# Tools to tap the Landsat archive for monitoring and validation

Robert E. Kennedy[1], Warren B. Cohen[2],

Zhiqiang Yang[1], Peder Nelson[1], Eric Pfaff[1]

*[1] Department of Forest Ecosystems and Society, Oregon State University*
*[2] USDA Forest Service Pacific Northwest Research Station*

*Landsat Science Team Meeting, January 19-21 2010*

# Landsat: A fly on the wall

- Landsat has witnessed the transformation of the Earth's surface
- Understanding the changes that have already occurred allows us to:
  - Test mechanisms for drivers of change
  - Build models to capture those mechanisms
  - Manage for what comes next
- Landsat's durability is a testament to the people in this room who have persevered and made it happen -- THANK YOU

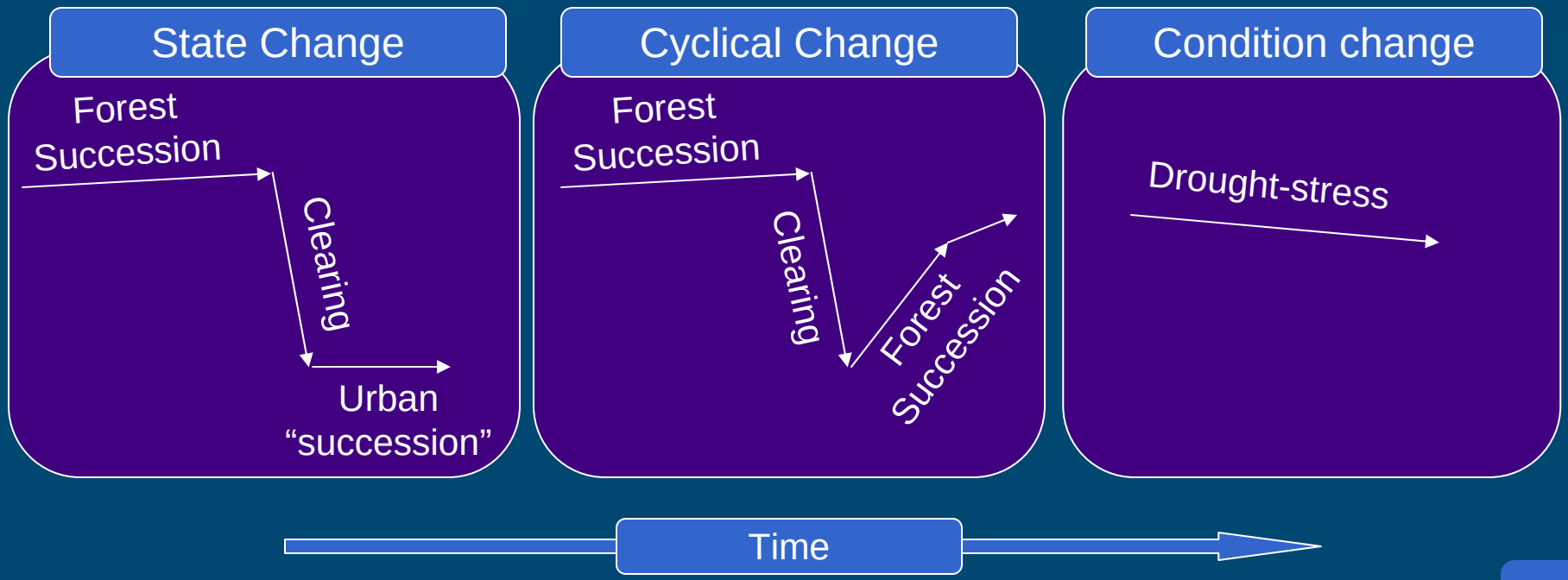
# Which changes do we care about?

- State change
  - Durable conversion from one landcover type to another: urbanization, deforestation, woody encroachment, vegetative community type change
- Cyclical change
  - Temporary conversion that eventually reverts to a prior condition: Disturbance, phenological change
- Condition change
  - Change in "vigor" or structure within a type: insect-related mortality, loss of productivity, increase in density

- *Land cover change means many different things*

# Our simplifying (?) view

- Change occurs *all the time*
  - Internal and external processes
  - Variable velocities and magnitudes
- Can be captured as vectors of change
- Critical events occur when direction changes



# Capturing change in the Landsat record

- Landsat data may have the potential to describe those vectors
  - Consistency, grain size, spectral properties, etc.
- Landscapes are spatial mosaics of these change dynamics, occurring simultaneously
  - Capturing only one type at a time may be inadequate

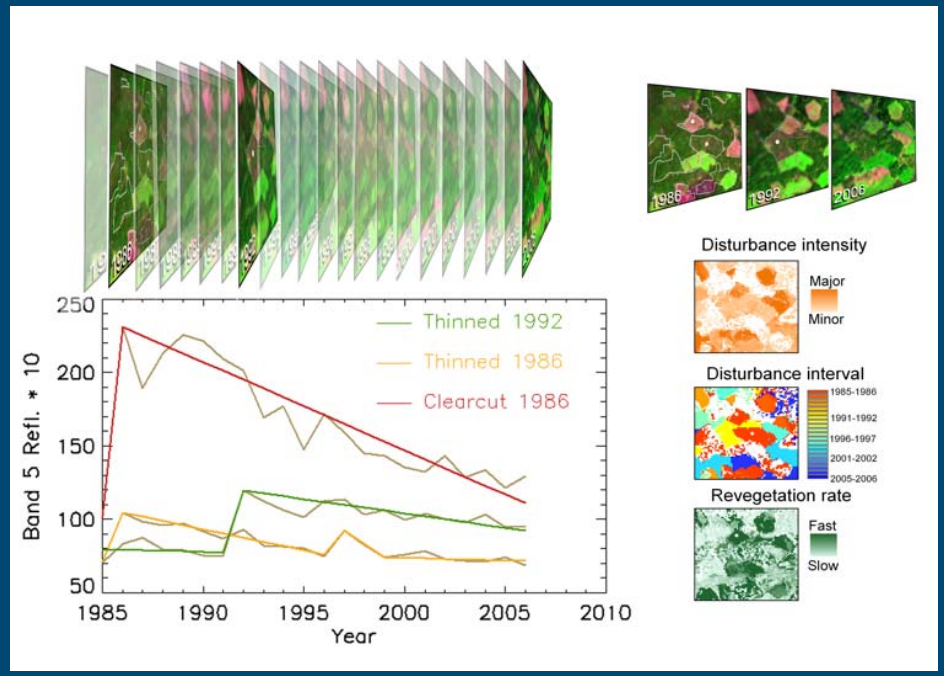
# Tools

- Our goal: Flexible tools to capture arbitrary change vectors
  - "Temporal segmentation"
- Automated, algorithm-based: LandTrendr
- Manual, interpreter-based: TimeSync

# LandTrendr: Landsat-based Detection of Trends in Disturbance and Recovery

Prepare stack of yearly imagery

Statistically identify and fit segments with consistent trends




Segment rules

Temporal-smoothing

Maps of change

Spectrally stable stack

Evaluation

Any spectrally-based operation: Classification, modeling, etc.

# Preparing stack: Pre-processing

- Geometric
  - Believe USGS
- Atmospheric
  - Current approach: COST, but any atmospheric correction base could be used
  - Relative normalization to a base is critical (MADCAL, Canty et al. approach)
- Cloud, shadow masking
  - Current approach: fairly dumb and labor intensive, but it works (uses cloud-free reference year and tasseled-cap differencing)
  - Any good approach could be used

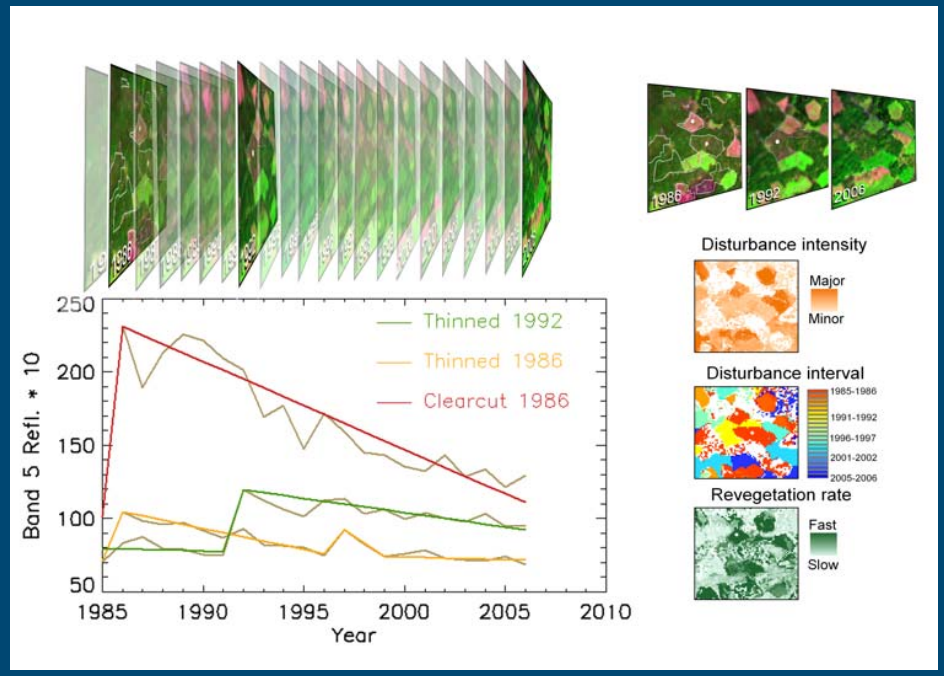
# Preparing stacks: On-the-fly mosaicking

- Stacks constructed from any number of images per year
  - In Pacific Northwest, regularly using 30 to 60 images to create 22-24 years of usable image stacks
  - Images prioritized based on proximity to median julian data of whole stack
  - *Consistency of date (phenological condition) trumps cloud-free imagery!*
    - "Give me three semi-cloudy images in July and August over one clean one in mid-September"
- SLC-off data?
  - Yes, give us as much as you want. We'll use every good pixel.

# LandTrendr: Landsat-based Detection of Trends in Disturbance and Recovery

Prepare stack of yearly imagery

Statistically identify and fit segments with consistent trends




Segment rules

Temporal-smoothing

Maps of change

Spectrally stable stack

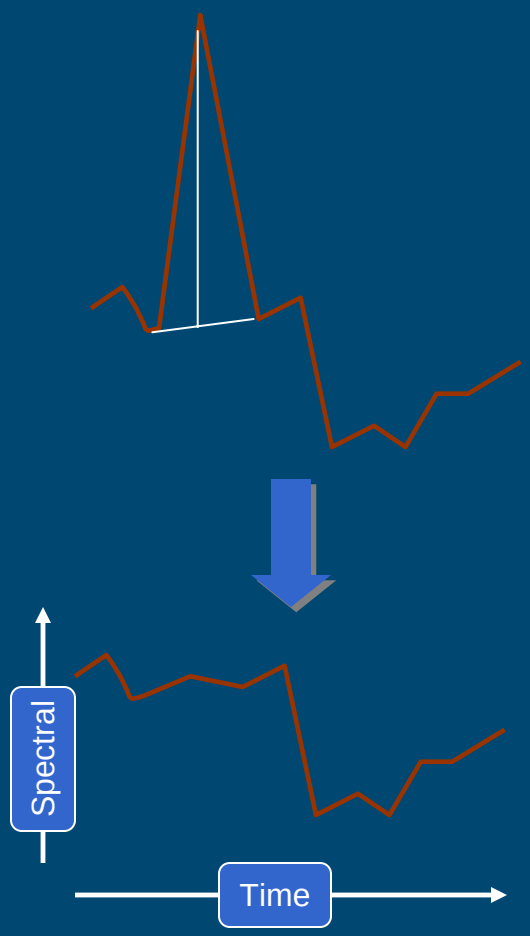
Evaluation

Any spectrally-based operation: Classification, modeling, etc.

