HARMONIC ANALYSIS OF DENSE TIME SERIES OF LANDSAT IMAGERY FOR MODELING CHANGE IN FOREST CONDITIONS

Barry Tyler Wilson[1]

Abstract—This study examined the utility of dense time series of Landsat imagery for small area estimation and mapping of change in forest conditions over time. The study area was a region in north central Wisconsin for which Landsat 7 ETM+ imagery and field measurements from the Forest Inventory and Analysis program are available for the decade of 2003 to 2012. For the periods 2003-2007 and 2008-2012, the monthly imagery was processed using harmonic analysis in order to capture seasonal trends in reflectance across spectral bands. A nonparametric modeling approach was used with predictor variables and field measurements at two points in time to predict change in live tree basal area. Predictions for individual plots poorly matched observations of change, however the resultant maps of change compared favorably to a purposive sample of locations of high predicted relative change, based on multi-date, high-resolution aerial photography. This suggests the need for a larger sample of plots or further tuning of the model.

INTRODUCTION

While the national forest inventory (NFI) conducted by the USDA Forest Service, Forest Inventory and Analysis (FIA) program is intended to address strategic-level questions about the forest resources across large geographic areas under a design-based mode of inference (Bechtold and Patterson 2005), there is increasing interest in using this information for reporting on and monitoring change in forest conditions over time for smaller areas within the population (McRoberts et al. 2010). By using auxiliary variables from data collected for all population units, such as those obtained from remote sensors, and shifting to a model-based mode of inference, dramatic gains in the precision of estimates can be achieved, though possibly at the expense of the unbiasedness assumption for the estimators (Gregoire 1998). Here we evaluate the use of dense time series of satellite imagery for predicting change in forest conditions by an examination of the linear regression of observed versus predicted values and a comparison of a purposive sample of areas on the map of predicted change to multi-date aerial photos.

[1] Research Forester (BTW), Northern Research Station, USDA Forest Service, 1992 Folwell Avenue, St. Paul, MN 55108. To contact BTW, call (651) 649-5189 or e-mail at barrywilson@fs.fed.us.

STUDY AREA

The study area was approximately 5.56 million acres in north central Wisconsin, corresponding to Web-enabled Landsat Data (WELD) tile H20V05 and described in the next section (Roy et al. 2010). The landscape includes a variety of land covers and uses, from agriculture in the south and southwest, deciduous and evergreen forest in the north and northeast, developed land around the city of Wausau in the central portion, as well as scattered wetlands, lakes, and rivers. The study area experienced a severe weather-related disturbance event in June of 2007, when a tornado traced a ½ mile-wide swath through part of Menominee County.

METHODS

The auxiliary data used in the study were dense Landsat time series images from the WELD project. WELD imagery are composites of the highest fidelity data, determined on a pixel-by-pixel basis, from all Landsat 7 ETM+ scenes collected over a compositing period. These composite images have been processed for the contiguous United States and Alaska over the decade of 2003-2012. The composite scenes have been orthorectified, transformed to top-of-atmosphere reflectance, and mosaicked into 5,000-by-5,000 pixel tiles.

NFI data from 1,446 plots were used in the study, including both forested and nonforested plots. Each of these plots was measured and then remeasured during the decade of 2003-2012, spanning two 5-year FIA measurement cycles in Wisconsin, a state that has a sampling intensity twice the base level of roughly 1 plot per 6,000 acres. Remeasurement of sample plots provides information about change in forest conditions in the population over the remeasurement period.

Since satellite-based sensors, such as Landsat 7's ETM+ instrument, detect reflectance from the Earth's surface, these data are expected to be closely correlated with land cover. Of the attributes measured on NFI plots, live tree basal area is considered to be one of those most correlated with tree canopy cover (Jennings et al. 1999). Therefore, the live tree basal area condition-level attribute was used as the attribute of interest. These summary values were calculated on all plots for both 5-year periods: measurements during 2003-2007 and remeasurements during 2008-2012. Nonforested conditions were assigned a value of 0. The plot-level data were calculated by multiplying each condition's plot proportion by its corresponding basal area value, then summing across all conditions. Differences in live tree basal area were computed for each plot by subtracting the summary value at the earlier time period from the summary value at the later time period.

