

## Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors

Curtis E. Woodcock<sup>a,b,\*</sup>, Scott A. Macomber<sup>a,b</sup>, Mary Pax-Lenney<sup>a,b</sup>, Warren B. Cohen<sup>c</sup>

<sup>a</sup>Department of Geography, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA

<sup>b</sup>Center for Remote Sensing, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA

<sup>c</sup>Forestry Sciences Laboratory, Pacific Northwest Research Station, USDA Forest Service, 3200 Jefferson way, Corvallis, Oregon 97331 USA

Received 19 January 2000; received in revised form 26 February 2001; accepted 30 April 2001

---

### Abstract

Landsat 7 ETM+ provides an opportunity to extend the area and frequency with which we are able to monitor the Earth's surface with fine spatial resolution data. To take advantage of this opportunity it is necessary to move beyond the traditional image-by-image approach to data analysis. A new approach to monitoring large areas is to extend the application of a trained image classifier to data beyond its original temporal, spatial, and sensor domains. A map of forest change in the Cascade Range of Oregon developed with methods based on such generalization shows accuracies comparable to a map produced with current state-of-the-art methods. A test of generalization across sensors to monitor forest change in the Rocky Mountains indicates that Landsat 7 ETM+ data can be combined with earlier Landsat 5 TM data without retraining the classifier. Methods based on generalization require less time and effort than conventional methods and as a result may allow monitoring of larger areas or more frequent monitoring at reduced cost. One key component to achieving this goal is the improved availability and affordability of Landsat 7 imagery. These results highlight the value of the existing Landsat archive and the importance for continuity in the Landsat Program. © 2001 Elsevier Science Inc. All rights reserved.

### 1. Introduction

One of the most valuable uses of Landsat imagery is for monitoring environmental change. To date a wide range of applications have been demonstrated ranging from forest change (Coppin & Bauer, 1994; Hall, Botkin, Strelbel, & Goetz, 1991; Jha & Unni, 1994; Vogelmann & Rock, 1988) to urbanization (Jenson & Toll, 1982; Seto, Song, & Woodcock, 1999) to agricultural expansion (Pax-Lenney, Woodcock, Collins, & Hamdi, 1996) and desertification (Coiner, 1980; Lunetta & Elvidge, 1998). The historical archive of Landsat imagery dating back to the launch of ERTS in 1972 provides a unique and invaluable data source for tracking change in landscapes. While use of that archive is increasing, its use is somewhat limited by the way in which Landsat data is used to monitor environmental change. Some change monitoring efforts are based on visual interpretation of pairs of Landsat images while others involve computer-based analysis. What almost all efforts to monitor environmental change using Landsat

imagery share is the need for local calibration or training of the change detection methods. For visual interpretation this requirement will never change, as the interpreter will always have to examine each pair of images to be included in a study. However, for computer-based analysis it is possible to imagine developing environmental monitoring methods applicable outside the immediate domain in which they were trained. As such, these methods would depend on *generalization*.

The purpose of this paper is to pursue the question of generalization with respect to the problem of monitoring change in temperate forests. The intent is to develop monitoring methods that employ generalization to extend the geographic range and the temporal frequency of their applicability, thus allowing monitoring of large areas at frequent time intervals at reasonable expense. To pursue this notion of generalization, it is helpful to characterize types and degrees of generalization. To help this discussion, let us clarify some simple terms. Let us define a *mapping process* as some algorithmic approach applied to an image or images to make a map or monitor change. Such a mapping process might be a classification algorithm, for example. Within change detection there have been a wide

---

\* Corresponding author.

variety of such mapping processes used (see Singh, 1989), ranging from image differencing to postclassification comparisons to multitime classification to change vector analysis. Similarly, let us define *training data* as whatever data might be necessary for the mapping process. For example, in image classification the training data might be examples of the desired map classes, or training sites. By analogy, any data used to help select a threshold for image differencing could be considered *training data*. Also, in this paper, a *scene* refers to a geographic location (such as a path and row location in the Landsat World Reference System) while an *image* refers to a specific image acquisition. Thus, there may be many images for any one scene.

Using the simple terms from above, the following types and levels of generalization can be defined. *Within-image* generalization refers to deriving training data from within the same image that will be mapped. This level of generalization is very limited, but has long been the standard approach to classification in remote sensing, particularly with fine resolution data such as Landsat or Systeme Pour l'Observation de la Terre (SPOT). Another level of generalization, *within-scene*, is temporal. Training data are derived from one image of a scene and the mapping process is applied to an image from a different date from the same scene. With respect to spatial generalization, *within-region* generalization involves applying training data to a mapping process in nearby scenes. A more extreme spatial generalization is *across-region*, where training data from one region of a continent is applied to mapping another. Generally, *within-* and *across-region* generalizations include temporal generalization as well. Similarly, it is possible to define *across-sensor* generalization as deriving training data from one sensor and applying it to a mapping process on data from another sensor.

The concept of generalization is not new; it has also been known as spectral extendibility and signature extension (Botkin, Estes, Macdonald, & Wilson, 1984). Perhaps the most well known test of generalization of vegetation signatures in the Landsat community is the Large Area Crop Inventory Experiment (LACIE). During the mid-1970s, LACIE evaluated the feasibility of spectral extendibility for determining wheat acreage. Classifiers trained with data from one segment within the US wheat-growing region were tested against nearby segments. While haze and sun-angle corrections improved cross-segment classification, in general, the approach was considered untenable (Minter, 1978). Since that time, there has been little attention paid to spectral signature extension (or generalization) with Landsat data. It is only recently as attention has turned to monitoring larger areas that interest in generalization, or signature extension, has returned (Cohen, Maersperger, Spies, & Oetter, 2001; Pax-Lenney, Woodcock, Gopal, & Macomber, 2001).