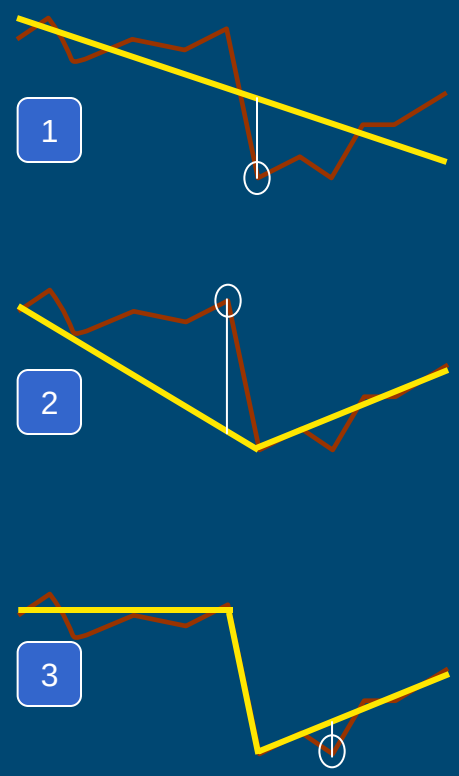
*Kennedy, Cohen et al. LST 2010*

# Identifying segments: Vertices

Eliminate spikes

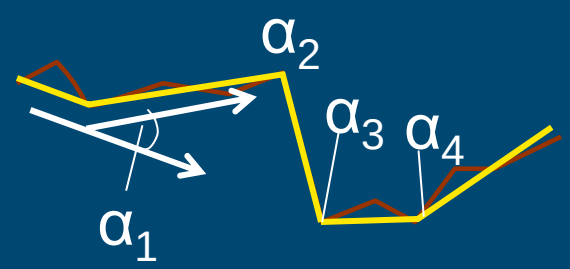



Identify potential vertices using regression deviation




etc Identify > *n*

Re-sort and cull back to *n* based on angle



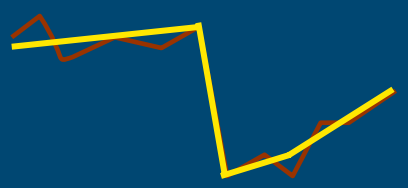

$\alpha_1 < \alpha_4 < \alpha_3 < \alpha_2$

Remove $\alpha_1$

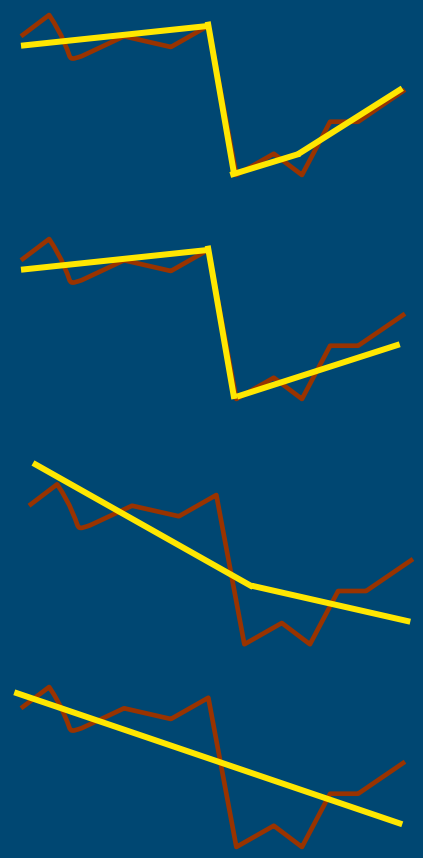
Result: Vertices defining maximum desired number of segments

# Identifying segments: Iterative fitting

Identify best path through all vertices using combination of regression or vertex-to-vertex connection



Iteratively cull vertices using segment-wise MSE



Calculate *p* of *f*-statistic (accounts for degrees of freedom but not temporal autocorrelation)

Pick best model using either lowest *p*-val, or allow more complex models win if nearly as good

Notes:

1. Regression works from left to right
2. Initial regression may be "free"
3. If left-to-right not effective, a second pass using full floating vertices is used

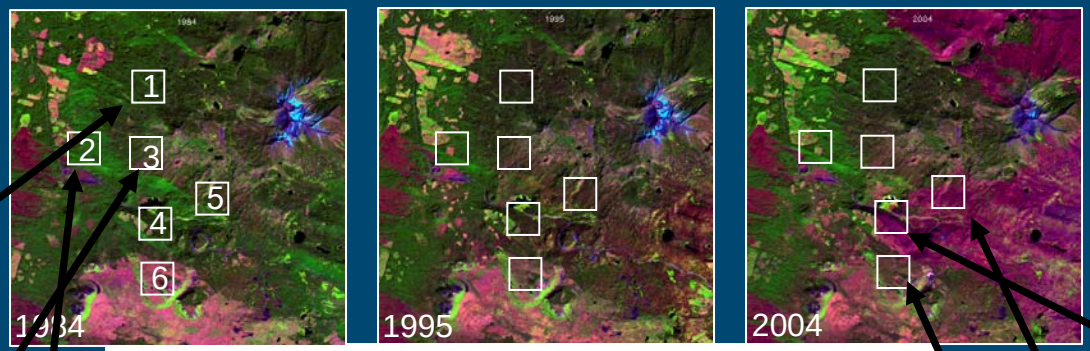
Results:

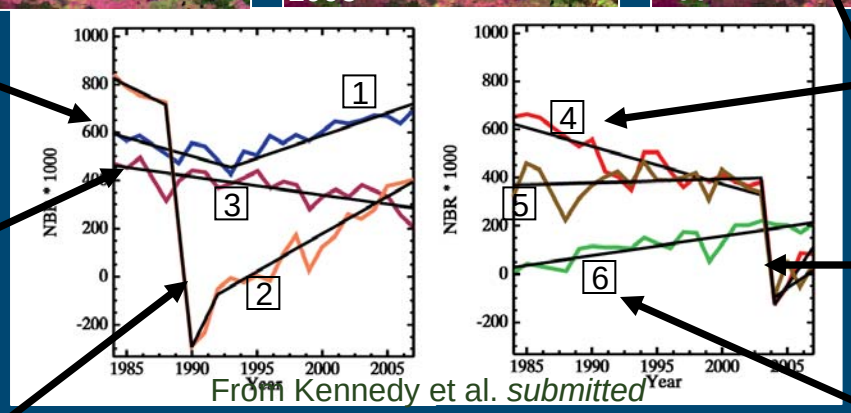
Vertices (x and y)

Fitted values of original index

Summary fitting statistics

# Examples of segmentation




From Kennedy et al. *submitted*

1: Insect, then regrowth

3: Insect-related mortality

2: Harvest, then regrowth

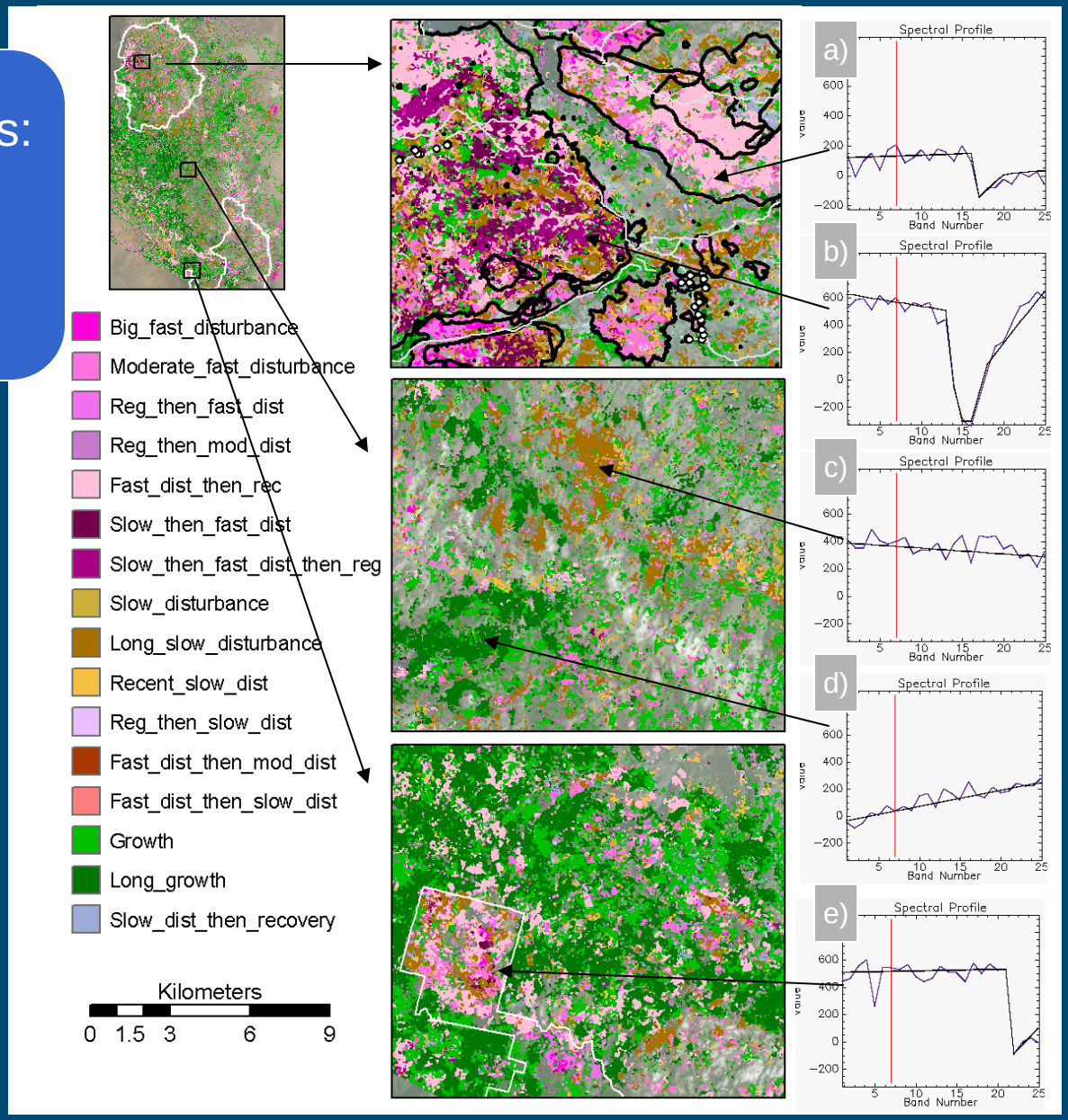
4: Insect, then fire

5: Growth, then fire

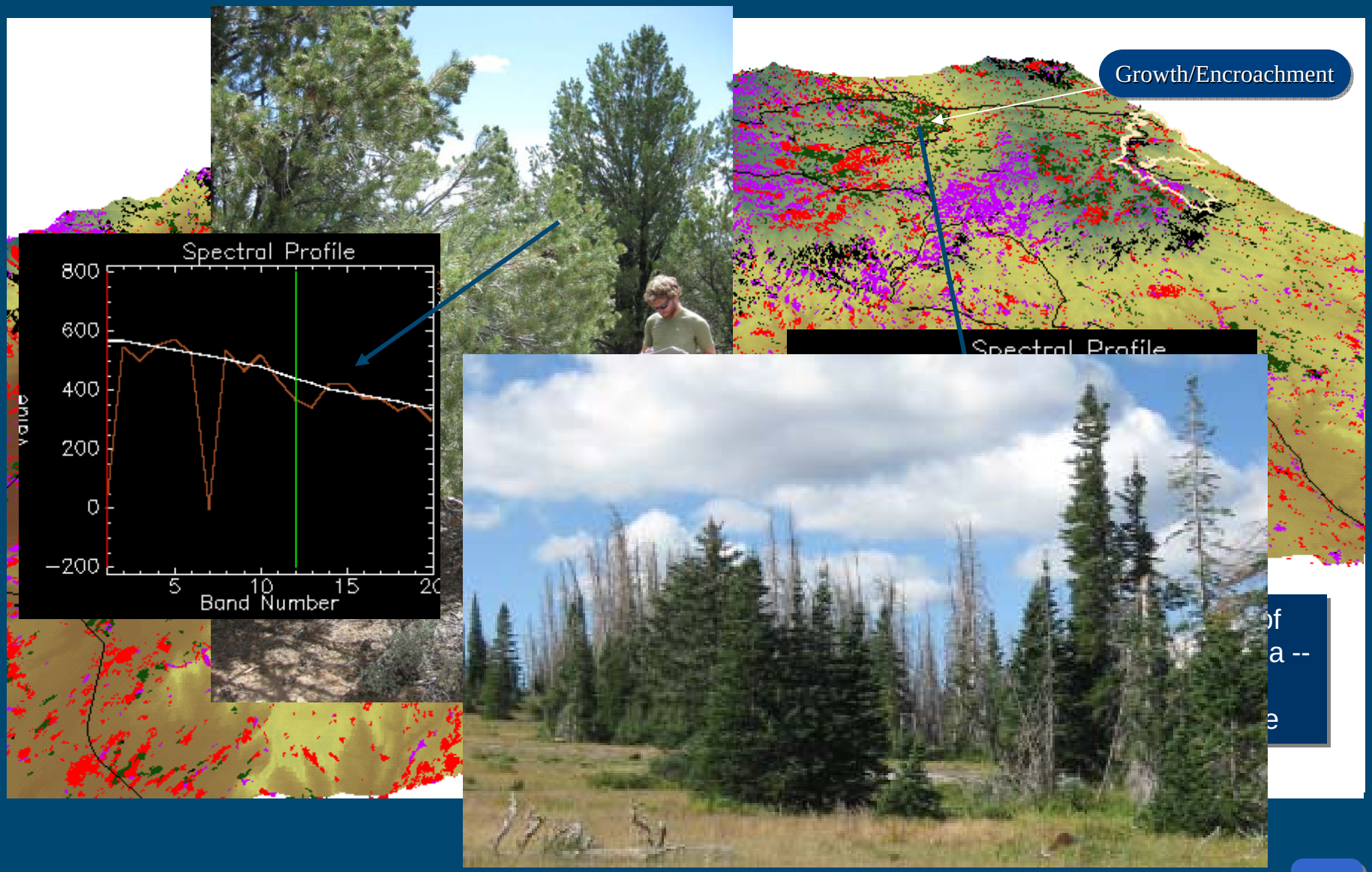
6: Growth

# Turning segments into maps

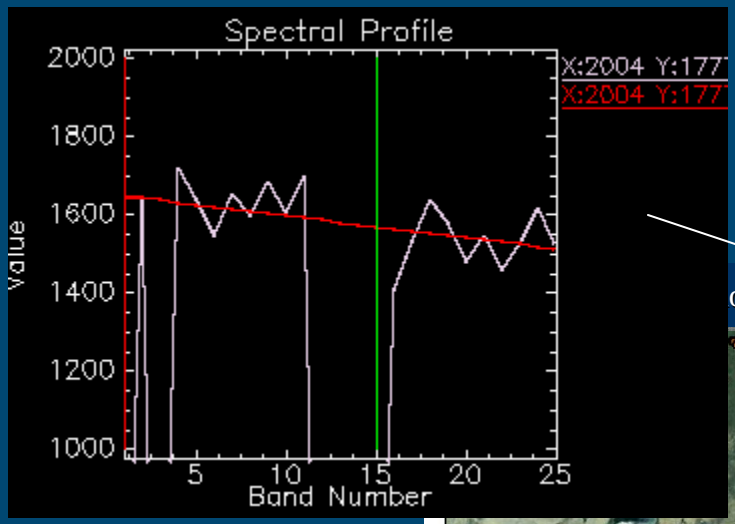
Change Labels: A map of the dynamic landscape