The WELD monthly composites for the entire decade of 2003-2012 for tile H20V05 were used for the study. For each monthly composite, the reflectance values from ETM+ were transformed to the first three Tasseled Cap (TC) components: brightness, greenness, and wetness (Huang et al. 2002). The monthly TC components were then compiled into individual stacks, by TC component and time period, resulting in six separate stacks. To account for the seasonality of the TC components, harmonic analysis was conducted separately on each stack fitting a Fourier series with two harmonics individually to each pixel in the stack via least squares regression (Sellers et al. 1996). To compensate for missing data in the time series due to clouds or the scan line corrector (SLC) failure, weighted regression was used, with the weight for each observation calculated as the inverse of the total number of observations over each respective 5-year period for the month of the given observation.

The Fourier coefficients calculated for each of the TC components for each of the time periods were bundled together to form a 30-layer stack of auxiliary variables for each pixel in the tile. To reduce dimensionality, principal components analysis (PCA) was conducted on the stack. Only the first eight principal components (PCs) were kept, accounting for roughly 93 percent of the original variance. Finally, because of the spatial mismatch in the size of the FIA plot relative to the size of the ETM+ pixel, a 3-by-3 pixel moving window was used to compute the focal mean for each of the PCs.

The 8-layer stack of focal means was used as the feature space with the k-nearest neighbors (kNN) estimator. The kNN estimator has been widely used with NFI and remote sensing data (Eskelson et al. 2009). It provides an estimate for each unsampled unit in the population as a weighted average of the observed response variable for the k-nearest sample units in the feature space.

Because the focus of this study was the utility of dense time series of satellite imagery for modeling change in forest conditions, the only tuning of the kNN estimator was to objectively determine the optimal value of *k* to minimize the root mean square error (RMSE) using the 1,446 observations of change. For this optimization a leave-one-out procedure was used, whereby the k-nearest neighbors for a given plot were found by holding the given plot out of the list of possible neighbors. Using this criterion, the optimal value of *k* was determined to be 9, resulting in a minimum RMSE of approximately 19.2 square feet per acre.

RESULTS

Scatterplots of observed versus predicted live tree basal area for all 1,446 remeasured plots show strong agreement for both time periods. Figure 1 depicts the results for period 2, with similar results for period 1 (not shown). However, a scatterplot of observed versus predicted change indicates poor agreement, as shown in Figure 2 (Piñeiro et al. 2008).

After masking out areas that were nonforested in the first time period (less than 10 square feet per acre of live tree basal area), comparisons of the map of relative change in live tree basal area (i.e., relative to the basal area in the first time period), shown in Figure 3, to multi-date aerial photography collected over the same period indicate strong agreement for a purposive sample of areas of high predicted relative change, an example of which is shown in Figure 4. The multi-date aerial photography was collected from online property tax information systems for two counties in the study area: Langlade County (2003, 2008, and 2010) and Price County (2005, 2006, 2008, 2010, and 2011).

DISCUSSION

The uncertainty in the individual estimates of live tree basal area for each period results in even greater uncertainty in their difference. Furthermore, the results shown in the scatterplots, combined with the strong agreement between the predicted map and aerial photography for the purposive sample, suggest that the issue may be related to the size and frequency of areas of change. Small relative increases in live tree basal are commonplace across the study area, corresponding to forest growth. Large relative decreases are much less common and are highly localized in extent, corresponding to forest disturbance events such as harvests, wildfires, and blow-down due to weather events.

This suggests that the sample of 1,446 remeasured plots may not be large enough to adequately characterize such rare disturbance events. One possible solution is to use a larger sample of plots, perhaps by including sample units from neighboring WELD tiles in the kNN estimation. Another option would be to adjust the value of *k* according to the unsampled target unit's location in feature space, with units in the interior having a larger value of *k* than those closer to the convex hull enclosing all sample units.