Despite the lack of success in early tests of signature extension, we once again propose this approach as a way to monitor large areas for forest change with Landsat imagery

and believe it feasible for several reasons. First, over the past 20 years atmospheric correction methods have been developed that will reduce some of the problems encountered in earlier research. Second, new kinds of image classification algorithms have been developed that are better suited for these problems, in particular decision trees and artificial neural networks. Previous research has shown neural networks to be more effective in forest environments than traditional, statistical approaches (e.g., Carpenter, Gajja, Gopal, & Woodcock, 1997; Carpenter, Gopal, Macomber, et al., 1999; Gopal & Woodcock, 1996). One reason is that neural network classifiers do not make unrealistic assumptions about the distributional properties of classes, as was the case in the days of LACIE. Third, we are starting with a simple problem, the identification of wholesale forest change due to causes such as harvesting or fire. In essence, we are attempting to do less in terms of the level of detail in information extracted from the imagery, but over larger areas. This tradeoff could help make it possible to monitor large areas for forest change at reasonable expense using imagery with fine enough spatial resolution to detect changes directly. The purpose of this paper is to present results of tests of generalization in time, space, and across Landsat sensors. These tests involve monitoring forest change in the Cascade Mountains of Oregon and the Rocky Mountains of Colorado using imagery from Landsat 5 TM (Thematic Mapper) and Landsat 7 ETM+ (Enhanced Thematic Mapper Plus).

## 2. Methods

At the heart of our efforts regarding generalization are trained artificial neural networks. While the intent is change detection, the fundamental process involved is image classification where the inputs are two dates of Landsat images (all six reflective bands) and the output classes are very simple: *forest change* and *no forest change*. *Forest change* is defined as areas that were forest in the early image and are no longer forest in the later image. *No forest change* is everything else. The classifier is Fuzzy ARTMAP (Carpenter et al., 1997), a neural network that uses match-based learning. We have achieved good classification results using Fuzzy ARTMAP in a number of contexts and situations (Abuelgasim, Ross, Gopal, & Woodcock, 1999; Carpenter, Gopal, Macomber, et al., 1999; Carpenter, Gopal, Martens, & Woodcock, 1999; Gopal, Woodcock, & Strahler, 1999). The approach to classification is supervised, and thus requires training data that characterizes the desired output classes. One of the strengths of Fuzzy ARTMAP is that it allows explicitly for “many-to-one” mapping, or the existence of many spectral subclasses within any desired output map class. This feature is essential to our application, as the *no forest change* class can include everything from forested areas which have not changed to all nonforest classes to areas that were not forest in the original image, but did

change in the second image. While the spectral variability of classes must be present in the training data, there is no requirement to sort the training data into the spectral subclasses associated with an output class such as *no forest change*. This process is handled internally by Fuzzy ARTMAP through the generation of multiple F2 nodes for the various output classes. Carpenter et al. (1997) provides the best description of the Fuzzy ARTMAP architecture as applied in classification of remote sensing images.

The training data for Fuzzy ARTMAP are simply examples of each class from registered pairs of Landsat TM images, or training sites. Preprocessing of the Landsat data consists of automated image-to-image registration, and atmospheric correction via dark object subtraction (Chavez, 1989). Atmospheric correction is necessary as the trained classifiers are ultimately applied to different images and scenes than those originally used for training. As such, the use of surface reflectances is necessary to minimize effects related to variable sun angles and atmospheric conditions between images. Our use of the initial, and somewhat simple, form of dark object subtraction is based on prior tests of the influence of various image-based atmospheric correction methods on generalization of mapping and change detection processes (Pax-Lenney et al., 2001; Song, Woodcock, Seto, Pax-Lenney, & Macomber, 2001). These studies have shown that the simpler methods of atmospheric correction often work as well or better than more recently developed methods that are more complicated. This surprising result points to the need for improved methods for minimizing the influence of the atmosphere on remote sensing images. This issue is particularly important for the process of generalization.

The output from Fuzzy ARTMAP is a per-pixel classification with just two classes. The per-pixel results tend to be noisy due primarily to the effects of minor misregistration. These effects include strings of pixels along distinct boundaries between forest and nonforest in areas where no change has occurred. There is also some speckle in the classification results, as is common to many per-pixel classification and change detection results. We apply two steps to minimize the noisy nature of the Fuzzy ARTMAP results. First, we use erosion followed by dilation. The most beneficial effect of the erosion is to remove isolated pixels or lines of pixels from the *forest change* class where no forest change has occurred. The erosion process also reduces the size of groups of contiguous pixels called *forest change*, which is undesirable. To offset this effect we follow the erosion process with a dilation process, which expands the size of groups of pixels remaining after the dilation (Russ, 1995). Second, we segment a two-date Band 5 image into an image of homogeneous landscape patches, or polygons (Woodcock & Harward, 1992). This step acknowledges that we are looking for forest change at the patch or stand scale. For all tests in this paper, the minimum polygon size is set to 2 ha, a common minimum size for polygons in forest maps. A threshold for the percentage of *forest change*

pixels within a polygon required for a polygon to be labeled *forest change* is determined for each image via analyst inspection of alternative thresholds. Experience has indicated that relatively low thresholds on the order of 10% are common, indicating the conservative nature of the neural network *forest change* classes.