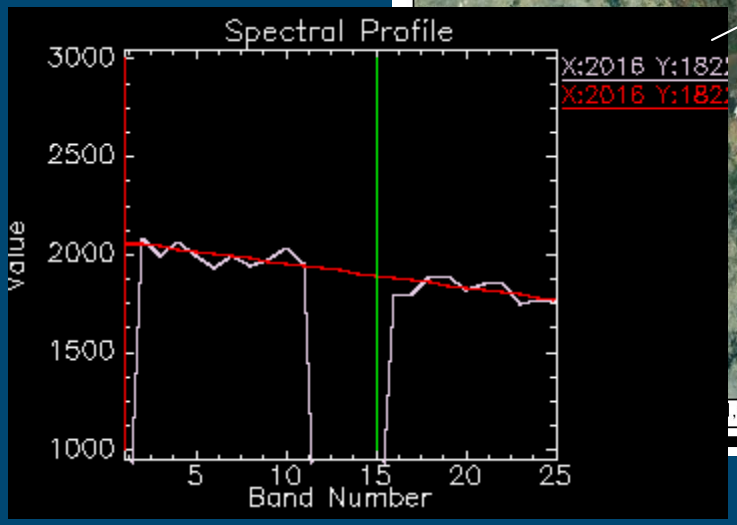


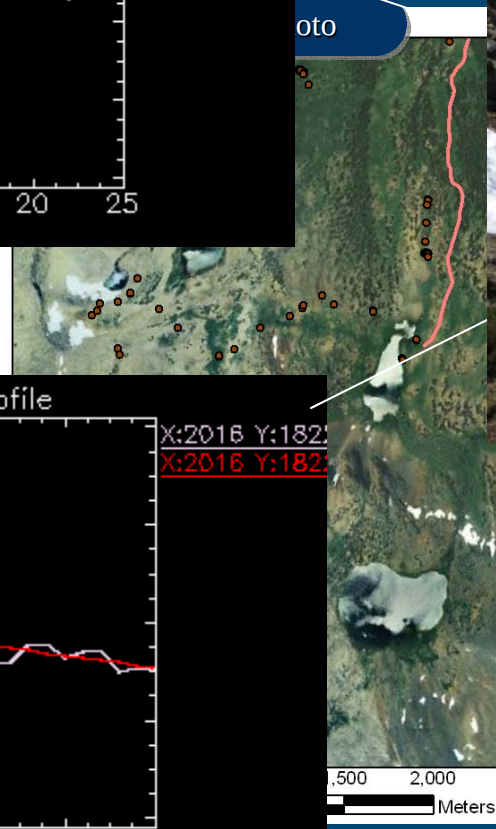
# Landscape dynamics on the Colorado Plateau



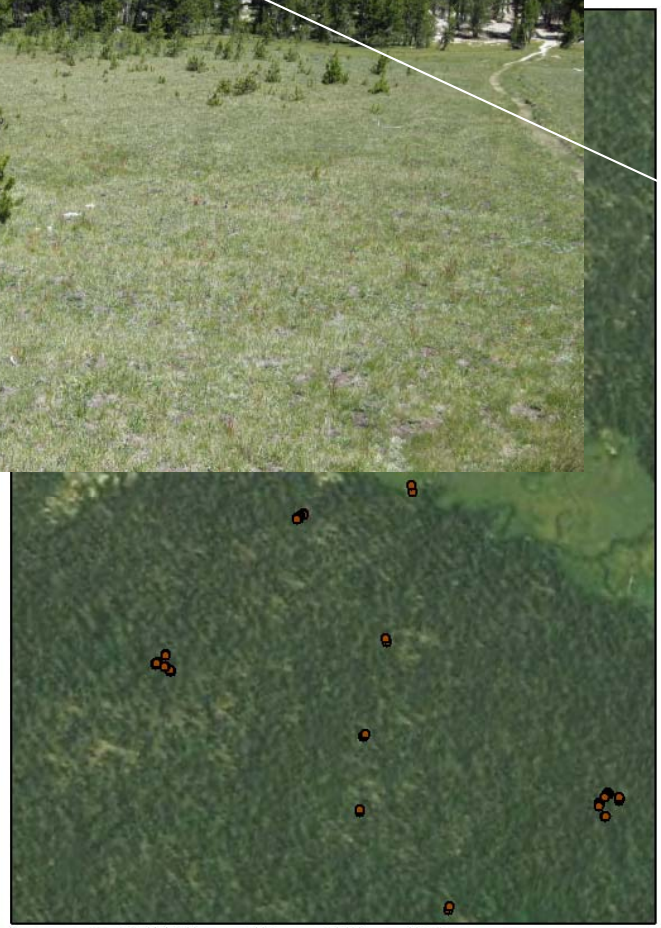
# Woody increase at treeline: Yosemite

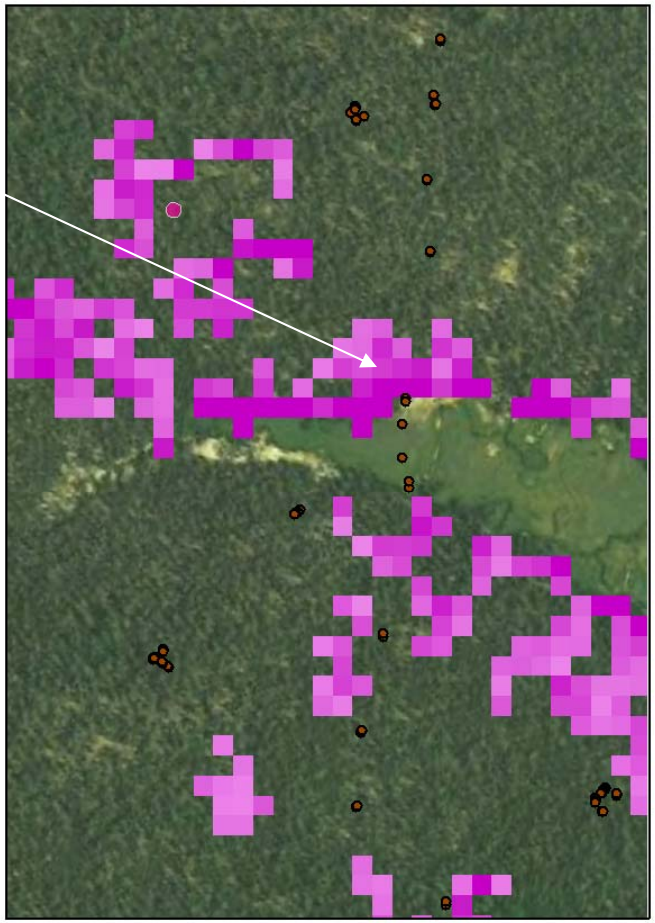


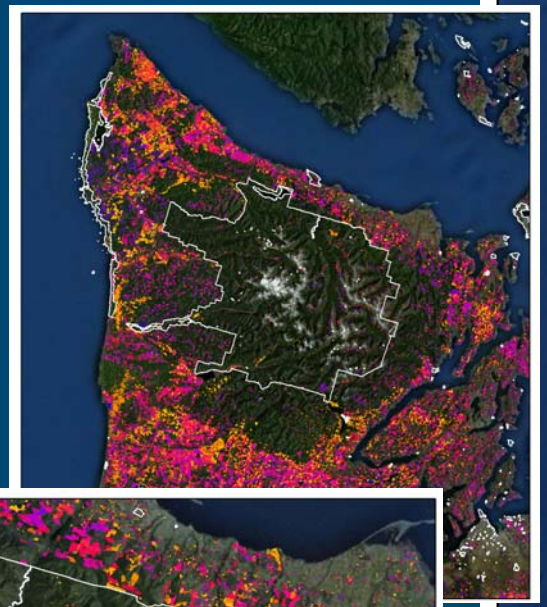



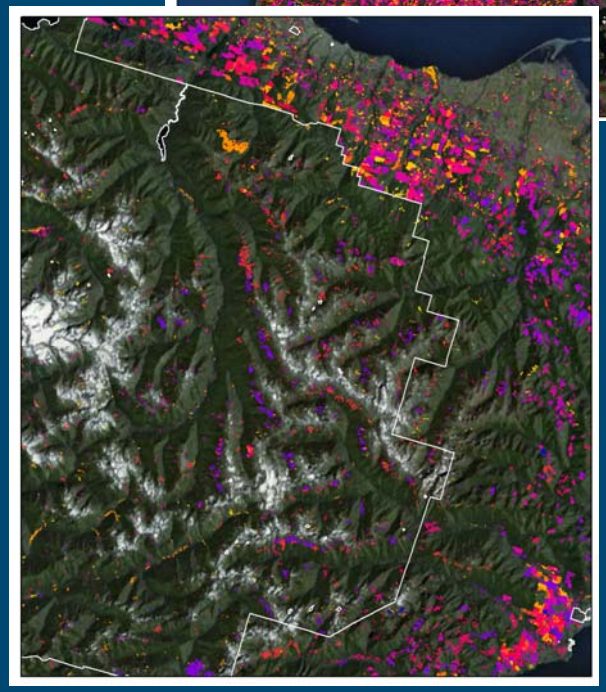
# Meadow encroachment: Yosemite

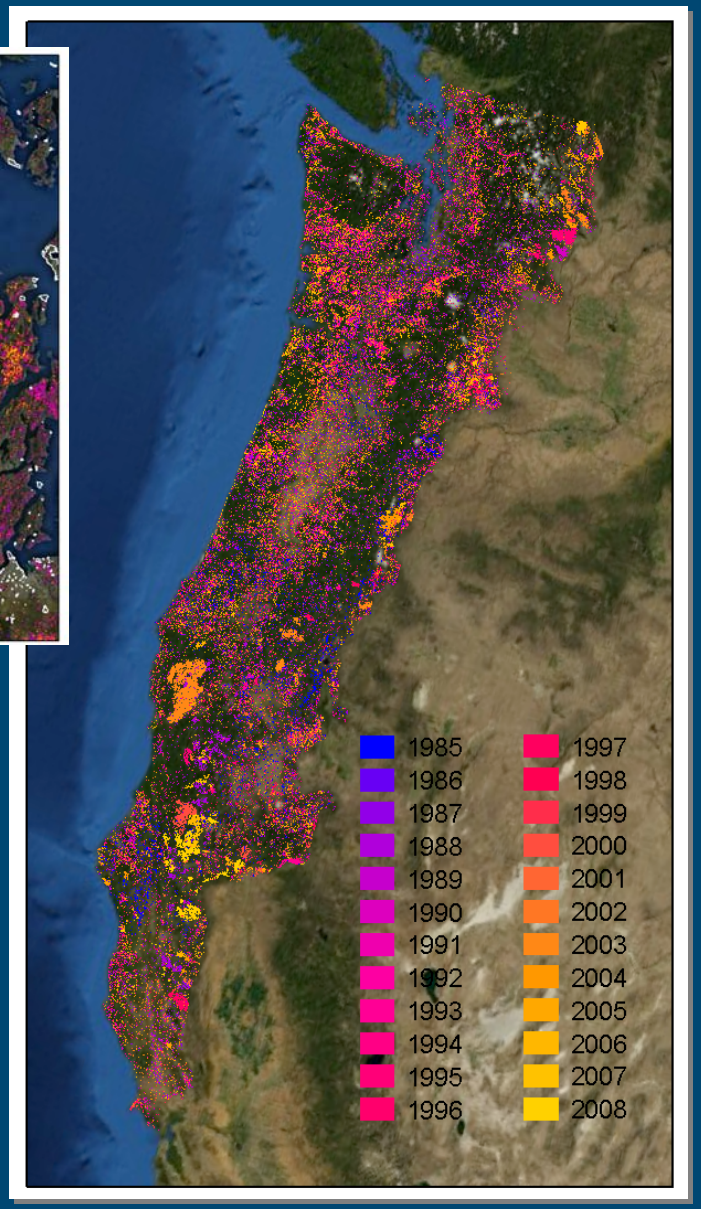


# Year of disturbance



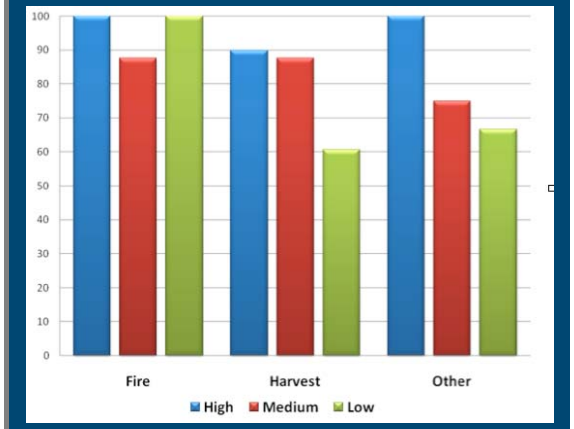





Project: Region 6 Effectiveness Monitoring Program for the Northwest Forest Plan (NWFP)