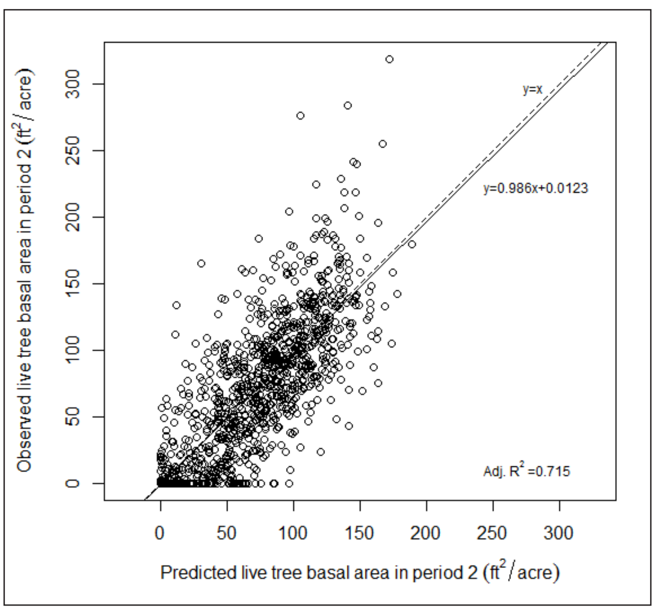

Figure 1—Scatterplot of observed vs. predicted live tree basal area for 1,446 plots in the second time period. The solid line is the linear regression of the data. The dashed line is the line y=x.

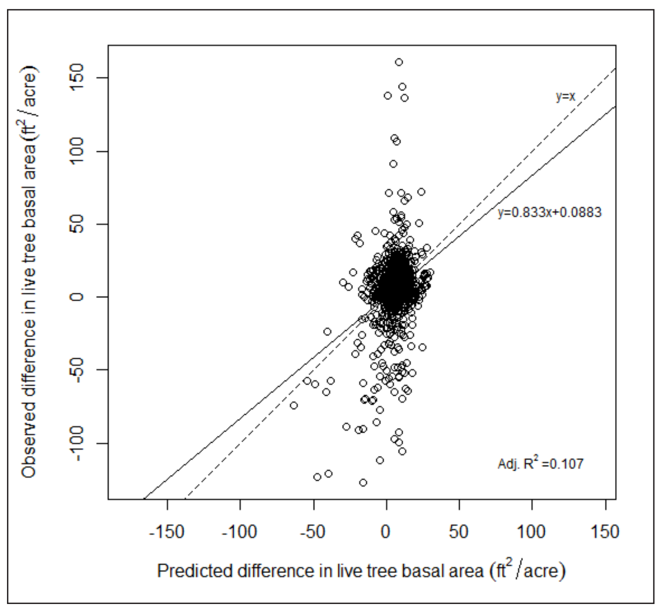

Figure 2—Scatterplot of observed vs. predicted difference in live tree basal area between time periods, for 1,446 remeasured plots. The solid line is the linear regression of the data. The dashed line is the line y=x.

