ISSN 1946-7664. MCFNS 2016
Vol. 8, Issue 1, pp. 4–13
Mathematical and Computational Forestry & Natural-Resource Sciences

AVAILABLE ONLINE AT HTTP://MCFNS.COM

Submitted: Feb. 10, 2016
Accepted: Mar. 26, 2016
Published: Mar. 30, 2016
Last Correction: Mar. 27, 2016

AUTOMATED ESTIMATION OF FOREST STAND AGE USING VEGETATION CHANGE TRACKER AND MACHINE LEARNING

Jobriath S. Kauffman, Stephen P. Prisley

Center for Natural Resources Assessment & Decision Support, Virginia Polytechnic Institute & State University, USA

Abstract. The ability to automatically delineate forest stands and determine their age is useful for natural resources professionals. Two common approaches to estimating forest area and age-class distributions are inventory-based methods, such as Forest Inventory and Analysis (FIA), and remote sensing based methods. Vegetation Change Tracker (VCT) is an algorithm that uses time series stacks of Landsat images to identify forest disturbances. However, additional computation is required to identify type of disturbance. This paper evaluates the usefulness of machine learning tools, such as support vector machine (SVM), for reclassifying VCT disturbances as stand-clearing disturbances or partial disturbances. Overall accuracy for a 2010 VCT disturbance map of the entire state of Virginia was determined to be 87 percent. 100 percent of 2010 Virginia clearcut harvests recorded in a reference dataset were classified as disturbances by VCT. Neighboring disturbed pixels, as classified by VCT, were clumped together and reclassified as stand-clearing disturbances or partial disturbances using SVM and variables for average disturbance magnitude and shape and size metrics of the clumped pixels, with an overall accuracy rate of 86 percent. The users and producers accuracy rates for stand-clearing disturbances were 88 percent and 95 percent respectively. In addition, an algorithm was developed in R for determining years since last stand-clearing disturbance for each pixel in a time series stack of reclassified VCT disturbance maps from 1984 to 2011. Neighboring pixels of the same age, in number of years since last stand-clearing disturbance, were clumped together and correspond, in general, to clearcut harvest boundaries.

Keywords: forest disturbance, harvest type, forest age, harvest delineation, automated, machine learning.

1 Introduction

The fact that there is value in mapping forest stands is unquestioned. Distributions of forest by stand age and type at various spatial scales provide valuable information for optimizing forest production and sustainability. Both field measured forest inventories and remotely sensed data have been used to estimate forest area and age. Each of these methods has its strengths and limitations. Precise field measured inventory estimates from sample plots are possible over large areas but are costly for fine scale estimates across a large area due to the large number of sample plots required. Remotely sensed data can be obtained just as easily for a specific point as it can for an entire image. Automated processing of these large amounts of data is an obstruction that is becoming easier to overcome. While cost effective methods for yearly, statewide, border-to-border mapping of forest by age and major species group have been elusive they are nonetheless obtainable.

The U.S. Forest Service Forest Inventory and Analysis program (FIA) comprises the most comprehensive field measurement of forest inventory today. The sampling process utilizes a hexagonal grid system placed over the forty-eight contiguous states. Each of the hexagons covers approximately six thousand acres. A forest inventory plot is randomly located within each hexagon. Trees are measured on four subplots within each plot, totaling approximately 1/6 of an acre (Bechtold and Patterson 2005).

Virginia contains approximately 4700 sampling hexagons. This large sample size allows for accurate estimates over large areas comprising the entire state or multiple counties. However, Figure 1 demonstrates that there can be a great amount of variation in forest distribution over a six thousand acre hexagon that cannot be represented on a small scale by one FIA plot. This highlights the value of smaller scale representation of timber products and biomass at the stand or pixel level.

Figure 1: Representation of variation in forest distribution over 6000 acre hexagons in central Virginia.

Timber product volume and biomass estimates at small scales are difficult to obtain but increasingly valuable for sustainable forest management. Maps of forest by major species groups are common, including nationwide land cover mapping such as the 30 meter pixel scale National Land Cover Database (NLCD) land cover map (Homer et al. 2015). Age estimates, in combination with forest type, from remote sensing data at the stand or pixel level can make a major contribution towards further refining volume estimates at this small scale.

A simple way to calculate "age" of a forest is to measure the number of years since the last clearcut. Algorithms using time series stacks of Landsat data, such as Vegetation Change Tracker (VCT), have proven to be reliable for detecting forest disturbances (Huang et al. 2010). After other known dark objects such as water and dark soils have been masked from Landsat images dating back to 1984 (Landsat 4), VCT uses forest training pixels identified from the forest peak in histograms of top of atmosphere reflectance in the near infrared and two short-wave infrared spectral bands (Huang et al. 2008, Huang et al. 2010). The means and standard deviations of these training pixels in the red and shortwave infrared bands are used to calculate an integrated forest z-score (IFZ) for each pixel in the image (Huang et al. 2010).