Data: > 500 individual Landsat scenes

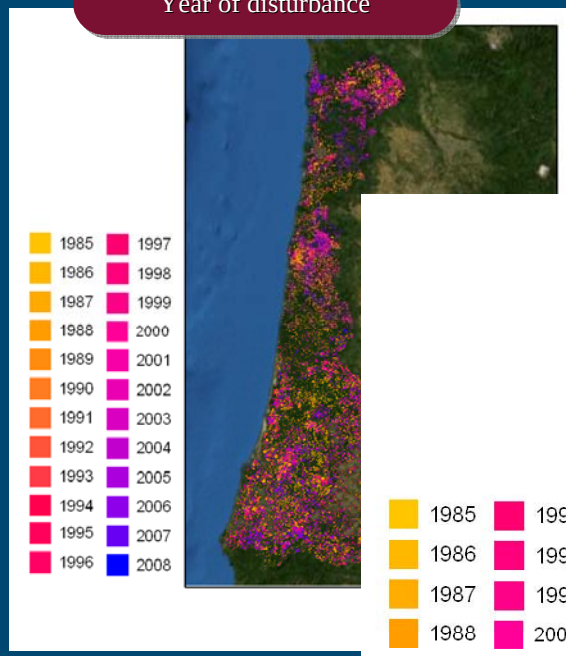
TimeSync Interpretation ongoing




*Kennedy, Cohen et al. LST 2010*

# Disturbance mapping: Time and magnitude

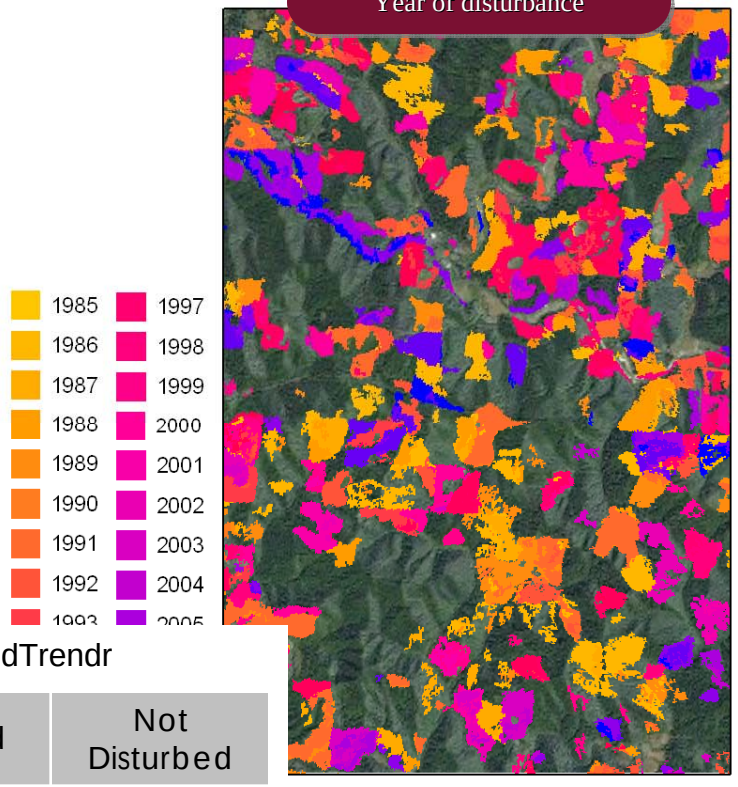
Year of disturbance



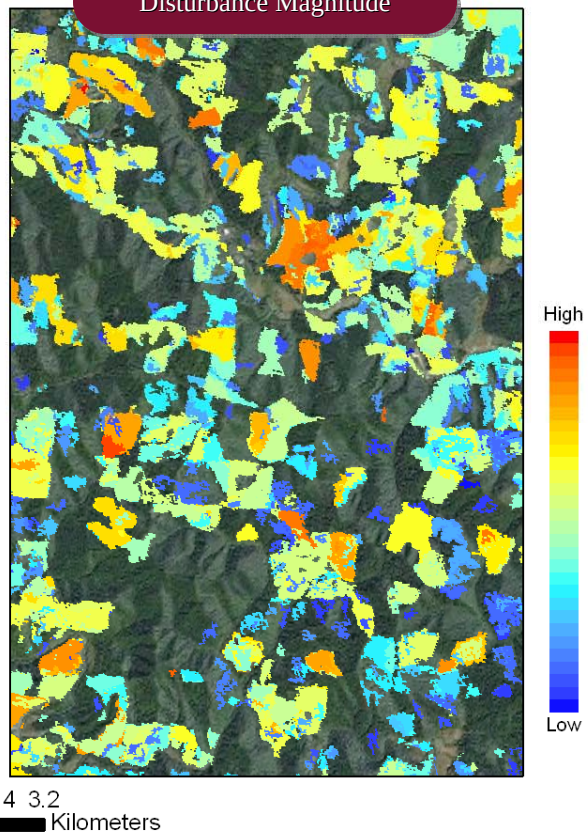

Disturbance Magnitude

Project: Quantify trends in harvest within Coastal Coho ecologically-significant unit

Year of disturbance




Disturbance Magnitude

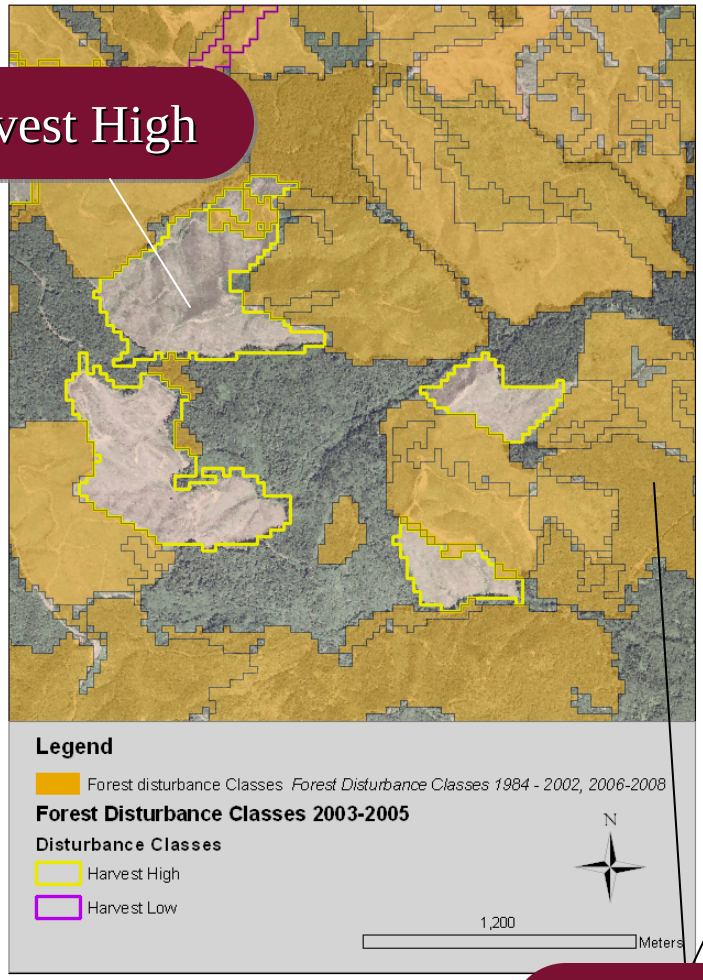



LandTrendr

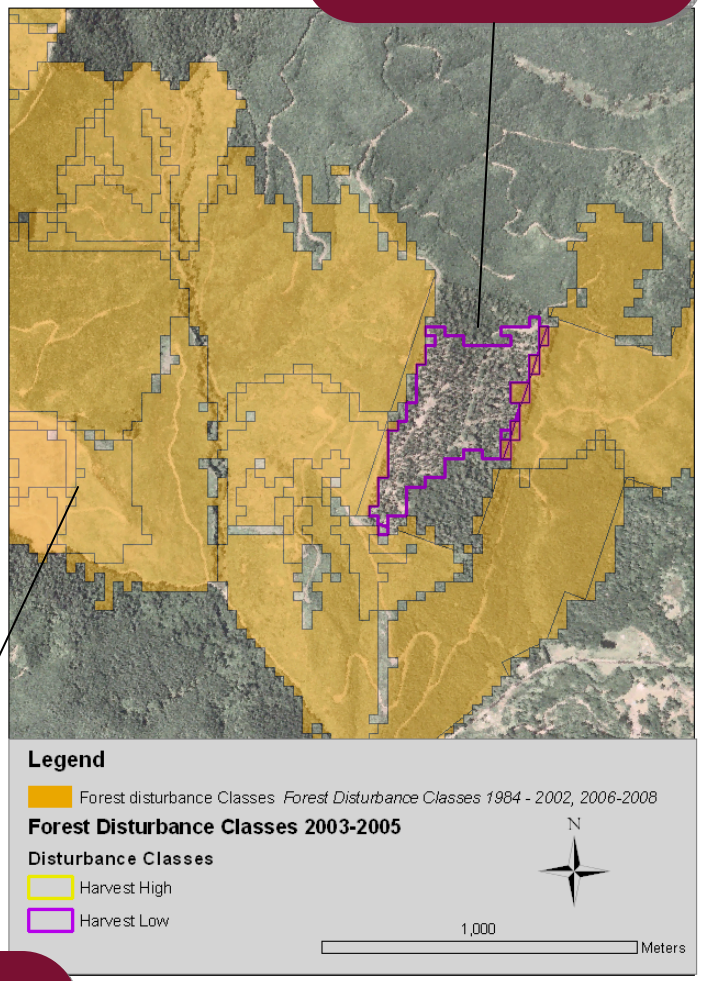
| TimeSync | Disturbed | Not Disturbed |
|---|---|---|
| Disturbed | 89 | 15 |
| Not Disturbed | 13 | 157 |

