Support Vector Machines and Chain Classification for large-area forest disturbance mapping in the Carpathians

Tobias Kuemmerle[1], Volker C. Radeloff[1], Jan Knorn[2], Andreas Rabe[2], Sebastian van der Linden[2], Mutlu Ozdogan[1], and Patrick Hostert[2]

[1] Department of Forest and Wildlife Ecology, University of Wisconsin-Madison

[2] Department of Geography, Geomatics Lab, Humboldt University Berlin, Germany

Background

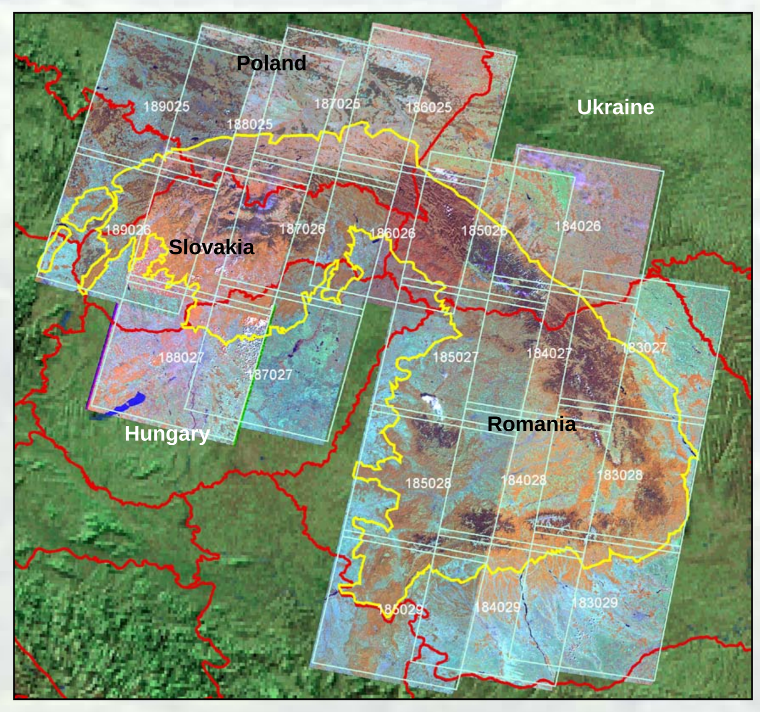
Large-area mapping

- New and exciting opportunities for LCLUC mapping due to free Landsat image archives
- Traditional approaches using Landsat images mostly focused on image pairs
- Most LCLUC studies to date have also assessed every Landsat footprint separately, which may not be feasible for large areas
- Overall, relatively few generalization efforts so far (but see Woodcock et al. 2001)
- We need approaches that
  - make full use of the Landsat archive without having to handle each image individually
  - minimize user input and allow for automatization, and
  - are transferable to larger areas

The Carpathian Mountains

- The Carpathians are one of Europe's last remaining large and undisturbed forests
- High biodiversity, rich cultural diversity, important ecosystem services, one of the last refuges for Europe's large mammals
- Drastic land use changes after the fall of the Iron Curtain
- Widespread land cover change, including
  - farmland abandonment
  - substantial, often undocumented logging
  - farmland parcelization
- Local cases studies map the rates and spatial patterns of Carpathian LCLUC
- Yet, Carpathian-wide assessments of LCLUC are lacking

Abandoned farmland in the Polish Carpathians (top) and clearcut in the Ukrainian Carpathians (left).



The Carpathian Mountains in Eastern Europe

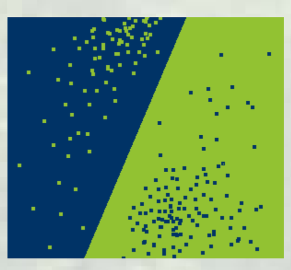
Objectives:

- Develop a robust forest disturbance (full canopy removal) mapping method, applicable to the entire Carpathians
- Use Support Vector Machines to generalize in time
- Use overlap areas between images to generalize in space

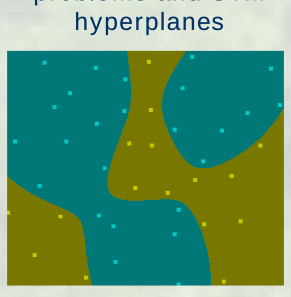
Change detection approach

Support Vector Machines (SVM)

- SVM Concept
  - Delineate two classes by fitting a separating hyperplane
  - Only training pixels describing class boundaries are important
  - Complexity in low-dimensional spaces is linearly separable in high-dimensional spaces
  - Use kernel functions to transform training data into high-dimensional spaces



