Automated Forest Cover Change Analysis

Chengquan Huang, Kuan Song, John Townshend, Samuel Goward, Department of Geography, University of Maryland, College Park, MD 20742

Paul Davis, Sunghee Kim, Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742

Jeffrey Masek, Biospheric Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771

Zhiliang Zhu, U.S. Geological Survey/EROS, Sioux Falls, SD 57198

Introduction

Forest cover change (FCC) affects many pressing global change issues, including carbon, hydrology, and biodiversity. Accurate estimate of change area requires fine spatial resolution (e.g., Landsat class) data. Global Landsat data sets called GeoCover have been assembled for three epochs around 1975, 1990, and 2000. A 2006 Mid-Decadal Global Land Survey (MDGLS) Landsat data set is forthcoming. Although the three GeoCover data sets are freely available and have been widely distributed, global FCC analysis still faces many challenges:

1. Each global Landsat data set is not a "single", consistent view of the earth's surface, but consists of numerous individual images.
2. Large variations exist among the images due to differences in solar geometry, atmospheric condition, and vegetation phenology.
3. The spectral signatures of both forest and non-forest land cover types have considerable geographic variability.
4. Thus it is difficult to apply a change analysis model developed for one image pair to another image pair.

Objectives

In this poster we describe an automated forest cover change analysis method called training data automation (TDA)-support vector machines (SVM). This method automatically delineates training samples based on individual Landsat images and uses them to train the SVM to generate localized change analysis models. Because each image pair is analyzed using a SVM model trained using locally derived training data, this method can address the above described challenges.

We have evaluated the change products produced using this method for locations selected from major forest biomes across the globe and tested it in a semi-production mode for two large regions.