In time series of yearly height of season IFZ values, forested pixels will remain persistently below a threshold IFZ value, while non-forested pixels will remain above the threshold or fluctuate above and below it (Huang et al. 2010). Thus, a sudden increase in a pixel's otherwise persistently low IFZ score indicates the timing of a forest disturbance within the time series. In this way, the VCT algorithm described by Huang et al. (2010) can be used to generate VCT products such as yearly disturbance maps at the 30 meter pixel level. The magnitude of these disturbances can also be calculated by finding the difference between a pixel's average IFZ score (or other index) and its IFZ score for the disturbance year. Normalized difference vegetation index (NDVI) and normalized burn ratio index (NBRI), and IFZ4 are also incorporated in the VCT algorithm and used to calculate similar measures of disturbance magnitude (Huang et al. 2010). NDVI measures photosynthetic capacity using the red and near-infrared bands. The calculation for NBRI is similar to NDVI but uses the near-infrared band and short-wave infrared band (band 7). Changes in NBRI can be used to measure burn severity. IFZ4 is calculated similarly to IFZ but uses only the near-infrared band. Maps of disturbance magnitudes of disturbed pixels measured each of these ways are in production for the contiguous United States.

Within secondary succession forests, especially those in which frequent harvest and regeneration occurs, groups of neighboring 30 meter pixels representing forest of the same age were most likely harvested together, or cleared by some other mechanism, at some point in the past. Clumping these neighboring pixels of the same age together can be used as a method for creating objects that conform to past harvest boundaries.

The Virginia Department of Forestry began keeping records of all harvests in Virginia in 2009 in order to facilitate inspections of best management practices (BMPs). The GPS point location of the first logging deck the BMP inspector comes to on the date of inspection is collected along with other harvest attributes. Typically, there are five or six thousand harvests per year in Virginia. Delineating these harvests with a GPS during inspection or by post-harvest photo interpretation is costly and time consuming. Therefore, efforts to automate this process and extend the records back to 1984, the first available year of VCT disturbance maps, are worthwhile.

The combination of VCT disturbance maps dating back to 1984 with several recent years of comprehensive harvest records creates a convenient avenue to extend the record historically, back to 1984, accurately placing their occurrence spatially and temporally. This process can be automated and used wherever VCT data is available, especially in areas dominated by harvest disturbances. Prior harvest locations are often the best indicators of where future harvests will occur. Therefore, in addition to extending harvest records historically, data of this type can be used to help forecast the locations of future harvests. Past and future harvest locations yield valuable information for many uses, including timber procurement, climate modeling, water resource modeling, and wildlife habitat analysis.

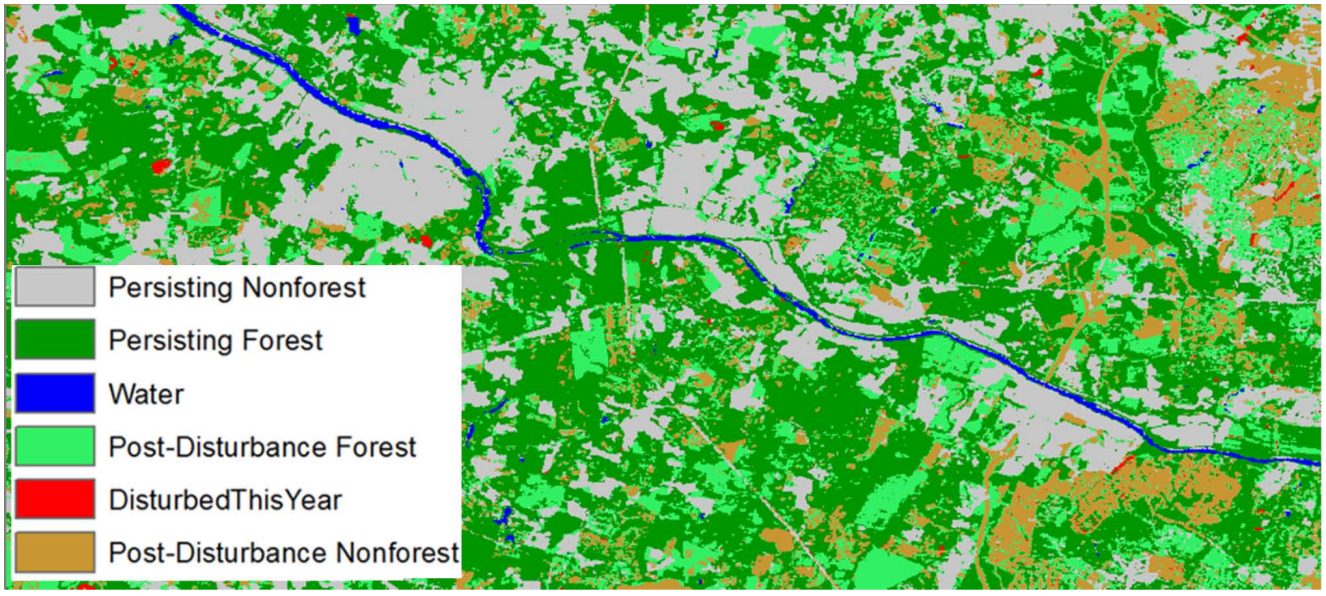

Figure 2: Portion of 2010 Virginia VCT map.

This research demonstrates procedures for using yearly VCT disturbance maps to create clumps of neighboring pixels that were disturbed collectively. These clumps can then be spatially linked to recent harvest records, and metrics related to shape, size, and average disturbance magnitude of each clump can be used to train machine learning tools used for classifying disturbances as either stand-clearing or partial. Throughout this paper the term "enhanced VCT" will be used when referring to VCT disturbance maps that have been reclassified to include only stand-clearing disturbances. Raster stacks of stand-clearing disturbance maps can be used to subsequently create an "age" map by calculating the number of years since the last stand-clearing disturbance. The process for creating a map of this type for Virginia will be described in the next section. VCT data products are anticipated to be available nationally, creating opportunities to repeat these methods anywhere in the contiguous United States, especially in areas dominated by harvest disturbances.