As is necessary with virtually all maps made with remotely sensed data, a postclassification editing step is necessary. This editing is done at the scale of polygons. An analyst reviews each forest change map and relabels individual polygons as needed using color composites of each image and a Band 5 difference image. The Band 5 difference image is particularly helpful for finding patches of forest change which have been omitted from the map.

### 3. Analyses and results

#### 3.1. Test I: The Cascade Range of Oregon

We made a map of forest change in the Cascade Range of Oregon from 1991–1995. The study area includes portions of six Landsat scenes (Table 1). One of the reasons we chose this region is because the USDA Forest Service (Region 6) recently mapped this area using more conventional methods based on unsupervised classification techniques where each pair of images was independently classified and labeled by image analysts. The USDA Forest Service generously made their data and final map available to us, which provided us with an opportunity to evaluate the results of methods based on generalization with the current state-of-the-art methods they employed.

To map change in the Cascades, a neural network was trained with data from one pair of images and applied to six different pairs of images. Thus this map involves generalization across time and space (*within-region*). The quantity of training data is small given the size of the region. A total of 35 training sites were used to train the net: 9 *forest change* and 26 *no forest change*.

A small window from the two maps and the images from which they are derived is shown in Fig. 1. A visual

Table 1  
Landsat Images Cascades Range, Oregon

WRS		Images	
Path	Row	Early (yyymmdd)	Late (yyymmdd)
<i>Training images</i>			
45	29	920803	950913
<i>Testing images</i>			
45	29	910902	950913
45	30	910801	950913
45	31	910801	950913
46	28	910909	950819
46	29	920810	950819
46	30	910909	950803