The TDA-SVM Method

1. Overview of TDA-SVM
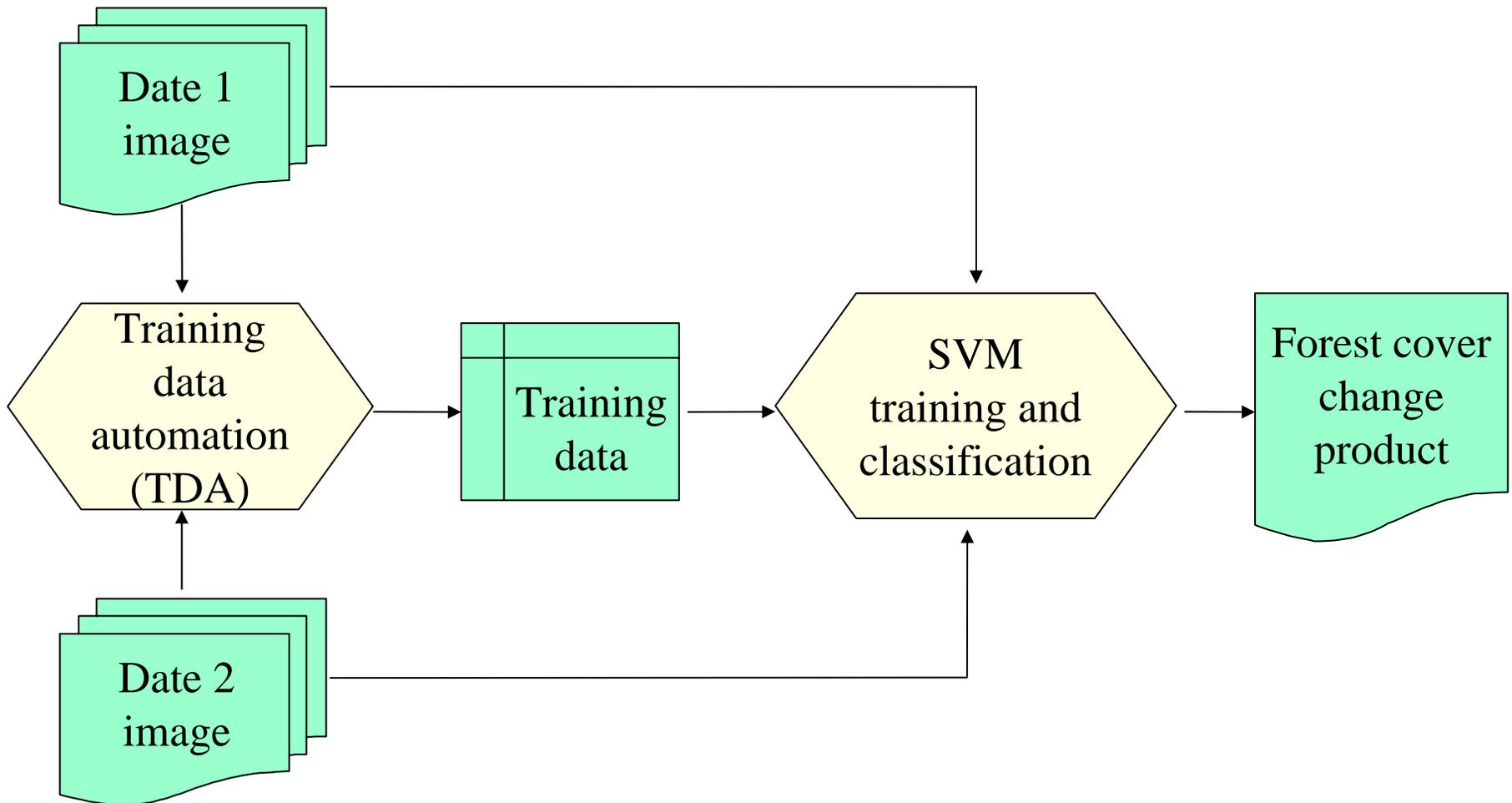
2. Training Data Automation (TDA)
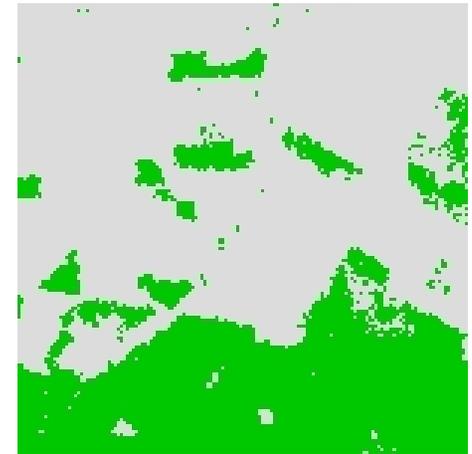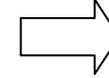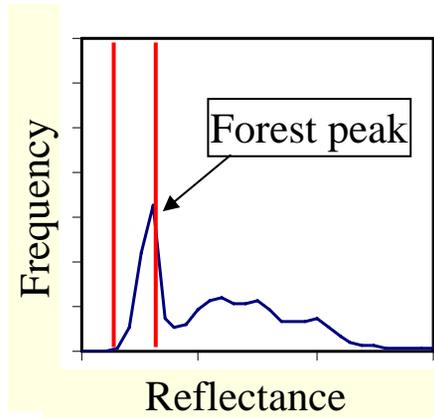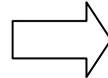
(a) Local image window (TM bands 5, 4, and 3 shown in red, green and blue)

(b) Forest peak in a histogram constructed using the red band and threshold values defined by the peak (red lines)

(c) Forest training pixels (in green color) delineated by the threshold values defined by the forest peak