Two binary classification problems and SVM hyperplanes



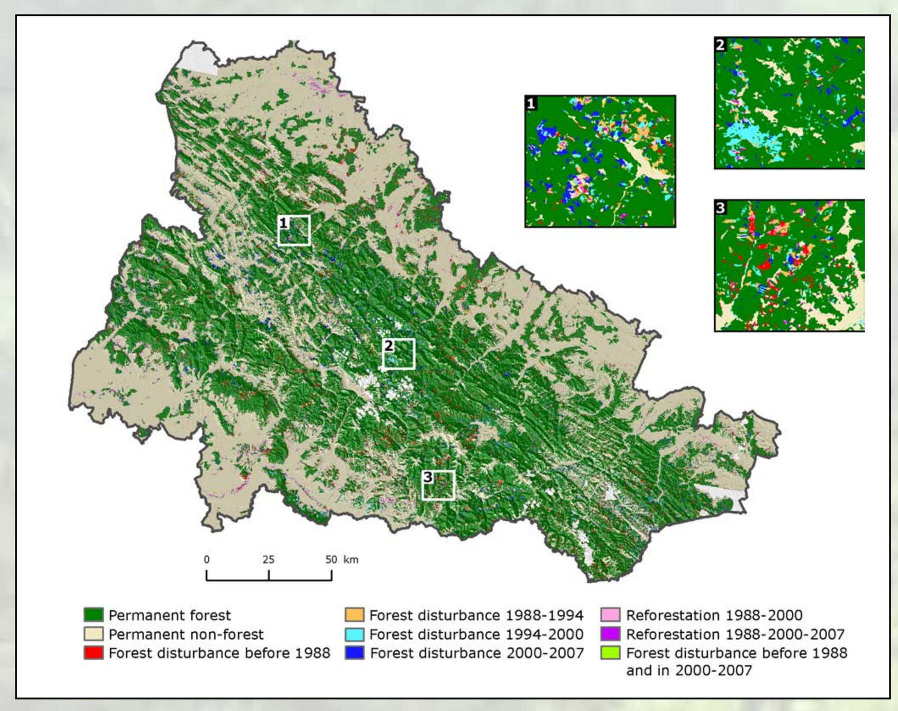
- Advantages of SVM
  - Can handle complex classes (typical for LCLUC)
  - Require potentially few training data
  - Often outperform other classifiers
  - Successful applications in forest mapping and change detection (e.g., Huang et al. 2008, Kuemmerle et al. 2008)

Generalization in time & change detection

- Forest/non-forest maps
  - Random sample of ground truth points based on GoogleEarth™ high-resolution images
  - discard points that are not constant in time (visual assessment of Landsat images)
  - SVM (C-SVM and Gaussian kernel function) to classify all images of a Landsat footprint
  - Automatic SVM parameterization and accuracy assessment (cross-validation)
- Change detection
  - Post-processing and rule-based identification of change trajectories
  - Independent assessment of disturbance detection rate

Results:

- SVM resulted in reliable forest/non-forest maps
- Mean overall accuracy = 97.88% (range 94.7-99.4%; kappa = 0.88-0.98)
- High disturbance detection rates (~89.4%)



Forest disturbances (full canopy removal) between 1988-2007 in the Ukrainian Carpathians.

- Relatively low numbers of training points (<500 per class) yielded robust classifications
- Next steps
  - Can active learning reduce the number of training points required?

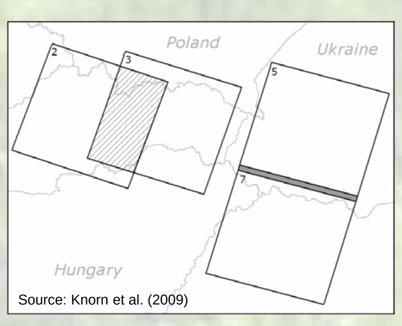
References

- The full SVM procedure is implemented in the free ENVI/IDL sotware imageSVM 2.0 (www.hu-geomatics.de)
- Kuemmerle, T., Chaskovskyy, O., Knorn, J. Kruhlov, I. Radeloff, V.C., Keeton, W.S., and Hostert, P. (2009): Forest cover change and illegal logging in the Ukrainian Carpathians in the transition period from 1988 to 2007. Remote Sensing of Environment, in press.

