

ESTIMATION OF FOREST DISTURBANCE INTENSITY FROM LANDSAT DATA IN NORTH CAROLINA AND SOUTH CAROLINA

Xin Tao, Chengquan Huang, Feng Zhao

Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA

ABSTRACT

Disturbance and regrowth are vital processes in determining the roles of forest ecosystem in carbon and biogeochemical cycles. The vegetation change tracker (VCT) algorithm derives the spectral disturbance magnitude based on the time series observations. While these spectral disturbance magnitudes are indicative of physical changes in tree cover or biomass, their quantitative relationships have yet to be established. This study focuses on estimating disturbance intensity measured by percent basal area removal using spectral data. Ground measurements from Forest Inventory Analysis (FIA) data are used for training and validation of the model. The overall R^2 between predicted disturbance intensity and the reference data is around 0.66, and the prediction uncertainty is 14% in North Carolina. A time series analysis shows that the stand clearing area remains relatively stable around 1143 km² in North and South Carolinas. The area of partial disturbance fluctuates heavily, and the average value is around 3287 km².

Index Terms— disturbance intensity, spectral disturbance magnitude, Landsat, model, validation

1. INTRODUCTION

Forests play an important role in carbon and biogeochemical cycles, and provide socio-ecological services to the human. Forest loss could lead to biodiversity loss, increased greenhouse emission, and soil deterioration [1]. Although considerable attention has been given to forestry because of their potential benefits and negative impacts, the forest system has not been well characterized in terms of its disturbance and regrowth history. A spatially and temporally explicit information on forest dynamics would bring a high level of transparency to the land use both economically and environmentally.

Early studies utilizing Landsat for disturbance mapping typically used spectral information from one or two images. A more comprehensive mapping approach is needed, utilizing as many points in time as possible and characterizing the disturbance magnitude and duration [2]. The vegetation change tracker (VCT) algorithm focuses on deviation using multiple instances of a presumed stable

condition to define when a persistent change has occurred to map the location, timing, and severity of forest disturbance events [3, 4]. The VCT algorithm has been evaluated across the United States with reported overall accuracy ranging between 77% and 86% [5].

The spectral disturbance magnitude data have been used to distinguish between disturbance types. For example, the disturbance magnitude in wetness, greenness, and brightness are used to distinguish the insect disturbance by defoliator and bark beetle [2]. Although spectral disturbance magnitude could be used as an indicator of physically based disturbance severity, it is hard to infer actual disturbance intensity from it. This study focuses on the estimation of disturbance intensity as measured by percent basal area removal from spectral disturbance magnitude data.

The objective of this study is to test how well the disturbance intensity could be modeled from spectral disturbance magnitude data, and what are the possible causes of the uncertainties. Section 2 introduces the data for disturbance intensity estimation and validation. Section 3 describes the methods for calculating reference disturbance intensity from FIA data and predicting disturbance intensity from spectral data. The predicted disturbance intensity are examined and validated with reference data in Section 4. The model is applied in North Carolina and South Carolina and the regional patterns and distributions of disturbance intensity are described in Section 5. The findings are concluded in Sections 6.

2. DATA

The data used in this study include spectral disturbance magnitude generated from the VCT algorithm and the validation data collected from FIA program. The FIA has over 125000 forest plots distributed across the United States with a nominal spacing of 5 km [6]. The data in North and South Carolinas are used in this study. A series of variables are measured for each plot, including tree diameter, height, status, etc. A unique sequence number is used to identify a site tree record. If a tree is remeasured from a previous annual inventory, there is a previous tree sequence number to link the tree to the previous inventory's tree record. With these information, we could build tree growth or disturbance

history for a plot by summarizing the information of all the trees in the plot.

The VCT algorithm generates two kinds of spectral disturbance magnitude data: delta variables and normalized before/after ratio (Table 1). The delta variables are the variables difference between two periods. The delta variables include the difference in forest index (FI), normalized difference vegetation index (NDVI), normalized burn ratio (NBR), udist calculated using band 4 (udB4) or using band 5 (udB5). The normalized before/after ratio is the variable difference between two periods normalized by the value in a period. The variables for calculating normalized before/after ratio include FI, NDVI, NBR, udB5, and normalized difference moisture index (NDMI).

3. METHOD

The flowchart for calculating reference disturbance intensity from FIA data and predicting from spectral disturbance magnitude data is shown in Fig. 1. The basal area are calculated for each tree and then summarized at the plot level. Denote the total basal area for live trees at time 1 as b_1 , the total basal area for the trees live at both time 1 and time 2 as b_{12} . The removed basal area is calculated from the difference between the live basal area at time 1 (b_1) and the basal area live at both time 2 and time 1 (b_{12}). We assume trees grow at the same rate, so the ratio between the removed basal area and initial live basal area at time 1 is the disturbance intensity from FIA data.