2.1 Forest TDA: Basic Concept (↑) and Detailed Steps (↓)

```
┌─────────────────┐
│    DN image     │
└─────────────────┘
         │                    ┌─────────────────────┐
         │        ◄───────────│  Standardize image  │
         ▼                    │     radiometry      │
┌─────────────────┐           └─────────────────────┘
│  TOA or surface │
│ reflectance image│          ┌─────────────────────┐
└─────────────────┘           │  Mask non-forest    │
         │        ◄───────────│    dark objects     │
         ▼                    └─────────────────────┘
┌─────────────────┐
│Local image window│
└─────────────────┘           ┌─────────────────────┐
         │                    │  Create histogram   │
  next   │        ◄───────────│ Identify forest peak│
 window  ▼                    │  Threshold image    │
┌─────────────────┐           └─────────────────────┘
│  Forest training │
│     samples      │──────────┐
└─────────────────┘           ▼
         ▲          ┌──────────────────┐     ┌─────────────────┐
         └──────────│   Consistency    │◄────│  Existing land  │
                    │    checking      │     │ cover products  │
                    └──────────────────┘     └─────────────────┘
```
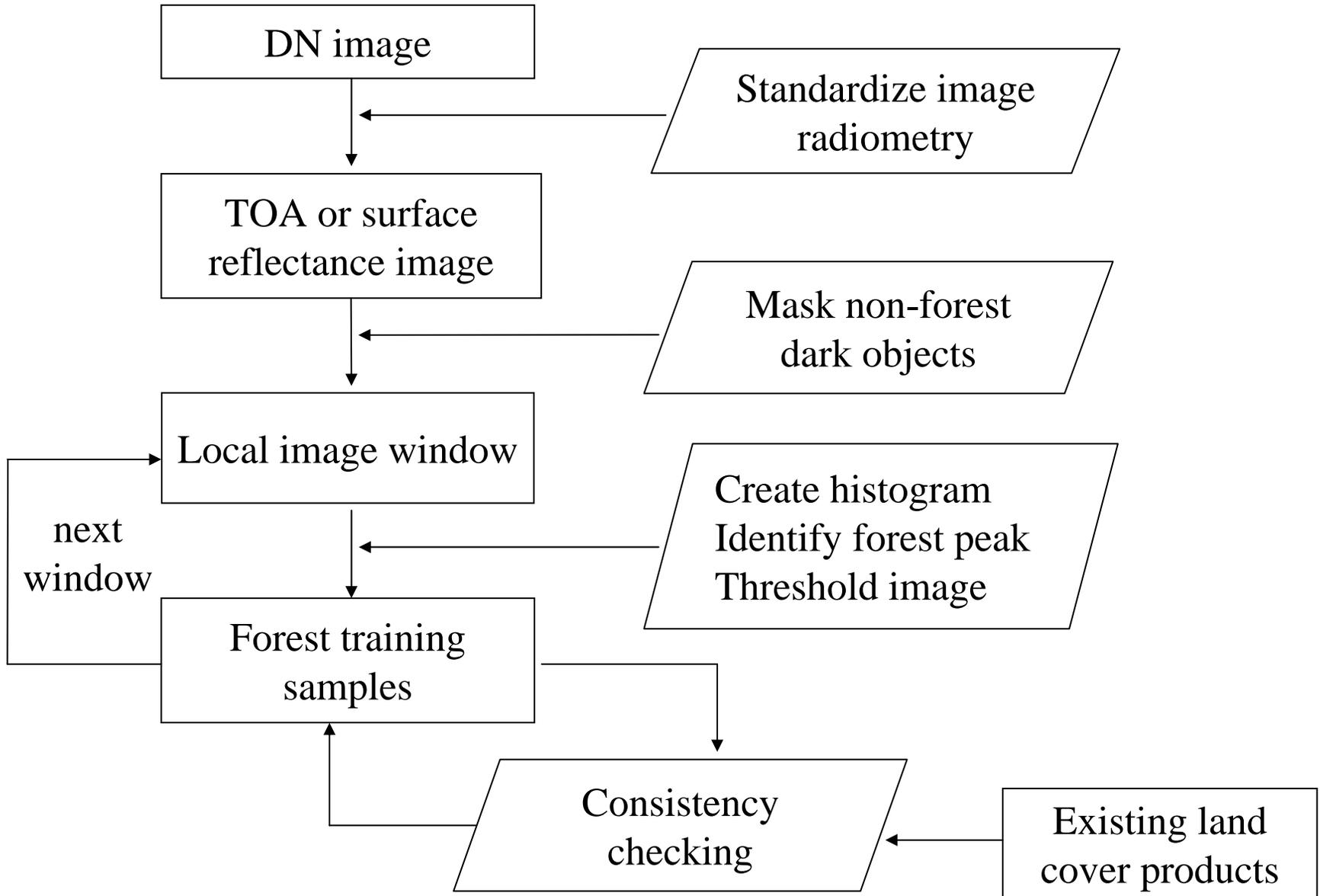
2.2 Non-forest TDA

An integrated forest index (*IFI*) is calculated using the mean and standard deviation of forest training samples to measure the likelihood of a pixel being a forest pixel:

$$IFI = \sqrt{\frac{1}{NB} \sum_{i=1}^{NB} \left(\frac{b_i - \overline{b_i}}{SD_i} \right)^2}$$

High *IFI* values indicate low likelihood of being forest. A non-forest training pixel is identified if its *IFI* value > a threshold value.

2.3 Less Pure Training Samples

Pixels having slightly higher *IFI* values but neighboring previously delineated forest training samples are identified as forest training samples.