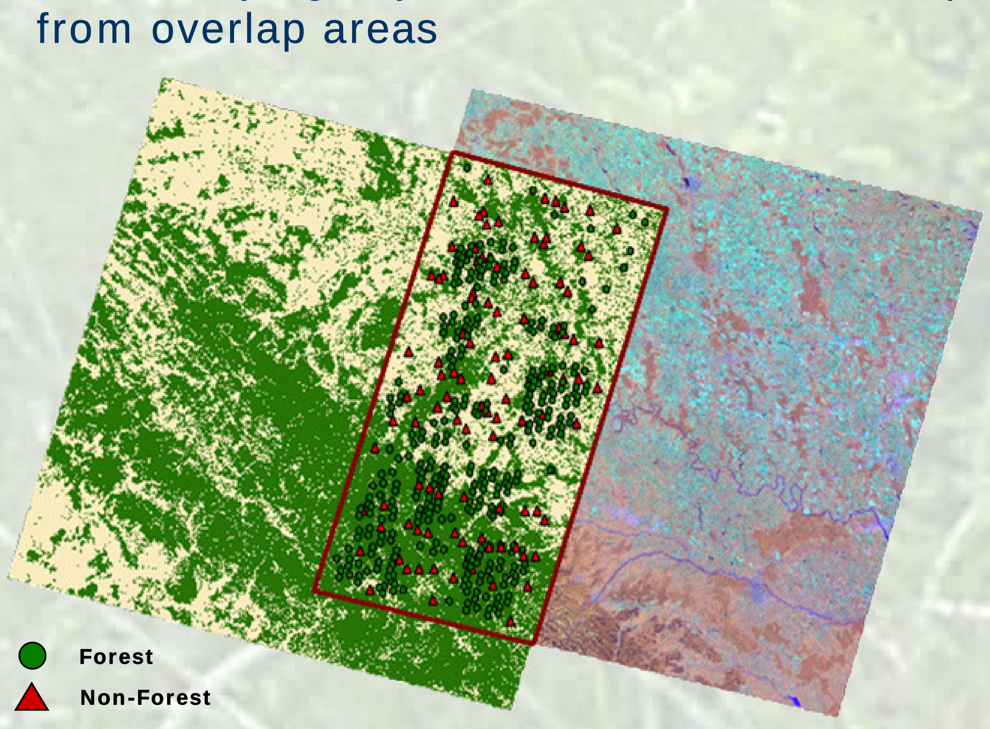
Large-area mapping

Generalizing space

- Mid-latitude Landsat footprints have substantial horizontal overlap → make use of these overlap areas to generalize in space



- If an initial classification exists, training data for classifying adjacent scenes can be sampled from overlap areas



- Does not require atmospheric correction or radiometric matching of images
- Can be applied along or across track
- Can be applied to image 'chains'

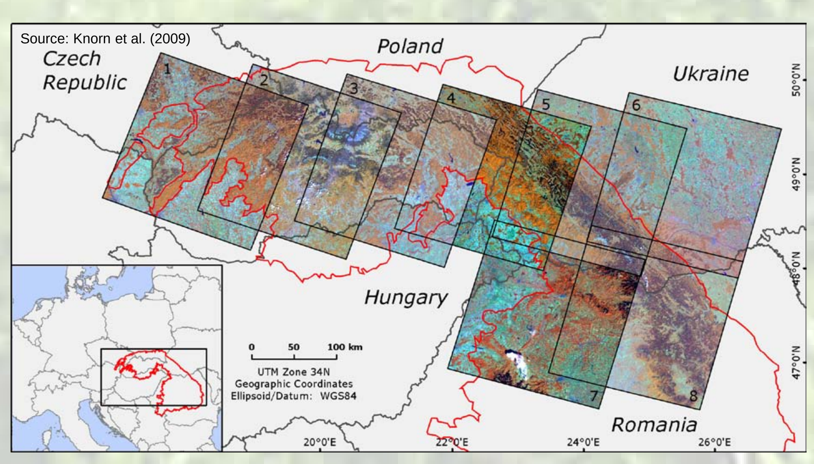


Image data used for the chain classification tests

- Classifier: Support Vector Machines
- Different chain lengths, starting points, directions of chain classification, etc
- Accuracy assessment based on 1,400 ground truth points from GoogleEarth™ per image
- Comparison to single-image SVM classifications using independent training data

Results:

- Even a chain of six images resulted only in a 5.1% accuracy loss compared to a reference classification for the last image
- Mean accuracy loss of 1.9%
- Dependency on starting point
- Some limitations, e.g. classes not well represented in overlap areas, low initial classification accuracy, or for images with haze
- Overall, chain classification appears to be a very promising tool for large-area mapping!
- Next steps
  - How does chain classification compare to signature extension?
  - Chain classification for change detection or more complex classification problems?
  - How important is the choice of the classifier (so far SVM)?

Reference

- Knorn, J., Janz, A., Radeloff, V.C., Kuemmerle, T. and Hostert, P. (2009): Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. Remote Sensing of Environment, in press.

Acknowledgements

- Humboldt Universität zu Berlin, Germany
- Alexander von Humboldt Foundation
- NASA Land Cover / Land Use Change Program
- University of Wisconsin-Madison
- European Science Foundation
- German Academic Exchange Service