2 Methods

2.1 2010 VCT disturbance map validation VCT disturbance maps and disturbance magnitude maps created from the VCT algorithm were obtained for all of Virginia. An accuracy assessment of the 2010 VCT disturbance map for Virginia was performed using before and after aerial photography, 2008 National Agriculture Imagery Program (NAIP) and 2012 NAIP respectively. Figure 2 depicts a portion of the 2010 VCT disturbance map. VCT yearly disturbance maps are classified into six groups: persisting non-forest, non-forest after a disturbance, persisting forest, forest after a disturbance, forest disturbed in the current year, and water. 100 sample points within each class were randomly chosen. For simplicity the two non-forest groups (200 points total), and the two forest groups (200 points total) were combined together. This assessment can be used to validate the accuracy of using the map to identify each of these classes, including forest and forest disturbances.

2.2 Evaluating the ability to detect clearcuts In conjunction with the 2010 VCT disturbance map accuracy assessment the ability of VCT to detect clearcut harvest was evaluated. VCT mapped all sample points that were clearcuts as disturbances. In order to add weight to this assessment the Virginia harvest records were used to look for possible errors outside of the sample used for validation in which VCT classified a clearcut harvest as something other than a disturbance. All harvest locations that did not intersect a VCT disturbance were inspected using the before and after aerial photography in order to find out if there was perhaps a stand-clearing disturbance that VCT missed. If VCT is doing reasonably well at detecting stand-clearing disturbances, the next step would be to reclassify all VCT disturbances as stand-clearing or not.

2.3 Reclassifying forest disturbances by harvest method It is known that VCT detects yearly forest disturbances using a Landsat time series stack of one scene per year during the height of growing season (Huang et al. 2010). Sometimes more than one scene is used

to allow for more cloud free pixels. The latest scene used for Virginia in 2009 was taken on September 14th, while the earliest scene taken for Virginia in 2010 was on June 3rd. Therefore, point locations of Virginia timber harvest data records from the Virginia Department of Forestry for inspections between these dates were intersected with the 2010 VCT disturbance maps covering Virginia. This encompasses all of the land area in Virginia within Landsat path/row scenes 14/34, 14/35, 15/33, 15/34, 15/35, 16/33, 16/34, 16/35, 17/33, 17/34, 17/35, 18/34, 18/35, 19/34, and 19/35.

Cells classified as disturbances within each VCT disturbance map were evaluated for adjacency to neighboring disturbed cells with the Queen's case of 8 directions (right, left, above, below, or diagonal) using the 'clump' function in the R 'raster' package (Hijmans 2015). In this manner, connected groups of neighboring disturbed pixels were assumed to be the result of the same forest disturbance. Therefore, they were clumped together and given a common identity. Since VCT detects both stand-clearing disturbances and partial disturbances, the ability to calculate stand "age" must be facilitated by reclassifying VCT disturbances as stand-clearing or not. Classification of VCT disturbances by type using machine learning tools, including SVM, is appropriate and has proven to be effective (Zhao et al. 2015).

Average VCT disturbance magnitudes as measured by IFZ, NDVI, NBRI, and IFZ4 were obtained for each disturbance clump. In addition, various shape and size metrics were calculated for each clump using the 'PatchStat' function in the 'SDMtools' R package (VanDerWal et al. 2014). Clearcut harvests tend to have higher disturbance magnitudes, larger areas, and less complex shapes as they conform to more linear parcel boundaries.

There were 1170 VCT disturbance clumps from 2010 that intersected with one of the harvest site point locations in the date range specified above. Half of the disturbance clumps were used for training and half for validation of three machine learning tools used to reclassify VCT disturbance clumps as stand-clearing or partial disturbances. k nearest neighbor (kNN; Meyer et al. 2015), support vector machine (SVM; Venables and Ripley 2002) , and the 'rpart' (Therneau et al. 2015) R package classification algorithms were trained using a somewhat arbitrary selection of variables, including disturbance magnitudes measured three ways (IFZ, NDVI, and NBRI), area, and fractal dimension index of the training clumps. Fractal dimension index is a measure of shape complexity. The kNN algorithm classifies a point in the feature space according to the majority class of a predefined number of nearest neighbors. SVM is also a supervised classification algorithm that maximizes the separation between two classes with a hyperplane. The 'rpart' package is an implementation of the Classification and Regression Trees (CART) algorithm, following Breiman et al., 1984 in most details. CART and rpart recursively split the feature space by finding the value of a variable that separates the training sample data into classes that minimize incorrect classification. In an effort to maximize automation of the reclassification process, no further effort was made to calibrate the models or select the most advantageous features for improving accuracy. For the kNN algorithm, k=9 was arbitrarily chosen for the number of nearest neighbors. Kernel methods for producing non-linear classifiers are possible with SVM, but the default 'linear' kernel was used for simplicity. Each of the three trained models was used to classify each member of the validation set as a stand-clearing disturbance or a partial disturbance, and the results were compared to the actual harvest data.

The trained machine learning classification tools were used to reclassify each clump in each disturbance map as a stand-clearing disturbance or not based on the majority class of the three models. This process was automated and repeated for the 15 Landsat scenes for the entire study area, each year for the 28 years from 1984 to 2011, using R and its "raster" package (R Core Team 2015, Hijmans 2015). The resulting maps of stand-clearing disturbances are what has been defined above as enhanced VCT disturbance maps.

