Annual Progress Report: April 2012 – March 2013

Land-Use And Land-Cover Changes In Temperate Forests Of European Russia: The Past, The Current, And The Future

Principal Investigator
Dr. Mutlu Ozdogan,
Center for Sustainability and the Global Environment (SAGE) &
Department of Forest and Wildlife Ecology
University of Wisconsin-Madison
1630 Linden Drive
Madison, WI 53706
ozdogan@wisc.edu
http://forestandwildlifeecology.wisc.edu/facstaff/ozdogan.html

Co-Investigators
Dr. Volker C. Radeloff
Department of Forest Ecology and Management
University of Wisconsin-Madison
1630 Linden Drive
Madison, WI 53706
radeloff@wisc.edu
http://forestandwildlifeecology.wisc.edu/facstaff/radeloff.html

Dr. David J. Lewis
Department of Economics
University of Puget Sound
1500 North Warner Street, CMB 1057,
Tacoma, WA 98416
djlewis@pugetsound.edu
http://www2.ups.edu/faculty/djlewis/

Collaborators
Dr. Dmitry E. Aksenov
Vice-Director, The Transparent World, non-profit partnership, Moscow, Russian Federation
picea2k@gmail.com

Dr. Alexander A. Maslov
Lab. of Forest Geobotany & Forest Soil Science, Institute of Forest Science, Russian Academy of Sciences, Moscow, Russian Federation
am@pyrola.msk.ru

Abstract

Our objective is to better understand the past, current, and future land-use and land-cover change (LCLUC) in the temperate forest zone of European Russia by integrating remotely sensed observations with social, economic, and environmental data within a well calibrated econometric model. The rationale for focusing on European Russia's temperate forests (including the temperate deciduous, coniferous and mixed forest zones in Russia, Belorussia, the Baltics, and the northern Ukraine) is twofold. First, LCLUC in this region is rapid. The breakdown of the Soviet Union and the subsequent socioeconomic changes weakened forest management and increased forest harvesting, as well as widespread agricultural abandonment, urban sprawl, and increasing forest fires. Unlike Russia's boreal forest, its temperate forests have received little attention by LCLUC science. This is unfortunate as temperate forests have much higher productivity and are under greater human pressure. Second, the Russian forest sector is changing rapidly, with increasing investments from foreign timber companies and changes in forest laws that shift control from the state to the regions. The Baltics experienced even more drastic changes when many forests reverted from public to private land ownership. What is unknown is how these changes in the forest sector have and will influence forest patterns and timber resource availability in our study area. We propose to fill this gap with a multi-part study including: i) Landsat TM/ETM+ based quantification of changes in forest cover during the transition period from the Soviet Union to today [1985-'90-'95-'00-'05]; ii) development of an econometric land-use/land-cover change model to relate past and current changes in forest cover to socioeconomic and environmental conditions; and iii) projections of amounts and patterns of changes in forest cover up to 2050 using the econometric model. The proposed research will also evaluate MDGLS Landsat and similar medium resolution datasets for temperate forest change monitoring. The principal contribution from the proposed investigation will be a comprehensive understanding of temperate forest dynamics in the European part of the former Soviet Union from 1985 to 2005, their socioeconomic causes, and projected future changes. As such, our proposal pertains mainly to the "Projections" component of the NRA. This research will provide an important regional assessment, and improve our understanding of coupled human-natural systems within three (GOFC-GOLD, IGBP/IHDP-GLP, and NEESPI) major international programs supported by the LCLUC Program. The project also pertains strongly to "management and protection of terrestrial ecosystems", "monitoring and conserving biodiversity", "management of energy resources", and to "supporting sustainable agriculture" of the GEOSS societal needs while contributing strongly to LCLUC Goals and Key Science Questions.

Keywords
Remote Sensing, change detection, forest change, land dynamics, Russia, economics, local, regional

Social science component: 40%

Project accomplishments during this performance period (April 2012 – March 2013)

1. Separating windfall forest disturbance from forest harvest: Remote sensing can provide accurate and timely information regarding forest disturbance in many ecoregions and at scales ranging from local to global and at many different temporal resolutions. Data from Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) instruments have been used for many of these studies because of (1) the favorable combination of spatial, spectral and temporal resolution, (2) the free availability of the and, (3) the long-term data record, which continues now thanks to the Landsat Data Continuity Mission. Despite these strengths, most forest disturbance mapping studies that utilize Landsat data, the derived change products only identify areas of 'forest disturbance', but do not discriminate among different types of disturbances. This has already been identified as a gap in remote sensing based forest disturbance studies. The lack of attribution to the type of disturbance often makes it difficult to interpret forest disturbance maps, especially when these data are used as inputs to carbon budget assessments or econometric analyses. For example, in our investigation of the drivers of forest change in European Russia, we were forced to equate forest disturbance with harvesting and as a result, natural disturbance was erroneously included in harvest estimates, which may have led to overestimation of harvested areas and dampened the significance of actual drivers of forest harvest.

