*BL*, *BT*) to the spectral data case *B* led to improved predictive accuracy and code compression. Of these two, the system that predicted both lifeform and Calveg labels slightly more accurately was the one that used spectral and terrain data (*BT*). When location information was included as well (*BLT*), the system created only 20% as

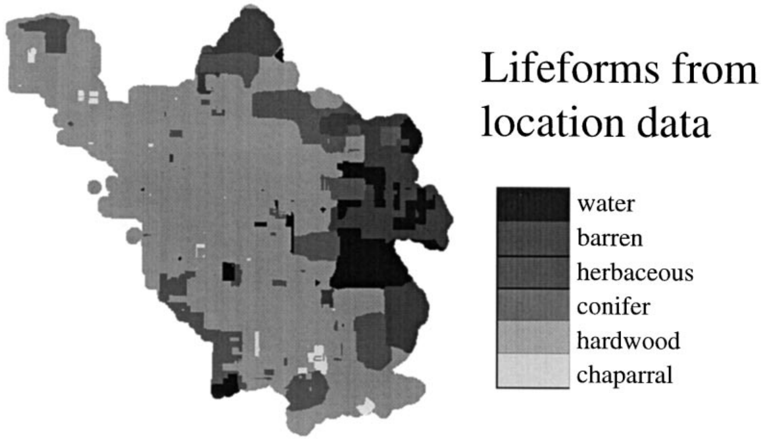
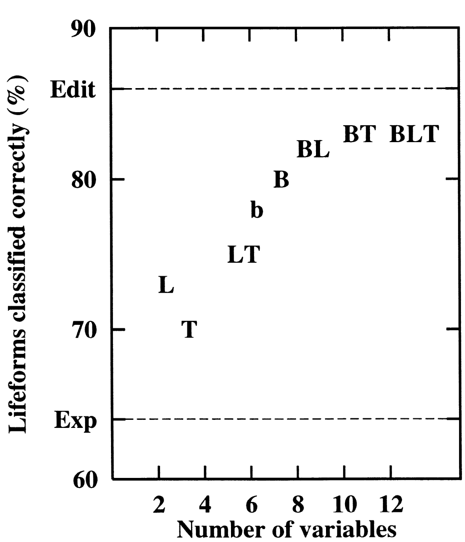


Figure 4. ARTMAP lifeform map from location data alone (*L*). The map is accurate at 73% of the field observation sites, yet contains major distortions. For example, the extended dark area near the right center of the map shows that too many pixels are labeled *water* simply because they are near a *water* site in an area that has few field observation labels (Fig. 2).

many internal categories as the one based on spectral data (*B*), while boosting Calveg discrimination from 47% to 57%. In general, ARTMAP predictive accuracy improved with the number of input components (Fig. 5).

The results for the ARTMAP tests with different inputs show interesting patterns which support the main hypotheses underlying the conventional methods. First, training the system on spectral data (*B*) resulted in high accuracies for lifeforms, but low accuracies for species associations. Similarly, the conventional method used by the USFS in Region 5 relied first on spectral data for lifeform classifications, and then on terrain relations for species associations (Fig. 1a). For both methods, the addition of terrain variables (*BT*) helped little for lifeform identification, but made more substantial contributions for the species association task.

Figure 5. Accuracy of ARTMAP lifeform predictions versus the number of components in the input vector. Input combination labels are defined in Table 3. Horizontal lines indicate the accuracy levels of the conventional method, before editing (64%—*Exp*) and after editing (83%—*Edit*).



During learning, an ARTMAP network creates a set of "rules," each of which could, in principle, employ any combination of input variables. Thus, the role of any single variable in a set of overall predictions is often difficult to assess. One could speculate, however, that, for the present example, the two location variables might have allowed the network to compute different relationships among species associations and terrain variables in different portions of the area being mapped. Location variables would then have provided the inputs needed for ARTMAP to learn the equivalent of the natural regions that have proved essential in applications of the conventional methods (Franklin et al., 1986; Woodcock et al., 1980).