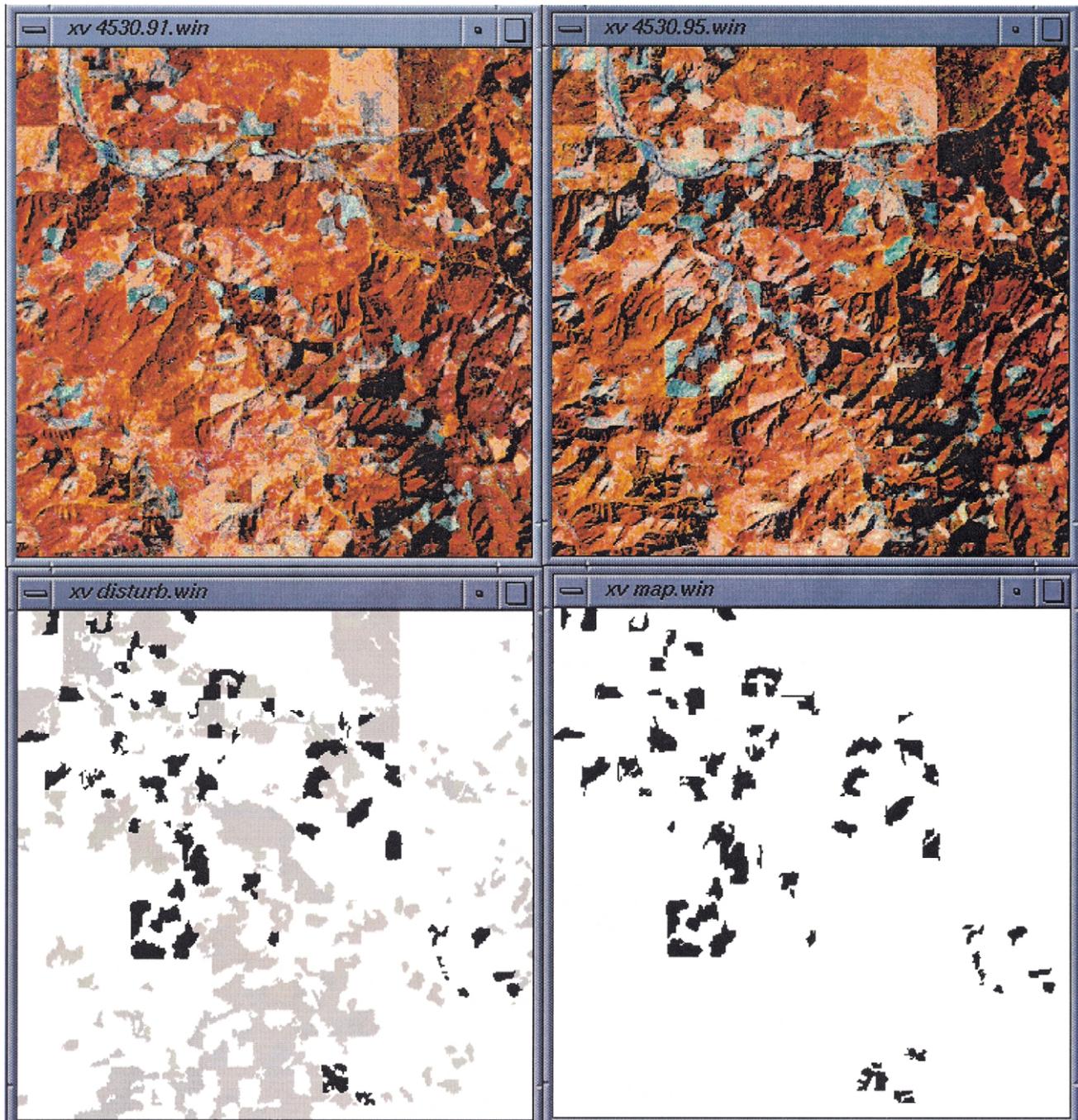


Fig. 1. This figure illustrates the strong similarities between the maps of forest change produced using traditional methods and generalization methods. The top two images are Landsat 5 TM images from 1991 (L) and 1995 (R) of a 15 km<sup>2</sup> region in the southern Cascades of western Oregon. The lower left image is a window from the USDA Forest Service map; black represents forest change between 1991–1995, grey represents forest change from earlier time periods. The lower right image is a window from the map produced using the generalization method; black represents change between 1991–1995.

comparison of the two maps shows strong correspondence between areas identified as forest change.

An accuracy assessment of the map was performed based on visual inspection by two examiners of a stratified, random sample of 536 3 by 3 pixel sites. For manifest forest change, such as clearcuts and forest fire, Cohen, Fiorella, Gray, Helmer, and Anderson (1998) demonstrated the validity of visual inspection as compared with fieldwork

for interpreting forest change in this environment. The *forest change* class covered 1.3% of the area within which 147 sites were randomly selected. For the *no forest change* sites we biased the sample to try to find areas of forest change that might have been missed by our change detection methods. For purposes of sampling, the area mapped as *no change* was separated into strata of *error least likely* and *error most likely* using a threshold in a Band 5 difference

image as an indicator of likely change. The *error least likely* stratum covered 91.7% of the image and 263 sites were drawn from this area; the *error most likely* stratum covered 7% of the image and 126 sites were drawn from this area. The results of the accuracy assessment are presented in Table 2a. In this table, the three map strata are: Change; No-Change (a) (error least likely); and No-Change (b) (error most likely).

The same accuracy assessment sites were also evaluated on the map made by the USDA Forest Service, and their accuracy is reported in Table 2b. The two sets of methods produced very similar overall accuracies, which indicates that the methods presented in this paper based on generalization are viable for monitoring forest change. More specifically, this result indicates that generalization *within-scene* and *within-region* are feasible for monitoring forest change.

There are lessons to be learned from examining the patterns of observed errors. First, misregistration in some local areas resulted in errors of commission in the *forest change* class. In part, this is a problem inherent in this particular set of images, which were obtained from the USDA Forest Service. To ensure we could directly compare our results with the existing map from the USDA Forest Service, we used the images directly as they were provided, which included image registration. We expect to be able to reduce this problem in the future with more precise image-to-image registration.

The second cause of errors of commission in the *forest change* class was the inclusion of areas of change that were not initially forest, such as agricultural lands and areas of shrubs or brush. The USDA Forest Service minimized this problem by overlaying a map of “forest” on the map of “change” to produce a map of “forest change.” We did not have access to this forest map. We expect that this problem can be greatly reduced in future work by increasing the number of training sites illustrating this sort of “nonforest change” in the *no change* class.

There are more errors of omission than commission in the *forest change* class. Many of these errors result from the segmentation process. At times, the segmentation process

produced very large polygons that included areas of both *change* and *no change* in the same polygon. Typically, in these large polygons the whole polygon is better characterized as *no change*. However, some of the 3 by 3 accuracy assessment sites that fall in these polygons are best characterized as *forest change*. In the future we can minimize this problem by reducing the maximum allowable size of polygons and by using a slightly different version of the segmentation algorithm that would include the per-pixel based *change/no change* output from the neural net along with the original spectral bands in the segmentation process (Woodcock, Collins, Jakabhazy, & Macomber, 1993). The idea that many of the observed errors are due to understandable and potentially correctable causes is encouraging.

Another relevant question concerns the time and effort required for the methods based on generalization relative to conventional methods. For the USDA Forest Service map, each pair of images was classified separately using unsupervised image classification and analyst labeling. For the time period covered in this analysis (1991–1995) there are six Landsat scenes required, and hence six separate image classifications were necessary. In contrast, for the methods based on generalization, a single neural network is trained and then applied to each of the six pairs of Landsat images. Thus the tradeoff is primarily between labeling six unsupervised image classifications and the identification of training sites from one pair of images that characterize two classes. The identification and digitizing of the training sites could be accomplished in a single morning’s work, which is much faster than the labeling of the unsupervised classifications. This tradeoff becomes more significant as the number of images requiring classification increases due to an increase in the size of the area to be mapped or the frequency with which forest change is to be monitored. The entire USDA Forest Service map covers four time periods and six scenes, and so ultimately required 24 separate image classifications. We only mapped one time period, but to add the other three time periods to the analysis there would be no need to retrain the classifier, and hence the efficiency grows.

The most time-consuming parts of the process are the postclassification editing and the segmentation process. Generally, the editing took one analyst 1 day per scene. Regardless of the classification method used, some postclassification editing will be required, so it does not figure significantly in the comparison of methods. The segmentation process is slow and required approximately 24 h of computing time per scene. While segmentation is time-consuming, it is necessary to convert the per-pixel classification output to a polygon-based map. Thus, this time-consuming step is not restricted to the generalization method.