The reference disturbance intensity data are combined with the spectral disturbance magnitude from VCT to train a random forest model. Random forest is built as an ensemble of regression trees [7]. In a random forest model, a small random sample of explanatory variables is selected at the root node and the best split is made using that limited set of variables. The processes of variable selection and split are repeated a large number of times. The final prediction is a plurality vote or average from prediction of all trees in the collection. A random forest model is used to predict disturbance intensity from delta variables, normalized before/after ratio, or a combination of them. The performances are evaluated using root mean square error (RMSE) and Person's correlation between observed and predicted disturbance intensity. The experiment design is shown in Table 1.

Table 1 Experiment design for estimating disturbance intensity from spectral disturbance magnitude data

Case 1	Case 2	Case 3
Delta variables: $FI_2 - FI_1$, $NDVI_2 - NDVI_1$, $NBR_2 - NBR_1$, $udB4_2 - udB4_1$, $udB5_2 - udB5_1$,	Normalized before/after ratio: $1 - FI_1 / FI_2$, $1 - NDVI_2 / NDVI_1$, $1 - NBR_2 / NBR_1$, $1 - udB5_1 / udB5_2$, $1 - NDMI_2 / NDMI_1$	Delta variables + Normalized before/after ratio

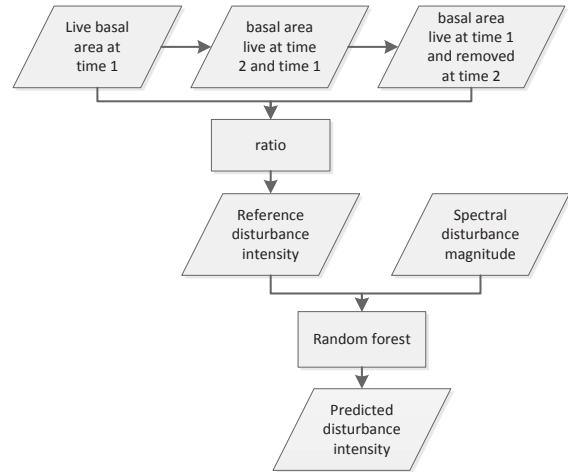


Fig. 1 Flowchart for calculating reference disturbance intensity from FIA data and predicting disturbance intensity from spectral disturbance magnitude data.

4. VALIDATION

The predicted disturbance intensity is examined with reference satellite images before and after disturbance, and the values are validated with reference FIA data.

4.1. Visual examination

A regional disturbance intensity in path 16 row 35 in year 2006 is shown in Fig. 2. The left is the disturbance intensity. There are higher disturbance intensity in the center of the disturbed patch, with a decreasing trend towards the boundary of the patch. The disturbance intensity map is compared with the closest high resolution Google earth image before and after the disturbance year 2006. A clear cut of forests is in the top left of the region where disturbance intensity is above 80%. The partial cut or selective logging is in the center of the region. A visual examination of the results prove the effectiveness of the disturbance intensity model mapping.

4.2. Validation with reference FIA data

In control experiments, we used only delta variables or normalized before/after ratio variables for prediction of disturbance intensity. The results are shown in the left two scatterplots of Fig. 3. The validation of the predicted disturbance intensity using both sets of variables are shown in the right scatterplot of Fig. 3. The results using either set of variables are not as good as the one using a combination of the two sets of variables regarding R^2 . The result indicate that the two sets of variables are both important for disturbance intensity prediction from spectral data.

There are good agreements between predicted and observed disturbance intensity between 10% and 90%. There are some overestimations of small values and

underestimations of large values in predicted disturbance intensity when compared with observed ones. The underestimation of high disturbance intensity, especially the underestimation of some stand clearing disturbance as partial disturbance, is caused by the plot located near the boundary of disturbed patches (top panel of Fig. 4). Because of the scale difference, the validation results at heterogeneous sites are expected to have a lower accuracy than those at homogenous sites [8, 9]. The overestimation of the low disturbance intensity below 10% could be caused by the disturbance occurred after the field measurements so that the reference disturbance intensity is small while stand clearing is mapped (bottom panel of Fig. 4). Overall, the R^2 between the predicted disturbance intensity and the reference data is around 0.66.



Fig. 2 Visual examination of the disturbance intensity with Google earth images. Left is disturbance intensity in 2006, middle is Google earth image on June 13, 2005, right is image on June 17, 2008.