2.4 Calculating age R was also used to create a time series stack of the enhanced VCT maps in order to calculate the "age" of each pixel by determining the number of years since the last stand-clearing disturbance. Neighboring pixels of the same "age" using the Queen's case of 8 directions were clumped together and the "Eliminate" filter function in Erdas Imagine (Leica Geosystems 2013) was used to clean edges and eliminate small clumps less than five pixels. Thus, clumps smaller than one acre were removed, and those pixels were back-filled with information from the surrounding clumps.

3 Results

The overall accuracy rate of the 2010 VCT disturbance map was estimated to be 87 percent. According to this assessment, an estimated 100 percent of the 2010 Virginia stand-clearing disturbances were classified as disturbances by VCT. Furthermore, clearcuts from 2010 harvest records were overlaid with 2010 VCT disturbances. This was done to insure that there was not a clearcut harvest at the location of the harvest record that VCT did not pick up. It was confirmed that VCT did not miss any clearcut harvests. GPS point locations of harvests are taken at the first logging deck that the forestry official comes to when inspecting the logging site. Logging decks are most often on the edge of

Table 1: Error Matrix with percent accuracy rates for machine learning classification of disturbances larger than 1 acre as recorded by VCT.

Method		Error Matrix		
	Classification Data	Reference Data		
SVM		Non-Clear-Cut	Clear-Cut	
	Non-Clear-Cut	77	23	0.77
	Clear-Cut	57	428	0.882
		0.575	0.949	0.863
rpart		Non-Clear-Cut	Clear-Cut	
	Non-Clear-Cut	76	30	0.717
	Clear-Cut	58	421	0.879
		0.567	0.933	0.850
kNN		Non-Clear-Cut	Clear-Cut	
	Non-Clear-Cut	64	34	0.653
	Clear-Cut	70	417	0.856
		0.478	0.925	0.822

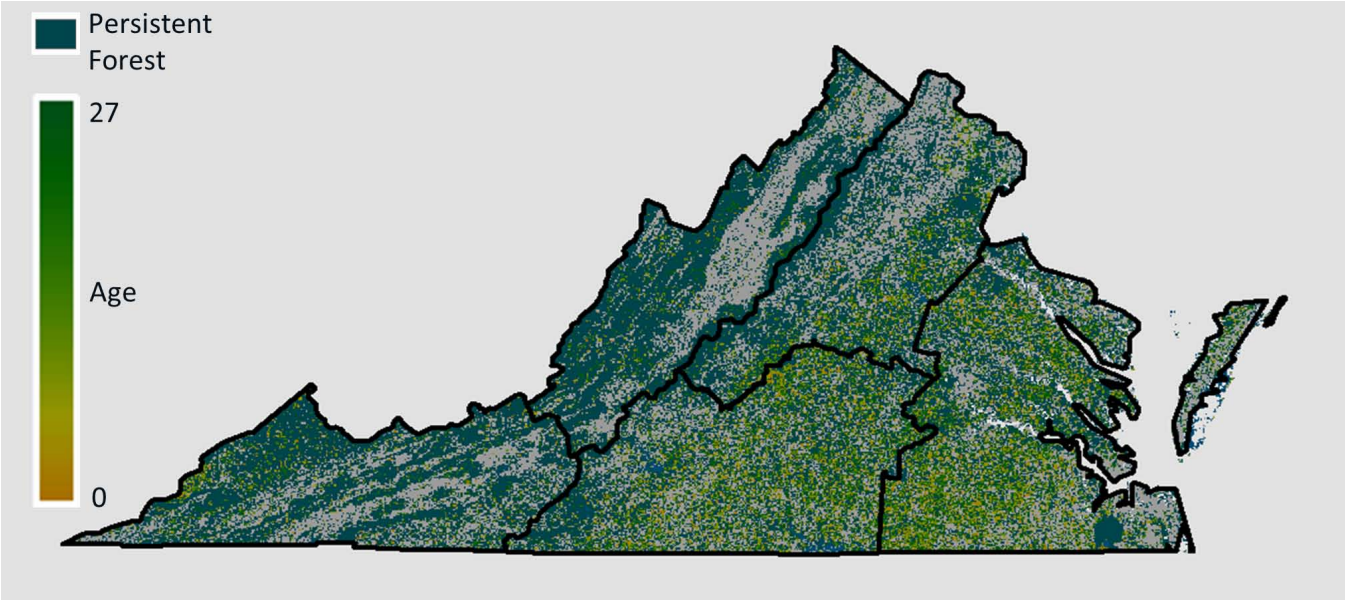

Figure 3: Virginia VCT "age" map enhanced by reclassifying disturbed pixels as stand-clearing or not.

the harvest site and are sometimes outside of the harvest site altogether, so it is not expected that all harvest point locations will be within the actual boundaries of the harvest.

The greatest overall reclassification accuracy rate was achieved with SVM at 86 percent (Table 1). SVM correctly classified 95 percent of clearcuts but misclassified 42.5 percent of partial harvests. The results for kNN and rpart were comparable.

An enhanced VCT "age" map for all of Virginia was created (Figure 3). This map shows age as the number of years since the last stand-clearing disturbance regardless of whether the cleared forest returned to forest or remained non-forest. The total number of acres that VCT considered persisting forest or disturbed forest for at least one year in the timespan from 1984 to 2011 is 17.6 million (Figure 4). This compares with 16.2 million acres if post-disturbance non-forest pixels as identified in the most recent VCT disturbance map are removed. Therefore approximately 8 percent of the pixels that are given a forest "age" in the enhanced VCT "age" map have either converted to non-forest, or have not regener-

ated to the point that VCT can discern it is forest, with IFZ values remaining above the forest threshold.