Analysis of ARTMAP Performance

The best performance of the ARTMAP systems was obtained by using all available spectral, location, and terrain input components (*BLT*). This final combination was most successful at discriminating Calveg classes, while also minimizing memory requirements. It correctly classified the Calveg species of 57% of test sites and the lifeforms of 83% of these sites.

Predictions of the ARTMAP (*BLT*) network are broken down by Calveg class in Figure 6, which depicts a test set confusion matrix. This format makes the system predictions more legible than if they were presented as a table of numbers, thus facilitating comparison among model variations. Each matrix column corresponds to a Calveg class as specified in the field observation data set (actual Calveg class), and rows correspond to the predicted Calveg classes. A column shows the distribution of ARTMAP predictions for all test set sites that were actually in a given Calveg class. The darker a matrix cell, the larger the fraction of that column's test sites that were predicted to be in that row's class. Cells on the diagonal indicate the fraction correctly predicted for each

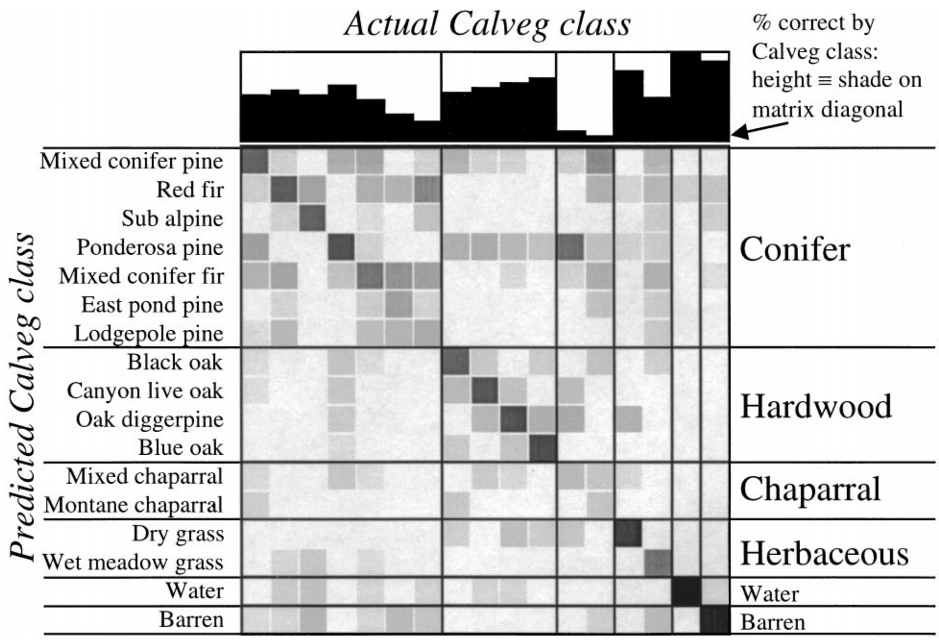


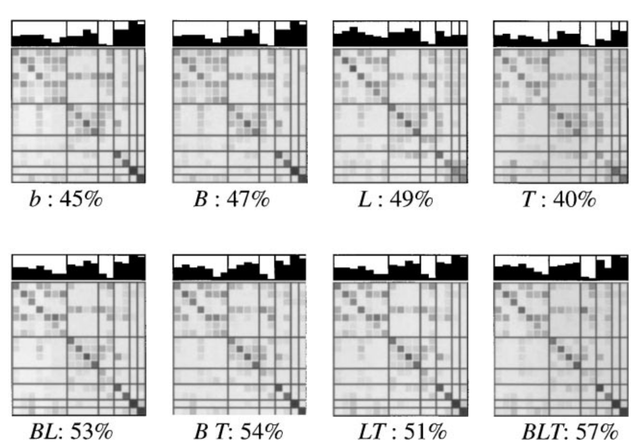
Figure 6. Calveg confusion matrix for the ARTMAP (*BLT*) system, which used all 12 input components. Columns correspond to actual Calveg classes and rows correspond to predicted classes. The darker the shading of a cell, the larger the fraction of sites of the actual Calveg class that were associated with the species class of the corresponding row. The bar graph (top) shows the fraction of sites of a given class that were correctly predicted as belonging to that class, ranging from 0% (no bar) to 100% (bar at full height). The heights of these bars provide a calibration of the matrix grey scale, since the darkness of a diagonal cell is proportional to the percent correct for the Calveg class of that column.