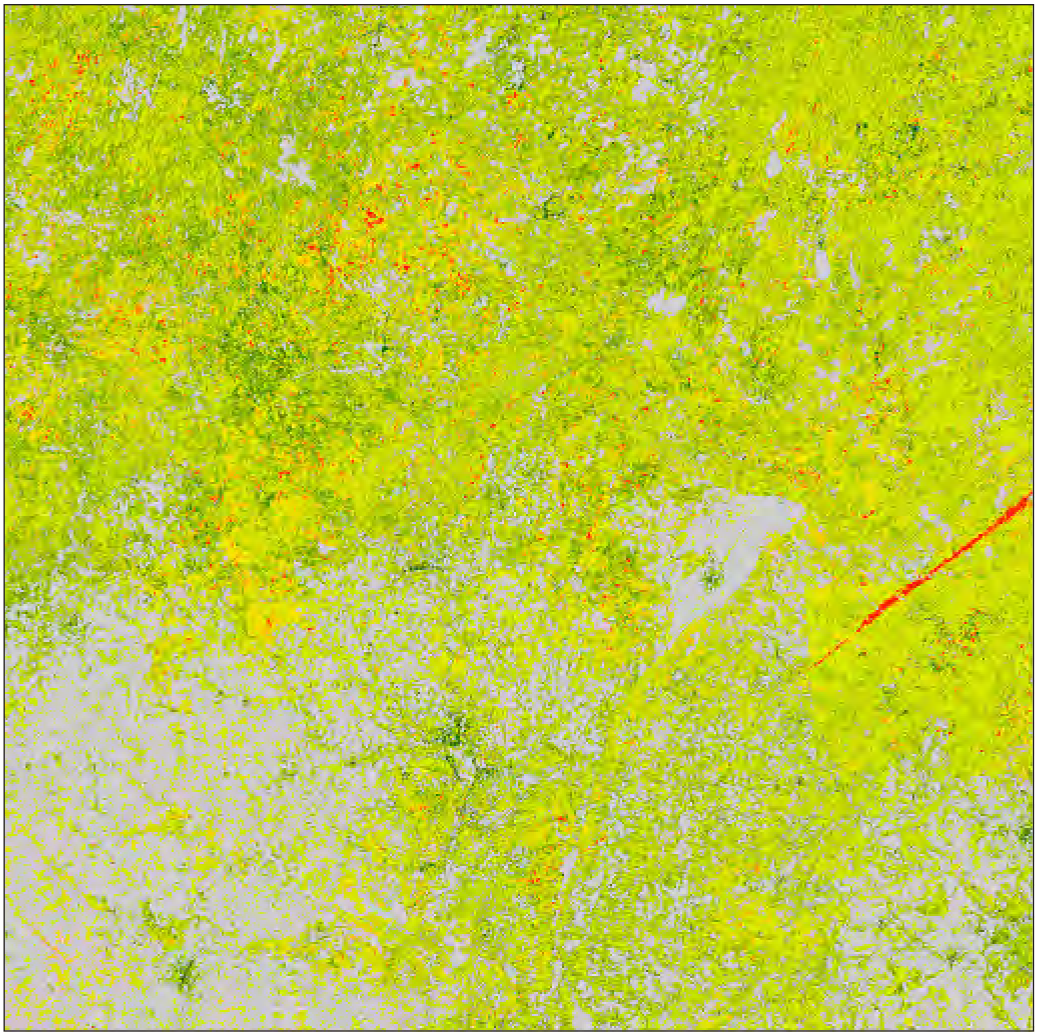

Figure 3—Map of the study area of predicted relative difference in live tree basal area, in percent of the predicted total for the first time period. Nonforest land (less than 10 square feet per acre of live tree basal area in the first time period) is gray, relative gain is green, relative loss is red, and no change is yellow. Darker shades of green and red indicate larger relative gains or losses than lighter tints.