The enhanced VCT estimate of forest acres after removing post-disturbance non-forest pixels is within the 95 percent confidence bounds of FIA forest acre estimates using the 2011 population evaluation group, and is less than 2 percent higher than the FIA point estimate (Miles 2016). It is also comparable to forest area estimates using the 2011 NLCD land cover map (Homer et al. 2015). The NLCD "shrub/scrub" class includes true shrubs and young trees less than 5 meters tall. Much of what is classified as shrub in Virginia is actually early successional forest or trees stunted from environmental conditions. Therefore two NLCD totals are shown in Figure 4. The first includes the deciduous forest, evergreen forest, mixed forest, and woody wetlands classes without shrub. The second also includes the "shrub/scrub" class. The NLCD estimates are very close to the lower and upper 95 percent confidence bounds of the FIA estimate, respectively.

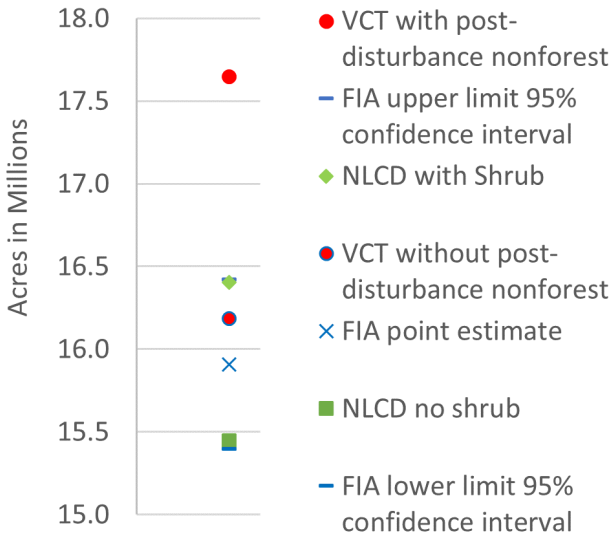

Figure 4: Estimates of forested acres in Virginia.

Figure 5 compares forest area estimates for the enhanced VCT with FIA in Virginia by age class (Miles 2016). These estimates generally coincide, but special note should be made where enhanced VCT estimates fall outside the confidence bounds of FIA. In addition to low area estimates when excluding the pixels that no longer look like forest in the 0-5, 6-10, and 21-25 year age classes, and the high estimates when including these pixels in the 0-5 and 11-15 year age classes, the enhanced VCT also overestimates forest acres when compared with FIA in the over 25 age group, even when post-disturbance non-forest pixels are excluded. The es-

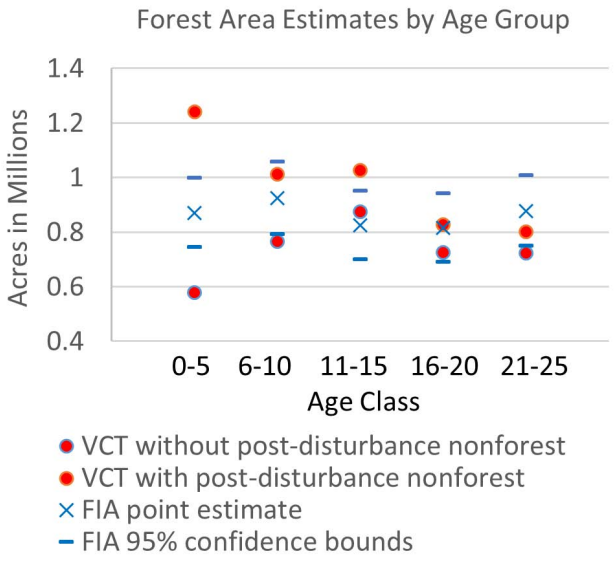

Figure 5: Forested acres in Virginia by age class, up to age 25.

timate without post-disturbance non-forest pixels in this age group is 12.5 million and rises to 12.7 million when these pixels are included. Possible reasons for these differences will be discussed in the next section.

Figure 6 shows the tendency of clumps of neighboring reclassified VCT disturbances of the same "age" to conform to actual harvest boundaries. These clumps also frequently conform to parcel boundaries (Figure 7). Figure 8 shows how the resulting enhanced VCT gets rid of isolated pixels and provides a more realistic determination of stand "age" by reclassifying disturbed pixels in a 2010 thinning as a partial disturbance, giving the accurate number of years since the most recent stand-clearing disturbance.

4 Discussion

This study demonstrates both the value of comprehensive harvest records for use in training and validating machine learning models as well as the ability to successfully classify disturbance clumps by harvest method utilizing shape metrics in an automated manner. It is the intent that similar automated methods will be used to reclassify VCT disturbances and create forest age maps and harvest boundaries in other areas of interest.