To remedy this lack of attribution in forest disturbance mapping, we developed a windfall disturbance mapping algorithm applied to Landsat data that exploits the success of the disturbance index for detecting wind-related forest damage in the temperate zone of European Russia (Landsat Path/Row 177/019, Figure 1) where temperate coniferous, broadleaf, and mixed forests dominate the landscape with Norway spruce (*Picea abies*) and Scots pine (*Pinus sylvestris*) being the most abundant coniferous species. The study site experiences frequent anthropogenic disturbances in the form of commercial harvests but also have history of at least one (and possible more) strong windfall event associated with storms. The specific windfall events we focused on occurred in October 2009 and July 2010 and were studied in detail by the Russian Forest Health Center.

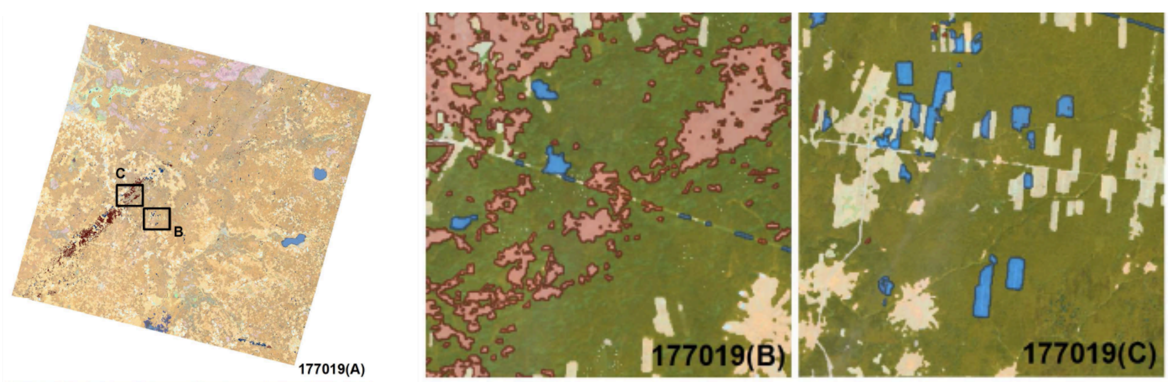

Figure 1. Test location of the windfall classification method. Study site is located in the temperate zone of European Russia. The blue polygons in B and C depict clear-cut harvests while the red polygons indicate areas affected by windfall activity.

To separate windfall-related disturbance from commercial forest harvest, we analyzed Landsat data from the year before and the year after each windfall event. Landsat imagery were pre-processed by converting digital numbers in to surface reflectance using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm. Cloud-free images were not available for both time points so we selected images with the least amounts of clouds and gap-filled them using other Landsat scenes from the same growing season. The cloud gap-filling was accomplished by first masking clouds and cloud shadows using the FMask tool. Then we filled the gaps of in the base-image using all other images from the respective growing season. The result was a nearly cloud-free image composite for both time points (2009 and 2011) before and after the storm event. We then classified the pre-disturbance image into 'forest' and 'non-forest' using a training data set generated automatically using the dark object approach and a modern supervised classifier.

Using these inputs, we mapped forest disturbance using the Disturbance Index (DI). The DI is a linear combination of normalized Tasseled-Cap bands. We combined the initial forest/non-forest map with the areas of forest disturbance and created a change-map for 2009-2011. To detect windfall disturbance, we first assumed that only two forms of disturbances occurred on the landscape: windfall and harvests. To extract training data for each disturbance type, we visually examined the Landsat imagery to determine how wind-related disturbance may be spectrally different from clear-cut harvests. Based on these observations we postulated that, compared to harvests, a wind-related disturbance site would have: 1) *Lower Tasseled Cap brightness values*: The Tasseled Cap brightness is a measure of the soil proportion in the signal and sensitive to the abundance of shade. After a recent clear-cut harvest soil is often exposed and shadows are rare, leading to high brightness values. In contrast, after a windfall event, biomass often remains, reducing soil reflectance and maintaining shadows. This would result in lower brightness values for windfall disturbance than clear-cut harvests; 2) *Higher Tasseled Cap wetness values*: The Tasseled Cap wetness provides information about the moisture content of a site. Major over- and understory removal, typical for a clear-cut harvest, strongly reduces tasseled cap wetness. Hence, a windfall disturbance will have on average a higher Tasseled Cap wetness value than a clear-cut harvest; and 3) *Lower short-wave infrared (SWIR) reflectance (Landsat band 5):* Similar to the Tasseled Cap wetness index, TM band 5 is sensitive to the amount of water in vegetation, but through an inverse relationship. On average, a windfall disturbance site would be expected to have lower SWIR then a clear-cut harvest due largely to more shadows in a windfall site.