### 3.2. Test II: The Colorado Rocky Mountains

As a second test of the generalization approach to monitoring forest change, the same methods were applied to an area of the Rocky Mountains in Western Colorado.

Table 2  
Cascade Mountains Forest Change 1991–1995 Accuracy Assessments

Map	Truth		User’s Accuracy (%)
	Change	No Change	
<i>a) Map based on Generalization</i>			
Change	138	9	93.9
No-Change (a)	8	255	97.0
No-Change (b)	5	121	96.0
Producer’s Accuracy (%)	91.4	97.7	95.9
<i>b) USDA Forest Service Map</i>			
Change	131	5	96.3
No-Change (a)	17	259	93.8
No-Change (b)	3	121	97.6
Producer’s Accuracy (%)	86.8	98.7	95.3

This test includes generalization across sensors: the neural net was trained with data from one pair of Landsat 5 TM images and it was applied to image pairs of Landsat 5 TM and image pairs combining Landsat 5 TM and Landsat 7 ETM+ data (Table 3).

The same basic steps applied in the Cascade Range of Oregon were applied to the Colorado Rockies. The images were registered and atmospherically corrected. A neural net was trained with local examples of *forest change* and *no forest change* and the net was applied to five pairs of images. The net output was eroded and dilated, and converted to a polygon-based image through segmentation. A *no change, error most likely* stratum was created for the accuracy assessment based on a Band 5 difference threshold. A cloud/cloud-shadow mask and a snow mask were applied using two independently trained networks. The main difference between the Cascade Range and Rocky Mountains analyses is that we did not create a map product for the Rockies. Therefore, we did not do the postclassification editing, which means that we would expect site-based accuracies for the Rockies to be lower than for the Cascades.

There are five separate accuracy assessments for the Rockies: one for each pair of images (Table 4). For each pair, we evaluated a stratified, random sample of  $3 \times 3$  pixel sites from each of the three strata: Change; No-Change (a) (error least likely); and No-Change (b) (error most likely). During the 1997–1999 period, there was little forest change in this region due to significant recent changes in land management policies and the number of sites selected in the *forest change* class reflects a shortage of available change sites.

Generally, accuracies are high for all five image pairs. Sources of *forest change* commission errors are similar to those found in the Cascade Range: (1) mixed polygons including both stable and changed forest; (2) areas of real change, but nonforested lands; and (3) snow and cloud contamination. In addition, there are three sites that are simply wrong. In the *no change, error least likely* class, there are no errors. Five of the six errors in the *no change, error most likely* class are change sites that we simply missed. At this point, no explanation is apparent.

Table 3  
Landsat Images Rocky Mountains, Colorado

WRS		Images	
Path	Row	Early (yyymmdd)	Late (yyymmdd)
<i>Training Images</i>			
34	32	920705	940625
<i>Testing Images</i>			
34	32	920705	940625
34	32	940625	970703
34	32	970703	990701 <sup>a</sup>
34	33	970703	990701 <sup>a</sup>
34	34	970703	990701 <sup>a</sup>

<sup>a</sup> Landsat 4 ETM+ data.

Table 4  
Rocky Mountains Forest Change Accuracy Assessments

Map	Truth	
	Change	No Change
<i>Scene Image 3432.9294</i>		
Change	43	7
No-Change (a)	0	50
No-Change (b)	0	50
<i>Scene Image 3432.9497</i>		
Change	45	5
No-Change (a)	0	50
No-Change (b)	1	49
<i>Scene Image 3432.9799</i>		
Change	17	1
No-Change (a)	0	50
No-Change (b)	3	47
<i>Scene Image 3433.9799</i>		
Change	6	2
No-Change (a)	0	50
No-Change (b)	2	48
<i>Scene Image 3434.9799</i>		
Change	11	0
No-Change (a)	0	50
No-Change (b)	0	50

More important to this study in the Rockies than the overall accuracies (as a standard for comparison is not available based on more conventional methods) is the pattern of accuracies between different combinations of training and application data. Training data were extracted from a pair of Landsat 5 TM images. The trained neural network was then applied to pairs of images that are either both Landsat 5 TM or combinations of Landsat 5 TM and Landsat 7 ETM+ (see Table 3). There are no immediately apparent differences in the overall accuracies of the tests between pairs of Landsat 5 TM images and the combinations of Landsat 5 TM and Landsat 7 ETM+ images. This initial success in generalization across sensors indicates the potential for *across-sensor* generalization within the Landsat program.

### 3.3. Test III: Generalization across regions

In an attempt to understand the limits to the range of generalization possible for monitoring forest change using Landsat, we did an exploratory test of the ability to generalize *across-regions*. For this test, we simply applied the neural network trained in the Cascades of Oregon to the images from the Rockies and vice versa. We did not attempt to make maps from these results, so postclassification editing of the results was not done. Instead, the distributions of raw neural net *forest change* pixels within the accuracy assessment sites that are truly forest change were calculated (Fig. 2). The intent of these distributions is to provide an