class. For example, the first column shows the distribution of Calveg class predictions for all test sites that should have been labeled *mixed conifer pine*, according to field observations. About half of these sites were correctly labeled, according to the bar at the top of the column; and *ponderosa pine* was the most common erroneous label, followed by *mixed conifer fir.* In the fourth column, the height of the bar graph shows that *ponderosa pine* is the conifer class that was most often labeled correctly, and this observation is confirmed by the darkness of the diagonal cell. The confusion matrix also shows that chaparral species were most commonly mislabeled as types of conifer.

The confusion matrix indicates how ARTMAP Calveg predictions were merged to make lifeform predictions. All test sites that were actually conifer were correctly labeled as *conifer* in the lifeform prediction task if the Calveg label was any type of conifer. In Figure 6, actual conifer sites correspond to the first seven columns of the matrix. Thus all predictions that appear as grey cells in the upper left-hand 7×7 submatrix collapse into the single lifeform prediction *conifer.* Errors correspond to grey cells that appear in rows 8–17. The Calveg confusion matrix provides more details about the causes of lifeform errors than would a lifeform confusion matrix alone. For example, the matrix shows that sites that are actually ponderosa pine (column 4) were placed in the correct Calveg class more often than any other conifer species, but the matrix also shows that many of the errors for these sites occurred as incorrect hardwood species labels (rows 8–11). All these errors would have contributed to lifeform errors, causing conifer sites to be mislabeled as *hardwood.* Conversely, a number of sites that were actually hardwood (columns 8–11) were mislabeled *ponderosa pine* at the Calveg level (row 4).

Figure 7 uses the confusion matrix format to display Calveg prediction results from each of the eight ARTMAP input combinations listed in Table 3. This figure shows that a system such as *b*, which uses only the six spectral bands as system input, produced a widespread distribution of off-diagonal grey cells. On the other hand, many of the errors for *b* are seen to be between Calveg classes that share the same lifeform, especially in the *conifer* and *hardwood* submatrices. This observation helps explain why the differences in predictive accuracy between, say, *b* and *BLT* were smaller for the lifeform predictions (78% vs. 83%) than for the Calveg predictions (45% vs. 57%).

Figure 7. ARTMAP confusion matrices for the eight input combinations (Table 3). The overall rate of correct Calveg test set predictions is shown for each case. The *BLT* matrix is the same as in Figure 6.



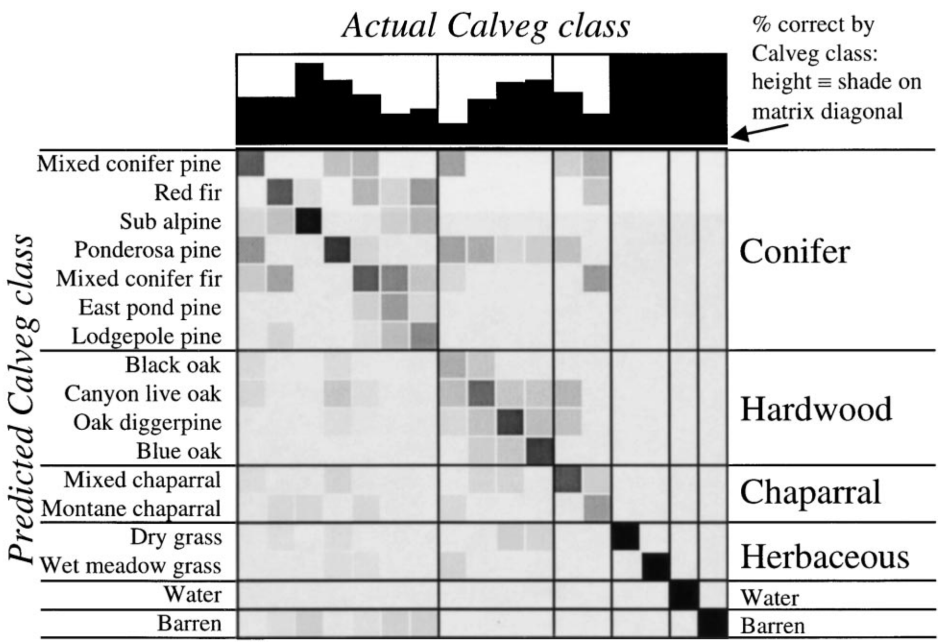


Figure 8. Calveg confusion matrix for the conventional (edited expert) map. The overall correct classification rate was 61%. The bar graph shows that classification for *herbaceous*, *water*, and *barren* classes was perfect.