Pixels having slightly lower *IFI* values but neighboring previously delineated non-forest training samples are identified as non-forest training samples

2.4 Spectral Signals for Change Classes

Date 1 forest training samples	Persisting forest →	Date 2 forest training samples

Forest loss

Forest gain

Date 1 non-forest training samples	Persisting non-forest →	Date 2 non-forest training samples
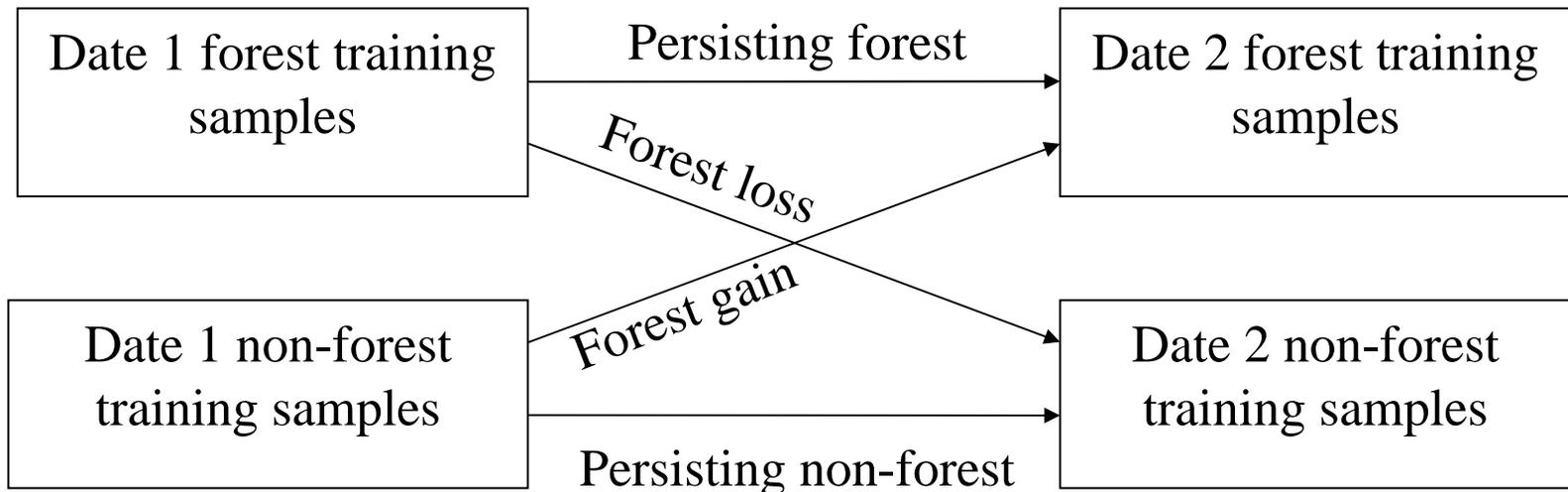
3. SVM

SVM defines optimal class boundaries using optimal separating hyperplanes, which are defined by minimizing the following function:

Linear boundary

Non-linear boundary

$$F(w) = \frac{1}{2}(w \cdot w)$$
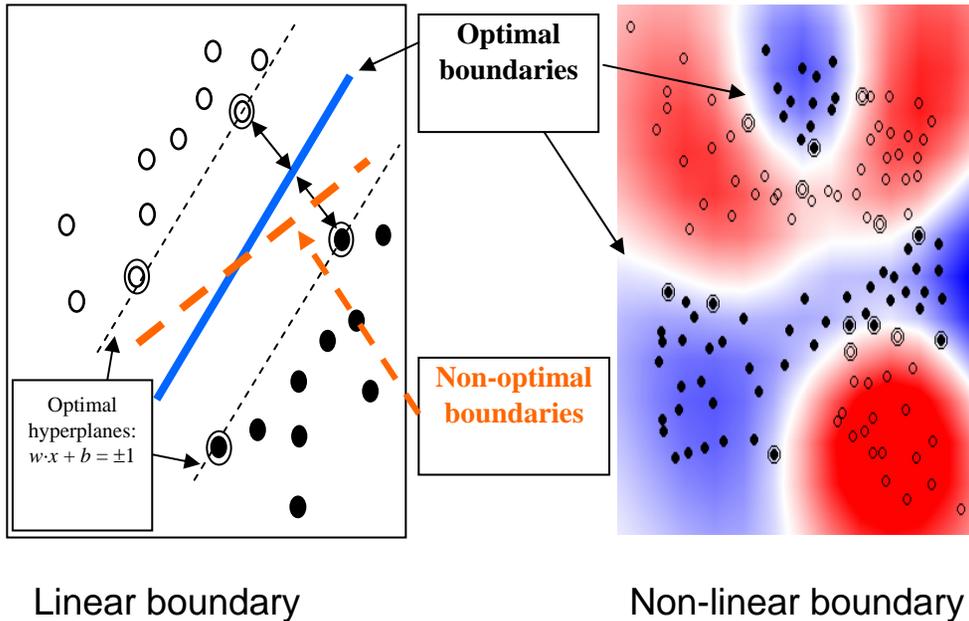
Other classifiers such as decision trees and neural networks are designed to locate the first class boundary that satisfies the training requirements, which mostly likely will not be the optimal boundary.