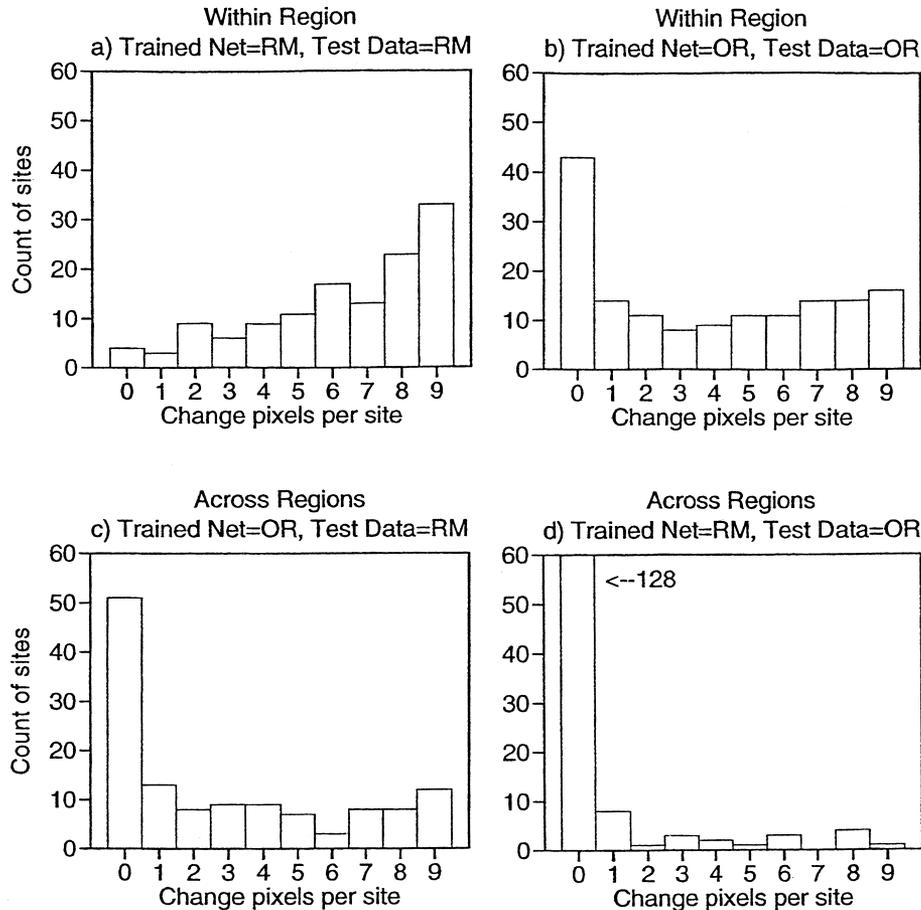


Fig. 2. These four charts compare *within-region* (a and b) and *across-region* (c and d) neural net results. In the Rocky Mountains region (RM), there were 128 change sites (test sites known to have changed), and in Oregon (OR), 151 change sites, each having 9 pixels. Ideally, for each change site the neural net would label all 9 pixels as change. Therefore, one measure of how well each trained neural net performs is how many change pixels it finds in each site, and by extension, how many change sites have 9 change pixels, 8 change pixels, etc. For the Rocky Mountains region, (a) shows what one would hope to see: when the neural net is trained *within-region*, most of the test sites contain *change* pixels. (c) shows that a neural net trained in a different place (*across-region* training) finds fewer change pixels per site. (b) and (d) show similar results for tests in Oregon.

indication of the performance of various combinations of training and testing locations. The *within-region* results are another way of looking at the results from the analyses described above for the Cascades in Oregon and the Rockies in Colorado. The *within-region* results for the Rockies most closely approximate the desired distribution, with many of the change sites having a high proportion of *forest change* pixels. For the Cascades, the *within-region* results are less impressive, yet effective for mapping as the accuracy assessment presented in Table 1 shows.

The *across-region* results illustrated in Fig. 2 are discouraging with respect to the potential for generalization *across-regions*. The results for the combination of training in the Rockies and testing in the Cascades are particularly grim, with 128 out of 151 sites of forest change showing no neural net *forest change* pixels. The distribution for the other across-region test (train in the Cascades and test in the Rockies) shows a somewhat improved distribution of *forest change* pixels, but there are still many sites where forest change is not found by the neural net classifier.

The relative performance of the two *across-region* tests are somewhat surprising. Since the *within-region* tests show that the neural net trained in the Rockies performed better than the one trained in the Cascades, an argument could be made for expecting the results of the application of the Rockies neural network in other regions to be better than the Cascades neural network. However, that pattern is not found in the data in Fig. 2. Instead, the results for both the *within-region* and *across-region* analyses are worse for the Cascades than the Rockies, which supports the idea that change is more difficult to identify in the Cascades.

#### 4. Discussion

The result from the Cascades which shows that methods based on generalization in time and space (*within-region*) provide comparable accuracies to state-of-the-art methods is encouraging with respect to the potential for large area monitoring of forest change using Landsat. While it is

important to acknowledge that the problem of forest change is relatively simple, the results presented here support the idea that generalization of mapping processes can be effective. It is our expectation that there will be tradeoffs between the level of detail of discrimination of surface properties and the ability to generalize mapping processes. Determining the limits of such tradeoffs will be a challenge for future research.

The results from the Cascades are particularly encouraging given that this analysis represents a “first attempt” from the point of view of generalization. The state-of-the-art methods based on training within the domain of application have evolved and developed over almost a quarter of a century. Surely with more experience and improvement in the methods used for generalization the results can be improved.