Comparing the Conventional Mapping Method and the ARTMAP Network

Figure 8 shows a confusion matrix for the edited expert map, which had an accuracy rate of 61% for Calveg predictions and 86% for lifeform predictions at the 1013 field observation sites. The corresponding rates for the ARTMAP system were 57% and 83%, respectively. Comparing the confusion matrices in Figures 6 and 8 reveals that the conventional method discriminated *chaparral* and *herbaceous* more accurately than did ARTMAP. Both methods discriminated *water* sites perfectly, and the *barren* class was also easily identified. Conventional discrimination of *conifer* and *hardwood* was similar to that of ARTMAP: the largest differences occurred in the *subalpine* conifer class (column 3), where the conventional system was more accurate; and in the *black oak* hardwood class (column 8), where ARTMAP was more accurate.

Figure 9 (a,b) shows the lifeform maps produced by ARTMAP and by the conventional method (with editing), across the entire forest. Note, in particular, the distribution of chaparral sites (yellow) on the expert map, and the relative absence of these site labels on the ARTMAP map. The confusion matrices in Figures 6 and 8 suggest that chaparral is responsible for the difference in performance rate between the two systems. A probable cause of the ARTMAP chaparral errors is the small number of training set sites available for these vegetation types (Table 1). This hypothesis is supported by Figure 10, which shows the ARTMAP classification accuracy as a function of the number of available field observation sites for each Calveg class. With some exceptions, an approximate correlation is visible between the number of training sites per class and the ARTMAP test set accuracy. Figure 10 indicates the relative ease of discriminating *water* and *barren*, which, because they are spectrally distinct, had high accuracy despite only moderate training set representation; and the relative difficulty of the classes *mixed–conifer–fir* and *mixed–conifer–pine*, which had only moderate accuracy despite an abundance of training sites.

Confidence Maps

Recall that ARTMAP predictions were the result of voting across five networks. Each network was trained on a unique ordering of the given training set, and the output class prediction was the one that received a plurality of votes. Voting helps an ARTMAP system to make robust predictions and to improve accuracy, but it can, in addition, be used to calculate a confidence index. Namely, the number of ARTMAP voters that agree with a prediction serves as a gauge of confidence in that prediction. For the Sierra National Forest map, when all five voters agreed on a single output class, the prediction was viewed with maximal confidence. At the other extreme, the system had minimal confidence in a prediction when each voter chose a different output class.

Figure 11 demonstrates that voter confidence was, in fact, a good measure of predictive accuracy. Plots show confidence assessments for each of the eight spectral, terrain, and location input variations (Table 3). The x-axis in each plot marks five confidence levels, or bins, based on the number of voting networks that agreed on the outcome. The bar graph shows the percent of test sites at each of the five confidence levels. The dashed line with diamonds shows the percent of predictions that were correct at each confidence level. In nearly all cases, accuracy increased with confidence: The only exceptions