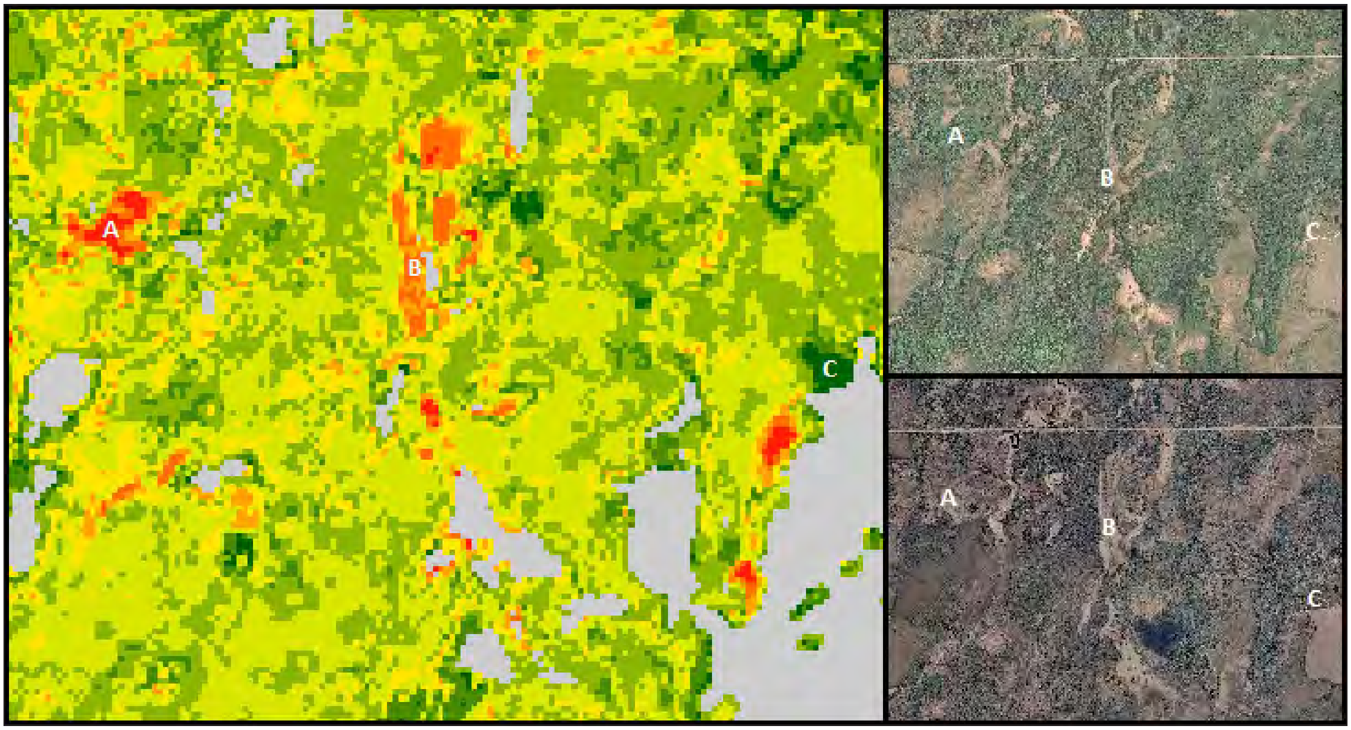

Figure 4—Detail from Figure 3 (left) and aerial photos from 2005 (top right) and 2010 (bottom right) of an area in Price County with patches of large relative gains and losses. Patch A was harvested earlier in the second time period than patch B. Patch C shows regrowth of an area harvested prior to the first time period.

LITERATURE CITED

Bechtold, W.A.; Patterson, P.L., eds. 2005. The enhanced forest inventory and analysis program - national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.

Eskelson, B.N.; Temesgen, H.; Lemay, V.; Barrett, T.M.; Crookston, N.L.; Hudak, A.T. 2009. The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. Scandinavian Journal of Forest Research. 24(3): 235-246.

Gregoire, T. G. 1998. Design-based and model-based inference in survey sampling: appreciating the difference. Canadian Journal of Forest Research. 28(10): 1429-1447.

Huang, C.; Wylie, B.; Yang, L.; Homer, C.; Zylstra, G. 2002. Derivation of a tasselled cap transformation based on Landsat 7 at-satellite reflectance. International Journal of Remote Sensing. 23(8): 1741-1748.

Jennings, S.B.; Brown, N.D.; Sheil, D. 1999. Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. Forestry. 72(1): 59-74.

McRoberts, R.E.; Tomppo, E.O.; Næsset, E. 2010. Advances and emerging issues in national forest inventories. Scandinavian Journal of Forest Research. 25(4): 368-381.

Piñeiro, G.; Perelman, S.; Guerschman, J.P.; Paruelo, J.M. 2008. How to evaluate models: observed vs. predicted or predicted vs. observed? Ecological Modelling. 216(3): 316-322.

Roy, D.P.; Ju, J.; Kline, K.; Scaramuzza, P.L.; Kovalskyy, V.; Hansen, M.C.; Loveland, T.R.; Vermote, E.F.; Zhang, C. 2010. Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. Remote Sensing of Environment. 114: 35-49.

Sellers, P. J.; Tucker, C. J.; Collatz, G. J.; Los, S. O.; Justice, C. O.; Dazlich, D. A.; Randall, D. A. 1996. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: The generation of global fields of terrestrial biophysical parameters from satellite data. Journal of Climate. 9(4): 706-737.