In an effort to maximize automation for ease and efficiency in creating future age maps and harvest boundaries, the full capacity of the machine learning tools for correctly classifying disturbances as clearcut or partial harvests was not realized. Features were chosen without testing to see which set might produce the highest accu-

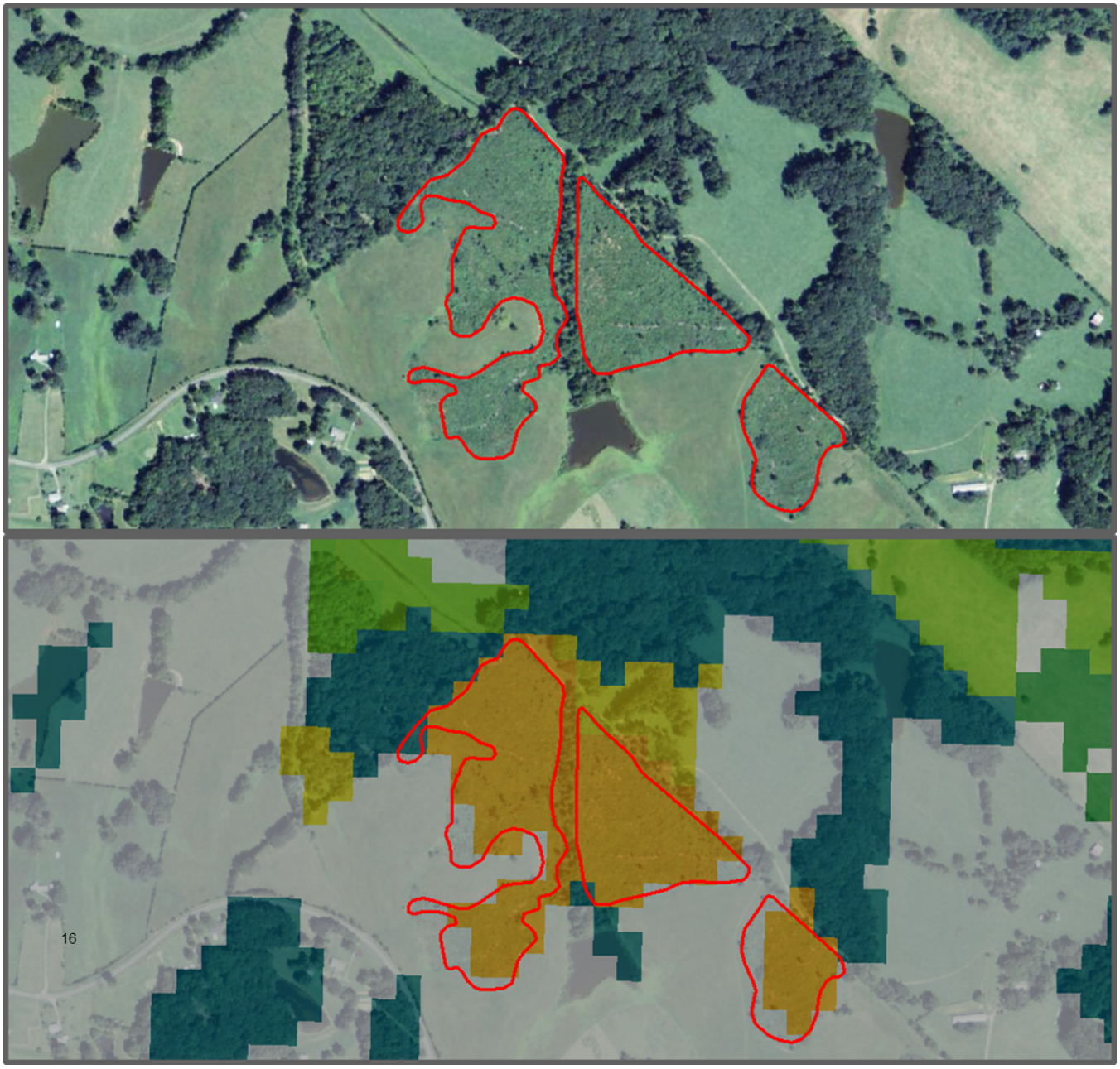

Figure 6: Depiction of ability of enhanced VCT to conform to harvest boundaries. Top: post-harvest NAIP aerial photography with photo-interpreted harvest boundaries; Bottom: Same image overlaid with semi-transparent enhanced VCT "age" raster.

racy rates. Also, algorithm parameters were not tuned to maximum accuracy. Some strategies that would not hinder automation while improving accuracy can and will be implemented. Four disturbance magnitudes and various shape and size metrics were calculated for each disturbance clump. One simple method for selecting from these variables would be to exclude one variable from each pair of highly correlated variables while keeping all of the remaining variables. The results here show that good accuracy rates can be obtained without tuning model parameters. k=9 was arbitrarily chosen for kNN, and the other algorithms required no parameter specification. Perhaps an arbitrary parameter specification, such as k=9, could be used for kNN in the future, or kNN could be replaced by another algorithm that does not require parameter specification.

Figure 7: Depiction of conformance of enhanced VCT stand-clearing disturbances to parcel boundaries.