# Disturbance classes: Examples

Harvest High

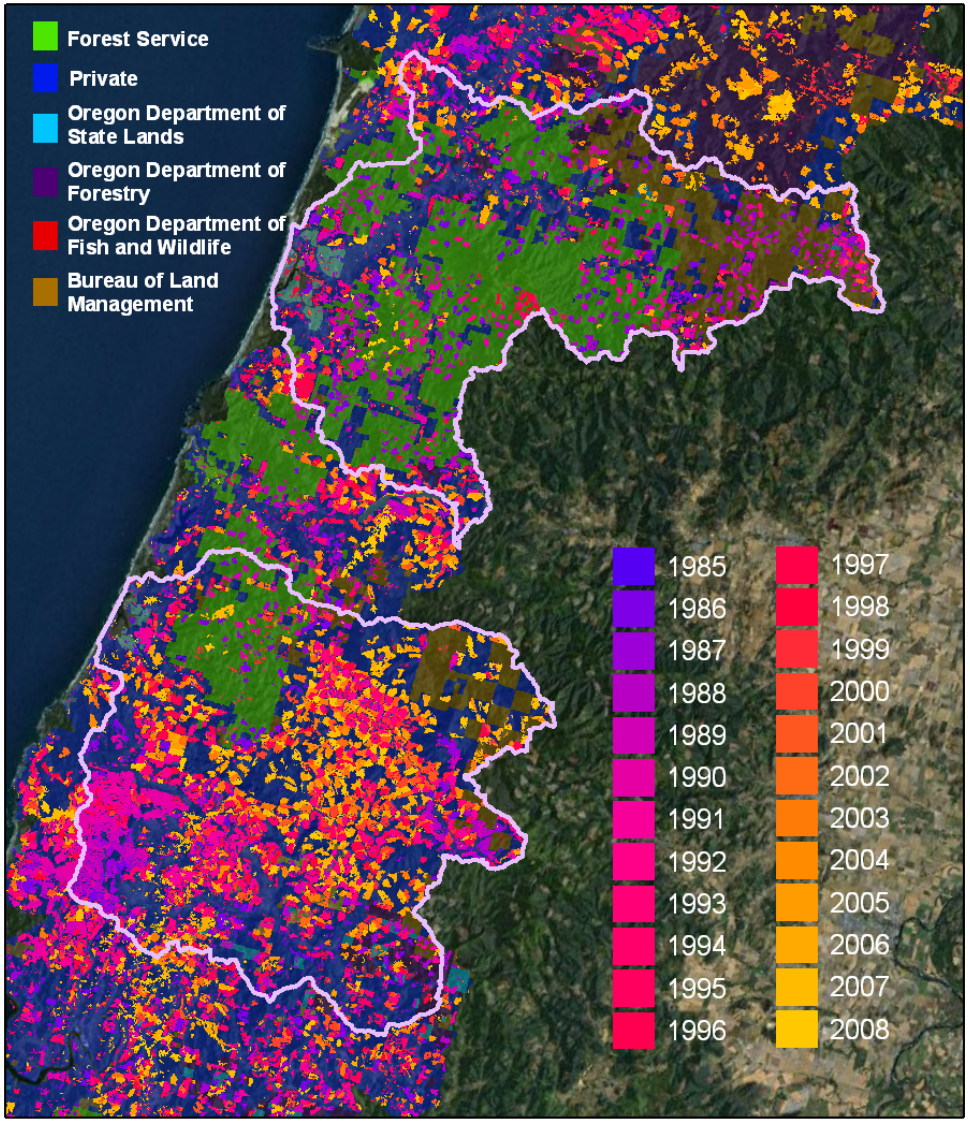


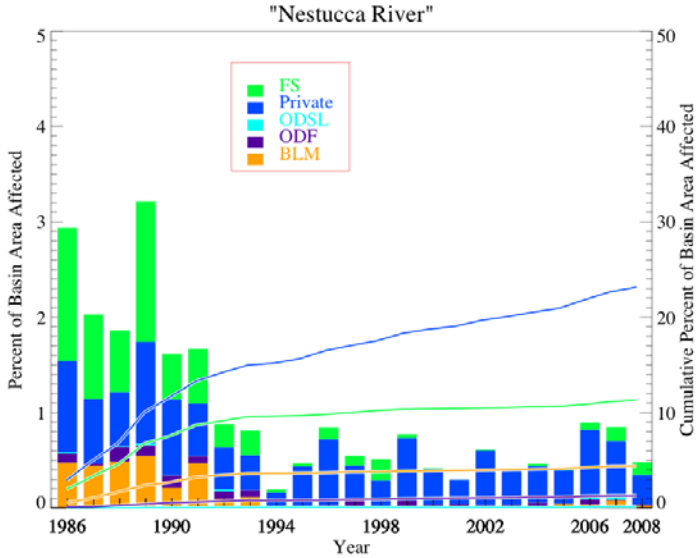

Harvest Low

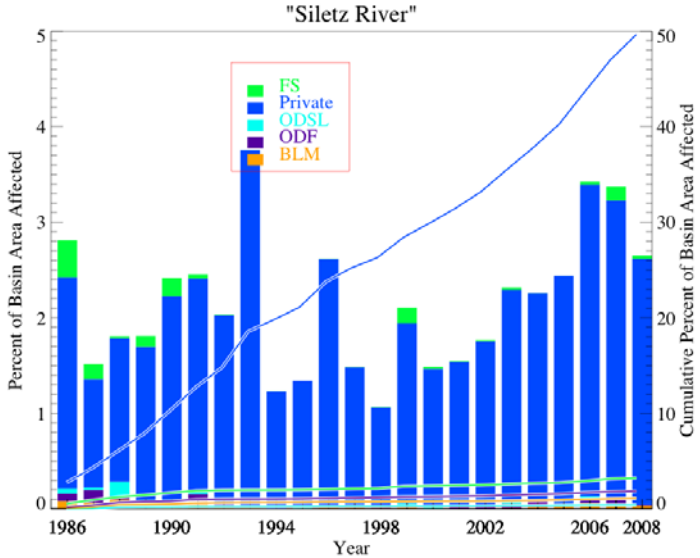



Prior Harvest

# Contrasting temporal patterns

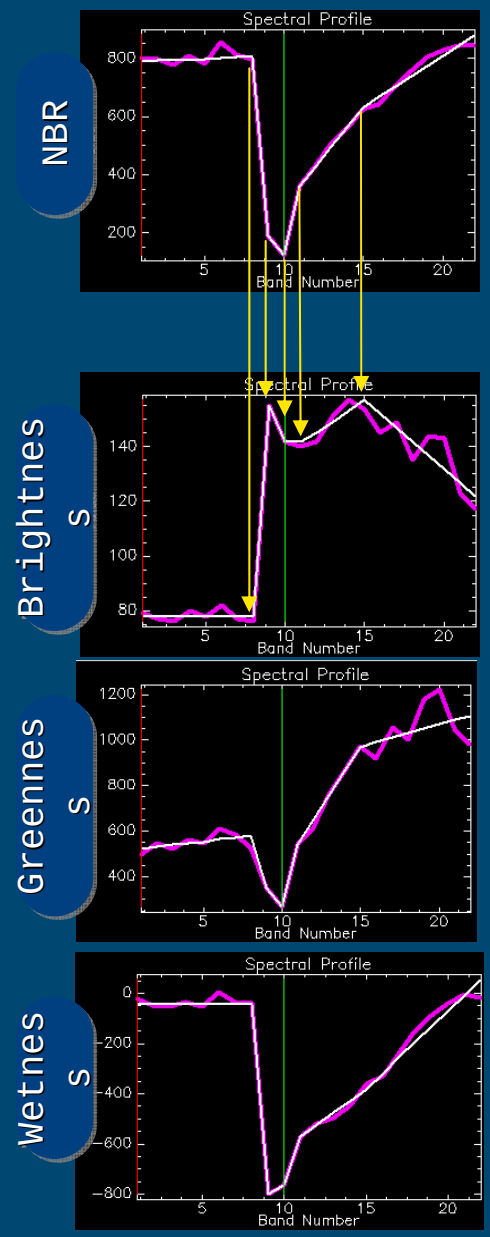


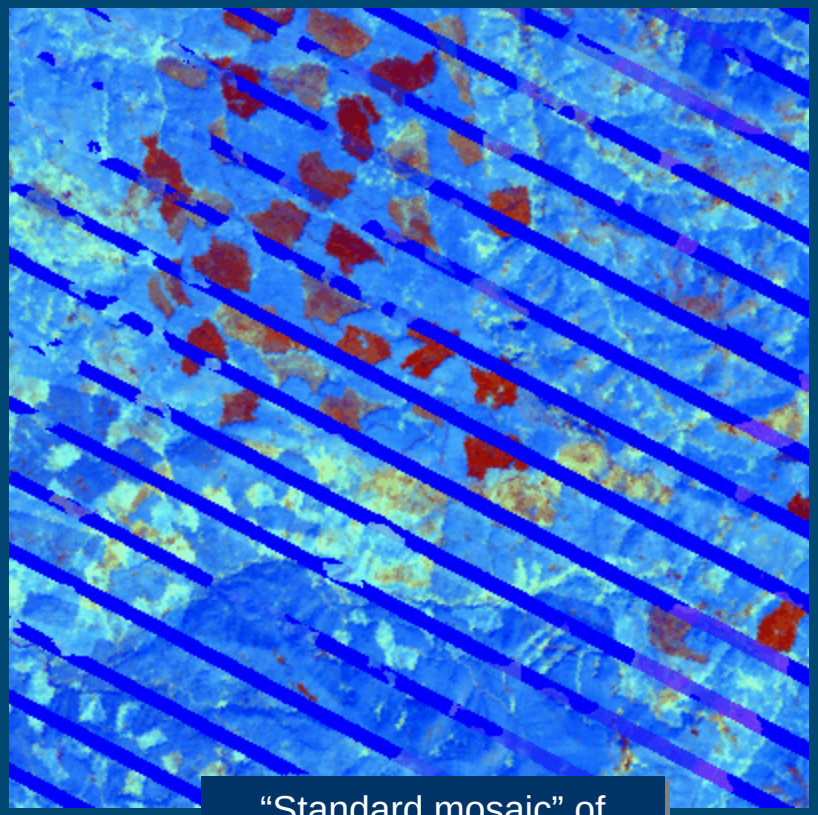



# LandTrendr: Temporal fitting

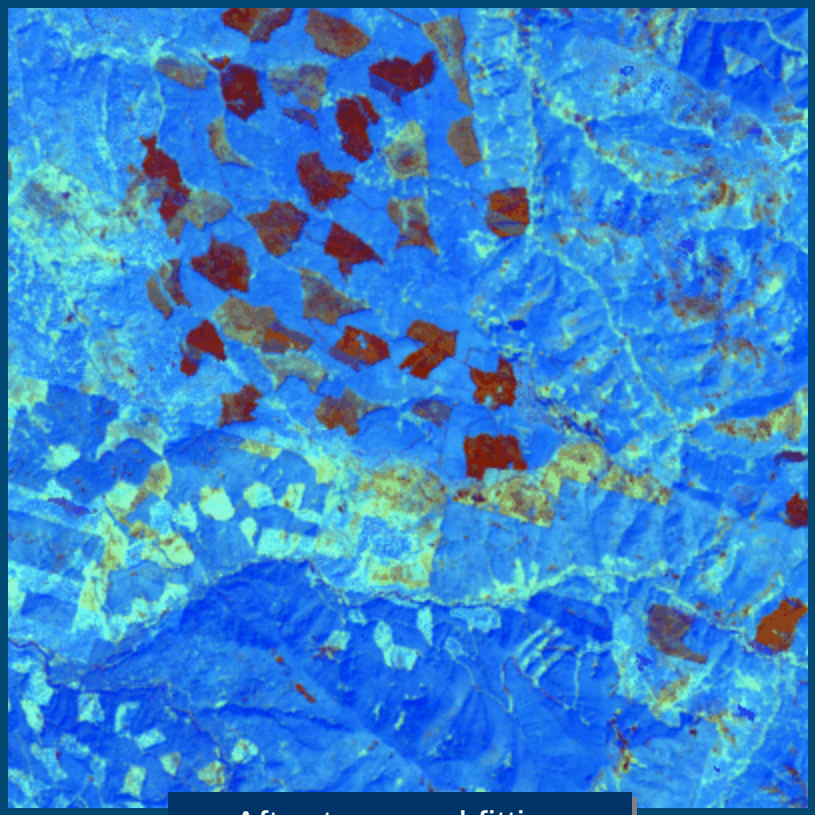
- Use segmentation of one band to identify "vertices" in time series
- Smooth between vertices in other bands
- Result: "Pseudo-images" with year-to-year noise removed, but actual change retained