Using normalized pixel values around a mean of zero following a standard Z-transformation, a histogram of all disturbed pixels will exhibit three main 'areas'. For example, in the case of band-5 reflectance, the locations of importance in the histogram are 1) the center, in which the spectral characteristics of windfall disturbance and clear cuts are essentially the same; 2) the left side of the histogram, which is dominated by 'windfall' pixels; and 3) the right side of the histogram, which is dominated by 'clear-cut harvest' pixels. The nature of a normal distribution makes it convenient to target these areas, which, in the case of band-5 reflectance, are located to the left (windfall) or to the right (clear-cut) of one standard deviation of the Gaussian distribution. We targeted pixels with these characteristics and extracted them as training data for the 'windfall' and 'clear-cut harvest' categories and supplied them to a modern classifier, using the six multi-spectral bands from Landsat and the same parameter-search method as for the initial forest/non-forest classification

Our results show that the separation between windfall and clear-cut disturbance was possible in over 75% of the disturbed area in the Russia site (Figure 2). Given the small number of studies that simultaneously classify windfall and clear-cut harvests using Landsat data, only a limited comparison to previous work can be made. Compared to the studies of fire disturbance and clear-cut harvests, our accuracies were generally lower. One likely explanation for the differences is that previous studies gathered training data with considerable user input. In contrast, our algorithm did not require any user intervention during the training process. It is well known that classifications that use manually identified training data generally result in higher overall accuracies compared to studies based on automated approaches. As such, the automation inherent in the algorithm is probably more prone to errors but comes with the advantage of not requiring manually collected training data.

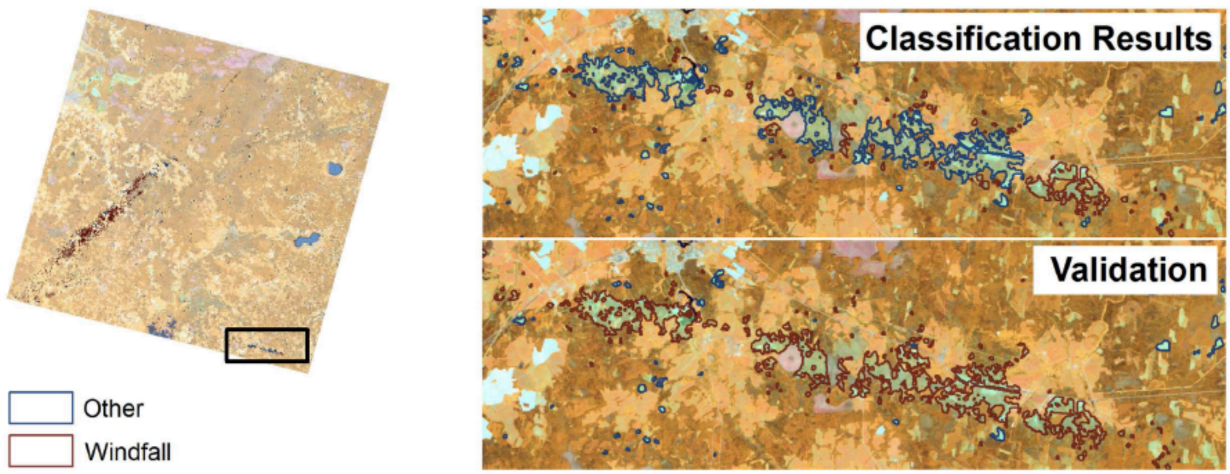

Figure 2. Validation of the training data to classify disturbance areas into 'windfall' and 'clear-cut harvest' in the spectral feature space in which they were generated.