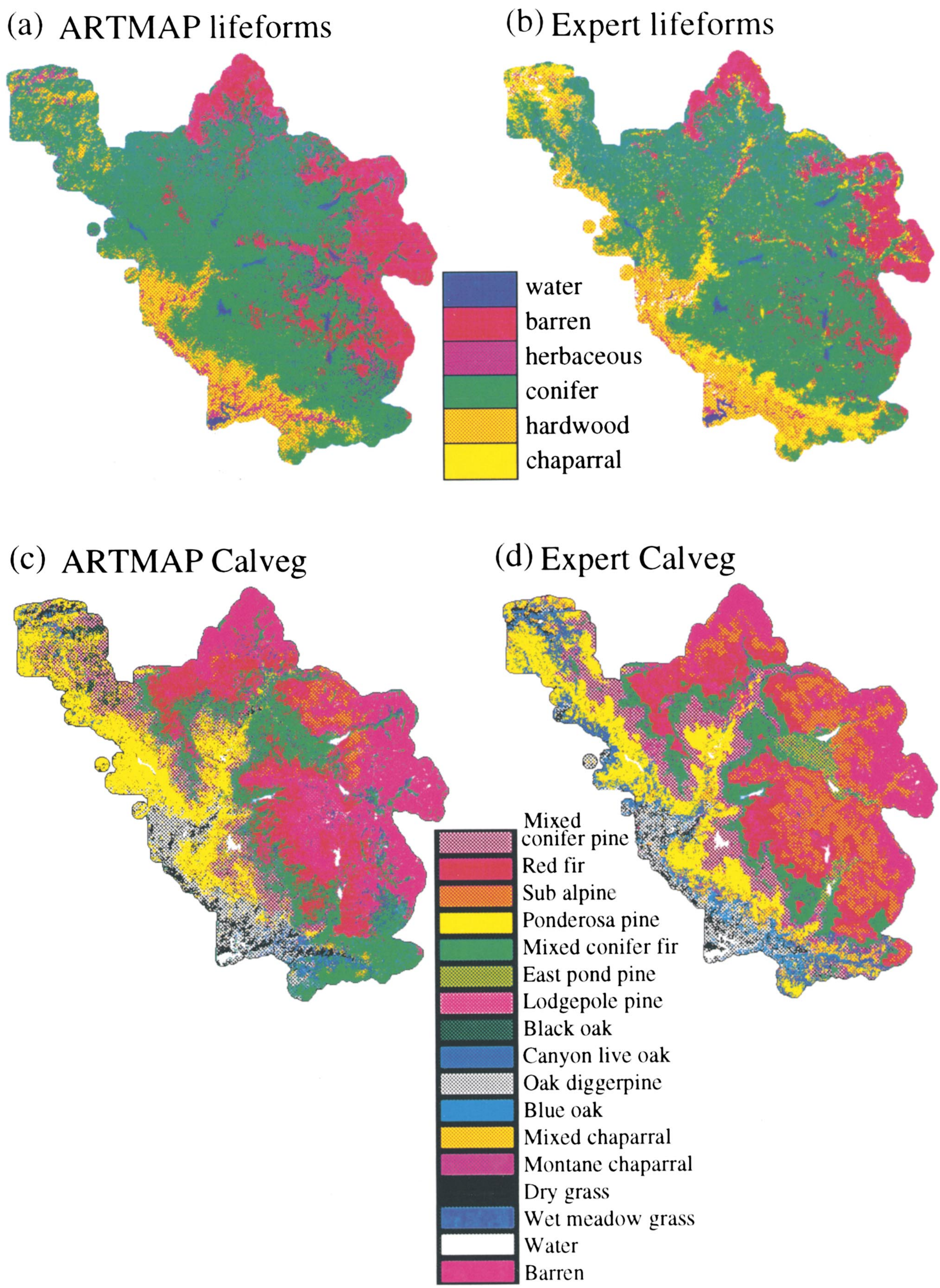


Figure 9. Lifeform maps: a) the ARTMAP map, which used all available spectral, location, and terrain components of the input data; b) conventional (edited expert) map. Calveg label maps of 17 vegetation classes produced by: c) ARTMAP; and d) conventional (edited expert) methods.