# Temporally-fit imagery

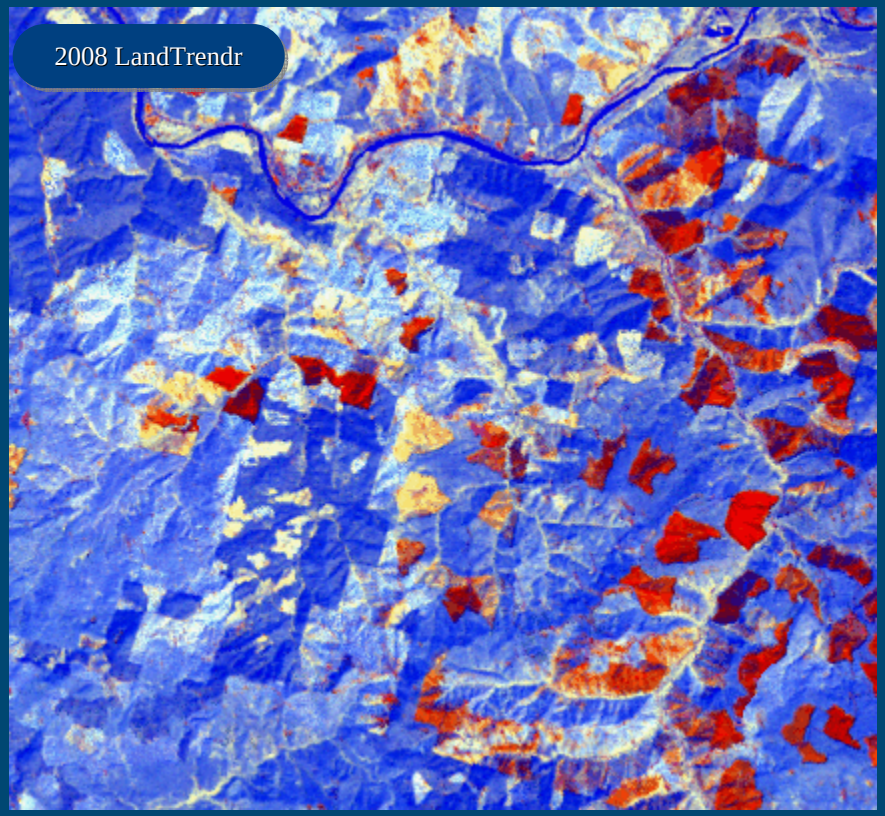


"Standard mosaic" of images within one year

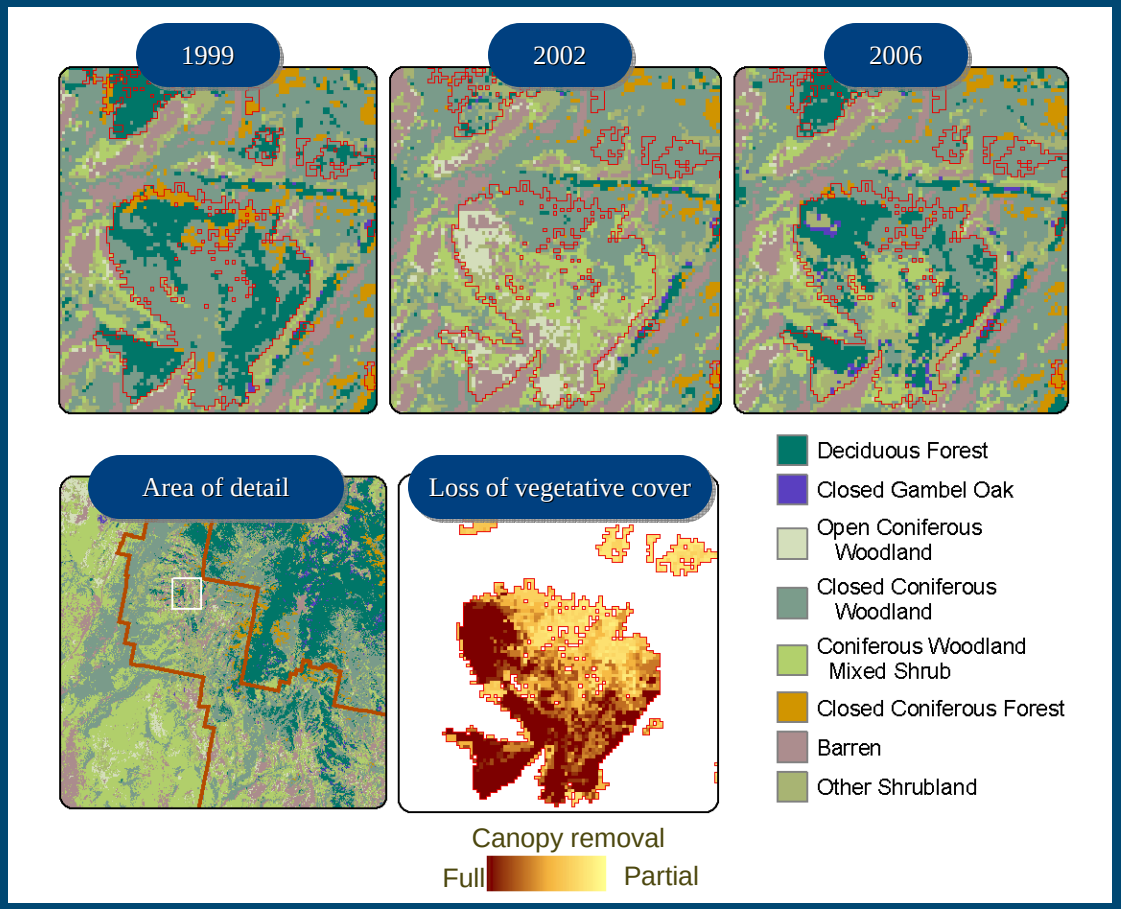


After temporal fitting

# LandTrendr: Temporal fitting



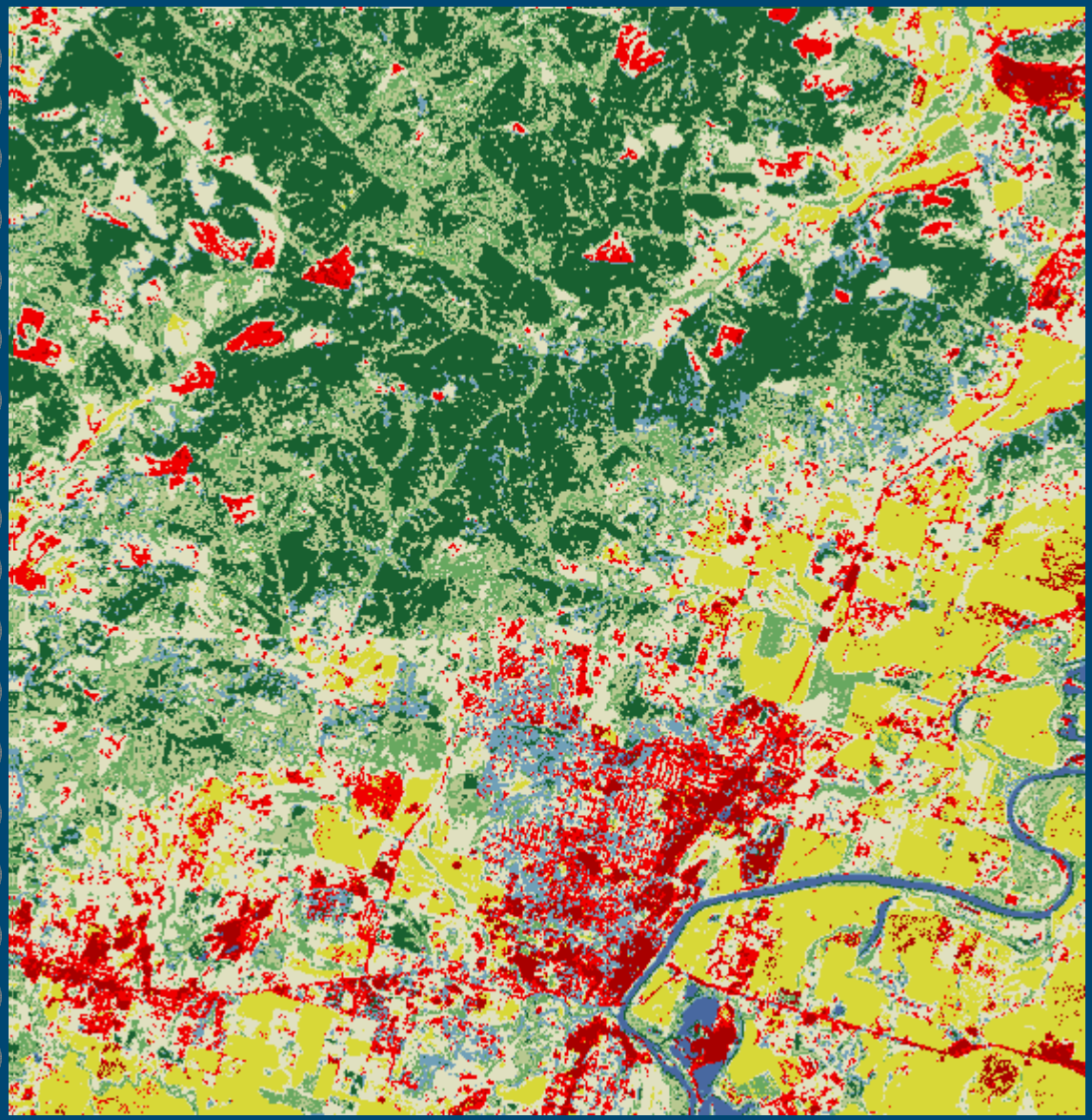
# Yearly classification: Fire effects




*Track fire effects using class labels familiar to users*

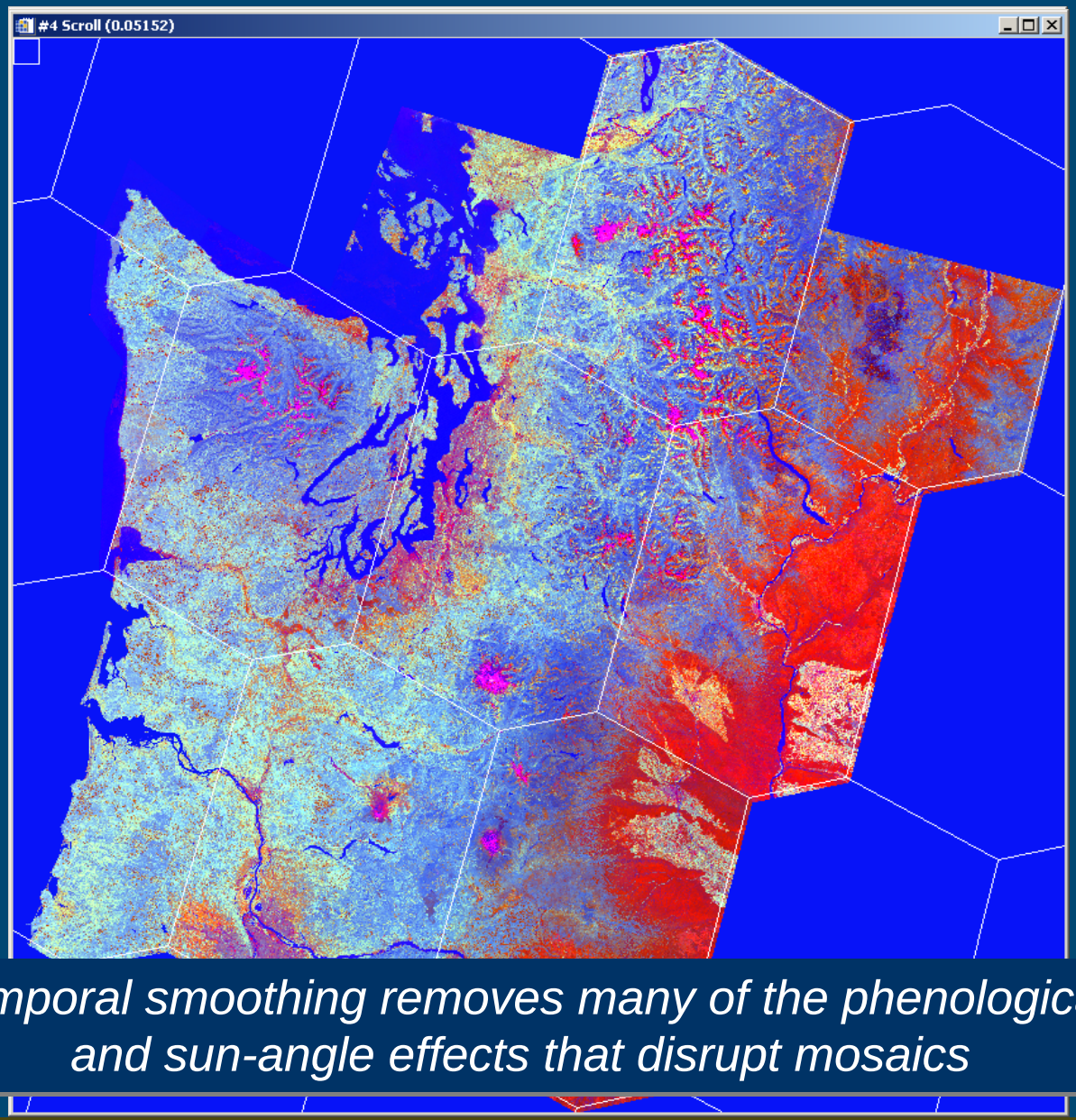
# Example: Application to NLCD mapping

1984
1986
1987
1988
1989
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003



2008
2007
2006
2005
2004

# Bonus: Mosaicking




*Temporal smoothing removes many of the phenological and sun-angle effects that disrupt mosaics*

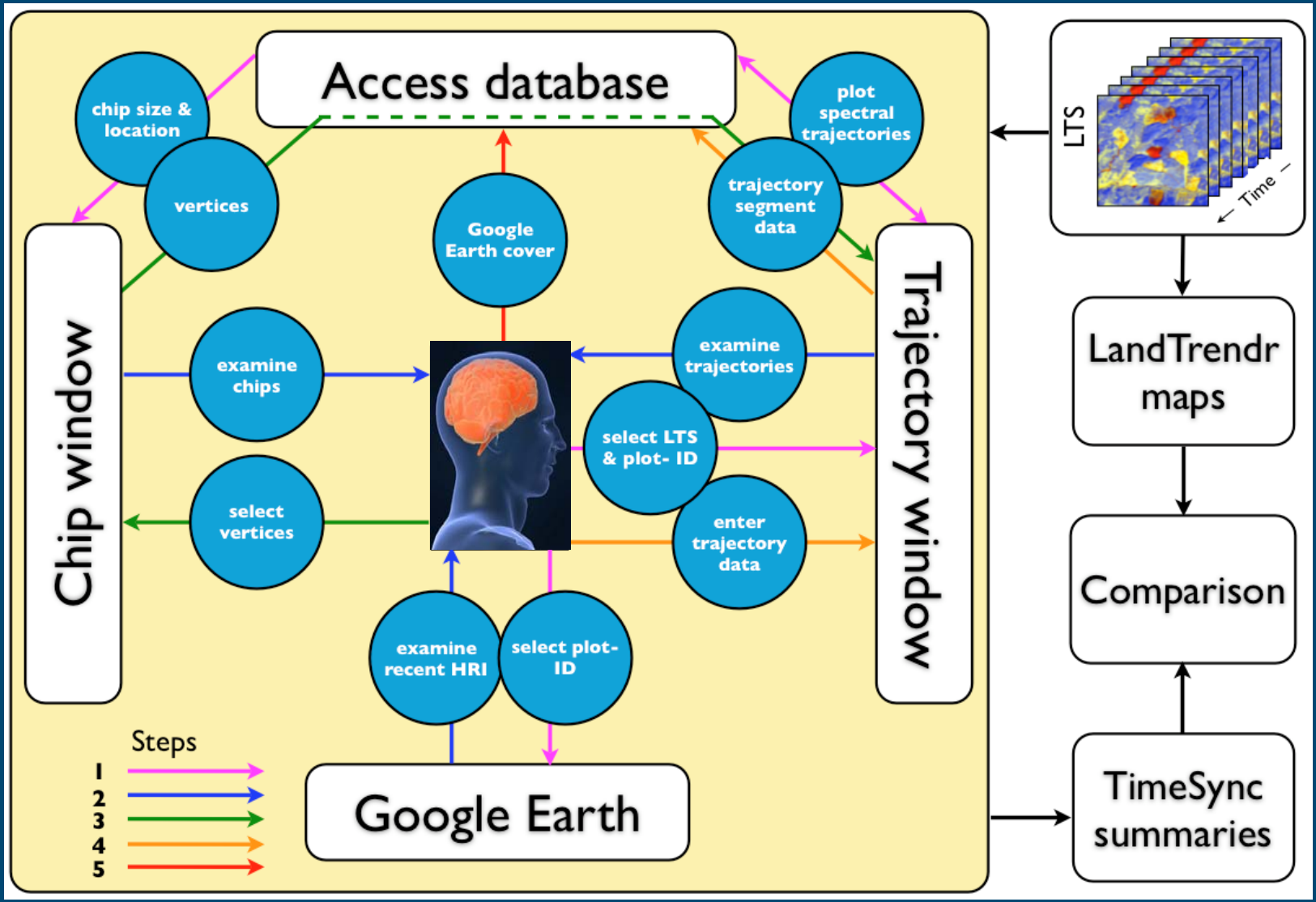
# How to validate? Corroborate?

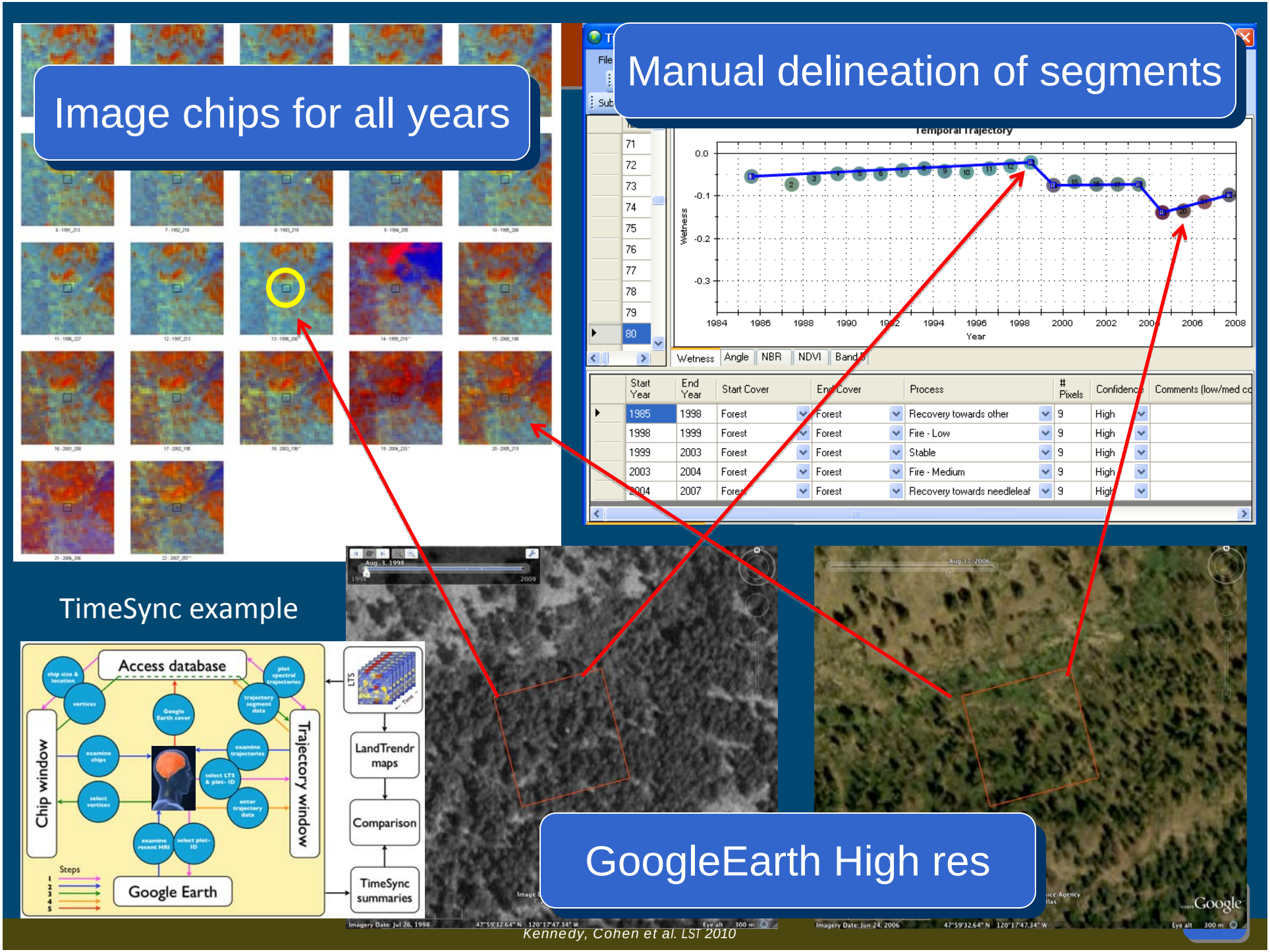
- LandTrendr, Vegetation Change Tracker (VCT; Huang 2009), CLASlite, STAARCH, etc. can create *yearly (sub-yearly) maps everywhere*
- EarthEngine will potentially allow anyone to create maps from new algorithms
- How good are they?
- *As algorithms and processing become more prevalent and easier to implement, validation will become more and more critical*