2. Large area automated forest harvest mapping

In recent years, a variety of new methods have been developed to process and extract forest disturbance information, including semi-automated methods. These methods rely on the availability of relatively dense acquisitions (annual or biennial) of image data, which may exist only in a few places (e.g. the United States), despite the recent opening of the Landsat archive. In other locations, such as the European Russia, characterization of land cover change still occurs at relatively sparse temporal intervals, relying on pairs of images from different years. In these locations, image acquisition strategies coupled with issues related to cloud cover and the Landsat 7 Scan-Line-Corrector (SLC)-Off problem preclude image analysis that requires dense time series. In regions with limited temporal coverage, other approaches are still needed that are sufficiently generalizable and repeatable to document forest cover changes through time.

During this performance period, we also developed an automated forest harvest detection algorithm applied to large areas. The algorithm is based on the traditional pair-wise change detection method for forest disturbance mapping and provides a set of procedures that automate the process of image-to-image information translation. The methods rely on the distinct spectral signature associated with forest disturbance captured in a pair of Landsat images. The approach also relies on the ability to automatically capture this signal without user interaction and use it in a robust supervised classification algorithm that can handle incorrectly labeled training data. The method is essentially a supervised classification exercise, but eliminates the need for manual image interpretation for extracting training data. It is specifically designed to work in situations where dense time stacks of Landsat imagery are not available and the user is forced to use pair-wise image analysis.

The methodology includes the following steps: a) preprocessing of the Landsat image pairs by masking all clouds, cloud shadows, and non-forested areas such as water and agriculture; b) automated extraction of training data from local windows using the Landsat shortwave-infrared (SWIR) band difference image with local windows; c) removal of incorrectly labeled instances in the training data through *n*-cross validation; d) classification of disturbed areas using a supervised classification algorithm; and e) per-pixel to per-region translation of classification results using segmentation. The two most important parts of the method are automated extraction of training data and elimination of incorrectly labeled training data.

The process of extracting training data automatically is accomplished by identifying thresholds in histograms created from local SWIR-band (Landsat Band 5) image windows. More specifically, given a pair of SWIR reflectance images from two different dates, subtraction of the second image date from the first date will yield large negative reflectance values in disturbed forested areas. This is because SWIR reflectance is often low in mature forests and high in disturbed areas. In contrast, pixels representing forest recovery/regrowth exhibit a large increase in reflectance values between the two dates. The overall distribution of the SWIR reflectance difference image is Gaussian with a mean value near zero, primarily because the majority of the pixels between the two dates exhibit no change and form the bulk of the distribution. In contrast, areas that experienced forest disturbance or recovery are in the negative (removal) and positive (regrowth) ends of the distribution. The training data to identify pixels as disturbed, regrowth, and no-change were extracted using thresholds applied in local image windows of the two-date SWIR reflectance difference image. A local window is defined as a square portion of an image, usually 400 by 400 pixels in size, extracted using a moving window approach. More specifically, in the absence of water, clouds, cloud shadows, and non-vegetated surfaces – which would already be masked out – forest disturbance and regrowth pixels are identified as pixels whose value exceeds the threshold defined by 1.5 times the standard deviations in either direction. Note that while the value of the threshold changes across windows and images, the number of standard deviations used to identify the threshold is fixed across all windows and images. This number was determined empirically to balance omission and commission errors, using a large number of Landsat footprints. The purpose of using local image windows is to select an appropriate threshold that is not easily determined from global image statistics.

The automatic extraction of training data can result in labeling errors arising from the selection of inappropriate thresholds, inaccurate masks, and mis-identification of forested areas. This step focuses on improving the quality of the training data by identifying and eliminating mislabeled training samples. To achieve this goal, we followed a filtering approach, in which the raw training data is passed through a number of classification algorithms through an *n*-fold cross-validation process that serve as a *filter* to remove mislabeled samples. An interesting feature of this approach is that it employs a consensus filter where only those samples that all of the individual classifiers tagged as mislabeled are discarded.

The overall accuracy of the forest change and stable forest classes using the automated classification across all footprints and time periods ranged from 40 to 99 percent with a mean value of 82.3±12.5 percent (N = 85). These overall accuracy values are quasi-normally distributed, but are positively skewed having the bulk of the quantities between 80 and 90 percent. In all cases, the derived forest change maps accurately capture the spatial distribution of forest disturbance as evidenced by both Landsat and NAIP imagery interpretation (Figure 3). Second, most errors are confined to commission errors, especially in landscapes with complex terrain. Third, non-ideal acquisition dates for image pairs such as an acquisition too early or too late in the season, as well as long periods of time between image dates reduce the reliability of maps derived from the automated procedure. While the automated approach successfully determined harvested conifer stands between the dates, some issues remain due to errors of omission given the leaf-off status of the deciduous stands. The successful launch of Landsat 8 will provide more opportunities to acquire data during leaf-on seasons as a potential and simple remedy to this issue.