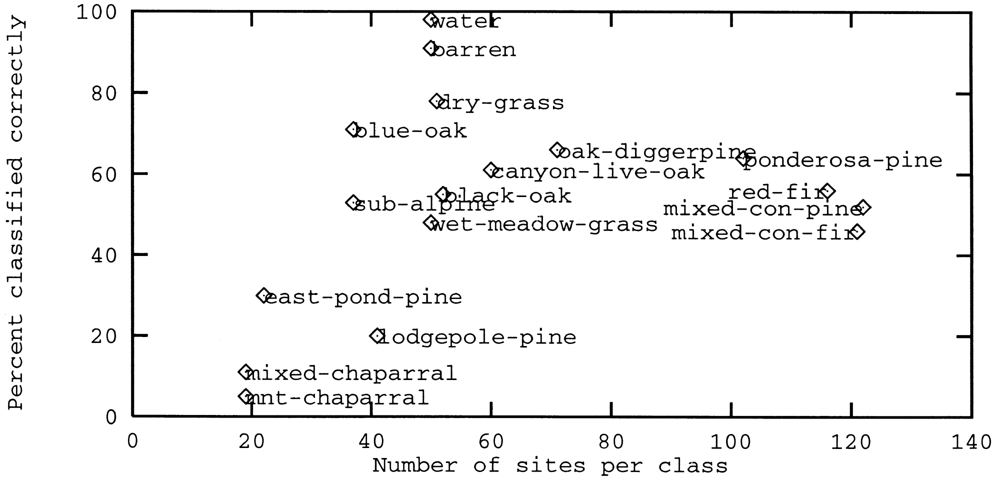


Figure 10. In ARTMAP computations, the rate of correct Calveg predictions within a given class tended to increase with the number of sites available in the field observation set.

to this rule are two cases (*T*, *BL*) where some rare low-confidence predictions happened to be correct. The dotted line with crosses marks the product of the other two data series, indicating accuracy relative to the number of test sites at each confidence level.

Figure 12 further demonstrates how ARTMAP predictive accuracy increased with confidence. Figure 6 shows the confusion matrix for this system, but does not indicate confidence. Figure 12 displays components of this confusion matrix according to the number of voters making the predictions. For example, predictions at 241 of the field observation sites were made on the basis of at least four, but fewer than five, votes. (Recall that fractional votes occurred in the case of ties; and that each field observation site took a turn as a test set site, according to the cross-validation protocol.) Accuracy increased with confidence level (left to right), as shown by the increasing heights of the bar graphs and the darkening shades of the diagonal cells.

When a voting ARTMAP system produces a vegetation map, the system also automatically produces a corresponding *confidence map.* At each pixel, this map indicates the number of voters agreeing on the predicted class. Figure 13 shows an ARTMAP (*BLT*) Calveg confidence map. Light areas indicate where the network was confident (4–5 votes), while dark areas indicate low confidence (1–2 votes) in network predictions. For example, the areas at the bottom of the map and in the upper left-hand corner show low-confidence regions. Much of the chaparral is located in these areas (Fig. 9b): Although ARTMAP failed to find many of these sites, the system did point correctly to locations where editing could most profitably be concentrated. Editing low-confidence sites, which accounted for less than 10% of the total, would be an efficient way to improve ARTMAP performance.

Figure 11. ARTMAP predictive accuracy by confidence bin, for each of the eight input combinations. Bar graph height shows the percent of sites at a confidence level, equal to the fraction of voting networks that agreed on the site prediction. The dotted line with diamonds shows the percent of sites in the bin where ARTMAP correctly predicted the Calveg class. The dashed curve with + symbols marks the product of the percent correct and the confidence level.

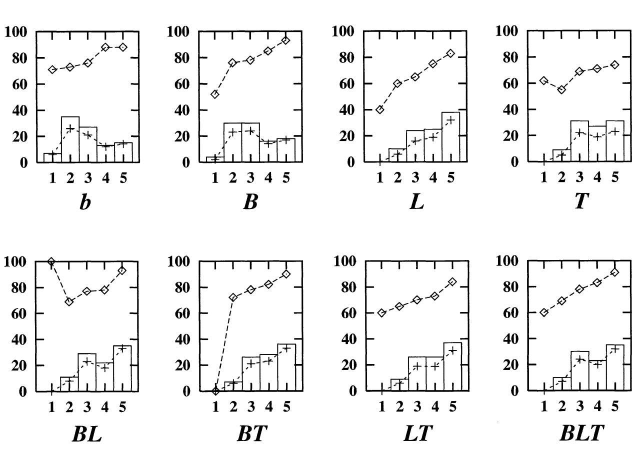


Figure 12. Confusion matrices for ARTMAP (*BLT*) Calveg predictions by confidence level. Prediction of most classes became more accurate as confidence increased.