# How to validate? Corroborate?

- For true validation of change:
  - Detailed information about change for every year, every where, distributed in a statistically sound manner, from 1985 to present
  - *Such datasets do not exist!*
- Our strategy:
  - Stage 1 validation: TimeSync
    - Use *imagery itself* to interpret dynamics, everywhere, all the time, in a statistically sound manner
  - Stage 2 validation: Opportunistic
    - Use any other reference dataset to validate TimeSync for a subset in time and space

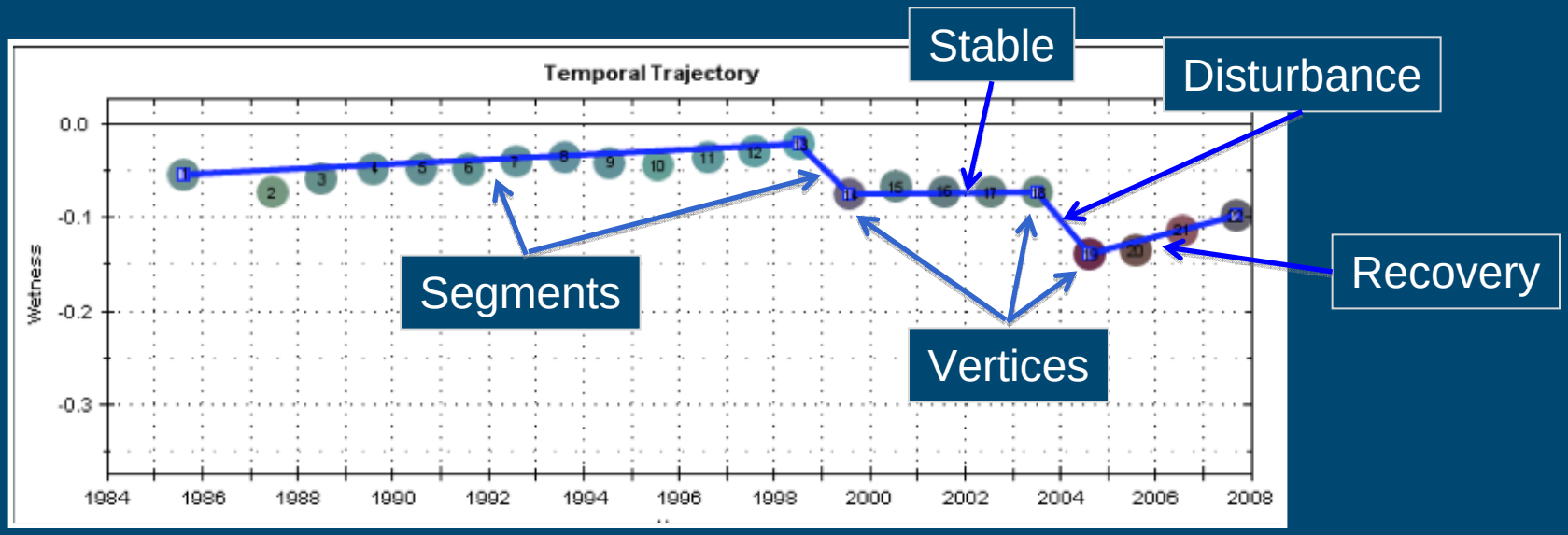
# Conceptual basis of TimeSync





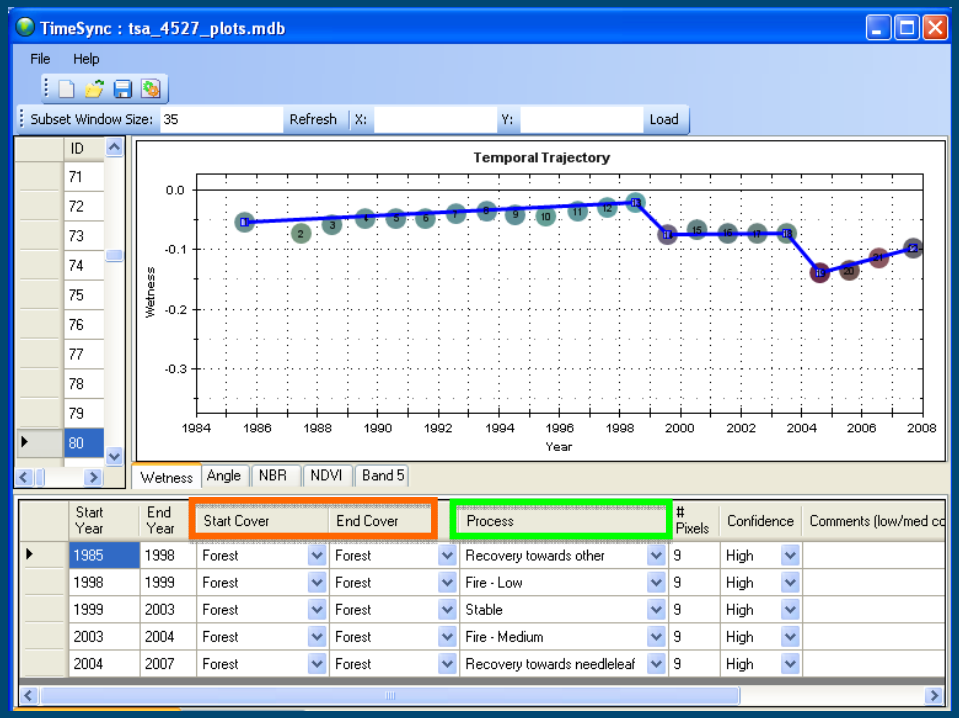

Kennedy, Cohen et al. LST 2010

# Segment-based change labeling




- Linear segments with start and end vertex dates, and labels (stable, disturbance, recovery)

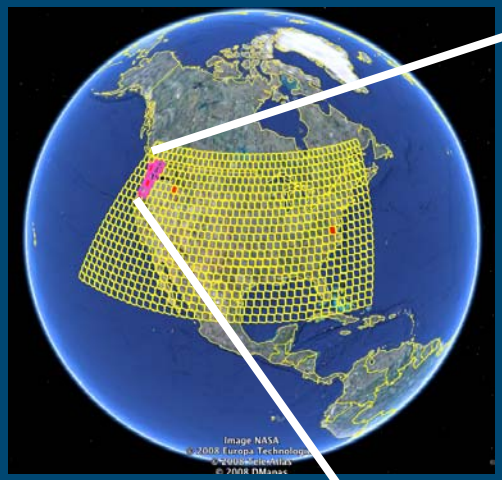
# TimeSync: Database population

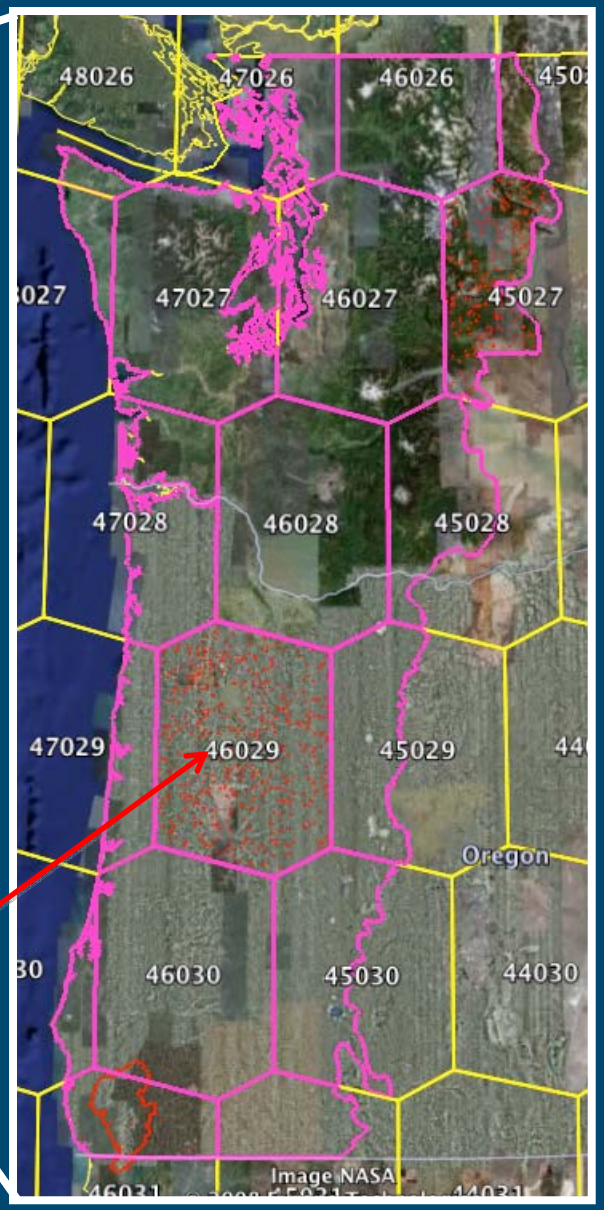



| *Vertices* | *Segments* |
|---|---|
| **Land Use/ Cover Class** | **Process*** |
| Forest | Stable |
| Needleleaf forest | Recovery |
| Broadleaf forest | Recovery towards needleleaf |
| Mixed forest | Recovery towards broadleaf |
| Woodland | Woody encroachment |
| Shrubland | Harvest |
| Grassland | Fire |
| Agriculture | Pathogen |
| Wetland | Blowdown |
| Urban | Other |
| Snow/ice | |
| Barren | |
| Water | |
| Other | |

- Any band or index
- # pixels affected
- Confidence and comment

# Allows robust sample design






- Any statistical design possible
- Design used for LandTrendr & the NWFP:
  - Stratification by TSA
  - Randomly locate ~600 plots
  - Interpret ~ 200 per full TSA