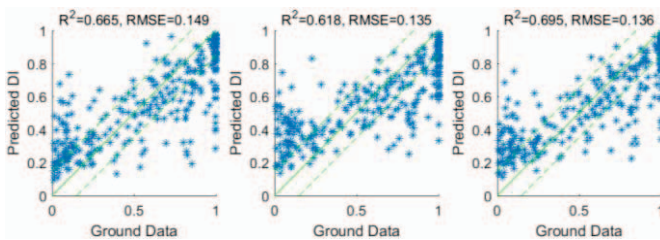


Fig. 3 Scatterplots between the modeled disturbance intensity and the FIA data in North Carolina using training data in Case 1 (left), Case 2 (middle), or a combination in Case 3 (right). The middle green line is $y = x$. The two other green lines are $y = x \pm 0.15$, respectively.

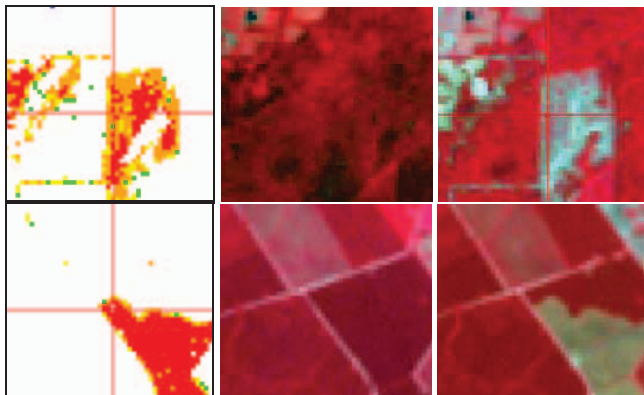


Fig. 4 The mapped disturbance intensity in 2005 (top left), the reflectance on August 9, 2004 (top middle), and July 27, 2005

(top right). Medium intensity disturbance at 59.3% is mapped but high intensity disturbance was observed by FIA at 96.2% during 2003-2013. The mapped disturbance intensity in 2011 (bottom left), the reflectance on June 24, 2010 (bottom middle), and May 26, 2011 (bottom right). High intensity disturbance at 75.1% is mapped but only low intensity disturbance at 10.3% was observed by FIA during 2005-May 24, 2011 previous disturbance. The reflectance images are RGB: 432.

5. PATTERNS AND DISTRIBUTIONS OF DISTURBANCE INTENSITY

The disturbance intensity are validated and examined at the plot scale, and the method is further applied at the regional scale in North Carolina and South Carolina. The disturbance intensity map in North and South Carolinas in 2007 is shown in Fig. 5. The high disturbance intensity region is mainly located in the east. The low disturbance intensity region is located in the center and mid-west of North Carolina and in the center of South Carolina.

The histograms of the mapped disturbance intensity in North Carolina and South Carolina in 2007 are shown in Fig. 6. There are two peaks of the disturbance intensity, around 30% and 90% in North Carolina and around 50% and 80% in South Carolina, which correspond to the most often occurred partial and stand clearing disturbance intensity. South Carolina has higher disturbance intensity than North Carolina.

The areas of stand clearing and partial disturbance in North and South Carolinas during 1986–2015 are shown in Fig. 7. The stand clearing area remains relatively stable in North and South Carolinas over time. The area of partial disturbance fluctuates heavily, with peaks in 1986, 1990, 1997, and 2007. The high value of partial disturbance area in North and South Carolinas could reach over 6000 km², while the low value is only 1919 km² in 2013. Overall, the average area of stand clearing is around 1143 km² and that of partial disturbance is around 3287 km² for North and South Carolinas. Due to the dominance of partial disturbance in the two regions, the total area of partial and stand clearing disturbance has similar pattern as that of the partial disturbance.

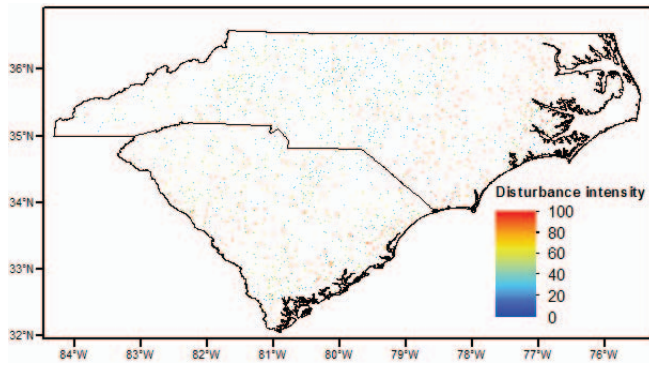


Fig. 5 Disturbance intensity map in North Carolina and South Carolina in 2007 from Landsat data (not FIA plot locations).