The results presented in this paper indicate several areas where improvements are needed. First, the per-pixel results from the neural network tend to be too conservative about finding forest change. This result is indicated by:

1. the low threshold of *forest change* pixels within polygons used to determine which polygons should be mapped as *forest change*;
2. the distributions of *forest change* pixels within the accuracy assessment sites identified as forest change, as shown in Fig. 2; and
3. the errors of omission in the *forest change* class exceed the errors of commission in the study in the Cascades.

Second, future research is needed to understand how the properties and amounts of training data for the various classes influence this result. For example, the full range of kinds of forest change found in the areas may not be represented in the training data. Finally, changing the relative amounts of training data within classes has been shown to influence the resulting size of classes in classifications (Borak, 1996; Carpenter, Gopal, Macomber, et al., 1999). Understanding these effects will be necessary before operational use of neural networks can be used most effectively for monitoring large areas.

The results from the Cascades also indicate a need for improved training data for the *no forest change* class. The accuracy assessment revealed that areas that were not initially forest but did change between 1991 and 1995 were frequently included inappropriately in the *forest change* class. The most common example was agricultural land, where change between time periods is the norm. We did not foresee this problem and did not include many examples of agricultural land in the training for the *no forest change* class. This should be changed in the future. Also, many of the errors resulted from polygons defined in the segmentation process including both areas of change and no change. Improvements in this step would improve the resulting maps.

The success of the initial attempts to generalize across Landsat sensors is particularly encouraging given the exist-

ing archive of Landsat imagery and points to the importance of continuity of Landsat-style measurements. This factor should be taken into account in plans for future Landsat missions or other missions intended to meet the needs for Landsat-style measurements.

Methods based on generalization hold the potential for rapid assessment of change over large areas and/or frequent time intervals. For the potential of this approach to be realized, large amounts of imagery will be necessary at reasonable costs, which is exactly what Landsat 7 is producing. Most important in this regard is the availability of images. Four components of the Landsat 7 program contribute dramatically to improved availability of images:

1. the rate at which images are being collected and archived (up to 250 images per day) is a dramatic improvement over the past;
2. the strategy for image acquisition implemented through the Long Term Acquisition Plan is helping ensure systematic, global, seasonal coverage;
3. data pricing based on the “cost of filling user orders” has dropped prices by almost an order of magnitude; and
4. the lifting of restrictions on distribution of images following purchase will allow sharing of imagery.

The net effect of improved availability of Landsat imagery will improve the ability to monitor environmental change.

The *across-region* results, while discouraging, need to be viewed as preliminary. One possible problem with the tests reported concerns the dates of the images in the two regions. The image pairs are closer in date within the Cascades and Rockies than they are between the two areas (see Tables 1 and 3). Thus, phenological differences between the two areas could be undermining the ability to generalize the classifiers. Considerably more work is warranted to determine if approaches which include training data from multiple geographic regions can produce neural networks that are more applicable across regions.

One issue of tremendous significance for generalization is the conversion of image data from the original units of digital numbers (or DNs) to a stable unit of measurement such as surface reflectance that can be compared across space and time. There are two key steps in this process: the conversion of DNs to radiances, and then conversion of radiances to surface reflectance. The second step is the most important, with the problem of compensating for varying atmospheric conditions between images being the most significant hurdle. Atmospheric correction in the absence of independent measurements of atmospheric conditions at the time of the satellite overpass is particularly challenging. While a number of new methods have shown promise for improved atmospheric correction (see for example, Liang et al., 1997; Wen, Tsay, Cahalan, & Oreopoulos, 1999), they do not always lead to better datasets for generalization of mapping processes (Song et al., 2001). Thus, it may prove

important to evaluate methods of atmospheric correction on the basis of their influence on mapping processes.

## 5. Conclusions

The accuracies achieved for detecting forest change in the Cascade Range in Oregon using methods based on generalization of training data in time and within a geographic region are similar to those achieved using state-of-the-art methods, which indicates that methods based on generalization are viable for large area monitoring of forest change. Methods based on generalization require less time and effort than conventional methods and as a result may allow monitoring of larger areas or more frequent monitoring at reduced cost. One key component to achieving this goal is improved availability and affordability of Landsat imagery, which Landsat 7 is providing.

Generalization across Landsat sensors was successful, indicating that imagery from Landsat 7 ETM+ can be used in combination with prior Landsat TM images for monitoring forest change without the need to retrain change detection algorithms such as the neural networks used in this study. This result highlights the value of the existing archive of Landsat imagery and the importance for continuity in the Landsat Program.

Preliminary efforts to generalize across geographic regions (from the Cascades of Oregon to the Rockies of Colorado and vice versa) were discouraging and help define current limits to the extent of generalization possible for monitoring forest change. Future work on this problem is necessary if generalization across such distances and ecological/topographic/climatic differences is to become viable.

## Acknowledgments

This research was funded by the NASA Landsat 7 Science Team, NASA Grant NAS5-3439.