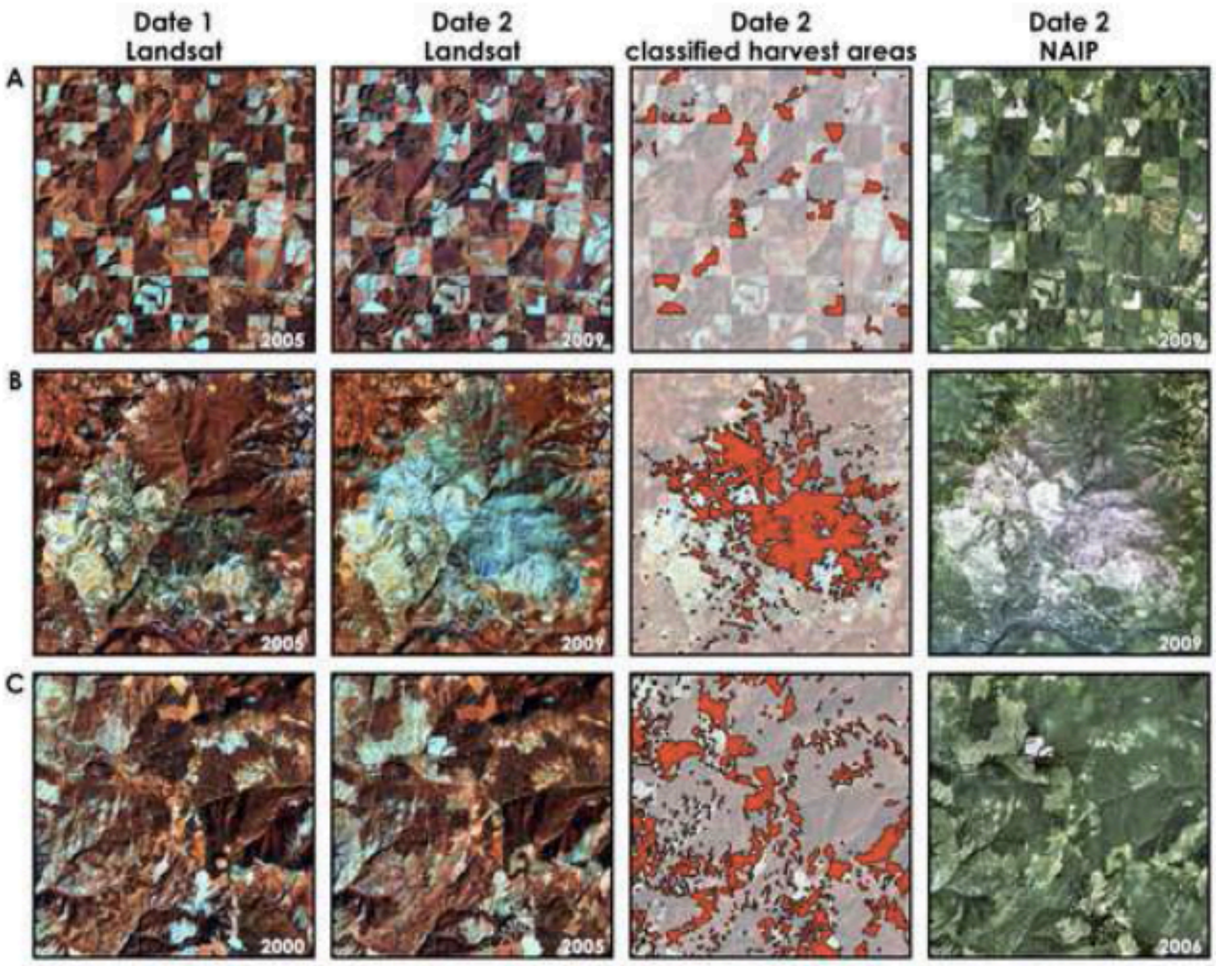

Figure 3. Samples used in the visual assessment of map accuracy. The left two columns are the Landsat images used in the change detection process. The third column is the derived change polygons displayed on the original Landsat images. The last column is NAIP imagery acquired in the year closest to the post-forest disturbance year. The locations are: (top). Oregon (path/row 46/30); (middle). Oregon (path/row 46/30); (lower) Washington (path/row 44/26).