The optimal class boundary is more likely to yield higher accuracy than any other non-optimal boundary when used to classify unseen samples.

$$K(x_i, x_j) = e^{-\gamma(x_i - x_j)^2}$$

SVM uses kernel functions to locate nonlinear optimal boundaries:

Test and Validation

Table 1. Accuracy values of TDA-SVM derived FCC products calculated using design-based accuracy assessment procedures

Accuracy measure	Persisting forest	Persisting non-forest	Forest loss	Forest gain
Eastern Virginia, U.S.A, WRS2 path 15/row 34: Overall accuracy = 92.7				
User's accuracy	96.7	94.9	88.2	74.5
Producer's accuracy	96.7	86.2	88.3	88.4
Southeastern Finland, WRS2 path 187/row 17: Overall accuracy = 94.8				
User's accuracy	96.7	96.1	80.6	95.9
Producer's accuracy	97.9	93.3	96.3	82.2
Central Mississippi, U.S.A., WRS2 path 22/row 37: Overall accuracy = 89.4				
User's accuracy	95.8	92.2	85.7	72.6
Producer's accuracy	88.0	95.3	80.7	88.9
Western Washington, U.S.A., WRS2 path 46/row27: Overall accuracy = 94.4				
User's accuracy	93.4	93.0	98.2	100.0
Producer's accuracy	99.5	99.6	100.0	70.6
Connecticut, U.S.A., WRS2 path 13/row 31: Overall accuracy = 96.5				
User's accuracy	98.5	95.2	88.2	80.5
Producer's accuracy	99.7	92.2	100.0	77.6
Eastern Paraguay, WRS2 path 224/row 78: Overall accuracy = 94.2				
User's accuracy	76.5	95.9	83.9	N/A
Producer's accuracy	93.0	91.3	80.4	N/A
Western Paraguay, WRS2 path 228/row 75: Overall accuracy = 89.6				
User's accuracy	89.8	96.4	85.1	N/A
Producer's	93.1	79.7	91.8	N/A

1. The TDA-SVM method was initially tested using GeoCover 1990 and 2000 images at 19 WRS2 locations selected from major forest biomes across the globe. For 5 of those areas IKONOS images were used to derive design-based accuracy estimates of the change products for the areas covered by the IKONOS images. In addition, a wall-to-wall forest cover change product of Paraguay was independently developed using conventional methods with intensive human inputs. That product was found highly accurate (overall accuracy ~95%) when evaluated using high resolution aerial photos, IKONOS, and QuickBird images. That product was used as reference to evaluate change maps produced by the TDA-SVM method for two Paraguay sites.

2. The remaining 12 sites were evaluated by visually examining the change products against the input Landsat images and available IKONOS images. Four visual assessment examples are shown here, including two areas where accuracy assessment was also performed (WRS2 path 46/row 27 and path 187/row 17).

Table 2. IKONOS images and reference samples (TM/ETM+ pixels) used to evaluate the FCC products produced by the TDA-SVM method

Study area	No. of IKONOS images used	Area covered by IKONOS images	Number of reference samples
Eastern Virginia, U.S.A, WRS2 path 15/row 34	1	115	265
Southeastern Finland, WRS2 path 187/row 17	3	220	379
Central Mississippi, U.S.A., WRS2 path 22/row 37	65	7468	296
Central western Washington, U.S.A., WRS2 path 46/row27	6	146	228
Connecticut, U.S.A., WRS2 path 13/row 31	12	2467	209