## References

- Abuelgasim, A. A., Ross, W. D., Gopal, S., & Woodcock, C. E. (1999). Change detection using adaptive fuzzy neural networks: environmental damage assessment after the Gulf War. *Remote Sensing of Environment*, 70 (2), 208–223.
- Borak, J. S. (1996). *Feature selection and land cover classification of a MODIS-like data set for a semiarid environment*. MA thesis, Department of Geography, Boston University, Boston, MA, 58 pp.
- Botkin, D. B., Estes, J. E., Macdonald, R. M., & Wilson, M. V. (1984). Studying the Earth's vegetation from space. *BioScience*, 34 (8), 508–514.
- Carpenter, G. A., Gajaja, M. N., Gopal, S., & Woodcock, C. E. (1997). ART neural networks for remote sensing: vegetation classification from Landsat TM and terrain data. *IEEE Transactions on Geoscience and Remote Sensing*, 35 (2), 308–325.
- Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., Woodcock, C. E., & Franklin, J. (1999a). A neural network method for efficient vegetation mapping. *Remote Sensing of Environment*, 70 (3), 326–338.
- Carpenter, G. A., Gopal, S., Martens, S., & Woodcock, C. E. (1999b). A neural network method for mixture estimation for vegetation mapping. *Remote Sensing of Environment*, 70 (2), 138–152.
- Chavez, P. S. Jr. (1989). Radiometric calibration of Landsat Thematic Mapper multispectral images. *Photogrammetric Engineering and Remote Sensing*, 55 (9), 1285–1294.
- Cohen, W. B., Fiorella, M., Gray, J., Helmer, E., & Anderson, K. (1998). An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, 64 (4), 293–300.
- Cohen, W. B., Maersperger, T. K., Spies, T. A., & Oetter, D. R. (2001). Modeling forest cover attributes as continuous variables in a regional context with Thematic Mapper data. *International Journal of Remote Sensing*, in press.
- Coiner, J. C. (1980). Using Landsat to monitor changes in vegetation cover induced by desertification processes. *Symposium on Remote Sensing of the Environment*, 3 (14), 1341–1347 (Ann Arbor, MI: Environmental Research Institute of Michigan).
- Coppin, P. R., & Bauer, M. E. (1994). Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features. *IEEE Transactions on Geoscience and Remote Sensing*, 32 (4), 918–927.
- Gopal, S., & Woodcock, C. E. (1996). Remote sensing of forest change using artificial neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 34 (2), 398–404.
- Gopal, S., Woodcock, C. E., & Strahler, A. H. (1999). Fuzzy neural network classification of global land cover from a 1-degree AVHRR data set. *Remote Sensing of Environment*, 67, 230–243.
- Hall, F. G., Botkin, D. B., Strebel, D. E., & Goetz, S. J. (1991). Large-scale patterns of forest succession as determined by remote sensing. *Ecology*, 72, 628–640.
- Jenson, J. R., & Toll, D. L. (1982). Detecting residential land use development at the urban fringe. *Photogrammetric Engineering and Remote Sensing*, 48, 629–643.
- Jha, C. S., & Unni, N. V. M. (1994). Digital change detection of forest conversion of a dry tropical Indian forest region. *International Journal of Remote Sensing*, 15 (13), 2543–2552.
- Liang, S., Fallah-Adl, H., Kalluri, S., JaJa, J., Kaufman, J. J., & Townshend, J. R. G. (1997). An operational atmospheric correction algorithm for Landsat Thematic Mapper imagery over the land. *Journal of Geophysical Research*, 102, 17,173–17,186.
- Lunetta, R. S., & Elvidge, C. D. (1998). *Remote sensing change detection: environmental monitoring methods and applications* (pp.318). Chelsea, MI: Ann Arbor Press.
- Minter, T. C. (1978). Methods of extending crop signatures from one area to another. *Proceedings. The LACIE Symposium. A Technical Description of the Large Area Crop Inventory Experiment (LACIE)*. October 23–26, 1978, Houston, TX.
- Pax-Lenney, M., Woodcock, C. E., Collins, J., & Hamdi, H. (1996). The status of agricultural lands in Egypt: the use of multitemporal NDVI features derived from Landsat TM. *Remote Sensing of Environment*, 56, 8–20.
- Pax-Lenney, M., Woodcock, C. E., Gopal, S., & Macomber, S. A. (2001). Monitoring Temperate Conifer Forests with Landsat TM: A new look at classification generalization. *Remote Sensing of Environment*, in press.
- Russ, J. C. (1995). *The image processing handbook* (2nd ed., pp.674). Ann Arbor, MI: CRC Press.
- Seto, K. C., Song, C., & Woodcock, C. E. (1999). Land-use land-cover change in the Pearl River delta, 1988–1996: evidence from Landsat TM. *Proceedings, 1999 ASPRS Annual Conference, May 19–21* (pp. 633–636). Portland, OR.
- Singh, A. (1989). Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, 10, 989–1003.

- Song, C., Woodcock, C. E., Seto, K., Pax-Lenney, M., & Macomber, S. A. (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment*, 75, 230–244.
- Vogelmann, J. E., & Rock, B. N. (1988). Assessing forest damage in high-elevation coniferous forests in Vermont and New Hampshire using Thematic Mapper data. *Remote Sensing of Environment*, 24, 227–246.
- Wen, G., Tsay, S. C., Cahalan, R. F., & Oreopoulos, L. (1999). Path radiance technique for retrieving aerosol optical thickness over land. *Journal of Geophysical Research*, 104, 31,321–31,332.
- Woodcock, C. E., Collins, J., Jakabhazy, V. D., & Macomber, S. A. (1993). *Technical manual: forest vegetation mapping methods designed for Region 5 of the US Forest Service*. Technical Paper Number 7, Boston University Center for Remote Sensing, 105 pp.
- Woodcock, C. E., & Harward, V. J. (1992). Nested-hierarchical scene models and image segmentation. *International Journal of Remote Sensing*, 13 (16), 3167–3187.