On average, the filtering process removed 20 percent (20 ± 7 %) of the original, automatically extracted training data. The removal of incorrectly-labeled training data significantly improved the classification accuracy (at the 0.05 level of significance using a paired t-test) in many cases. The primary improvements occurred through a reduction of false positives, or areas falsely identified as forest change when there was none (Figure 4). The changes in the landscape are clearly manifested in both the before- and after-harvest images, as well as in the Band 5 difference image. Without filtering, most of the coniferous stands located in the center of the scene are labeled as harvested, although visual inspection of the Landsat images clearly indicates that this is not the case (Figure 6a, lower left panel). When this image is classified using the filtered training data, this stand is not longer a part of the forest change category, which is the correct expectation (Figure 6a, lower left panel). Finally, the image in the lower right panel highlights the differences between classification results achieved with and without filtering.

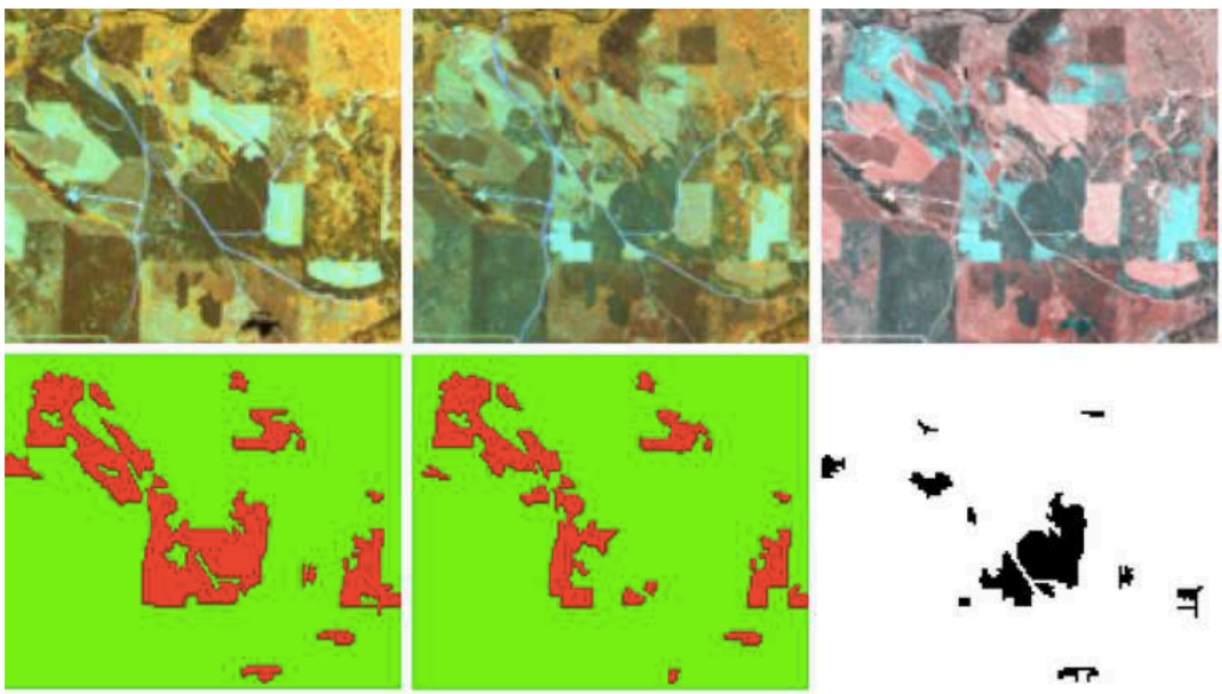

Figure 4. The effects of filtering mislabeled training data on classification accuracy. (top left) pre-harvest image bands 453 as RGB; (top center) post-harvest image bands 453 as RGB; (top right) band 5 difference image; (bottom left) filtered result; (bottom center) unfiltered result; and (bottom right) difference between the results.

The requirements for the method to produce appropriate forest change products include peak growing season acquisition of two images that form the pair, accurate masking of forest and non-forested areas, and a robust methodology to screen clouds and cloud shadows. When these conditions are met, the method can be used to produce forest change maps rapidly and with accuracies on par with those delivered by traditional change detection methods. One of the advantages of the proposed approach is the speed: a pair of Landsat images can be processed to create a change map in under an hour. Finally, the automated approach can screen incorrectly labeled instances from the dataset used to train the supervised classification algorithm, which improves the classification accuracy by reducing commission errors.