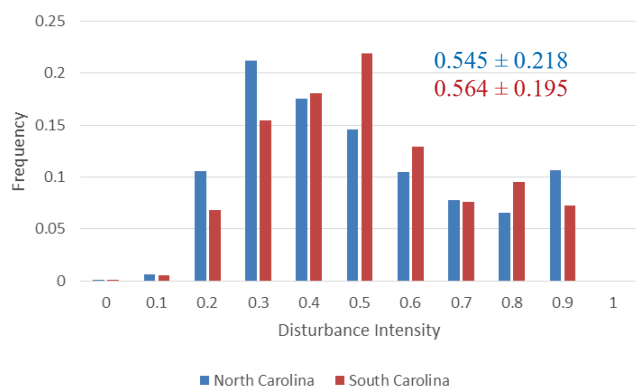


Fig. 6 Histograms of disturbance intensity in North Carolina and South Carolina in 2007. The numbers are the regional mean and standard deviation.

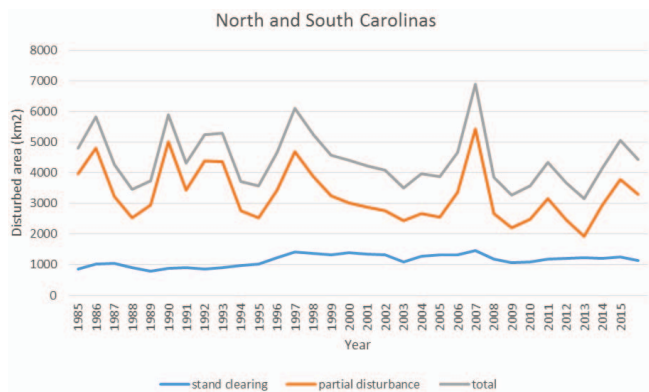


Fig. 7 Disturbed area in North and South Carolinas over time.

6. CONCLUSIONS

The disturbance intensity could be predicted from spectral disturbance magnitude with an accuracy of 14% in North Carolina. Overall, the R^2 between the predicted disturbance intensity and the reference data is around 0.66. The stand clearing area remains relatively stable around 1143 km² in

North and South Carolinas than the partial disturbance area. The area of partial disturbance fluctuates heavily, and the average value is around 3287 km² for North and South Carolinas. Due to the dominance of partial disturbance, the total disturbance area has similar pattern as partial disturbance area in the two regions.

7. ACKNOWLEDGEMENTS

This study contributes to the North American Carbon Program, with grant support from NASA's Land Cover and Land Use Change, Terrestrial Ecology, Carbon Cycle Science, and Applied Sciences Programs. Additional support was provided by the U.S. Geological Survey and USDA Forest Service. We thank Elizabeth Burrill for assisting with access to the FIA plot data, which was made possible through an MOU with the FIA program.

8. REFERENCES

- [1] J. W. Petranka, M. E. Eldridge, and K. E. Haley, "Effects of timber harvesting on southern Appalachian Salamanders," *Conservation Biology*, vol. 7, pp. 363-377, Jun 1993.
- [2] C. Senf, D. Pflugmacher, M. A. Wulder, and P. Hostert, "Characterizing spectral-temporal patterns of defoliator and bark beetle disturbances using Landsat time series," *Remote Sensing of Environment*, vol. 170, pp. 166-177, Dec 1 2015.
- [3] C. Q. Huang, S. N. Goward, J. G. Masek, N. Thomas, Z. L. Zhu, and J. E. Vogelmann, "An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks," *Remote Sensing of Environment*, vol. 114, pp. 183-198, Jan 2010.
- [4] C. Q. Huang, S. N. Goward, K. Schleeweis, N. Thomas, J. G. Masek, and Z. L. Zhu, "Dynamics of national forests assessed using the Landsat record: Case studies in eastern United States," *Remote Sensing of Environment*, vol. 113, pp. 1430-1442, Jul 2009.
- [5] N. E. Thomas, C. Q. Huang, S. N. Goward, S. Powell, K. Schleeweis, and A. Hinds, "Validation of North American Forest Disturbance dynamics derived from Landsat time series stacks," *Remote Sensing of Environment*, vol. 115, pp. 19-32, Jan 2011.
- [6] W. B. Smith, "Forest inventory and analysis: a national inventory and monitoring program," *Environmental Pollution*, vol. 116, pp. S233-S242, 2002.
- [7] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5-32, Oct 2001.
- [8] X. Tao, S. L. Liang, T. He, and H. R. Jin, "Estimation of fraction of absorbed photosynthetically active radiation from multiple satellite data: Model development and validation," *Remote Sensing of Environment*, vol. 184, pp. 539-557, Oct 2016.
- [9] X. Tao, S. L. Liang, and D. D. Wang, "Assessment of five global satellite products of fraction of absorbed photosynthetically active radiation: Intercomparison and direct validation against ground-based data," *Remote Sensing of Environment*, vol. 163, pp. 270-285, Jun 2015.