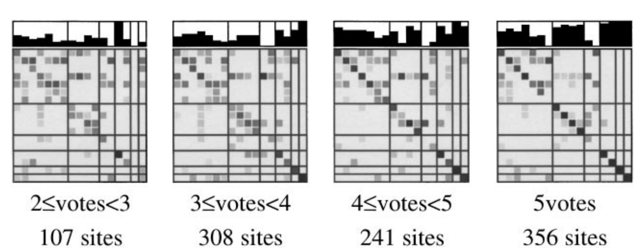


Figure 13. ARTMAP Calveg confidence map. Dark shades indicate locations of low-confidence sites, where votes were split among several different predictions. Lakes and barren areas were among the regions predicted most confidently (white).

Benefits of the ARTMAP System

With respect to the present study, the primary question for a map developer concerns the relative benefits of the ARTMAP system for operational purposes. Based on the results presented in this article and on the authors' collective experience using both ARTMAP and conventional mapping systems, the following points can be made.

First, both ARTMAP and the conventional system require training, but that training appears in very different forms. For the conventional system, training is of two types. One type is labeling unsupervised classes by analysts to produce a lifeform map; the second type is defining natural regions within the Forest, and then, for each such region, calibrating the terrain rules for species associations within each of the lifeforms. For ARTMAP, the training requires only a set of field observations in order to calibrate the system in a single, computer-based classification step.

From a number of perspectives, the requirements for ARTMAP are preferable. First, the collection of a set of training sites is easier, faster, and probably more useful in the long run than the special-purpose training required for the conventional method. One reason relates to the common need to update these maps following the completion of an accuracy assessment. With the conventional maps, the training used to label lifeform classes is not preserved, and hence has to be recreated for any future attempts to improve the map. In contrast, the training sites used for ARTMAP can be preserved and augmented as necessary to improve maps incrementally at any future time, without requiring that the original training data be available for an entirely new training process.

A second advantage is the relative simplicity of the ARTMAP approach compared to the conventional methods. As illustrated in Figure 1, the conventional methods require two tracks of multiple steps, while the ARTMAP system is trained with a single, automated step. This sort of simplicity is highly desirable, as it makes the new method faster, less expensive, and easier to learn. In addition, the ARTMAP method does not require the level of knowledge of the mapping region that is needed for successful implementation of the conventional method: It is more difficult to define natural regions, and the terrain rules within those regions, than it is to collect a set of labeled field observation sites.

A third advantage is the production of a confidence map by the ARTMAP system. This map has immediate value as an editing guide, to improve the efficiency of a tedious and slow process. The confidence map can also be passed along with the vegetation maps to users who may benefit from knowledge of uncertainty.

A final advantage is the reasonable expectation of future improvements in the accuracy of the maps. Without editing, the ARTMAP maps are of comparable accuracy to the conventional maps after they have been edited. Following light editing, guided by the confidence maps, the ARTMAP results would be expected to become considerably more accurate.

DISCUSSION

Figure 9(c,d) displays Calveg maps of the Sierra National Forest produced by ARTMAP and by the conventional method, with editing. Some differences are apparent, in particular in the case of *subalpine* and *chaparral.* Though correct identification of 17 Calveg classes is a challenging problem, the maps are qualitatively similar.

The ARTMAP neural network learned to classify vegetation stands as one of 17 Calveg types from knowledge of spectral, terrain, and location variables across all pixels in the stand. Once the Calveg labels of field observation sites had been collected, ARTMAP carried out the entire task of training and map production in a matter of hours. In contrast, producing unedited expert maps required about 6 months for developing heuristics and simple ecological models based on forest visits, followed by nearly as much time spent in painstaking editing of the uncorrected maps. In all, it took the equivalent of about 1 year's effort to produce the expert map.

The neural network approach yielded maps with accuracies that exceeded those of unedited expert maps and that approached the accuracies of the edited expert maps. Training ARTMAP with the field observation data took about 1 hour. Voting automatically provides a confidence index, which may be used to guide future editing. The ARTMAP neural network method was therefore found to have drastically increased the efficiency of the mapping process.