3. Modeling green leaf phenology using Dynamic Time Warping and all available Landsat data

Characterizing green leaf phenology is an important measure to describe the development of vegetation over the year and thus offers ways to characterize the interaction between climate and the biosphere. Remote sensing from satellites allows for the observation of green leaf phenology across large spatial scales because of the standardized and repeated measurements, and has become popular in the past. While data from the MODIS instrument are used to generate phenological products, many applications would benefit from a higher spatial resolution dataset such as those maybe provided by Landsat data. However, a major problem with the Landsat data is that, while the spatial and spectral resolution of Landsat satellites appear to be ideal to describe green leaf phenology, the temporal resolution (i.e., the number of images per year that are available for the same location) is less than imperfect; the 16-day repeat cycle makes extraction of phenologically meaningful information from a single year challenging if not impossible. Our goal was to revisit this problem and combine all available Landsat data into a single year, following dynamic time warping (DTW) of non-target year images.

When combining multi-year imagery within one time-series, the challenge is to remove the effect of differences in green leaf phenology between years. In other words, contrary to highlighting phenological differences across years, the goal is to eliminate the between-year variation in vegetation phenology to create a synthetic one-year-phenology using Landsat data. The DTW method can help to overcome the between-year differences. DTW was originally developed for purposes of speech recognition, but there is an increasing number of applications of DTW in remote sensing questions, mostly of time-series problems. In this work we applied the DTW idea to re-align all available Landsat imagery between 2002 and 2012.

The aim of DTW is to stretch and compress two time series locally to ultimately make them as similar as possible. The underlying idea is that, while the two time series are potentially very different at a certain day of the year, their overall evolution is very similar. For example, in case of the phenological variation throughout the year, the green-up stage might be slightly offset in different years. However, in every year there will be a green-up stage of vegetation, and this green-up stage will always precede the maturity stage of the vegetation. In other words, across multiple different years, the phenological evolution over the year will likely always follow the double logistic function, but the parameters of the function will vary across years. Thus, given a reference phenology (e.g., for the year 2005) and the phenology of a second, different year (e.g., 2007), the goal is to find for each day in the 2007-phenology the corresponding day of year in the reference phenology. We call the corresponding days of the two time series 'Day-of-Year-matches' (DOY_m), which can be interpreted, for example, that May 15th of the year 2007 corresponds phenologically to May 25th of the reference year (2005). Thus, using DTW we translate the chronological time series into a phenological time series.

One main criterion to apply DTW is to have a sufficiently accurate measure to describe the phenology of the two time series. In this work, we used the MODIS Enhanced Vegetation Index (EVI) to find the DOY_m for each day of year. Using MODIS EVI, we then re-organized a Landsat time series, consisting of all available Landsat images between 2002 and 2012 to create a new synthetic time series of a high temporal resolution that corresponds to the green leaf phenology of any reference year.

Merging all available Landsat data from all three types (TM5, ETM+, ETM+-SLC-off) between 2002 and 2012 increased the data that were available in the time series from an average of 24.9 images per year to an average of 274 images. Qualitatively, applying DTW to re-aligning Landsat imagery made the resulting Landsat EVI time series appear much more similar to the reference EVI time series from MODIS (Figure 5). Comparing the Landsat time series and MODIS time series, all keydates showed a higher correlation coefficient compared to the keydate-comparison between Landsat and the PhenoCam time series, and this was consistent across all keydates. The average differences between keydate-estimates were also comparatively lower between the Landsat and MODIS time series than between the Landsat and PhenoCam time series. Our models revealed that between Landsat and MODIS, the Green-Up dates were on average 1.3 days apart (standard deviation 1.04 days), the Start-of-Season date 1.2 days (0.84), the Maturity date 1.32 days (1.31), the Senescence date 2.34 days (3.54).