Table 3. 19 study areas where the the TDA-SVM method was initially tested

path/row	Location	WWF biome
231/67	Eastern Rondonia, Brazil	Tropical & subtropical moist broadleaf forests
226/69	Central Mato Grosso, Brazil	Tropical & subtropical moist broadleaf forests
224/78	Eastern Paraguay	Tropical & subtropical moist broadleaf forests
20/48	Western Guatemala	Tropical & subtropical moist broadleaf forests
228/75	Western Paraguay	Tropical & subtropical dry broadleaf forests
131/47	Northwestern Thailand	Tropical & subtropical dry broadleaf forests
22/37	Central Mississippi, U.S.A.	Temperate broadleaf & mixed forests
15/34	Eastern Virginia, U.S.A.	Temperate broadleaf & mixed forests
13/31	Connecticut, U.S.A.	Temperate broadleaf & mixed forests
23/28	Northwestern Michigan, U.S.A.	Temperate broadleaf & mixed forests
19/27	Southeastern Ontario, Canada	Temperate broadleaf & mixed forests
116/31	Northeastern China	Temperate broadleaf & mixed forests
34/32	North Central Colorado, U.S.A.	Temperate conifer forests
46/27	Central western Washington, U.S.A.	Temperate conifer forests
47/27	Western Washington, U.S.A.	Temperate conifer forests
192/27	Western Austria	Temperate conifer forests
23/26	South central Ontario, Canada	Boreal forests/Taiga
34/23	Western Manitoba, Canada	Boreal forests/Taiga
187/17	Southeastern Finland	Boreal forests/Taiga

Western Washington, U.S.A. (path 46/row27)

Western Manitoba, Canada (path 34/row 23)

Southeastern Finland (path 187/row 17)
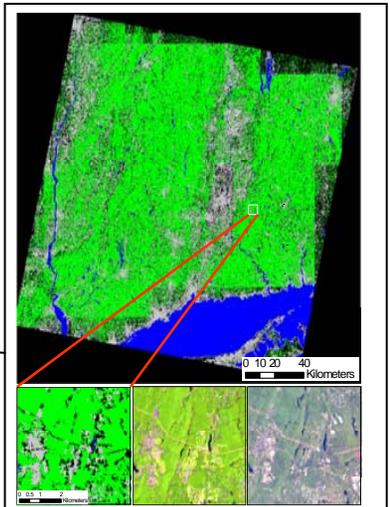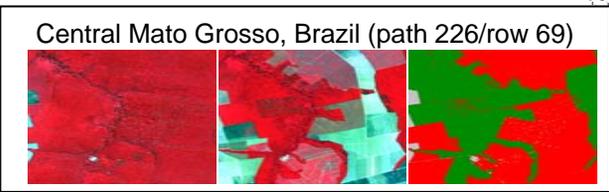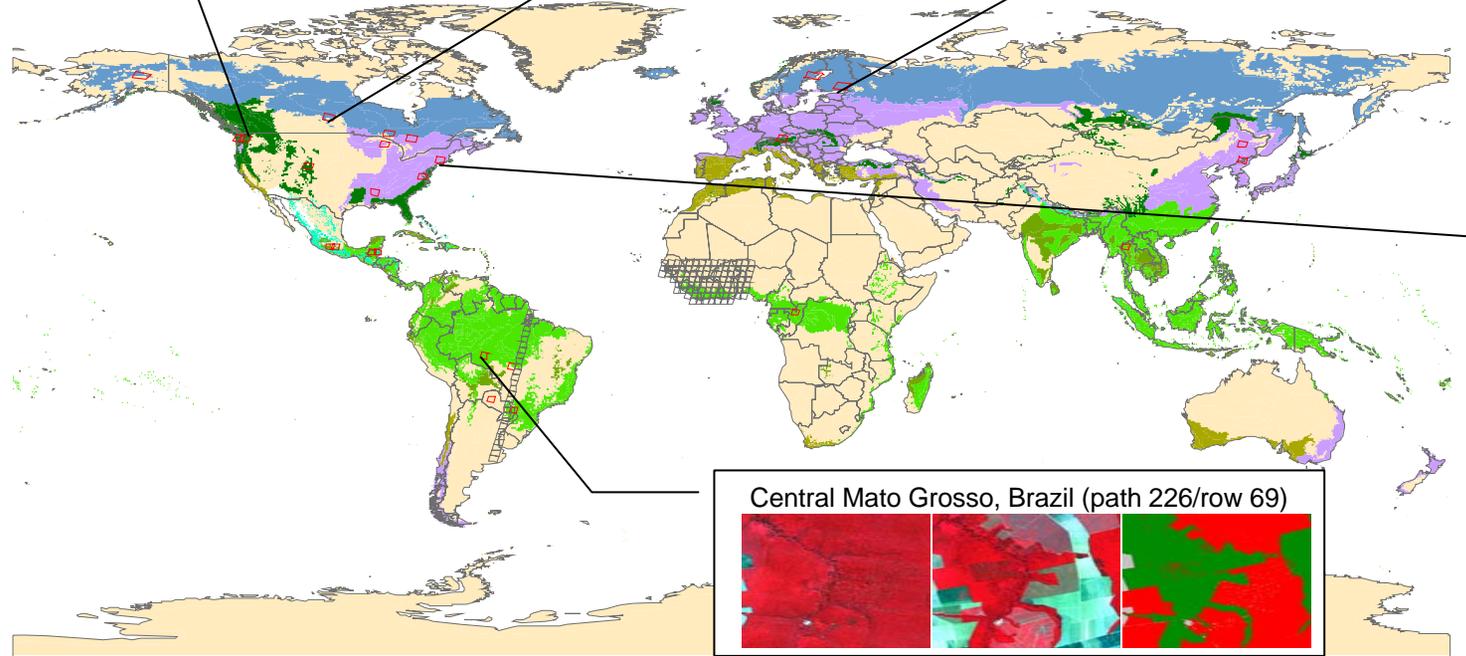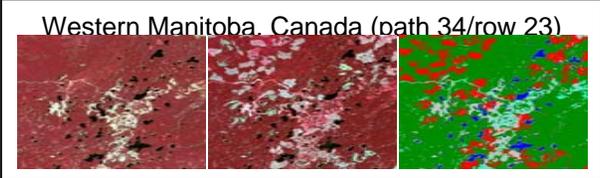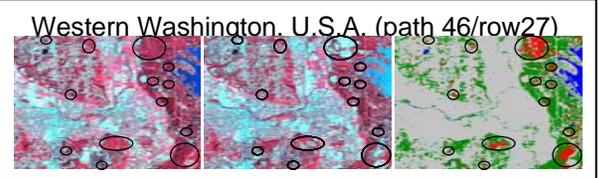
Example TDA-delineated forest (green) and non-forest (gray) training samples for the Connecticut, U.S.A. site (path 13/row 31) and visual assessment at full resolution (bottom left) using the Landsat image (bottom middle, R/G/B = bands 5, 4, and 3) and an IKONOS image (bottom right, R/G/B = IKONOS red, green, and blue bands).