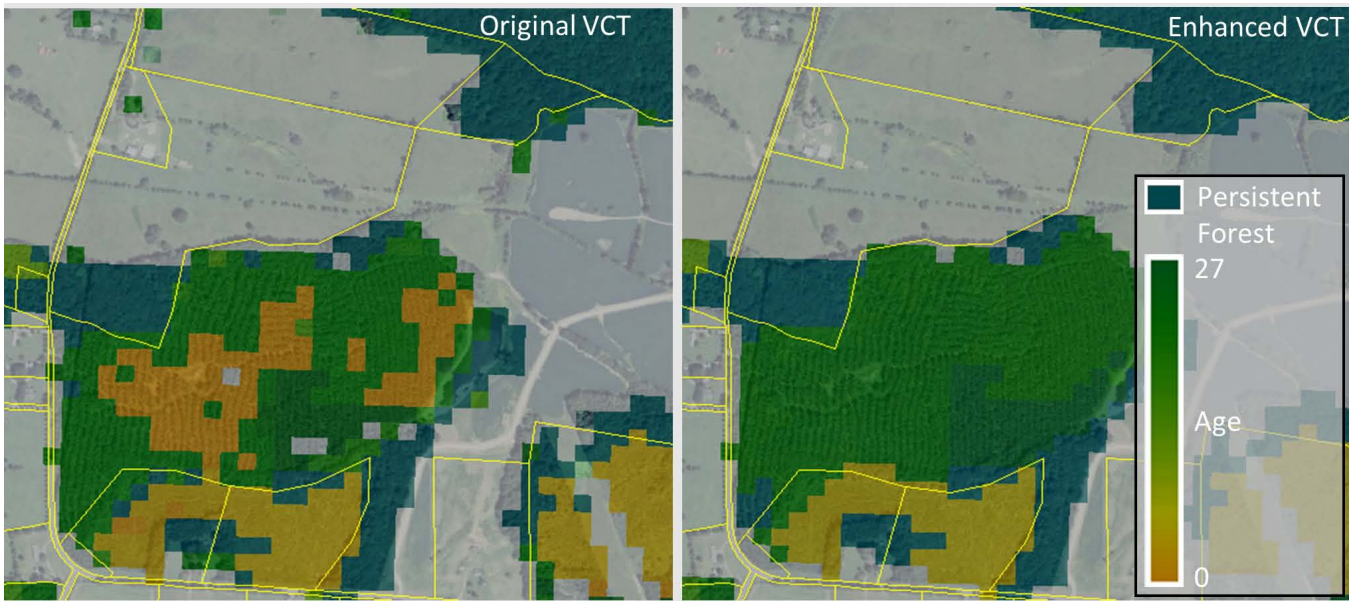

Figure 8: The enhanced VCT map reduces salt and pepper effects and more accurately depicts stand age by reclassifying disturbances such as this 2010 thinning as a partial disturbance before calculating years since most recent stand-clearing disturbance.

It is possible that the method used to sample clumps of disturbed pixels is biased towards clearcuts over partial harvests. Partial harvests often appear as multiple disjoint clumps in the same parcel rather than one solid clump like most clearcuts. Thus, it seems that there is a smaller chance that the logging deck of a partial harvest will intersect a VCT disturbance clump. The actual proportion of disturbances that are stand-clearing is likely to be somewhere between the proportion observed in the sample and 50 percent. It is unlikely to be below 50 percent because most partial harvests are followed by a clearcut, but it is impossible for a stand to be partially harvested after it has already been cleared. An improved sampling process can overcome this deficiency by intersecting harvests with parcels and then parcels with disturbances. Otherwise, correcting for the bias in this

sample indicates that if VCT detects all stand-clearing disturbances and an equal area of partial disturbances, the overall map accuracy rate could be as low as 76 percent. On the other hand, by including parcel data, it is likely that the number of disjoint clumps in a parcel could be a useful feature to help to distinguish partial harvests and improve accuracy. Partial harvests are more likely to appear as several disjoint clumps within the same parcel, while clearcuts most often appear as one contiguous clump.

In addition, the model used to classify harvest disturbances as stand-clearing or not is also used for other disturbances detected by VCT. This would indicate that the sample is biased towards harvest disturbances. The concern here seems to be minimal because harvests are the overwhelming disturbance type by area in Virginia. Other disturbances are likely to be so small that they will not affect stand age. Nonetheless, opportunities exist to reclassify VCT disturbances using auxiliary GIS data. These include conformance with parcel boundaries, shape metrics of disturbance clumps, LiDAR forest structure metrics, and variables such as distance to road and slope. Harvests are unlikely to occur in areas that are far from a road or with steep slope.

It should be noted that because a percentage of the pixels with an associated age have converted to non-forest and the number of years until regeneration begins depends on forest type and management intensity, there is room for improvement in the enhanced VCT "age" map. An analysis of the percentage of pixels that were classified as cleared according to the methods of this paper and return to forest according to VCT in one-year increments after disturbance is underway. Percentages can be calculated by forest type and/or ecoregion in order to shed light on differences in time to detect regeneration using VCT.

When considering the area estimates in Figure 4 and Figure 5 it is important to note that there are differences in how forest is defined by VCT, FIA, and NLCD (Huang et al. 2010, Bechtold and Patterson 2005, Homer et al. 2015). VCT uses a threshold cutoff of an integrated measure of number of standard deviations above the mean a pixel's brightness in the red and two shortwave infrared Landsat bands is compared to the average of a forest sample. FIA imposes area, width, stocking, and use restrictions which are not accounted for with VCT. The figures for NLCD do not include developed open space which can include vegetation planted in developed areas for recreation, erosion control, or aesthetic purposes. In addition the shrub/scrub class can include forest areas where there are young trees in an early successional stage that are less than 5 meters tall. However, this class and the woody wetlands class can also include true shrubs, of which there is very little in Virginia. It is reasonable to believe that the total forest acres estimate for the enhanced VCT exceeds FIA because there are less restrictions on what it defines as forest.

Furthermore, higher estimates for enhanced VCT over FIA are not consistent across age classes when excluding VCT classified post-disturbance non-forest. This reveals a time lag after a disturbance before VCT can detect a return to forest. It is conceivable that the reclassified enhanced VCT stand-clearing disturbance maps and the enhanced VCT age map derived from them can be intelligently combined with information from NLCD or a similar land cover map to arrive at a more accurate age map. For instance, a disproportionately high number of pixels that are classified as forest by VCT but grassland by NLCD are in the 0-5 age group. Thus, VCT forest pixels estimated to be older according to enhanced VCT and are also classified as grassland according to NLCD are less likely to be forest with the correct age estimate. Perhaps some of these pixels were incorrectly classified by VCT as forested and some are actually young forest with incorrect older age estimates.