# Example: Comparing LandTrendr and TimeSync

- Evaluate algorithm performance many ways

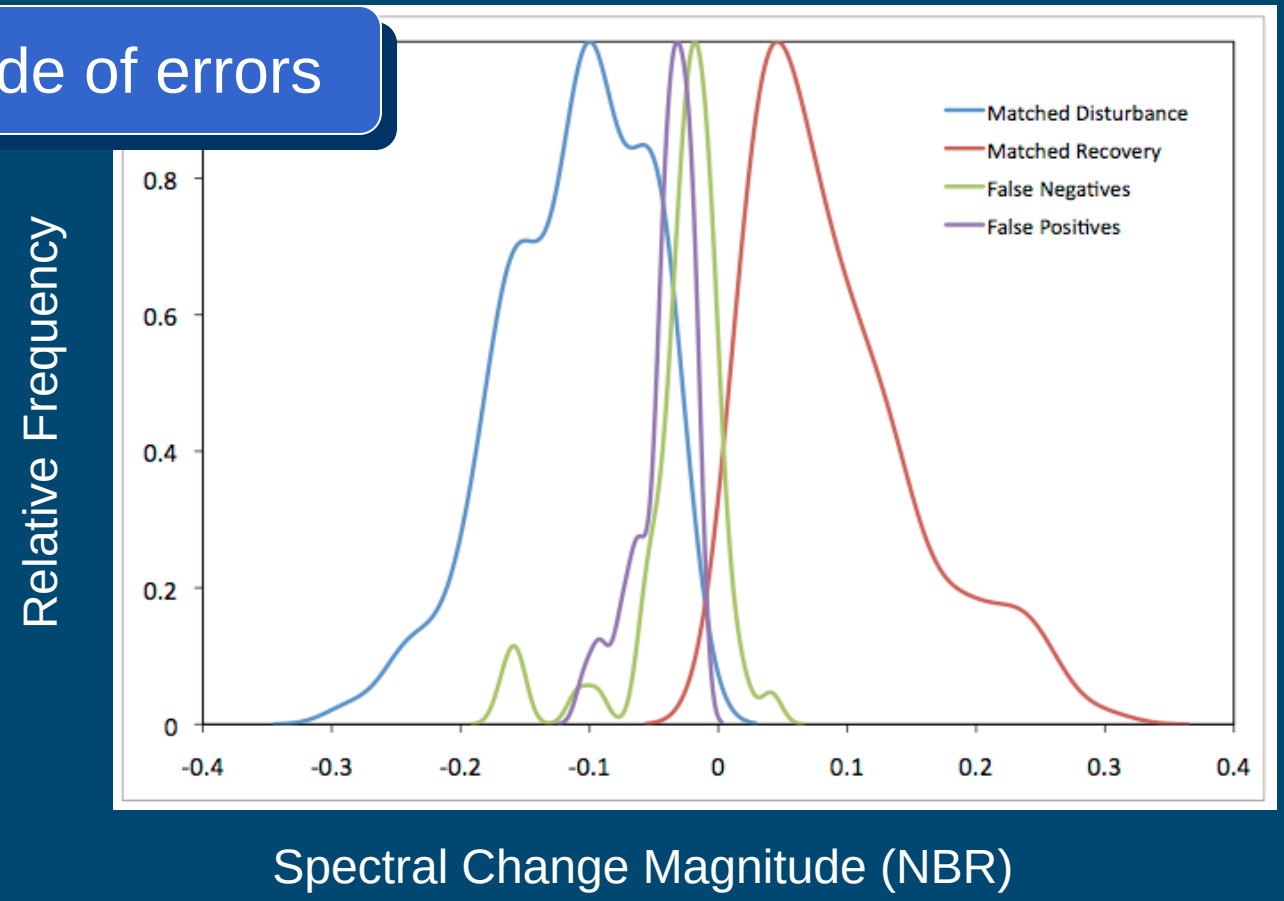
Vertex timing

**Algorithm (NBR)**

| TimeSync | Disturbance | Recovery | Stable | None | LT Omission | LT Omission by Type | LT Omission by Intensity |
|---|---|---|---|---|---|---|---|
| *Harvest - High* | 45 | 0 | 1 | 4 | **0.10** | | |
| *Harvest - Medium* | 28 | 0 | 2 | 8 | **0.26** | | |
| *Harvest - Low* | 14 | 3 | 4 | 17 | **0.63** | | |
| *Harvest - All* | 87 | 3 | 7 | 29 | | **0.31** | |
| *Fire - High* | 6 | 0 | 0 | 1 | **0.14** | | |
| *Fire - Medium* | 17 | 0 | 0 | 2 | **0.11** | | |
| *Fire - Low* | 5 | 0 | 0 | 0 | **0.00** | | |
| *Fire - All* | 28 | 0 | 0 | 3 | | **0.10** | |
| *Pathogen - Medium* | 4 | 0 | 0 | 2 | **0.33** | | |
| *Pathogen - Low* | 4 | 0 | 0 | 4 | **0.50** | | |
| *Pathogen - All* | 8 | 0 | 0 | 6 | | **0.43** | |
| *Other - High* | 2 | 0 | 0 | 0 | **0.00** | | |
| *Other - Low* | 1 | 0 | 0 | 2 | **0.67** | | |
| *Other - All* | 3 | 0 | 0 | 2 | | **0.40** | |
| *All - High* | 53 | 0 | 1 | 5 | | | **0.10** |
| *All - Medium* | 49 | 0 | 2 | 12 | | | **0.22** |
| *All- Low* | 24 | 3 | 4 | 23 | | | **0.56** |

# TimeSync: Characterize agreement & disagreement

## Magnitude of errors



# Stage 2: Validate the validation

- How well can Landsat + Interpreter brain capture real change?
- Comparison with other databases:
  - USDA Forest Service and USDI Bureau of Land Management activities databases
  - USDA Forest Service Forest Health Monitoring overflights
  - Monitoring trends in burn severity (MTBS) maps
- Overlay and compare calls

# Stage 2: Examples

## Comparison w/ MTBS

| | MTBS (Intensity) | | | | |
|---|---|---|---|---|---|
| **TimeSync** | *Low* | *Medium* | *High* | *Not Disturbed* | *Agreement* |
| *Harvest* | 1 | | | | n/a |
| *Fire* | 8 | 10 | 4 | 5 | 0.81 |
| *Not Disturbed* | | | | 360 | 1 |
| *Agreement* | 0.89 | 1 | 1 | 0.99 | 0.98 |

* Below MTBS size threshold; confirmed as fire by alternate dataset

## Comparison w/ agencies

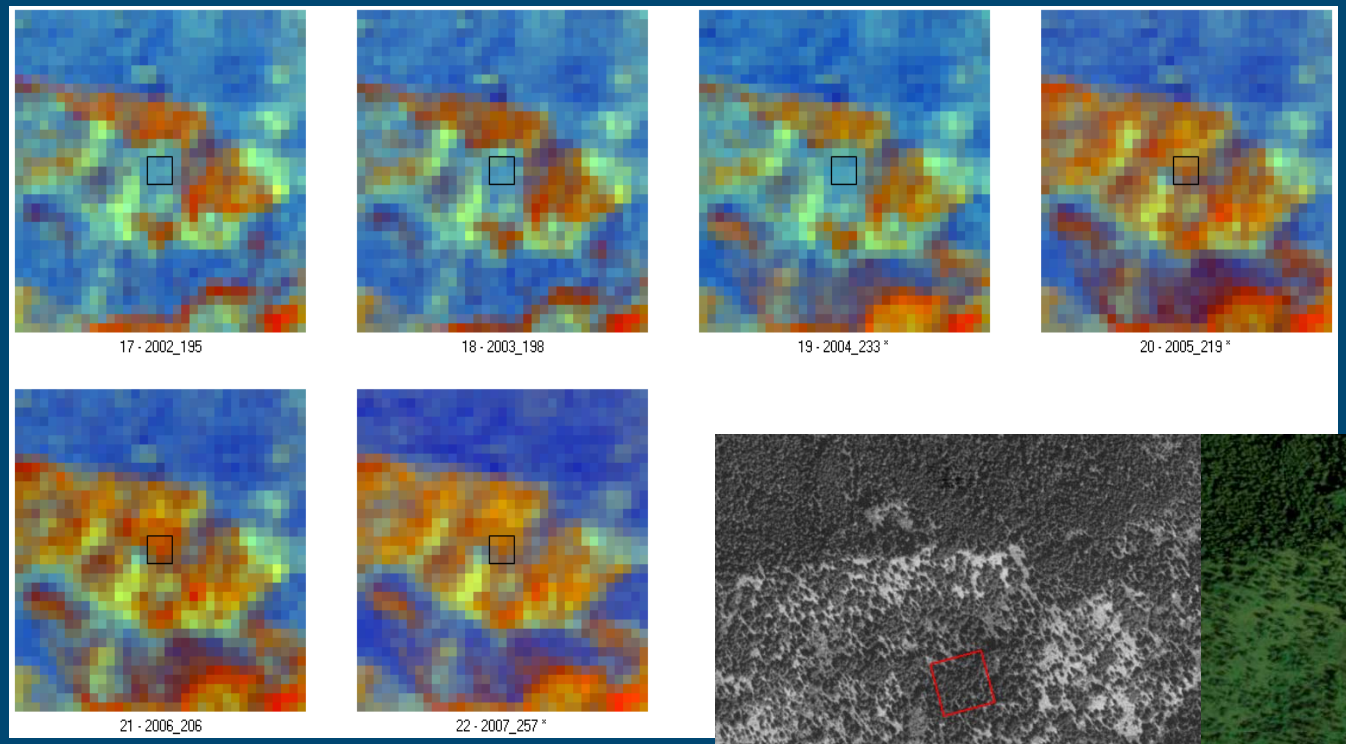
| | Integrated FACTS - FOI - FHM | | | | | | |
|---|---|---|---|---|---|---|---|
| **TimeSync** | *Harvest (commercial)* | *Harvest (pre-commercial)* | *Understory Burn* | *Insect* | *Road* | *Not Disturbed* | *Agreement* |
| *Harvest* | 8 | 2 | | | | 10 | 0.5 |
| *Insect* | | | | 7 | | 1 | 0.88 |
| *Road* | | | | | | 3 | 0 |
| *Not Disturbed* | 1 | 7 | 2 | | | 143 | 0.93 |
| *Agreement* | 0.89 | 0.22 | 0 | 1 | 0 | 0.91 | 0.87 |

TimeSync false negative

Agency database false negative

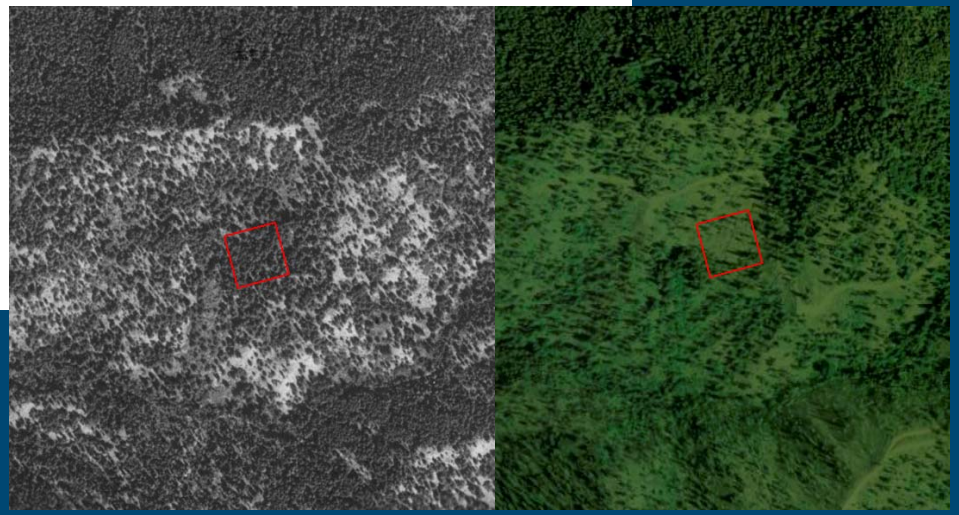
# Agency database false negative

- Database missing harvest




Partial time-series

2006

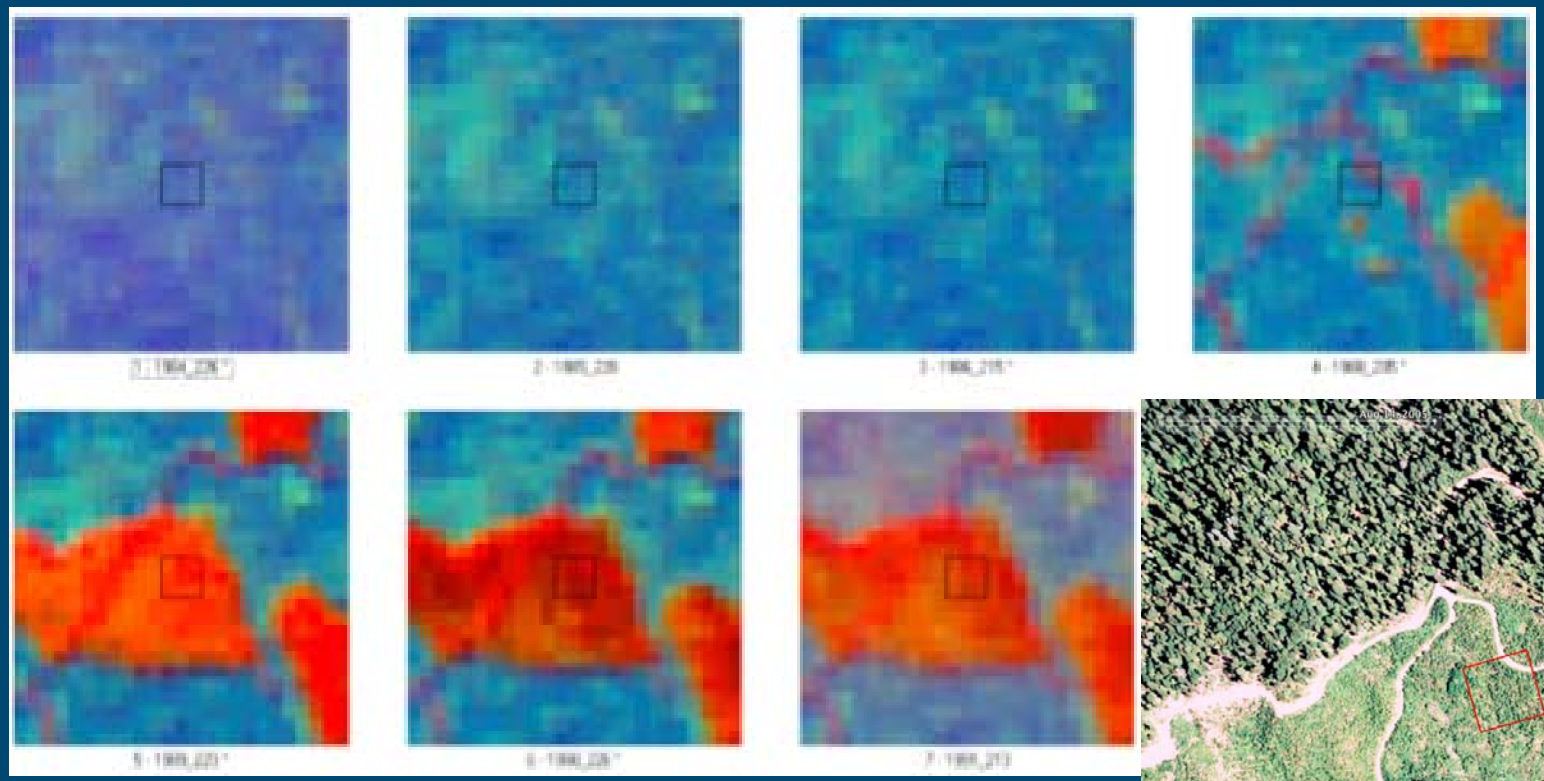


TimeSync plot: 45-27 #38

1998

# Agency database false negative

- Database missing road



Partial time-series

TimeSync plot: 45-29 # 22

# Summary

- LandTrendr:
  - An automated algorithm approach to segment the temporal time series of Landsat data
- TimeSync:
  - A manual interpretation and segmentation approach to summarize the Landsat archive at points
- Philosophy:
  - Segmentation of temporal dimension appears to capture much of the landscape dynamics

# Which changes do we care about?

- State change
  - Durable conversion from one landcover type to another: urbanization, deforestation, woody encroachment, vegetative community type change
- Cyclical change
  - Temporary conversion that eventually reverts to a prior condition: Disturbance, phenological change
- Condition change
  - Change in "vigor" or structure within a type: insect-related mortality, loss of productivity, increase in density

- *Land cover change means many different things*

# Summary

- Movement from scene-based to pixel-based view of the archive
- Validation is critical, and new algorithms go beyond limits of existing reference data
  - Truth data are themselves approximations

*Key point: The landscape is dynamic, and Landsat data appear to have captured a rich record of that dynamism*

Thank you.