Central Mato Grosso, Brazil (path 226/row 69)

19 WRS2 where FCC products were assessed
105 WRS2 analyzed by TDA-SVM

WWF Biomes

Tropical & Subtropical Moist Broadleaf Forests
Tropical & Subtropical Dry Broadleaf Forests
Tropical & Subtropical Coniferous Forests
Temperate Broadleaf & Mixed Forests

Temperate Conifer Forests
Boreal Forests/Taiga
Tropical & Subtropical Grasslands, Savannas & Shrublands
Temperate Grasslands, Savannas & Shrublands
Flooded Grasslands & Savannas
Montane Grasslands & Shrublands
Tundra
Mediterranean Forests, Woodlands & Scrub

Deserts & Xeric Shrublands
Mangroves
Lake
Rock and Ice

Geographic distribution of test areas (WRS2 tiles) within the World Wildlife Foundation (WWF) Biomes, along with example results for 4 test areas. In each example the 1990 image (first) and 2000 image (second) are shown with R/G/B = bands 4, 3, 2. In the derived change map (third), persisting forest, persisting nonforest, persisting water, forest loss and forest gain are shown in green, gray, blue, red, and cyan respectively.

3. The TDA-SVM method was used to produce change products using GeoCover 1990 and 2000 images for two large areas: a north-south stripe across South America (WRS2 path 225/rows 59-86) and an east-west stripe along the western African coast (WRS2 paths 191-204/rows 51-57). It took a Windows XP desktop PC about 3 weeks to analyze a total of 105 pairs of Landsat images. A mosaic of the derived change maps for WRS2 path 225/rows 62-68 is shown to the right.

Summary

A TDA-SVM method has been developed to automate forest cover change analysis using Landsat data. This method has been tested in 19 areas selected from all major forest biomes across the globe and produced very accurate results.