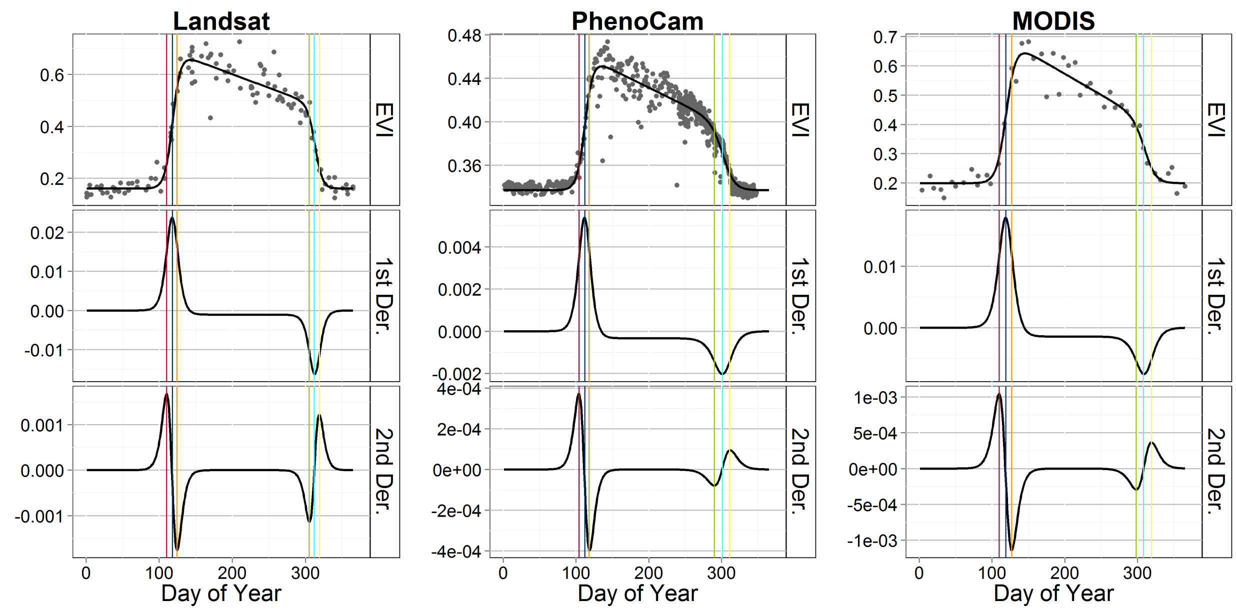

Figure 5. Example of the three phenology-profiles (Landsat, PhenoCam and MODIS) and the estimation strategy of the keydates. The top row represents the EVI/ExGm time series, whereas the middle and bottom row represent the first and second derivative of the phenology-profiles. The keydates are (chronologically with the year): green-up date (GU, red line), Start-of-Season date (SoS, blue line), Maturity date (Mat, orange line), Senescence date (Sen, green line), the End-of-Season date (EoS, light blue line), and Dormancy date (Dorm, yellow line).

Green leaf phenology is an important measure to describe the dynamic of vegetation throughout a year. Characterizing green leaf phenology at a high spatial resolution might help to map mixed forest stands more accurately or to improve our understanding of other ecological questions such as the interplay of phenological evolution and bird migration. In this performance period, we developed an approach to model green leaf phenology at the resolution of Landsat satellites. There is substantial agreement between the Landsat phenology and the MODIS reference phenology as well as the phenology on the ground from PhenoCam data. This makes the presented green-leaf phenology product being a considerable alternative to existing remote sensing based products, though at a much higher spatial resolution.

Publications resulting from this grant (* indicates a student)

*Wendland, K.J., Lewis, D.J., Alix-Garcia, J., Ozdogan, M., Baumann, M., and Radeloff, V.C., 2011. Regional- and district-level drivers of timber harvesting in European Russia after the collapse of the Soviet Union, *Global Environmental Change*, 21, 1290-1300.

*Baumann, M., Ozdogan, M., Kuemmerle, T., Wendland, K., Esipova, E., and Radeloff, V., 2012. Using the Landsat record to detect forest-cover changes during and after the collapse of the Soviet Union in the temperate zone of European Russia, *Remote Sensing of Environment*, 124: 174-184. http://dx.doi.org/10.1016/j.rse.2012.05.001.

Ozdogan, M., A Practical and Automated Approach to Large Area Forest Disturbance Mapping with Remote Sensing, *PLOS One, in press.*

*Baumann, M., Ozdogan, M., Peter T Wolter, P.T., Krylov, A., Vladimirova, N., and Radeloff, V.C., Remote sensing of windfall disturbance with Landsat, *Remote Sensing of Environment, in press.*

*Wendland, K.J., Baumann, M., Lewis, D., Sieber, A., Radeloff, V.C. Protected area effectiveness in European Russia during and after the collapse of the Soviet Union. *Land Economics*, *in review.*

Baumann, M., Ozdogan, M., Richardson, A.D., and Rodeloff, V.C., Modeling green leaf phenology using Dynamic Time Warping and all available Landsat data, to be submitted to Remote Sensing of Environment*.