5 Conclusion

While the enhanced VCT product is a good proxy for "age" and generally conforms to harvest boundaries, some work remains. Despite the overwhelming presence of secondary succession forest in Virginia due to clearcut harvesting practices in the south, a majority of the forest in Virginia has not been disturbed since 1984. Therefore its age cannot be precisely determined, and large clumps of undisturbed forest cannot be broken up into stand-sized pieces using reclassified VCT disturbances. These stand-sized pieces would be a good basis for modeling future change. It may be possible to use variables such as LiDAR height and structure metrics along with other remotely sensed and auxiliary GIS data to create these stand-sized units.

Additional work needs to be done to generate identities for unique harvests rather than unique clusters. Partially forested parcels may include disjoint clumps of harvested pixels even if the harvest was a clearcut. Therefore parcel data should be combined with VCT data to identify disjoint clearcut harvest clumps that are most likely part of the same clearcut harvest. This could be important for modeling efforts in which only harvests that meet a minimum area threshold are considered. For instance, small harvests may be excluded when modeling commercial wood supply. Further processing should be done to combine adjacent clumps that differ by one VCT year because a single harvest spanned consecutive VCT years.

Even without addressing these imperfections, reclassifying clumps of disturbed pixels in yearly VCT maps

as stand-clearing disturbances or not based on average disturbance magnitude, shape, and size results in a good proxy for "age" and objects that conform to harvest boundaries. The net result is an historical record of harvest boundaries that can also be used to predict when and where future harvests will occur. A decades long historical record begins to shed light on the impact of variables such as policy change, social and cultural values, and ownership demographics on harvesting practices, although their overall impact may not be known for many more decades. Estimates of biomass or timber volume across time provide valuable data for procurement foresters, landscape ecologists, climate scientists, water resource experts, and many others.

Acknowledgments

We are very grateful to Chengquan Huang for generously sharing the data that this work is based on. Special thanks to the anonymous reviewers whose dedication and insightful feedback significantly improved our manuscript. Thanks also to the Center for Natural Resources Assessment and Decision Support and its partners for their support of this project.

References

Bechtold, W.A., and P.L. Patterson. 2005. The enhanced Forest Inventory and Analysis Program–national sampling design and estimation procedures. USDA For. Serv. Gen. Tech. Rep. SRS-GTR-80. 88 p.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. Classification and regression trees. Wadsworth. 368 p.

Hijmans, R.J. 2015. raster: Geographic data analysis and modeling. R package version 2.4-18. Last accessed online on Jan. 19, 2016, at: http://CRAN.R-project.org/package=raster.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States–representing a decade of land cover change information. Photogrammetric Engineering and Remote Sensing. 81(5):345–354.

Huang, C., L.S. Davis, and J.R.G. Townshend. 2002. An assessment of support vector machines for land cover classification. International Journal of Remote Sensing. 23(4):725–749.

Huang, C., K. Song, S. Kim, J.R.G. Townsend, P. Davis, P. Masek, and S.N. Goward. 2008. Use of a dark object concept and support vector machines to automate forest cover change analysis. Remote Sensing of Environment. 112:970–985.

Huang, C., S.N. Goward, J.G. Masek, N. Thomas, Z. Zhu, and J.E. Vogelman. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sensing of Environment. 114:183–198.

Leica Geosystems. 2013. ERDAS Imagine version 2014. Atlanta, Georgia.

Meyer, D., E. Dimitriadou, K. Hornik, A. Weingessel, and F. Leisch. 2015. Misc functions of the Department of Statistics, Probability Theory Group (formerly: E1071), TU Wien. R package version 1.6-7. Last accessed online on Jan. 19, 2016, at: http://CRAN.R-project.org/package=e1071.

Miles, P.D. 2016. Forest Inventory EVALIDator web-application version 1.6.0.03. USDA For. Serv., Northern Research Station, St. Paul, MN. Last accessed online on Jan. 26, 2016, at: http://apps.fs.fed.us/Evalidator/evalidator.jsp.

Pal, M., and P. Mather. 2005. Support vector machines for classification in remote sensing. International Journal of Remote Sensing. 26:1007–1011.

R Core Team. 2015. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Last accessed online on Jan. 19, 2016, at: https://www.R-project.org/.

Therneau, T., B. Atkinson, and B. Ripley. 2015. rpart: Recursive partitioning and regression trees. R package version 4.1-10. Last accessed online on Jan. 19, 2016, at: http://CRAN.R-project.org/package=rpart.

VanDerWal, J., L. Falconi, S. Januchowski, L. Shoo and C. Storlie. 2014. SDMTools: Species Distribution Modelling Tools: Tools for processing data associated with species distribution modelling exercises. R package version 1.1-221. Last accessed online on Jan. 19, 2016, at: http://CRAN.R-project.org/package=SDMTools.

Venables, W.N., and B.D. Ripley. 2002. Modern applied statistics with S, 4th ed. Springer, New York. 498 p.

Zhao, F., C.Q. Huang, and Z.L. Zhu. 2015. Use of Vegetation Change Tracker and support vector machine to map disturbance types in Greater Yellowstone ecosystems in a 1984-2010 Landsat time series. IEEE Geoscience and Remote Sensing Letters. 12(8):1650–1654.