With a demonstrated capability to analyze 105 pair of images on a desktop PC within just 3 weeks, this method will greatly increase the feasibility of global forest cover change analysis using Landsat data.

WRS2 path 225

Row 62
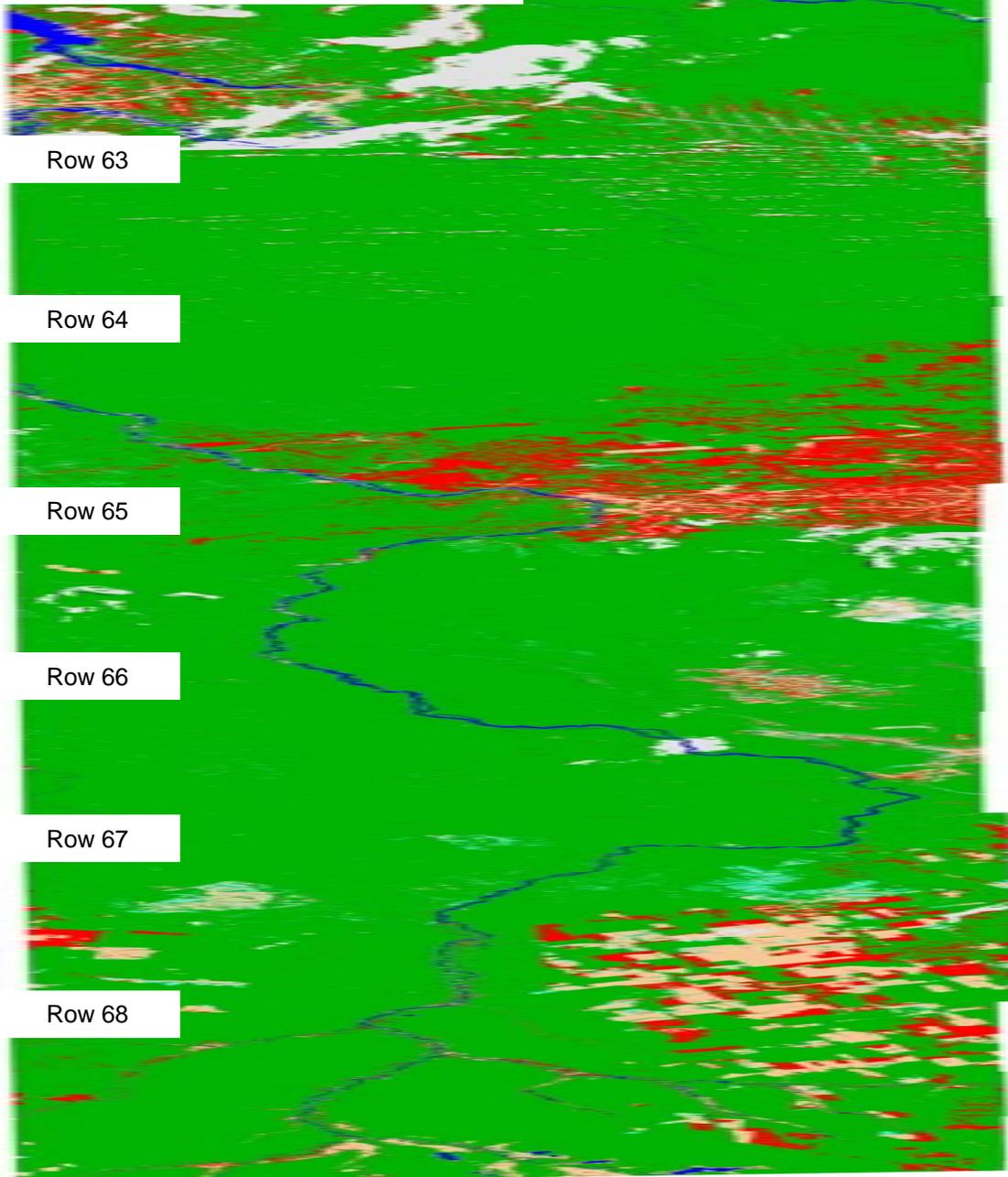
Row 63

Row 64

Row 65

Row 66

Row 67

Row 68

0   10   20   30   40   Kilometers
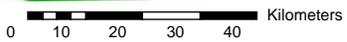
Legend

water
shadow
cloud

Forest gain
Persisting forest
Forest loss

Persisting non-forest

Acknowledgement