This research was supported in part by the National Aeronautic and Space Administration (NASA NAS5-31369 and NASA NAS5-3439), the National Science Foundation (NSF SBR-95-13889), the Office of Naval Research (ONR N00014-95-1-0409 and ONR N00014-95-1-0657), and the U.S. Forest Service Region 5 Remote Sensing Laboratory, Sacramento. The authors thank Dena Simons for assembling the data set.

REFERENCES

Aspinall, R., and Veitch, N. (1993), Habitat mapping from satellite imagery and wildlife survey data using a Bayesian modeling procedure in GIS. *Photogramm. Eng. Remote Sens.* 59:537–543.

Bauer, M. E., Burk, T. E., Ek, A. R., et al. (1994), Satellite inventory of Minnesota forest resources. *Photogramm. Eng. Remote Sens.*, 60, 287–298.

Carpenter, G. A., and Markuzon, N. (1998), ARTMAP-IC and medical diagnosis: Instance counting and inconsistent cases. *Neural Netw.* 11:323–336.

Carpenter, G. A., Grossberg, S., and Reynolds, J. H. (1991), ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Netw.* 4:565–588.

Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J. H., and Rosen, D. B. (1992), Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Trans. Neural Netw.* 3:698–713.

Carpenter, G. A., Gjaja, M. N., Gopal, S., and Woodcock, C. E. (1997), ART neural networks for remote sensing: vegetation classification from Landsat TM and terrain data. *IEEE Trans. Geosci. Remote Sens.* 35:308–325.

Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., and Woodcock, C. E. (1999), A neural network method for mixture estimation for vegetation mapping. *Remote Sens. Environ.* 70:138–152.

Cohen, W. B., Spies, T. A., Swanson, F. J., and Wallin, D. O. (1995), Landcover on the western slopes of the central Oregon Cascade range. *Int. J. Remote Sens.* 16:595–596.

Congalton, R. G., Green, K., and Teply, J. (1993), Mapping old-growth forest on National Forests and park lands in the Pacific Northwest from remotely sensed data. *Photogramm. Eng. Remote Sens.* 59:529–535.

Franklin, S. E., and Wilson, B. A. (1991), Vegetation mapping and change detection using SPOT MLA and Landsat imagery in Kluane National Park. *Can. J. Remote Sens.* 17:2–17.

Franklin, S. E., Logan, T. L., Woodcock, C. E., and Strahler, A. H. (1986), Coniferous forest classification and inventory using Landsat and digital terrain data. *IEEE Trans. Geosci. Remote Sens.* GE-24:139–149.

Gopal, S., and Fischer, M. M. (1997), Fuzzy ARTMAP—a neural classifier for multispectral image classification. In *Recent Developments in Spatial Analysis: Spatial Statistics, Behavioural Modelling and Computational Intelligence* (M. M. Fischer and A. Getis, Eds.), Springer-Verlag, Heidelberg, pp. 306–335.

Gopal, S., Woodcock, C. E., and Strahler, A. H. (1999), Fuzzy ARTMAP classification of global land cover from the 1 degree AVHRR data set. *Remote Sens. Environ.* 67:230–243.

Matyas, W. J., and Parker, I. (1980), CALVEG mosaic of existing vegetation of California, Regional Ecology Group, U.S. Forest Service, Region 5, San Francisco, 27 pp.

Mosier, C. I. (1951), Symposium: The need and the means for cross-validation. 1. Problem and designs of cross-validation. *Ed. Psychol. Measure.* 11:5–11.

Richards, J. A. (1993), *Remote Sensing Digital Image Analysis: An Introduction*, 2nd ed., Springer-Verlag, Berlin, 340 pp.

Woodcock, C. E., Strahler, A. H., and Logan, T. L. (1980), Stratification of forest vegetation for timber inventory using Landsat and collateral data. In *14th International Symposium on Remote Sens. Environ.*, San Jose, Costa Rica, pp. 1769–1787.

Woodcock, C. E., Collins, J., Gopal, S., et al. (1994), Mapping forest vegetation using Landsat TM imagery and a canopy reflectance model. *Remote Sens. Environ.* 50:240–254.